Interlaboratory analytical performance studies; a way to estimate measurement uncertainty[*]

ELŻBIETA ŁYSIAK-PASTUSZAK

Institute of Meteorology and Water Management,
Maritime Branch,
Waszyngtona 42, PL–81–342 Gdynia, Poland;

e-mail: Elzbieta.Lysiak-Pastuszak@imgw.pl

Abstract

Comparability of data collected within collaborative programmes became the key challenge of analytical chemistry in the 1990s, including monitoring of the marine environment. To obtain relevant and reliable data, the analytical process has to proceed under a well-established Quality Assurance (QA) system with external analytical proficiency tests as an inherent component. A programme called Quality Assurance in Marine Monitoring in Europe (QUASIMEME) was established in 1993 and evolved over the years as the major provider of QA proficiency tests for nutrients, trace metals and chlorinated organic compounds in marine environment studies. The article presents an evaluation of results obtained in QUASIMEME Laboratory Performance Studies by the monitoring laboratory of the Institute of Meteorology and Water Management (Gdynia, Poland) in exercises on nutrient determination in seawater. The measurement uncertainty estimated from routine internal quality control measurements and from results of analytical performance exercises is also presented in the paper.

The complete text of the paper is available at http://www.iopan.gda.pl/oceanologia/

1. Introduction

'Quality' became the buzzword of the 1990s. This was nothing new to those who had been making measurements as part of their routine daily tasks. However, with the development of large international monitoring programmes over the years, it became clear that the measurements performed on the same sample by one laboratory often bore no resemblance to the values obtained by another laboratory, and concern was expressed over the unsatisfactory degree of interlaboratory comparability of data (Nehring 1994). Objective, reliable data on contaminants in water, sediments and biota are needed for their environmental significance, and for the ensuing political measures to improve the quality of the natural environment in general and the marine environment in particular (Topping 1992). Therefore the acquisition of relevant and reliable data is an essential component of all research and monitoring programmes associated with marine environmental protection. To obtain such data, the whole analytical process must proceed under a well established Quality Assurance (QA) programme.

Problems relating to the unreliability of data produced for the International Monitoring Programmes of the Joint Monitoring Group (JMG) of the Oslo and Paris Commissions (OSPARCOM) and the North Sea Task Force (NSTF) led scientists to approach the European Commission for funding for a 'Quality Programme' to address this problem. In direct response to these requirements a programme called Quality Assurance in Marine Monitoring in Europe (QUASIMEME) was born in late 1992 (QUASIMEME 1993). Over the years QUASIMEME has evolved as the major provider of QA proficiency testing schemes for mandatory determinands: nutrients, trace metals and chlorinated organic compounds.

2. Structure of QUASIMEME Laboratory Performance Studies (LPS)

The co-ordination centre of the programme is the QUASIMEME Project Office at FRS Marine Laboratory, Aberdeen. The QUASIMEME LPS operates under the IUPAC/ISO/AOAC international protocol for Laboratory testing and is organised according to ISO Guide 43 (QUASIMEME 2003a, b). The determinands included in the QUASIMEME LPS are those requested by the national and international marine monitoring programmes. It has to be stressed here that the Helsinki Commission 'urged the Contracting Parties to participate in Quality Assurance exercises whenever possible and decided that the participation in QA-exercises is mandatory for the laboratories providing BMP (Baltic Monitoring Programme) data

on parameters included in the mandatory part of the Baltic Monitoring Programme' (Helsinki Commission 1993, Łysiak-Pastuszak 1996).

The current membership of the QUASIMEME LPS is by payment of a fee for each group of test materials required. The annual subscription covers two exercises per year. Each exercise contains a number of samples, e.g. the group for nutrient determination in estuarine water includes 3 spiked samples of estuarine water (salinity 8–12 PSU; UNESCO 1981) and the fourth sample is a natural seawater with typical Baltic properties, and for heavy metal analyses, each exercise contains 2 samples of test material.

