PD CEN/TS 16800:2015

BSI Standards Publication

Guideline for the validation of physico-chemical analytical methods

bsi.

...making excellence a habit.™

National foreword

This British Standard is the UK implementation of CEN/TS 16800:2015.

The UK participation in its preparation was entrusted to Technical Committee EH/3/2, Water analysis.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

Published by BSI Standards Limited 2015

ISBN 978 0 580 87777 3

ICS 13.060.50

Compliance with a British Standard cannot confer immunity from legal obligations.

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 December 2015.

Amendments/corrigenda issued since publication

Date	Text affected

TECHNICAL SPECIFICATION

SPÉCIFICATION TECHNIQUE

TECHNISCHE SPEZIFIKATION

CEN/TS 16800

December 2015

ICS 13.060.50

English Version

Guideline for the validation of physico-chemical analytical methods

Lignes directrices pour la validation des méthodes d'analyse physico-chimiques

Anleitung zur Validierung physikalisch-chemischer Analysenverfahren

This Technical Specification (CEN/TS) was approved by CEN on 14 March 2015 for provisional application.

The period of validity of this CEN/TS is limited initially to three years. After two years the members of CEN will be requested to submit their comments, particularly on the question whether the CEN/TS can be converted into a European Standard.

CEN members are required to announce the existence of this CEN/TS in the same way as for an EN and to make the CEN/TS available promptly at national level in an appropriate form. It is permissible to keep conflicting national standards in force (in parallel to the CEN/TS) until the final decision about the possible conversion of the CEN/TS into an EN is reached.

CEN members are the national standards bodies of Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom.

EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

CEN-CENELEC Management Centre: Avenue Marnix 17, B-1000 Brussels

Ref. No. CEN/TS 16800:2015 E

Contents

Page

European foreword

This document (CEN/TS 16800:2015) has been prepared by Technical Committee CEN/TC 230 “Water analysis”, the secretariat of which is held by DIN.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. CEN [and/or CENELEC] shall not be held responsible for identifying any or all such patent rights.

According to the CEN-CENELEC Internal Regulations, the national standards organizations of the following countries are bound to announce this Technical Specification: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom.

Introduction

Environmental monitoring of chemical substances is increasingly carried out within a European framework, and there is concern about the comparability of data at the European level. In particular methods used for the monitoring of substances with recent interest have often not been properly validated either in-house (i.e. within a single laboratory) or at the international level.

These issues may be addressed by adopting a harmonized approach towards method development and validation. The main objective of this document is to provide a common European approach to the validation of chemical methods for the respective monitoring of chemical substances in a broad range of matrices. Although the development of this approach was triggered by the needs for monitoring of emerging pollutants, it is of general nature and can be applied to the measurement of a wide range of substances in a variety of matrices.

This guidance takes into account the different requirements for the level of method maturity and validation at different stages of the investigation or regulation of chemical substances.

In the case of a specific monitoring task, this protocol will guide the user through the following steps:

— classification of existing methods with respect to their status of validation, and the selection of the appropriate validation approach;

— development of a method so as to extend its application; for example, if a method for determining a required target compound in a particular matrix is available, but is not suitable for the same compound in a different matrix of interest;

— the validation procedures to be carried out in order to effectively demonstrate the validation status of a selected method according to the two approaches adopted.

Many (national and international) standards currently contain in their scope a statement like “this method is applicable from a concentration level of xx µg/l or yy mg/kg dry matter”, without any statement how this concentration level was established. When the limit of quantification (LOQ) is evaluated using the procedure of this Technical Specification, there is a possibility that it does not meet the lower limit of the claimed range.

1 Scope

This Technical Specification describes an approach for the validation of physico-chemical analytical methods for environmental matrices.

The guidance in this document addresses two different validation approaches, in increasing order of complexity. These are:

a) method development and validation at the level of single laboratories (intra-laboratory validation);

b) method validation at the level of several laboratories (between-laboratory or inter-laboratory validation), with a focus on methods that are sufficiently mature and robust to be applied not only by a few expert laboratories but by laboratories operating at the routine level.

The concept of these two approaches is strictly hierarchical, i.e. a method shall fulfil all criteria of the first level before it can enter the validation protocol of the second level.

This Technical Specification is applicable to the validation of a broad range of quantitative physico-chemical analytical methods for the analysis of water (including surface water, groundwater, waste water, and sediment). Analytical methods for other environmental matrices, like soil, sludge, waste, and biota can be validated in the same way. It is intended either for analytical methods aiming at substances that have recently become of interest or for test methods applying recently developed technologies.

The minimal requirements that are indispensable for the characterization of the fitness for purpose of an analytical method are: selectivity, precision, bias and measurement uncertainty. The aim of validation is to prove that these requirements are met.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 78-2, *Chemistry — Layouts for standards — Part 2: Methods of chemical analysis*

ISO 5725, *Chemistry — Layouts for standards — Part 2: Methods of chemical analysis*

ISO 11352:2012, *Water quality — Estimation of measurement uncertainty based on validation and quality control data*

ISO 21748:2010, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*

ISO/IEC Guide 99:2007, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC Guide 99:2007 (VIM) and the following apply.

3.1
accepted reference value
value that serves as an agreed-upon reference for comparison, and which is derived as:

a) a theoretical or established value, based on scientific principles;

b) an assigned or certified value, based on experimental work of some national or international organization;

c) a consensus or certified value, based on collaborative experimental work under the auspices of a scientific or engineering group;

d) when a), b) and c) are not available, the expectation of the (measurable) quantity, i.e. the mean of a specified population of measurements

[SOURCE: ISO 3534-2:2006, definition 3.2.7]

3.2
accuracy
closeness of agreement between a test result and the accepted reference value

[SOURCE: ISO 3534-2:2006, definition 3.3.1]

Note 1 to entry: The term accuracy, when applied to a set of test results, involves a combination of random components (usually expressed by a precision measure) and a common systematic error or bias component (usually expressed by a measure for trueness).

Note 2 to entry: The technical term "accuracy" should not be confused with the term ‘trueness’ (see definition of "trueness").

3.3
analyte
substance to be analysed (chemical species or physical parameter)

Note 1 to entry: The quantity of an analyte is the measurand (3.15).

3.4
bias
difference between the expectation of a test result or measurement result and a true value

Note 1 to entry: Bias is the total systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the accepted reference value is reflected by a larger bias value.

Note 2 to entry: The bias of a measuring instrument is normally estimated by averaging the error of indication over an appropriate number of repeated measurements. The error of indication is the: "indication of a measuring instrument minus a true value of the corresponding input quantity".

Note 3 to entry: In practice, accepted reference value is substituted for the true value.

[SOURCE: ISO 3534-2:2006, definition 3.3.2]

3.5
blank
sample or test scheme without the analyte known to produce the measured signal

Note 1 to entry: Use of various types of blanks enable assessment of which proportion of the measured signal is attributable to the measurand and which proportion to other causes. Various types of blank are available (see definition of reagent blank and blank sample).

3.6
calibration
operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication

Note 1 to entry: A calibration may be expressed by a statement, calibration function, calibration diagram, calibration curve, or calibration table. In some cases, it may consist of an additive or multiplicative correction of the indication with associated measurement uncertainty.

Note 2 to entry: Calibration should not be confused with adjustment of a measuring system, often mistakenly called "self-calibration", nor with verification of calibration.

[SOURCE: ISO/IEC Guide 99:2007, definition 2.39]

3.7
certified reference material
CRM
reference material, accompanied by documentation issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceabilities, using valid procedures

[SOURCE: ISO/IEC Guide 99:2007, definition 5.14]

3.8
fitness for purpose
degree to which data produced by a measurement process enables a user to make technically and administratively correct decisions for a stated purpose

3.9
intermediate precision
precision under intermediate precision conditions

[SOURCE: ISO 3534-2:2006, definition 3.3.15]

3.10
intermediate precision conditions
conditions where test results or measurement results are obtained with the same method, on identical test/measurement items in the same test or measurement facility, under some different operating conditions

Note 1 to entry: There are four elements to the operating condition: time, calibration, operator and equipment.

[SOURCE: ISO 3534-2:2006, definition 3.3.16 and ISO 11352:2012, definition 3.10]

3.11
limit of detection
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 β, given a probability α of falsely claiming its presence

Note 1 to entry: IUPAC recommends default values for α and β equal to 0,05.

Note 2 to entry: The abbreviation LOD is sometimes used.

Note 3 to entry: The term “sensitivity” is discouraged for ‘detection limit’.

[SOURCE: ISO/IEC Guide 99:2007, definition 4.18]

Note 4 to entry: The LOD is the lowest concentration of measurand in a sample that can be detected, but not necessarily quantitated under the stated conditions of the test.

3.12
limit of quantitation
lowest concentration of a measurand that can be determined with acceptable precision under the stated conditions of the test

3.13
reporting limit
specific concentration at or above the limit of quantification that is reported to the client with a certain degree of confidence

Note 1 to entry: The reporting limit is often defined on a project-specific basis. If the reporting limit is set below the limit of quantification by the client, method modification is required.

[SOURCE: ISO/TS 13530:2009, 4.4.7]

3.14
linearity
ability of the method to obtain test results proportional to the concentration of measurand

Note 1 to entry: The linear range is by inference the range of measurand concentrations over which the method gives test results proportional to the concentration of the measurand.

[SOURCE: EURACHEM Guide]

3.15
measurand
quantity intended to be measured

Note 1 to entry: The specification of a measurand requires knowledge of the kind of quantity, description of the state of the phenomenon, body, or substance carrying the quantity, including any relevant component, and the chemical entities involved.

Note 2 to entry: In chemistry, “analyte”, or the name of a substance or compound, are terms sometimes used for “measurand”. This usage is erroneous because these terms do not refer to quantities.

[SOURCE: ISO/IEC Guide 99:2007, definition 2.3]

3.16
measurement
process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity

[SOURCE: ISO/IEC Guide 99:2007, definition 2.1]

3.17
measurement uncertainty
non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used

[SOURCE: ISO/IEC Guide 99:2007, definition 2.26].

3.18
outlier
member of a set of values which is inconsistent with the other members of that set

Note 1 to entry: ISO 5725-2 specifies the statistical tests and the significance level to be used to identify outliers in trueness and precision experiments.

[SOURCE: ISO 5725-1:1994, definition 3.21]

3.19
precision
closeness of agreement between independent test results obtained under stipulated conditions

Note 1 to entry: Precision depends only on the distribution of random errors and does not relate to the true value or the specified value.

Note 2 to entry: The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Less precision is reflected by a larger standard deviation.

[SOURCE: ISO 3534-2:2006, definition 3.3.4]

Note 3 to entry: „Independent test results“ means results obtained in a manner not influenced by any previous result on the same or similar test object. Quantitative measures of precision depend critically on the stipulated conditions. Repeatability and reproducibility conditions are particular sets of extreme conditions.

3.20
proficiency testing
evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons

[SOURCE: EN ISO/IEC 17043:2010, definition 3.7]

3.21
quality assurance
part of quality management focused on providing confidence that quality requirements will be fulfilled

[SOURCE: EN ISO 9000:2015, definition 3.3.6]

Note 1 to entry: A major part of quality assurance is quality control.

