EXPRESSION OF MEASUREMENT UNCERTAINTY FOR QUANTITATIVE TESTING

Date of implementation: 16.12.2022

HISTORY OF THE DOCUMENT

Revision and date of approval	Reason for revision	Extent of the revision
EA-4/16 **December 2003**	Application of document EA-4/16. The concerned institutions and auditors/experts are requested to take note of this document on the EA website.	
BELAC 2-108 Rev 0 **CC 19.12.2006**	Document EA-4/16, in its original English version, is formally included in the BELAC documents as document BELAC 2-108.	
BELAC 2-108 Rev1 **CC** **01.12.2022**	Revision of the document following the repeal of EA-4/16:2003 and publication of ILAC G17:01/2021 (ILAC Guidelines for Measurement Uncertainty in Testing)	Full document

EXPRESSION OF MEASUREMENT UNCERTAINTY FOR QUANTITATIVE TESTING

1. OBJECTIVES OF THE DOCUMENT AND NORMATIVE REFERENCES

The standard EN ISO/IEC 17025: 2017 (§ 7.6) requires a test laboratory to identify the contributions to the measurement uncertainty and to determine or estimate the measurement uncertainty if the test method precludes a precise determination of the measurement uncertainty.

This document specifies the requirements of the standard regarding the expression of the measurement uncertainty for testing and takes into account the provisions of ILAC G17:2021 and concepts of EA 4/16:2003 (particularly those based on GUM[1]).

.

2. RECIPIENTS

- Members of the Coordination Commission
- Members of the Accreditation Board
- Accreditation secretariat
- Assessors/Experts for the testing laboratories
- Accredited testing laboratories

[1] Evaluation of measurement data — Guide to the expression of uncertainty in measurement (JCGM 100:2008 - GUM 1995 with minor corrections)

3. INTRODUCTION

- The Guide to the Expression of Uncertainty in Measurement (GUM) is internationally recognised as the master document on measurement uncertainty. Therefore, consistency with the GUM is generally required for specific guidance or recommendations for the evaluation of measurement uncertainty in any field of application within accreditation.

- In general, the GUM is also applicable in testing, although there are decisive differences between measurement and testing procedures. The very nature of some testing procedures may make it difficult to apply the GUM strictly. Section 8 provides guidance on how to proceed in such cases.

- Wherever possible, accredited testing laboratories are required, when reporting the uncertainties associated with quantitative results, to do so in accordance with the GUM. A basic requirement of the GUM is the use of a model for the evaluation of uncertainty. The model should include all quantities that can contribute significantly to the uncertainty associated with the test result. There are circumstances, however, where the effort required developing a detailed model is unnecessary. In such a case other identified guidance should be adopted, and other methods based, for example on validation and method performance data be used.

- To ensure that clients benefit fully from laboratories' services, accredited testing laboratories have developed appropriate principles for their collaboration with clients. Clients have the right to expect that the test reports are factually correct, useful and comprehensive. Depending on the situation, clients are also interested in quality features, especially
  - the reliability of the results and a quantitative statement on this reliability, i.e. uncertainty
  - the level of confidence of a conformity statement about the product that can be inferred from the testing result and the associated expanded uncertainty.

  Other quality features such as repeatability, intermediate precision reproducibility, trueness, robustness and selectivity are also important for the characterisation of the quality of a test method.

- Knowledge of the measurement uncertainty of test results is fundamentally important for laboratories, their customers and all parties using and interpreting these results.

4. SCOPE OF APPLICATION

- This document is intended to provide guidance for the evaluation[2] and reporting of measurement uncertainty in quantitative testing. It is applicable to all areas of testing covered by the ILAC Arrangement in testing. This document is also relevant in some parts of medical examination (ISO 15189) as well as other kinds of conformity assessment where testing is performed.

- Estimation of sampling uncertainty is not addressed in the present document. For more information on this topic, other sources can be consulted [13,14,15]

- For the evaluation of uncertainty in calibration, BELAC 2-107 (EA–4/02[3]) shall be consulted.

- This document does not deal with the use of uncertainty in conformity assessment : for that topic reference is made to ILAC G8[4] and JCGM 106[5]. Also guidance from Eurolab and Eurachem is available about this subject [6,7,8].

5. TERMS AND DEFINITIONS

- The terms and definitions of the “International Vocabulary of Metrology – Basic and General Concepts and Associated Terms” (VIM)[9] and ISO/IEC 17025 are applied.

6. BASIC PRINCIPLES

- The statement of uncertainty of measurement should contain sufficient information for comparative purposes;
- The GUM and ISO/IEC 17025 form the basic documents for the evaluation of measurement uncertainty but it has proved necessary to produce sector specific guidance taking due care to the nature of the specific sector. While the GUM principles can be followed in most fields of physical measurements, for microbiological or chemical quantities, more relevant references have been

