

Manual for Marine Monitoring in the

COMBINE

Programme of HELCOM

Part B

General guidelines on

quality assurance

for monitoring in the Baltic Sea



Last updated: 31.03.2006

PART B. GENERAL GUIDELINES ON QUALITY ASSURANCE FOR MONITORING IN THE BALTIC SEA

B.1. Introduction	1
B.1.1 Need for Quality Assurance of Analytical Procedures in Marine Monitoring.....	1
B.1.2 Objective	2
B.1.3 Topics of Quality Assurance	2
B.1.4 Units and conversions	3
B.2. The Quality System	6
B.2.1 General.....	6
B.2.2 Scope.....	6
B.2.3 Organization, Management and Staff.....	7
B.2.3.1 Organization	7
B.2.3.2 Management	7
B.2.3.3 Staff	8
B.2.4 Documentation	8
B.2.5 Laboratory Testing Environment	8
B.2.6 Equipment.....	9
B.2.7 Quality Audit	10
B.3. Specifying analytical requirements.....	10
B.3.1 General.....	10
B.3.2 Variable of Interest	11
B.3.3 Type and Nature of the Sample and its Environment.....	11
B.3.4 Concentration Range of Interest.....	11
B.3.5 Permissible Tolerances in Analytical Error.....	11
B.3.6 Technical note on the QA of the determination of co-factors.....	12
Co-factors, Definition, and Use.....	12
Co-factors in Biota Analysis	12

Co-factors in Water Analysis	13
QA Information to Support the Data.....	13
References.....	13
B.4. Validation of analytical methods	13
B.4.1 General.....	13
B.4.2 Validation	14
B.4.2.1 Selectivity	14
B.4.2.2 Sensitivity	15
B.4.2.3 Limit of detection, limit of quantification	15
General	15
Definitions	15
References	17
B.4.2.4 Range.....	17
B.4.2.5 Accuracy	18
B.4.2.5.1 Estimating random errors	19
B.4.2.5.2 Estimating systematic errors (biases).....	20
B.4.2.6. Measurement uncertainty	21
Introduction.....	21
Definitions	21
Procedures to Estimate Uncertainty	22
Uncertainty in dependence on analyte levels.....	23
Reporting of Uncertainty.....	24
B.4.3 Validation of data (new chapter)	24
B.4.3.1 Data screening.....	24
B.4.3.2 Treatment of suspect and missing data	24
B.5. Routine quality control (use of control charts).....	25
B.5.1 General.....	25
B.5.2 Control of trueness	25
B.5.2.1 X-charts.....	26
B.5.2.2 Blank control chart	26
B.5.2.3 Recovery control chart	27

B.5.3 Control of precision.....	27
B.5.3.1 R-chart	28
B.5.4. Control charts with fixed quality criteria (target control charts)	28
B.5.5. Control charts for biological measurements.....	29
B.5.6. Interpretation of control charts, out-of-control situations.....	30
B.5.7. Selection of suitable control charts	31
B.5.8. References.....	32
B.6. External quality assessment	32
B.7. Definitions.....	33

Foreword

This document provides an introduction to quality issues, in general, and quality assurance in Baltic marine monitoring laboratories, in particular. The guidelines are intended to assist laboratories in starting up and operating their quality assurance systems. For laboratories with existing quality systems, the guidelines may give inspiration for issues that can be improved. The guidelines contain information for all levels of staff in the marine laboratory.

Sections [1](#), [2](#), [3](#) and [6](#) together with [Annexes B-1](#) (Quality manual) and [B-3](#) (Quality audit) give guidance on organizational technical quality assurance principles that are relevant to *administrative managers*.

Sections [1](#), [2](#), [5](#) and [6](#) with [Annexes B-1](#) (Quality manual), [B-7](#) (Reference materials), and [B-3](#) (Quality audit), regarding the implementation and operation of a quality system, are the main sections of relevance for *quality managers*.

For *technical managers* all sections in the main part of the document are relevant. The guidelines provide technical managers with a description of the principles concerning how to introduce and maintain the technical aspects of quality assurance.

It is believed that *analysts* will find all of the guidelines and annexes relevant regarding optimization of their analytical work. The applicability of [Annexes B-6](#) (Sampling), [B-10](#) (Technical notes on nutrients) and [B-12](#) (Technical notes on contaminants) will, however, depend on the specific job description of each analyst.

It is the intention of the guidelines that other members of the *staff connected to the Monitoring Programme* can find use for specific parts of the guidelines.

B.1. INTRODUCTION

B.1.1 NEED FOR QUALITY ASSURANCE OF ANALYTICAL PROCEDURES IN MARINE MONITORING

It has been seen that, although there has been considerable improvement in analytical procedures over the past two decades, it has been obvious that a large number of European laboratories which still had difficulties in providing reliable data in routine work (Topping, 1992; HELCOM, 1991; ICES, 1997). Topping based his conclusion on the results of a series of external quality assessments of analysis (generally referred to as intercomparison exercises), organized over the last twenty years by the International Council for the Exploration of the Sea (ICES), and which have shown that there are large interlaboratory differences.

As a consequence of improperly applied measures to assure the quality of analytical data, information about variations of levels both in space and time is often uncertain or misleading, and the effects of political measures to improve the quality of the marine environment cannot be adequately assessed. Therefore, the acquisition of relevant and reliable data is an essential component of any research and monitoring programme associated with marine environmental protection. To obtain such data, the whole analytical process must proceed under a well-established Quality Assurance (QA) programme. Consequently, the HELCOM Environment Committee (EC) at its fifth meeting (HELCOM, 1994) recommended that: 'all institutes reporting data to BMP/CMP shall introduce in-house quality assurance procedures'.

In addition, the following principles of a quality assurance policy were formulated:

QUALITY ASSURANCE POLICY OF THE HELSINKI COMMISSION (HELCOM, 1995)

1. Contracting Parties acknowledge that only reliable information can provide the basis for effective and economic environmental policy and management regarding the Convention area;
2. Contracting Parties acknowledge that environmental information is the product of a chain of activities, constituting programme design, execution, evaluation and reporting, and that each activity has to meet certain quality requirements;
3. Contracting Parties agree that quality assurance requirements be set for each of these activities;
4. Contracting Parties agree to make sure that suitable resources are available nationally (e.g., ships, laboratories) in order to achieve this goal;
5. Contracting Parties fully commit themselves to following the guidelines, protocols, etc., adopted by the Commission and its Committees in accordance with this procedure of quality assurance.

The Contracting Parties shall clearly declare, in relevant data reports, if they fail to fulfil the recommendations of the Manual. If alternative methods are being used, proof shall be given that the results are comparable with results generated from methods described in the Manual. The supplier of the data has then the responsibility to proof the comparability of the methods. HELCOM Monitoring and Assessment Group (HELCOM MONAS) will ultimately decide if the data could be incorporated in the HELCOM Database. If agreement could not be reached, HELCOM/ICES Steering Group on Quality Assurance

of Chemical Measurement (SGQAC) will be given the task to evaluate the method and give advice to HELCOM MONAS.

All Contracting Parties have nominated persons responsible for quality assurance in all laboratories reporting to the monitoring programmes.

All institutes/laboratories should participate in regular (annual) intercomparison exercises, arranged in the Baltic community and laboratories should take part in proficiency testing schemes, e.g. the QUASIMEME-II. As new certified reference materials become (commercially) available these might be used by all participating institutes or laboratories.

The results of intercomparison exercises and the analyses of certified standards should be reported together with the monitoring data according to procedures to be decided by EC MON. It should also be noted that it is possible for laboratories to authorize the QUASIMEME and the BEQUALM office to report the individual laboratories performance data directly to the data host for HELCOM.

The monitoring laboratories should have a QA/QC system that follows the requirements of EN ISO/IEC 17025 "General requirements for the competence of testing and calibration laboratories" (formerly EN 45001 and ISO Guide 25). Participating laboratories are encouraged to endeavour the obtainment of official accreditation (or certification) for the variables on which they report data in accordance with COMBINE.

In order to assist laboratories in setting up their quality assurance system the general advice that follows in this chapter is applicable.

B.1.2 OBJECTIVE

The objective of the Manual outlined here is to support laboratories working in marine monitoring to produce analytical data of the required quality. The Manual may also help to establish or improve quality assurance management in the laboratories concerned. The technical part of the Manual provides advice on more practical matters. The Manual will, for chemical variables, not focus on sampling in detail, since this will be dealt with at a later stage. The details for biological sampling are found in Annexes C-4 to C-12.

B.1.3 TOPICS OF QUALITY ASSURANCE

In practice, Quality Assurance applies to all aspects of analytical investigation, and includes the following principal elements:

- A knowledge of the purpose of the investigation is essential to establish the required data quality.
- Provision and optimization of appropriate laboratory facilities and analytical equipment.
- Selection and training of staff for the analytical task in question.
- Establishment of definitive directions for appropriate collection, preservation, storage and transport procedures to maintain the integrity of samples prior to analysis.

