ORIGINAL ARTICLE

A C G publications

J. Chem. Metrol. 8:1 (2014) 1-12

journal of chemical metrology

Comparison of measurement uncertainty estimates using quality control and validation data

Teemu Näykki[*1], Bertil Magnusson[2], Irja Helm[3], Lauri Jalukse[3], Tero Väisänen[4] and Ivo Leito[3]

[1]Finnish Environment Institute, SYKE, Environmental Measurement and Testing Laboratory, Hakuninmaantie 6, 00430 Helsinki, Finland
[2]SP Technical Research Institute of Sweden, Chemistry, Materials and Surfaces, 50115 Borås, Sweden
[3]University of Tartu, Institute of Chemistry, Ravila 14a, 50411 Tartu, Estonia
[4]Finnish Environment Institute, SYKE, Environmental Measurement and Testing Laboratory, Paavo Havaksen tie 3, 90570 Oulu, Finland

(Received May 8, 2014 Revised June 10, 2014; Accepted June 16, 2014)

Abstract: A study was organised where consistency of measurement uncertainty estimations between laboratories in Finland, Sweden and Estonia was surveyed. For all laboratories a file was delivered containing the same set of fictional quality control and validation results for the measurement of total nitrogen in waste water. The laboratories were asked to evaluate the measurement uncertainty using quality control and validation data with the Nordtest approach, using the free MUkit software for measurement uncertainty estimation developed by SYKE. A total of 21 laboratories participated in the survey.

Attention was paid to handling of the data, e.g. selecting the concentration ranges for uncertainty estimation, choosing the appropriate approach among those proposed in the Nordtest guide used for uncertainty estimation, choosing the way in which the uncertainty was reported (absolute or relative) and the outcomes of the measurement uncertainty estimations. Most of the laboratories estimated measurement uncertainty for more than one concentration range. The majority also reported measurement uncertainty in relative numbers, even in the low concentration range, where it is advised for most instrumental methods to perform calculation with absolute values. As measurement uncertainty was reported as relative values, it was heavily underestimated at the lowest concentration levels.

However, the measurement uncertainty estimates were consistent between the laboratories, and variability of relative uncertainty estimates was small (within ± 2% units from the average value). This indicates that with the same data and with the unified uncertainty estimation approach, laboratories are able to achieve the same expanded measurement uncertainty. Therefore, the unified estimation of measurement uncertainty is a way of improving the comparability of analysis results between laboratories.

Keywords: Measurement uncertainty; Nordtest approach; Survey; MUkit; Quality control; Validation.

1. Introduction

Numerous guides are available for the calculation of measurement uncertainty in chemical analysis. The general principles are described in the GUM [1]. Other guides that can be regarded

[*] Corresponding author: E-Mail: teemu.naykki@ymparisto.fi ; Phone: + 358 (40) 734 5080; Fax: + 358 (9) 495 913.

The article was published by Academy of Chemistry of Globe Publications
www.acgpubs.org/JCM © Published 06/24/2014
Online ISSN 1307-6183

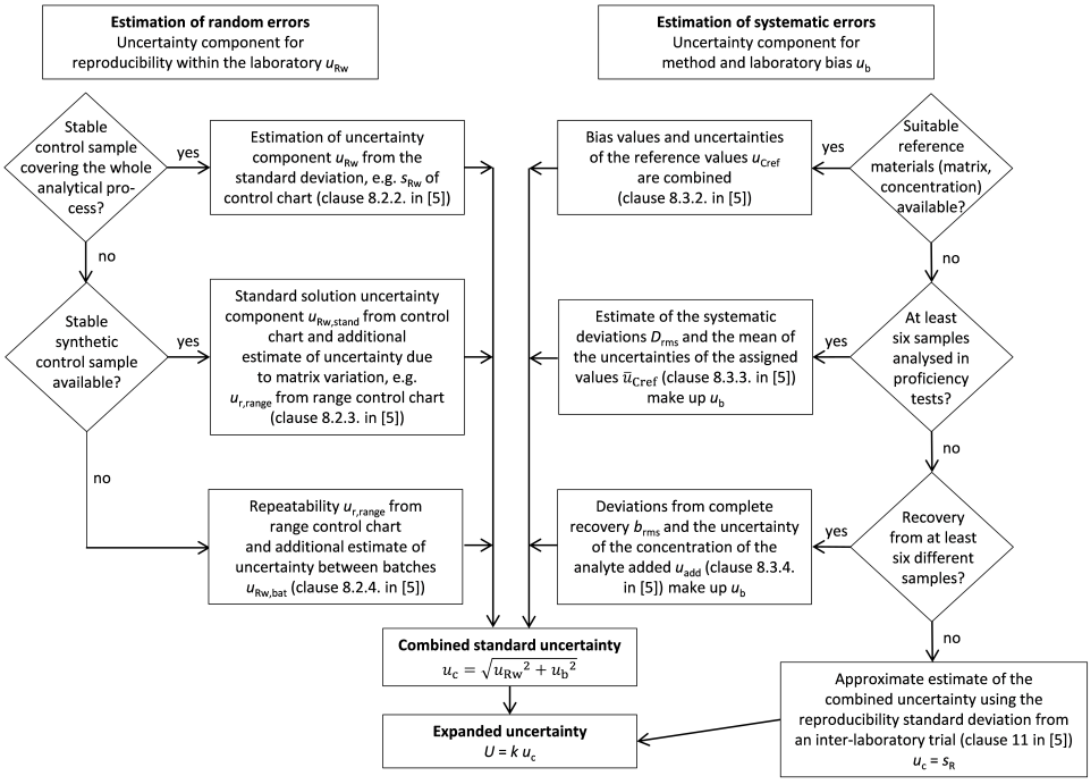
relevant include the Eurachem/CITAC QUAM guide [2], Eurolab [3] and Nordtest [4] Technical Reports and the ISO Standard 11352 [5]. However, many laboratories feel that uncertainty estimation is a laborious and intellectually challenging task to perform. For routine methods that have already been in use for a long time in the laboratory, the experimental approaches [4, 5] for estimating measurement uncertainty can be applied. Software for uncertainty estimates based on quality control and validation data following the Nordtest approach is freely available in Excel® format from the University of Stuttgart [6] and as a separate program, MUkit, from SYKE [7]. Both software packages work with uncertainty calculations in either relative or absolute units.