[2] The term evaluation has been used in preference to the term estimation. The former term is more general and is applicable to different approaches for uncertainty. This choice is also made to be consistent with the vocabulary used in GUM.
[3] EA-4/02:2022 Evaluation of the Uncertainty of Measurement in calibration.
[4] ILAC G8:09/2019 Guidelines on Decision Rules and Statements of Conformity
[5] JCGM 106:2012 Evaluation of measurement data – The role of measurement uncertainty in conformity assessment.
[6] Eurolab Technical report n° 01/2017 : Decision rules applied to conformity assessment
[7] Eurachem/CITAC Guide : Use of Uncertainty Information in Compliance Assessment (2nd Edition, 2021)
[8] Eurolab Cook Book 8: Determination of Conformance with Specifications or Limit Values.
[9] JCGM 200:2012 International vocabulary of metrology – Basic and general concepts and associated terms (VIM) (available from www.BIPM.org)

developed (e.g. EURACHEM/CITAC guidance[10], ISO 19036:2019[11], ISO 29201:2012[12]) Meanwhile also guidances related to measurement uncertainty arising from sampling have been developed (EURACHEM/EUROLAB/CITAC/Nordtest/AMC Guide[13], Nordtest Technical Report 604[14], NEN 7776[15])

- In some areas of testing in which uncertainty cannot be expressed as an expanded uncertainty for the test result (e.g. qualitative testing or examinations) other means for evaluation of measurement uncertainty, such as a probability for false positive or false negative test results, may be more relevant[16].

  For quantitative measurements where the final results are expressed in a qualitative way (e.g. pass/fail), evaluation of measurement uncertainty is still applicable.
- The basic requirement should be either an estimation of the overall uncertainty, or identification of the major components followed by an attempt to estimate their size and the size of the combined uncertainty;
- The basis for the estimation of uncertainty of measurement is to use existing experimental data (quality control charts, validation, round robin tests, PT, CRM…), external data (handbooks, scientific publications, calibration certificates, CRM certificates etc.) or expert judgement ;
- The required depth of the uncertainty estimations may be different in different technical fields. Factors to be taken into account include:
  - common sense;
  - influence of the uncertainty of measurement on the result (appropriateness of the determination);
  - appropriateness;
  - risk level.
- When the estimation of the uncertainty of measurement does not include all relevant sources of uncertainty, any report of the uncertainty should make this clear;

---

[10] Quantifying Uncertainty in Analytical Measurement, 3rd Edition (2012)
[11] ISO 19036:2019 Microbiology of the Food Chain – Estimation of Measurement Uncertainty for Quantitative Determinations
[12] ISO 29201:2012 Water Quality – The Variability of Test Results and the Uncertainty of Measurement of Microbiological Enumeration Methods
[13] EURACHEM/EUROLAB/CITAC/Nordtest/AMC Guide (2019) Measurement uncertainty arising from sampling: A guide to methods and approaches, Second Edition (available from www.eurachem.org)
[14] Nordtest Technical Report 604 (2020) Uncertainty from sampling - A Nordtest Handbook for Sampling Planners on Sampling Quality Assurance and Uncertainty Estimation (available from www.nordtest.info)
[15] NEN 7776:2021 Milieu, voedingsmiddelen, diervoeders, minerale vaste brandstoffen en vaste biobrandstoffen - Bijdrage van bemonstering aan meetonzekerheid
[16] EURACHEM / CITAC Guide Assessment of performance and uncertainty in qualitative chemical analysis (2021)

7. BRIEF SUMMARY OF THE GUM

- The GUM is based on sound theory and provides a consistent and transferable evaluation of measurement uncertainty and supports metrological traceability. The following paragraphs provide a brief interpretation of the basic ideas and concepts.

- Three levels in the GUM can be identified. These are basic concepts, recommendations and evaluation procedures. Consistency requires the basic concepts to be accepted and the recommendations to be followed.

- The basic evaluation procedure presented in the GUM, the law of propagation of uncertainty, applies to linear or linearised models.It should be applied whenever appropriate, since it is straightforward and easy to implement. However, for some cases more advanced methods such as the use of higher-order expansion of the model or the propagation of probability distributions may be required.

  - The basic concepts in uncertainty evaluation are
    - the knowledge about any quantity that influences the measurand is in principle incomplete and can be expressed by a probability density function (PDF) for the values attributable to the quantity based on that knowledge
    - the expectation value of that PDF is taken as the best estimate of the value of the quantity
    - the standard deviation of that PDF is taken as the standard uncertainty associated with that estimate
    - the PDF is based on knowledge about a quantity that may be inferred from
      - repeated measurements—Type A evaluation
      - scientific judgement based on all the available information on the possible variability of the quantity—Type B evaluation.

  - The evaluation procedure comprises four parts:
    - Derivation of the model of the measurement. Because in general this is the most difficult part of the evaluation, the use of a cause-effect-relationship linking the input quantities to the measurand is recommended. A guide concerned with the development of a measurement model and the practical use of the model has been developed by the Joint Committee for Guides in Metrology[17]. The GUM concentrates on measurement models in the form of measurement functions having a single output quantity. The extension to any number of output quantities in a general measurement model is treated in JCGM 102[18].

---

[17] JCGM GUM-6:2020 Guide to the expression of uncertainty in measurement — Part 6: Developing and using measurement models (10)

[18] JCGM 102:2011 Evaluation of measurement data – Supplement 2 to the "Guide to the expression of uncertainty in measurement" – Models with any number of output quantities,

- The provision of probability density functions (PDFs) for the input quantities to the model, given information about these quantities. In many cases in practice, it is necessary to specify only the expectation value and standard deviation of each PDF, i.e. the best estimate of each quantity and the standard uncertainty associated with that estimate. The assignment of probability distributions to the input quantities in a measurement model is considered in JCGM 101 and JCGM 102 [19,18]
- Propagation of uncertainty. The basic procedure (the law of propagation of uncertainty) can be applied to linear or linearised models, but is subject to some restrictions. JCGM 101:2008 provides detailed information on using Monte Carlo Method as an implementation of the propagation of distributions
- Stating the complete result of a measurement by providing the best estimate of the value of the measurand, the combined standard uncertainty associated with that estimate and an expanded uncertainty (Section 10).

