TECHNICAL SPECIFICATION

ISO/TS 13530

First edition
2009-03-15

Water quality — Guidance on analytical quality control for chemical and physicochemical water analysis

Qualité de l'eau — Lignes directrices pour le contrôle de qualité analytique pour l'analyse chimique et physicochimique de l'eau

Reference number
ISO/TS 13530:2009(E)

© ISO 2009

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

COPYRIGHT PROTECTED DOCUMENT

© ISO 2009

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

 © ISO 2009 – All rights reserved

Contents

Page

© ISO 2009 – All rights reserved

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In other circumstances, particularly when there is an urgent market requirement for such documents, a technical committee may decide to publish other types of document:

— an ISO Publicly Available Specification (ISO/PAS) represents an agreement between technical experts in an ISO working group and is accepted for publication if it is approved by more than 50 % of the members of the parent committee casting a vote;

— an ISO Technical Specification (ISO/TS) represents an agreement between the members of a technical committee and is accepted for publication if it is approved by 2/3 of the members of the committee casting a vote.

An ISO/PAS or ISO/TS is reviewed after three years in order to decide whether it will be confirmed for a further three years, revised to become an International Standard, or withdrawn. If the ISO/PAS or ISO/TS is confirmed, it is reviewed again after a further three years, at which time it must either be transformed into an International Standard or be withdrawn.

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

ISO/TS 13530 was prepared by Technical Committee ISO/TC 147, *Water quality*, Subcommittee SC 2, *Physical, chemical and biochemical methods*.

This first edition of ISO/TS 13530 cancels and replaces ISO/TR 13530:1997, which has been technically revised.

 © ISO 2009 – All rights reserved

Water quality — Guidance on analytical quality control for chemical and physicochemical water analysis

1 Scope

This Technical Specification provides comprehensive guidance on within-laboratory and between-laboratory quality control for ensuring the production of results with a known level of accuracy in the analysis of waters.

This Technical Specification is applicable to the chemical and physicochemical analysis of all types of waters. It is not intended for application to the analysis of sludges and sediments (although many of its general principles are applicable to such analysis) and it does not address the biological or microbiological examination of water. Whilst sampling is an important aspect, this is only briefly considered.

Analytical quality control, as described in this Technical Specification, is intended for application to water analysis carried out within a quality-assurance programme. This Technical Specification does not address the detailed requirements of quality assurance for water analysis, which can be found in the EURACHEM/CITAC Guide (2002) [20].

The recommendations of this Technical Specification are in agreement with the requirements of established quality-assurance documentation (e.g. ISO/IEC 17025).

This Technical Specification is applicable to the use of all analytical methods within its field of application, although its detailed recommendations may require interpretation and adaptation to deal with certain types of determinands (for example, non-specific determinands, such as suspended solids or biochemical oxygen demand, BOD). In the event of any disparity between the recommendations of this Technical Specification and the requirements of a standard method of analysis, the requirements of the method should prevail.

2 Normative references

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

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

ISO 5725 (all parts), *Accuracy (trueness and precision) of measurement methods and results*

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

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

ISO 13528:2005, *Statistical methods for use in proficiency testing by interlaboratory comparisons*

ISO/IEC 17025:2005, *General requirements for the competence of testing and calibration laboratories*

ISO Guide 35, *Reference materials — General and statistical principles for certification*

© ISO 2009 – All rights reserved

ISO/IEC Guide 43-1, *Proficiency testing by interlaboratory comparisons — Part 1: Development and operation of proficiency testing schemes*

ISO/IEC Guide 43-2, *Proficiency testing by interlaboratory comparisons — Part 2: Selection and use of proficiency testing schemes by laboratory accreditation bodies*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1 Terms related to measurement methods

3.1.1
validation
confirmation by examination and the provision of objective evidence that the particular requirements for the specific intended use are fulfilled

[ISO/IEC 17025:2005]

3.1.2
accuracy
closeness of agreement between a test result or measurement result and the true value

NOTE 1 In practice, the accepted reference value (3.2.6) is substituted for the true value.

NOTE 2 The term “accuracy”, when applied to a set of test or measurement results, involves a combination of random components and a common systematic error or bias component.

NOTE 3 Accuracy refers to a combination of trueness and precision.

[ISO 3534-2:2006]

3.1.3
bias
difference between the expectation of a test result or measurement result and a true value

[ISO 3534-2:2006]

3.1.4
trueness
closeness of agreement between the expectation of a test result or a measurement result and a true value

NOTE 1 The measure of trueness is usually expressed in terms of bias.

NOTE 2 Trueness is sometimes referred to as “accuracy of the mean”. This usage is not recommended.

NOTE 3 In practice, the accepted reference value is substituted for the true value.

[ISO 3534-2:2006]

3.1.5
precision
closeness of agreement between independent test/measurement results obtained under stipulated conditions

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

NOTE 2 The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results or measurement results. Less precision is reflected by a larger standard deviation.

© ISO 2009 – All rights reserved

NOTE 3 Quantitative measures of precision depend critically on the stipulated conditions. Repeatability conditions and reproducibility conditions are particular sets of extreme stipulated conditions.

[ISO 3534-2:2006]

3.1.6
limit of detection
output signal or value above which it can be affirmed with a stated level of confidence, for example 95 %, that a sample is different from a blank sample containing no determinand of interest

[ISO 6107-2:2006]

3.1.7
limit of quantification
stated multiple of the limit of detection, for example, two or three times the limit of detection, at a concentration of the determinand that can reasonably be determined with an acceptable level of accuracy and precision

NOTE Limit of quantification can be calculated using an appropriate standard or sample, and may be obtained from the lowest calibration point on the calibration curve (excluding the blank).

[ISO 6107-2:2006]

3.1.8
analytical run
group of measurements or observations carried out together, either simultaneously or sequentially, without interruption, on the same instrument by the same analyst using the same reagents

NOTE 1 An analytical run may consist of more than one batch of analyses. During an analytical run, the accuracy and precision of the measuring system is expected to be stable.

NOTE 2 Definition taken from Reference [33] in the Bibliography.

3.1.9
batch of analyses
group of measurements or observations of standards, samples and/or control solutions which have been performed together in respect of all procedures, either simultaneously or sequentially, by the same analysts using the same reagents, equipment and calibration

3.2 Terms related to measurement results

3.2.1
error of measurement
test result or measurement result minus the true value

NOTE 1 In practice, the accepted reference value is substituted for the true value.

NOTE 2 Error is the sum of random errors and systematic errors.

NOTE 3 Adapted from ISO 3534-2:2006.

3.2.2
systematic error of result
component of the error of result which, in the course of a number of test results or measurement results, for the same characteristic or quantity, remains constant or varies in a predictable manner

NOTE Systematic errors and their causes can be known or unknown.

[ISO 3534-2:2006]

© ISO 2009 – All rights reserved

3.2.3
random error of result
component of the error of result which, in the course of a number of test results or measurement results, for the same characteristic or quantity, varies in an unpredictable manner

NOTE It is not possible to correct for random error.

[ISO 3534-2:2006]

3.2.4
true value
value which characterizes a quantity or quantitative characteristic perfectly defined in the conditions which exist when that quantity or quantitative characteristic is considered

NOTE The true value of a quantity or quantitative characteristic is a theoretical concept and, in general, cannot be known exactly.

[ISO 3534-2:2006]

3.2.5
conventional true value
value of a quantity or quantitative characteristic which, for a given purpose, may be substituted for a true value

NOTE A conventional true value is, in general, regarded as being sufficiently close to the true value for the difference to be insignificant for the given purpose.

[ISO 3534-2:2006]

3.2.6
accepted reference value
value that serves as an agreed-upon reference for comparison

NOTE The accepted reference value 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 technical group;

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

[ISO 3534-2:2006]

3.2.7
certified reference material
reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence

NOTE Definition taken from Reference [36] in the Bibliography.

3.2.8
metrological traceability
property of a measurement result relating the result to a stated metrological reference through an unbroken chain of calibrations of a measuring system or comparisons, each contributing to the stated measurement uncertainty

[ISO/IEC Guide 99:2007]

© ISO 2009 – All rights reserved

3.3 Terms related to uncertainty

3.3.1
uncertainty of measurement
non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used

[ISO/IEC Guide 99:2007]

3.3.2
standard uncertainty
$u(x_i)$
uncertainty of the result x_i of a measurement expressed as a standard deviation

[ISO/IEC Guide 98-3:2008]

3.3.3
combined standard uncertainty
$u_c(y)$
standard measurement uncertainty that is obtained using the individual standard measurement uncertainties associated with the input quantities in a measurement model

[ISO/IEC Guide 99:2007]

3.3.4
expanded uncertainty
U
product of a combined standard measurement uncertainty and a factor larger than the number one

NOTE 1 The factor depends upon the type of probability distribution of the output quantity in a measurement model and on the selected coverage probability.

NOTE 2 The term “factor” in this definition refers to a coverage factor.

[ISO/IEC Guide 99:2007]

3.3.5
coverage factor
k
numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty

NOTE A coverage factor is typically in the range from 2 to 3.

[ISO/IEC Guide 98-3:2008]

4 Performance characteristics of analytical systems

4.1 Introduction

ISO/IEC 17025 requires the validation of methods; the validation process is described in detail in EURACHEM Guide (1998) [18]. Primary validation is part of the development of a new analytical method, and is performed during the standardization of the method. Important points to be considered for the primary validation of an analytical method are the following:

— scope of the method;

— calibration;

© ISO 2009 – All rights reserved

— limit of detection/limit of quantification;

— interferences;

— estimation of accuracy (trueness and precision);

— uncertainty of measurement;

— robustness;

— fitness for purpose.

According to ISO/IEC 17025, the laboratory shall confirm that it can correctly operate standard methods before applying these methods. This procedure is called secondary validation, where emphasis is laid on calibration and interferences, as well as the laboratory limit of quantitation and measurement uncertainty. In 4.2 to 4.8, the topics calibration and quantification, matrix effects and measurement uncertainty are dealt with especially.

4.2 Scope of the method

A clear definition should be given of the forms of the substance that are determined by the procedure and also, when necessary to avoid ambiguity, those forms that are not capable of determination. At this point, it is worth emphasizing that the analyst's selection of an analytical method should meet the user's definition of the determinand. Non-specific determinands need the use of rigorously stipulated analytical methods in order to obtain reliable and comparable results.

Many substances exist in water in a variety of forms or “species”, and many analytical systems provide a differential response to the various forms. For example, when a separation of “dissolved” and “particulate” material is required, special care is necessary to define precisely the nature and pore-size of the filter to be used.

A precise description of the types and natures of samples is important before the analytical system can be chosen. The precautions to be taken when a sample is analysed will depend to a high degree on the sample. The analyst needs information that is as complete as possible on sample types, concentration levels and possible interferences. The scope should contain a clear statement of the types of sample and sample matrices for which the procedure is suitable. If necessary, a statement should also be made of important sample types and matrices for which the procedure is not suitable.

The range of application corresponds to the lowest and highest concentrations for which tests of precision and bias have been carried out using the system without modification. Where an extension can be used to enable the examination of samples containing concentrations greater than the upper limit, such as by analysis after dilution, then it should be regarded as a different procedure but whose performance characteristics may be inferred from the values quoted for the original.