The marine chemistry laboratory of the Institute of Meteorology and Water Management (IMWM; Gdynia, Poland) is the institution responsible for the monitoring of Polish marine areas of the southern Baltic Sea within the HELCOM COMBINE programme, hence its participation in QUASIMEME LPS is mandatory (Łysiak-Pastuszak 1996, Danowska et al. 2000). The history of IMWM participation goes back to 1994; here the results of nutrient analyses in seawater are presented and evaluated.

The analyses of test materials for nutrients include the determination of the sum of nitrate and nitrite ($NO_3 + NO_2 = TOxN$), nitrite, ammonia, phosphate, silicate, total nitrogen and total phosphorus. The colorimetric methods used for nutrient determination are based on Grasshoff et al. (1983). The results of the analyses are returned to the QUASIMEME co-ordination centre for statistical evaluation.

3. Test results

On receipt of the data collector – copies of the official QUASIMEME diskettes sent to laboratories for data submission – the data are inspected for inconsistencies, determinations reported as less than the limit of detection, and values reported as zero. The following data assessment conforms with the international standards given in the ISO/IEC guide (ISO 1996) to allow comparison of data between exercises. Each determinand for each test material is given an assigned value.

Robust statistics (Cofino & Wells 1994) is used to calculate the means and the between-laboratory standard deviations for the measurements. These parameters are used to assess the performance of the group of participating laboratories as a whole. The performance of individual laboratories is assessed with assigned values and independently-established allowable errors are targed. The robust statistical method of the Analytical Methods Committee (AMC 1989) has been applied to evaluate all laboratory performance study results. This is a non-parametric approach, which determines the between-laboratory robust standard deviation of the determinations.

Z-scores are calculated as a simple method of giving each participant a normalised performance score for bias; the method was adopted as a standard for ISO/IUPAC (Thompson & Wood 1993, ISO 1996). Z-scores for each laboratory and each determinand are calculated as follows:

$$Z = \frac{(x_1 - \overline{x})}{S_b},$$

where

x_1 – the reported value for the analyte concentration in the test sample (from a given laboratory),

$\overline{x}$ – the assigned value,

S_b – the target value performance (usually 6% or 12.5%).
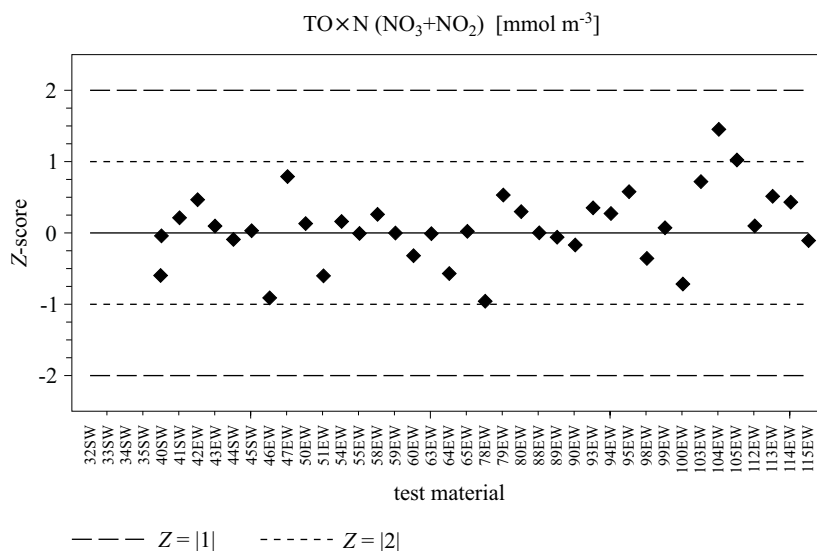
The Z-scores are interpreted as follows:

$|Z| \leq 2$ satisfactory,

$2 < |Z| < 3$ questionable,

$|Z| > 3$ unsatisfactory.