- The GUM provides guidance on stating a complete result of a measurement in its section 7, titled "Reporting uncertainty". Section 10 in this document follows the recommendations of the GUM and ILAC G17 and provides some more detailed guidance. Note that the GUM permits the use of either the combined standard uncertainty $u_c(y)$ or the expanded uncertainty $U(y)$, i.e. the half width of an interval having a stated level of confidence, as a measure of uncertainty. However, if the expanded uncertainty is used, one must state the coverage factor k, which is equal to the value of $U(y)/u_c(y)$.

- For the evaluation of the uncertainty associated with the measurand Y one needs only to know
  - the model, $Y = f(X_1,..., X_N)$,
  - the best estimates x_i of all input quantities X_i and
  - the uncertainties $u(x_i)$ and the correlation coefficients $r(x_i,x_j)$ associated with x_i and with x_i and x_j.

  The best estimate x_i is the expected value of the PDF for X_i, $u(x_i)$ is the standard deviation of that PDF and $r(x_i,x_j)$ is the ratio of the covariance between x_i and x_j and the product of the standard deviations.

  To state the combined standard uncertainty $u_c(y)$ associated with the measurement result y, no further knowledge of the PDF is required. To state the half width of an interval having a stated level of confidence, i.e. an expanded uncertainty, it is necessary to know the PDF. This requires more knowledge since the two parameters, expectation value and standard deviation, do not fully specify a PDF unless it is known to be Gaussian.

[19] JCGM 101:2008 Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method (see clause 2),

Section 10 provides guidance on obtaining the expanded uncertainty in those cases where a Gaussian PDF is not assumed for the measurand Y.

8. USE OF VALIDATION AND METHOD PERFORMANCE DATA FOR UNCERTAINTY EVALUATION

8.1. SOURCES OF METHOD PERFORMANCE AND VALIDATION DATA

- The observed performance characteristics of test methods are often essential in evaluating the uncertainty associated with the results (Section 7). This is particularly true where the results are subject to important and unpredictable effects, which can best be considered as random effects, or where the development of a comprehensive mathematical model is impractical. Method performance data also very frequently includes the effect of several sources of uncertainty simultaneously and its use may accordingly simplify considerably the process of uncertainty evaluation. Information on test method performance is typically obtained from
    - data accumulated during validation and verification of a test method prior to its application in the testing environment
    - interlaboratory studies according to the ISO 5725 series
    - accumulated quality control (that is, check sample) data
    - proficiency testing schemes

- This section provides general guidance on the application of data from each of these sources.

8.2. DATA ACCUMULATED DURING VALIDATION AND VERIFICATION OF A TEST METHOD PRIOR TO APPLICATION IN THE TESTING ENVIRONMENT

- In practice, the fitness for purpose of test methods applied for routine testing is frequently checked through method validation and verification studies. The data so accumulated can inform the evaluation of uncertainty for test methods. Validation studies for quantitative test methods typically determine some or all of the following parameters:

    - **Precision**. Studies within a laboratory will obtain precision under repeatability conditions and intermediate conditions, ideally over time and across different operators and types of test item. The observed precision of a testing procedure is an essential component of overall uncertainty, whether determined by a combination of individual variances or by a study of the complete method in operation.

- **Bias**. The bias of a test method is usually determined by studying relevant reference materials or test samples. The aim is typically to identify and eliminate significant bias. In general, the uncertainty associated with the determination of the bias is an important component of overall uncertainty.

- **Linearity**. Linearity is an important property of methods used to make measurements over a range of values. Correction for significant non-linearity is often accomplished by the use of non-linear calibration functions. Alternatively, the effect is avoided by the choice of a restricted operating range. Any remaining deviations from linearity are normally sufficiently accounted for by the use of overall precision data. If these deviations are negligible compared with the uncertainties associated with calibration, additional uncertainty evaluation is not required.

- **Capability of detection**. The lower limit of operability of a test method may be established. The value obtained is not directly relevant to the evaluation of uncertainty. The uncertainty in the region at or near this lower limit is likely to be significant compared with the value of the result, leading to practical difficulties in assessing and reporting uncertainty. Reference to appropriate documentation on the treatment and reporting of results in this region is accordingly recommended[21].

- **Selectivity and specificity**. These terms relate to the ability of a test method to respond to the appropriate measurand in the presence of interfering influences, and are particularly important in chemical testing. They are, however, qualitative concepts and do not directly provide uncertainty information, though the influence of interfering effects may in principle be used in uncertainty evaluation[20] [12].

- **Robustness or ruggedness**. Many method development or validation protocols require that the sensitivity to particular parameters be investigated directly. Ruggedness data can therefore provide information on the effect of important parameters, and is particularly important in establishing whether a given effect is significant[21].

- Experimental studies of method performance should be carried out carefully. In particular:

  - Representativeness is essential: as far as possible, studies should be conducted to provide a realistic survey of the number and range of effects operating during normal use of the method, as well as covering the range of

[20] EURACHEM / CITAC Guide CG 4, Quantifying Uncertainty in Analytical Measurement (third edition) 2012
[21] EURACHEM, The Fitness for Purpose of Analytical Methods (ISBN 0- 948926-12-0) 1998

values and sample types within the scope of the method. Estimates of precision covering a wide variety of sources of variation are particularly appropriate in this respect.
    - Where factors are suspected to interact, the effect of interaction should be taken into account. This may be achieved either by ensuring random selection from different levels of interacting parameters, or by careful systematic design to obtain both variance and covariance information.
    - In carrying out studies of overall bias, it is important that the reference materials and values are relevant to the materials under routine test.