- Use of suitable pre-treatment procedures prior to analysis of samples, to prevent uncontrolled contamination and loss of the determinand in the samples.
- Validation of appropriate analytical methods to ensure that measurements are of the required quality to meet the needs of the investigations.
- Conduct of regular intralaboratory checks on the accuracy of routine measurements, by the analysis of appropriate reference materials, to assess whether the analytical methods are remaining under control, and the documentation and interpretation of the results on control charts.
- Participation in interlaboratory quality assessments (proficiency testing schemes, ring-tests, training courses) to provide an independent assessment of the laboratory's capability of producing reliable measurements.
- The preparation and use of written instructions, laboratory protocols, laboratory journals, etc., so that specific analytical data can be traced to the relevant samples and *vice versa*.

B.1.4 UNITS AND CONVERSIONS

This notes summarizes the units that should be used for data submission within the COMBINE programme, and also gives the relevant formulas for conversion between different commonly used units.

References are made to the appropriate annexes of the COMBINE Manual.

Please note that the units dm^3 and cm^3 are used throughout the note, although the units l (litre) and ml (millilitre) would be equally correct.

Part 1: Units

Parameter	Symbol	Unit	Comment
Temperature	t	C	see Annex C-2
Salinity	S		see Annex C-2 according to the current definition of the Practical Salinity Scale of 1978 (PSS78)
Secchi depth (light attenuation)		m	see Annex C-2
Current speed		cm/s	see Annex C-2
Current direction			report as compass directions; see Annex C-2
Dissolved Oxygen	DO	cm^3/dm^3	see Annex C-2

Parameter	Symbol	Unit	Comment
Oxygen saturation			reported as fraction (%), see Annex C-2
Hydrogen Sulphide		$\mu\text{mol}/\text{dm}^3$	see Annex C-2
Nutrients		$\mu\text{mol}/\text{dm}^3$	as N, P or Si; see Annex C-2
Total P and N	TP/TN	$\mu\text{mol}/\text{dm}^3$	see Annex C-2
pH			NBS-scale; see Annex C-2
Alkalinity		mmol/dm^3	as carbonate; see Annex C-2
Particulate and dissolved organic matter (TOC, POC, DOC and PON)		$\mu\text{mol}/\text{dm}^3$	as C or N; see Annex C-2
Humic matter			depending on way of calibration; see Annex C-2
Heavy metals in water		ng/dm^3 or pg/dm^3	dissolved
Halogenated organics in water		ng/dm^3	
PAHs in water		ng/dm^3	
Heavy metals in biota		$\mu\text{g}/\text{kg}$	wet weight
Halogenated organics in biota		$\mu\text{g}/\text{kg}$ or ng/kg	wet weight, reported together with lipid content
Total suspended matter load		mg/dm^3	
Chlorophyll a	Chl-a	mg/m^3	see Annex C-4
Primary production (as carbon uptake)		$\text{mg}/\text{m}^3\cdot\text{h}$	see Annex C-5
Phytoplankton species			see Annex C-6
abundance		Counting units/ dm^3	
biomass		mm^3/dm^3	
Mesozooplankton			see Annex C-7

Parameter	Symbol	Unit	Comment
abundance		Individuals/m ³	
biomass		mm ³ /m ³ ; mg/m ³	
Macrozoobenthos			see Annex C-8
abundance		Counting units/m ²	
biomass		g/m ²	dry or wet weight

Part 2: Conversions

Parameter	From	To	Formula or multiplication factor
Any compound	g/dm ³	mol/dm ³	(g/dm ³)/molar weight
	mol/dm ³	g/dm ³	(mol/dm ³)* molar weight
	μmol/kg	μmol/dm ³	(μmol/kg)*density; density determined from salinity, temperature and pressure
	μmol/dm ³	μmol/kg	(μmol/dm ³)/density; density determined from salinity, temperature and pressure
Dissolved oxygen	mg/dm ³	cm ³ /dm ³	0.7
	cm ³ /dm ³	mg/dm ³	1.429
	μmol/dm ³	cm ³ /dm ³	11.196
	cm ³ /dm ³	μmol/dm ³	0.0893
	mg/dm ³	μmol/dm ³	0.06251
	μmol/dm ³	mg/dm ³	15.997
	DO	Oxygen saturation	see Grasshoff et al., Methods of Seawater Analysis, 2nd or 3rd edition
	Oxygen saturation	DO	see Grasshoff et al., Methods of Seawater Analysis, 2nd or 3rd

Parameter	From	To	Formula or multiplication factor edition
Hydrogen sulphide	$\mu\text{mol}/\text{dm}^3$	Negative oxygen	– 0.044001 (multiplication factor)
	Negative oxygen	$\mu\text{mol}/\text{dm}^3$	– 22.727 (multiplication factor)

B.2. THE QUALITY SYSTEM

B.2.1 GENERAL

'Quality system' is a term used to describe measures which ensure that a laboratory fulfills the requirements for its analytical tasks on a continuing basis. A laboratory should establish and operate a Quality System adequate for the range of activities, i.e., for the type and extent of investigations, for which it has been employed.

The Quality System must be formalized in a Quality Manual which must be maintained and up-to-date. A suggested outline of a Quality Manual is given in Annex B-1. Some comments and explanations are given in this section.

The person responsible for authorization and compilation of the Quality Manual must be identified, and an identification of holders of controlled copies should be listed in the manual.

The Quality System must contain a statement of the intentions of the laboratory top management in relation to quality in all aspects of its work (statement on Quality Policy).

For chemical variables guidance on the interpretation of ISO/IEC/EN 17025 'General Requirements for the Competence of Testing and Calibration of Laboratories' (formerly EN 45001 and ISO Guide 25) was given by a joint international EURACHEM/WELAC Working Group (EURACHEM/WELAC, 1992). Specific guidance to Analytical Quality Control for Water Analysis was elaborated by European - CEN/TC 230 (EN 14996) - as well as by international - ISO/TC 147 SC 7 (ISO/TR 13530) - standardisation authorities. All these publications have been taken into consideration when drafting these guidelines. References, which deal with specific aspects of quality assurance of chemical measurements, are cited in the text.

B.2.2 SCOPE

The laboratory's scope should be formulated in terms of:

- the range of products, materials or sample types tested or analysed;
- the types of tests or analyses carried out;
- the specification of method/equipment/technique used;
- the concentration range and accuracy of each test and analysis.

B.2.3 ORGANIZATION, MANAGEMENT AND STAFF

B.2.3.1 ORGANIZATION

The Quality System should provide general information on the identity and legal status of the laboratory and should include a statement of the technical role of the laboratory (e.g., employed in marine environmental monitoring).

The following information must be included in an organizational chart:

- Technical Manager, Quality Manager, and any deputies;
- general lines of responsibility within the laboratory (including the relationship between management, technical operations, quality control and support services);
- the lines of responsibility within individual sections of the laboratory;
- the relationship between the laboratory and any parent or sister organizations.
- The appropriate chart should show that, for matters related to quality, the Quality Manager has direct access to the highest level of management at which decisions are taken on laboratory policy and resources, and to the Technical Manager.

B.2.3.2 MANAGEMENT

Job descriptions, qualifications, training and experience are necessary for:

- Technical Manager,
- Quality Manager,
- other key laboratory managerial and technical posts.

Job descriptions should include:

- title of job and brief summary of function,
- person or functions to whom jobholder reports,
- person or functions that report to jobholder,
- key tasks that jobholder performs in the laboratory,
- limits of authority and responsibility.

The Technical Manager. The Quality System should include a statement that the post-holder has overall responsibility for the technical operation of the laboratory and for ensuring that the Quality System requirements are met.

The Quality Manager. The Quality System should include a statement that the post-holder has responsibility for ensuring that the requirements for the Quality System are met continuously and that the post-holder has direct access to the highest level of management at which decisions are taken on laboratory policy or resources, and to the Technical Manager.

The Quality System should state explicitly the Quality Manager's duties in relation to control and maintenance of documentation, including the Quality Manual, and of specific procedures for the control, distribution, amendment, updating, retrieval, review and approval of all documentation relating to the calibration and testing work of the laboratory.

B.2.3.3 STAFF

The laboratory management should define the minimum levels of qualification and experience necessary for engagement of staff and their assignment to respective duties.

Members of staff authorized to use equipment or perform specific calibrations and tests should be identified.

The laboratory should ensure that all staff receive training adequate to the competent performance of the tests/methods and operation of equipment. A record should be maintained which provides evidence that individual members of staff have been adequately trained and their competence to carry out specific tests/methods or techniques has been assessed. Laboratory managers should be aware that a change of staff might jeopardize the continuation of quality.