3.22
quality control
part of quality management focused on fulfilling quality requirements

[SOURCE: EN ISO 9000:2015, definition 3.3.7]

3.23
quantity
property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference

[SOURCE: ISO/IEC Guide 99:2007, definition 1.1]

3.24
working range
interval, being experimentally established and statistically proved by the calibration of the method, between the lowest and highest quantity possibly measured by the method

Note 1 to entry: The lowest possible limit of a working range is the limit of quantification of an analytical method.

3.25
reagent blank
all reagents used during the analytical process (including solvents used for extraction or dissolution) are analysed in isolation in order to check whether they contribute to the measurement signal

Note 1 to entry: The measurement signal arising from the measurand can then be corrected accordingly.

3.26
recovery
extent to which a known, added quantity of determinant in a sample can be measured by an analytical system

Note 1 to entry: It is calculated from the difference between results obtained from spiked and unspiked aliquots of sample, and is usually expressed as a percentage.

[SOURCE: ISO 6107-8:1993/Amd 1:2001-12]

3.27
reference material
RM
material, sufficiently homogeneous and stable with respect to one or more specified properties, which has been established to be fit for its intended use in a measurement process

[SOURCE: ISO Guide 30:2015, definition 2.1.1, modified – Notes to entry have not been included.]

3.28
repeatability
precision under repeatability conditions, i.e. conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time

[SOURCE: ISO 3534-2:2006, definition 3.3.5 and 3.3.6 , modified – These definitions were combined.]

3.29
reproducibility
precision under reproducibility conditions, i.e. conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment

[SOURCE: ISO 3534-2:2006, definition 3.3.10 and 3.3.11, modified – These definitions were combined.]

3.30
residual
difference between the observed response and that predicted by a calibration function

3.31
robustness
measure of capacity of a procedure to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage

3.32
sample
totality of a homogeneous analysis material with an identical composition or quality (similar to term batch)

[SOURCE: ISO/TS 20612:2007, definition 3.3]

3.33
blank sample
matrix with no measurand

Note 1 to entry: They are difficult to obtain but such materials are necessary to give a realistic estimate of interferences that would be encountered in the analysis of test samples.

3.34
selectivity
ability of a method to determine accurately and specifically the measurand of interest in the presence of other components in a sample matrix under the stated conditions of the test

3.35
sensitivity
change in the response of a measurand divided by the corresponding change in the stimulus

3.36
traceability
property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty

[SOURCE: ISO/IEC Guide 99:2007, definition 2.41]

3.37
trueness
closeness of agreement between the average value obtained from a large series of test results and an accepted reference value

Note 1 to entry: In the present document, trueness will be expressed in terms of bias.

Note 2 to entry: Trueness should not be confused with the term ‘accuracy’ (see definition of "accuracy").

3.38
validation
verification, where the specified requirements are adequate for an intended use

[SOURCE: ISO/IEC Guide 99:2007, definition 2.45]

Note 1 to entry: This process is used to asses that a method is fit for its intended purpose. It includes:

— establishing the performance characteristics, advantages and limitations of a method and the identification of the influences which may change these characteristics, and the extent of such changes;

— a comprehensive evaluation of the outcome of this process with respect to the fitness for purpose of the method.

4 Concept

4.1 The concept of two validation levels

The requirements for methods used for monitoring of chemical measurands depend on:

— the extent of the intended or requested monitoring activity; and

— the potential of the available methods that may be used for monitoring a specific substance (group).

In some cases, fully developed methods used by routine laboratories may already exist. More frequently, in the case of new substances or newly developing methods there will be a lack of information on the extent to which the methods have been fully developed and validated. In order to cover most eventualities, two distinct (and hierarchical) levels of method validation are described in this document.

4.2 First level - Validation 1 (V1)

Validation 1 addresses method development (new development or extension of an existing method application to new matrices) and method validation at the intra-laboratory level. The endpoint of Validation 1 is a method with a complete internal validation for the intended purpose at the level of a single laboratory. The endpoint of Validation 1 is identical with the starting point of Validation 2.

4.3 Second level - Validation 2 (V2)

Validation 2 addresses method validation at the inter-laboratory level. The main issue is to demonstrate that the method possesses sufficient inter-laboratory performance and is applicable for use at the level of routine laboratories. This also comprises the development and control of key aspects of method documentation and method usability.

NOTE If a method has successfully passed the Validation 2 procedures, it is usually suitable for standardization at the European level.

4.4 Method validation using a modular approach

4.4.1 Validation modules

The goal of any validation activity is to demonstrate the applicability of an analytical method to the intended purpose. Any analytical method ideally provides reliable and comparable data when carried out on similar samples, independent from external modifications of operating conditions, e.g. different laboratories or operators. This fitness-for-purpose shall be demonstrated through experimental, well-documented evidences. This comprises a set of general principles and criteria that are applicable to physico-chemical test methods. These principles and criteria are organized in four modules (A to D).

This Technical Specification follows a hierarchical approach with respect to the two validation levels. This means that for second level validation (V2) all requirements and criteria of the first level validation (V1) shall be fulfilled.

4.4.2 Module A: Test method definition, documentation and general requirements

The aim of Module A is to check before starting experimental investigation, that the laboratory has extensively taken into account all requirements for performance of the intended method. This step is essentially documental.

It is essential that an unambiguous definition of the method, the intended scope of application, its scientific basis, its purpose and a statement about the need for the test method, and criteria for acceptable method performance be documented. This should also include, if known, a description of the relationship (mechanistic or empirical) between the measurand and the response in the investigated system, e.g. electronic signal. Furthermore, a detailed documented protocol for the method should be available. This should include a description of all materials and reagents, a description of what is to be measured, and how this is to be carried out.

The criteria to be fulfilled in Module A at validation level V1 and at level V2 are different. They are presented in Table 1. A detailed description and structure for Module A is given in Annex A.

4.4.3 Module B: Applicability domain and pre-validation

Experimental investigation starts with Module B.

Test methods may be developed, optimized and validated for measuring a substance or an operationally defined parameter within a specified scope. This scope may include restrictions on the sampled media, the matrix, the concentration range of the measurand, and other limits of the scope of application. The suitability of the method drafted according to Module A is investigated on the whole scope in module B.

It implies that, when an extension of the scope of a validated test method is to be investigated, Module B shall entirely be carried out on the additional parts of the new scope. The criteria to be fulfilled in Module B at validation level V1 and at level V2 are identical. They are presented in Table 1. A detailed description and structure for Module B is given in Annex B.

4.4.4 Module C: Intra-laboratory performance

The basic performance characteristics of the test method within one laboratory should be determined in order to evaluate its reliability, relevance and limitations. These performance characteristics may include (among others) parameters such as precision, bias, selectivity and traceability.

Module C is related to validation level V1. A detailed description and structure for Module C is given in Annex C.

4.4.5 Module D: Inter-laboratory performance

Module D is related to validation level V2. Validation level V1 shall be completed before entering validation level V2 investigations.

In order to be applicable for large-scale European-wide purpose, a test method for pollutants shall show sufficient inter-laboratory performance at the level of routine laboratories. Therefore it is essential that a test method has a high degree of robustness and usability. Furthermore, the completeness and clarity of the method description (protocol) are more important than when a method is to be tested within a single laboratory. The protocol shall be accompanied by detailed control procedures and should not be open to misinterpretation.

The principal tool used to evaluate the inter-laboratory performance of a method is an inter-laboratory comparison involving the analysis of identical test items across all participating laboratories. These inter-laboratory studies are usually designed with the aim of sorting the effects of inter- and intra-laboratory variation on the measures used to characterize and evaluate the performance characteristics of the test method. Appropriated tools and procedures to establish the value of measures for inter-laboratory performance criteria can be found in ISO 5725-2, ISO 5725-5 and ISO 13528. Validation studies carried out as a part of European standardization activities are examples of this level of validation.

A detailed description and structure for Module D is given in 7.4 and Annex D.

4.5 Method classification

In order to select the appropriate validation protocol, potential candidate methods shall be classified according to their level of maturity and validation. This may be considered as an abridged version of a retrospective validation study. Guidance on method classification with respect to the two validation levels is given in this section. However, this should only be used to provide a quick and approximate estimate of the maturity and validation status of the method.

If a classified method has been subject to the appropriate validation protocol, detailed work on either retrospective or prospective validation studies may reveal some inadequacies within the method. In this case it may be necessary to investigate certain modules of the lower validation level. This process may eventually lead to a downgrading of the method by a validation level.

Table 1 provides guidance on how to classify existing methods with respect to the two levels of validation. As a result of the classification, the user should be directed to the appropriate validation protocol. The validation modules shall be used to identify the criteria that shall be fulfilled at the endpoint of each validation protocol. If a method fails to fulfil one or more mandatory criteria assigned to the modules of the respective validation level, the method should be placed in the lower level of validation maturity. In Table 1, a '+' indicates that the respective criterion shall be fulfilled by the candidate method in order to be considered as validated at the respective level, and '(+)' means that the fulfilment of this criterion is not mandatory, but is, at least, highly recommended.

Table 1 — Method classification - requirements for the two validation levels

Criteria for the type of test method to be classified	**Required at validation level**	
	V1	**V2**
Module A – Test method definition and documentation		
Definition of need	+	+
Purpose		
Development of knowledge	+	+
Regulatory purpose		(+)
Scientific basis		
Defined mechanism/effect	+	+
or Scientific proof of relationship between a measured signal effect and the measurand in the investigated system	+	+
Documentation (Protocol)		
with sufficient information for a researcher with special expertise to use the method	+	+
with detailed information sufficient for a trained analyst		+
according to ISO 78-2 (standard-like)		+
with detailed control procedures and performance criteria		+
statistics available (record of survey of performance characteristics)	[a]	+
Dissemination		
Grey literature	+	+
Peer-reviewed publication		(+)

Criteria for the type of test method to be classified	**Required at validation level**	
	V1	**V2**
National, European or International Standard		(+)
Module B – Applicability domain		
Applicability		
to the measurand	+	+
to the matrix of interest	+	+
to the environmental compartment of interest	+	+
Modules C and D – Intra- and Interlaboratory Performance		
Matching the performance characteristics required (e.g. from the regulator or other 'ordering' party)		
shown by one laboratory only	+	+
shown by inter-laboratory study		+
+ indicates that the respective criterion shall be fulfilled by the candidate method in order to be considered as validated at the respective level (+) means that the fulfilment of this criterion is not mandatory, but is, at least, highly recommended		
[a] In the early stages of the implementation of a newly developed method, the laboratory has not enough data to conduct statistical evaluation. Nevertheless, the laboratory should start collecting data to be able to have enough for V2 level validation.		

5 Documentation of the validation process

All validation steps shall be documented. In order to facilitate this process and to ensure a common documentation format, templates for documentation (in the form of tables) are presented in the respective validation protocols (see Annex A to Annex D). A harmonized set of documentation templates ensures that the documentation of the validation process is comprehensible and traceable.

Such templates also enable a quick evaluation of the validation status of the method or the identification of gaps that need to be bridged.

Four different templates shall be used for documentation of the validation process. These templates correspond to the four validation modules A, B, C and D which are defined and described in Annexes A to D. Therefore, the extent of documentation and the number of templates to be completed depends on the level of validation a method has passed (see Table 2).