- Careful experimental design is accordingly invaluable in ensuring that all relevant factors are duly considered and properly evaluated.

- The general principles of applying validation and performance data to uncertainty evaluation are similar to those applicable to the use of performance data (above). However, it is likely that the performance data available will adequately cover fewer contributions. Correspondingly further supplementary estimates will be required. A typical procedure is:
    - Compile a list of relevant sources of uncertainty. It is usually convenient to include any measured quantities held constant during a test, and to incorporate appropriate precision terms to account for the variability of individual measurements or the test method as a whole. A cause and effect diagram[21] is a very convenient way to summarise the uncertainty sources, showing how they relate to each other and indicating their influence on the uncertainty associated with the result
    - Assemble the available method performance and calibration data
    - Check to see which sources of uncertainty are adequately accounted for by the available data. It is not generally necessary to obtain separately the effects of all contributions; where several effects contribute to an overall performance figure, all such effects may be considered to be accounted for. Precision data covering a wide variety of sources of variation are therefore particularly useful as they will often encompass many effects simultaneously (but note that in general precision data alone are insufficient unless all other factors are assessed and shown to be negligible)
    - For any sources of uncertainty not adequately covered by existing data, either seek additional information from the literature or existing data (certificates, equipment specifications, etc.) or, plan experiments to obtain the required additional data.

8.3. INTERLABORATORY STUDY OF TEST METHODS PERFORMANCE ACCORDING TO THE ISO 5725 SERIES OR EQUIVALENT

- Interlaboratory studies according to the ISO 5725 series typically provide the repeatability standard deviation s_r and reproducibility standard deviation s_R (both as defined in ISO 3534-1[22]) and may also provide an estimate of trueness (measured as bias with respect to a known reference value). The application of these data to the evaluation of uncertainty in testing is discussed in detail in ISO 21748[23] . The general principles are:

  i) Establishing the relevance of method performance data to measurement results from a particular measurement process. Section 9.2 of this document provides details of the measures required.
  ii) Establishing the relevance of method performance data to the test item by identifying differences in sample treatment, sampling, or expected level of response between the laboratory's test item and those test items examined in a collaborative study. An adjustment of the reproducibility standard deviation to take account of, for example, changes in precision with level of response may be necessary.
  iii) Identifying and evaluating the additional uncertainties associated with factors not adequately covered by the interlaboratory study (see further).
  iv) Using the principles of the GUM to combine all the significant contributions to uncertainty, including the reproducibility standard deviation (adjusted if necessary), any uncertainty associated with the laboratory component of bias for the test method, and uncertainties arising from additional effects identified in iii).

- These principles are applicable to test methods that have been subjected to interlaboratory study. For these cases, reference to ISO 21748 is recommended for details of the relevant procedure. The EURACHEM/CITAC guide[20] also gives guidance on the application of interlaboratory study data in chemical testing.

- The additional sources that may need particular consideration are:
  - **Sampling**. Collaborative studies rarely include a sampling step. If the method used in-house involves sub-sampling, or the measurand is a bulk property of a small sample, the effects of sampling should be investigated and their effects included
  - **Pre-treatment.** In most studies, samples are homogenised, and may additionally be stabilised, before distribution. It may be necessary to investigate and add the effects of the particular pre-treatment procedures applied in-house

[22] ISO 3534-1:2006 Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability
[23] ISO 21748:2017 Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty evaluation

- **Method bias.** Method bias is often examined prior to or during interlaboratory study, where possible by comparison with reference methods or materials. Where the bias itself, the standard uncertainties associated with the reference values used, and the standard uncertainty associated with the estimated bias are all small compared with the reproducibility standard deviation, no additional allowance need be made for the uncertainty associated with method bias. Otherwise, it will be necessary to make such allowance.
- **Variation in conditions.** Laboratories participating in a study may tend to steer their results towards the means of the ranges of the experimental conditions, resulting in underestimates of the ranges of results possible within the method definition. Where such effects have been investigated and shown to be insignificant across their full permitted range, however, no further allowance is required.
- **Changes in sample type.** The uncertainty arising from samples with properties outside the range covered by the study will need to be considered.

8.4. TEST OR MEASUREMENT PROCESS QUALITY CONTROL DATA

- Many test or measurement processes are subject to control checks based on periodic measurement of a stable, but otherwise typical, test item to identify significant deviations from normal operation. Data obtained in this way over a long period of time provide a valuable source of data for uncertainty evaluation. The standard deviation of such a data set provides a combined estimate of variability arising from many potential sources of variation. It follows that if applied in the same way as method performance data (above), the standard deviation provides the basis for an uncertainty evaluation that immediately accounts for the majority of the variability that would otherwise require evaluation from separate effects.

- Quality control (QC) data of this kind will not generally include sub-sampling, the effect of differences between test items, the effects of changes in the level of response, or inhomogeneity in test items. QC data should accordingly be applied with caution to similar materials, and with due allowance for additional effects that may reasonably apply.

