



OA03

## MEASUREMENT UNCERTAINTY IN SAMPLING AND CHEMICAL TESTING

### Table of contents

1	GENERAL .....	2
2	THE STANDARD SIST EN ISO/IEC 17025 .....	2
3	INTERPRETATION OF TERMS .....	3
4	MEASUREMENT UNCERTAINTY IN CHEMICAL TESTING .....	7
4.1	Introduction .....	7
4.2	Procedure for evaluating measurement uncertainty according to GUM .....	7
4.3	Evaluation of measurement uncertainty using validation and quality control data .....	10
4.4	Calculation of total (laboratory) measurement uncertainty based on within-laboratory reproducibility and bias uncertainty .....	12
4.5	Reporting of the results of quantitative testing .....	13
4.6	Sampling .....	13
5	CHANGES WITH REGARD TO PREVIOUS REVISION .....	18
6	TRANSITORY PROVISIONS .....	18
7	CONTROL OF THE DOCUMENT .....	19
8	REFERENCES .....	19

## 1 GENERAL

This document shall be considered as expert interpretation of the requirements of the standard SIST EN ISO/IEC 17025 [Ref. 1] for application in laboratories which perform sampling and/or chemical testing. It is intended as guidelines for the laboratories to meet the requirements for accreditation as well as for assessors to assess these laboratories.

The approaches and evaluation methods for measurement uncertainty are already known and established in most chemical testing areas. In cases where the nature of the test method does not allow exact evaluation of measurement uncertainty (according to the procedures from GUM [Ref. 5]), and the laboratory makes the evaluation of measurement uncertainty by taking into account the theoretical fundamentals of the test method and the capacities found in practice [Ref. 1, 7.6.3], it may rely on the standard ISO 21748 [Ref. 6] and the ISO 5725 series of standards [Ref. 7–12].

In SIST EN ISO/IEC 17025:2017, the requirements for evaluation of contributions arising from sampling are also explicitly mentioned. The methodologies can be expected to evolve further. Therefore, the laboratories performing sampling should estimate the contribution to the measurement uncertainty arising from sampling on the basis of the available data and taking into account the currently valid approach in the sub-area concerned, and at the same time follow and take into account the development of new approaches. In Chapter 4 below, a relatively easy and generally applicable approach to the evaluation of measurement uncertainty arising from sampling and/or chemical testing is described.

## 2 THE STANDARD SIST EN ISO/IEC 17025

SIST EN ISO/IEC 17025:2017 lays down the requirements regarding the evaluation of measurement uncertainty under sub-clause 7.6; however, measurement uncertainty is also considered in other clauses (e.g. 3.8, 6.4.1, 6.4.5, 6.4.6, 6.5.1, 7.2.1.1, 7.2.2.1 f), 7.2.2.3, 7.5.1, 7.8.3.1 c) and 7.8.5 f)). The standard requires of the laboratories to use appropriate methods and procedures for evaluation of measurement uncertainty [Ref. 1, 7.2.1.1]. In Notes, SIST EN ISO/IEC 17025 refers to additional information in ISO/IEC Guide 99:2012 [Ref. 3], in the ISO 5725 series of standards [Ref. 7–12] and in ISO 21748 [Ref. 6].

The following standpoints can be identified from the above-mentioned clauses:

- In every testing, all the components contributing to measurement uncertainty that are appropriate to be evaluated shall be identified.
- Meaningful evaluation of measurement uncertainty needs to be carried out based on existing knowledge of the method and on experience.
- Existing data and knowledge may be used (e.g. data obtained in validation/verification of the method, control charts from internal quality control, results of participation in interlaboratory comparisons, results of reference material testing, information obtained from References).
- When the result does not refer only to the sample as received by the laboratory, the contribution of sampling to the measurement uncertainty needs to be included in the evaluation of measurement uncertainty of the result. [Ref. 1, 7.6.1]. The laboratory reporting of sampling shall state in its report the information needed to evaluate that contribution [Ref. 1, 7.8.5 f)].

- Verification or validation of a method also includes evaluation of measurement uncertainty based on the theoretical principles and experimental data as well as evidence that the laboratory meets a target measurement uncertainty [Ref. 1, 7.2.2.1 f), 7.2.2.3].
- When selecting measuring equipment, or defining the requirements regarding calibration of the equipment, and in selecting reference materials, the contribution to measurement uncertainty of the result of testing arising from the use of that equipment should be kept in mind [Ref. 1, 6.4.1, 6.4.5, 6.5.1].
- The laboratory shall ensure that the technical records contain all the data affecting the test result, and the measurement uncertainty related to that result [Ref. 1, 7.5.1].
- The laboratory shall report measurement uncertainty when required so by the customer; when it affects the validity or use of the results for the intended purpose; and in particular, when the measurement uncertainty affects compliance with the limit value [Ref. 1, 7.8.3.1 c)].

### 3 INTERPRETATION OF TERMS

SIST-V ISO/IEC Guide 99:2012 was used in Slovenian definitions, which is a translation of JCGM 200:2012 [Ref. 3]. Where the term is not noted in [Ref. 3], the terminology from JCGM 100:2008 [Ref. 5] and the terminology from [Ref. 32] is used, where the terminology refers to chemical testing and is upgraded by Eurachem Guide [Ref. 4].

**Measurand** [Ref. 3] is the quantity to be measured.

**EXAMPLE:** Vapour pressure of a given sample of water at 20 C.

**NOTE:** The specification of a measurand may require statements about other quantities such as time, temperature and pressure. Note 4 (from SIST ISO 99): In chemistry, “analyte”, or the name of a substance or compound, are terms sometimes used for ‘measurand’. This usage is erroneous because these two terms do not refer to quantities.

**Matrix** [Ref. 2] are all components in a sample except the analyte.

**Verification** [Ref. 1] is provision of objective evidence that a given item fulfils specified requirements.

**Validation** [Ref. 3] is verification that the particular requirements for a specific intended use are fulfilled. Validation of a method is the detailed procedure of experimental verification and the documentation of relevant evidence that the method is appropriate for solving a particular analytical problem. Standard methods are validated for the intended use. All non-standard and standard methods, which are used outside the scope foreseen by the standard, need to be validated. Also, all sampling and sample handling methods not making part of the standard method need to be validated.

**Reference quantity value,  $\mu$**  [Ref. 3] is the quantity value used as a basis for comparison with values of quantities of the same kind. It corresponds to the true value of a result, when this value is known or identifiable. In some cases, the true value is determined by consensus, i.e., ascribed.

