OSPAR COMMISSION

Protecting and conserving the North-East Atlantic and its resources

JAMP Guidelines for estimation of a measure for uncertainty in OSPAR monitoring

Agreement 2011-3
(Source HASEC 11/12/1, Annex 5)

Introduction

1. In assessments of contaminants in biota and sediment done for the OSPAR area quality assurance (QA) information reported by the laboratory is presently used to estimate the uncertainty. This uncertainty is utilised to give statistical weight to individual data points. These weights are then used in the further assessment process.

2. This measure of uncertainty is estimated from information reported to the ICES database by the analysing laboratories: This information contains:

- Results of CRM
- Performance in Quasimeme proficiency testing.

3. The processing of this information is tedious and takes time that would be better used for the actual assessment. Presently quite some experience has been built and the estimation of the uncertainty could as well or even better be done by the laboratories themselves. Laboratories are generally all accredited under ISO 17025 or work on the same quality level and have more information available then presently reported to ICES. Estimating the uncertainty by the laboratory therefore can include a broader set of QA information. Another advantage of this approach is that a measure for uncertainty can also be made available for monitoring years where Quasimeme information is absent. Moreover results from other proficiency testing schemes (PTS) can be included and also a value for uncertainty can be obtained for parameters not in the Quasimeme PTS.

4. Methods to estimate measuring uncertainty in the laboratory are widely and freely available but this document gives some guidelines to make sure that some degree of harmonisation is obtained. This document does not describe how QA measures should be applied in a laboratory but only how to merge available QA information to a realistic merged estimation of uncertainty. Note that this document does not consider bias separately but essentially includes it. It is likely that estimations of standard deviations will always be an approximation and may be more variable than the reported analytical data. Note also that this action is not a competition for the lowest value; the best standard deviation is a realistic one. The uncertainty is finally expressed as “expanded uncertainty” equal to twice the standard deviation. This expanded uncertainty should be reported in the same units as the reported value and not as a coefficient of variation.

5. Laboratories that have already implemented a system according to the approach of the Eurachem/CITAC Guide CG 4 and are able to give uncertainty as standard deviations for individual data may not need this guideline. This guideline is partly based on existing documents and the process followed for the OSPAR assessments using information that previously was reported to ICES. This guideline lists in a pragmatic way the ‘*to do-s*’ for the estimation of individual uncertainties for data reported to ICES using several information sources available in the laboratory.

Available QA information

6. Laboratories usually validate applied analytical methods used by examining the properties in terms of limit of detection (LoD), precision and accuracy following guidance given in literature. Eurachem/CITAC

Guide CG 4 suggests that uncertainty of different steps in the analytical procedure are investigated that can be combined to an overall uncertainty. Once methods are validated and in use for monitoring, during application several quality control (QC) measures are undertaken to monitor the quality in time. These are for example:

- Control charts of internal or external reference materials
- Duplicate analyses
- Data obtained from analysis of certified reference materials (CRM)
- Results from proficiency testing schemes.

7. In Eurachem terms these are all already "combined uncertainties" although not all from the same level. Control chart and duplicate data are mainly representing precision and when including all variables e.g. time, chemist and equipment, it can at best give the intermediate precision. Results from CRM and PTS will include also bias. Merging all four QC measures will give an acceptable measure for uncertainty.

8. Two sources of variation are considered. At high concentration the uncertainty will be dominated by a relative error but for data close to the LoD there is a constant error not related to the concentration in the sample. It is recognised that this constant error can be intake dependent but also sample specific depending on matrix disturbances.

9. So to estimate the uncertainty two components are required:

1. a coefficient of variation - *v*, also referred to as relative standard deviation, valid at higher concentration
2. a constant error – *d*, at concentrations close to the LOD.

10. Below is described how these two factors in the uncertainty can be estimated from QC measures and finally combined to give a final uncertainty individual for each measured concentration.

Determination of the coefficient of variation

11. Provided the concentration in the sample material is sufficiently high, e.g. 3 times above the LoD the variability is dominated by a relative component. For results lower than 3 times the LoD the absolute component in the error will dominate and those data should be excluded in calculation of the coefficients of variation.

Control charts

12. For control charts the same sample, internal reference material (IRM) or reference material obtained elsewhere, is analysed on a regular basis or with each batch of samples. If a certified value is provided, the relative error from the control chart data, v_{IRM}, is estimated by:

$$v_{IRM} = \frac{1}{C_{IRM}} \sqrt{\frac{1}{n_{IRM}} \sum_{i=1}^{n_{IRM}} \left(C_{meas}^{i} - C_{IRM}\right)^2} \qquad \textbf{eq. 1a}$$

where C_{meas}^{i} is the concentration of measurement *i* of the IRM, C_{IRM} is the certified value, and n_{IRM} is the number of measurements. If a certified value is not provided, the relative error is estimated by

$$v_{IRM} = \frac{1}{C_{IRM}} \sqrt{\frac{1}{n_{IRM} - 1} \sum_{i=1}^{n_{IRM}} \left(C_{meas}^{i} - C_{IRM}^{AVG}\right)^2} \qquad \textbf{eq. 1b}$$

where C_{IRM}^{AVG} is the average concentration of the measurements.