Table 2 — Method documentation - requirements for each validation level

Documentation required	***Method validated at level***	
	V1	***V2***
Module A (Template Table A.1)	+	+
Module B (Template Table B.1)	+	+
Module C (Template Table C.1)	+	+
Module D (Template Table D.1)		+

Module A and Module B shall contain general information on the method (e.g. its definition, and its applicability domain), whereas Module C and Module D correspond to the specific validation tasks carried out at the level of V1 and V2, respectively. The documentation templates (at least those

corresponding to modules C and D) may therefore also be used as a preview of the validation tasks which shall be carried out at the respective validation level.

The documentation of the method validation process may not be confused with the method description, although some of the information in the two types of documents may be similar or even identical. Information from Module A (Table A.1) and Module B (Table B.1) can be used to compile the information for the method description. At the V1 level, the information given in modules A and B, together with an appropriate reference to the (scientific) literature, may be sufficient as method description, but at the higher level the requirements for the description of the method increase. Therefore, a more comprehensive set of criteria for the method description has been compiled for the V2 level, and should be followed in the preparation of the method description.

If a method enters a higher validation level, the information in modules A and B may need to be updated, because more information has been or needs to be gathered on specific requirements or abilities of the method, or requirements for the method and its performance characteristics may change. Therefore, a method successfully validated up to the V2 level will usually be accompanied by a set of modules recording the history of the validation process of this particular method.

6 Validation 1 (V1): Intra-Laboratory Validation

6.1 General

The Validation V1 protocol covers the scenario where for a given (group of) measurand(s) a method is available but

— is either not applicable to the matrices or compartments of interest (pre-validation), or

— its suitability for the intended purpose with respect to certain performance criteria has not been sufficiently tested and proven.

Protocol V1 describes guidance for the intra-laboratory validation of methods, parameters and criteria that are needed to establish a chemical method at the intra-laboratory level.

The key performance parameters that require attention during the intra-laboratory validation vary according to the measurement requirement and method. Nevertheless, commonly important parameters are listed in 6.4.1. These are based on the earlier described validation Module A (test method definition, documentation and general requirements), Module B (applicability domain and pre-validation), and Module C (intra-laboratory performance).

Module A focuses on the requirements of the method and the information about the method which is needed. These requirements are compared to the application domain of the method, which is described in Module B, and with the intra-laboratory performance characteristics described in Module C.

6.2 Module A: Test method definition, documentation and general requirements

In this module (see Annex A for the structure) general information on the methods shall be provided such as:

— external requirements;

— title of the method;

— beginning and end of validation procedure;

— responsible party;

— scientific basis of the method;

— method definition;

— requirements on devices, reagents, experimental conditions.

Most of the parameters listed are easy to understand and short descriptions of the terms are provided. Some parameters need more attention, and these are discussed in more detail in the following sections.

6.3 Module B: Applicability domain and pre-validation

In this module, information on the applicability of the validated method shall be given. It covers the measurands, matrix and samples, and sampling.

6.4 Module C: Intra-laboratory performance

6.4.1 General

In Module C, the intra-laboratory performance characteristics of the method shall be provided (see Annex C for the structure of the module). For the requirements, see ISO 5725 (all parts) and ISO 11843. Module C shall include detailed information on:

— application range;

— robustness;

— selectivity and interferences;

— accuracy: bias and precision;

— calibration and traceability;

— calibration function (e.g. linearity) and sensitivity;

— limits;

— uncertainty of measurement.

6.4.2 Bias

In order to evaluate a method with respect to its bias, an accepted reference value (often referred to as "conventional true value") is essential. Ideally, the accepted reference value is established for a so-called certified reference material (CRM). For pollutants with recent interest that are to be monitored using methods validated at the V1 level, the availability of CRMs is unlikely. In the absence of a suitable CRM, consensus mean or median values of ring test samples are often used as an estimate of accepted reference values. At the V1-level, neither CRMs nor ring test results may be available.

As an alternative, spiking a sample with a known amount of measurand and analysing the sample before and after spiking offers a means of determining recovery. The recovery shall be calculated on at least 2 levels, as described in ISO 11352:2012, 8.3.2 or 8.3.4.

The spiking level may influence the bias of the method when using this approach. Lower spike concentrations will result in a larger relative bias.

A homogeneous batch of spiked sample material – a so-called internal reference material – may be prepared that can also be used for other steps in the validation process. The sample matrix should be very similar to or mimic the matrix type of that used when the method was developed; the spiking of tap-water to establish the bias of a method used for waste water would not be appropriate. The internal reference material shall be aged before use, in order to allow partition equilibrium to take place. The

laboratory shall document the aging, including the selection of aging time. Aging should not last less than two days.

NOTE Aging is not required for volatile compounds.

Further guidance can be found in the ICH documents [3], [4] and [5].

Another way to estimate bias is to compare the new method with a well-characterized reference method.

The single recovery A_i for the i^{th} fortification level is then calculated as the difference between the measured concentrations in the spiked sample and in the unspiked sample related to the amount added to the sample:

$$A_i = \frac{x_{c+\Delta c,i} - x_{c,i}}{\Delta c} \quad (1)$$

where

$x_{c+\Delta c,i}$ is the measured value of the measurand in the spiked sample,

$x_{c,i}$ is the measured value of the measurand in the sample before spiking,

Δc is the added amount of measurand.

Calculate the recovery A:

$$A = \sum_{i=1}^{n} \frac{A_i}{n} \quad (2)$$

The procedure outlined above is applicable to all chemical methods. For operationally defined methods, the measurement of a reference material shall be preferred, according to ISO 11352:2012, 8.3.2.

6.4.3 Precision

Precision can be divided into repeatability and reproducibility. At the V1 level, only intermediate precision, and if necessary repeatability, is appropriate. The precision should be estimated following repeated analysis of samples, preferably at different concentrations levels. In practice, the spiked sample used for the estimation of the bias (see 6.4.2) should be used for the precision determination, the average value of the outcome being used for the estimation of the bias and the variation for the estimation of the precision. The test set-up described in ISO 11352 shall be used.

NOTE For other matrices the principles described in ISO 11352 can also be applied.

To establish intermediate precision, test results should be independent, i.e. a new calibration solution should be prepared from a different batch of the calibration standard used previously, in order to take into account variations in calibrant purity, weighing and diluting errors, etc.

If repeatability shall be established, carry out the tests with the test set-up mentioned above under repeatability conditions. If repeatability and intermediate precision are needed, perform two test set-ups.

The repeatability, or the intermediate precision, depending of the set-up data used, can be calculated at each investigated level i:

$$C_{V,i} = \frac{s_i}{\bar{x}_i} \quad (3)$$

where

$$s_i = \sqrt{\frac{1}{(p-1)}\sum_{j=1}^{p}\left(x_j - \overline{x}_i\right)^2} \quad (4)$$

$$\overline{x}_i = \sum_{j=1}^{p}\frac{x_j}{p} \quad (5)$$

where

$C_{V,i}$ is the variation coefficient C_V of the x_j in the i^{th} concentration level;

p is the number of repetitions for each concentration level;

x_j is the measured value of the measurand in the j^{th} repetition in the concentration level, in the unit defined for the measurand;

$\overline{x}_i$ is the average of the j repetitions in the i^{th} concentration level, in the unit defined for the measurand.

When several concentration levels show alike repeatability, it may be appropriated to combine all the measurement results of the set up for these results in a global C_V:

$$C_V = \frac{s}{\overline{x}_i} \quad (6)$$

where

$$s = \sqrt{\frac{1}{n(p-1)}\sum_{i=1}^{n}\sum_{j=1}^{p}\left(x_{i,j} - \overline{x}_i\right)^2} \quad (7)$$

$$\overline{x}_i = \sum_{j=1}^{p}\frac{x_{i,j}}{p} \quad (8)$$

where

C_V is the variation coefficient of the measured values $x_{i,j}$ of a set-up;

n is the number of concentrations levels;

p is the number of repetitions for each concentration level;

$x_{i,j}$ is the measured value of the measurand in the j^{th} repetition in the i^{th} concentration level, in the unit defined for the measurand;

$\overline{x}_i$ is the average of the measured values $x_{i,j}$ of a set-up, in the unit defined for the measurand.

6.4.4 Calibration data and function

6.4.4.1 Working range

The linearity or working range should be established, mathematically.

When assessing the linearity of instrumental methods, the parameter related to linearity which is used in subsequent uncertainty calculations is the lack of fit. This is determined from the residuals of the fit of the calibration data to the calibration curve (see ISO 8466-1). Calibration calculations are generally carried out using computer programs that can manage both linear and non-linear functions. The linearity or working range can be established by analysing measurand solutions possessing a wide range of concentration levels.

6.4.4.2 Calibration stability

The susceptibility of a method for instabilities of the calibration should be investigated. The level and frequency of re-calibration should be predefined and meet acceptable criteria. Results of robustness tests (see C.7) should also be considered when setting up the frequency of re-calibration. A lower frequency probably will increase the measurement uncertainty. As long as the method stays fit for the intended use, a less frequent re-calibration may be sufficient.

In many cases, the need for a re-calibration can be identified by periodic measurement of a limited number of standards or CRMs. It should be described how the stability of the calibration has been checked, and which QA/QC measures are necessary to control calibration stability when applying the method.

6.4.5 Limits and application range

Limits should be evaluated in the scope of the fitness for purpose of the method, and in relation with the concentration range. Precision at reporting level should fit the customer's needs for further interpretation.

The limit of a method may be established by different approaches. In practice the terms limit of detection (LOD) and limit of quantification (LOQ) are widely used. ISO/TS 13530 provides options to estimate the LOD of a method, and for its verification. An example is given by Formula (9):

$$x_{\mathrm{LOD}} = 3s_0 + x_{\mathrm{Bl}} \tag{9}$$

where

x_{LOD} is the limit of detection;

s_0 is the standard deviation of the outlier-free results of a blank sample;

x_{Bl} is the mean concentration of the blank sample.

NOTE 1 ISO/TS 13530 provides also a method for the estimation of the LOQ based on multiple of LOD. Examples for the verification of the LOQ are given in ISO/TS 13530:2009, Annex A.

The LOQ represents a concentration of the measurand that can reasonably be determined with an acceptable level of uncertainty. Verification of the LOQ is essential if routine samples frequently show analyte concentrations near the limit of quantification.

In order to verify an estimated LOQ, the level of uncertainty should be defined. Spiked representative blank samples at this concentration level are used for verification. A sufficient amount of the spiked blank matrix sample, or, if no blank sample is available, of a representative synthetic solution, is prepared and divided in $n \geq 5$ test portions. The complete analytical process is conducted on each test portion under conditions of intermediate precision, and the average measured value and the standard deviation are calculated. If the uncertainty of results is smaller than or equal to the pre-set uncertainty the LOQ is verified.

NOTE 2 Routine operation of a mean control chart using a blank sample at LOQ level as control sample is a good way to collect a sufficient number of data, e.g. $n = 20$, for a reliable verification of the LOQ.