As previously discovered in the study of Finnish proficiency test (PT) participants [8], laboratories reported very different measurement uncertainties for the same measurement. Many reported low expanded measurement uncertainties (k = 2) in the measurement of inorganic analytes, some of them even lower than 5% in the optimal analyte concentration range, when a realistic expanded measurement uncertainty would have been 10 – 20% when looking at the laboratories' performance in PT. In addition, some laboratories reported their expanded measurement uncertainties being between 20 and 40% or even higher, some in the range of 50 – 100%, which can only be realistic estimates at the lower end of the measurement range. It was concluded that roughly half of the laboratories either over- or underestimated the measurement uncertainty [8], but underestimation was the most common. Underestimation of measurement uncertainty is in fact very common, and the component missing in the uncertainty calculations can be called "dark uncertainty" [9]

The Finnish Environment Institute SYKE, the Technical Research Institute of Sweden SP and the University of Tartu launched a study in 2013, in which consistency of the measurement uncertainty estimations in laboratories in Finland, Sweden and Estonia was surveyed. The aim of the study was to investigate the variation in uncertainty estimates when laboratories use the same set of data and the same computer software for uncertainty calculations. The data set contained hypothetical, but very realistic routine quality control and proficiency test results from testing laboratories. The laboratories were advised to use MUkit measurement uncertainty software for their uncertainty estimations.

MUkit is practical, user-friendly, open-source, freeware measurement uncertainty estimation software package [7, 8]. The calculations are based on results from validation and internal quality control. The software is mainly based on the Nordtest measurement uncertainty handbook [4] using the single laboratory approach, i.e. the calculations are based on data from only one laboratory. The Nordtest handbook [4] describes two main approaches for the estimation of combined standard uncertainty u_c. The first one includes estimation of u_c according to reproducibility standard deviation s_R reported in a standard method from an inter-laboratory trial [10]. The other one, which the MUkit software is based on, is the estimation of 1) the uncertainty component from within-laboratory reproducibility u_{Rw} (also called intermediate precision), and 2) the uncertainty component due to possible method and laboratory bias u_b. u_{Rw} reflects the random error component covering method repeatability and day-to-day variation while u_b describes the systematic effects of method and individual laboratory. Both of these uncertainty components can be conveniently estimated from quality control and validation data [11], for example, thus significantly reducing the need for performing dedicated experiments for estimating detailed uncertainty contributions and thereby simplifying the uncertainty estimation in routine laboratories [8].

The Nordtest approach, using quality control and validation data, describes several different ways of estimating u_{Rw} or u_b for each concentration levels (Scheme 1). For u_{Rw}, a laboratory can choose to use 1) the analysis results of "control sample covering whole analytical procedure" or 2) the analysis results of "control samples and routine sample replicates". For u_b, a laboratory can choose whether to use 1) analysis results of one or more "Certified Reference Materials (CRM)", 2) the results of "proficiency tests PT" or 3) the results of "recovery tests" [4]. All these possible choices may result in different measurement uncertainty estimates calculated by the Nordtest approach.