**True quantity value** [Ref. 3] is the quantity value consistent with the definition of a quantity, contained, with a stated probability, within an interval described by measurement uncertainty. True value corresponds to ideal execution of a measurement and also has its own measurement uncertainty. When

the measurement uncertainty of the true value is negligible in comparison with the measurement uncertainty of the result, a single true value can be assumed.

**Measurement accuracy** [Ref. 3] is a measure of how individual measurement result deviates from the true value. It is a qualitative concept related to the verification of an evaluated measurement uncertainty. 'Measurement uncertainty' is the appropriate quantitative concept.

**Measurement trueness** [Ref. 3] is a measure of how the average of a large number of measurements carried out deviates from the true value. It is a qualitative concept related to the verification of an estimated bias; derivations from the estimate are provided in ISO 5725. 'Bias' is the appropriate quantitative concept for evaluation.

**Measurement error** [Ref. 3] is the deviation of individual test result from the true value.

**Systematic measurement error** [Ref. 3] is a component of measurement error that with increasing the number of replicate measurements remains constant or varies in a predictable manner.

**Random measurement error** [Ref. 3] is a part of deviation that decreases with repeating measurements, the average value of replicate measurements being closer to the true value.

**Bias** [Ref. 5] is the difference between the average of a large number of replicate tests and the reference value. It is usually determined using appropriate reference materials (similar matrix and concentration range).

**Bias of a measuring instrument** [Ref. 3] is the difference between the average of a large number of readings of a measuring instrument and the reference value.

**NOTE:** The bias of a measuring instrument is normally estimated by averaging the error of indication over an appropriate number of repeated measurements.

**Measurement precision** [Ref. 3] is a qualitative concept representing the capability of repeating a measurement under pre-specified conditions: repeatability conditions of measurement [Ref. 3], intermediate precision conditions of measurement [Ref. 3], or reproducibility conditions of measurement [Ref. 3]. 'Standard deviation' is the appropriate quantitative concept [Ref. 3] ( $s_r$  – for repeatability,  $s_{(RW)}$  – for intermediate precision, and  $s_R$  – for reproducibility).

**Interlaboratory comparison, ILC** [Ref. 1] is the organisation, performance and evaluation of tests on same or similar test items where two or more laboratories collaborate according to predetermined conditions.

**Proficiency testing, PT** [Ref. 1] is the verification of a laboratory's competence to perform tests with interlaboratory comparison.

**Reference standard** [Ref. 3] is the standard, generally having the highest metrological quality, available at a given location or in a given organisation, performing as a basis for measurements at that location.

**Reference material, RM** [Ref. 3] is the material or substance sufficiently homogeneous and one or more of whose properties are sufficiently well established to be used for the calibration of an apparatus,

the assessment of a measurement method, or for assigning values to materials. Reference materials need not necessarily have assigned quantity values, in which case they can only be used for determination of repeatability.

**NOTE:** Reference material may be in the form of pure gas, liquid, solid, or a mixture thereof. Examples are water for the calibration of viscometers, sapphire as a heat-capacity calibrant in calorimetry, and solutions used for calibration in chemical analysis.

**Certified reference material, CRM** [Ref. 3] is a reference material, accompanied by documentation (certificate) issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceabilities, using valid procedures.

**Measurement uncertainty** [Ref. 3] is the parameter, associated with the result of a measurement, that characterizes the dispersion of the quantity values that could reasonably be attributed to the measurand. Within this dispersion range lies, with a certain degree of probability, the true value of the result.

**NOTES:**

- 1 The parameter may be, for example, a standard deviation (or a specified multiple of it), or the half-width of an interval having a stated level of confidence.
- 2 Measurement uncertainty comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of series of measurements and can be characterized by experimental standard deviations (Type A). The other components, which can also be characterized by standard deviation, are evaluated from assumed probability distributions (normal, rectangular, triangular) based on experience or other information (Type B).
- 3 It is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion.

**Standard uncertainty,  $u$**  [Ref. 5] is the uncertainty of the result of a measurement expressed as a standard deviation.

**Combined standard uncertainty,  $u_c$**  [Ref. 5] of a measurement result is obtained by combining all the uncertainty components evaluated using the law of propagation of uncertainty.

**Expanded uncertainty,  $U$**  [Ref. 5] is an interval within which lies the measurement result with a defined level of confidence.  $U$  is obtained by multiplying the combined standard uncertainties by a coverage factor  $k$ . The choice of the factor  $k$  depends on the level of confidence (for a level of confidence of 95%, when the normal distribution applies,  $k = 2$ ).

**Calibration** [Ref. 3] is a set of operations to establish, under specified conditions, the relation between the quantity values indicated by a measuring instrument or measuring system, or values represented by a material measure or a reference material, and the corresponding quantity values provided by measurement standards.

**NOTES:**

- The result of a calibration permits either the assignment of values of measurands to the indications or the determination of corrections with respect to indications.

- A calibration may also determine other metrological properties such as the effect of influence quantities.
- The result of a calibration may be recorded in a document, sometimes called calibration certificate or calibration report.

EXPLANATORY NOTE:

Distinction should be made between:

- verification of an instrument, which is usually performed by a competent external operator who may, for the purpose of verifying the instrument, also carry out calibration of the measuring system or individual parts thereof, and issue a report of the verification carried out, which serves to the laboratory as evidence that the characteristics of the instrument are in compliance with the manufacturer's specifications, and
- daily within-laboratory calibration by means of reference materials, which ensures metrological traceability of the result [Ref. 1, 6.5.3].

**QC sample** [Ref. 4] by its characteristics, it is as similar to CRM as possible, except that it is prepared using within-laboratory procedures. True value is known, measurement uncertainty is the lowest possible, it is sufficiently stable and it contains the relevant matrix.

**Linearity** [Ref. 4] account should be taken of the measurement uncertainties related to the calibration curve.

**Limit of detection, LOD** [Ref. 2] is a measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a component in a material is  $\beta$ , given a probability  $\alpha$  of falsely claiming its presence. Usually,  $\beta = \alpha = 0.05$ . In a limit of detection, one can confirm with certain degree of probability the presence of a component, but the quantity cannot be measured with sufficient reliability.

**Limit of quantitation, LOQ** [Ref. 4] is the lowest quantity value that can be measured with sufficient accuracy. Accuracy should correspond to the use of the measurement result.

NOTES: Although the LOQ value is not directly related to the evaluation of measurement uncertainty, the uncertainty in the concentration range close to LOQ is constant and it limits the usability of the test method at concentrations lower than LOQ, as the measurement uncertainty of the measurement result may become too high for its use.