The Z-scores obtained by the marine chemistry laboratory of IMWM in QUASIMEME intercomparison exercises on nutrient determination in sea water are presented in Figs 1–7.
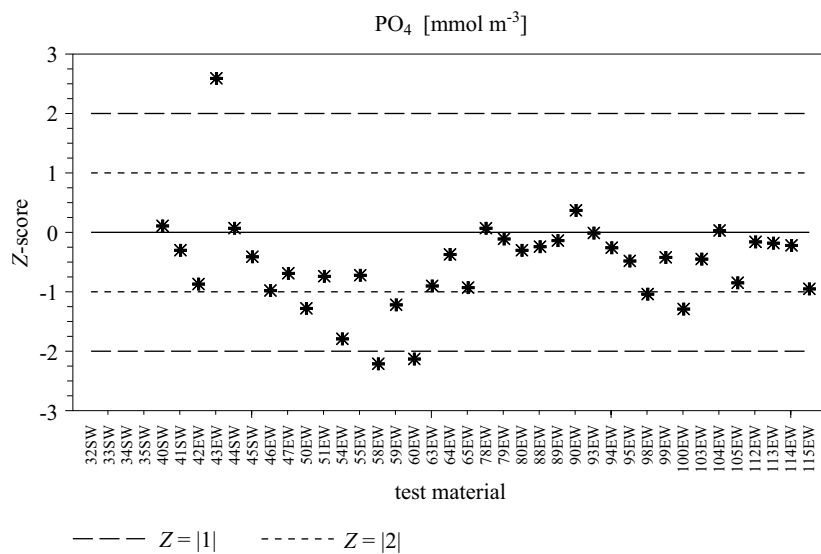

Fig. 1. Results of QUASIMEME tests on nitrate+nitrite determination in seawater and estuarine water obtained by the IMWM in 1994–2003

NO$_2$ [mmol m^{-3}]
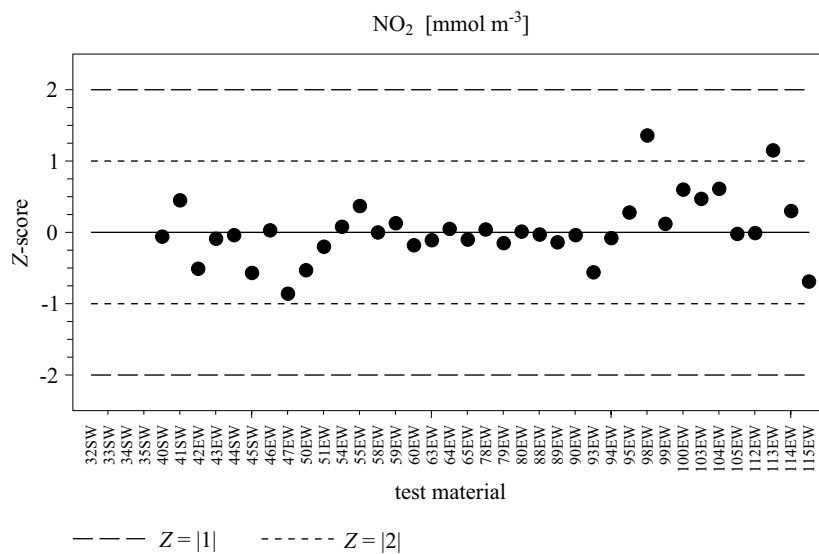
Fig. 2. Results of QUASIMEME tests on nitrite determination in seawater and estuarine water obtained by the IMWM in 1994–2003