- Data points from QC data that gave rise to rejection of measurement and test results and to corrective action should normally be eliminated from the data set before calculating the standard deviation.

8.5. PROFICIENCY TESTING DATA

- Proficiency tests are intended to check periodically the overall performance of a laboratory, and are best used for that purpose . A laboratory's results from its participation in proficiency tests can accordingly be used to check the evaluated uncertainty, since that uncertainty should be compatible with the spread of results obtained by that laboratory over a number of proficiency test rounds.

- In general, proficiency tests are not carried out sufficiently frequently to provide good estimates of the performance of an individual laboratory's implementation of a test method. Additionally, the nature of the test items circulated will typically vary, as will the expected result. It is thus difficult to accumulate representative data for well- characterised test items. Furthermore, many schemes use consensus values to assess laboratory performance, which occasionally lead to apparently anomalous results for individual laboratories. Their use for the evaluation of uncertainty is accordingly limited. However, in the special case where
    - the types of test items used in the scheme are appropriate to the types tested routinely
    - the assigned values in each round are traceable to appropriate reference values, and
    - the uncertainty associated with the assigned value is small compared with the observed spread of results,

  the dispersion of the differences between the reported values and the assigned values obtained in repeated rounds provides a basis for an evaluation of the uncertainty arising from those parts of the measurement procedure within the scope of the scheme.

- Systematic deviation from traceable assigned values and any other sources of uncertainty (such as those noted in connection with the use of interlaboratory study data obtained in accordance with the ISO 5725 series) must also be taken into account.

- It is recognised that the above approach is relatively restricted. Proficiency testing data may have wider applicability in providing a preliminary estimate of uncertainty in some circumstances.

8.6. SIGNIFICANCE OF UNCERTAINTY CONTRIBUTIONS

- Not all the uncertainty sources identified during an uncertainty evaluation will make a significant contribution to the combined uncertainty; indeed, in practice it is likely that only a small number will. Those few clearly need careful study to obtain reliable estimates of their contributions. A preliminary estimate of the contribution of each component or combination of components to the uncertainty should therefore be made, by judgement if necessary, and attention paid to those that are most significant.

- In deciding whether an uncertainty contribution can be neglected, it is important to consider
  - The relative sizes of the largest and the smaller contributions. For example, a contribution that is one fifth of the largest contribution will contribute at most 2% of the combined standard uncertainty
  - The effect on the reported uncertainty. It is imprudent to make approximations that materially affect the reported uncertainty or the interpretation of the result
  - The degree of rigour justified for the uncertainty evaluation, taking into account the client and regulatory and other external requirements identified, for example, during contract review.

8.7. USE OF PRIOR STUDY DATA

- In order to use the results of prior studies of the method to evaluate the uncertainty, it is necessary to demonstrate the validity of applying prior study results. Typically, this will consist of:
  - Demonstration that a precision comparable to that obtained previously can be achieved
  - Demonstration that the use of the bias data obtained previously is justified, typically through the determination of bias on relevant reference materials (see, for example, ISO Guide 33[24]), by satisfactory performance on relevant proficiency schemes, or other interlaboratory comparisons
  - Continued performance within statistical control as shown by regular QC sample results and the implementation of effective analytical quality assurance procedures.

- Where the conditions above are met, and the method is operated within its scope and field of application, it is normally acceptable to apply the data from prior studies (including validation studies) directly to uncertainty evaluations in the laboratory in question.

- For methods operating within their defined scope, when the reconciliation stage shows that all the identified sources have been included in the validation study or when the contributions from any remaining sources have been shown to be negligible, the reproducibility standard deviation sR may be used as the combined standard uncertainty.

[24] ISO Guide 33:2015 Reference materials — Good practice in using reference materials

- If there are any significant sources of uncertainty that are not included in the validation study their contribution is evaluated separately and combined with sR to obtain the overall uncertainty.

9. REPORTING RESULTS OF A TEST

- ISO/IEC 17025:2017 §7.8.3.1 c) requires laboratories to include measurement uncertainty (presented in the same unit as that of the measurand or in a term relative to the measurand (e.g. percent)) in the test report when:
  - it is relevant to the validity or application of the test results;
  - a customer's instruction so requires, or
  - the measurement uncertainty affects conformity to a specification limit.

  Laboratories are encouraged to evaluate carefully the situations where reporting measurement uncertainty can help the interpretation of test results, in order to conform to 7.8.3.1 c).

  In the following examples, it will normally be necessary to report measurement uncertainty in order to comply with 7.8.3.1 c), if the laboratory is not required to report a statement of conformity:
    - *Environmental tests conducted regularly and where conformity to a specification limit is assessed by the customers. Such cases may be mandated by legislation or be voluntary. In order for customers to assess if a test parameter is subject to change and poses a risk for not complying with the regulation, the measurement uncertainty needs to be known. The measurement uncertainty is necessary for the customers to make a qualified decision, e.g., on changes to their water or waste water treatment facilities.*
    - *Product tests where a product is tested for conformity to a specification. In such cases the test result may be quantitative as well as pass/fail. In both cases the reporting of measurement uncertainty should be important for a customer to assess the risk of product failure for an item near the specification limit. This is particularly relevant if the customer is the product manufacturer.*

  It is however recognized that there are situations where the requirement for reporting of measurement uncertainty may not be obvious, e.g., the laboratory cannot be sure about the end use of the test results and the customer also does not explicitly require MU to be reported. In such cases, customary reporting of measurement uncertainty in testing can help the laboratory to fulfil its responsibility under ISO/IEC 17025:2017.