Example (see also Fig. 1): The laboratory has evaluated the measurement uncertainty of a test method at 10% in a range from 1.0 to 10 mg/L. They set the lower limit of the testing range – defined as LOQ – at 0.20 mg/L. At 0.20 mg/L, the measurement uncertainty is not 0.02 mg/L, i.e., 10% as in higher concentration, but 0.05 mg/L (25%). In an additional evaluation of measurement uncertainty, the laboratory established that measurement uncertainty can be defined as: 0.2–0.5 mg/L, 0.05 mg/L; 0.5–10 mg/L 10%. At the 0.5 mg/L limit between the two evaluation ranges, the measurement uncertainty is the same (10% and 0.05 mg/L, respectively). Should measurement uncertainty, separately evaluated in two or more ranges, not be the same, it should be established which evaluation provides a more realistic estimate at the limit of the two ranges and, as shown in Fig. 1, adjust the evaluated measurement uncertainty as a function of concentration.

**Ruggedness, Robustness** [Ref. 4] method development or validation studies require the research of sensitivity of a method to changes of some parameters. These data may provide useful information

about the impact of significant parameters and to establishing whether their impact on measurement uncertainty is material.

## 4 MEASUREMENT UNCERTAINTY IN CHEMICAL TESTING

### 4.1 Introduction

For the purpose of validation or verification of a test method, laboratories shall identify the acceptable measurement uncertainty of the results. Doing so they shall take into account:

- the requirements and limitations of the test method,
- the requirements of the client [Ref. 1, 7.2.2.3],
- the limit values on which decisions on conformity to a specification are based.

When the result to be reported by the laboratory refers to the complete population under examination, the evaluation of measurement uncertainty of the measurement result shall include contributions from sampling. When the result refers to a laboratory sample, the contributions from sampling need not be taken into account.

Laboratories may use different ways of evaluating measurement uncertainty of the results. In general, the degree of rigour needed in an evaluation of measurement uncertainty is related to the degree of risk. Should the measurement uncertainty not be acceptable for the client, or should it be too high for decision on compliance with the specification, the laboratory must try to reduce the measurement uncertainty by identifying the highest contributions to the uncertainty by improving the performance of the test method (for example, by a longer measurement time, larger sample, better CRM, ...).

The procedure for evaluating measurement uncertainty according to GUM is often not feasible (due to the complexity of the process or to ignoring the process model), and the measurement uncertainty is evaluated using the validation and quality control data.

The most useful practical strategy of an evaluation of measurement uncertainty is presented in Fig. 1 of ISO 11352 [Ref. 22]. In using Fig. 1 of ISO 11352, it is necessary to comply with the general recommendation from GUM [Ref. 5], after the correction of a well-defined bias, and the EURACHEM GUIDE, 7.16. [Ref. 17]. Laboratories shall constantly check the bias following the method [Ref. 12, Table 7] for all actual matrixes, analyse the data obtained and take action, as appropriate. Then measurement uncertainty shall be evaluated as described in Chapter 4.3 hereof. An evaluated measurement uncertainty shall be verified, where possible, through participation in ILCs, or else by other forms of assuring validity of the results.

### 4.2 Procedure for evaluating measurement uncertainty according to GUM

Identification of the test method's stages and the use of the Cause/Effect ("Fishbone") Diagram could provide a useful approach to evaluating measurement uncertainty and presenting the sources and components of measurement uncertainty. Various types of samples, the matrix and various concentration ranges of the measurand should be taken into account when evaluating measurement uncertainty.

Often individual stages of test methods are common to different test methods. In such cases, the evaluation of measurement uncertainty for individual stage can be used for evaluating combined measurement uncertainty with all the methods in which these stages are used during the testing

procedure. The steps of the procedure of evaluating measurement uncertainty in chemical testing shall be as follow:

- 1 Definition of measurand,
- 2 Identification of sources of measurement uncertainty,
- 3 Quantification of the components of uncertainty,
- 4 Calculation of combined and expanded uncertainty.

**Step 1:** Definition of measurand

The measurand should be clearly defined. The usual measurand in chemical analysis is the concentration of a particular analyte in the matrix. The relation between the measurand and the input quantities shall be determined (model equation). Corrections for known systematic deviations should be included, where possible.

**Step 2:** Identification of sources of measurement uncertainty

Any possible sources of uncertainty should be identified.

Follow some examples:

- incomplete definition of measurand,
- sampling,
- sample transport and storage,
- sample preparation for analysis,
- measurement conditions and environmental conditions,
- operators,
- changes in the test procedure,
- measuring equipment,
- reference materials.

**Step 3:** Quantification of the components of uncertainty

First of all, find out how measurement uncertainty differs with concentration in at least three (recommendable five) [Ref. 6, p. 15] properly selected concentration ranges, a sufficient number of repetitions (e.g. from preliminary data, for example from repeatability, reproducibility within the laboratory, differences in the duplicates of real samples). General guidance is that measurement uncertainty is absolute in a range up to ten times the LOD [Ref. 19, p. 19], and it does not change as a function of concentration. In a concentration range higher than ten times the LOD, however, the relatively expressed measurement uncertainty does not change as a function of concentration (Fig. 1). This is used to plan the execution of those parts of testing range in which measurement uncertainty is evaluated. In general, measurement uncertainty is evaluated in at least three parts of the testing range: in the area around LOQ, in the concentration area around ten times the LOD, and in the higher concentration areas. Determine concentration dependency separately for repeatability and within-laboratory reproducibility. The easiest way of presentation is by a graph, while for establishing the relation, the basic functions from [Ref. 6, Clause 8.5] should be used. The most frequent relation is of the type  $s = a + b \cdot c$ ,  $s$  – standard deviation,  $c$  – concentration,  $a$  and  $b$  – constants.