The concentration range of interest can have a marked effect on the choice of analytical technique; of primary concern is the smallest concentration of interest.

4.3 Calibration

4.3.1 Basics of calibration and quantification

A calibration function is determined from information values y_i, obtained by measuring given standard concentrations x_i. The resulting standard deviation of the method s_{x0} or confidence interval is a performance characteristic of this calibration, not a performance characteristic of the quantification. Quantification of the analyte content of samples using a calibration function is based on interpolation. The most important performance characteristic for quantification is the estimation of measurement uncertainty (see 4.6) where the uncertainty of calibration is one of the contributions.

 © ISO 2009 – All rights reserved

Depending on the kind of the functional relationship between the analyte concentration and the measurement response, different mathematical or statistical tools can be used. Mathematical and statistical models are subject to different assumptions; therefore, not only is the performance of the analytical equipment essential but also the kind of mathematical approach. The models should help the analyst to find a clear and reliable functional relationship for calibration. The models should not limit the capabilities of the analytical equipment. Therefore, the selection of the calibration model should be undertaken carefully, taking into consideration the fitness for purpose and the measurement uncertainty required.

— Model of linear regression (ISO 8466-1): the classical model for calibration with the limitations of normal distribution of the responses and a homogeneity of variances over the working range. Normally, this homogeneity of variances is only given in a working range of one or at maximum two decades of concentration. This fact limits the applicability.

— Model of non-linear second-order calibration functions (ISO 8466-2): after discovering a significant non-linearity. This model has the same limitations as linear regression: normal distribution of the responses and homogeneity of variances over the working range.

— Model of weighted regression, weighting the information value with the reciprocal variance, calibration over more than a decade of concentration range is possible.

— Calibration over more than one decade of concentration with at least two concentrations, e.g. for inductively coupled plasma/mass spectrometry (ICP/MS). Linearity must have been checked beforehand with a minimum of five concentrations, e.g. by graphical presentation.

NOTE Detailed instructions on calibration procedures are given in References [26] and [28] in the Bibliography.

All these kinds of calibration should be tested for their contribution to measurement uncertainty. For example, analyse N times, as a minimum in triplicate, an independent standard or a certified reference material within the concentration range of interest, and calculate the results in accordance with the applied calibration method. The standard uncertainty component, s_{p}, representing the deviation of the single results, ρ_{i}, from the reference value, ρ, is calculated as follows:

$$s_{\mathrm{p}} = \sqrt{\frac{\sum_{i=1}^{N}\left(\rho_{\mathrm{i}} - \rho\right)^2}{N-1}} \tag{1}$$

where

s_{p} is the standard uncertainty component from calibration;

N is the number of replicate measurements.

N should be chosen with respect to the confidence required. s_{p} should be compared with the respective target measurement uncertainty to set the tolerable amount.

4.3.2 Calibration strategies

4.3.2.1 Basic or instrument calibration

This type of calibration is carried out without matrix components and without sample preparation; for calibration, pure standard solutions are used. This is inexpensive and directly suited for quantification, if matrix components do not change the slope and the intercept of the calibration function significantly. Normally this calibration is more precise than a calibration including all sample preparation steps.

© ISO 2009 – All rights reserved

4.3.2.2 Calibration with matrix material

This kind of calibration is comparable with basic calibration (4.3.2.1). No sample preparation steps are carried out. Instead of the pure solutions of the standards, solutions of the standard substances in analyte-free matrix or artificial matrix are used.

4.3.2.3 Calibration over the total procedure

This type of calibration includes all sample preparation steps, and the calibration solutions are prepared with a representative matrix material. It is suitable to show whether matrix components or sample preparation steps do or do not change the slope and intercept of the calibration function compared to the basic calibration function. Frequently, it is not easy to find a representative matrix material to use it as a basis of calibration solutions. In water analysis, this means that blank free natural water representing the possible matrices of other natural waters has to be found.

If the matrix components significantly change intercept and slope of the calibration function compared to the basic calibration, it is possible to use this calibration over the total procedure for quantification.

4.3.3 Internal standardization

4.3.3.1 General aspects

The use of internal standardization for the quantification of concentrations minimizes possible errors made both during injection and by sample losses during sample pre-treatment steps, and also differences in the final sample extract volumes and changes in recoveries caused by matrix effects. Substances used as internal standards should have the following properties.

— Chemical-physical properties should be the same concerning the error-prone procedure steps which should be corrected. If the total procedure should be controlled by the internal standardization, isotopic labelled compounds are recommended; if only final volumes or detection should be controlled, other substances, which are representative concerning these steps, can be used.

— There should not be any measurement interferences with the internal standards.

— No occurrence of the internal standard, neither in real samples nor as blanks, which cannot be avoided.

— Concentration of the added internal standard: in the dynamic range of the method, preferably in the same concentration range as the analytes.

— They should have similar intensities of responses as the analytes.

It is recommended to carry out the internal calibration as a basic calibration, because all multiplicative matrix effects are corrected by the internal standard if it has the properties listed above. Otherwise, an internal calibration over the total procedure has the same disadvantages as described in 4.3.2.3, often there is a lack of representative matrix material. As with all the other possibilities for correcting matrix effects, internal standardization cannot overcome additive matrix effects as well.

4.3.3.2 Calibration with internal standards

Calculation of the calibration function is usually available as an option in the quantification programs of most manufacturers' data analysis software.

Adjust the concentrations according to the sensitivity of the equipment used and the range of determinations required. Evaluate the linear range and, subsequently, set up a calibration. Establish the linear function as a basic calibration of the pairs of values $y_i / y_{is,i}$ and $\rho_i / \rho_{is,i}$ of the measured series using Equation (2):

© ISO 2009 – All rights reserved

$$\frac{y_{\mathrm{i}}}{y_{\mathrm{is,i}}} = a_{\mathrm{i}} \frac{\rho_{\mathrm{i}}}{\rho_{\mathrm{is,i}}} + b_{\mathrm{i}} \tag{2}$$

where

y_{i} is the measured response of substance i; the unit depends on the evaluation, e.g. area value;

ρ_{i} is the mass concentration of substance i (external standard) in the working standard solution, e.g. in nanograms per millilitre, ng/ml;

a_{i} is the slope of the calibration function of substance i; the unit depends on the evaluation, e.g. area value millilitres per nanogram, area value·ml/ng.

b_{i} is the ordinate intercept of the calibration curve; the unit depends on the evaluation, e.g. area value;

$y_{\mathrm{is,i}}$ is the measured response of the internal standard for the substance i; the unit depends on the evaluation, e.g. area value;

$\rho_{\mathrm{is,i}}$ is the mass concentration of the internal standard, for the substance i, e.g. in nanograms per millilitre, ng/ml.

Calibration can be done as linear regression or as a two point calibration over more than one decade after the previous linearity check.

4.3.3.3 Quantification with internal standards

Add a known amount of the internal standards to the sample prior to sample preparation. Adjust this amount of the internal standards in such a manner that the mass concentration $\rho_{\mathrm{is,i}}$ in the final volume, e.g. of the extract, is nearly the same in the prepared samples as in the calibration solutions. Use the same solvent composition for the standard solutions and the samples.

Calculate the mass concentration $\rho_{\mathrm{i,sample}}$ of the substance using Equation (3).

$$\rho_{\mathrm{i,sample}} = \frac{\frac{y_{\mathrm{i,sample}}}{y_{\mathrm{is,i,sample}}} - b_{\mathrm{i}}}{a_{\mathrm{i}}} \times \frac{m_{\mathrm{is,i}}}{V_{\mathrm{sample}}} = \frac{\rho_{\mathrm{i,sample\ extract}}}{\rho_{\mathrm{is,i,sample\ extract}}} \times \frac{m_{\mathrm{is,i}}}{V_{\mathrm{sample}}} \tag{3}$$

where

$y_{\mathrm{i,sample}}$ is the measured response, e.g. peak area, of substance i in the sample extract;

$y_{\mathrm{is,i,sample}}$ is the measured response, e.g. peak area, of the internal standard, for substance i, of the sample;

$\rho_{\mathrm{i,sample\ extract}}$ is the mass concentration of substance i in the sample extract, e.g. in nanograms per millilitre, ng/ml; usually calculated by the software;

$\rho_{\mathrm{is,i,sample\ extract}}$ is the mass concentration of the internal standard in the sample extract, for substance i, e.g. in nanograms per millilitre, ng/ml; usually reported by the software;

$\rho_{\mathrm{i,sample}}$ is the mass concentration of substance i in the water sample, e.g. in micrograms per litre, µg/l;

$m_{\mathrm{is,i}}$ is the mass of the added internal standard substance, e.g. in micrograms, µg;

V_{sample} is the sample volume, in litres, l;

© ISO 2009 – All rights reserved

a_i is defined in Equation (2);

b_i is defined in Equation (2).

4.3.3.4 Determination of recoveries of the internal standards

It is necessary to control the recovery of the internal standards for the total procedure for each sample.

This is possible by comparing the mean response from the internal standard in the calibration solutions with the response obtained from the prepared sample. To achieve this, it is essential, that the final volume of the prepared sample is known for calculating the theoretical final concentration of the internal standard in the prepared sample. In routine use, the theoretical concentration of the internal standards in the prepared samples is the same as in the calibration standard solutions, therefore no additional effort results.

Alternatively, a second internal standard, e.g. an injection standard can be used for the calculation of recoveries of internal standards. This is a useful procedure if the final volumes after the sample preparation vary, as frequently happens after enrichment to very small final volumes. This second internal standard is added to the calibration solution and to the final prepared sample prior to the sample measurement. The final concentration shall be the same for the calibration solution and the prepared sample with a theoretical final volume. Now it is possible to obtain recoveries directly by comparing the responses of the internal and the second internal (injection) standards obtained in the calibration with those obtained from prepared samples, e.g. extracts.

$$A_{\text{is,i,sample}} = \frac{y_{\text{is,i,sample}} \times y_{\text{is,inj,calibration}}}{y_{\text{is,i,calibration}} \times y_{\text{is,inj,sample}}} \times 100 \quad (4)$$

where

$A_{\text{is,i,sample}}$ is the recovery of the internal standard, for substance i, in percent, %;

$y_{\text{is,i,sample}}$ is the measured response, e.g. peak area, of the internal standard, for substance i, in the sample;

$y_{\text{is,i,calibration}}$ is the measured response, e.g. peak area, of the internal standard, for substance i, in the calibration solution;

$y_{\text{is,inj,sample}}$ is the measured response, e.g. peak area, of the injection internal standard, in the sample;

$y_{\text{is,inj,calibration}}$ is the measured response, e.g. peak area, of the injection internal standard, in the calibration solution.

A control chart plot of the recoveries of the internal standards or another kind of documentation should indicate that the recovery rates lie within defined control limits.

4.4 Limit of detection, limit of quantification

4.4.1 General aspects

In broad terms, the limit of detection is the smallest amount or concentration of an analyte in the test sample that can be reliably distinguished from zero or blank (Reference [31] in the Bibliography). For some physical parameters, e.g. pH and redox potential, the concept of limit of detection does not apply and no attempt should be made to determine it for these parameters.