Scheme 1. Schematic approach for the estimation of measurement uncertainty for one concentration level. Scheme is reproduced from Figure 1 in ISO 11352 [5] courtesy of ISO, Geneva

2. Participant questionnaire

In connection with the measurement uncertainty comparison study, a questionnaire was sent to the 21 laboratories to identify the uncertainty estimation and other quality control procedures performed in laboratories. Sixteen laboratories responded to the questionnaire. Laboratories were both public routine (n=6) and public research (n=1) laboratories, private laboratories (n=5) and industrial laboratories (n=4).

The survey revealed that all of the laboratories, even industrial, estimate measurement uncertainties for their analysis results. This may be regarded as an excellent result. In a similar survey conducted in 2011, 9% of the respondents indicated that they do not perform uncertainty calculations [8]. At that time, the laboratories not calculating measurement uncertainties operated mainly in the industrial sector.

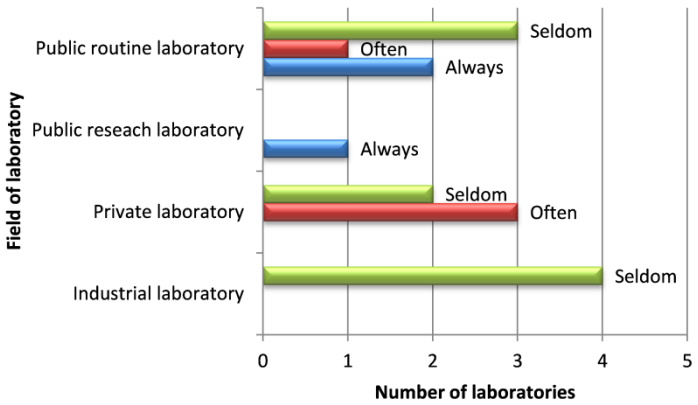


Figure 1. Number of laboratories reporting measurement uncertainty to their customers. Respondents are grouped according to their field of operation

Similarly to the survey conducted in 2011 [8], private, public routine and public research laboratories report measurement uncertainty to their customers more often than industrial laboratories (Fig. 1). Many laboratories stated that their customers either do not want to know about measurement uncertainty, or that reporting leads to confusion.

Fourteen laboratories out of the 16 that responded based their measurement uncertainty calculations on the Nordtest TR 537 [4]. Other guides used most frequently were the Eurachem/CITAC QUAM guide [2] and guides from national institutes or universities (Fig. 2). Most laboratories (88%) reported that they routinely apply quality control charts as one component for measurement uncertainty estimation and 25% use MUkit software for the final calculation of measurement uncertainty.

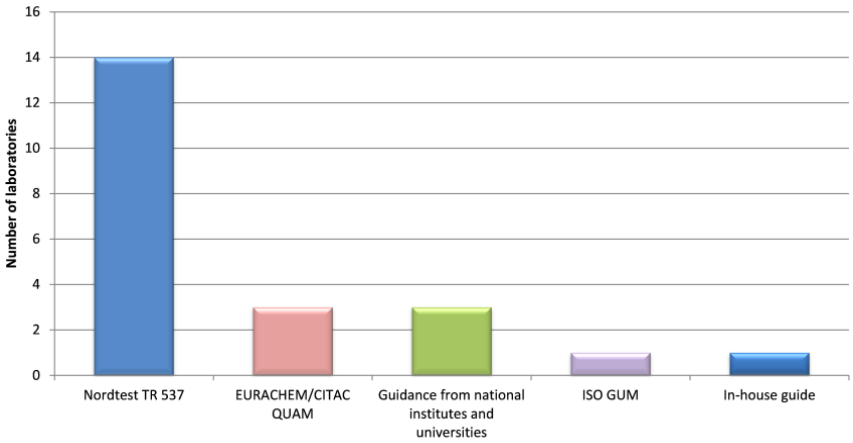


Figure 2. The guidance used for measurement uncertainty estimations. Five laboratories reported that they used more than one guide

3. A study to compare measurement uncertainty estimations

A total of 21 laboratories participated in the survey at the beginning of 2014, with participating laboratories in Finland (*n*=16), Estonia (*n*=3) and Sweden (*n*=2). The scope of the study was to investigate the different measurement uncertainty estimations carried out by the laboratories.

Table 1. General information on the analysis method of total nitrogen in waste water

Analyte	Total nitrogen (N_{tot})
Matrix	Waste water
Analysis method	In-house method based on standard EN ISO 11905-1, Determination of nitrogen - Part 1: Method using oxidative digestion with peroxodisulfate
Analyser used	Skalar SAN++
Sample pre-treatment	Oxidation with peroxodisulfate in autoclave, 120 °C and 30 minutes
Range of detection	up to 5 mg/L

An Excel® file was sent to the participants which contained the set of fictional analysis results of the measurement of total nitrogen in waste water, and the laboratories were asked to calculate the measurement uncertainty using the MUkit software following the Nordtest approach. The data set contained general information on the fictional analysis method (Table 1), measurement results of waste water routine sample replicates at different concentration levels, results of control samples/certified reference materials (CRM), and results of proficiency tests (PT). A summary of the data delivered is described in Table 2. The Excel® file is available as Electronic Supplementary Material.