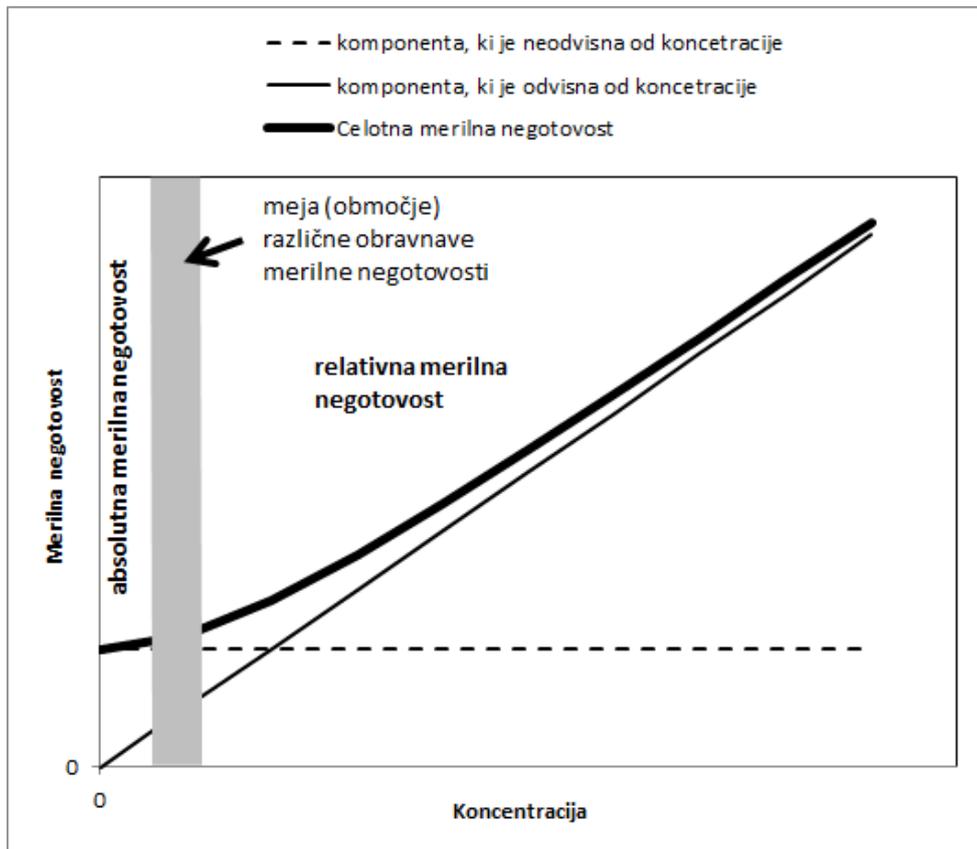


Figure 1: Concentration dependency of measurement uncertainty of a majority of instrumental test procedures (an exception being, for example, determination of pH, ...).

Translation of graph:

komponenta, ki je neodvisna od koncentracije – *component independent of concentration*

komponenta, ki je odvisna od koncentracije – *component dependent of concentration*

celotna merilna negotovost – *Total measurement uncertainty*

meja (območje) različne obravnave merilne negotovosti – *limit (range) of different treatment of measurement uncertainty*

relativna merilna negotovost – *relative measurement uncertainty*

koncentracija – *concentration*

merilna negotovost – *Measurement uncertainty*

absolutna merilna negotovost – *absolute measurement uncertainty*

The extent of the components of uncertainty related to the identified potential sources of uncertainty shall be estimated by carrying out the appropriate experiments or from other available information; the individual components should, however, not be taken into account more than once. It is often possible to estimate or define a contribution to uncertainty related to several separate sources. It is also important to determine whether the available data cover all sources of uncertainty, or to plan additional experiments to provide all necessary information.

**Step 4:** Calculation of combined and expanded measurement uncertainty

The information obtained under Step 3 above consists of a number of quantified contributions to the measurement uncertainty related to individual sources or combined effects of several sources. The contributions shall be expressed as standard uncertainty and combined in accordance with appropriate rules (law of propagation of uncertainty) to provide the combined standard uncertainty. An appropriate coverage factor shall be used for the expanded measurement uncertainty. It should be identified, during the process of identification of individual sources of uncertainty, which sources essentially contribute to the combined measurement uncertainty. Practice will show that they are usually few. For those that are available, reliable information should be obtained. A preliminary estimation of the contribution of each

component or each combination of components to the measurement uncertainty should be made, and attention should be paid to those that are more important. The sources of measurement uncertainty and the values of individual components shall be documented.

When deciding which of the identified components of measurement uncertainty should be comprised individually in the final evaluation, the following should be taken into consideration:

- the effect of reporting of the measurement uncertainty; when considerable material or other consequences are derived from the stated measurement uncertainty or interpretation of the results related to it, approximate estimates of uncertainty should not be stated,
- the degree of rigour in evaluating the measurement uncertainty based on the requirements of the client, as well as legal and other requirements.

Uncertainty also arises from assuring traceability of the measurement results and it often has several components. The uncertainty of all the components of traceability of a measurement procedure (e.g. balances, thermometers, volumetric equipment, reference materials, ...) should be taken into account in the evaluation of measurement uncertainty.

### **4.3 Evaluation of measurement uncertainty using validation and quality control data**

#### 1) Evaluation of reproducibility contribution within the laboratory ( $R_w$ )

The characteristics of method capacity are essential in evaluating measurement uncertainty of the results. In practice, the feasibility of a test method for the expected use is checked through verification or validation of the method. The information obtained can be used for evaluating the measurement uncertainty.

Normally, repeatability within the laboratory ( $R_w$ ) is the essential component of the result of analysis, taking into account the way of carrying out the experiments, as shown in Fig. 1, that is the performance, *mutatis mutandis*, of the evaluation of measurement uncertainty in different parts of the testing range.

An estimate of a component of uncertainty can be obtained in several ways:

- a) By means of a corresponding control sample

Use stable control samples which have the matrix and analyte concentration levels similar to those of the test samples. The analysis of the control sample shall cover the complete analytical process including sample preparation, taking into account the findings from Step 3; usually, performance at three concentration levels will suffice.

A longer time period (a framework of one year), different operators and equipment and all types of test sample matrixes need to be taken into account. Estimate the uncertainty component using standard deviation in repeatability conditions within the laboratory  $s_{(R_w)}$ .

$$U_{(R_w)} = s_{(R_w)} \text{ [Ref. 21]} \quad (1)$$

- b) Without a corresponding control sample

When no control samples with identical matrix are available and a standard solution is used for the control, which does not cover the impact of the matrix, evaluate the measurement uncertainty on the basis of parallel analyses of real samples. Use a specified share of real samples and analyse them in duplicates by including the conditions within the laboratory reproducibility, for example: carry out first

analysis upon receiving the sample using analyte 1, and analysis of second parallel of the same sample after the longest time of storing the sample using analyte 2. In the meantime, carry out re-calibration, when possible, provided that this is the object of the test method. In this way the difference of the replicates in the conditions within the laboratory reproducibility is determined. The evaluation of measurement uncertainty based on parallel measurements (from control charts of differences) of real samples shall comprise different matrixes and concentrations, which cover the complete accredited scope of testing. It is important that at least 8 series of parallel determinations (duplicates) are combined in individual part of the concentration range and in individual group of different matrixes, if there are several matrixes [Ref. 22].