There is much diversity in the way in which the limit of detection and limit of quantification of an analytical system is estimated. Most approaches are based on multiplication of the within-batch standard deviation of results of typical matrix blanks or low-level material or the multiplication of the standard deviation of the method, $s_{x\text{O}}$, by a factor. These statistical inferences depend on the assumption of normality, which is at least

 © ISO 2009 – All rights reserved

questionable at low concentrations. Notwithstanding this, in method validation, a simple definition, leading to a quickly implemented estimation of the detection limit, can be applied.

The different ways of estimating the limit of detection and limit of quantification which are described in the following subclauses are optional. With the recommended minimum degrees of freedom, the value of the limit of detection is quite uncertain, and may easily be in error by a factor of 2. Where more accurate estimates are required, more complex calculations should be applied. For special cases, see ISO 11843, Parts 1 to 4.

The essential step after the estimation of the limit of detection and limit of quantification is the verification. The analyst needs to prove that he is able to detect and, respectively, quantify the analyte at the estimated limits in the respective matrix. If the criteria given in 4.4.6 are fulfilled, the estimated limits, e.g. the limit of detection according to 4.4.2, are verified. After verification, the limit of detection and limit of quantification of different laboratories can be compared, e.g. with quality targets.

4.4.2 Limit of detection based on standard deviation of results of blank samples

The limit of detection can be estimated as:

$$x_{\mathrm{LD}} = 3s_0 + x_{\mathrm{Bl}} \tag{5}$$

where

x_{LD} is the limit of detection;

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

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

If the applied method of calculation of results comprises the subtraction of a matrix blank, the second term, x_{Bl}, has to be ignored.

The precision estimate, s_0, should be based on at least 10 independent complete determinations of analyte concentration in a typical matrix blank or low-level material, with no censoring of zero or negative results. For that number of determinations, the factor of 3 corresponds to a significance level of $\alpha = 0{,}01$.

4.4.3 Limit of detection based on the standard deviation of the method

For analytical methods which have a linear calibration function determined alternatively to the way described in 4.4.2, the limit of detection can be estimated as:

$$x_{\mathrm{LD}} = 4s_{x\mathrm{o}} \tag{6}$$

where

x_{LD} is the limit of detection;

$s_{x\mathrm{o}}$ is the standard deviation of the method (from calibration).

4.4.4 Limit of detection based on baseline noise

For methods which show a baseline noise, the limit of detection, x_{LD}, can be estimated as the concentration of the analyte at a signal/noise ratio $S/N = 3$ after blank correction (if possible).

4.4.5 Limit of quantification

The limit of quantification represents a concentration of the determinand that can reasonably be determined with an acceptable level of accuracy. Usually it is arbitrarily taken as a fixed multiple of the detection limit.

© ISO 2009 – All rights reserved

For method validation, the limit of quantification, x_{LQ}, can be estimated as:

$$x_{LQ} = 3\, x_{LD} \tag{7}$$

The factor $k = 3$ corresponds to a relative result uncertainty of approximately 33 %.

4.4.6 Verification of the limit of detection and the limit of quantification in the matrix

Verification of the limit of quantification is vitally important if routine samples frequently show analyte concentrations near the limit of quantification. In that case, investigations for verification of the limit of quantification should be performed regularly in routine analysis.

For verification of the limit of detection and limit of quantification, spiked blank matrix samples at these concentration levels and blank matrix samples shall be analysed in the same manner as real samples, i.e. under within-laboratory reproducibility conditions.

If the mean response of the samples spiked at the limit of detection level is greater than the maximum blank value, the limit of detection is verified.

If the uncertainty of results for the samples spiked at the limit of quantification level is smaller than or equal to the relative precision corresponding to the factor k, the limit of quantification is verified.

NOTE At the limit of quantification, the uncertainty component bias can be neglected, as the blank is included in the estimation of the limit of detection and limit of quantification.

Examples for the verification of the limit of detection and the limit of quantification are given in Annex A.

4.4.7 Reporting limit

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

4.5 Interferences and matrix effects

4.5.1 General considerations on matrix effects

An important source of random and systematic error in results is the presence of constituents of a sample other than the determinand that cause an enhancement or a suppression of the analytical response, so-called matrix effects.

Matrix effects have an essential influence on the measurement uncertainty. It is possible to correct systematic matrix effects, e.g. to correct the multiplicative matrix effects by standard addition, internal standardization (4.3.3) or a calibration with matrix material (4.3.2.2). In the case of additional matrix effects, the source should be determined. The main sources are poor selectivity of the analytical method and blanks. A control of blanks often is no problem, the enhancement of selectivity is mostly connected with higher expenses on the instrumental equipment. After examining the sources of additional matrix effects, sometimes a subtraction is possible (e.g. blanks).

Most analytical techniques produce accurate results with standard solutions at the optimal concentration.

4.5.2 Quantification for samples showing matrix effects when using basic calibration

If the matrix cannot be defined, as often happens in water analysis, it is not possible to test whether the slope and intercept will be changed significantly. In these cases, it is possible to use the basic calibration, but in combination with additional tests on matrix reference material or recovery examinations on real samples.

© ISO 2009 – All rights reserved

Recovery rates shall be considered in the final results. This may increase the measurement uncertainty by the uncertainty component of the recovery rate. These tests are also suited to demonstrate whether recoveries in different sample matrices are constant, e.g. by keeping a recovery control chart (see 6.4.3). After that, it is possible to reduce the expense and to determine the recoveries only once in every batch. Alternatively, internal standardization (see 4.3.3) should be used.

When using basic calibration, blank determinations over the total procedure have to be performed regularly. If there is no possibility to avoid blanks, single measurements, as well as the detection limit and limit of quantification, should be corrected by the mean of several blank measurements. Measurement uncertainty increases in this case because of the addition of the uncertainty of the blank correction.

4.5.3 Recovery tests

Recovery tests are essential and should be carried out regularly, also in routine practice. The object of the test is to identify bias from certain sources occurring in the analysis of real samples. A known quantity of the determinand is added to a real sample, forming the spiked solution, and the two are analysed, the difference in concentrations found being used to calculate the recovery. This is repeated n times and the mean differences are compared statistically with the theoretically expected recovery of 100 %.

Since the spiked solution is made up by adding a fixed quantity of standard solution to a fixed quantity of real sample, the calculation of its recovery can be made as follows:

Standard solution: volume V, concentration c_a (known)

Real sample: volume V_1, concentration c_r (measured)

Spiked sample: volume $(V + V_1)$, concentration c_s (measured)

Recovery, η percentage of the added determinand which is determined:

$$\eta = \frac{c_s(V + V_1) - c_r V_1}{c_a V} \times 100\ \% \tag{8}$$

It should be emphasized that η, calculated from $m \cdot n$ values (i.e. m batches, n replicate analyses in each batch), is only an estimate of the true mean recovery.

The standard deviation, s, is calculated from the m daily mean recoveries (each calculated from, for example, two replicates); s refers therefore to the standard deviation of m daily mean recoveries.

The standard error, s_η, of η is calculated from

$$s_\eta = \frac{s}{\sqrt{m}} \tag{9}$$

where m is the number of values on which s is based.

The standard error s_η is, in fact, the standard deviation of an estimate of the mean (as opposed to the standard deviation of a single observation). The true mean can be expected to lie within $\pm\, t_{0,05}\, s_\eta$ of the estimated recovery η with 95 % confidence.

It should be noted that the spiking recovery test is fairly limited in the information it yields. For example, if bias is found in the standard solution results, it is quite probable that it will also occur in the spiking recovery results and yield no additional information. It only assumes importance when significant bias does not occur elsewhere. In this case, the implication is a cause of bias in the real sample only, and this usually implies interference proportional to the concentration of the determinand. (Clearly an interference effect of absolute magnitude would not affect the difference between spiked and real samples.) In the case of unsatisfactory spiking recovery, it is advisable to check the precision of the real and spiked results, particularly if the spiked

© ISO 2009 – All rights reserved

solution has not been freshly prepared for each analysis. If either of the two solutions shows signs of deterioration, this could easily produce an unsatisfactory spiking recovery.

NOTE As an alternative to the described recovery test, a recovery function, which is explained in Reference [26] in the Bibliography, can be determined.

4.6 Accuracy (trueness and precision) and uncertainty of measurement

The general term accuracy is used to represent the combination of trueness and precision. Accuracy is a measure of the total displacement of a result from the true value ("total error"); trueness is a measure of systematic errors, precision is a measure of random errors (see Annex B). The determination of the precision can be realized on different levels. The bandwidth comprises, on one side, the repeatability which is measured in a single laboratory (one person, same equipment, short time), on the other side, the reproducibility which is calculated from the results of an interlaboratory comparison which is executed within the method validation process and is published in method standards.

Estimates of precision from one batch of analyses ("within-batch precision") may give an overoptimistic idea of the precision of results produced during routine analysis. For this reason, precision should be estimated from analyses taken from separate batches, spread over a suitable period ("reproducibility within-laboratory"). It is desirable to design the test so that a satisfactory estimate of the dominant source of error is obtained. The experimental design recommended for general use is to make $n = 2$ determinations in $m = 8$ to 11 batches of analysis. Such a design provides estimates of within- and between-batch standard deviations with approximately equal numbers of degrees of freedom. Examples of analysis of variances are given in References [26], [27] and [28] in the Bibliography.

Trueness is closely connected with the demonstration of measurement traceability which is a requirement of ISO/IEC 17025. The procedure for demonstrating traceability and for using appropriated reference material is reported in many guides, e.g. References [21] and [36] in the Bibliography. It should be mentioned that the demonstration of traceability in chemical analysis of water is not easy or can only be achieved in part because of the use of "empirical" methods and because of the complexity of the matrices.

Following the International Standard ISO/IEC 17025, accredited laboratories shall have procedures for the estimation of measurement uncertainty available and under certain conditions they shall state the measurement uncertainty. Uncertainties in an analytical process may be estimated using different procedures according to the destination of the result itself. In any case, when associating an uncertainty to a result, the analytical laboratory must indicate the approach which has been chosen.

The formal approach to measurement uncertainty estimation calculates a measurement uncertainty estimate from an equation or mathematical model. The procedures described as method validation are designed to ensure that the equation used to estimate the result is a valid expression embodying all recognized and significant effects upon the result. The basic document in which this approach is described is ISO/IEC Guide 98-3. Because this guide is hard to comply with in practice, some different standards and guidelines (ISO/TS 21748 and References [19], [23] and [32] in the Bibliography) have been published to give support to the implementation of the concept of measurement uncertainty for routine measurements in laboratories. In this standard, preference is given to reduced procedures (so-called "top-down models") to estimate the uncertainty of measurement.

In Annex C, two different top-down models to estimate the uncertainty of measurement are described. Generally random and systematic errors have to be considered. Therefore the determination of uncertainty is based on validation data which represent the within-laboratory reproducibility and the method and laboratory bias. Data from internal quality control should be preferably used to calculate measurement uncertainty if there is a choice between these data and results of collaborative studies.

4.7 Robustness

A measurement programme, for example, a river survey, may often include a high number of very different types of samples, which may be analysed in several laboratories. For this reason, routine analytical laboratories often prefer robust, multi-purpose analytical techniques applicable to a broad range of samples.