B.2.4 DOCUMENTATION

Necessary documentation includes:

- a clear description of sampling equipment;
- a clear description of all steps in the sampling procedure;
- a clear description of the analytical methods;
- a strict keeping of ship and laboratory journals;
- instrument journals;
- protocols for sample identification;
- clear labelling of samples, reference materials, chemicals, reagents, volumetric equipment, stating date, calibration status, concentration or content as appropriate and signature of the person responsible.

B.2.5 LABORATORY TESTING ENVIRONMENT

Samples, reagents and standards should be stored and labelled so as to ensure their integrity. The laboratory should guard against deterioration, contamination and loss of identity.

The laboratory should provide appropriate environmental conditions and special areas for particular investigations.

Staff should be aware of:

- the intended use of particular areas,
- the restrictions imposed on working within such areas,
- the reasons for imposing such restrictions.

B.2.6 EQUIPMENT

As part of its quality system, a laboratory is required to operate a programme for the necessary maintenance and calibration of equipment used in the field and in the laboratory to ensure against bias of results.

General service equipment should be maintained by appropriate cleaning and operational checks where necessary. Calibrations will be necessary where the equipment can significantly affect the analytical result.

The correct use of equipment is critical to analytical measurements and this equipment must be maintained, calibrated and used in a manner consistent with the accuracy required of data. For certain chemical analysis, one should consider that measurements can often be made by mass rather than by volume.

Particularly for trace analyses, contamination through desorption of impurities from, or uncontrolled determinand losses through sorption on, surfaces of volumetric flasks can be significant. Therefore, special attention should be paid to the selection of appropriate types of material (quartz, PTFE, etc.) used for volumetric equipment and its proper cleaning and conditioning prior to analysis.

Periodic performance checks should be carried out at specific intervals on measuring instruments (e.g., for response, stability and linearity of sources, sensors and detectors, the separating efficiency of chromatographic systems).

The frequency of such performance checks will be determined by experience and based on the need, type and previous performance of the equipment. Intervals between checks should be shorter than the time the equipment has been found to take to drift outside acceptable limits and should be given in the equipment list.

B.2.7 QUALITY AUDIT

The ISO 9000 and ISO 14000 series of International Standards emphasize the importance of audits as a management tool for monitoring and verifying the effective implementation of an organizations quality and/or environmental policy. Audits are also an essential part of conformity assessment activities such as accreditation.

So it is stated in ISO/IEC 17025:2005 General Requirements for the Competence of Testing and Calibration Laboratories that the laboratory shall periodically and in accordance with a predetermined schedule and procedure conduct

- internal audits of its activity to verify that its operations continue comply with the requirements of the quality system and the International Standard,
- review about the laboratory's quality system testing and/or calibration activities: ensure their continuing suitability effectiveness, and to introduce necessary changes or improvements.

The audit is a systematic, independent and documented process for obtaining audit evidence and evaluating it objectively to determine the extent to which the audit criteria are fulfilled (EN ISO 19011:2002).

Internal audits, sometimes called first-party audits, are conducted by the organization itself for management review and other internal purposes, and may form the basis for an organization's self-declaration of conformity. External audits include those generally termed second- and third-party audits. Second-party audits are conducted by parties having an interest in the organization, such as customers. Third-party audits are conducted by external, independent auditing organizations, such as accreditation bodies.

Further information on auditing is available in EN ISO 19011:2002.

Arrangements for implementing a program for internal audits may be based upon a check list developed by APLAC (APLAC, 2004), which is attached as Annex B-3 to these Guidelines.

B.3. SPECIFYING ANALYTICAL REQUIREMENTS

B.3.1 GENERAL

The objective of analytical investigations in chemistry is to obtain information about materials or systems concerning their specific qualitative and quantitative composition and structure (Danzer, 1992).

The objectives of analytical investigations in biology are to measure rates in activity concentrations of biological variables and to make taxonomical determinations.

Before the analyst starts an analytical investigation, the intended use of the data must be explicitly stated. That is, the minimum quality requirement the data must meet to make it useful for a given purpose should

be established for every measurement situation. Careful specification of analytical requirements and critical consideration of data quality objectives are vital when designing analytical programmes.

Environmental analytical measurements are developed for a variety of purposes, such as the determination of the fate of a component in the context of biogeochemical studies, or the determination of the environmental concentration of a component for use in environmental risk assessment.

The broad range of applications of analytical data requires different analytical strategies, and the accuracy of the data obtained must be adequate for each use. A failure to pay proper attention to this topic can endanger the validity of an analytical programme, since the analytical results obtained may be inadequately accurate and lead to false conclusions.

Based on these considerations, the following parameters should be discussed and evaluated before an investigation is carried out:

- the variable of interest,
- the type and nature of the sample,
- the concentration range of interest,
- the permissible tolerances in analytical error.

B.3.2 VARIABLE OF INTEREST

Frequently, particularly for chemical variables, a single method may be used for analysis of a variable in a wide variety of matrices. However, one has to recognize that many variables exist in different matrices in a variety of chemico-physical forms, and most analytical methods provide a different response to the various forms. Therefore, particular care must be exercised that the variable of interest is clearly defined and the experimental conditions selected allow its unambiguous measurement.

B.3.3 TYPE AND NATURE OF THE SAMPLE AND ITS ENVIRONMENT

A precise description of the type and nature of the sample is essential before the analytical method can be selected. Suitable measures and precautions can only be taken during sampling, sample storage, sample pretreatment and analysis, if sufficient knowledge about the basic properties of the sample is available. There may be other, non-analytical factors to consider, including the nature of the area under investigation.

B.3.4 CONCENTRATION RANGE OF INTEREST

It is important that samples of a definite type and nature have been characterized by the concentration range of the variable. If such information is not given, needless analytical effort may be expended or, vice versa, insufficient effort may jeopardize the validity of the analytical information gained.

B.3.5 PERMISSIBLE TOLERANCES IN ANALYTICAL ERROR

Taylor (1981) pointed out that 'the tolerance limits for the property to be measured are the first condition to be determined. These are based upon considered judgement of the end user of the data and present the best estimate of the limits within which the measured property must be known, to be useful for its intended purpose'...'Once one has determined the tolerance limits for the measured property, the permissible tolerances in measurement error may be established'.

In the whole analytical chain, there are systematic errors (biases) and random errors, as indicated by the standard deviation. The bounds representing the sum of both must be less than the tolerance limits defined for the property to be measured, if the analytical data are to be useful.

B.3.6 TECHNICAL NOTE ON THE QA OF THE DETERMINATION OF CO-FACTORS

CO-FACTORS, DEFINITION, AND USE

A co-factor is a property in an investigated sample, which may vary between different samples of the same kind, and by varying may affect the reported concentration of the determinand. Thus, the concentration of the co-factor has to be established in order to compare the determinand concentrations between the different samples (e.g., for the purpose of establishing trends in time or spatial distribution) by normalization to the co-factor.

By the definition given above, it is understood that the correct establishment of the co-factor concentration is just as vital to the final result and the conclusions as is the correct establishment of the determinand concentration. Thus, the co-factor determination has to work under the same QA system, with the same QA requirements and the same QC procedures, as any other parts of the analytical chain. It is also vital that QA information supporting the data contains information on the establishment and use of any co-factors.

CO-FACTORS IN BIOTA ANALYSIS

Dry weight

Freeze-drying or heat drying at 105 °C can be used. Dry to constant weight in both cases. By constant weight is meant a difference small enough not to significantly add to the measurement uncertainty.

Lipid content

The method by Smedes (1999), which uses non-chlorinated solvents and has been demonstrated to have high performance, is recommended. This method is a modification of the Bligh and Dyer (1959) method, and can be performed using the same equipment. The two methods have been shown to give comparable results.

Physiological factors

Age, sex, gonad maturity, length, weight, liver weight, etc., are important co-factors for species of, for example, fish. For more information, see Section D.5 of the COMBINE Manual.

CO-FACTORS IN WATER ANALYSIS

Particulate material

Determined by filtration through a filter according to the ISO 11923:1997 standard.

Organic carbon

The method recommended is described in Annex C-2 of the COMBINE Manual.

Salinity

Salinity (and temperature) may be defined as a co-factor in investigations where mixing of different water masses is studied or takes place. The same standard oceanographic equipment as described in the Technical Note on Salinity is used, and the performance requirements will also be the same.

QA INFORMATION TO SUPPORT THE DATA

When reporting data that have been normalized to a co-factor, or where the co-factor data are reported along with the results, always supply the following information:

- type of co-factor (parameter),
- analytical method for the co-factor,
- uncertainty in the co-factor determination,
- how the co-factor has been used (if it has),
- results from CRMs and intercomparison exercises (on the co-factor).

REFERENCES

Smedes, F. 1999. Determination of total lipid using non-chlorinated solvents. *The Analyst*, 124: 1711.