NH$_4$ [mmol m^{-3}]
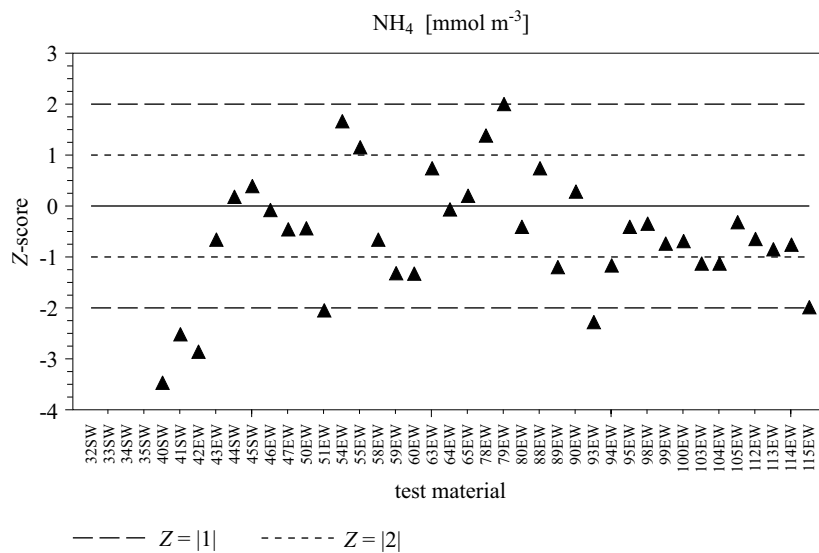
Fig. 3. Results of QUASIMEME tests on ammonia determination in seawater and estuarine water obtained by the IMWM in 1994–2003

PO$_4$ [mmol m^{-3}]

Fig. 4. Results of QUASIMEME tests on phosphate determination in seawater and estuarine water obtained by the IMWM in 1994–2003
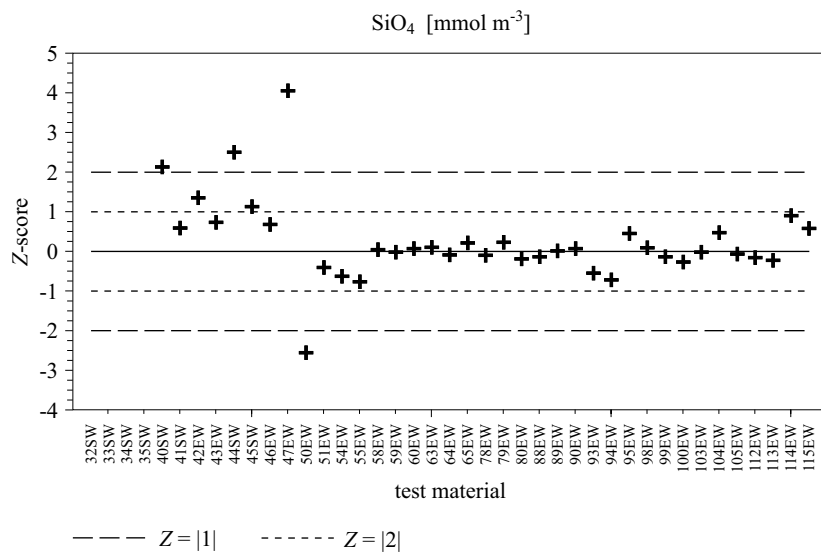
SiO$_4$ [mmol m^{-3}]

Fig. 5. Results of QUASIMEME tests on silicate determination in seawater and estuarine water obtained by the IMWM in 1994–2003

P-tot [mmol m^{-3}]

Fig. 6. Results of QUASIMEME tests on total phosphorus determination in seawater and estuarine water obtained by the IMWM in 1994–2003

N-tot [mmol m^{-3}]