Individual level calculation:

$$S(R_w) = \sqrt{\frac{\sum (s_i)^2}{n}} \quad (2)$$

$s_i$  – standard deviation of individual duplicate analyses

$n$  – number of duplicate analyses (minimum  $n = 8$  duplicate analyses, a total of 16 measurements)

$$\text{or} \\ S(R_w) = \frac{\sum ABS(D)_i}{n * 1.128} \quad (3)$$

ABS(D) – absolute value of differences of parallel measurements »range«

$n$  – number of duplicate analyses (minimum  $n = 8$  duplicate analyses, a total of 16 measurements)

$$U(R_w) = S(R_w) \quad (4)$$

## 2) Evaluation of a measurement uncertainty component due to bias from the true value

Evaluation of measurement uncertainty due to bias from the true value can be carried out using certified reference materials (CRMs), taking part in proficiency tests (PTs), or on the basis of efficiency data (control sample).

It should be ensured that the bias of the method is under control (systematic measurement error being negligible with respect to the random measurement error), which is carried out according to the manner in [Ref. 12, Table 7], where the prescribed positioning of the X central line of the CRM chart (control sample) is at the value of 0% deviation and the use of type 2s or 3s control limits. Should you find out that the bias is not under control, make a correction of the results [Ref. 5, p. 5], or carry out correction procedures in the performance of the standard method [Ref. 6, p. 13]. The choice of the way depends on whether the bias is reliably identified as systematic (permanent) bias [Ref. 5, p. 5].

If the bias can be proved to be under control, the remaining bias, including its measurement uncertainty, may be included in the estimate of the total measurement uncertainty using the procedure from [Ref. 22].

Usually, the same CRM (control sample, ILC) as used for proving that bias is under control is not used for evaluation of a measurement uncertainty component due to bias.

Calculate  $s_r$  and bias using the following formulae:

$$S_r = \sqrt{\frac{\sum (x_i - x_{\text{aver}})^2}{n-1}} \quad (5)$$

$$\text{bias} = x_{\text{aver}} - \mu \quad (6)$$

$x_i$  – result of individual measurement

$x_{\text{aver}}$  – average value of measurements

$n$  – minimum 6 measurements recommendable to meet the condition

$$\sqrt{\frac{s_r^2}{n}} < 0,2 * s_R \text{ [Ref. 6, p. 11] } \quad (7)$$

$s_r$  – repeatability of  $n$  measurements;  $s_R$  – reproducibility (based on the test method data or on ILCs)

$\mu$  – reference value of the result

Two components of bias need to be evaluated [Ref. 21, p. 19]:

- root mean square of individual deviations from the true value ( $\text{RMS}_{\text{bias}}$ ),
- average standard uncertainty of the ascribed/certified true value –  $u(\text{Cref})$ .

$$\text{RMS}_{\text{bias}} = \sqrt{\frac{\sum \text{bias}_i^2}{n_{\text{CRM}}}} \quad (8)$$

$$u_{\text{bias}} = \sqrt{\text{RMS}_{\text{bias}}^2 + u(\text{Cref})^2} \quad (9)$$

$n_{\text{CRM}}$  is the number of CRMs used (control samples, ILCs)

If, exceptionally, only one CRM is used, then:

$$u_{\text{bias}} = \sqrt{\text{bias}^2 + \frac{s_{\text{bias}}^2}{n} + u(\text{Cref})^2} \quad (10)$$

$n$  – minimum 6 measurements of CRM; it is recommendable that the number of measurements performed meets the condition from equation 7.

### Interlaboratory comparative testing

The results of a laboratory taking part in a proficiency test may be used to determine the component of measurement uncertainty due to deviation from the true value, under the following conditions:

- at least 6 participations within a reasonable time interval,
- the types of samples used in an ILC scheme shall be similar to the real samples,
- the accepted values shall be traceable to a suitable reference value, and
- the uncertainty of the accepted value shall be small in comparison with the actual dispersion of the results.

In such cases the dispersion of differences between the reported and the accepted values in repeated proficiency tests would represent the basis for evaluation of measurement uncertainty for that part of testing which is included within the scope of the scheme. The systematic deviation from the traceable accepted values and other sources of measurement uncertainty should be taken into account.

### The efficiency of standard addition (control sample)

The efficiency of standard addition of measurand to a sample can be used to estimate the uncertainty component due to deviation from the true value in methods which do not include correction due to the efficiency in the procedure. For that purpose, experiments need to be performed in such a way that the uncertainty of determination of the efficiency is negligible (less than 1/3–1/5 in comparison with the efficiency expressed in the form of error).

## 4.4 Calculation of total (laboratory) measurement uncertainty based on within-laboratory reproducibility and bias uncertainty

$$u_c(\text{laboratory}) = \sqrt{u_{(Rw)}^2 + u_{\text{bias}}^2} \quad (11)$$

$$U(k = 2, \text{laboratory}) = 2 * u_c(\text{laboratory}) \quad (12)$$

## 4.5 Reporting of the results of quantitative testing

Quantitative tests provide the value expressed in SI units, if possible. When reporting the measurement uncertainty, the expanded measurement uncertainty ( $U$ ) should be stated with the defined degree of confidence, or the combined standard measurement uncertainty ( $u_c$ ). Usually, reference to the procedure used to evaluate the measurement uncertainty is also stated. If all the sources identified as important in a certain area are not considered in the evaluation of measurement uncertainty, this should be clearly stated when reporting. The number of decimals in the stated measurement uncertainty should reflect the practical capability of the measurement. Rarely more than two significant decimal places are reported. Also, the numerical value for the result should be rounded in such a way that the last decimal place corresponds to the last decimal place of the measurement uncertainty. Example: When the result is 123.456, and the evaluated measurement uncertainty was 2.27, the result should be given as a rounded value 123.5±2.3.

## 4.6 Sampling

### 4.6.1 Interpretation of terms

**Sampling programme**, [Ref. 23], **testing programme** [Ref. 28]: collection of all the facts referring to the planning and carrying out examination of a population. It contains a much greater amount of data than a sampling plan and can provide the basis for preparing individual sampling plans [Ref. 28], or it can in itself be the basis for carrying out sampling, as it contains all the data, including practical guidance for the sampler [Ref. 23].

**Sampling plan** [Ref. 28]: instructions for carrying out individual sampling or several samplings. It contains all the necessary information for carrying out one or several samplings with a predetermined or identified measurement uncertainty from sampling.

**Population** [Ref. 28], **sampling target** [Ref. 19]: material that we wish to examine (object of sampling). Examples of population: a part of or complete water body; total quantity of waste; total quantity of flue gases; and similar objects of examination. The scope of the examined population should be defined when drawing up sampling/testing programmes.