© ISO 2009 – All rights reserved

A “robust” or “rugged” analytical procedure means a procedure so designed that the accuracy of analytical results is not appreciably affected by small deviations from the experimental design prescribed by the analytical method. The use of robust procedures is of great help in achieving reliable results in routine laboratories. Robust methods will produce similar results for the same sample when used in independent laboratories. The most robust procedure is the preferable choice, if the procedure meets the user's requirements. There is no simple numerical value indicating the robustness, but results from interlaboratory trials should be used to illustrate the robustness of a procedure. Special responsibility falls on experts improving or standardizing methods to produce robust techniques.

The need for complete and clear specification of analytical procedures should be stressed. The method should specify all details regarding analysis, equipment, calibration, calculation of results, etc. and also include any details on sampling, sample handling and preservation, any digestion step or other specific pretreatment of samples. Any optional operations should be specifically noted.

Details on the design of ruggedness tests can be found in Reference [18] in the Bibliography.

4.8 Fitness for purpose

Fitness for purpose is the extent to which the performance of a method matches the criteria, agreed between the analyst and the end-user of the data, describing the end-user’s needs. For instance, the errors in data should not be of a magnitude that would give rise to incorrect decisions more often than a defined small probability, but they should not be so small that the analyst is involved in unnecessary expenditure. Analytical fitness-for-purpose criteria may be expressed either in terms of acceptable combined measurement uncertainty or acceptable individual performance characteristics. Sometimes, non-analytical requirements will be important factors in deciding whether a method is fit for purpose, for example, whether results will be available in time for critical actions to be taken.

5 Choosing analytical systems

5.1 General considerations

Generally, standardized methods should be applied in water analysis, for legal reasons. To choose the most suitable method for the analytical task (analyte in question, matrix, and concentration range), the scope of the standard should be considered, as well as possible interfering substances in the samples.

Sampling of waters and effluents is carried out in order to provide information on their qualities. This information may be used for different reasons, for example:

— legal control of discharges;

— environmental monitoring;

— process-control of treatment plant performance;

— evaluation of taxes and charges based on actual emissions.

The user's needs are of primary importance. It is the responsibility of the user to define precisely the objectives of the measurement programme and to help to choose the measurement techniques to be used. The following topics should be defined in the measuring programme:

a) definition of the quantitative information required;

b) definition of the determinands;

c) location, time and frequency of sampling;

d) requirements for analytical results;

© ISO 2009 – All rights reserved

e) use of data and data handling routines, including statistical calculations;

f) introduction of a quality-assurance programme.

All analytical work should be based on a sound and precisely defined measurement programme, providing the analyst with representative and stable samples. The inclusion of a quality-assurance system implies the production of data of stated quality. This is partially attained by analytical quality-control activities which establish whether random and systematic errors are within prescribed limits.

5.2 Practical considerations

When discussing the requirements with the user and selecting suitable analytical systems to fit the measuring programme, the following practical points should be considered:

— the frequency of sampling and the total number of samples on each occasion;

— the maximum period between sampling and analysis, in relation to sample stability;

— the maximum period between sampling and the user's need for the results;

— the volume of sample available;

— automatic or manual techniques;

— equivalent analytical methods;

— robustness and description of the proposed method;

— applicability of the proposed method in the laboratory concerned with respect to cost, speed, etc.

Regarding these practical considerations, factors such as convenience, speed and cost may have a great influence on the final selection of analytical systems. When analysis is required infrequently, it may be necessary to adopt a different approach from that used for regular, frequent determinations. It is still essential that the most appropriate action is taken to ensure control of the measurement process and to provide an estimate of analytical accuracy (see Clause 9).

6 Intralaboratory quality control

6.1 General

Clause 4 deals with the evaluation of the capabilities of an analytical technique in order to judge its likely suitability for a particular application. This clause describes the procedure to be adopted when the system is put into routine use, sometimes called “tertiary validation” or “internal quality control” (IQC) or “routine analytical quality control” (routine AQC).

Having chosen an analytical system capable of being used to produce results of adequate accuracy, the next stage is to establish control over the system and to monitor routine performance. The aim is to achieve a continuing check on the variations observed in routine analysis and to provide a demonstration of satisfactory performance of the method.

The principal tool in routine quality control is the control chart.

© ISO 2009 – All rights reserved

6.2 Terms relating to within-laboratory quality control

Control sample: Sample material whose analytical results are used to construct control charts, for example, standard solutions, real samples, blank samples.

Control value: Value entered on the control chart.

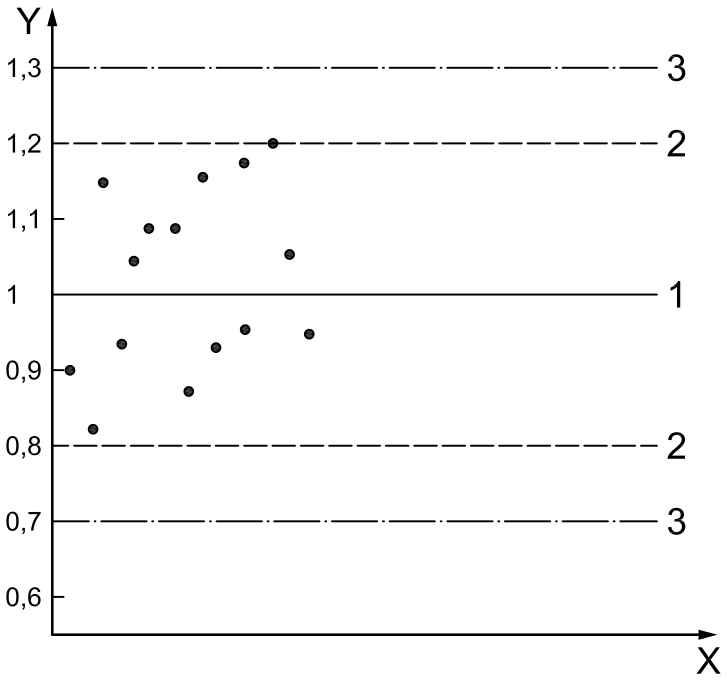
Response value: Value obtained by application of the measurement process.

Analytical result: Value reported as defined in the method. It is derived from the response by application of the calibration.

6.3 Control of accuracy

The use of control charts facilitates the control of accuracy and whether it can be maintained. The simplest form of control chart is one in which the results of the individual measurements made on a control sample are plotted against a time series.

This type of chart (for example, see Figure 1) provides a check on random and systematic error (from the spread of results and their displacement). It is an easy procedure to be used by the analyst because it is simple to plot and no data processing is needed. It is useful when the size of analytical batches is variable or when batches consist of a small number of determinations. Individual result charts are used widely and often form the mainstay of a laboratory's approach to control charting.



Key

X data intervals
Y concentration found for standard solution (mg/l)
1 expected concentration
2 warning limit
3 action limit

Figure 1 — Example of Shewhart control chart for a standard solution

© ISO 2009 – All rights reserved

However, this type of chart may produce false out-of-control values if random error does not follow the normal distribution. For these reasons, a range of more specialized types of charts has been devised. Table 1 gives an overview of the different types of control charts which are described below.

Table 1 — Types of control charts and their application

Type of control chart	Control of trueness	Control of precision
Mean control chart with certified matrix samples	yes	yes
Mean control chart with synthetic samples (standard solutions)	limited because matrix effects are not detected	limited, should be combined with range control chart
Blank control chart	limited	no
Recovery control chart with real samples	yes	yes
Range control chart with real samples	no	yes

Normally, control charts are constructed using statistically calculated warning and action limits. It is also possible to fix limits based on customer or legislative targets or on target measurement uncertainties (see 6.8).

6.4 Control of trueness

6.4.1 General

One way of assessing systematic error is to participate regularly in interlaboratory trials, but these are too infrequent and the results take too long to process for routine day to day control.

As a routine procedure for controlling systematic error, the use of Shewhart control charts (ISO 8258) based on single results or on the mean, spiking recovery and analysis of blanks is recommended.

NOTE A range control chart is most commonly used in conjunction with a mean chart constructed from the same data. The combined use of mean and range charts gives greater control over both systematic and random errors than the use of a single-result control chart.

6.4.2 Mean control chart

For trueness control, standard solutions, synthetic samples or certified real samples may be analysed using a Shewhart chart of individual or mean values.

The analysis of standard solutions serves only as a check on calibration. If, however, solutions with a synthetic or real matrix are used as control samples, the specificity of the analytical system under examination can be checked, provided an independent estimation of the true value for the determinand is available. A useful alternative is to use a typical sample matrix containing none of the determinand and to spike it with a known amount of determinand.

The respective control sample should be analysed a fixed number of times ($\geqslant 1$) in each batch of samples and the mean result entered in the mean control chart.

Certified reference samples (if suitable ones are available) should be analysed with routine samples as a check on trueness. A restricted check on systematic error by means of recovery control charts is often made instead (6.4.3).

NOTE A single-result chart is different from a mean chart for individual results. It does not require the same number of replicates per batch of analyses. For single-result charts, all individual results for variable numbers of control determinations in a batch are plotted. This is a big disadvantage of the single-result chart, as it is difficult to differentiate between within- and between-batch precision.

© ISO 2009 – All rights reserved

6.4.3 Recovery control chart

The recovery control chart is used as a check on systematic errors arising from matrix interferences. A separate control chart for each type of matrix is required in water analysis, because samples of strongly varying matrix composition, such as surface water and municipal and industrial waste water, may be subject to errors of differing sizes and natures.

The recovery control chart, however, provides only a limited check on trueness because the recovery tests will identify only systematic errors which are proportional to determinand concentration; bias of constant size may go undetected.

6.4.4 Blank control chart

The blank control chart represents a special application of the mean control chart.

The blank control chart may help to identify the following sources of error:

— contamination of reagents;

— contamination of reaction vessels and of the measurement system;

— instrumental faults (for example baseline drift).

It is appropriate, therefore, to analyse a blank solution at the beginning and at the end of each batch of samples. The blank values thus obtained are then entered on the blank control chart.

6.5 Control of precision

6.5.1 General

There are five ways of estimating the precision of analytical results in routine analysis:

— use of a range control chart (within-batch errors only) (6.5.2);

— replicate determinations (6.5.3);

— standard addition (6.5.4);

— use of a difference control chart (6.5.5);

— use of a mean control chart (between-batch errors only) (6.4.2).

6.5.2 Range control chart

A range control chart is used to control the within-batch precision of an analytical method. In addition, it allows some assessment of errors caused by calibration drift. The standard deviation for a certain analytical result can be estimated from an existing range control chart, provided the matrix of the sample under examination is similar to that of control samples chosen for the range control chart. The range of the sample in question may also be determined and entered on the control chart as well, in order to prove that an out-of-control situation does not exist.

Estimation of the standard deviation, s, from range control charts:

$$s = \frac{\bar{R}}{d_2} \tag{10}$$

© ISO 2009 – All rights reserved

where

d_2 is a factor (see Table 2);

$\overline{R}$ is the mean range.

Table 2 — d_2 values

n	d_2
2	1,128
3	1,693
4	2,059
5	2,326
6	2,534
7	2,704
8	2,847
9	2,970
10	3,078
n is the number of replicate analyses in each batch (see ISO 8258 and Reference [13] in the Bibliography).	