Table 2. Summary of the quality control and PT results provided to the survey participants for total nitrogen

Source of the data	Concentration (mg/L)	Number of results	Additional information
Replicates from routine samples	0.16 – 4.99	115	Number of duplicate pairs. See chart for *R*-% values in Fig. 5
Control sample/CRM$_1$	0.5	75	Confidence interval of the reference value ± 0.01 mg/L (k = 2)
Control sample/CRM$_2$	2.5	123	Confidence interval of the reference value ± 0.04 mg/L (k = 2)
Control sample/CRM$_3$	4.0	75	Confidence interval of the reference value ± 0.05 mg/L (k = 2)
Proficiency tests	0.28 – 4.20	12	PT results analysed over several years and from different waste water matrices. Assigned values, between laboratory standard deviation and the number of participants were also given

3.1. Uncertainty estimation suggested by survey organiser based on the study data supplied to the participants

The data set was divided into two ranges according to the behaviour of random variation as the function of total nitrogen concentration for replicate analysis (see Fig. 5). According to the visual examination of Figure 5, 1.0 mg/L was chosen for the limit concentration, so the low range covered concentrations of 0.16 – 1.0 mg/L and the high range concentrations of >1.0 – 5.0 mg/L.

u_{Rw} was estimated by pooling standard deviations obtained from 1) routine sample replicates and 2) synthetic control samples in the low and high ranges. In this way, part of the repeatability component will be included twice, but usually it is small in comparison to the between-days variation [4]. Including routine sample replicate results for calculation, the laboratories incorporate random variation resulting from sample inhomogeneity, for example, due to particles in the test samples, which could be regarded as an important factor specifically in waste water analysis, where the

particles contain nitrogen. When also including synthetic control sample results in the calculation of u_{Rw}, variation includes results from different days e.g. change of analyst and influence of changes in the instrument calibrations. These factors are not included in the variation of the routine sample replicate measurements.

u_b was estimated using certified reference materials. In this study, PT results from the years 2007-2012 were also available, but as stated earlier PTs are inferior to CRMs for estimation of bias. If the PT results were collected over a time period of several years, and the CRM results available were analysed during the most recent year, the bias estimate based on CRM results should better reflect the laboratory's current performance. It should be also noted that uncertainty may be overestimated, if proficiency test results are used for bias estimates in the case where the random component "between laboratory standard deviation" is high and the number of participating laboratories is low. Then the uncertainty of the assigned value will become too high, leading to an overestimated bias component. In this study, bias estimate increased only slightly (0.03 mg/L higher at low range and 1.5% higher at high range) if PT results were used instead of CRMs.

Table 3 presents a summary of expanded measurement uncertainty estimation (k = 2) based on the authors' experience. The measurement uncertainty was 0.11 mg/L for a low concentration range (0.16 – 1.0 mg/L) and 14% for a high concentration range (>1.0 – 5.0 mg/L).

Table 3. Summary of *one possible* set of MUkit uncertainty calculations for total nitrogen measurement in waste water proposed by the authors according to the Nordtest approach [4]. For the low concentration range, uncertainty is expressed as an absolute value (mg/L), and as a relative value for the high range (%). The expanded uncertainty value is reported with two significant figures and is rounded upwards.

Component of Uncertainty	**Low Range (0.16 – 1.0 mg/L)**	**High Range (>1.0 – 5.0 mg/L)**
Within-laboratory Reproducibility		
Standard deviation from synthetic control samples	0.035 mg/L[a]	4.66%[b]
Standard deviation from routine sample replicates	0.034 mg/L	4.49%
$\boldsymbol{u_{Rw}}$	**0.049 mg/L**	**6.47%**
Bias		
bias (b) from CRM analysis results	0.015 mg/L[a]	1.72%[c]
s_b	0.035 mg/L	-
u_{cref}	0.005 mg/L	0.7%
$\boldsymbol{u_b}$	**0.017 mg/L**	**1.87%**
u_c	0.052 mg/L	6.7%
$\boldsymbol{U}$ **(**$\boldsymbol{k}$ **= 2)**	**0.11 mg/L**[d]	**14%**[d]