If no requirements on the LOQ value and characteristics have been pre-set, then a maximum allowed tolerance of ± 60 % on the defined LOQ is suitable [23], as described by Formula (10) and Formula (11) and shown in Figure 1.

$$X_{\mathrm{LOQ}} - 2 \times s_{\mathrm{LOQ}} > LOQ - 0{,}6 \times LOQ \tag{10}$$

$$X_{\mathrm{LOQ}} + 2 \times s_{\mathrm{LOQ}} < LOQ + 0{,}6 \times LOQ \tag{11}$$

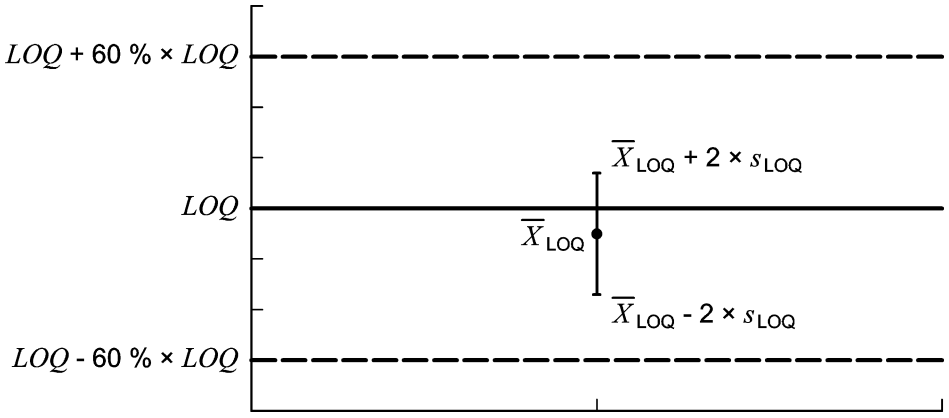

Figure 1 — Graphical representation of a validated LOQ

6.4.6 Selectivity

Selectivity should be taken into consideration at the V1 level, albeit more in qualitative terms of the attention that has been paid to possible interferences than in terms of quantification.

Selectivity is a method performance characteristics that is difficult to quantify. It is necessary to establish that the signal produced by the measurement system, or other measured property, is actually attributed to the measurand, and not produced by accident or coincidence or due to the presence of chemically or physically similar compounds. The selectivity of a method can usually be investigated by studying its ability to measure the substance of interest compared to selected interferences which have been introduced in the sample (i.e. those interferences thought likely to be present in samples or which, from expert knowledge, are known to be likely interferences for the method). Another indirect way is checking the bias of a method (e.g. by analysing a CRM). However, often a CRM is not available for chemicals of recent interest, and a possible interference should also be present in the CRM. If a bias consistent with the fitness for purpose of the method can be demonstrated by analysis of CRMs that also contain representative amounts of interfering compounds, then the method should be selective.

Where it is unclear whether interferences are present, the selectivity of the method can be investigated by studying and comparing this and other different, independent methods/techniques to measure the measurand.

Some chemical methods are more prone to interferences from matrix constituents than other methods. Some interferences will be known to the method developer, and the method should therefore be investigated by adding known amounts of suspected interfering compounds to samples comprising different matrices. Interference effects can lead to an increased response (indicating potential false-positive or increased results), due to signal enhancement, or may lead to a decreased response (indicating potential false-negative or decreased results), due to signal suppression.

Analyses of spiked and non-spiked samples with and without interfering substances allow to distinguish between matrix effects (change of sensitivity) and background interferences.

6.4.7 Robustness

Robustness for routine use can be regarded as multiple:

— robustness toward sample variation in one laboratory (matrix);

— robustness towards small variations in environmental and/or operational conditions;

— robustness toward variability in implementation by several different laboratories.

The major objective of the collaborative trial, which is to investigate the robustness of the method toward variability in implementation by different laboratories, is not required for V1 validation.

Robustness of the method toward sample variation shall be investigated during method development before collaborative trials. This approach helps also in saving costs by avoiding to run costly trials on not sufficiently investigated methods.

The capacity of an analytical method to remain unaffected by small variations in environmental and/or operational conditions provides an indication of its reliability during normal usage. This can be tested using a systematic set of experiments that introduce small but deliberate changes to the experimental conditions of the method, and by observing (either in a qualitative or quantitative way) how these changes affect the final result by determining the relative standard deviations of e.g. the spike sample used in C1.

With regard to the deliberate changes that are introduced, these can be different instrumental settings, reagents, materials, amounts of sample material, exposure times, etc. Eventually, this approach should provide information about the most critical conditions that affect the performance and reliability of the method.

To examine the effect of the variation in the environmental conditions on the results, a “factorial design” approach could be applied as described in Annex F. The advantage of this approach is that information can be provided on which environmental/operational conditions significantly affect the results.

6.4.8 Measurement uncertainty

Measurement uncertainty is an important parameter of a measurement, which is indispensable in a comparable data quality objective. It aids the interpretation of results, allows the users of measurement to render possible comparison of different measurements, potentially obtained using different measurements methods.

If requirements on maximum uncertainty are defined beforehand, the measurement uncertainty checked during validation, covers all aspects of accuracy.

ISO 11352 provides a suitable procedure to establish measurement uncertainty for an analytical method.

7 Validation 2 (V2): Inter-Laboratory Validation

7.1 General

The V2 validation protocol covers the completion of the external validation of a test method. If a method is intended to be used at the level of routine laboratories, the variability aspects of a method across a number of routine laboratories needs to be fully evaluated. This is carried out by means of an inter-laboratory study with the focus on method validation. This inter-laboratory study shall be performed under conditions that are representative for monitoring a measurand by routine laboratories, in order to enable a realistic assessment of the method performance under routine conditions.

The inter-laboratory validation at V2 level also encompasses an evaluation of the usability of the method. This requires a close examination of aspects such as “ease of use”, robustness, completeness of the method description, and sufficient description of the necessary quality control steps and acceptance criteria.

7.2 Method definition and description

The objective of the V2 validation protocol is to facilitate and accelerate the validation and establishment of methods that are not only fit for the desired purpose but also suitable for harmonization or standardization across Europe. An important pre-condition for a method to be applicable to routine laboratories, and to become a European Standard, is a detailed and unambiguous method description. The method description should ensure clarity in procedural detail and minimize or eliminate the risk of misinterpretation.

Detailed requirements on the structure and degree of detail of a method description at V2 level are defined in Annex E. These requirements mainly originate from ISO 78-2, a standard that defines the requirements of international standards describing chemical test methods. The structure given in Annex E is an example and may be changed if necessary .

The method description from the V1 level (together with the information in the documentation templates A to C) should be used as a starting point for the preparation of the method description at the V2 level.

7.3 Module C: Intra-laboratory performance

The validation at the intra-laboratory level has been done at the V1 level. Nevertheless, the participating laboratories should also carry out a basic internal validation of the method in order to participate successfully in the transferability study at the V2 level. This should be performed and documented according to the appropriate parts of the V1 protocol, in Annex C and its sections. As the participating laboratories have to adhere to the method description provided by the organizing laboratory, several sections of Annex C may be skipped by the participating laboratories (e.g. calibration procedures and treatment of raw data will usually be prescribed by the method description).

7.4 Module D: Inter-laboratory performance

7.4.1 General

See Annex D for the structure of this module. In the following the requirements for this module are described.

In order to be applicable for large-scale European monitoring programmes, a test method for pollutants shall show sufficient inter-laboratory performance (preferably across Europe) at the level of routine laboratories.

The principal tool used to evaluate the inter-laboratory performance of a method is an inter-laboratory comparison involving the analysis of identical test items across all participating laboratories. A collaborative study to evaluate inter-laboratory performance at the V2 level requires a sufficient number of participating laboratories (see 7.4.2.1) and geographical coverage. These inter-laboratory studies should be designed with the aim of minimizing the effect of intermediate variation on the measures used to characterize and evaluate the performance characteristics of the test method. Details of the tools and procedures to establish the value of measures for inter-laboratory performance criteria may be different depending on the type of method and the measurand. However, the general approach is similar in most cases, and is outlined in this section and its sections.

The focus of an inter-laboratory study at this level of validation is to validate the method and to assess its applicability by routine laboratories, and not to evaluate the proficiency or capability of the participating laboratories. Nevertheless, the objective of such an inter-laboratory performance study requires the integration of a number of elements from proficiency testing schemes. Therefore, some of

the references that are given in this section deal with the design or evaluation of inter-laboratory trials for proficiency testing.

Annex D provides an overview of the requirements on such an inter-laboratory study, on the information to be compiled and on the tasks to be performed. The structure of Annex D shall also be used as a template for the documentation of the validation process. The sections given in Annex D are discussed in more detail, and guidance is given on minimum requirements for specific aspects of a V2 inter-laboratory study.

7.4.2 General set-up of the inter-laboratory study

7.4.2.1 Participating laboratories

The inter-laboratory study shall only be performed if a sufficient number of laboratories can participate. A high number of participating laboratories increases the reliability of the statistically based conclusions. The minimum number of data sets required for a statistically reliable analysis of the full scope of the inter-laboratory variability of a method is usually recognized as being $l \geq 8$. Therefore the number, l, of remaining laboratories after outliers elimination shall not fall below 8. In order to allow for potential outliers and those laboratories unable, for whatever reason, to submit data, it is recognized that l should be at least in the range of 10 to 12. A minimum number of replicate measurements, $n \geq 3$, should be made.

7.4.2.2 Criteria for participation

The laboratories participating in an inter-laboratory study at the V2 level of validation shall be laboratories using the method at the level of exercise it is intended to: e.g. a significant proportion of routine laboratories participants is required to validate a method intended for routine application. These laboratories should preferably run a quality management system according to the requirements of EN ISO/IEC 17025.

The participating laboratories shall assure that the method description (protocol) is strictly adhered to, and all technical conditions described in the protocol are fulfilled. Tools for statistical evaluation of precision and trueness measures (preferably according to ISO 5725-2) shall be available. Participating laboratories should disclose, at the latest during the training phase (see 7.4.2.3), existing information on the internal validation and use of the method to be validated at the V2 level.

Participating laboratories should represent a cross section of the regional areas in which the use of the method is intended.

7.4.2.3 Training phase

The inclusion of a training phase before the actual inter-laboratory study is highly recommended. This training phase is a pre-condition for the successful performance of the laboratories participating in the study. The objective of the training phase is to enable the participants to gain sufficient experience to enable the adequate and successful performance of the test method. According to the application range of the test method, at least two samples of different concentration (in the lower and upper range of application) should be examined by the participating laboratories in the training phase. The training phase can be accomplished either with spiked sample material, or with standard solutions (provided by the organizer), which are added to a sample matrix by the participant(s). If available, the use of (certified) reference materials shall be preferred. It is not sufficient to perform a training phase with standards only.

In this phase, calibration standards can be provided by the organizer. Together with the exercise samples, information on the concentration of the measurand(s) and on target values for internal performance characteristics (e.g. precision data) shall be provided. At least one sample with a concentration unknown to the participants should be included. In the training phase, the organizing laboratory shall be prepared to provide technical support for the method upon request from the

participant(s). Requests for advice shall immediately initiate a review of the method description with respect to the clarity and coverage of the protocol. If, in the training phase, laboratories do not achieve the required internal performance characteristics of the method, this suggests a problem within the method. In this case, sufficient efforts should be made to rectify the problem prior to the actual inter-laboratory study.