Duplicate measurements

13. In many labs it is common practice to analyse samples in duplicate spread over different batches and chemists. These duplicates give a realistic measure for intermediate precision. Again, provided the result is more than 3 times larger than the LoD the relative variability v_{dup} is calculated by:

$$v_{dup} = \sqrt{\frac{2}{n_{dup}} \sum_{i=1}^{n_{dup}} \left(\frac{C_A^i - C_B^i}{C_A^i + C_B^i} \right)^2} \qquad \textbf{eq. 2}$$

where C_A^i and C_B^i are the duplicate values for sample i, and n_{dup} is the number of duplicates.

Certified reference materials

14. For CRMs different situation can occur. One or more reference materials can be analysed just once or multiple times in the monitoring year. Further both the repeatability and bias can range from poor to excellent with considerable bias. All this variation is combined into a coefficient of variation from the relative difference of the measured (C_{meas}^i) and certified value (C_{Cert}^i) using:

$$v_{CRM} = \sqrt{\frac{1}{n_{CRM}} \sum_{i=1}^{n_{CRM}} \left(\frac{C_{meas}^i - C_{Cert}^i}{C_{Cert}^i} \right)^2} \qquad \textbf{eq. 3}$$

where n_{CRM} is the total number of measurements. Because the relative difference is used data from multiple CRMs can be combined.

Proficiency testing schemes

15. Results from proficiency testing schemes (PTS) are treated equally as CRMs following:

$$v_{PTS} = \sqrt{\frac{1}{n_{PTS}} \sum_{i=1}^{n_{PTS}} \left(\frac{C_{meas}^i - C_{\mathrm{Ref}}^i}{C_{\mathrm{Ref}}^i} \right)^2} \qquad \textbf{eq. 4}$$

Data from different proficiency testing schemes are combined and n_{PTS} is the number of PTS.

16. Note that Z-scores are not used here as they may be based on different systems, e.g. Z=1 can correspond to 12.5% but also other values could be used. The method described here is universal for all PTSs.

Merging results of different QA measures

17. The results from the above control measures can be merged in several ways. One option is to weight each coefficient of variation by the corresponding degrees of freedom to give an overall coefficient of variation:

$$v_C = \sqrt{\frac{n_{IRM} v_{IRM}^2 + n_{dup} v_{dup}^2 + n_{CRM} v_{CRM}^2 + n_{PTS} v_{PTS}^2}{n_{IRM} + n_{dup} + n_{CRM} + n_{PTS}}} \qquad \textbf{eq. 5}$$

However, our recommendation is to give each coefficient of variation equal weight, since they each represent a different type of QA information. This gives:

$$v_C = \sqrt{\frac{v_{IRM}^2 + v_{dup}^2 + v_{CRM}^2 + v_{PTS}^2}{n}} \qquad \textbf{eq. 6}$$

where n is the number of QA measures available (items in the numerator).

Estimating the constant variability

18. A first estimate of the constant error at low concentrations can be deduced from the LoD, since this is how the LoD is defined. Further information can be obtained from other QA measures. Blanks or results for parameters that are around the detection limit in IRMs, duplicates, CRMs and PTS's can all contribute to a realistic estimate of the constant error. Note that here only data close to the LoD (e.g. < 2xLoD) should be used and that also data lower than the LoD are used, including negative values.

19. The constant error can be deduced from the LoD or underlying data. By definition, the limit of detection (LoD) is set to three times the standard deviation of the blank or samples with very low concentrations. That means that d will usually equal:

$$d_{LoD} = \frac{LoD}{3} \qquad \textbf{eq. 7}$$

20. The IRM sample applied and also the CRMs used may contain compounds at very low concentrations close to the LoD and the standard deviation (d_{IRM}) of the obtained results (including below LoD and negative) can be calculated resulting in a d_{IRM} and d_{CRM} . Results from different CRMs can be combined. This gives

$$d_{IRM} = \sqrt{\frac{1}{n_{IRM}} \sum_{i=1}^{n_{IRM}} \left(C^{i}_{meas} - C_{IRM}\right)^2} \qquad \textbf{eq. 8a}$$

or

$$d_{IRM} = \sqrt{\frac{1}{n_{IRM} - 1} \sum_{i=1}^{n_{IRM}} \left(C^{i}_{meas} - C^{AVG}_{IRM}\right)^2} \qquad \textbf{eq. 8b}$$

depending on whether C_{IRM} is provided, or estimated from the data (C^{AVG}_{IRM}), respectively, and

$$d_{CRM} = \sqrt{\frac{1}{n_{CRM}} \sum_{i=1}^{n_{CRM}} \left(C^{i}_{meas} - C^{i}_{Cert}\right)^2} \qquad \textbf{eq. 9}$$