[a] Control sample (CRM_1) measurement results were applied.
[b] Control sample (CRM_2) measurement results were applied, since the standard deviation of its results was higher than for CRM_3.
[c] Since two CRMs (CRM_2 and CRM_3) were used for bias estimate, Root Mean Square of bias (RMS_b) is calculated thus: $RMS_b = \sqrt{\frac{\sum(b_i)^2}{M}}$ (Eq. 1), where b_i is the bias estimate obtained from individual CRM_i measurement results, i=2 and 3, M is the number of different CRMs used.
[d] The expanded uncertainty is rounded up.
NOTE:
s_b is the standard deviation of the CRM measurement results applied for bias estimate.
u_{cref} is the standard uncertainty of the reference value of the CRM. In cases where several CRMs used, it is the average of individual u_{cref} values

u_b for the low range using one CRM: $u_b = \sqrt{b^2 + \left(\frac{s_b}{\sqrt{n}}\right)^2 + u_{cref}^2}$ (Eq. 2), where n is the number of CRM measurement results

u_b for the high range using two CRMs: $u_b = \sqrt{RMS_b^2 + u_{cref}^2}$ (Eq. 3)

Typically, the aim is to have "joint" sub-ranges, i.e. the same value of uncertainty on both sides of the limit concentration between the low and high ranges. Otherwise, the uncertainty function will undergo a break, as can be seen from the triangle and circular signs at a concentration of 1 mg/L in Figure 3. As a solution, it is advised to adjust the limit concentration from 1.0 mg/L to 0.8 mg/L, where relative uncertainty equals absolute uncertainty.

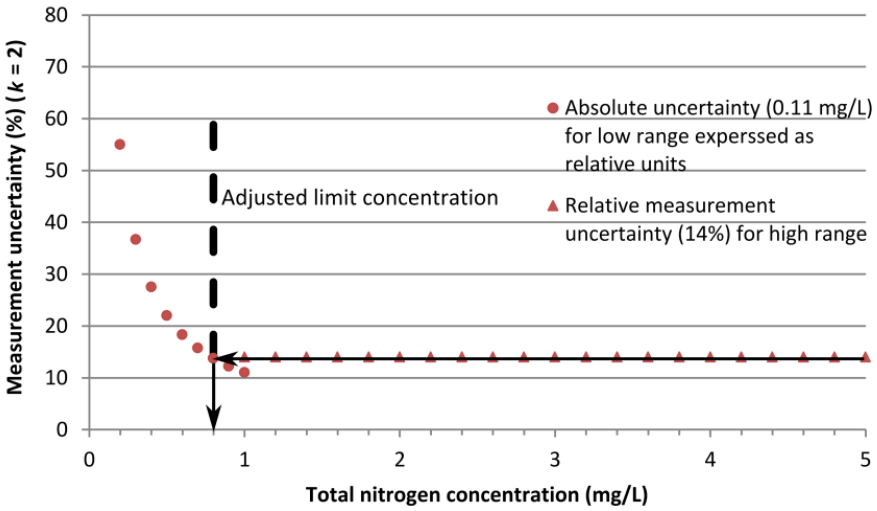


Figure 3. Setting up joint sub-ranges for the measurement uncertainty function. The limit value was adjusted from 1.0 mg/L to 0.8 mg/L

A more detailed uncertainty estimation report (MU_Ntot.muk2) is available from the ENVICAL SYKE website [7]. The file can be viewed and modified by MUkit (Measurement Uncertainty kit) software version 1.9.5.0 or higher, which is also freely available from the ENVICAL SYKE website [7].

4. Results and discussion - Observations according to the study

4.1. Concentration ranges

The measurement uncertainty will normally vary in concentration in terms of the instrumental methods used. In the lower concentration ranges the absolute measurement uncertainty is usually constant, while at higher concentrations the relative uncertainty is constant [2, 4, 12]. Therefore, it is advised to divide the measurable concentration range into parts, and use either fixed relative measurement uncertainty or absolute uncertainty. Nevertheless, six of the laboratories (29%) estimated measurement uncertainty for only one range. The other laboratories divided the measurement uncertainty estimation in two or three ranges (Fig. 4).