6.5.3 Estimation of precision with replicate analysis

The analyst should perform replicate analyses of the sample in question to obtain higher reliability of the final result, especially in those cases where the contravention of a threshold value is to be proved. From the data obtained, the standard deviation valid for the matrix in question can be estimated.

Additionally, the performance of replicate determinations offers two further advantages: firstly, coarse errors (outliers) can be detected and, secondly, the analytical error can be reduced.

If a high precision of results is required, the number of replicate analyses should be $n > 6$.

The standard deviation for replicate measurements is calculated as:

$$s = \sqrt{\sum_{i=1}^{i=n} \frac{(x_i - \overline{x})^2}{(n-1)}} \tag{11}$$

where

x_i is the single result of measurement;

$\overline{x}$ is the mean of results;

n is the number of single results.

6.5.4 Standard addition

The estimation of the standard deviation from the range control chart, or with replicate analysis, can help to identify a matrix-dependent imprecision.

© ISO 2009 – All rights reserved

The application of the method of standard addition, whilst helping to control trueness, can tend to degrade precision compared with direct determination. This is the price paid for control over systematic error. The method of standard addition should be applied with caution. It is essential that the linear range of the method be established.

6.5.5 Difference control chart

The difference control chart is a chart of the difference, D ($R_1 - R_2$), in the results of analysis of two portions of the same sample. R_1 and R_2 are the results for the first and second portions respectively. It is essential always to subtract the second result from the first and to plot the difference including sign. The expected value for the chart is zero. In all other ways the chart is constructed as for a mean control chart. This type of chart is useful when a control solution of known or reproducible value is not available. It is also useful when sample homogeneity is a major source of error. The main disadvantage of this type of chart is the dependence of the standard deviation, and therefore the control limits, on the concentration. This problem may be overcome by plotting the percentage difference instead of the absolute difference.

6.6 Principles of applying control charts

6.6.1 Choice of control samples

The choice of control samples depends on the matrix, the analytical method and the accuracy required. Advantages and disadvantages of the several types of control samples are described in References [26] and [27] in the Bibliography.

In general, the following control samples are used.

For mean control charts:	Solutions of the determinand in water, preferably real samples, stable for at least one control period
For blank control charts:	Water samples with a sufficiently small concentration of the determinand or purified water
For recovery control charts:	Real samples with and without addition of the determinand
For range control charts:	Real samples; in special cases, solutions of the determinand in water
For difference control charts	Real samples

6.6.2 Construction of control charts

6.6.2.1 Construction of mean, blank and difference control charts

6.6.2.1.1 Preliminary estimation of warning and action limits

At least 20 control values x_i are required for a trial period to estimate the following tentative control parameters. They are obtained by analysing the control sample on at least 10 working days in duplicate.

From the control values, x ($n \geqslant 20$), estimate the statistical parameters:

— control value x_i (i.e. depending on the type of control chart: single result, mean of the replicate analyses, single blank or single difference of the ith batch);

— mean ($\bar{x}$):

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{i=n} x_i \qquad (12)$$

© ISO 2009 – All rights reserved

— standard deviation (s), see Equation (11) in 6.5.3;

— upper warning limit and lower warning limit:

$$(\mathrm{UW}, \mathrm{LW}) = (\bar{x}) \pm 2s \tag{13}$$

where

UW is the upper warning limit;

LW is the lower warning limit;

— upper action limit and lower action limit:

$$(\mathrm{UA}, \mathrm{LA}) = (\bar{x}) \pm 3s \tag{14}$$

where

UA is the upper action limit;

LA is the lower action limit.

The control chart is constructed in a coordinate system with the ordinate “concentration” and the abscissa “time of analysis” and/or “batch number”. The numerical values for mean, warning limit and action limit are plotted on the ordinate and drawn as lines parallel to the abscissa in the control chart.

6.6.2.1.2 Routine operation

The control value should be obtained at least once per batch of analyses. The frequency with which control values are obtained within a batch lies in the responsibility of the laboratory and should be related to the risks of important errors and the seriousness of their likely consequences. At regular intervals, the control chart should be examined for changes in mean and standard deviation.

The Harmonised Guidelines [30] give recommendations for frequencies of analysis of control samples.

In the long-term operation of a control chart, the question arises as to whether or not to update the estimate of mean and standard deviation used to generate the action and warning limits and, if so, how this might best be done. The guiding principle should be that the chart is intended to detect (with known risks of making the wrong decision) departures from the existing state of statistical control. Including the latest data in the overall estimates of mean and standard deviation may not be sufficient to allow this aim to be fulfilled.

The following approach to monitoring for changes is applied.

It is assumed that the last 60 data points are a homogeneous set and that the issue is whether or not these points are of the same precision as that implied by the initial choice of control limits.

It is also assumed that the normal practice is to base the action and warning limits on a mean and standard deviation derived from all available data points (including the latest). Data points corresponding to “out-of-control” situations for which a definite cause has been identified should not, of course, be included in the calculations.

Review the last 60 data points on the chart. If there are between 1 and 6 (inclusive) cases where the $2s$ warning limits have been exceeded, there is no clear evidence that the precision of analysis has changed. No revision of the chart is required except, as usual, the incorporation of new data points into the estimates of s and $\bar{x}$.

© ISO 2009 – All rights reserved

If there are either no cases where the warning limits have been exceeded or more than 6 cases, it may be concluded with approximately 90 % confidence that the precision has changed (improved or degraded, respectively) and that a revision of the action and warning limits is needed.

In this case, recalculate the control limits on the basis of the mean and standard deviation of the last 60 points and proceed as usual.

Whenever new control limits are calculated as a result of a change in precision, review the new standard deviation (and where appropriate the bias implied by the new mean) against the accuracy targets which apply to the analyses in question. Take corrective action if necessary.

The above procedure need not be carried out each time a new data point is generated. This check on the validity of the current control limits might be worthwhile after, for example, 20 successive points have been plotted, although any obvious changes in the operation of the chart would warrant immediate concern.

6.6.2.2 Construction of recovery control charts

The design and the criteria of decision of the recovery control chart are similar to those of the mean control chart.

6.6.2.2.1 Preliminary tests

For the construction of a recovery control chart, it is recommended to run a trial period of tests.

The control variable to be plotted is η_i:

$$\eta_i = \left(x_a - x_o\right) \times 100/c_a \tag{15}$$

where

x_a is the analytical result (for example concentration) of the determinand in the spiked sample;

x_o is the analytical result (for example concentration) of the determinand in the original sample;

c_a is the concentration or mass respectively of the spiked determinand. This assumes negligible dilution of the sample by the spiked addition.

After completion of the trial period, the following statistical characteristics are derived from the recoveries, η_i ($n \geqslant 20$):

— mean recovery ($\bar{\eta}$);

— standard deviation of the mean recovery (s_η);

— upper warning limit and lower warning limit (UW, LW);

— upper action limit and lower action limit (UA, LA).

Calculation:

$$\bar{\eta} = \frac{1}{n}\sum_{i=1}^{i=n} \eta_i \ (\%) \tag{16}$$

$$s_\eta = \sqrt{\sum_{i=1}^{i=n} \frac{(\eta_i - \bar{\eta})^2}{(n-1)}} \ (\%) \tag{17}$$

© ISO 2009 – All rights reserved

$$\mathrm{UW} = \bar{\eta} + 2s_{\eta} \tag{18}$$

$$\mathrm{LW} = \bar{\eta} - 2s_{\eta} \tag{19}$$

$$\mathrm{UA} = \bar{\eta} + 3s_{\eta} \tag{20}$$

$$\mathrm{LA} = \bar{\eta} - 3s_{\eta} \tag{21}$$

6.6.2.2.2 Further processing

The recovery chart is constructed and maintained in the same way as that described in 6.6.2.1. For the calculation of the statistical parameters, x and s should be replaced with η and s_{η}, respectively.

6.6.2.3 Construction of range control charts

At least 20 control values ($n \geqslant 20$) are required for the pre-period. The control value is the relative range, $R\mathrm{rel}_j$:

$$R\mathrm{rel}_j = \frac{x_{i\max} - x_{i\min}}{\bar{x}_i} \cdot 100\ \% \tag{22}$$

where j is the jth batch of i replicates.

with:

$$\bar{x}_i = \frac{1}{n}\sum_{i=1}^{i=n} x_i \tag{23}$$

where

x_i is the individual analytical result of the respective control sample;

n is the number of replicate determinations of the respective control sample.

After the completion of a preliminary test period, the relative range values, $R\mathrm{rel}_i$ ($n \geqslant 20$), are used to calculate the following statistical parameters:

$$\bar{R}\mathrm{rel} = \frac{1}{n}\sum_{j=1}^{j=n} R\mathrm{rel}_j\ (\%) \tag{24}$$

$$\mathrm{UA} = \bar{R}\mathrm{rel} \cdot D_{\mathrm{UA}}\ (\%) \tag{25}$$

$$\mathrm{LA} = \bar{R}\mathrm{rel} \cdot D_{\mathrm{LA}}\ (\%) \tag{26}$$

Several calculation models may be used to estimate the action limits for this type of control chart. For application in routine work, it is recommended that only the upper action limits (UA) be calculated, warning limits can be calculated additionally (see Reference [26] in the Bibliography). When performing replicate determinations (duplicate to six-fold), the lower action limit (LA) is identical to the abscissa (zero-line).

© ISO 2009 – All rights reserved

The numerical values for the factors D_{UA} and D_{LA} for $P = 99{,}7$ % are:

	Duplicate determination	Three-fold determination	Four-fold determination	Five-fold determination
D_{UA} ($P = 99{,}7$ %)	3,267	2,575	2,282	2,115
D_{LA} ($P = 99{,}7$ %)	0	0	0	0

For further numerical values for the factors D_{UA} and D_{LA}, refer to ISO 8258 or Reference [13] in the Bibliography.

6.6.3 Interpretation of control charts, out-of-control situations

The quality-control chart is intended to identify changes in random or systematic error.

The following criteria for out-of-control situations are recommended for use with Shewhart charts:

— 1 control value being outside the action limit; or

— 2 consecutive values outside warning limits; or

— 7 consecutive control values with rising tendency; or

— 7 consecutive control values with falling tendency; or

— 10 out of 11 consecutive control values being on one side of the central line.

The following out-of-control situations apply to the range type of control chart if:

— a range $R\text{rel}_j$ falls outside the upper action limit; or

— a range $R\text{rel}_j$ falls below the lower action limit (valid only for LA $>$ 0); or

— 7 consecutive control values show an ascending/descending tendency; or

— 7 consecutive control values lie above the mean range, $R\text{rel}$.

A cyclic variation of ranges may be observed, for example, by a regularly scheduled maintenance of an analytical instrument or by re-preparation of reagents.

The Harmonised Guidelines [30] give rules that can be applied for simultaneous consideration of two control charts.

6.7 Conclusions

6.7.1 Measures to be taken in out-of-control situations

An out-of-control situation occurring on a control chart is strong evidence that an uncontrolled change has occurred in the analysis and it is likely that an important error applies to the analysis of the routine samples. It is very important to immediately take steps to identify and eliminate the cause of the error in order to re-establish control over the performance of the analytical system before further analysis is undertaken. For fast and effective identification of the source of analytical error, the approach described in the following subclauses is recommended.