Bligh, E.G., and Dyer, W.J. 1959. *Canadian Journal of Biochemical Physiology*, 37: 911.

ISO. 1997. Water quality—Determination of suspended solids by filtration through glass-fibre filters. ISO 11923:1997.

B.4. VALIDATION OF ANALYTICAL METHODS

B.4.1 GENERAL

On the basis of the specifications developed in the items under Section 3, the method must now be examined to determine whether it actually can produce the degree of specificity and confidence required.

Accordingly, the objective of the validation process is to identify the performance of the analytical method and to demonstrate that the analytical system is operating in a state of statistical control.

When analytical measurements are 'in a state of statistical control', it means that all causes of errors remain the same and have been characterized statistically.

B.4.2 VALIDATION

Validation of an analytical method is the procedure that 'establishes, by laboratory studies, that the performance characteristics of the method meet the specifications related to the intended use of the analytical results' (Wilson, 1970; EURACHEM/WELAC, 1992).

Performance characteristics include:

- selectivity,
- sensitivity,
- range,
- limit of detection,
- accuracy (precision, bias).

These parameters should be clearly stated in the documented method description so that the suitability of the method for a particular application can be assessed.

In the following, a brief explanation and, where appropriate, guidance on the estimation of these parameters is given.

B.4.2.1 SELECTIVITY

Selectivity refers to the extent to which a particular component in a material can be determined without interference from the other components in the material. A method which is indisputably selective for a variable is regarded as specific.

Few analytical methods are completely specific for a particular variable. This is because both the variable and other substances contribute to the analytical signal and cannot be differentiated. The effect of this interference on the signal may be positive or negative depending upon the type of interaction between variable and interfering substances.

The applicability of the method should be investigated using various materials, ranging from pure standards to mixtures with complex matrices.

- Each substance suspected to interfere should be tested separately at a concentration approximately twice the maximum expected in the sample (use Student's t-test to evaluate).
- Knowledge of the physical and chemical mechanisms of interference operative in the particular method will often help to decide for which substances tests should be made.

Interference effects causing restrictions in the applicability of the analytical method should be documented.

B.4.2.2 SENSITIVITY

Sensitivity is the difference in variable concentration corresponding to the smallest difference in the response by the method that can be detected at a certain probability level. It can be calculated from the slope of the calibration curve.

Most analytical methods require the establishment of a calibration curve for the determination of the (unknown) variable concentration. Such a curve is obtained by plotting the instrumental response, y , versus the variable concentration, x . The relationship between y and x can be formulated by performing a linear regression analysis on the data. The analytical calibration function can be expressed by the equation $y = a + bx$, where b is the slope or response and a is the intercept on the y -axis.

As long as the calibration curve is within the linear response range of the method, the more points obtained to construct the calibration curve the better defined the b value will be. A factor especially important in defining the slope is that the measurement matrix must physically and chemically be identical both for samples to be analysed and standards used to establish the calibration curve.

B.4.2.3 LIMIT OF DETECTION, LIMIT OF QUANTIFICATION

GENERAL

Limit of detection, quantification or application are validation parameters which describe the sensitivity of an analytical methods with regard to the detection and quantification of a certain analyte. Therefore, a number of publications recently provided different approaches to define and calculate these measures by instrumental or mathematical approaches (DIN 32645, 1994; EURACHEM, 1992; Geiß and Einax, 2000; ICH, 1996; ISO 11843, 1997-2003; ISO/CD 13530, 2003; IUPAC, 1997, 2002).

DEFINITIONS

In broad terms, the limit of detection (LOD) is the smallest amount or concentration of an analyte in the test sample that can be reliably distinguished from zero (IUPAC, 2002). For analytical systems where the application range does not include or approach it, the LOD does not need to be part of a validation.

There has been much diversity in the way in which the limit of detection of an analytical system is defined. Most approaches are based on multiplication of the within-batch standard deviation of results of blanks by a certain factor. These statistical inferences depend on the assumption of normality, which is at least questionable at low concentrations (ISO/WD 13530, 2003).

Limit of quantification (LOQ) is a performance characteristic that marks the ability of an analytical method to adequately “quantify” the analyte. Sometimes that LOD is arbitrarily defined as a relative standard deviation RSD (commonly RSD = 10%), sometimes the limit is arbitrarily taken as a fixed multiple (typically 2-3) of the detection limit. This quite arbitrary setting of LOQ does not consider that measurements below

such a limit are not devoid of information content and may well be fit for purpose. Hence, it is preferable to try to express the uncertainty of measurement as a function of concentration and compare that function with a criterion of fitness for purpose agreed between the laboratory and the client or end-user of the data (IUPAC, 2002).

Lower Limit of Application (LLOA) is an agreed criterion of the fitness for purpose for the monitoring of priority hazardous substances within the EU Water Framework Directive (2000/60/EC). The LLOA shall be defined: $LLOA \geq LOQ$. The LLOA refers to the lowest concentrations for which a method has been validated with specified accuracy (AMPS, 2004). For methods which need calibration the lowest possible LLOA is equal to the lowest standard concentration (ISO/CD 13530, 2003).

The LLOA is required to be equal or lower than 30% of the defined Environmental Quality Standards (EQS). This ensures that priority hazardous substance concentrations around the proposed EQS can be measured with an acceptable measurement uncertainty of $\leq 50\%$ (AMPS, 2004).

Calculation of LOD for methods with normally distributed blank values

The LOD shall be calculated as:

$$LOD = 3 s_0$$

where

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

The precision estimate s_0 shall 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 $\alpha = 0,01$.

Note that with the recommended minimum degrees of freedom, the value of the limit of detection is quite uncertain, and may easily vary by a factor of 2. Where more rigorous estimates are required more complex calculations should be applied ISO 11843 (1997-2003).

Calculation of LOD for chromatographic methods

There are several options for the determination of LOD/LOQ for chromatographic methods:

- The LOD is defined as the concentration of the analyte at a signal/noise ratio $S/N=3$.
- Measure concentrations in a very low level sample e.g. 10 times and calculate standard deviation
- Spike analyte-free sample and measure e.g. 10 times, and then calculate standard deviation
- Dilute natural low level sample extract to achieve the required concentration. Then measure e.g. 10 times and calculate standard deviation

The proposed options are arranged according to their appropriateness.

Calculation of LOQ

The limit of quantification (LOQ) is the smallest amount or concentration of analyte in the test sample which can be determined with a fixed precision, e.g. relative standard deviation $s_{rel} = 33,3 \%$. Usually it is arbitrarily taken as a fixed multiple of the detection limit (IUPAC, 2002).

For method validation the LOQ shall be calculated as: $LOQ = 3 LOD$

The factor of 3 corresponds to a relative standard deviation $s_{rel} = 33,3 \%$.

For verification of the LOQ a spiked sample at this concentration level shall be analysed in the same manner as real samples. The analytical result must be in the range of $LOQ \pm 33,3\%$.

REFERENCES

AMPS (2004) Draft final report of the expert group on analysis and monitoring of priority substances (AMPS), EAF(7)-06/01

DIN 32645 (1994) Nachweis-, Erfassungs- und Bestimmungsgrenze

EURACHEM (1992) The Fitness for Purpose of Analytical Methods; A Laboratory Guide to Method Validation and Related Topics, EURACHEM, www.eurachem.ul.pt/guides

Geiss S and Einax JW (2001) Comparison of detection limits in environmental analysis – is it possible?, Fresenius J. Anal. Chem. 370, 673-678

ICH (1996) Validation of Analytical Procedures: Methodology, EMEA, <http://www.emea.eu.int/pdfs/vet/vich/059198en.pdf>

ISO/CD 13530 (2003) Water quality - Guide to analytical quality control for water analysis

ISO 11843 (1997-2003) Capability of detection, Part 1-4

IUPAC (1997) Compendium of Chemical Terminology, Second edition, Edited by A D McNaught and A Wilkinson

IUPAC (2002) Harmonized guidelines for single-laboratory validation of methods of analysis, Pure Appl. Chem., 74, 835-855

B.4.2.4 RANGE

The range of the method is defined by the smallest and greatest variable concentrations for which experimental tests have actually achieved the degree of accuracy required.

The concentrations of the calibration standards must bracket the expected concentration of the variable in the samples.

It is recommended to locate the lower limit of the useful range at $x_B + 10s_B$, where x_B is the measured value for the blank, and s_B is the standard deviation for this measurement.

The range extends from this lower limit to an upper value (upper limit) where the response/variable concentration relationship is no longer linear.

B.4.2.5 ACCURACY

The term 'accuracy' is used to describe the difference between the expected or true value and the actual value obtained. Generally, accuracy represents the sum of random error and systematic error or bias (Taylor, 1981).