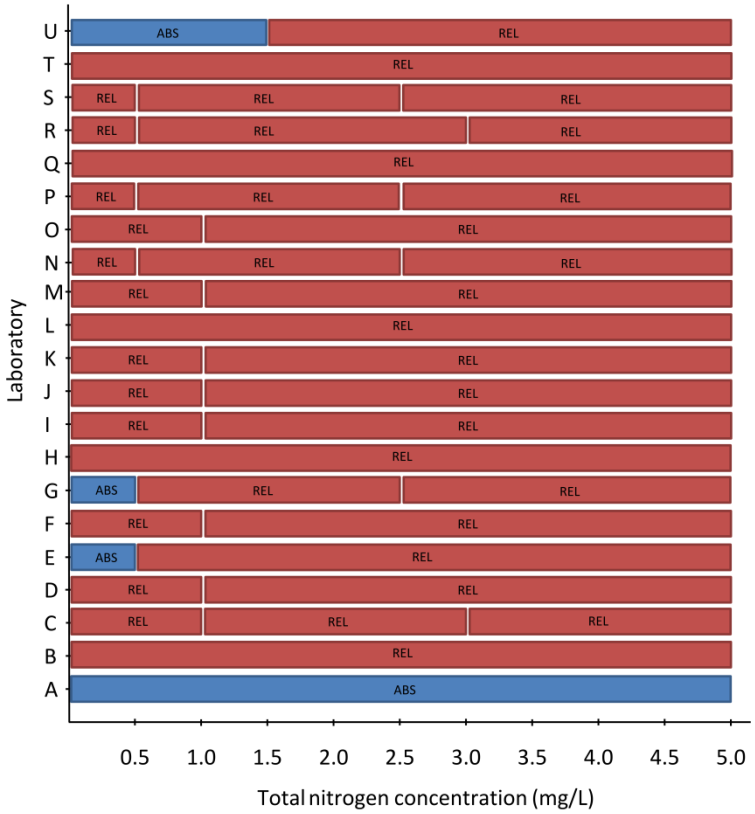


Figure 4. Laboratories' choices for the measurement uncertainty estimation as a function of total nitrogen concentration. ABS = uncertainty estimation in absolute value (mg/L). REL = uncertainty estimation in relative value (%). Number of ranges used for measurement uncertainty estimation at different laboratories. Industrial laboratories A – D; private laboratories E – L; public research laboratories M – N; public routine laboratories O – U

The analysis results in the given data set ranged from ca. 0.2 mg/L to 5.0 mg/L. Most of the laboratories indicated the low (or lowest) range to cover total nitrogen concentrations up to 0.5 – 1.5 mg/L. Typically, the decision for limit concentration between the low and high ranges was based on the visual study of behaviour of relative random variation using duplicate analysis as the function of total nitrogen concentration (Fig. 5). The concentration range was divided at the point where random variation started to increase or remained constant, respectively. An example of setting the limit concentration is shown with a dashed line in Fig. 5.

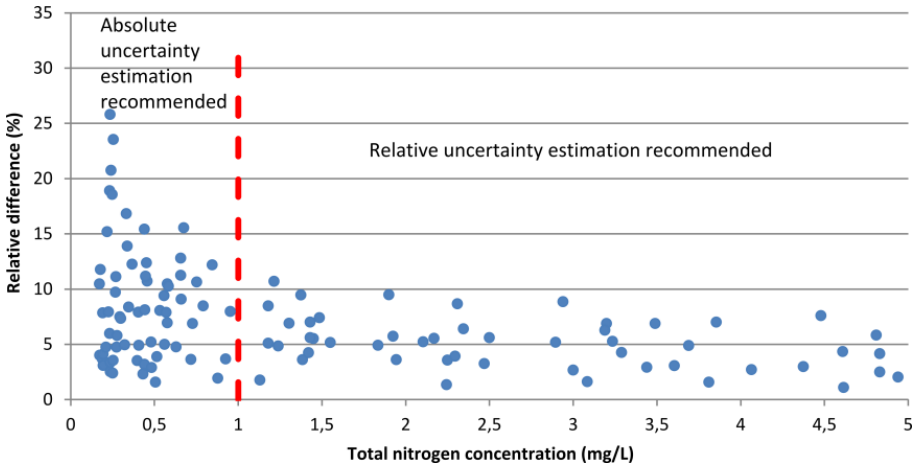


Figure 5. Study of relative difference (max-min) as a function of total nitrogen concentration. Each dot is calculated from the difference of routine sample duplicate results as a percentage of their mean value (*r*-% or *R*-% value [11]). The dashed line signifies the limit concentration between the low and high ranges.

According to Nordtest TR 537 [4], at low concentration levels, it is better to use an absolute uncertainty rather than a relative for most instrumental methods, as relative numbers tend to become extreme at very low concentrations. Though most of the laboratories studied how the random variation behaved as a function of concentration (Fig. 5), only four laboratories estimated measurement uncertainty as absolute values in the low concentration range. One of the laboratories estimated absolute uncertainty also for high concentration levels (Fig. 4).

4.2. Approaches selected for the calculation of u_{Rw}

According to Nordtest TR 537 [4] and ISO 11352 [5], within-laboratory reproducibility u_{Rw} can be calculated from the measurement results of a synthetic control sample, which has a similar matrix to routine samples, and the control sample covers the whole analytical process. When stable routine samples are not available, another option is to use a synthetic control sample together with routine sample replicate results for the calculation of u_{Rw} (Scheme 1).