**Objectives of testing programme** [Ref. 28]: population information sought.

**Sample** [Ref. 30]: a representative part of the population with regard to the purpose of examination.

**Discrete sample** [Ref. 25, 30]: a sample taken in a certain moment.

**Grab sample** [Ref. 25, 30]: a sample taken at random, see additional comment in [Ref. 30], or a spot sample [Ref. 30] taken according to a pre-set plan which includes the identification of the location and/or time of sample collection, see additional comment in [Ref. 30]; a sample taken from the population being examined (final sampling target) or increment [Ref. 28]: a sample taken at a certain time and a certain place from the population being examined, representing only a part of meeting the sampling target.

**Composite sample** [Ref. 28, 30]: two or more increments or sub-samples mixed together in a homogeneous way.

**Proportional sampling** [Ref. 25]: sampling made at a frequency adapted to the flow rate of the liquid being sampled in the case of discrete sampling; or in the case of continuous sampling, the rate of the pumping by a sampling pump is adapted to the flow rate of the liquid being sampled.

**Time-proportional sampling** [Ref. 26]: sampling where a specified quantity of partial sample is sampled at a predetermined frequency.

**Field sample** [Ref. 28]: complete sample taken from two or more increments without sub-sampling.

**Field duplicate**: two field samples taken in two complete sampling procedures, so that the place and time inhomogeneity of the population is included in the most appropriate way.

**Laboratory sample** [Ref. 28]: a sample received by the laboratory, smaller or equal in quantity to a field sample.

**Test sample** [Ref. 28]: part of a sample prepared from a laboratory sample that serves for the performance of individual testing.

**Sub-sample** [Ref. 28]: a homogeneous sample taken from individual partial sample or from a composite sample to be used in further work. Most often it means a sample taken from a sample.

**Representative sample** [Ref. 28]: a sample having properties that can be ascribed, with a certain degree of probability, to the population.

**Reserve sample** [Ref. 31]: a sample stored in a controlled environment for the needs of repeating the analysis in case of dispute.

**Storage sample** [Ref. 32]: a sample stored in a controlled environment for the needs of repeating the analysis.

#### **4.6.2 Measurement uncertainty evaluation for sampling**

Measurement uncertainty arising from sampling usually exceeds the amount of measurement uncertainty from laboratory testing. It is more difficult to evaluate the measurement uncertainty from sampling than that arising from laboratory testing, since also the time inhomogeneity in the property of the population being examined needs to be included in the estimate of measurement uncertainty from sampling, where relevant, not just the spatial inhomogeneity on the object of sampling during sampling. Also, in some cases it is necessary to include an estimate of spatial inhomogeneity across the entire population in the estimate of measurement uncertainty, even across the part that is inaccessible for technical or other reasons. Such are, for example, batch wastewater discharges; changes in the composition of waste due to changes in production; changes in gas emissions due to the start and different intensity of operation of production lines; water bodies of underground water, where only a limited number of sampling spots is available; and the like.

Examples of evaluations are shown in [Ref. 19, 24 and 34].

The easiest way of estimating measurement uncertainty from sampling is to make two independent collections of at least 8 field samples in the material being sampled according to a repeated full sampling protocol, and – when we wish to also include the uncertainty related to different samplers – each individual field sample taken by another sampler.

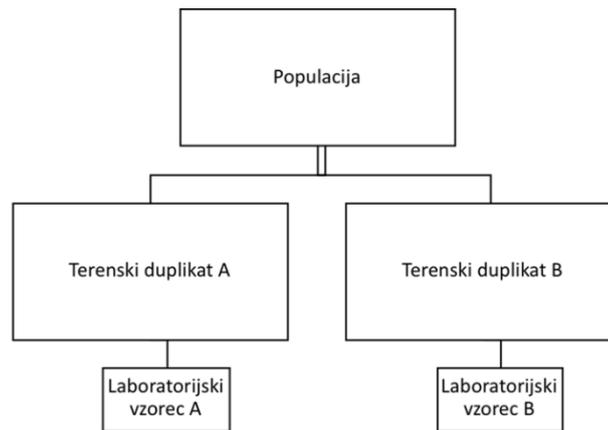


Figure 2 Least demanding determination of measurement uncertainty of sampling (simultaneous or – in case that the population does not significantly change with time – consecutive collection of two field samples according to an established sampling plan).

Translation:

Populacija – Population

Terenski duplikat A – Field duplicate A

Terenski duplikat B – Field duplicate B

Laboratorijski vzorec A – Laboratory sample A

Laboratorijski vzorec B – Laboratory sample B

It is important, when determining measurement uncertainty from sampling, to carry out the repetition of the second collection of the field sample independently of the collection of the first field sample. The uncertainty related to the time and spatial distribution of the measurand in the population during the sampling should be included in the collection (as presented in examples 1–3). Any major changes in time and spatial distribution of the measurand (e.g. major technological and seasonal changes; spatial distribution in which part of the analyte in the population cannot be reached for sampling, or part of the population is inaccessible, all for technical reasons, etc.) should be treated separately from the basic estimate of measurement uncertainty of sampling. Also, the legal requirements need to be taken into account, which may prescribe the method of sampling, for example, during maximum loads. In such case, the reported measurement uncertainty of sampling shall be related to collection at maximum load, and the part of the uncertainty arising from changes in operation shall not be determined.

In case that  $s(\text{laboratory})$  from chapter 4.6.2 hereof cannot be estimated from validation data, the experiments shown in Fig. 2 shall be carried out as shown in Fig. 3 or 4.

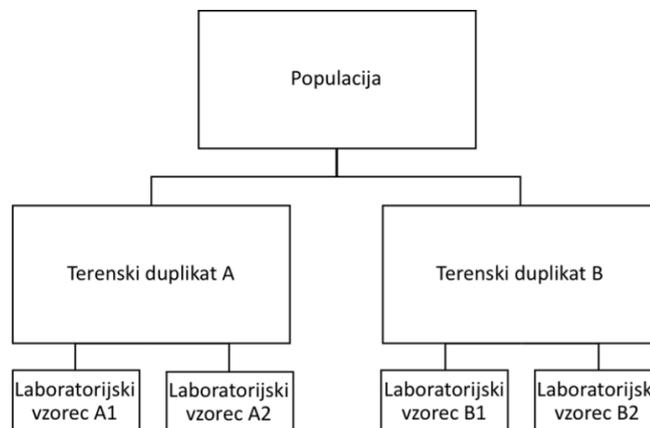


Figure 3 Determination of measurement uncertainty of sampling, when the laboratory part of uncertainty  $s(\text{laboratory})$  cannot be estimated (chapter 4.6.2 hereof.)