Fig. 7. Results of QUASIMEME tests on total nitrogen determination in seawater and estuarine water obtained by the IMWM in 1994–2003
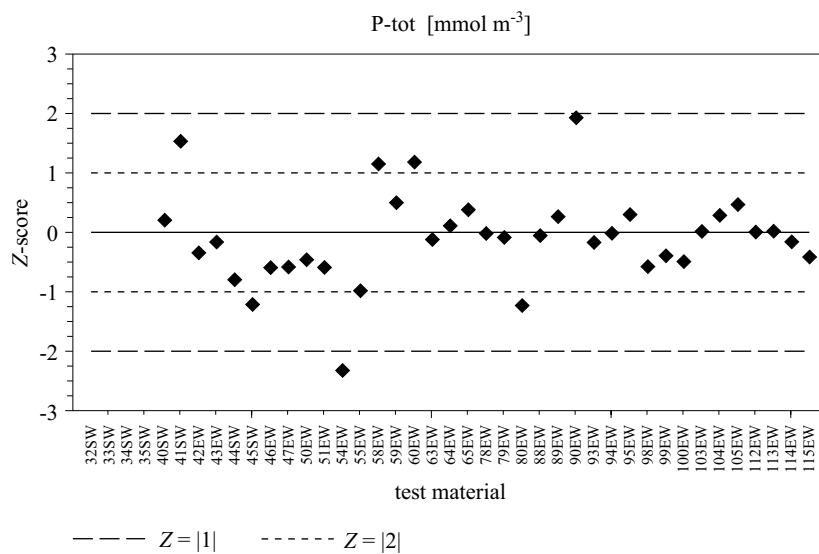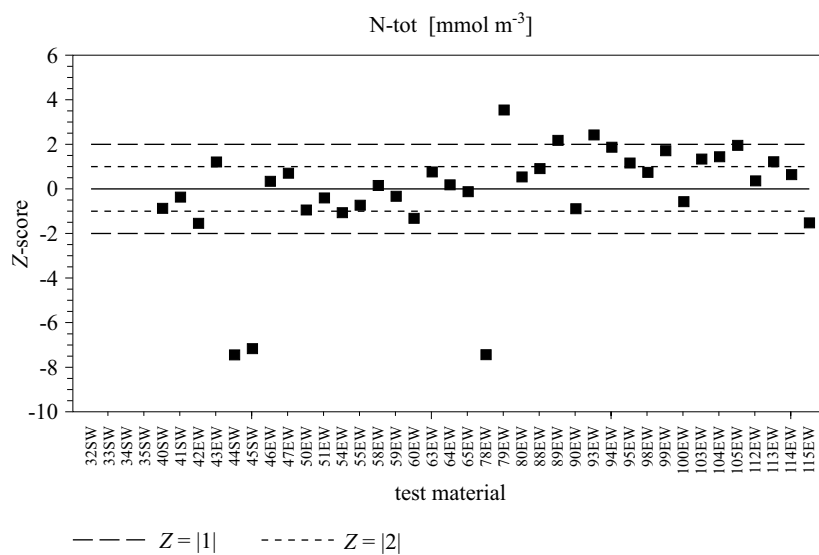
The overall performance characteristics of the results are given in Table 1.

Table 1. Characteristics [%] of the overall performance by the IMWM chemical laboratory in QUASIMEME LPS on nutrients in sea water in 1994–2003

Analyte	$Z > 2$	$Z \leq |2|$	$Z \leq |1|$
TOxN ($NO_3 + NO_2$)	–	100	94.4
Nitrite	–	100	94.4
Ammonia	16.7	83.3	72.2
Phosphate	9.3	91.7	77.8
Silicate	11.1	88.9	80.6
N-tot	16.7	83.3	66.7
P-tot	2.8	97.2	80.6

Over a ten year period, more than 80% of test results are in the range $Z \leq |2|$; this result is satisfactory and the percentage of questionable or unsatisfactory results is less than 20%. Moreover, only in the case of total nitrogen determination (N-tot) does the percentage of results within the range $Z \leq |1|$ fall below 70. The results are considered very satisfactory.

4. Measurement uncertainty

Measurement uncertainty can be estimated by identifying all possible sources of uncertainty associated with a method, quantifying uncertainty components and calculating total uncertainty by combining the individual uncertainty components following appropriate mathematical rules (EU-RACHEM 2003). Another approach utilises the data from routinely undertaken internal quality control measurements, e.g. the results of replicate analyses of certified reference materials (CRM) or laboratory reference materials (LRM). In this case, the relative combined uncertainty of a method is obtained from the contribution from the precision of the method, expressed as the relative standard deviation, and the contribution from trueness of the method, expressed as the relative recovery of the analyte from internal control measurements (EURACHEM 2003).

The relative combined standard uncertainty $u_c(y)$ (y = e.g. phosphate) is expressed as

$$u_c(y) = \sqrt{RSD_{obs}^2 + u(\overline{R}_{\text{LRM}})^2}.$$

The relative standard deviation of the mean concentration, RSD_{obs}, is expressed as

$$RSD_{obs} = \frac{S_{obs}}{\overline{C}_{obs}},$$

where

$\overline{C}_{obs}$ – mean of replicate analyses of the LRM,

S_{obs} – standard deviation of the results from the replicate analyses of the LRM.