7.4.3 The inter-laboratory study

7.4.3.1 General

Apart from the care taken and the verifications made during the development of a method, its performance can be linked with local characteristics, such as the laboratory apparatus brand used, or mastery of the staff. Therefore, a realistic assessment of performance (accuracy, precision, range of application) cannot be obtained by a single laboratory. It is necessary to conduct an inter-laboratory exercise bringing together representative laboratories of various conditions of implementation of the method. These exercises are expensive, but avoid an erroneous performance estimate of the method. Due to their cost, it is imperative to program these exercises very carefully.

7.4.3.2 Materials: selection, preparation and pre-testing of samples

The number of materials to be investigated shall be selected according to the fitness for purpose of the method subject to validation. A “material” is a compound/concentration level/matrix combination. If blind duplicates are used as replicates (see D.3.2), these shall be considered as one material only (they are not independent). Also blanks or negative controls in a given matrix shall be considered as a “material”.

At the V2 level, the inter-laboratory variability with respect to all potential routine applications of the method shall be evaluated. The study shall therefore encompass all compounds and matrices for which the method is intended to be used. At the V2 level, the study shall be performed with samples that are representative of the actual sample composition under realistic routine conditions, i.e. at this level it is not sufficient to work with simplified matrices.

Furthermore, all of the intended application range (from the lower to the upper concentration limit) shall be covered. If pre-set requirements exist with respect to the lowest concentration that is to be determined (within a certain target uncertainty) at least one material in the inter-laboratory study shall cover this concentration. The same principle shall be applied if there are any requirements or target values for an upper (concentration) limit. If a multi-compound test method is to be validated, the type and number of compounds in the samples shall be representative of actual scenarios in which the method will be applied in an environmental context (e.g. for monitoring).

The samples shall be provided by the organizing party (in case of quantitative chemical methods) or prepared internally by the participating laboratories. If samples are to be provided by the organizer, information on homogeneity and stability of the samples shall also be provided. If necessary, the samples shall be preserved and stabilized in a way that ensures homogeneity and stability up to the agreed period of the study. Homogeneity and stability testing of the samples should be undertaken by the organizer, e.g. according to ISO 13528:2015, Annex B; or [9] (Appendix I and II). The sample quantity or amount should be adjusted to accommodate the required number of analytical replicates).

7.4.3.3 Replicates

The most common practice is to use known replicates, i.e. the participating laboratories should be requested to perform the analysis of the same material in several independent runs. In this case, a minimum of three replicate measurements should be made.

An alternative strategy to obtain precision data is to repeat the measurement of randomly coded blind duplicate (or triplicate) test samples. In this case, only one measurement per test sample is required,

and it is more effective to utilize resources for the analysis of more levels and/or materials rather than for increasing the number of replicates for the individual material.

The minimum numbers of participants and replicate measurements are dependent on the acceptable uncertainty of the estimates for repeatability and reproducibility standard deviations. These estimates can differ considerably from their true values if only a small number of laboratories (e.g. ≈ 8) take part in the collaborative study. An increase in the number of laboratories by 2 or 3 yields only a small reduction in the uncertainties if the number of participants is already larger than 20.

More detailed guidance on this issue is provided in the ISO 5725 series, in particular in ISO 5725-1:1994 and ISO 5725-4:1994, 4.5. These documents provide equations and tables that can be used to adjust the number of replicates and participants to the specific requirements.

7.4.3.4 Performance of the inter-laboratory study

As in the training phase, each participating laboratory should be provided with a supplementary standard to evaluate the calibration carried out within each laboratory. At this validation level, the concentration of the standard shall not be disclosed to the participants. A tolerance level for the correctness of the calibration shall be pre-set in advance. In case of multi-compound methods and standards, a minimum number of compounds to be calibrated, identified and quantified correctly shall also be defined in advance.

Transport and storage of the samples before analysis shall be compliant with the respective recommendations outlined in the method description. A fixed time-scale for carrying out the measurements shall be agreed. Forms for the transmission of results and experimental details shall be provided by the organizing party.

7.4.3.5 Reference data

One of the most essential requirements for a successful validation study is the availability of a reliable assigned value, preferably supported by some quantitative information on the degree of reliability or variability (e.g. in terms of a standard uncertainty).

The assigned value at the V2 level can be determined in one of five ways (see ISO 13528 and EN ISO/IEC 17043:2010, B.2), as described in order of priority:

1) certified reference values;

2) reference values (measured in a real sample matrix, calibrated against a certified reference material);

3) consensus values from expert laboratories (e.g. from transferability study at V2 level);

4) formulation and calculation from the amounts used;

5) consensus values from participants (e.g. mean of the valid results from all participants at V2 level).

Guidance on the estimation of the standard uncertainty of the assigned value for the approaches given above can be found in ISO 13528. If an assigned value from option number 5 is used, then it will not be possible to investigate a systematic bias of the method.

7.4.4 Statistical analysis and calculation of the results

7.4.4.1 General

There are several appropriate ways for a statistical evaluation of the inter-laboratory data. Therefore, the basic principles of the most common approaches are outlined, and some guidance is given on the advantages and disadvantages of the presented approaches.

Calculation of accuracy in terms of precision and bias measures should follow the recommendations of the ISO 5725 series, in particular ISO 5725-2. These "classical" statistical procedures are widely disseminated and accepted, and are applicable to data from a wide range of method types. The ISO 5725 series of standards also provides guidance on statistical outlier tests.

Robust statistical methods, such as the algorithms given in ISO 5725-5, ISO 13528 or the Q-method outlined in ISO/TS 20612, may also be applied to calculate some of the statistical measures given below. These methods have the advantage of being less influenced by single outliers or of a non-normal distribution of results.

7.4.4.2 Results of each laboratory

The following results shall be calculated for each laboratory (and each material):

— number of valid replicate measurements;

— laboratory mean;

— intra-laboratory standard deviation.

7.4.4.3 Results of the inter-laboratory study

The following data shall be calculated and tabulated for each material; detailed guidance for calculation is given in the ISO 5725 series:

— number of valid results;

— number of laboratories after outlier elimination;

— number of eliminated outliers;

— number of eliminated laboratories;

— mean;

— assigned value and its standard uncertainty, if known;

— recovery rate (ratio of mean to assigned value, in %) or bias;

— repeatability standard deviation (s_r);

— coefficient of variation of repeatability ($C_{V,r}$);

— reproducibility standard deviation(s_R);

— coefficient of variation of reproducibility relative standard deviation ($C_{V,R}$).

7.4.5 Evaluation of the fitness for purpose

7.4.5.1 Bias

The trueness of the method shall be evaluated in order to investigate the potential for systematic bias in the method.

The method bias b can be estimated by comparing the mean of analytical results $\overline{x}$ (with uncertainty $u_{\overline{x}}$) with the assigned value x_{av} (uncertainty u_{av})

$$b = \overline{x} - x_{av} \quad (12)$$

NOTE The identification of a bias is not possible when the consensus value from the participants is chosen as the assigned value.

The standard uncertainty of the bias results from the combination of the above mentioned uncertainties:

$$u_b = \sqrt{u_{\overline{x}}^2 + u_{av}^2} \quad (13)$$

With the coverage factor k=2 (for about 95 % confidence) the expanded uncertainty is:

$$U_b = 2 \times u_b \quad (14)$$

As long as the bias b is lower than its expanded uncertainty u_b the bias may be regarded as not significant.

If this test indicates a significant difference between the mean from the validation study and the assigned value, it shall be evaluated whether the fitness-for-purpose of the method is put at risk by this bias. If the bias is within acceptable limits with regard to the intended purpose, this should be clearly documented. Otherwise, the method subject to validation fails to fulfill the requirements at the V2 level, and needs to be improved by modification or optimization of some or all of the procedures.

The evaluation given above shall be performed for each concentration level investigated in the inter-laboratory study. Requirements on a method are often expressed for a specific minimum concentration level, above which the method shall fulfil the respective criteria. The lowest concentration level at which the method fulfils the requirements on bias shall be clearly identified.

7.4.5.2 Precision

The same principle for evaluating the bias (see D.5.1) shall be applied in evaluating the precision data of the method, which has been defined in advance (and documented in the respective template of module A, section A.1). The reproducibility standard deviation, s_R, and where appropriate, also the repeatability standard deviation, s_r, shall be compared with the requirements on precision measures. Usually this can be carried out without applying any statistical test (simply by observing whether the respective standard deviation is larger or smaller than the required precision).

7.4.5.3 Measurement uncertainty

If requirements on the method have been defined in terms of target values for measurement uncertainty (MU), e.g. for a specific concentration level, the measurement uncertainty shall be re-calculated, as described in ISO 21748 to provide the input from the inter-laboratory results.

This approach revisits the uncertainty sources determined in earlier validation stages, replacing these with those that have been addressed by this V2 inter-laboratory study. For example, if the inter-laboratory study can be considered to have covered a suitable range of conditions for a given influence factor (for example an extraction stage in the analysis of a sample) which has been determined individually in a sensitivity study, then this component of the MU may be excluded as it will have been covered within the bias and precision studies of the inter-laboratory studies. The resultant calculation then reduces to the equation given in ISO 21748:2010, 5.3 which in simple terms combines, as a sum of squares, the reproducibility standard deviation calculated from the terms determined in ISO 5725-2 for the collaborative study with any terms addressing uncertainty sources not covered within the scope of the inter-laboratory study.

To compare the MU to pre-set requirements, it is important that the MU is expressed at a stated level of confidence. If the pre-set requirements do not explicitly state this should otherwise be assumed to be with a level of confidence of 95 %, implying a coverage factor, k , of k =2.

7.4.5.4 Application range

Evaluation of other performance criteria (such as limits of detection and quantification, application range etc.) should have been carried out at the lower validation level. Nevertheless, after the completion of the inter-laboratory study at V2 level, it should be checked whether the validation results from V1 are consistent with the results from the V2 study. It may be the case that the required performance may not be achievable by all the routine laboratories. If this is the case, this may indicate a clear limitation on the applicability of the method at the level of routine laboratories. Therefore, discrepancies between the internal validation data of the participants and the data obtained at V1 and V2 levels should be listed in this section, and should be regarded as a clear indication for a limited usability of the method at routine level.

7.4.5.5 Usability of the method at the routine level

If the respective ratio of outlier values (single laboratories) or outlying laboratories is more than 25 % of the total data (or of the number of participants) this may indicate that the method is not yet applicable at the level of routine laboratories. Reports where participating laboratories were unable to submit results or to comply with the requirements of the method shall be included in the calculation of the outlier ratio.

Depending on the specific statistical methods that have been used to calculate the statistical data, outlying values or laboratories with insufficient performance may not have been eliminated, and therefore no number or ratio of eliminated values or laboratories may be available. In this case, a calculation of z-scores (or z_u-scores) should be performed (according to ISO 13528 or ISO/TS 20612).

Target standard uncertainty values are needed to calculate the required laboratory z-scores, and these should be derived from either the pre-set requirements on the method or using uncertainty data that have been generated in V1 and refined in V2 activities. The latter approach is to be preferred, especially in cases where there is significant uncertainty in the reference material used for the study. This can be taken into account by including a term related to the uncertainty of the reference material in the target standard deviation.

If necessary, it may also be possible to calculate target standard uncertainties using an appropriate model, e.g. the Horwitz function for quantitative chemical methods [20]. In general, z-scores outside the range from -2 to 2 should be used to provide an indication of the ratio of eliminated data.