Random errors arise from uncontrolled and unpredictable variations in the conditions of the analytical system during different analyses. Fluctuations in instrumental conditions, variations of the physical and chemical properties of sample or reagent taken on different occasions, and analyst-dependent variations in reading scales are typical sources causing random errors.

The term 'precision' should be used when speaking generally of the degree of agreement among repeated analyses. For numerical definition of this degree of agreement, the parameter standard deviation or relative standard deviation should be used.

Systematic errors or biases originate from the following sources:

a) instability of samples between sample collection and analysis

Effective sample storage, sample stabilization and sample preservation, respectively, are essential to ensure that no losses or changes of the physical and chemical properties of the variable occur prior to analysis. Effective sample stabilization methods exist for many variables and matrices, but they must be compatible with the analytical system being employed, and with the particular sample type being analysed.

b) deficiencies in the ability to determine all relevant forms of the variable

Many variables exist in different matrices in a variety of physical and/or chemical forms ('species'). The inability of the analytical system to determine some of the forms of interest will give rise to systematic negative deviations from the true value, if those forms are present in the sample.

c) biased calibration

Most instrumental methods require the use of a calibration function to convert the primary analytical signal (response) to the corresponding variable concentration. Generally, calibration means the establishment of a function by mathematically modelling the relationship between the concentrations of a variable and the corresponding experimentally measured values.

An essential prerequisite when establishing a calibration function is that the sample and calibration standards have similar matrices and are subject to the same operational steps of the analytical method, and that identical concentrations of the variable in standards and sample give the same analytical response.

d) incorrect estimation of the blank

It is common practice to correct quantitative analytical results for a constant systematic offset, denoted the 'blank'. A definite answer must be found to what the true blank in an analysis is, in order to make correction for the blank satisfactory.

A good review of several kinds of 'blank' and their use in quantitative chemical analysis was given by Cardone (1986a, 1986b).

Principally, it is important to realize that a 'blank' is the response from a solution containing all constituents of the sample, except the variable, processed through all procedural steps of the method under study. The analyst must know that the size of the blank and its influence on the analytical result can only be assessed if the sample matrix has been adequately approximated and the whole analytical process has been considered.

B.4.2.5.1 ESTIMATING RANDOM ERRORS

The within-batch standard deviation, s_w , represents the best precision achievable with the given experimental conditions, and is of interest when the analyst is concerned with the smallest concentration difference detectable between two samples.

The between-batch standard deviation, s_b , is a measure of the mutual approximation of analytical results obtained from sequentially performed investigations of the same material in the same laboratory.

The total standard deviation, s_t , is calculated from the formula $s_w^2 + s_b^2$. It is of interest to analysts concerned with the regular analysis of samples of a particular type in order to detect changes in concentration.

A realistic approach to estimate s_w and s_b is to perform n determinations on a representative group of control samples in each of m consecutive batches of analysis.

The experimental design recommended to estimate s_w , s_b and s_t is to make n replicate analyses per batch in a series of m different batches. The design should be modified according to practical experience gained from the analytical method tested. In particular, when s_w is assumed to be dominant, $n=4$ to 6 could be chosen. The product $n \cdot m$ should not be less than 10 and should preferably be 20 or more.

Analysis of variance (ANOVA) allows identification of the different sources of variation and calculation of the total standard deviation s_t . A general scheme of ANOVA (after Doerffel, 1989) is given in the following paragraphs.

Source of variability	Sum of squareyys	Degrees of freedom	Mean squares (variances)	Variance components
Between batches	$QS_1 = n_j(x_j - \bar{x})^2$	$f_1 = m - 1$	$s_{bm}^2 = QS_1 / m - 1$	$s_{bm}^2 = n_j s_b^2 + s_w^2$
Within batches (analytical error)	$QS_2 = (x_{ij} - \bar{x}_j)^2$	$f_2 = m(n_j - 1)$	$s_w^2 = QS_2 / m(n_j - 1)$	
Total	$QS_1 + QS_2$	$f = mn_j - 1$		

m = number of batches of analysis;

n_j = number of replicate analyses within a batch;

x_j = mean of j th batch;

\bar{x} = overall mean;

x_{ij} = j th replicate analytical value in i th batch

s_w^2 = estimate of within batch variance

s_{bm}^2 = estimate of the variance of the batch means

$F = s_{bm}^2/s_w^2$ is tested against the tabled value $F(P = 0.05; f_1; f_2)$.

If the test is significant, i.e., $F > F(P=0.05; f_1; f_2)$, the between batch variance s_b^2 can be estimated as

$$s_b^2 = (s_{bm}^2 - s_w^2)/n_j$$

Carry out F-test to see if s_b is significantly larger than s_w .

If the testing value $s_b^2/s_w^2 < F(f_b, f_w, 95\%)$, one can conclude that s_b is only randomly larger than s_w . In this case $s_t = s_w$.

If the testing value $s_b^2/s_w^2 > F(f_b, f_w, 95\%)$, one can conclude that s_b significantly influences the total standard deviation.

Accordingly, the estimate of the total variance of a single determination is $s_t^2 = s_b^2 + s_w^2$.

For routine analysis, it is recommended that s_b does not exceed the value of s_w by more than a factor of two.

A step-wise approach to scrutinize experimental design and to optimize analytical performance may be necessary. This process might be repeated iteratively until target values of s_w , s_b and s_t , respectively, are attained.

B.4.2.5.2 ESTIMATING SYSTEMATIC ERRORS (BIASES)

A) USING AN INDEPENDENT ANALYTICAL METHOD

The analyst can test for systematic errors in the analytical procedure under investigation by using a second, independent analytical method (Stoepler, 1991). A t-test can be carried out to check for differences in the measured values obtained (on condition that the precision of both methods applied is comparable). A significant difference between the results obtained by both procedures indicates that one of them contains a systematic error. Without further information, however, it is not possible to say which one.

B) USING CERTIFIED REFERENCE MATERIAL (CRM)

An analytical procedure should be capable of producing results for a certified reference material (CRM) that do not differ from the certified value more than can be accounted for by within-laboratory statistical fluctuations.

In practice, when performing tests on CRM, one should ensure that the material to be analysed and the certified reference material selected have a similar macrocomposition (a similar matrix) and approximately similar variable concentrations.

C) PARTICIPATION IN INTERCOMPARISON EXERCISES

In an intercomparison exercise, the bias of the participating laboratory's analytical method is estimated with respect to the assigned value X for the concentration of the variable in the sample which was distributed to participants. The assigned value X is an estimate of the true value and is predetermined by some 'expert' laboratories. In some instances, X is a consensus value established by the coordinator after critical evaluation of the results returned by the participants. The bias is equal to the difference between the variable concentration x reported by the participant and the variable concentration X assigned by the coordinator.

If a target standard deviation s representing the maximum allowed variation consistent with valid data can be estimated, the quotient $z = (x - X)/s$ is a valuable tool for appropriate data interpretation. If z exceeds the value of 2, there is only a 5 percent probability that the participating laboratory can produce accurate data (Berman, 1992).

B.4.2.6. MEASUREMENT UNCERTAINTY

INTRODUCTION

The principal aim of analytical work is to gain information on the material under investigation. This information always constitutes a probability distribution determined by a random error and a systematic error inherent in the analytical procedure used. A systematic error can act as an additive or as a multiplicative shift. Systematic errors are superimposed by the random error. Analytical practice shows that there is always some doubt about the correctness of a stated result, even when all the suspected sources of error have been taken into account and the appropriate corrections have been applied. This is due to the uncertainty regarding the correction factors and the uncertainty arising from random effects, which cannot be eliminated, although they can be reduced by increasing the number of observations. Consequently, a measurement cannot be properly interpreted without the knowledge of the uncertainty associated with the result.

The concept of expressing or estimating the uncertainty of measurements was developed to inform the final users of the analytical data concerning how much allowance must be made for the possibility that repetition of the test will give a different value (Horwitz, 1998). This information is particularly necessary when analytical results are not used by the data originator, as is the regular case in the assessment of data from environmental monitoring program. This technical note provides information on how the uncertainty of measurement of the analytical methods used in the COMBINE program of HELCOM can be estimated, so that it would be possible to judge whether or not the accuracy (trueness and precision) of the method meets the requirements of this program. It should be taken into account that the requirements on accuracy depend on the aims and the purpose of the monitoring program.

DEFINITIONS

The “Guide to the expression of uncertainty in measurement” (JCGM 100: 2008) defines Measurement Uncertainty as a parameter, associated with the result of a measurement, which characterises the dispersion of the values that could reasonably be attributed to the measurand.