- Customary reporting of measurement uncertainty in testing has several advantages:
  - Only after taking measurement uncertainty into account, a deviation between two test results can objectively be judged to be compliant or non-compliant.

  - Reporting measurement uncertainty allows users to assess if the test results are fit for purpose (i.e. if measurement uncertainty is adequately low or smaller than the target measurement uncertainty).
  - The need for repetitive and redundant tests is reduced when reported measurement uncertainties are initially taken into account.
  - Reported measurement uncertainties provide information of the performance of a test method both in a laboratory and across laboratories and allows for development and improvement of standardized methods.
  - Laboratories will not on a case-by-case basis be asked by their customers for additional information of measurement uncertainties and will not have to determine when the measurement uncertainty is necessary for interpretation of test results and when it is not.
  - Customary reporting consolidates measurement uncertainty evaluation

- When customary reporting of measurement uncertainty in testing is not made, it will be important to assess how the borderlines between reporting and non-reporting of measurement uncertainty are established. Such borderlines may be connected to a decision rule [4, 5, 6, 7]

- Laboratories are encouraged to discuss with their stakeholders and regulators the intended use of the reported results and the relevance of evaluating and/or reporting measurement uncertainty. Laboratories should also take into account requirements of stakeholders of regulators when evaluating and/or reporting measurement uncertainty.

- The methods used to calculate the result and its uncertainty shall be available including:
  - Sufficient documentation of the data analysis to enable a repetition of any calculation if necessary and to show how the uncertainty is calculated;
  - All corrections and constants used in the analysis, and their sources;

- Whenever either a component of measurement uncertainty, including that arising from sampling, cannot be reasonably evaluated or the relevant requirement is not applicable then this should be clarified in the test report. For example, in the case of sampling, the disclaimer may be: “The measurement uncertainty arising from sampling is not included in the expanded measurement uncertainty”.

- ISO/IEC 17025 §7.8.2.1 l) requires that a test report contains a clear statement to the effect that the results relate only to the items tested, calibrated or sampled. To extrapolate those results beyond the item tested, for example where the item is a representative sample taken from a batch, this shall only be done by means of opinions and interpretations. In this case, the laboratory shall ensure that :

- this is adequately discussed, defined, documented and understood during contract review
  - the sampling is performed under accreditation and an estimation of sampling uncertainty is available
  - the sampling is representative for the batch
  - the ISO/IEC 17025 requirements related to opinions and interpretations are met

- The number of decimal digits in a reported uncertainty should always reflect practical measurement capability. In view of the process for evaluating uncertainties, it is rarely justified to report more than two significant digits. Often a single significant digit is appropriate. Similarly, the numerical value of the result should be rounded so that the last decimal digit corresponds to the last digit of the uncertainty. The normal rules of rounding can be applied in both cases. For example, if a result of 123.456 units is obtained, and an uncertainty of 2.27 units has resulted from the evaluation, the use of two significant decimal digits would give the rounded values 123.5 units ± 2.3 units. Rounding should always be carried out at the end of the process in order to avoid the effects of cumulative rounding errors.

- When measurement uncertainty is reported, it should normally be the expanded measurement uncertainty based on the coverage probability of approximately 95% and the coverage factor k needed to achieve the probability. It is understood that coverage probabilities other than 95% may be better suited to particular circumstance. To this, an explanatory note should be added, which may have the following content: "*The reported expanded measurement uncertainty is stated as the combined standard measurement uncertainty multiplied by the coverage factor k = [value used] such that the coverage probability corresponds to approximately [the desired coverage probability]%.*".
  This statement will depend on the nature of the probability distribution; some examples are presented below. All clauses below that relate to a 95% level of confidence require modification if a different level of confidence is required.

  Normal distribution

  It is generally safe to assume a normal distribution from the viewpoint of providing a coverage interval at the 95% level of confidence when the model is linear in the input quantities and one of the following three possibilities applies:

  1. There is a single, dominant contribution to the uncertainty, which arises from a normal distribution, and the corresponding degrees of freedom exceed 30.
  2. The three largest uncertainty contributions are of comparable size.
  3. The three largest contributions are of comparable size, and the effective degrees of freedom[25] exceed 30.

[25] The effective degree of freedom can be estimated by one of the following:

Under these circumstances the following statement can be made:
The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor k = 2, which for a normal distribution provides a level of confidence of approximately [*] 95%.

[*] For the purposes of this document the term approximately is interpreted as meaning effectively or for most practical purposes.

Note: Normality should NOT be assumed if the measurement model is significantly non-linear in the region of interest, particularly if uncertainties in input values are large compared with the input values themselves. Under these circumstances, reference to more advanced texts, e.g. the GUM, is necessary.

t-distribution

The t-distribution may be assumed if the conditions for normality (above) apply but the degrees of freedom is less than 30. Under these circumstances the following statement (in which the appropriate numerical values are substituted for XX and YY) can be made:
The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor k = XX, which for a t-distribution with ν_{eff} = YY effective degrees of freedom provides a level of confidence of approximately 95%.

Dominant (non-normal) contributions in a Type B evaluation of uncertainty

If the uncertainty associated with the measurement result is dominated by a contribution resulting from an input quantity that is non-normal and that contribution is so large that a normal or t-distribution is not obtained when the quantity is convolved with the remaining input quantities, special consideration should be given to obtaining a coverage factor that will provide a level of confidence of approximately 95%.