7.4.5.6 Final conclusion

A statement on the fitness for purpose of the method shall be given, summarizing any relevant restraint and limitation of the method. The results of the inter-laboratory study shall be summarized and evaluated with regard to the fitness for purpose of the method (on the basis of the preset requirements).

If only a partial validation of the method has been achieved (e.g. only for a limited application range or only a part of the investigated compounds or matrices), this shall be documented in this section. Any limitations with regard to the desired applicability domain or method performance shall also lead to an update of the respective information in Module A (see Table A.1) and Module B (see Table B.1). If such limitations exist, the internal validation data of the participating laboratories should be checked for discrepancies between V1 and V2 data. Furthermore, it should be checked whether some of the limitations are due to any insufficiencies in the method description, in order to enable a targeted refinement of the method or the method description and eventually a recurrence of the validation activities where appropriate.

7.5 Documentation, publication and standardization

The organizer of the inter-laboratory is responsible for the documentation of the interlaboratory validation study.

Based on the results from the inter-laboratory study and the feedback from the participants, the method description shall be revised if necessary. The results of the validation should be published, preferably in electronic form and should be made accessible to all participating laboratories.

Annex A
(normative)

Module A: Test method definition, documentation and general requirements

Table A.1 — Structure and requirements for test method definition, documentation and general requirements as part of a intra-laboratory validation

Module A		
A.1	**Prerequisite external requirements**	External requirements are used to evaluate if the validation of the method is successful. A clear description of pre-set requirements and specifications of the method are therefore needed. Information on the aim of the method shall be provided, e.g. which compound or end-point is measured at which concentration level, in which matrix. Furthermore, the application range of the methods shall cover the expected range of interest, and any requirements on the measurement uncertainties of data produced by the method shall be documented (if any exist).
A.1.1	Objectives and task	Specify the (pre-set) objectives of the method (measurement application for which the method is being considered)
A.1.2	Requirements and specifications	Documentation of the pre-set requirements e.g. in terms of target values for method performance: — target compound or end-point — application range — matrix — measurement uncertainty
A.2	**Title of the method**	Brief but unambiguous title, e.g. "Determination of volatile aliphatic and aromatic hydrocarbons in the range C_6 – C_{10} in waste water by pentane extraction using GC-FID".
A.3	**Beginning and end of validation procedure**	Start and end date
A.4	**Responsible party**	Institute or person, including full contact data
A.5	**Scientific basis of the method**	Description of the reaction and/or detection principle(s), if necessary supported by reaction formulae and separation principles. A short description of the detection principle, shall be provided. Reference to literature may also be used to provide more detailed information. If another method is used as the starting point for the validation, a reference to the method shall be given. Where the method provides a measurement of a surrogate measurand, the relationship to the required measurand in the environmental context, shall be described.
A.6	**Method definition**	In the method definition a bibliographic reference (if appropriate) or source of the detailed method shall be provided (A 6.1). In the following sections (A 6.2 to A.6.5), a brief description of the experimental setup, the main pre-treatment, sample treatment and analysis steps shall be given (e.g. sieving, extraction, etc.)

Module A		
A.6.1	Method description / SOP	(Bibliographic) reference (if applicable) or source where the detailed method description (with a degree of detail according to the respective validation level) can be obtained.
A.6.2.	Experimental setup	Requires a brief description, with no duplication of the method description or SOP
A.6.3.	Sample preparation and pre-treatment	Indicate whether a specific pre-treatment of the environmental sample is needed (e.g. sieving, centrifugation, filtration)
A.6.4	Sample measurement	Give a brief description of the sample measurement technique
A.7	**Requirements for devices, reagents, experimental conditions**	This section shall provide information on the physical requirements of the method such as instruments, reagents and medium needed for the determination of an pollutant. The quality of reagents used is an important issue in relation to background levels of the target compounds. If certain reagents of a specific source or quality should not be used this shall also be mentioned. Reagents used in the collection of the sample shall be included in this Section.
A.7.1	Instruments/devices	Type of measurement devices; specify any requirements on certain materials (e.g. specific separation phases) and instruments (e.g. resolution, sensitivity)
A.7.2	Environmental conditions	Environmental conditions under which the test shall be conducted (temperature, light, moisture), if this is of relevance to the method.
A.7.3	Reagents	purity of applied reagents have specific requirements for the purity of reagents been identified? are in-house purification procedures for reagents necessary? influence on blank values
A.7.4	Matrix / Medium	The matrix in which the test is conducted (e.g. water, sediment, soil) and whether an artificial medium is needed (e.g. reconstituted water, artificial sediment or soil). In this case, the composition of the medium should be described. In all cases, the physico-chemical characteristics of the medium are to be indicated (e.g. pH, hardness, water retention capacity of soils).
A.8	**Health and Safety**	Information on Health and Safety aspects of the method shall be provided (if required).

Annex B
(normative)

Module B: Applicability domain and pre-validation

Table B.1 — Structure and requirements on the applicability for an intra-laboratory validation

Module B		
B.1	**Measurands**	Detailed information shall be given on the measurands that are covered by the method. For a chemical method the full chemical name (not the brand name) and if possible a CAS number and EC number shall be provided. If information on additional parameters is needed, this information should also be given. Indicate whether all requirements from A.1 are met.
B.2	**Matrix and samples**	
B.2.1	Type of matrix	Indicate the matrices for which the validation has been successfully performed, provide information on specific limitations of the method with respect to the matrix composition (e.g. detailed description of matrix composition, or if a method is only suitable for soils with less than 20 % organic carbon, or the method is only suitable for drinking water). Refer to A.1. This may be updated depending of the results from module C.
B.2.2	Sampling	Refer to specific sampling procedures and precautions which shall be applied to obtain the sample materials used in the validation process, including information on material of containers and sources of error, etc. Describe and give references to specific sampling procedures or precautions that shall be carried out to obtain the sample materials used in the validation process. Information on the material of containers used and sources of contamination, e.g. some target compounds may also be present in the sample containers or sampling equipment. The use of field and laboratory blank samples shall be described if required by the method.
B.2.3.	Sample characteristics	Information on the origin and main composition of the samples used for the validation procedure(s) shall be provided, e.g. amount of organic carbon, suspended particulate matter.
B.2.4	Sample stability and preservation, including transport	Describe all measures taken to stabilize/preserve the samples, if necessary. If available, give advice on suitable/unsuitable techniques, the influence of sample storage, and specific requirements. Describe how samples should be transported. If specific requirements are needed for sample transport, e.g. cooling of samples, these should be defined here. Issues of possible contamination and sample integrity should be documented.
B.3	**Expandability of the method (optional)**	Indicate expected future fields of application (extending the applicability to other matrices or working ranges). Refer to A.1 if appropriate

Annex C
(normative)

Module C: Intra-laboratory performance

Table C.1 — Structure and requirements for intra-laboratory performance validation

Module C		
C.1	**Bias**	Describe the approach used to check bias of the method, and provide the result(s).
C.1.1	Reference materials	State the type of reference material(s) used; in-house or commercially available material (manufacturer); details on spiking solutions and spiked sample matrices, requirements on the uncertainty of the reference materials; characterization of the matrix (e.g.: pH, conductivity, SPM content, TOC, origin).
C.1.3	Recovery	— How was the recovery determined? — For what types of samples/matrices? — What is the relation between concentration range and recovery rate?
C.1.4	Comparability with other methods	Provide results obtained with this method compared to another one, if available, in order to compare the results of the developed/validated method.
C.2	**Precision**	Describe the approach used to check precision; provide results on intermediate precision measures.
C.2.1	Type of samples used for validation	For example real samples, reference materials, "synthetic" samples, reference substance(s).
	Design of experiments to determine precision	According to ISO 11352
C.3	**Calibration**	
C.3.1	Type of calibration	The type of calibration used in the validation process shall be explained and justified (e.g. standard addition, internal/external standards). The scope of the calibration should be described, i.e. which parts of the method (if any) are not covered by the calibration.
C.3.2	Calibration substances	Give details on type, composition, origin and quality of substances used for calibration
C.3.3	Calibration data and function	Description of the handling of the raw data; How have the raw data been treated, e.g. — evaluation of a calibration function according to ISO 8466-1 or ISO 8466-2? — has homogeneity of variances been checked? — what type of calibration function has been used (linear, logarithmic, polynomial)?

Module C		
C.3.4	Calibration stability	How has the stability of the calibration been checked? What are the results? Can recommendations be given on recalibration frequency?
C.4	**Traceability**	Are the calibration (or calibration standards) or spiking solutions traceable to national or international standards? if yes: provide details of the traceability chain. Describe how this relates to the traceability of the method as a whole. If not, describe the source of calibration.
C.5	**Limits and application range**	What are the lower (and probably upper) limits of application? How have they been determined? Where required express lower limits as quantification limits and detection limits.
C.6	**Selectivity, interferences, and discriminative ability**	Check for interfering compounds. Check for discriminative ability (if applicable to the method).
C.7	**Robustness**	Has robustness been checked, e.g. by systematic experiments (deliberate variation of specific parameters)? If yes, what have been the results? Indicate the most sensitive parameters and their impact (can be either qualitative, semi-quantitative of quantitative).
C.8	**Uncertainty of measurement**	How has the uncertainty of measurement been calculated? Which approach has been used? Provide the results.
C.9	**Final evaluation**	Are all requirements defined in A.1 met?

Annex D
(normative)

Module D: Requirements for an inter-laboratory validation study

Table D.1 — Module D - Requirements for an inter-laboratory validation study

Module D		
D.1	**General set-up of the inter-laboratory study**	
D.1.1	Organizing party	Who is responsible for the organization of this inter-laboratory study? Does the organizer meet the requirements for such an inter-laboratory study? EN ISO/IEC 17043 provides a suitable description.
D.1.2	Announcement/ Dissemination	How was the information on this study disseminated? Which measures have been taken to address a representative cross-section of routine-laboratories that are active in the specific field? EN ISO/IEC 17043 provides a suitable description.
D.1.3	Participating laboratories	Number of laboratories involved. Contact data of the participating laboratories
D.1.4	Criteria for participation	How many of the V1 laboratories are involved? How many countries are involved? Does this reflect a cross section of those states where the particular pollutant that is to be measured by the validated method is an issue? Have other criteria / conditions for participation been pre-set, and were they fulfilled?
D.1.5	Time frame	Start and end dates of the inter-laboratory study (including the complete schedule, i.e. dates of the announcements, sample delivery, completion of analysis, submission of results, finalization of the evaluation)
D.1.6	Number of investigated alternatives/options	Were any variants/options of the method investigated? How many (by how many participants)? In this case, a separate documentation and evaluation of each variant is necessary
D.2	**Training phase**	
D.2.1	Participants	With how many participants?
D.2.2	Type of sample material	What type of sample material was used for the training phase; refer to B2 if necessary?
D.2.3	measurands	Does the training phase address the full scope of the method, or only a representative subgroup of compounds?
D.2.4	Examined concentration levels	How many concentration levels were examined? What were the expected or assigned values, what are the actual data from participants?
D.2.5	Standards	Have calibration solutions been provided? How many solutions or samples with concentrations known or unknown to the participants?