Translation:

Populacija – Population

Terenski duplikat A – *Field duplicate A*  
Terenski duplikat B – *Field duplicate B*  
Laboratorijski vzorec A1 – *Laboratory sample A1*  
Laboratorijski vzorec A2 – *Laboratory sample A2*  
Laboratorijski vzorec B1 – *Laboratory sample B1*  
Laboratorijski vzorec B2 – *Laboratory sample B2*

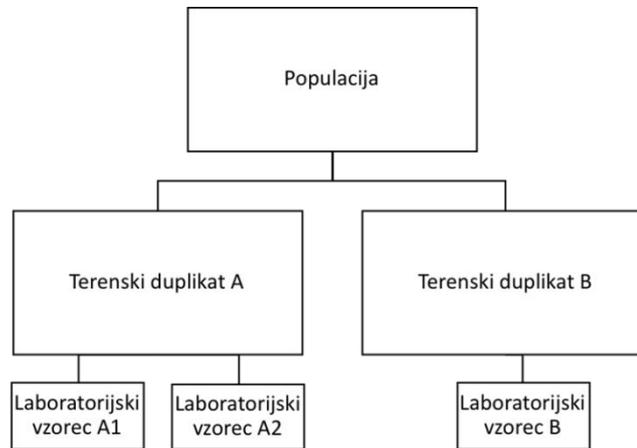


Figure 4 Alternative determination of measurement uncertainty of sampling, when the laboratory part of uncertainty  $s(\text{laboratory})$  cannot be estimated (chapter 4.6.2 hereof).

Translation:

Populacija – *Population*

Terenski duplikat A – *Field duplicate A*

Terenski duplikat B – *Field duplicate B*

Laboratorijski vzorec A1 – *Laboratory sample A1*

Laboratorijski vzorec A2 – *Laboratory sample A2*

Laboratorijski vzorec B – *Laboratory sample B*

When estimating  $s(\text{laboratory})$  from the validation data, i.e., without using additional experiments such as presented in Fig. 2, carefully consider whether all the contributions have been included, i.e., all the contributions related to the full preparation of the sample being tested, taken from the laboratory sample. When not, the experiments presented in Fig. 3 and 4 need to be carried out, or the validation shall be updated.

Follow the descriptions of some examples of experiments, which can be taken into account, *mutatis mutandis*, also in other types of sampling.

**Example 1:** When determining measurement uncertainty from sampling of a 24 hour composite sample of wastewater, two sampling devices must be installed in such a way that they do not take the increment at the same time, but rather in a certain time lag; the most reasonable being that the second sampling device takes the sample after the lapse of half of the time of the frequency set for the first sampling device.

**Example 2:** When determining measurement uncertainty from sampling waste, the time change in the population shall be considered, and when it is negligible, consecutive collection of two field samples shall be carried out on the population according to the sampling plan. The network of collection spots of increments shall be planned so as to ensure the maximum difference between the position of the two networks (e.g., opposite position, see example in [Ref. 19, p. 36]). When time change is present in the population, the collection should be made as in Example 1 according to the sampling plan.

**Example 3:** When determining measurement uncertainty in measuring air emissions, the evaluation shall be carried out in the manner of operation of technological devices, as provided by law.

Field samples from Examples 1–3 shall be treated as usual samples, including any decrease in the sample quantity in the field, collection of subsamples for special types of analysis, and all the procedures not comprised in the evaluation of the laboratory part of measurement uncertainty.

The procedure in Fig. 2 does not include uncertainty due to bias, except for the bias due to different samplers. Neither does it include evaluation of major time changes in the population property being examined, except for changes during the performance of sampling.

It is appropriate to divide populations into similar groups with respect to the properties being examined, when feasible, and to plan experiments within a group. For each group, it is appropriate to find parameters where measurement uncertainty from sampling is interdependent, and to estimate measurement uncertainty for a group of parameters. When the measurement uncertainty from sampling is found to be significant in comparison to the measurement uncertainty from chemical testing, the sources of measurement uncertainty should be found as described under [Ref. 24], and regular control of sampling should be introduced (e.g., regular collection of field duplicates to representative parameters), and where appropriate (i.e., where sampling of duplicates is made with sufficient frequency on similar population), the data should be statistically evaluated using a control chart of differences. Avoid the construction of control cards of differences of sampling duplicates from populations that differ between themselves in terms of properties or concentration range (see Fig. 1).

An estimate of measurement uncertainty from sampling shall include the results from regular control of sampling quality according to the procedures under [Ref. 19, 24 and 34].

When at least 8 independent collections of field duplicates are gathered, evaluation of measurement uncertainty shall be performed by combining individual differences of the sampled duplicates, as shown below. Care should be taken here in which part of the concentration range lie the results that are the object of evaluation of measurement uncertainty from sampling (see also Fig. 1, separate treatment of experiments in different parts of concentration range is necessary).

**Possibility 1:** Homogeneous variances (most often for results in the concentration range up to 10 times the LOD [Ref. 19]).

- Method 1 (from individual standard deviations between duplicates,  $n$  – number of duplicates 8 or more):

$$s(\text{total}) = \sqrt{\frac{\sum s_i^2}{n}} \quad (13)$$

- Method 2 (from D ranks and using a constant of 1.128, which applies to the ratio between  $s$  and  $D$  for duplicate experiments,  $n$  – number of duplicates 8 or more):

$$s(\text{total}) = \frac{\sum \text{ABS}(D)_i}{n * 1,128} \quad (14)$$

**Possibility 2:** Non-homogeneous variances, relative values of standard deviations/ranks (most often for the results in the concentration range higher than 10 times the LOD of the test procedure).

- Method 1 (from individual standard deviations ( $s_r$ ) between duplicates,  $n$  – number of duplicates 8 or more)

$$s_r(\text{total}) = \sqrt{\frac{\sum s_{ri}^2}{n}} \quad (15)$$

- Method 2 (from individual relative ranks ( $d$ ) among duplicates,  $x_1$  and  $x_2$  – results of measurement of duplicates,  $n$  – number of duplicates 8 or more):

$$d = \frac{2 \cdot D}{X_1 + X_2} \quad (16)$$

$$s_r(\text{total}) = \frac{\sum \text{ABS}(d)_i}{n \cdot 1,128} \quad (17)$$

When estimating  $s(\text{laboratory})$  from the validation data, i.e., without using additional experiments such as presented in Fig. 2, the measurement uncertainty from sampling is calculated either from the relative ( $s_r$ ) or absolute standard deviations ( $s$ ), in the following way [Ref. 19]:

$$s(\text{sampling}) = \sqrt{s(\text{total})^2 - s(\text{laboratory})^2} \quad (18)$$

When  $s(\text{laboratory})$  cannot be estimated from validation data, then it is calculated from the difference between the parallels of sample A and sample B, as shown in Fig. 3 and 4, respectively. In doing so, the determined standard deviation divided by  $\sqrt{2}$  is used for  $s(\text{laboratory})$ , as it is used for determining the difference between A and B averages of two determinations [Ref. 34].