© ISO 2009 – All rights reserved

6.7.1.1 Initial investigation to identify gross errors or deviations from the analytical procedure

The analysis of the control sample is repeated, strictly following the analytical method and avoiding possible gross errors. If the new result of the control sample shows that the method is under control again, it may be assumed that the method of analysis had not strictly been observed on the previous batch of analyses or that a gross error had occurred. The entire batch should then be re-analysed.

If, however, the result of the analysis of the control sample is erroneous but reproducible, a systematic error is very likely to exist.

6.7.1.2 Elimination of systematic errors

To check for systematic errors, several different trueness-control samples are analysed. To detect errors depending on the reagents or the method, control samples should be used whose concentrations cover the entire measuring range. As a minimum, a trueness-control sample in the lower part and one in the upper part of the working range should be used. In the event of a systematic error with results predominantly higher or lower than the actual values, a step by step examination should be performed to find the reason for this bias. Exchanging experimental parameters, such as reagents, apparatus or staff, might help to identify quickly this type of error.

6.7.1.3 Improving precision

The precision can also be improved by a step-by-step approach to find the causes of random error.

The total precision of an analytical method can be improved by examining its individual procedural steps to find the one which contributes most to the total error.

6.7.2 Plausibility control

There could be errors which may not be detected by a statistical approach to quality control. In the majority of such cases, this concerns errors influencing individual analyses in a batch, but not ones before or after. This type of error can only be revealed by means of plausibility controls, i.e. checks on the observed value in relation to expectations based on previous knowledge. Such knowledge may be based on chemical considerations, for example, checks on the equivalence of anions and cations in a sample, or a prior expectation, for example, that chemical oxygen demand (COD) will be greater than biological oxygen demand (BOD).

A successful approach to plausibility control requires that appropriate background information is available. The procedure of plausibility control may be subdivided into two parts:

— information/harmonization;

— control (for details, see Reference [26] in the Bibliography).

Plausibility control may form a worthwhile additional check to supplement routine analytical quality control. A large proportion of failures on the basis of plausibility control (which is not mirrored by routine quality control) suggests an inadequate routine system of quality control or a system which is not stable in its operation.

6.7.3 Further corrective action

In the event of repeatedly occurring out-of-control situations being detected in the control charts, the initial tests for implementation of analytical quality control, as described in Clause 4, should be performed with the matrix in question if the out-of-control situations cannot be remedied by simpler actions, such as exchange of vessels, apparatus or reagents.

© ISO 2009 – All rights reserved

6.8 Control charts with fixed quality critera (target control charts)

Contrary to the classical control charts of the Shewhart type described in 6.4, the target control charts operate without statistically evaluated values. The bounds for this type of control chart are given by external prescribed and independent quality criteria. A target control chart (for the mean, the true value, the blank value, the recovery rate, the range) is appropriate if

— there is no normal distribution of the values from the control sample (i.e. blank values),

— there are not enough data available for the statistical evaluation of the bounds, or

— there are external and internal prescribed bounds which should be applied to ensure the quality of analytical values.

6.8.1 Control samples

The control samples for the target control charts are the same as for the classical control charts as described in 6.4 to 6.6.

6.8.2 Definition of bounds

The bounds for target control charts can be taken from one of the following sources:

— requirements from legislation;

— standards of analytical methods and requirements for internal quality control (IQC);

— the (at least) laboratory-specific precision and trueness of the analytical value which had to be ensured;

— the valuation of laboratory-internal known data of the same sample type (see 6.4).

The chart is constructed with an upper and lower bound. A pre-period is inapplicable. The target control chart of the range needs only the upper bound.

6.8.3 Out-of-control situation

The analytical method is out-of-control if the analytical value is higher or lower than the respective prescribed bounds.

6.8.4 Measures in out-of-control situations

The measures are the same as described in 6.7.1.

6.8.5 Identification of changes in analytical quality

If applying the target control charts, the analytical method is formally out-of-control only if the analytical values are outside the defined bounds. Nevertheless, trends in the analytical quality should be identified and steps should be taken against them. Helpful hints are given in 6.7.

7 Quality control in sampling

Careful attention to the soundness of sampling and sample handling systems is essential if data of adequate accuracy are to be obtained, and it is therefore necessary to ensure that appropriate control tests are applied to these aspects of the overall process, as well as to analysis. Whilst many issues concerning sampling and sample handling lie outside the scope of this Technical Specification, it is fitting to point out the important role of quality control of sampling and sample handling.

© ISO 2009 – All rights reserved

It is emphasized that, as with the analytical stage, the initial selection of soundly based sampling procedures is of fundamental importance. Indeed, given the difficulty of assessing by practical tests many of the potential errors which may arise during sampling, the need for careful initial selection of equipment and procedures is probably even more crucial than in analysis. Similarly, control tests of sampling and sample handling have the same basic objectives as their counterparts in analysis, namely to ensure that any important deterioration of the accuracy of results arising from these steps is detected as rapidly as possible, so that corrective action can be taken.

Guidance on quality control and quality assurance of sampling is given in ISO 5667-14 and Reference [22] in the Bibliography.

8 Interlaboratory quality control

The aims of interlaboratory testing may be divided into two groups: collaborative studies and proficiency testing. Within both groups, several subgroups may be identified. The most important kinds of interlaboratory tests are as follows:

— collaborative studies for validation of a candidate method for standardization, as specified in ISO 5725;

— interlaboratory tests to determine a consensus (certified) value for the composition of a reference test material, as specified in ISO Guide 35;

— proficiency testing, as specified in ISO/IEC Guide 43 and ISO 13528;

— collaborative studies to estimate the accuracy of data produced by a group of laboratories which share a common interest using the Youden method [35], known as the paired sample technique.

For further details, see ILAC Guide 13 [29] and IUPAC Harmonized Protocol [34], as well as guidelines [12], [14], [15], [16], [24] and [25] in the Bibliography.

9 Quality control for lengthy analytical procedures or analysis undertaken infrequently or on an ad hoc basis

9.1 Quality control for lengthy analytical procedures

Some multi-stage analytical procedures, for example, the determination of trace organic contaminants, are capable of producing relatively few results at a time. This raises the question of how to implement quality control measures which were initially put into practice with high-throughput techniques. The argument that because organic analyses are time-consuming they should not be subject to performance tests of the same complexity as, for example, nutrient determinations is unsound.

An analytical result which takes hours to produce should be supported by performance and quality control information of at least the same reliability as that associated with “simple” determinations. Indeed, because trace analysis is subject to greater uncertainty and is more costly to repeat, it can be argued that proportionally more effort needs to be directed towards quality control. The maxim that relatively few results of known and adequate accuracy are better than many results of unknown and probably inadequate accuracy remains true.

The stated approach to tests of precision and recovery should not be regarded as an ideal only attainable under favourable circumstances. Rather, it is the minimum of testing which will provide a modestly reliable indication of performance. For trace analysis, there is a strong case for expanding the range of samples tested to include checks on precision and recovery from samples of differing matrices. Where limit of detection is of special interest, it is particularly important that a pooled estimate is obtained from many batches of tests. Replicate determinations performed on a single occasion are likely to give an unreliable and probably optimistic estimate.

© ISO 2009 – All rights reserved

Similarly, the approach to routine quality control should follow the recommendations given in Clause 6. Particular attention should be paid to the implementation of recovery control charts or other means of monitoring and controlling recovery through the whole process.

Annex D gives an example for a lengthy gas chromatographic/mass spectrometric (GC-MS) procedure.

9.2 Analysis undertaken infrequently or on an ad hoc basis

The procedures recommended for preliminary performance tests (Clause 4) and routine quality control (Clause 6) are most easily put into practice for analyses which are carried out regularly and often. It is necessary to consider what approach to quality control should be adopted for analyses which may be performed infrequently or which may be undertaken only once. The same considerations apply to analyses carried out over a short period in relatively few batches.

Two main features distinguish this type of analysis from frequent, regular determinations.

Firstly, any quality control activity is likely to take up a relatively large proportion of the total analytical effort compared with routine analyses. This is inconvenient and expensive, but it is a consequence of organizing analysis in this way. It should not be used as an excuse to avoid evaluation of the analytical system. Any analytical system used to produce data should be tested to provide an estimate of its performance. Not to test would be to provide data of unknown accuracy. This is unacceptable to users of analytical data. Tests as described in Clause 4 are recommended as a means of providing background performance data for all analytical systems.

Secondly, it is not possible to establish and maintain a state of statistical control in relatively few batches of analysis. This is an important drawback of not carrying out frequent, regular batches of analysis. It may be a consideration why analytical work might be subcontracted to laboratories having reason to perform the determination in question frequently. However, when analyses are carried out on a one-off or ad hoc basis the following approach is recommended.

The following quality-control measures should be applied:

— checks on spiking recovery in the matrix of interest;

— replicate analyses of samples;

— use of a difference control chart (see 6.5.5);

— use of field and procedural blanks;

— confirmation of the calibration using material from an independent source;

— use of reference materials [where an appropriate certified reference material (CRM) is available] as blind controls.

The proportion of samples analysed more than once should not be less than 20 % but could be as large as 100 % in the case of very small batches or highly important analyses. Single analysis of samples is an acceptable approach only when a state of statistical control can be established and maintained.

The Harmonized Guidelines [31] recommend that all test materials (samples) be analysed in duplicate. In addition, the use of spiking or recovery tests or use of a formulated control material, with different concentrations of analyte if appropriate, is recommended. If possible, procedural blanks should also be carried out. As no control limits are available, the estimates of bias and precision obtained should be compared with values derived from fitness for purpose.

© ISO 2009 – All rights reserved

Annex A
(informative)

Verification of the limit of detection and the limit of quantification

A.1 Fundamental considerations

What to measure?	What to estimate? (statistically or graphically)	Which criteria should be fulfilled?
Limit of detection		
$n \geqslant 3$ blank matrix samples $n \geqslant 3$ spiked matrix samples at the concentration level of the limit of detection which has to be verified	Maximum blank signal Mean of signals for spiked matrix samples at detection limit	Mean of signals from spiked matrix samples with detection limit concentration should give signals higher than the maximum blank signal.
Limit of quantification		
$n \geqslant 3$ spiked matrix samples at the concentration level of the limit of quantification which has to be verified	Computation of the concentration x via calibration function Computation of the confidence range of the mean concentration Other possibilities to estimate uncertainty of measurement	Uncertainty should be smaller than $1/k$ ($k > 1$) of limit of quantification.

A.2 Example for the verification of the limit of detection

Concentration	Replicates			Maximum	Mean
	y_{i1}	y_{i2}	y_{i3}	y_{max}	$\bar{y}_i$
Blank	0,001	18,196	13,387	18,196	—
Detection limit $x_{dl} = 3$ ng/l (spiked)	15,573	19,684	25,432	—	20,230

As the mean response of the matrix samples spiked at the limit of detection level is greater than the maximum blank, the estimated limit of detection (e.g. according to 4.4.2, 4.4.3 or 4.4.4) can be considered as verified for the respective sample matrix.