Module D		
D.2.6	Evaluation	Criteria for successful participation in the training phase; number of unsuccessful participants
D.3	**Inter-laboratory study**	
D.3.1	Materials	How many materials are included in the study? Are all relevant matrices covered? Are all relevant compounds and the whole application range covered? Are materials provided by the organizer or prepared by the participants? What measures have been taken to ensure sufficient homogeneity and stability of the samples?
D.3.2	Replicates	Number and type of replicates (e.g. known replicates or coded blind replicates) used in the study
D.3.3	Performance of the study	Has a supplementary standard been provided? Has a tolerance level for correctness of the calibration of the participants been pre-set? How has transport & storage of samples been performed? What was the timeframe for carrying out the analysis?
D.3.4	Reference data	How has the assigned value been determined?
D.4	**Statistical analysis and calculation of the results**	
D.4.1	Statistical Analysis	What statistical tools & approaches have been used for the statistical analysis of the data? Have outlying results or laboratories been identified? How have outliers been treated?
D.4.2	Calculation of the results	Calculate and present the final results — of each laboratory — of the whole inter-laboratory study — for each material or concentration level
D.5	**Evaluation of the fitness-for purpose**	
D.5.1	Bias	Are the requirements on the bias met? Does the method have a significant bias? Is the fitness for purpose put at risk by the bias? Can the bias be traced to a particular issue, i.e. can it be shown to be systematic?
D.5.2	Precision	Are the precision measures derived from the inter-laboratory study within an acceptable range (cf. pre-set requirements)?
D.5.3	Measurement Uncertainty	If requirements have been defined in terms of measurement uncertainty, can values obtained by the method fulfil the requirements on measurement uncertainty?
D.5.4	Application range	Are the other requirements on the method (as documented in templates A and B) met by all routine laboratories (if there are any that are not covered by D.5.1 to D.5.3)? Are there stronger limitations to the use of the method that had not been foreseen at the lower validation level (e.g. exclusion of specific matrices, or applicability for a limited number of compounds in case of a multi-compound method)?
D.5.5	Usability	How many outlying results and/or outlying laboratories have been identified? Is the method fully applicable (with results meeting all requirements) by the majority of the participating routine laboratories?
D.5.6	Conclusion	Final conclusion on the fitness for purpose of the method

Annex E
(informative)

Structure and content of the documentation for a validation study (V2)

Table E.1 — Structure and documentation requirements at the V2 validation level

No.	Content and the required degree of detail
E.1	**Title**
	The title shall express clearly and unambiguously (i) the test objects to which the method can be applied, (ii) the substances to be measured, and (iii) the nature or principle of the determination.
E.2	**Introduction**
	Additional information on the technical content of the method description or any other background information on the method (or its "history" with respect to the development of the method) should be included in this section
E.3	**Warnings**
	If any of the reagents or samples used in this method are known to be hazardous either to human health or to the environment, these hazards shall be clearly identified here. Appropriate precautions and safety measures shall also be described.
E.4	**Scope**
	The scope shall state succinctly the chemical method and specifically the test objects to which it applies. If applicable, it shall state the detection limit and/or the limit beyond which the method can no longer be relied upon. The information in this section should enable the user to judge quickly whether the method is applicable to the task or purpose for which it is intended, or whether certain restrictions exist. These restrictions shall take into account the potential presence and extent of other components in the types of samples to be investigated, and of their limiting contents. Relevant information regarding possible interferences shall also be provided. If it is necessary to provide modifications to the basic method e.g. to ensure the elimination of certain interfering factors, these modifications should preferably be treated as special cases. These special cases shall be indicated in the "Scope" clause, and the corresponding modifications shall be described in the "Special Cases" clause (see E.17 in this table).
E.5	**References (optional)**
	This clause shall list those references which are necessary for the proper application of the method. Documents that have served as references in the preparation of the method description should be listed in the bibliography, at the end of the document.
E.6	**Terms and Definitions (optional)**
	This clause shall give any definitions of terms used in the text that facilitate its understanding. At this level of method validation, the terminology shall as far as possible conform to the terminology of European or international standards (CEN, ISO), and reference should be made to existing ISO or CEN definitions.
E.7	**Principle**
	This clause indicates the essential steps in the method used, the basic principles and the properties of which use is made and, if appropriate, the reasons justifying the choice of certain procedures.
E.8	**Reactions (optional)**
	If knowledge about the essential reactions is necessary to understand the method description or for the calculation of the results, these reactions shall be indicated here (supported by reaction formulae, if possible). Reactions can be chemical reactions or physiological effects/mechanisms

No.	Content and the required degree of detail
E.9	**Reagents, materials, media**
	This section shall list (with a sequential reference number) all reagents, materials and media used during the test, together with their essential characteristics (concentration, density, etc.). In addition, this section shall specify, if necessary, their degree of purity (for chemicals) and/or other relevant details. If they exist, Chemical Abstract Service Registry numbers (CAS numbers) and EC numbers of all chemicals shall be given. If necessary, any precautions and conditions to be taken/applied in storing the reagents, and the time period for which they may, or should, be stored/acclimated, shall also be specified. All necessary preliminary test procedures (e.g. to verify the absence of an interfering component in a reagent) shall also be defined and described in this section
E.9.1	Products used in their commercially available form
	In the list of reagents, materials, and media, products used in their commercially available form shall be described unambiguously, giving the particulars necessary for their identification (e.g. the chemical name, the chemical formula, the concentration, the CAS number and EC number).
E.9.2	Products to be prepared by the laboratory
E.9.2.1	Solutions of defined concentration
	The concentration of all solutions which are to be prepared by the laboratory shall be given in an unambiguous form. Solvents for the preparation and/or dilution of solutions shall be clearly defined. Requirements on the quality and/or purity of solvents shall also be defined. If a solution is prepared by dilution of another specified solution, the conventions outlined in ISO 78-2 how to describe the dilution procedure shall be observed.
E.9.3	Reference substances
	Any reference substances or reference materials that are required or recommended should be listed, and appropriate details given (see 9.1).
E.10	**Apparatus**
	This clause shall list (with a sequential reference number) the names and significant characteristics (e.g. material properties) of all the apparatus and equipment (other than standard laboratory apparatus) to be used during the analysis or test. If appropriate, reference shall be made to existing European or international standards e.g. concerning laboratory glassware and related apparatus, or to other relevant international standards or internationally acceptable documents. It is advisable to illustrate, by means of a diagram, special types of apparatus and to indicate the way in which they are assembled. Special requirements on any apparatus that is critical to the method shall be given in this section, especially if they play a significant role in the procedure or if they constitute an important factor in the safety, precision and/or bias of the method. Pre-treatment or cleaning procedures of the apparatus should also be described in this section. Any checking of the functioning of the (assembled) apparatus shall be described in the “Procedure” section, preferably in a sub-clause titled “preliminary test” or “check test”.

No.	Content and the required degree of detail
E.11	**Sampling**
E.11.1	Sampling procedure
	For many occasions, it may be sufficient to refer to the relevant European or international standard dealing specifically with the sampling. If no appropriate standard exists, the sampling clause may include a sampling plan and a sampling procedure, giving guidance on the following issues: - how to obtain a representative sample that can be used for the intended test method - how to avoid or minimize undesirable changes occurring to the sample - required minimum number, mass or volume of sample(s) - sampling equipment - handling of samples - characteristics and material of the containers for sample collection and storage
E.11.2	Preparation of the test sample.
	This clause shall give all relevant information necessary for the preparation of the test sample from which the test portions will be drawn. This test sample is usually prepared from the laboratory sample or field sample as specified in 11.1. For details on the sample terminology (Laboratory sample, test sample, test portion) see ISO 78-2. In each case, all the steps in the preparation shall be stated (e.g. drying, crushing, grinding, sieving etc.) together with appropriate information (e.g. particle size distribution, approximate mass or volume) on the required characteristics of the sample thus prepared. If necessary, details of any containers used for storage, and the storage conditions shall be given.
E.12	**Procedure**
	The "procedure" section may be divided into as many sub-sections or clauses as there are operations or sequences of operations to be carried out. Each operation or sequence of operations shall be described unambiguously and concisely. If the number of steps in the procedure is large, it is recommended to use subdivisions in the sub-clauses (point numbering system), with each element corresponding to a given operation and including all indispensable preliminary operations. If the method or a specific sequence of operations within the method is already given in a European or International Standard, this shall be indicated. In such cases, it may be sufficient to indicate modifications of or deviations from the standard operations. If there are risks or hazards during the procedure for which special precautions are necessary, a statement shall be included at the beginning of the clause. If necessary, more detailed advice on safety procedures and first-aid measures can be given in an annex.
E.12.1	Preparation of the test portion
	Describe how the test portion is prepared from the test sample (or the laboratory sample, if the two are the same). It shall state the method of determining the mass or volume of the test portion (e.g. weighing). It shall state the mass or volume or amount of other discrete units, and the tolerance with which this needs to be measured.
E.12.2	Blank test
	Indicate whether a blank test is necessary or advisable to verify the purity of the reagents or the cleanliness of the laboratory environment or apparatus. If this is the case, this sub-clause shall indicate all the conditions for carrying out this blank test. The blank test should usually be carried out in parallel with and under the same conditions as the actual determination, following the same procedure, using the same quantities of all the reagents and using the same apparatus as in the determination, but without any test portion.
E.12.3	Preliminary test or check test
	If it is necessary to perform any preliminary checks e.g. of the apparatus, all details necessary to carry out these checks should be given in this sub-clause.

No.	Content and the required degree of detail
E.12.4	Determinations, measurement or tests
	Each sequence of operations shall be described adequately and unambiguously. The test shall be set out in an easily readable form in suitable sub-clauses and paragraphs, in order to facilitate the description, the understanding and the application of the procedure. If the product resulting from one of the steps is to be retained and used as a test portion in a later procedure, this shall be clearly stated and identified.
E.12.5	Calibration
	If the method requires any apparatus to be calibrated, this operation shall be the subject of a separate sub-clause located at the most appropriate point in the "procedure" clause. This sub-clause shall describe all necessary operations to be carried out in detail, including requirements on traceable reference materials and calibration artefacts. The frequency of calibration and QA/QC criteria for the calibration (e.g. acceptability criteria or performance criteria) shall also be defined in this sub-clause. If several steps in the calibration procedure are identical to those of the determination procedure, one of the two sub-clauses shall make reference to the other in order to avoid the duplication of redundant information.
E.13	**Calculation**
	This section shall describe all issues of data treatment, including the procedures to calculate the final (reported) result. If comprehensive procedures of data treatment need to be performed prior to the calculation (e.g. selection and/or correction of chromatographic signals or peaks), detailed guidance on these steps shall be given. If the application of complex procedures like sophisticated mathematical or statistical models is required (e.g. fitting a non-linear model by regression analysis), reference can also be made to external sources where these procedures are described in detail (preferably references which are wide-spread and easily accessible (e.g. international standards). In particular, information shall be given on: — the units in which the result is to be expressed — the formula(s) used for the calculation — the meaning of the algebraic symbols used in the formula(s) — the units in which all used quantities are expressed — the number of decimal places or significant figures to which the result is to be given
E.14	**Interpretation of results**
	If the result of the method needs specific interpretative steps guidance on the interpretation should be given here.
E.15	**Performance characteristics**
	This section shall include information on all performance characteristics of the method derived from validation work. For detailed information, reference should be made to the documentation of the validation work, which may be part of an annex (see E.19 in this table).
E.16	**Quality assurance and control**
	This section shall provide a full description of — expected QA/QC procedures — acceptance criteria — control measures and remedial actions to take if these measures indicate that the method is not under control.
E.17	**Special cases (optional)**
	Essential information on special cases that has not been given in the preceding sections may be placed here.
E.18	**Test report**
	This section shall specify in detail the reporting requirements for the method which fully describe the results and supporting QA/QC information enabling an audit trail to be carried out (see the requirements of EN ISO/IEC 17025).