In this case, measurement uncertainty is calculated as follows:

$$s(\text{sampling}) = \sqrt{s(\text{total})^2 - \frac{s(\text{laboratory})^2}{2}} \quad (19)$$

Derivation of measurement uncertainty of sampling:

$$s(\text{sampling}) = u(\text{sampling}) \quad (20)$$

$$U(k=2) \text{ of sampling} = 2 \cdot u(\text{sampling}) \quad (21)$$

**$s(\text{total})$**  – standard deviation for within-laboratory reproducibility of sampling and the laboratory part of analyses, determined from 8 independent collections of duplicate field samples

**$s(\text{sampling})$**  – within-laboratory reproducibility of sampling, equivalent to  $u(\text{sampling})$

**$s(\text{laboratory})$**  – standard deviation for within-laboratory reproducibility of the laboratory part of testing

Attention:  $s(\text{laboratory})$  is not equal to  $u_c(\text{laboratory})$ , i.e., the total combined measurement uncertainty of the laboratory part.

$$U(\text{total}, k=2) = 2 \cdot \sqrt{u_c(\text{laboratory})^2 + u(\text{sampling})^2} \quad (22)$$

#### NOTES:

Also, the calculation procedure using ANOVA or RANOVA can be used in estimating the measurement uncertainty arising from sampling, which is more appropriate when we wish to exclude an excessive impact of outliers. [Ref. 19].

All the phases need to be included in the estimation of measurement uncertainty of sampling, i.e. from sampling itself to sample handling and the laboratory sample preparation procedure. Internal control of sampling procedures should be carried out in compliance with [Ref. 27] or an equivalent procedure. Participation in interlaboratory comparisons of sampling is necessary in order to include bias in the estimate of measurement uncertainty from sampling.

## 5 CHANGES WITH REGARD TO PREVIOUS REVISION

References have been updated.

## 6 TRANSITORY PROVISIONS

N/A.

## 7 CONTROL OF THE DOCUMENT

A valid copy of this document shall be located in i4 (SA's information system). A clean copy shall be published on SA's website, and available in printed form at SA's head office.

Individual copies may be controlled in physical form. The recipients or places of storage shall be shown in records on issuance of the document.

Other printouts and copies of the document shall have informative nature and shall not be considered as controlled copies. The validity of these documents should be checked in i4 or on SA's website.

## 8 REFERENCES

1. SIST EN ISO/IEC 17025: 2017; General requirements for the competence of testing and calibration laboratories.
2. Eurachem Guide to Quality in Analytical Chemistry; An aid to accreditation, 3rd edition, 2016.
3. SIST-V ISO/IEC Guide 99:2012 (JCGM 200:2012: International vocabulary of metrology – Basic and general concepts and associated terms, VIM, 2012).
4. Eurachem Guide, The Fitness for Purpose of Analytical Methods; A Laboratory Guide to Method Validation and Related Topics, 2014.
5. ISO/IEC Guide 98:2008 (JCGM 100:2008; Evaluation of measurement data – Guide to the expression of uncertainty in measurement, 2008).
6. SIST ISO 21748:2017; Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty evaluation.
7. SIST ISO 5725-1:2003; Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions.
8. SIST ISO 5725-2:2020; Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method.
9. SIST ISO 5725-3:2003; Accuracy (trueness and precision) of measurement methods and results – Part 3: Intermediate measures of the precision of a standard measurement method.
10. SIST ISO 5725-4:2020; Accuracy (trueness and precision) of measurement methods and results – Part 4: Basic methods for the determination of the trueness of a standard measurement method.
11. SIST ISO 5725-5:2003/C1:2006; Accuracy (trueness and precision) of measurement methods and results – Part 5: Alternative methods for the determination of the precision of a standard measurement method.
12. SIST ISO 5725-6:2003; Accuracy (trueness and precision) of measurement methods and results – Part 6: Use in practice of accuracy values.
13. SIST ISO 3534-1: 2008; Statistics, Vocabulary, Symbols – Part 1: General statistical terms and terms used in probability.
14. SIST EN ISO 9000:2015; Quality management systems – Fundamentals and vocabulary.
15. Estimation of measurement uncertainty in chemical analysis, online course, University of Tartu.
16. Analytical measurement: measurement uncertainty and statistics, 2013.
17. EURACHEM/CITAC Guide, Quantifying Uncertainty in Analytical Measurement, 2012.

18. EUROLAB Technical Report No. 1/2006, Guide to the Evaluation of Measurement Uncertainty for Quantitative Test Results, 2006.
19. EURACHEM/CITAC Guide, Measurement uncertainty arising from sampling. A guide to methods and approaches, 2019.
20. A2LA G104; Guide for Estimation of Measurement Uncertainty in Testing, 2019.
21. NT TR 537, Handbook for calculation of measurement uncertainty in environmental laboratories, 2017.
22. SIST ISO 11352:2013; Water quality – Estimation of measurement uncertainty based on validation and quality control data.
23. SIST EN ISO 5667-1:2007; Water quality – Sampling – Part 1: Guidance on the design of sampling programmes and sampling techniques.
24. SIST EN ISO 5667-14:2017; Water quality – Sampling – Part 14: Guidance on quality assurance and quality control of environmental water sampling and handling.
25. SIST EN ISO 5667-13:2012; Water quality – Sampling – Part 13: Guidance on sampling of sludges.
26. SIST ISO 5667-20:2010; Water quality – Sampling – Part 20: Guidance on the use of sampling data for decision making – Compliance with thresholds and classification systems, Guidance on the use of sampling data for decision making – Compliance with thresholds and classification systems.
27. SIST ISO 5667-24:2018; Water quality – Sampling – Part 24: Guidance on the auditing of water quality sampling.
28. SIST EN 14899:2006; Characterization of waste – Sampling of waste materials – Framework for the preparation and application of a Sampling Plan.
29. ISO 14488:2007; Particulate materials – Sampling and sample splitting for the determination of particulate properties.
30. SIST ISO 6107-2:2007; Water quality – Vocabulary.
31. SIST EN ISO 11074:2015; Soil quality – Vocabulary.
32. <https://www.iso.org/obp/ui> A collection of terms from ISO standards.
33. <http://www.electropedia.org/> A collection of terms from IEC standards.
34. NT TR 604, Uncertainty from sampling; A Nordtest Handbook, 2020.