A.3 Example for the verification of the limit of quantification

In the following, the equation for the calculation of the required precision is developed. For the limit of quantification, a factor of $k = 3$ corresponds to a relative measurement uncertainty of approximately 33 %. This means that the half width of the double-sided confidence interval of the measurement result for the spiked matrix sample, Δx, shall be not more than 33 %.

The general equation for the correlation of uncertainty and limit of quantification is:

$$\frac{\Delta x}{x_{LQ}} = \frac{1}{k} \text{ with } \Delta x = \frac{t_{f;\alpha} \cdot s}{\sqrt{n}} \quad \text{(A.1)}$$

© ISO 2009 – All rights reserved

where

Δx is the half width of the confidence interval of the measurement result of the spiked sample;

x_{LQ} is the estimated limit of quantification (concentration of spiked sample);

k is the factor for calculating the limit of quantification as a multiple of the limit of detection, here: $k = 3$;

$t_{f;\alpha}$ is the table value of the t-distribution for $f = n - 1$ and $\alpha = 0{,}05$ ($P = 95$ %);

n is the number of analyses of the spiked sample;

α is the significance level;

s is the standard deviation of the results of the spiked sample.

For $k = 3$, Equation (A.2) applies:

$$\frac{t_{f;\alpha} \cdot s}{x_{LQ} \cdot \sqrt{n}} = \frac{1}{3} \tag{A.2}$$

Conversion for s results in:

$$s = \frac{x_{LQ} \cdot \sqrt{n}}{3 \cdot t_{f;\alpha}} \tag{A.3}$$

Depending on the number of analyses of the spiked sample, the value of the required standard deviation for verification of the limit of quantification differs. For statistical reasons, verification is more likely to succeed with increasing number of analyses. Therefore, a number of analyses between $n = 3$ and $n = 5$ is recommended.

EXAMPLES

If the spiked sample was analysed $n = 3$ times, the following applies:

$s = 0{,}134\ x_{LQ}$ with $t = 4{,}303$

If the spiked sample was analysed $n = 4$ times, the following applies:

$s = 0{,}210\ x_{LQ}$ with $t = 3{,}182$

If the spiked sample was analysed $n = 5$ times, the following applies:

$s = 0{,}268\ x_{LQ}$ with $t = 2{,}776$

The limit of quantification is verified if the standard deviation of the results of the spiked sample is less than or equal to the calculated value.

© ISO 2009 – All rights reserved

Annex B
(informative)

The nature and sources of analytical errors

B.1 General

The following clauses provide a succinct discussion of the nature and origin of errors in analytical results for waters and effluents. Further information on many of the topics covered is given elsewhere in this Technical Specification, and the subject is also discussed extensively in Reference [28] in the Bibliography.

B.2 Nature of errors

B.2.1 Total error

The total error, E, of an analytical result, R, is defined as the difference between that result and the true value, T, i.e.

$$E = R - T$$

B.2.2 Random error

Repeated analysis of identical portions of the same, homogeneous sample does not, in general, lead to a series of identical results [1)]. Rather, the results are scattered about some central value. The scatter is attributed to random error, so called because the sign and magnitude of the error of any particular result vary at random and cannot be predicted exactly. Precision is said to improve as the scatter becomes smaller, i.e. as random error decreases, and imprecision is therefore a synonym for random error.

Because random errors are always present in analytical results, statistical techniques are necessary if correct inferences regarding true values are to be made from the results.

Terms such as “repeatability” and “reproducibility” have specialized meanings in the context of interlaboratory collaborative trials. In this Technical Specification, random error is quantified in terms of the standard deviation, σ. Since exact measurement of the standard deviation generally requires an infinite number of repeated results, only estimates, s, of σ will usually be obtainable. The number of degrees of freedom (DF) of the estimate provides an indication of its worth; as the number of degrees of freedom increases, the random error of the estimate itself, s, decreases.

B.2.3 Systematic error

Systematic error (or bias) is present when there is a persistent tendency for results to be greater, or smaller, than the true value. The mean of n analytical results for identical portions of a stable, homogeneous sample approaches a definite limiting value, μ, as n is increased indefinitely. When μ differs from the true value, T, results are said to be subject to systematic error or bias, β, where:

$$\beta = \mu - T$$

1) This may not be true when the discrimination of the analytical system is coarse. However, the apparent perfect concordance of repeated results in such a situation is illusory, because samples differing in concentration will also give the same results.

© ISO 2009 – All rights reserved

Because an indefinitely large number of determinations cannot be made on a single sample, the effect of random error prevents exact determination of μ, and hence also of β. Only an estimate, $\hat{x}$, of μ will generally be available, so that only an estimate, b of β can be obtained.

As the systematic error or bias of results decreases, trueness is said to increase.

B.3 Sources of error

The distinction between random and systematic errors is important for two reasons: first, because they have different effects on the use to be made of analytical results, and second, because they usually have different origins.

B.3.1 Causes of random error

Random errors arise from uncontrolled variations in the conditions of the analytical system [2)] during different analyses. The nature of such variations include, for example, differences in the volume of sample or reagent taken on different occasions, fluctuations in temperature, either in time, or across the different sample positions in a heating bath, block or oven, fluctuations in instrumental conditions (for example, in temperatures, fluid flowrates, voltages and wavelengths) and operator-induced variations in reading scales. Variations from batch to batch, in the extent to which the calibration function represents the true calibration for that batch, also give rise to between-batch random errors, whereas a consistent calibration error across many batches gives rise to systematic error, see below.

Whilst many of these factors causing random errors can often be more closely controlled to achieve better precision, they can never be totally eliminated, so that all results are subject to some degree of random error.

B.3.2 Causes of systematic error

There are five general sources of systematic error (if clear blunders by the analyst in following the written method and bias introduced by the sample collection itself are both excluded).

These are as follows.

a) Instability of samples between sample collection and analysis

   This is a potentially important source of error in many cases, and evidence should always be obtained, either from the literature or by a direct test, to ensure that unacceptable bias is not introduced by this factor. Effective sample stabilization procedures are available for many determinands, but they should be compatible with the analytical system being employed, and with the particular sample type being analysed.

   For guidance on the preservation of samples, see ISO 5667-3.

b) Inability to determine all relevant forms of the determinand

   Many substances exist in water in a variety of physical and/or chemical forms (or “species”). For example, iron can exist in both dissolved and particulate forms, and within each of those physical categories a variety of chemical species may be present, for example, free ions and complexes, including those of different oxidation states, in the dissolved phase. An inability of the analytical system to determine some of the forms of interest will give rise to a bias when those forms are present in samples.

   Some determinands are overall properties of a sample, rather than a particular substance, for example, biochemical oxygen demand (BOD). Such determinands are called “non-specific” and have to be carefully

2) The analytical system is the combination of all factors, analyst, equipment, method, reagents, etc., involved in producing analytical results from samples.

© ISO 2009 – All rights reserved

defined by specifying the use of a particular analytical method. The so-called “dissolved” fractions of, for example, trace metals, are also non-specific in the sense that the type and pore-size of filter to be used in their determination should be clearly specified.

c) Interferences

Few analytical methods are completely specific for the determinand. Response to another substance (for example, response to iron by a spectrophotometric procedure for manganese based on formaldoxime) will give rise to biased results when that substance is present in samples, and it is important that the effects of all such interferents likely to be present in samples are known before a new method is applied routinely.

In some cases, the effect of another substance is to alter the chemical state of the determinand such that it is not measured by the method being used, for example, the presence of fluoride will cause aluminium complexes to form, which may not be measured by an ion-selective electrode. Such an effect can be regarded as an interference upon the determination of total dissolved aluminium, or as a failure to recover all forms of dissolved aluminium. Although it more strictly falls into the latter category, the effect, and others like it, may be most conveniently treated as an interference when data on performance characteristics are being obtained or reported.

d) Biased calibration

Most methods require the use of a calibration function (explicit or implicit) to convert the primary analytical response for a sample to the corresponding determinand concentration. If the samples and calibration standards are treated in exactly the same manner (and provided that the materials used to prepare the calibration standards are of adequate purity) no systematic error should arise from the calibration. (It has been noted in B.3.1 that any variations in the correctness of the calibration from batch to batch will be manifested as between-batch random errors.)

If, however, samples and calibration standards are treated differently, this can represent a potentially serious source of error. Thus, for example, a method prescribing some form of pre-concentration of the determinand from samples, but employing direct calibration with standards not taken through the pre-concentration step, will give rise to negative bias if the pre-concentration recovery is less than 100 %. In such cases, evidence should be obtained on the accuracy of the prescribed calibration, or the difference in treatment of samples and standards should be eliminated.

Impurity of the material used to prepare calibration standards is, of course, another potential cause of biased results.

e) Biased blank

The same considerations as in d) above apply to blanks. There is, however, another source of bias arising from blank correction. If the water used for the blank contains the determinand, results for samples will be biased low by an equivalent amount unless a correction for the determinand content of the blank water is applied. Ideally, however, a source of blank water should be obtained, such that the determinand content is negligible in comparison with the concentration in samples.

© ISO 2009 – All rights reserved

Annex C
(informative)

Estimating the measurement uncertainty

C.1 Different approaches for estimating the uncertainty

There are in principle two different top-down models to estimate the uncertainty of measurement. In the Nordtest approach [32] the uncertainty component of the random errors will be estimated from the internal quality control and the uncertainty component of the systematic errors from proficiency tests, measurement of CRM or validation studies. The second procedure, described in detail in ISO/TS 21748, uses the results of collaborative studies established under the principles of ISO 5725. Here the reproducibility standard deviation is the decisive figure (see C.3).

C.2 NORDTEST approach

In general, the uncertainty contributions from two kinds of error, random and systematic errors, are determined independently by using suitable validation data. The procedure for the determination of measurement uncertainty essentially consists of the steps shown in Figures C.1 and C.2.