The relative standard uncertainty of the recovery $u(\overline{R}_{\mathrm{LRM}})$ is calculated using the following equation:

$$u(\overline{R}_{\mathrm{LRM}}) = \overline{R}_{\mathrm{LRM}}\sqrt{\left(\frac{S_{obs}^2}{n\,\overline{C}_{obs}^2}\right) + \left(\frac{u(C_{\mathrm{LRM}})}{C_{\mathrm{LRM}}}\right)^2},$$

where

C_{LRM} – the certified concentration of the analyte in the LRM (or CRM),

n – the number of replicates ($n \geq 10$, Barwick & Ellison 2000),

$u(C_{\mathrm{LRM}})$ – the standard uncertainty of the certified concentration for the LRM with the mean recovery

$$\overline{R}_{\mathrm{LRM}} = \frac{\overline{C}_{obs}}{C_{\mathrm{LRM}}}.$$

It is normally assumed that $\overline{R}_{\mathrm{LRM}}$ does not differ significantly from 1; otherwise a correction for recovery has to be made (Barwick & Ellison 2000). A statistical significance test (t-test) is performed to test whether the recovery differs significantly from 1; t is calculated using the equation

$$t = \frac{|1| - \overline{R}_{\mathrm{LRM}}}{u(\overline{R}_{\mathrm{LRM}})}.$$

Because the number of degrees of freedom associated with $u(\overline{R}_{\mathrm{LRM}})$ are usually unknown, t is compared with k, the coverage factor that will be applied in the calculation of the expanded uncertainty (Barwick & Ellison 2000).

If $t < k$ (recommended coverage factor $k = 2$ gives an interval containing approximately 95% of the distribution values, i.e. 95% confidence interval) it can be assumed that the recovery is not significantly different from 1.

Using the recommended coverage factor $k = 2$, the expanded uncertainty, $U(\mathrm{LRM})$, is expressed as

$$U(y) = k\,u_c(y) = 2\,u_c(y).$$

The results of the expanded uncertainty calculated for nutrient analyses from internal control data are listed in Table 2.

Table 2. Expanded uncertainty [%] of nutrient analyses in seawater as determined from routine internal control measurements (LRM) and QUASIMEME LPS

Test \ Analyte	TOxN	NO_2	NH_4	PO_4	SiO_4	N-tot	P-tot
U_{LRM}	14.9	25.8	56.2	22.8	17.6	19.1	11.4
U_{LPS}	15.5	62.1*	79.8*	22.9	33.8*	36.3*	38.1*

The measurement uncertainty of a particular method can also be estimated using the results of intercomparison exercises (Magnusson et al. 2004, HELCOM 2004). Estimation of measurement uncertainty using reproducibility data from interlaboratory studies (LPS) is based on the relative reproducibility of standard deviation. The averaged relative reproducibility of standard deviation, expressed as the coefficient of variation $CV(y)$ for the intercomparison study on the given nutrient determination in seawater, can be equated with the combined standard uncertainty $u_c(y)$ and is given by

$$CV(y) = \sqrt{\frac{\sum x_i^2}{n}} = u_c(y),$$

where
x_i – relative standard deviation in the given round or exercise,
n – number of exercises.
Using the recommended coverage factor $k = 2$ the expanded uncertainty $U(y, LPS)$,

$$U(y, \text{ LPS}) = k\, u_c(y) = 2\, u_c(y).$$

Uncertainty values obtained from QUASIMEME Laboratory Performance Studies (round 30, exercise 531, August–November 2002 – test samples 103EW, 104EW and 105EW, and round 32, exercise 545, January–April 2003 – test samples 112EW, 113EW, 114EW and 115EW) for nutrient determination in seawater are compared in Table 2.