For an additive model, i.e. when the measurand can be expressed as a linear combination of the input quantities, the PDF for the measurand can be obtained by convolving, i.e. propagating, the PDFs for the input quantities. Even in this case, and almost always when the model is non-linear, the mathematics required can, however, be difficult. A practical approach is to make the assumption that the resulting distribution will be little different in form from that of the dominant component.
In many cases a rectangular distribution will be assigned to a dominant non-normal input quantity. In such a case a rectangular distribution can then be assigned to the measurand. An expanded uncertainty at the 95% level of

- taking the effective degree of freedom for a single, dominant contribution
- using the Welch-Satterthwaite formula given in the GUM and EA-4/02
- (approximately) by taking the number of degrees of freedom for the largest contribution.

confidence can be obtained by multiplying the combined uncertainty by 0.95√3 = 1.65.

Under these circumstances the following statement can be made:
The reported expanded uncertainty is dominated by a single component of uncertainty for which a rectangular probability distribution has been assumed. A coverage factor of 1.65 (= 0.95√3) has therefore been used in order to provide a level of confidence of approximately 95%.

- The test result can usually be expressed as y ± U. However there may be situations where the upper and lower bounds are different; for example if cosine errors are involved. If such differences are small then the most practical approach is to report the expanded uncertainty as ± the larger of the two. However, if there is a significant difference between the upper and lower values they should be evaluated and reported separately. This may be achieved, for example, by determining the shortest coverage interval at the desired level of confidence in the PDF for the measurand. For example, for an uncertainty of +6.5 units and –6.7 units, for practical purposes ± 6.7 units could simply be stated. However, if the values were +6.5 units and –9.8 units they should be separated, e.g. +6.5 units; –9.8 units.

10. STEPWISE IMPLEMENTATION OF THE UNCERTAINTY CONCEPT

- It is recognised that the knowledge of mathematical modelling and the determination of the various influence factors is generally different in different testing fields. This aspect has to be taken into account when implementing ISO/IEC 17025.

- Laboratories cannot in general be expected to initiate scientific research to assess the uncertainties associated with their measurements and tests. The respective requirements should be adapted according to the current state of knowledge in the respective testing field. If a mathematical model as a basis for the evaluation of measurement uncertainty is not available, laboratories can
    - list those quantities and parameters that are expected to have a significant influence on the uncertainty and estimate their contribution to the overall uncertainty
    - use data concerning repeatability or reproducibility that might be available from validation, internal quality assurance or interlaboratory comparisons
    - refer to data or procedures given in the relevant testing standards
    - combine the approaches mentioned above.

- Laboratories should strive to refine their uncertainty evaluations, where appropriate, taking into account for instance
    - recent data from internal quality assurance in order to broaden the statistical basis for the uncertainty evaluation
    - new data from the participation in interlaboratory comparisons or proficiency tests
    - revisions of the relevant standards
    - specific guidance documents for the respective testing field.

- Laboratories should select the most suitable approach for their area and evaluate measurement uncertainty to the extent appropriate to the intended use.

11. ADVANTAGES OF UNCERTAINTY EVALUATION FOR TESTING LABORATORIES

There are several advantages linked with the evaluation of measurement uncertainty in testing, although the task can be time-consuming:

- Measurement uncertainty assists in a quantitative manner in important issues such as risk control and the credibility of test results
- A statement of measurement uncertainty can represent a direct competitive advantage by adding value and meaning to the result
- The knowledge of quantitative effects of single quantities on the test result improves the reliability of the test procedure. Corrective measures may be implemented more efficiently and hence become more cost-effective
- The evaluation of measurement uncertainty provides starting points for optimising the test procedures through a better understanding of the test process
- Clients such as product certification bodies need information on the uncertainty associated with results when stating compliance with specifications
- Calibration costs can be reduced if it can be shown from the evaluation that particular influence quantities do not substantially contribute to the uncertainty.