The following definitions apply (EURACHEM/CITAC, 2000):

Standard uncertainty

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

Combined standard uncertainty

$u_c(y)$ standard uncertainty of the result y of a measurement when the result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighed according to how the measurement results varies with these quantities

Expanded standard uncertainty

U quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand, an expanded uncertainty is calculated from a combined standard uncertainty u_c and a coverage factor k using $U = k \cdot u_c$

Coverage factor

k numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty, the choice of the factor k is based on the level of confidence desired, $k = 2$ for an approximate level of confidence of 95 %

PROCEDURES TO ESTIMATE UNCERTAINTY

Generally, there are two main approaches to estimate the uncertainty of analytical measurements or an analytical procedure, respectively.

-“bottom-up” approach

Using the error budget model (JCGM 100: 2008), the combined standard uncertainty can be calculated as the square root of the sum of squares of all individual error components in the form of standard deviations. This “bottom-up” approach assumes that an analytical method can be structured into small, simple steps, and that an individual standard uncertainty can be attributed to all of these steps, sometimes based on a best guess of experienced analysts.

NOTE:

The “bottom-up” approach to estimate the uncertainty of analytical measurements seems to be rather impractical (Horwitz, 1998), because it does not use the results from analytical quality control, e.g. control

charts, analysis of reference materials or participation in proficiency testing schemes. Therefore, if information on the uncertainty of analytical data generated in the COMBINE programme is needed, the "top-down" approach according to ISO 11352 should be preferred.

- "top-down" approach

A „top-down“ view on estimating the measurement uncertainty is described in ISO 11352 (2012-07) based on the NORDTEST Report TR 537 (2004). According to this standard, the combined standard uncertainty u_c is characterised as the within-laboratory reproducibility u_{RW} (e.g. from control charts) combined with the method and laboratory bias u_{bias} (e.g. from analysis of suitable reference materials, results from interlaboratory comparisons or recovery experiments):

$$u_c = \sqrt{u_{RW}^2 + u_{bias}^2}$$

NOTE:

Annex B of the ISO 11352 (2012) "Water quality - Estimation of measurement uncertainty based on validation and quality control data" gives detailed examples for the calculation of the measurement uncertainty from validation data (e.g. using reference material, using data from proficiency tests and using a standard solution as quality control sample).

UNCERTAINTY IN DEPENDENCE ON ANALYTE LEVELS

For results near the limit of quantification, the uncertainty is often found to be constant and can therefore be expressed as an absolute value. When results are well above the limit of quantification the uncertainty is often proportional to the analyte concentration and can therefore be expressed as a relative value.

To allow for both proportionality of uncertainty and the possibility of an essentially constant value with level, the following general expression is used (EURACHEM/CITAC, 2000):

$$u(x) = \sqrt{s_0^2 + (x \cdot s_1)^2}$$

where

s_0 represents a constant contribution to the overall uncertainty and

s_1 is a proportionality constant.

When the result is far from zero (i.e. well above the limit of quantification) and there is clear evidence that the uncertainty changes proportionally with the level of analyte, the term $x \cdot s_1$ dominates. Under these

circumstances s_0 may reasonably be recorded as zero and s_1 is simply the uncertainty expressed as a relative standard deviation.

REPORTING OF UNCERTAINTY

The expanded uncertainty $U = k \cdot u_c$ (usually with $k = 2$) should be reported for individual monitoring parameters as percent relative uncertainty or as absolute uncertainty (in the form of a standard deviation) together with information on how it was determined.

B.4.3 VALIDATION OF DATA (NEW CHAPTER)

Data validation is defined as the inspection of all the collected data for completeness and reasonableness, and the elimination of erroneous values. To validate, correct and evaluate data a broad range of different tools are provided. To each time series a set of so-called plausibility checks can be assigned to. These plausibility checks can be defined for a particular time range or season.

This step of data validation transforms raw data into validated data. The validated data are then processed to produce the summary reports you require for data assessment and reporting. In principle, it is necessary that persons who are involved in data validation have enough experiences and knowledge of measurements, expected results and environmental conditions.

There are essentially two parts to data validation, data screening and the treatment of suspect and missing data.

B.4.3.1 DATA SCREENING

The first part uses a series of validation routines or algorithms to screen all the data for suspect (questionable and erroneous) values. A suspect value deserves scrutiny but is not necessarily erroneous. The result of this part could be a data validation report that lists the suspect values and which validation routine each value failed.

B.4.3.2 TREATMENT OF SUSPECT AND MISSING DATA

The second part requires a case-by-case decision on what to do with the suspect values, retain them as valid, reject them as invalid, or replace them with redundant, valid values (if available). This part is where judgment by a qualified person familiar with the monitoring equipment and local conditions is needed.

Before proceeding to the following sections, you should first understand the limitations of data validation. There are many possible causes of erroneous data. The goal of data validation is to detect as many significant errors from as many causes as possible. Catching all the subtle ones is impossible. Therefore, slight deviations in the data can escape detection. Properly exercising the other quality assurance components of the monitoring program will also reduce the chances of data problems.

A. Data screening

To screen data the following list shows a selection of supported plausibility checks:

1. Check the completeness of the collected data. Check if there are any missing data values. Check if the number of data fields is equal to the expected number of measured parameters.
2. Check the correctness of the defined data format, expected interval range and measurement units.
3. Range tests. The measured data are compared to allowable upper and lower limiting values.
4. Relational tests. This comparison is based on expected relationship between various 30 ICES STGQAC Report 2006 parameters.
5. Trend tests. These tests are based on the rate of change in a value over time.

B. Treatment of suspect and missing data

After the raw data are subjected to all the validation checks, what should be done with suspect data? Some suspect values may be real, unusual occurrences while others may be truly bad. Here are some guidelines for handling suspect data:

1. Generate a validation report that lists all suspect data. For each data value, the report should give the reported value, the date and time of occurrence, and the validation criteria that it failed.
2. A qualified person should examine the suspect data to determine their acceptability.
3. If there are suspect values go back to the raw data in the laboratory and check all analytical steps and quality assurance tools for relevant investigations.
4. Compare suspect values with earlier data from the database or with other information
5. Repeat the analysis if it is possible.

References

AWS Scientific, Inc. (1997) Data Validation, Processing, and Reporting. In: Wind Resource Assessment Handbook. www.awsscientific.com

B.5. ROUTINE QUALITY CONTROL (USE OF CONTROL CHARTS)

B.5.1 GENERAL

According to international standard, e.g. ISO 17025, a defined analytical quality must be achieved, maintained, and proven by documentation. The establishment of a system of control charts is a basic principle applied in this context. For further information for control charts refer to ISO/TR 13530 (1997).

B.5.2 CONTROL OF TRUENESS

As a routine procedure for controlling systematic error, the use of Shewhart control charts based on the mean, spiking recovery and analysis of blanks is recommended.

B.5.2.1 X-CHARTS

Synonyms for X-chart are X-control chart, mean control chart, average control chart or xbar control chart.

For trueness control, standard solutions, synthetic samples or certified real samples may be analysed using a Shewhart chart of 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 determinant is available.

A simple X-chart is constructed in the following way:

- The respective control sample should be analysed later on a regular basis with each batch of unknown environmental samples or, if a large number of unknowns is run in a batch, one control sample for each 10 or 20 unknowns.
- Analyse the control sample at least ten times for the given variable. The analyses should be done on different days spread over a period of time. This enables a calculation of the total standard deviation (s_t).
- It is advisable to analyse certified reference samples (if suitable ones are available and are not too expensive) with routine samples as a check on trueness. A restricted check on systematic error by means of recovery control charts is often made instead (5.2.2).
- Calculate the mean value (\bar{x}), the standard deviation (s) and the following values: $x + 2s$, $x - 2s$, $x + 3s$, $x - 3s$. Use these data to produce the plot.

If the data follow a normal distribution, 95 % of them should fall within $x \pm 2s$ (between the Upper Warning Limit and Lower Warning Limit) and 99.7 % should fall within $x \pm 3s$ (between the Upper Action Limit and Lower Action Limit).

B.5.2.2 BLANK CONTROL CHART

The blank control chart represents a special application of the X-chart (mean control chart). The following (constant) systematic error sources may be identified by the blank control chart:

- contamination of container for sampling, sample storage and sample pre-treatment;
- contamination of reagents, reaction vessels or laboratory equipment used during analysis.

Generally, the simultaneous determination of the blank value would be required for each analysis. Since this requirement can seldom be met due to the considerable effort, it appears reasonable to determine a

minimum of two blank values during the series of analyses (at the beginning and at the end of each batch of samples).

B.5.2.3 RECOVERY CONTROL CHART

Synonyms of recovery control charts are Control Charts for Spiked Sample Recovery, Spiked-sample (control) chart or Accuracy charts.

In marine chemistry the control charts for spiked sample recovery are especially useful when the sample matrix can be suspected of causing interferences that have an influence on the analytical response. They are useful in trace metal analysis and in nutrient analysis where the sample matrix can affect the chemical reaction of the signal response.