No.	Content and the required degree of detail
E.19	Annexes (optional)
	Annexes may be used to provide supporting information, e.g. the completed forms A to D as a history of the method and its validation maturity.
E.20	**Bibliography**
	References may be given at the point in the text at which they are referred to, or in Clause 5, or in a separate bibliography at the end of the document. Recommendations of ISO 690 should be followed.

Annex F
(informative)

Robustness testing by systematic variation of influencing factors

F.1Design of experiment

Robustness testing by critical-point analysis evaluates how small changes in the method conditions affect the measurement result. The aim is to identify and, if necessary, better control method conditions that might otherwise lead to variation in measurement results, when measurements are carried out at different times or in different laboratories. It can also be used to improve precision and bias. Any stable and homogeneous sample within the scope of the method can be used for robustness testing experiments.

Robustness testing can be carried out by considering each effect separately, by repeating measurements after varying a particular condition by a small amount and controlling the other conditions appropriately.

Since for a well-developed method, most of the effects can be expected to be small, it is possible to vary several parameters at the same time. The design described here allows seven independent factors to be examined in only eight experiments.

Let *A*, *B*, *C*, *D*, *E*, *F* and *G* denote the nominal level and the alternative values are *a*, *b*, *c*, *d*, *e*, *f* and *g*. The chosen levels may be the extreme values of the parameter, e.g. the two extremes of temperature likely to be encountered during use of the method. Table F.1 shows eight combinations of these letters which shows a balance between the capital and lower case letters. The table shows the values, for the seven factors, to be used when running the eight experiments. The results from the experiments are shown as *l*, *m*, *p*, *w*, *v*, *x*, *y*, *z*.

Table F.1 — Example for an experimental design

Experiment number	Method Parameter							Observed result
1	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*l*
2	*A*	*B*	*c*	*D*	*e*	*f*	*g*	*m*
3	*A*	*b*	*C*	*d*	*E*	*f*	*g*	*p*
4	*A*	*b*	*c*	*d*	*e*	*F*	*G*	*w*
5	*a*	*B*	*C*	*d*	*e*	*F*	*g*	*v*
6	*a*	*B*	*c*	*d*	*E*	*f*	*G*	*x*
7	*a*	*b*	*C*	*D*	*e*	*f*	*G*	*y*
8	*a*	*b*	*c*	*D*	*E*	*F*	*g*	*z*

F.2Calculation

For each factor calculate the difference between the average of the results with the factor at its nominal value and the average of the results obtained with the value at the alternative level. To find for example if changing factor '*A*' to 'a' has an effect we calculate Δ_A (Formula (F.1)). Experiments 1 - 4 had the factor at the nominal level, '*A*', and experiments 5 - 8 had the factor at the alternative level, 'a'. Inspection of the table shows that with this combination the effect of the other factors cancels out.

$$\Delta_A = \frac{l+m+p+w}{4} - \frac{v+x+y+z}{4} \quad \text{(F.1)}$$

The seven factors can be dealt with in a similar way by grouping the nominal level for that factor and subtracting the alternative level.

To calculate if any of the differences Δ_A to Δ_G are statistically significant a statistical test (Student-*t* test with e.g. 95 % level of confidence) is applied, comparing the difference with the expected precision of the method standard deviation, *s*.(Formula (F.2)).

If

$$(\Delta)_i > s \cdot t / \sqrt{2} \quad \text{(F.2)}$$

the change from nominal to alternative is significant. The results of the test will be misleading if the factors investigated are not independent.

Bibliography

[1] ISO 690, *Information and documentation — Guidelines for bibliographic references and citations to information resources*

[2] ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

[3] ISO 3534-2:2006, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

[4] ISO 5725-1:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*

[5] ISO 5725-2, *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*

[6] ISO 5725-3, *Accuracy (trueness and precision) of measurement methods and results — Part 3: Intermediate measures of the precision of a standard measurement method*

[7] ISO 5725-4:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 4: Basic methods for the determination of the trueness of a standard measurement method*

[8] ISO 5725-5:1998, *Accuracy (trueness and precision) of measurement methods and results — Part 5: Alternative methods for the determination of the precision of a standard measurement method*

[9] ISO 8466-1:1990, *Water quality — Calibration and evaluation of analytical methods and estimation of performance characteristics — Part 1: Statistical evaluation of the linear calibration function*

[10] ISO 8466-2:2001, *Water quality — Calibration and evaluation of analytical methods and estimation of performance characteristics — Part 2: Calibration strategy for non-linear second-order calibration functions*

[11] ISO 11843, *Capability of detection*

[12] ISO 13528:2015, *Statistical methods for use in proficiency testing by interlaboratory comparisons*

[13] ISO/TS 20612:2007, *Water quality — Interlaboratory comparisons for proficiency testing of analytical chemistry laboratories*

[14] ISO/TS 13530:2009, *Water quality — Guidance on analytical quality control for chemical and physicochemical water analysis*

[15] EN ISO/IEC 17025, *General requirements for the competence of testing and calibration laboratories (ISO/IEC 17025:2005)*

[16] EN ISO/IEC 17043:2010, *Conformity assessment — General requirements for proficiency testing (ISO/IEC 17043:2010)*

[17] EN ISO 9000:2015, *Quality management systems — Fundamentals and vocabulary (ISO 9000:2015)*

[18] ISO Guide 30:2015, *Reference materials — Selected terms and definitions*

[19] ISO/IEC Guide 98-3, *Uncertainty of measurement — Part 3 : Guide to the expression of uncertainty in measurement (GUM 1995)*

[20] NEN 7777, *Environment and food — Performance characteristics of measurement methods*

[21] NF T90-210 (2009-05-01), *Water quality — Protocol for the initial method performance assessment in a laboratory*

[22] IUPAC (1997): *The compendium of analytical nomenclature ("IUPAC orange book")*.

[23] IUPAC. The international harmonized protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report). *Pure Appl. Chem.* 2006, **78** (1) pp. 145–196

[24] Eurachem (2014): *EURACHEM Guide: The fitness for purpose of analytical methods — a laboratory guide to method validation and related topics*.

[25] ICH-Q2A (1995): *Guideline for industry: Text on validation of analytical procedures*. http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html

[26] ICH-Q2B (1996a): *Guidance for industry: Validation of analytical procedures: Methodology*. http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html

[27] ICH-Q6B (1996b): *Harmonised tripartite guideline. Specifications: Test procedures and acceptance criteria for biotechnological/biological products.* http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html

[28] ALBERT R., HORWITZ W. A Heuristic Derivation of the Horwitz Curve. *Anal. Chem.* 1997, **69** pp. 789–790

[29] COFINO W.P., VAN STOKKUM I.H.M., WELLS D.E., ARIESE F., WEGENER J.W.M., PEERBOOM R.A.L. *A new model for the inference of population characteristics from experimental data using uncertainties.* Application to interlaboratory studies. *Chemom. Intell. Lab. Syst.* 2000, **53** pp. 37–55

[30] DE BOER J., COFINO W.P. First world-wide interlaboratory study on polybrominated diphenylethers (PBDEs). *Chemosphere*. 2002, **46** pp. 625–633

[31] ELLISON S.L.R., WILLIANS A., eds. *EURACHEM/CITAC Guide*. Quantifying Uncertainty in Analytical Measurement, Third Edition, 2012., available from http://www.eurachem.org

[32] PRICHARD E. *"Guidelines for Laboratories carrying out measurements where the results will be used to implement the Water Framework Directive (2000/60/EC)"*, SWIFT-WFD Report D12 Part 1 (2007)

[33] TRAVERNIERS I., DE LOOSE M., VAN BOCKSTAELE E. Trends in quality in the analytical laboratory. II. Analytical method validation and quality assurance. *TRAC*. 2004, **23** (8) pp. 535–552

[34] VON HOLST C., MULLER A., BJORKLUND E., ANKLAM E. In-house validation of a simplified method for the determination of PCBs in food and feedingstuffs. *Eur. Food Res. Technol.* 2001, **213** p. 154

This page deliberately left blank

This page deliberately left blank

British Standards Institution (BSI)

BSI is the national body responsible for preparing British Standards and other standards-related publications, information and services.

BSI is incorporated by Royal Charter. British Standards and other standardization products are published by BSI Standards Limited.

About us

We bring together business, industry, government, consumers, innovators and others to shape their combined experience and expertise into standards -based solutions.

The knowledge embodied in our standards has been carefully assembled in a dependable format and refined through our open consultation process. Organizations of all sizes and across all sectors choose standards to help them achieve their goals.

Information on standards

We can provide you with the knowledge that your organization needs to succeed. Find out more about British Standards by visiting our website at bsigroup.com/standards or contacting our Customer Services team or Knowledge Centre.

Buying standards

You can buy and download PDF versions of BSI publications, including British and adopted European and international standards, through our website at bsigroup.com/shop, where hard copies can also be purchased.

If you need international and foreign standards from other Standards Development Organizations, hard copies can be ordered from our Customer Services team.

Subscriptions

Our range of subscription services are designed to make using standards easier for you. For further information on our subscription products go to bsigroup.com/subscriptions.

With **British Standards Online (BSOL)** you'll have instant access to over 55,000 British and adopted European and international standards from your desktop. It's available 24/7 and is refreshed daily so you'll always be up to date.

You can keep in touch with standards developments and receive substantial discounts on the purchase price of standards, both in single copy and subscription format, by becoming a **BSI Subscribing Member**.

PLUS is an updating service exclusive to BSI Subscribing Members. You will automatically receive the latest hard copy of your standards when they're revised or replaced.

To find out more about becoming a BSI Subscribing Member and the benefits of membership, please visit bsigroup.com/shop.

With a **Multi-User Network Licence (MUNL)** you are able to host standards publications on your intranet. Licences can cover as few or as many users as you wish. With updates supplied as soon as they're available, you can be sure your documentation is current. For further information, email bsmusales@bsigroup.com.

BSI Group Headquarters

389 Chiswick High Road London W4 4AL UK

Revisions

Our British Standards and other publications are updated by amendment or revision.

We continually improve the quality of our products and services to benefit your business. If you find an inaccuracy or ambiguity within a British Standard or other BSI publication please inform the Knowledge Centre.

Copyright

Useful Contacts:

Customer Services
Tel: +44 845 086 9001
Email (orders): orders@bsigroup.com
Email (enquiries): cservices@bsigroup.com

Subscriptions
Tel: +44 845 086 9001
Email: subscriptions@bsigroup.com

Knowledge Centre
Tel: +44 20 8996 7004
Email: knowledgecentre@bsigroup.com

Copyright & Licensing
Tel: +44 20 8996 7070
Email: copyright@bsigroup.com

bsi.

...making excellence a habit.™