Two laboratories (10%) calculated u_{Rw} using the approach "control sample covering the whole analytical procedure". The uncertainty may be underestimated if the laboratory calculates the u_{Rw} using only the results of control samples which have a less complex matrix than the routine samples, i.e. with lower levels (or absence) of interfering substances. In this case, the within-laboratory reproducibility may become too optimistic when the variation resulting from sample inhomogeneity, for example, is not included in the uncertainty budget [8].

4.3. Approaches selected for calculation of u_b

According to Nordtest TR 537 [4] and ISO 11352 [5], in order to calculate the uncertainty component due to possible bias, u_b, the laboratory can choose whether to use measurement results of one or more Certified Reference Materials (CRM), the results of PTs or the results of recovery tests. In this study, recovery test results were not distributed to the laboratories.

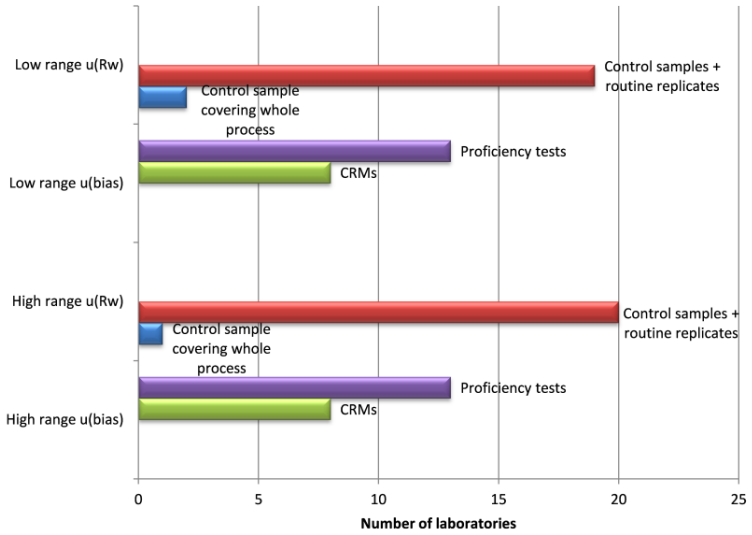


Figure 6. Approaches used by laboratories for the estimation of u_{Rw} and u_b for low and high concentration levels

Many of the respondents in the 2011 survey [8] used proficiency test results for calculating the u_b. Interestingly, this was also true for the current study, where the majority of the laboratories selected using proficiency test results for bias estimates (Fig. 6), although it is clearly stated in the Nordtest guide that in most cases PTs are inferior to CRMs for this purpose. The reasons for this are twofold. Firstly, when using a CRM, the bias can be determined much more reliably, using several replicates over time, while PT results simply provide a difference. This difference includes not only the systematic component, but also the random component. Using PT results for estimating the bias component of uncertainty leads to overestimated uncertainties. This is a drawback of the Nordtest approach. However, when no CRMs similar to test samples are available, the usage of PT results is still better than not estimating uncertainty at all. In this case the CRMs are similar to the test samples, but with no particles. The variation due to particles in test samples can be estimated very effectively using duplicate determinations.

4.4. Uncertainties estimated by laboratories

As seen from Fig. 4, only a few laboratories reported the absolute value for measurement uncertainty at low concentration ranges. The average of the reported absolute expanded measurement uncertainties (k = 2) was 0.17 mg/L. Fig. 7 presents the relative expanded measurement uncertainty estimates (k = 2) of the laboratories that estimated measurement uncertainty for more than one concentration range. Most of the laboratories used MUkit software for the estimation of measurement uncertainty as requested. Few laboratories calculated statistics (e.g. standard deviation) by other means and inserted the calculated value to the MUkit software. Unfortunately there were calculation or typing errors, which dramatically affected the uncertainty estimations of these laboratories. These were treated as outliers and were not included in Fig. 7. All these errors would have been avoided if laboratories copied and pasted source data directly to the software.

The average of the expanded measurement uncertainty (k = 2) for the low range was 22.7%, and for the high range it was 13.4%. Most of the laboratories were well within ± 2% units from the average value in both ranges (Fig. 7). The average result of the survey participants for the high concentration range (13.4%) is the same as the uncertainty suggested by the survey organiser. For the low range, the measurement uncertainty was underestimated. As the relative uncertainty increases at lower concentrations, the estimated measurement uncertainty becomes an underestimate at the limit of

quantification. For example, at a concentration of 0.5 mg/L, expanded measurement uncertainty of 22.7% equals 0.11 mg/L, which is in line with the uncertainty estimation previously suggested by the survey organiser. At a concentration level of 0.2 mg/L, the same relative expanded measurement would be only 0.05 mg/L as an absolute value, which is heavily underestimated.