The control chart for spiked sample recovery can be constructed as follows:

- Use the same spike concentration in all series of the same variable, concentration range and matrix.
- Select and analyse a natural sample in each analytical series.
- Spike by adding to the sample a known concentration of the analyte to be determined, and re-analyse. If possible, use a CRM concentrate.
- Calculate the measured difference in concentration by subtraction and correction for dilution from spiking.
- Calculate the percent recovery (%R) for each spiked sample for a given test, matrix, range.
- Calculate the mean %R by taking the %R's and dividing by the total number (n) of %R's (outlier excluded).
- Calculate the total standard deviation (s_t) on the basis of at least ten analytical series.
- Calculate the following values: $R + 2s_t$, $R - 2s_t$, $R + 3s_t$, $R - 3s_t$. Use these data to produce the plot.

With the presumption that the measured recoveries are normally distributed, the data should be distributed within the same limits as described for the X-charts (see B.5.2.1).

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 determinant concentration; bias of constant size may go undetected.

B.5.3 CONTROL OF PRECISION

There are four ways of controlling the precision of analytical results in routine analysis:

- use of the mean control chart (5.2.1);
- use of a range control chart (5.3.1);

- estimation of precision with replicate analysis (5.3.2);
- standard addition (5.3.3).

B.5.3.1 R-CHART

Synonyms for R-chart are Range (control) chart or Precision chart.

R-charts are used for graphing the range or the relative percent difference (RPD) of analytical replicate or matrix spike duplicate results.

It is common practice in analytical laboratories to run duplicate analyses at frequent intervals as a means of monitoring the precision of analyses and detecting out-of-control situations in R-charts. This is often done for determinants for which there are no suitable control samples or reference materials available.

The R-chart can be constructed using the following method:

- Calculate the relative range (R_{rel}) for each replicate analysis for a given natural sample of the same matrix:

$$R_{rel} = \frac{x_{max} - x_{min}}{\bar{x}} * 100\%$$

with the mean value (\bar{x}) of the replicate set $\bar{x} = \frac{1}{n} * \sum_{i=1}^n x_i$

- Calculate the mean relative range $\overline{R_{rel}}$ by summing all R_{rel} and dividing by the total number (n) of replicate sets (outliers excluded).

- Calculate the Upper Action Limit (UA): $UA = \overline{R_{rel}} * D_{UA}$

- The R_{rel} for each replicate analysis and the UA are drawn on the chart.

When performing replicate determinations (duplicate to six-fold), the lower action limit (LA) is identical with the abscissa (zero-line).

The numerical values for the factor D_{UA} 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

NOTE: For further numerical values for the factor UA refer to Funk et al. (1992).

B.5.4. CONTROL CHARTS WITH FIXED QUALITY CRITERIONS (TARGET CONTROL CHARTS)

In the contrary to the under clauses B.5.2 and B.5.3 described classical control charts of the SHEWHART type the target control charts operate without statistically evaluated values. The bounds for this type of

control charts are given by external prescribed and independent quality criterions. 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)
- the Shewhart or range control charts show persisting out of control situations
- there are not enough data available for the statistical evaluation of the bounds
- there are external prescribed bounds which should be applied to ensure the quality of analytical values.

The control samples for the target control charts are the same as for the classical control charts as described in clauses B.5.2 and B.5.3.

The bounds are given by:

- 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-intern known data of the same sample type.

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.

The analytical method is out-of-control if the analytical value is higher or lower than the respective prescribed bounds. The measures are the same as described in clauses 5.2 and 5.3.

B.5.5. CONTROL CHARTS FOR BIOLOGICAL MEASUREMENTS

For the quality control while measuring biological variables the Shewhart charts (in these cases R-charts where the criteria for evaluation of testing results is based on statistically calculated values are used.

The control chart for duplicate samples can be constructed as follows:

- For bacterioplankton, phytoplankton and mesozooplankton, run every 10th sample or at least one sample per batch as duplicate, counting two sub samples from the same sample (ca 10% of all samples).
- For chlorophyll-a run one duplicate sample within every batch of samples and calculate the range (R) with $R = |x_1 - x_2|$,

where x_1 and x_2 are concentrations of chlorophyll-a in duplicate samples. For other biological variables the difference in abundance of organisms and/or biomass is calculated.

- Calculate the relative range (R_{rel}) for each duplicate analysis

$$R_{rel} = \frac{x_{max} - x_{min}}{\bar{x}} * 100\%$$

with the mean value (\bar{x}) $\bar{x} = \frac{1}{n} * \sum_{i=1}^n x_i$ of the replicate set

- Calculate the mean relative range $\overline{R_{rel}}$ by summing all R_{rel} and dividing by the total number (n) of replicate sets (outliers excluded).
- Calculate the Upper Action Limit (UA): $UA = \overline{R_{rel}} * D_{UA}$ in the same way as in chapter 5.3.1
- For duplicate determination the Upper Action Limit (UA) is then: $UA = 3,267 R_{rel}$
- The lower action limit (LA) is identical with the abscissa (zero-line).
- The R_{rel} for each replicate analysis and the UA are drawn on the chart.

The blank control chart can also be used for biological measurements, e.g. for the determination of chlorophyll a .

B.5.6. INTERPRETATION OF CONTROL CHARTS, OUT-OF-CONTROL SITUATIONS

The results of analyses of reference material analysed with each batch of environmental samples indicate whether the errors fall within acceptable limits.

The quality control charts 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 UA or LA; or
- 2 consecutive values outside warning limit UW or LW; 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 R-chart if:

- a range RPD falls outside the upper action limit; or
- a range RPD 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 RPD.

For control charts with fixed quality criteria (target control charts) the analytical method is out-of-control if the analytical value is higher or lower than the respective prescribed bounds.

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.

B.5.7. SELECTION OF SUITABLE CONTROL CHARTS

The table shows which control samples are suitable for checking trueness and precision.

Type of control chart	Trueness	Precision
X-chart with standard solution	Restricted	Yes
X-chart with Certified reference materials	Yes	Yes
X-chart with Laboratory reference materials/intercomparison samples	Yes	Yes
Blank control chart	Restricted	No
Recovery control chart with real sample	Yes	Yes
R-chart	No	Yes

Control charts for the determination of sum parameters:

- R-chart as Shewhart or as target control chart for the whole range of the method whereas it should be worked with various matrices and concentration levels
- Blank control charts as Shewhart chart for an (approximated) normal distribution or otherwise as a target control chart

Control charts for the determination of single parameters:

- X-chart as Shewhart or as target control chart for the whole range of the method whereas the control sample concentration should be in the middle of the range
- Blank control charts as Shewhart chart for an (approximated) normal distribution or otherwise as a target control chart

Control charts for multi-parameter methods:

- The selection of suitable control charts for multi-parameter procedures such as gas chromatography and optical emission spectrometry should depend on whether the selected measured variable is particularly problematic, representative or relevant.

B.5.8. REFERENCES

ASTM Manual 7a. 2002. Manual on Presentation of Data and Control Chart Analysis. Seventh Edition

Funk, W., Donnevert, G. and Dammann V. 1992. Quality Assurance in Analytical Chemistry. Verlag Chemie, Weinheim

Hovind, H. 2002. Internal Quality Control - Handbook for Chemical Analytical Laboratories. Norwegian Institute for Water Research. <http://www.niva.no>

ISO/TR 13530. 1997. Water quality – Guide to analytical quality control for water analysis.

B.6. EXTERNAL QUALITY ASSESSMENT

For marine environmental monitoring programmes, it is essential that the data provided by the laboratories involved are comparable. Therefore, activities like participation in an external quality assessment schemes, ring tests, taxonomical workshops and use of external specialists by the laboratories concerned should be considered indispensable.

While the use of a validated analytical method and routine quality control (see above) will ensure accurate results within a laboratory, participation in an external quality assessment or proficiency testing scheme provides an independent and continuous means of detecting and guarding against undiscovered sources of errors and acting as a demonstration that the analytical quality control of the laboratory is effective.

Generally, proficiency testing, ring tests, etc. are useful to obtain information about the comparability of results, and ensures that each of the participating laboratories achieves an acceptable level of analytical accuracy.

Details of the development and operation of proficiency testing schemes are outlined in ISO Guide 43. An overview of the structure and an assessment of the objectives of proficiency testing have been given by the Analytical Methods Committee (1992).

An approach known as the paired sample technique, which has been described by Youden and Steiner (1975), provides a valuable means of summarizing and interpreting in graphical form the results of interlaboratory comparison exercises.

Most ring tests and proficiency testing schemes are based on the distribution of samples or identical sub-samples (test materials) from a uniform bulk material to the participating laboratories. The test material

must be homogeneous and stable for the duration of the testing period. Amounts of the material should be submitted that are sufficient for the respective determinations.