The measurement uncertainty values in TOxN and PO_4 analyses estimated from laboratory performance studies are in good agreement with the data obtained from routine internal control. The discrepancy in the uncertainty values marked with asterisks results from analyses of test material in 2003, which contained natural Baltic Sea water with nitrite and nitrate at concentrations near the limit of detection, and low concentrations of total nitrogen and total phosphorus. These low concentrations proved extremely difficult to determine correctly by the participating laboratories and contributed significant errors, hence the large measurement uncertainty

resulting from this exercise. The results of the IMWM laboratory obtained in both these tests were very good (test samples 112EW, 113EW, 114EW and 115EW in Figs 1–7).

The results of external analytical performance control presented here provide evidence that the measurements carried out within the monitoring programme are under control and provide highly reliable data for the relevant environmental assessments.

References

AMC (Analytical Methods Committee), 1989, *Report on an experiment test of recommendation for the conduct and interpretation of co-operative trial*, Analyst, 114, 1489 pp.

Barwick V. J., Ellison S. L. R., 2000, *Development and harmonisation of measurement uncertainty principles – Part (d): Protocol for uncertainty evaluation from validation data*; VAM Project (version 5.1, January 2000), http://www.vam.org.uk/publications

Cofino W. P., Wells D. E., 1994, *Design and evaluation of the QUASIMEME interlaboratory performance studies: a test case for robust statistics*, Mar. Pollut. Bull., 29, 149–158.

Danowska B., Drgas N., Łysiak-Pastuszak E., Piechowska B., 2000, *Elements of quality assurance system in nutrient and dissolved oxygen analyses*, Oceanol. Stud., 4, 15–25.

EURACHEM, 2003, *EURACHEM/CITAC Guide: quantifying uncertainty in analytical measurement*,
http://www.uni-stuttgart.de/eurachem/pdf

Grasshoff K., Ehrhardt M., Kremling K. (eds.), 1983, *Methods of sea water analysis*, Verl. Chemie, Weinheim, 419 pp.

HELCOM, 2004, *Manual for the COMBINE Baltic Monitoring Programme*, http://www.helcom.fi/ec.html

Helsinki Commission, 1993, *Report of the 14th meeting of the Helsinki Commission, 2–4 February, Helsinki 1993*, paragraph 5.24, p. 7.

ISO (International Organization for Standardization), 1996, *ISO/IEC Guide 43–1. Proficiency testing by interlaboratory comparisons*, ISO, Geneva.

Łysiak-Pastuszak E., 1996, *On quality assurance procedures in monitoring analyses of nutrients*, Oceanol. Stud., 3, 3–13.

Magnusson B., Näykki T., Hovind H., Krysell M., 2004, *Handbook for calculation of measurement uncertainty in environmental laboratories*, NORDTEST Rep. TR 537, 2nd edn., Espoo (Finland), 41 pp.

Nehring D., 1994, *Quality assurance of nutrient data with special respect to the HELCOM Baltic Monitoring Programme*, Baltic Sea Environ. Proc., 58, 5–21.

QUASIMEME (Quality Assurance of Information for Marine Environmental Monitoring in Europe), 1993, QUASIMEME Bull., 1 (March 1993), QUASIMEME Project Office, SOAFD Mar. Lab., Aberdeen, http://www.quasimeme.marlab.ac.uk

QUASIMEME, 2003a, *QUASIMEME Laboratory Performance Studies – Progress Report No 7 to the Advisory Board prepared by the Scientific Assessment Group, June 2003*, (final version revised October 2003), http://www.quasimeme.marlab.ac.uk

QUASIMEME, 2003b, *QUASIMEME protocols and reports*, http://www.quasimeme.marlab.ac.uk

Thompson M., Wood R., 1993, *The international harmonised protocol for the proficiency testing of (chemical) analytical laboratories IUPAC/ISO/AOAC*, J. Pure Appl. Chem., 65, 2123–2144.

Topping G., 1992, *The role and application of quality assurance in marine environmental protection*, Mar. Pollut. Bull., 25, 61–66.

UNESCO, 1981, *Background papers and supporting data on the Practical Salinity Scale 1978*, UNESCO Tech. Pap. Mar. Sci. No 37.