21. Duplicate measurements of samples with concentrations around the LoD allow the calculation of:

$$d_{dup} = \sqrt{\frac{1}{2n_{dup}} \sum_{i=1}^{n_{dup}} \left(C^{i}_{A} - C^{i}_{B}\right)^2} \qquad \textbf{eq. 10}$$

Note that input data should not be rounded.

22. Data from PTS are generally of limited use concerning the constant error as they are analysed only once and likely no reference value is available for very low concentrations. If a reference value is given, e.g. the reference labs operate with lower LoDs, your result should be in agreement.

23. The different d values are merged to give:

$$d_C = \sqrt{\frac{(d_{LoD})^2 + (d_{IRM})^2 + (d_{dup})^2 + (d_{CRM})^2}{n}} \qquad \textbf{eq. 11}$$

where d_c is the combined constant error. If there is interference, for example due to co-elution in GC, the laboratory can use expert judgment to increase d_C for a sample or set of samples. It is better to increase d_c in this way, although it includes some subjectivity, than report an uncertainty that is an underestimate.

Application

24. If C is the concentration measured, its standard deviation (s_C) is calculated from the standard error and the variation coefficient through:

$$s_C = \sqrt{d_C^2 + v_C^2 C^2}$$ **eq. 12**

The expanded uncertainty (effectively a 95% confidence interval) is then given by

$$U_C = 2s_C$$ **eq. 13**

25. Doing so will likely reveal that some d_C or v_C are missing. Some compounds were always lower than 3 times the LoD and no v_C was obtained, while others never came close to the LoD. How to deal with this is best explained by an example.

26. Considering mussel tissue in which the CB 28 is always close or at LoD in IRMs and CRMs. On the other hand the CB 153 was never near the LoD. A pragmatic solution is to use the d_C of CB 28 for CB153 and similarly the v_C of CB 153 could be used for CB 28. Since both parameters are analysed in the same extract such an exchange is allowed for d_C provided it is supported by more or less equal blanks for CB 28 and CB 153. An exchange of the v_C should be supported by equal variation of the recovery for standards that also passes the analytical procedure. Otherwise this information can be used to adjust them upwards. Note that when compounds are close to LoD in IRMs and CRMs they are like also low in real samples and the same for compounds with higher concentrations. In other words: when a parameter is always far above the LoD the constant standard error is of less importance and consequently its reliability is less relevant. The reverse applies for parameters with values close to LoD.

27. If uncertainty exists about the level of the d_C or v_C they should best be rounded upwards. Similarly when interferences in specific samplers may cause uncertainty higher as estimated above it is possible to increase the reported uncertainty by a factor based on expert judgement. This allows these data still to be reported but they will get appropriate weight in the assessment. However the deviation is clearly caused by a significant bias the data should not be reported.

Reporting

28. From 2010 for each individual data value the uncertainty field should be filled with the expanded uncertainty from eq. 12 and eq. 13 (U_C) using the constant error and coefficient of variation in the best possible way.

29 The expanded uncertainty should be reported in the same units as the data value. Three significant figures should be sufficient.

30. This approach to reporting uncertainty with the measured value means that values below LoD can be reported as well, since the associated uncertainty will give the value the appropriate weight in the assessment.

References and further reading

CEMP Assessment Manual for contaminants in sediment and biota

Draft 25 March 2008. www.ospar.org

EURACHEM/CITAC GUIDE: Quantifying Uncertainty in Analytical Measurement

2nd Edition, EURACHEM / CITAC 2000, http://www.measurementuncertainty.org/mu/QUAM2000-1.pdf

Measurement uncertainty revisited: Alternative approaches to uncertainty evaluation, 2007 European Federation of National Associations of Measurement, Testing and Analytical Laboratories Technical Report No. 1/2007;
http://www.eurolab.org/docs/technical%20report/Technical_Report_Measurement_Uncertainty_2007.pdf

Internal quality control, Handbook for chemical laboratories, Nordest report TR 569, Nordic Innovation Centre.
http://www.nordicinnovation.net/nordtestfiler/tr569_ed_3.pdf