ANNEX : LIST OF OTHER USEFUL RESOURCES - INFORMATIVE

- AFNOR FD X 07-021 Métrologie et application de la statistique – Aide à la démarche pour l'évaluation et l'utilisation de l'incertitude des mesures et des résultats d'essais (1999) (Aid in the procedure for estimating and using uncertainty in measurements and test results
- Adriaan M.H. van der Veen and Maurice G. Cox Good practice in evaluating measurement uncertainty. Compendium of examples (2021)
- Ellison, S. ISO uncertainty and collaborative trial data. Accred Qual Assur 3, 95–100 (1998)
- Ellison, S., Barwick, V. Estimating measurement uncertainty: reconciliation using a cause and effect approach. Accred Qual Assur 3, 101–105 (1998).
- Ellison S. Uncertainties in qualitative testing and analysis Accred Qual Assur (2000) 5:346–348
- Tsimillis, K.C. Measurement uncertainty: requirements set in the accreditation standards. Accred Qual Assur 23, 109–114 (2018)
- Ríos, A., Barceló, D., Buydens, L. et al. Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'. Accred Qual Assur 8, 68–77 (2003)
- IFCC-IUPAC Recommendations 2017 Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences, Pure Appl. Chem. 90 (2018) 913–935
- EURACHEM/EUROLAB/CITAC/Nordtest/AMC Guide (2019) Measurement uncertainty arising from sampling: A guide to methods and approaches, Second Edition (available from www.eurachem.org)
- EURACHEM / CITAC Guide Setting and Using Target Uncertainty in Chemical Measurement (2015)Nordtest Technical Report 604 (2020) Uncertainty from sampling - A Nordtest Handbook for Sampling Planners on Sampling Quality Assurance and Uncertainty Estimation (available from www.nordtest.info)
- Nordtest Technical report 537 Handbook for calculation of measurement uncertainty in environmental laboratories (2017)
- http://mechem.rd.ciencias.ulisboa.pt/ms-excel-spreadsheet-for-automatic-selection-of-significant-digits/
- UKAS M3003 The Expression of Uncertainty and Confidence in Measurement (Edition 4: October 2019) (available from www.ukas.com)
- UKAS LAB12 The Expression of Uncertainty in Testing (Edition 4, April 2022) (available from www.ukas.com)
- DAKKS Guidance for estimation of measurement uncertainty according to the requirements of DIN EN ISO/IEC 17025 for testing laboratories in the subject of chemical analytics in the fields of health-related consumer protection, agricultural sector, chemistry and environment (71 SD 4 016_e | Revision 1.0| 19 January 2017)
- COFRAC LAB GTA 86 Recommandations pour la mise en œuvre de la norme NF EN ISO/IEC 17025 en vue de l'accréditation des Laboratoires
- EA-4/02 Expression of the uncertainty of measurement in Calibration
- IEC Guide 115 Application of uncertainty of measurement to conformity assessment activities in the electrotechnical sector
- EN ISO 20988:2007 Air quality — Guidelines for estimating measurement uncertainty
- EUROLAB Technical report 1/2002 Measurement uncertainty in testing
- EUROLAB Technical report 1/2007 Measurement uncertainty revisited : alternative approaches to uncertainty evaluation
- ISO 11453 Statistical interpretation of data - Tests and confidence intervals relating to proportions (1996)
- ISO 13752 Air quality - Assessment of uncertainty of a measurement method under field conditions using a second method as reference (1998)
- ISO 11352Water quality — Estimation of measurement uncertainty based on validation and quality control data (2012)
- ISO 16269-7 Statistical interpretation of data - Part 7: Median - Estimation and confidence interval (2001)
- ISO 3534-2 Statistics - Vocabulary and symbols - Part 2: Applied statistics (2006)
- ISO 3534-3 Statistics - Vocabulary and symbols - Part 3: Design of experiments (2013)
- ISO 5479 Statistical interpretation of data - Tests for departure from the normal distribution (1997)
- ISO 6879 Air quality - Performance characteristics and related concepts for air quality measuring methods (1995)
- ISO 8466-1 Water quality - Calibration and evaluation of analytical methods - Part 1: Linear calibration function (2021)

- ISO 8466-2 Water quality - Calibration and evaluation of analytical methods and estimation of performance characteristics - Part 2:Calibration strategy for non-linear second order calibration functions (2001)
- ISO 9169 Air quality - Definition and determination of performance characteristics of an automatic measuring system (2006)
- ISO 11222 Air quality – Determination of the uncertainty of the time average of air quality measurements (2002)
- ISO 14956 Air quality — Evaluation of the suitability of a measurement procedure by comparison with a required measurement uncertainty (2002)
- ISO 10017 Guidance on statistical techniques for ISO 9001:2015 (2021)
- ISO/TS 13530 Water quality - Guidance on analytical quality control for chemical and physicochemical water analysis(2009)
- ISO 13843 Water quality - Requirements for establishing performance characteristics of quantitative microbiological methods (2017)
- ISO/TR 20461 Determination of uncertainty for volume measurements made using the gravimetric method (2000)
- ISO 5168 Measurement of fluid flow - Procedures for the evaluation of uncertainties (2005)
- NEN 7777 Milieu en voedingsmiddelen - Prestatiekenmerken van meetmethoden (2011)/C1 (2012)
- NEN 7778 Milieu - Gelijkwaardigheid van meetmethoden (2003)/C1 (2014)
- FD V 03-116 Analyse des produits agricoles et alimentaires. Guide d'application des données métrologiques (AFNOR) (2001)
- NIST Technical Note 1297 Guidelines for evaluating and expressing uncertainty of NIST measurement results
- NEN 7779 Milieu, voedingsmiddelen en diervoeders - Meetonzekerheid (2018)
- IANZ AS TG 5 Chemical and Microbiological Testing Laboratories Measurement Uncertainty, Precision and Limits of Detection
- SAC Technical Guide 2 - A Guide on Measurement Uncertainty in Chemical & Microbiological Analysis (2019)
- NATA Specific Accreditation Guidance - Infrastructure and Asset Integrity - Measurement Uncertainty in Geotechnical Testing (2018)
- ISO/IEC GUIDE 98-1 Uncertainty of measurement — Part 1: Introduction to the expression of uncertainty in measurement (2009)
- ISO/IEC GUIDE 98-3:2008/SUPPL 1:2008/COR 1:2009 Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) (2008) — Suppl 2008
- ISO/IEC GUIDE 98-3 Uncertainty of measurement — Part 3: guide to the expression of uncertainty in measurement (GUM:1995) — supplement 2: extension to any number of output quantities (2008)/Suppl 2011
- ISO/IEC GUIDE 98-4 Uncertainty of measurement — Part 4: role of measurement uncertainty in conformity assessment (2012)
- ISO/IEC Guide 98-6 Uncertainty of measurement — Part 6: developing and using measurement models (2021)