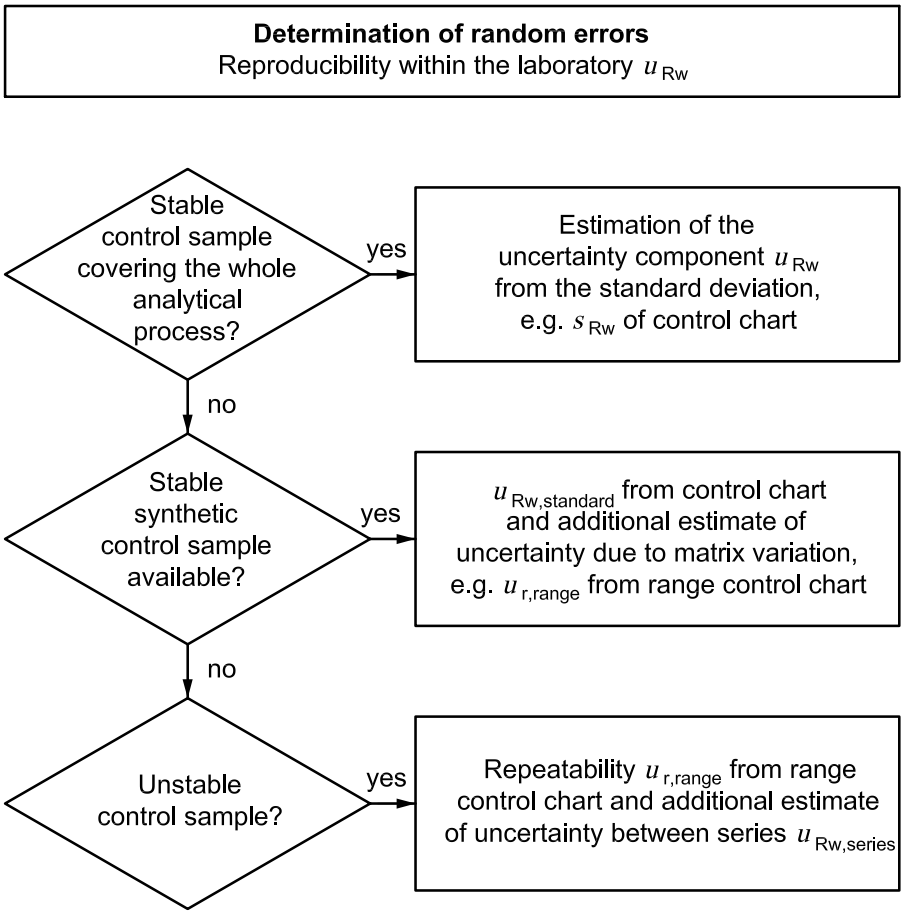


Figure C.1 — Estimation of the random component of measurement uncertainty

© ISO 2009 – All rights reserved

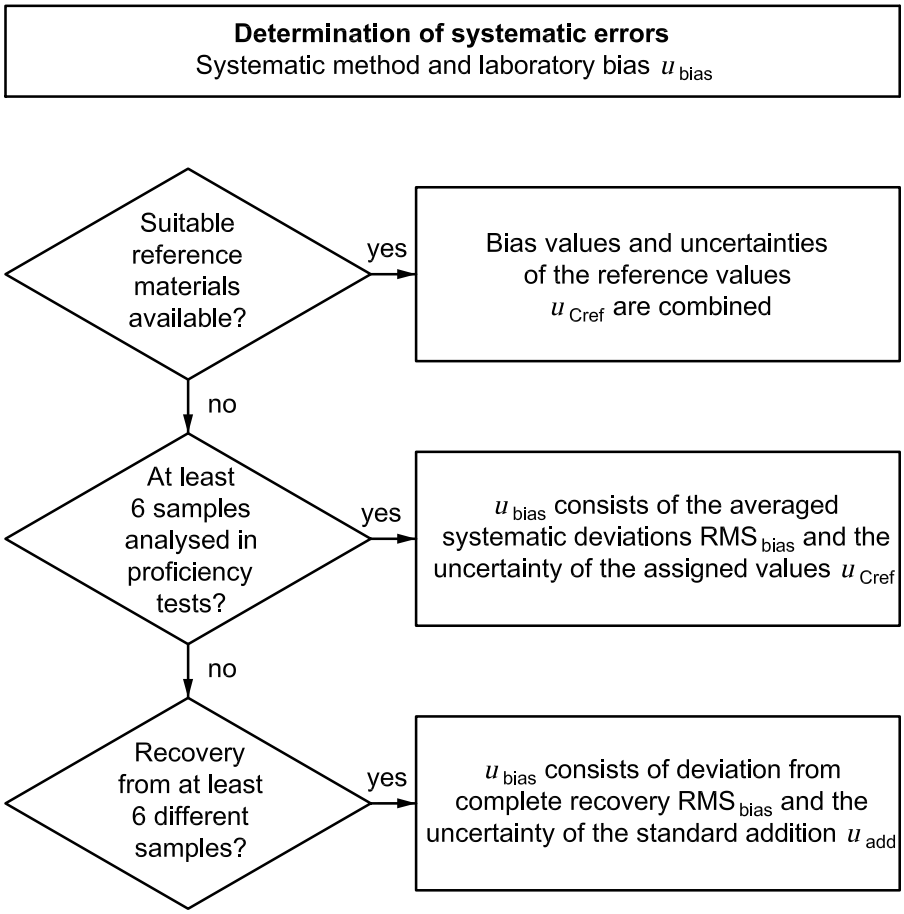


Figure C.2 — Estimation of the systematic component of measurement uncertainty

The combined standard uncertainty u_c is calculated as the square root of the quadratic sum of the random component u_{Rw} and the systematic component u_{bias}:

$$u_c = \sqrt{u_{Rw}^2 + u_{bias}^2} \tag{C.1}$$

Finally, the combined uncertainty is multiplied by the factor 2 to get the expanded uncertainty.

C.3 Estimation of measurement uncertainty from reproducibility standard deviation

If other data are not available, measurement uncertainty can be roughly estimated using the reproducibility standard deviation s_R from interlaboratory trials. For this purpose, s_R is multiplied by a factor of 2 and the result is an estimate for the expanded uncertainty, U (for more details see ISO/TS 21748).

A second step is to prove whether the internal precision data of the laboratory coincide with the repeatability standard deviation of the collaborative study. If they differ significantly it is recommended that the internal repeatability standard deviation, s_i, is used instead of the repeatability standard deviation, s_r, of the collaborative study. This can lead to higher or lower values of the measurement uncertainty.

© ISO 2009 – All rights reserved

Annex D
(informative)

Example for performing quality control for lengthy analytical procedures

Table D.1 is based on experiences of a laboratory which uses gas chromatography/mass spectrometry (GC/MS). Each week the system runs on 2 subsequent days and produces 60 results (including replicates, calibration and blank values). The remaining 3 days per week are used for preparation and evaluation of the analyses.

Table D.1 — Analytical quality control (AQC) for a GC-MS method

Step No.	Time after starting weeks	What to do to estimate precision	What to do to estimate trueness
1	0	Replicate analyses ($N \geqslant 6$) of selected analytes and matrices	Recovery tests with the most important analytes
2	1 to 4 or 5	Stepwise replacement of replicate analyses by range control charts or mean control charts (if there is no matrix influence)	1) Stepwise introduction of mean control charts or, if matrix influences the results, recovery rate control charts [a] 2) Blank control charts [a]
3	5 and later	Selected control charts [a]	Relevant control charts (mean, recovery rate, blank control charts) [a]
Overall, the AQC expenditure is reduced to about 33 %.			

[a] If there is no normal distribution, the corresponding target control charts can be used.

© ISO 2009 – All rights reserved

Bibliography

[1] ISO 5667-3, *Water quality — Sampling — Part 3: Guidance on the preservation and handling of water samples*

[2] ISO 5667-14, *Water quality — Sampling — Part 14: Guidance on quality assurance of environmental water sampling and handling*

[3] ISO 6107-2:2006, *Water quality — Vocabulary — Part 2*

[4] ISO 8258, *Shewhart control charts*

[5] ISO 11843-1:1997, *Capability of detection — Part 1: Terms and definitions*

[6] ISO 11843-2:2000, Capability of detection — Part 2: Methodology in the linear calibration case

[7] ISO 11843-3:2003, *Capability of detection — Part 3: Methodology for determination of the critical value for the response variable when no calibration data are used*

[8] ISO 11843-4:2003, *Capability of detection — Part 4: Methodology for comparing the minimum detectable value with a given value*

[9] ISO/TS 21748:2004, *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*

[10] ISO/IEC Guide 98-3:2008, *Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement* (GUM:1995)

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

[12] ASTM E691-05, *Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method*

[13] ASTM MNL 7A, *Manual on Presentation of Data and Control Chart Analysis*, Seventh Edition (2002)

[14] ASTM E1763-06, *Standard Guide for Interpretation and Use of Results from Interlaboratory Testing of Chemical Analysis Methods*

[15] ASTM E1301-95(2003), *Standard Guide for Proficiency Testing by Interlaboratory Comparisons*

[16] ASTM E1601-98(2003)e1, *Standard Practice for Conducting an Interlaboratory Study to Evaluate the Performance of an Analytical Method*

[17] BARWICK, V.J., ELLISON, S.L.R.: (VAM Project 3.2.1 Development and harmonisation of measurement uncertainty principles, part (d)): *Protocol for uncertainty evaluation from validation data* (Jan. 2000)

[18] EURACHEM Guide (1998), *The Fitness for Purpose of Analytical Methods: A laboratory guide to method validation and related topics*, LGC (Teddington) Ltd.

[19] EURACHEM/CITAC Guide (2000), Q*uantifying Uncertainty in Analytical Measurement*, 2nd edition

[20] EURACHEM/CITAC Guide (2002), *Guide to Quality in Analytical Chemistry — An Aid to Accreditation*

[21] EURACHEM/CITAC Guide (2003), *Traceability in Chemical Measurement*

© ISO 2009 – All rights reserved

[22] EURACHEM/CITAC Guide (2007), *Measurement uncertainty arising from sampling — A guide to methods and approaches*

[23] Eurolab Technical Report (2002), *Measurement Uncertainty in Testing*

[24] European Cooperation for Accreditation of Laboratories EAL (1996), *Interlaboratory Comparisons* (EAL-P7)

[25] European Cooperation for Accreditation (2001), *Use of profciency testing as a tool for accreditation in testing* (EA-03/04)

[26] FUNK, W., DAMMANN, V., DONNEVERT, G. (2007), *Quality Assurance in Analytical Chemistry*, second, completely revised and enlarged edition, Wiley-VCH, Weinheim

[27] GARDNER, M.J., WILSON, A.L., and CHEESMAN, R.J. (1989) *A Manual on Analytical Quality Control for the Water Industry*, NS30, WRc Medmenham

[28] HUNT, D.T.E. and WILSON, A.L. (1986), *The Chemical Analysis of Water. General Principles and Techniques*, Second edition, The Royal Society of Chemistry, London

[29] ILAC-G13:2000: *Guidelines for the Requirements for the Competence of Providers of Proficiency Testing Schemes* (http://www.ilac.org/)

[30] IUPAC: Harmonised Guidelines for Internal Quality Control (IQC) in Analytical Chemistry Laboratories, ISO/IUPAC/AOAC, *Pure and Applied Chemistry*, Vol. 67, No. 4, pp. 649-666, 1995

[31] IUPAC: Harmonized guidelines for single-laboratory validation of methods of analysis, *Pure and Applied Chemistry*, Vol. 74, pp. 835-855, 2002

[32] MAGNUSSON, B., NÄYKKI, T., HOVIND, H., KRYSELL, M: *Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories*. NORDTEST Report TR 537, Edition 2, 2004 (http://www.nordicinnovation.net/nordtestfiler/tec537.pdf)

[33] National Committee for Clinical Laboratory Standards, *Internal Quality Control Testing: Principles and Definitions*, http://www.westgard.com/quest7.htm

[34] THOMPSON, M.; ELLISON, S.L.R., WOOD, R.: The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories (IUPAC Technical Report); *Pure and Applied Chemistry*, Vol. 78, No. 1, pp. 145-196, 2006

[35] YOUDEN, W.J. (1972), Graphical diagnosis of interlaboratory test results. *Journal of Quality Technology*, Vol. 4, No. 1, pp. 29-33

[36] EEE/RM/062: *The selection and use of reference materials. A basic guide for laboratories and accreditation bodies* (2002). Available from the Eurachem secretariat or website (http://www.eurachem.org)

© ISO 2009 – All rights reserved

ICS 13.060.45

Price based on 39 pages

© ISO 2009 – All rights reserved