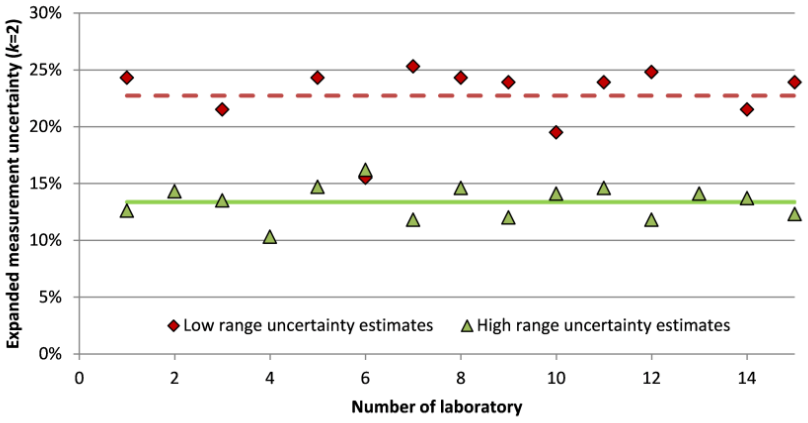


Figure 7. Relative expanded measurement uncertainty estimates calculated by laboratories for fictional total nitrogen determination in waste water. The dashed line signifies the average of the low range uncertainty estimates. The thick line describes the average of the high range uncertainty estimates. For laboratory number 6, the symbols are overlapping.

However, the results indicate that the unified estimation of measurement uncertainty is a way of improving the comparability of measurement results between laboratories, when their uncertainty estimates are more comparable. At the same time measurement uncertainties estimated in different laboratories are improved when the number of under- and overestimations of measurement uncertainties is reduced. This will also lead to better usability of the measurement results for decision-making processes.

5. Conclusions

The unified estimation of measurement uncertainty is the way to improve the comparability of measurement results between laboratories. The problem seems to be that many laboratories estimate measurement uncertainty for only one concentration range. Further understanding is also needed for the way the concentration ranges are chosen and which approach is selected for uncertainty estimation. Additionally, very few laboratories seemed to understand that for concentration levels close to the limit of quantification, the absolute measurement uncertainty should be used. As the relative uncertainty increases at lower concentrations, the estimated measurement uncertainty becomes an underestimate at the limit of quantification.

Acknowledgements

This work was partially supported by Graduate School “Functional materials and technologies” at the University of Tartu, Estonia, having received funding from the European Social Fund under project 1.2.0401.09-0079. The authors also wish to acknowledge the registered association ‘Maa- ja vesitekniikan tuki ry’ for financial support.

Markku Ilmakunnas (Proftest SYKE) is acknowledged for his support in conducting the survey. All participating laboratories are warmly acknowledged for the time they spent on the survey.

Supporting Information

Supporting Information accompanies this paper on http://www.acgpubs.org/JCM

References

[1] JCGM 100 (2008). Evaluation of measurement data - Guide to the expression of uncertainty in measurement (GUM).
[2] Eurachem/CITAC Guide CG 4 (2012). Quantifying Uncertainty in Analytical Measurement, 3rd edition.
[3] Eurolab (2007). Measurement Uncertainty Revisited: Alternative Approaches to Uncertainty Evaluation.
[4] B. Magnusson, T. Näykki, H. Hovind and M. Krysell (2011). Nordtest Technical Report 537 - Handbook for the calculation of measurement uncertainty in environmental laboratories, 3rd edition, Nordic Innovation, Oslo, Norway.
[5] ISO 11352:2012. Water quality — Estimation of measurement uncertainty based on validation and quality control data, ISO 2012.
[6] www.iswa.uni-stuttgart.de/ch/aqs/download/freeware.en.html
[7] MUkit website: http://www.syke.fi/en-US/Services/Calibration_services_and_contract_laboratory/MUkit__Measurement_Uncertainty_Kit
[8] T. Näykki, A. Virtanen and I. Leito (2012). Software support for the Nordtest method of measurement uncertainty evaluation, *Accred Qual Assur*. **17**, 603–612.
[9] S.L.R. Ellison and M. Thompson (2011). Dark Uncertainty, *Accred Qual Assur*. **16**, 483–487.
[10] ISO 21748:2010, Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation, ISO 2010.
[11] H. Hovind, B. Magnusson, M. Krysell, U. Lund and I. Mäkinen (2011). Nordtest Technical Report 569 - Internal Quality Control - Handbook for Chemical laboratories, 4th edition, Nordic Innovation, Oslo, Norway.
[12] B. Magnusson and M. Koch (2012). Use of characteristic functions derived from proficiency testing data to evaluate measurement uncertainties, *Accred Qual Assur*. **17**, 399–403.