The samples are analysed by the different laboratories independently of one another, each under repeatable conditions. Participants are free to select the validated method of their choice. It is important that the test material is not treated in any way different from the treatment of samples ordinarily analysed in the laboratory. In this way, the performance established by the proficiency testing results will reflect the actual performance of the laboratory.

Analytical results obtained in the respective laboratories are returned to the organizer where the data are collated, analysed statistically, and reports issued to the participants.

B.7. DEFINITIONS

In the following, a summary of the technical/scientific terms used in this document is given. Sections are mentioned when the terms have been explained in the text. Definitions are provided for terms not explained in the text.

Accuracy. See Section [4.2.5](#).

Analytical method. The set of written instructions completely defining the procedure to be adopted by the analyst in order to obtain the required analytical result (Wilson, 1970).

An **analytical system** comprises all components involved in producing results from the analysis of samples, i.e., the sampling technique, the 'method', the analyst, the laboratory facilities, the instrumental equipment, the nature (matrix, origin) of the sample, and the calibration procedure used.

Biological variables are chlorophyll *a*, primary production measurements, bacterioplankton, phytoplankton, zooplankton, macrozoobenthos, phytobenthos and fish.

HELCOM BMP. Baltic Monitoring Programme.

Blank control chart. See Section [5.5](#).

Calibration is the set of operations which establish, under specified conditions, the relationship between values indicated by a measuring instrument or measuring system, or values represented by a material measure, and the corresponding known values.

HELCOM CMP. Coastal Monitoring Programme.

CRM (Certified Reference Material) is a material one or more of whose property values are certified by a technically valid procedure, accompanied by or traceable to a certificate or other documentation which is issued by a certifying body.

Cusum Charts. See Section [5.4](#).

Detection limit. See Section [4.2.3](#).

Errors. See Sections [4.2.5](#), [4.2.5.1](#) and [4.2.5.2](#).

External quality assessment. See [Section 6](#).

LCL. Lower control limit.

LRM. Laboratory Reference Material.

Matrix. The totality of all components of a material including their chemical, physical and biological properties.

Performance characteristics of an analytical method used under given experimental conditions are a set of quantitative and experimentally determined values for parameters of fundamental importance in assessing the suitability of the method for any given purpose (Wilson, 1970).

Proficiency testing is the determination of the laboratory calibration or testing performance by means of interlaboratory comparisons. **Quality.** Characteristic features and properties of an analytical method/analytical system in relation to their suitability to fulfill specific requirements.

The term **Quality Assurance** involves two concepts: **Quality control** and **Quality assessment**.

- **Quality control** is 'the mechanism established to control errors', and **quality assessment** is 'the system used to verify that the analytical process is operating within acceptable limits' (ACS Committee, 1983; Taylor, 1981).
- **Quality assessments** of analyses, generally referred to as intercomparison exercises, have been organized over the last twenty years by the International Council for the Exploration of the Sea (ICES).
- **Quality audits** are carried out in order to ensure that the laboratory's policies and procedures, as formulated in the Quality Manual, are being followed.
- **Quality Manual** is a document stating the quality policy and describing the quality system of an organization.
- **Quality policy** forms one element of the corporate policy and is authorized by top management.
- **Quality system** is a term used to describe measures which ensure that a laboratory fulfills the requirements for its analytical tasks on a continuing basis.

Range. See Section [4.2.4](#), Ring test - See proficiency testing

Selectivity. See Section [4.2.1](#).

Quality Manager. The Quality System should include a statement that the post-holder has responsibility for ensuring that the requirements for the Quality System are met continuously and that the post-holder has direct access to the highest level of management at which decisions are taken on laboratory policy or resources, and to the Technical Manager.

Technical Manager. The Quality System should include a statement that the post-holder has overall responsibility for the technical operation of the laboratory and for ensuring that the Quality System requirements are met.

Traceability. Results obtained from an analytical investigation can only be accurate if they are traceable. Traceability of a measurement is achieved by an unbroken chain of calibrations connecting the measurement process to the fundamental units. In most instances, when analyses are carried out, the chain is broken because due to the sample pretreatment and preparation the original material is destroyed. In order to approach full traceability, it is necessary to demonstrate that no loss or contamination has occurred during the analytical procedure.

Traceability to national or international standards can be achieved by comparison with certified reference standards or certified reference materials, respectively, the composition of which must simulate to a high degree the sample to be analysed. Consequently, if analytical results for a certified reference material are in agreement with the certified values, it should be realized that owing to discrepancies in composition between certified reference material and sample, there is still a risk that the results on real samples may be wrong.

UCL. Upper control limit.

Validation of an analytical method is the procedure that 'establishes, by laboratory studies, that the performance characteristics of the method meet the specifications related to the intended use of the analytical results' (EURACHEM/WELAC, 1992).

X-charts. See Section [5.2.1](#).

References

ACS (American Chemical Society Committee on Environmental Improvement. Keith, L., Crummet, W., Deegan, J., Libby, R., Taylor, J., Wentler, G.), 1983. Principles of Environmental Analysis. Analytical Chemistry, 55: 2210-2218.

Analytical Methods Committee, 1987. Recommendations for the detection, estimation and use of the detection limit. Analyst, 112:199-204.

Analytical Methods Committee, 1992. Proficiency testing of analytical laboratories: Organization and statistical assessment. Analyst, 117: 97-104.

Berman, S.S., 1992. Comments on the evaluation of intercomparison study results. ICES Marine Chemistry Working Group, Doc. 7.1.2.

Cardone, M., 1986a. New technique in chemical assay calculations. 1. A survey of calculational practices on a model problem. Analytical Chemistry, 58: 433-438.

Cardone, M., 1986b. New technique in chemical assay calculations. 2. Correct solution of the model problem and related concepts. Analytical Chemistry, 58: 438-445.

- CEN/TC 230, 1995. Guide to Analytical Quality Control for Water Analysis. Document CEN/TC 230 N 223, Document CEN/TC 230/WG 1 N 120. Final Draft, April 1995. 80 pp.
- Danzer, K., 1992. Analytical Chemistry -- Today's definition and interpretation. *Fresenius Journal of Analytical Chemistry*, 343: 827-828.
- Doerffel, K., 1989. *Statystyka dla chemików analityków (Statistics in Analytical Chemistry)*. WNT, Warszawa, p. 101.
- EN 14996, 2006. Water quality - Guidance on assuring the quality of biological and ecological assessments in the aquatic environment
- EURACHEM/WELAC (Cooperation for Analytical Chemistry in Europe/Western European Legal Metrology Cooperation), 1992. Information Sheet No. 1 (Draft): Guidance on the Interpretation of the EN 45000 series of Standards and ISO Guide 25. 27 pp.
- Gorsuch, T., 1970. *The destruction of organic matter*. Pergamon Press, Oxford.
- Grasshoff, K., 1976. *Methods of seawater analysis*. Verlag Chemie, Weinheim, New York.
- HELCOM, 1991. Third Biological Intercalibration Workshop, 27-31 August 1990, Visby, Sweden. *Balt. Sea Environ. Proc.* No. 38.
- HELCOM, 1994. Environment Committee (EC). Report of the 5th Meeting, Nyköping, Sweden, 10-14 October 1994. EC5/17, Annex 9.
- HELCOM, 1995. Helsinki Commission (HELCOM). Report of the 16th Meeting, Helsinki, Finland, 14-17 March 1995. HELCOM 16/17, Annex 7.
- ICES, 1997. ICES/HELCOM Workshop on Quality Assurance of Pelagic Biological Measurements in the Baltic Sea. ICES CM 1997/E:5.
- ISO IEC GUIDE 43, 1997. *Proficiency Testing by Interlaboratory Comparisons - Development and Operation of Proficiency Testing Schemes*.
- ISO/TR 13530, 1997. *Water quality- Guide to analytical quality control for water samples*.
- IUPAC, 1978. Nomenclature, symbols, units and their usage in spectrochemical analysis - II. *Spectrochimica Acta*, Part B, 33: 242.
- Long, G.L. and Winefordner, J.D., 1983. Limit of detection. A closer look at the IUPAC definition. *Analytical Chemistry*, 55: 712A-724A.
- Stoeppler, M., 1991. Analytical chemistry of metals and metal compounds. *In Metals and Their Compounds in the Environment*. Ed. by E. Merian. VCH Weinheim. pp. 105-206.
- Taylor, J.K., 1981. Quality assurance of chemical measurements. *Analytical Chemistry*, 53: 1588A-1596A.
- Topping, G., 1992. The role and applicability of quality assurance in marine environmental protection. *Marine Pollution Bulletin*, 25: 61-66.

Wilson, A.L., 1970. The performance of analytical methods. *Talanta*, 17: 21-44 (Part I and II).

Youden, W.J. and Steiner, E.H., 1975. Statistical Manual of the AOAC. *Journal of the Association of Official Analytical Chemists*. Washington, D.C.