

Monitoring and Assessment Series

CEMP Assessment Manual

Co-ordinated Environmental Monitoring Programme Assessment Manual for contaminants in sediment and biota



**OSPAR Commission
2008**

The Convention for the Protection of the Marine Environment of the North-East Atlantic (the “OSPAR Convention”) was opened for signature at the Ministerial Meeting of the former Oslo and Paris Commissions in Paris on 22 September 1992. The Convention entered into force on 25 March 1998. It has been ratified by Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Luxembourg, Netherlands, Norway, Portugal, Sweden, Switzerland and the United Kingdom and approved by the European Community and Spain.

La Convention pour la protection du milieu marin de l'Atlantique du Nord-Est, dite Convention OSPAR, a été ouverte à la signature à la réunion ministérielle des anciennes Commissions d'Oslo et de Paris, à Paris le 22 septembre 1992. La Convention est entrée en vigueur le 25 mars 1998. La Convention a été ratifiée par l'Allemagne, la Belgique, le Danemark, la Finlande, la France, l'Irlande, l'Islande, le Luxembourg, la Norvège, les Pays-Bas, le Portugal, le Royaume-Uni de Grande Bretagne et d'Irlande du Nord, la Suède et la Suisse et approuvée par la Communauté européenne et l'Espagne.

© OSPAR Commission, 2008. Permission may be granted by the publishers for the report to be wholly or partly reproduced in publications provided that the source of the extract is clearly indicated.

© Commission OSPAR, 2008. La reproduction de tout ou partie de ce rapport dans une publication peut être autorisée par l'Editeur, sous réserve que l'origine de l'extrait soit clairement mentionnée.

ISBN 978-1-906840-20-4
Publication Number No. 379/2008

Contents

Executive Summary	4
1. Introduction	5
1.1 Background	5
1.2 Scope of the manual	5
2. Selection of bases for expressing concentrations	6
2.1 Introduction.....	6
2.2 Conversions of bases for field data	6
2.3 Conversion of assessment criteria to preferred bases	6
3. Methods used for the determination of Background Concentrations	9
3.1 Introduction.....	9
3.2 Background concentrations of contaminants in sediment	9
3.3 Background Concentrations of lead in sediment.....	11
3.4 Background concentrations for PAHs in sediments	13
3.5 Background concentrations of contaminants in biota.....	16
4. Application of Background Concentrations i.e. the derivation and use of Background Assessment Concentrations (BACs)	17
4.1 Introduction.....	17
4.2 Testing whether concentrations are near background or close to zero.....	17
4.3 Setting Background Assessment Concentrations	19
4.4 Technical method for deriving of BACs	21
5. Methods for normalisation of contaminant concentrations in sediments	22
5.1 Introduction.....	22
5.2 Normalisation and required parameters	22
6. Trend analysis	27
6.1 Statistical analyses prior to trend analysis	27
6.2 Uncertainty in biota analysis.....	27
6.3 Uncertainties in sediment analysis	28
7. Method used for trend analysis of time series	35
7.1 Introduction.....	35
7.2 Theory	36
References	36
8. Presentation of temporal trend assessments	37
9. Power of temporal trend programmes to detect changes in concentrations of contaminants	39

Executive Summary

This CEMP Assessment Manual is a companion document to the series of annual assessments of data on reported by OSPAR Contracting Parties on hazardous substances in the marine sediments and biota under the OSPAR Coordinated Environmental Monitoring Programme (CEMP). The manual has been developed by the experts in the OSPAR Working Group on Monitoring (MON) to provide detailed information on the assessment methodologies that are employed in the preparation of these assessments.

The Manual gives details of the methods used for selection of bases, normalisation of concentrations in sediments, temporal trend analysis and determination of background concentrations. Details of the presentation techniques used by MON in OSPAR assessments are also provided.

The manual will be updated as the techniques used by MON are matured.

Récapitulatif

Le manuel d'évaluation du CEMP est un document accompagnant les séries d'évaluations annuelles de données envoyées par les Parties contractantes à OSPAR sur les substances dangereuses dans les sédiments marins et la biote dans le cadre du Programme Coordonné de surveillance continue de l'environnement (CEMP). Ce manuel a été élaboré par les experts du groupe de travail OSPAR sur la surveillance (MON) pour fournir les informations détaillées sur les méthodologies d'évaluation qui sont employées dans la préparation des évaluations.

Le manuel détaille les méthodes utilisées pour la sélection des bases, pour la normalisation des concentrations dans les sédiments, pour l'analyse temporelle des tendances et pour la détermination des concentrations de bruit de fond. Les détails sur la présentation des techniques utilisées par MON pour les évaluations OSPAR sont également fournis.

Le manuel sera mis à jour lorsque les techniques utilisées par MON seront matures.

1. Introduction

1.1 Background

One of the primary tasks of OSPAR MON is to undertake assessments of data on the concentrations of contaminants in the environment and of their biological effects. The work is directed at the OSPAR policy objectives for hazardous substances:

“In accordance with the general objective, the objective of the Commission with regard to hazardous substances is to prevent pollution of the maritime area by continuously reducing discharges, emissions and losses of hazardous substances with the ultimate aim of achieving concentrations in the marine environment near background values for naturally occurring substances and close to zero for man-made synthetic substances.”

The MON assessments of CEMP data have emphasized the role of the assessments of temporal trends in monitoring data as indicators of the progress being made towards achieving near background, or close to zero, concentrations of contaminants, as appropriate. The assessments have been based on data held in the ICES Environmental Database and its successors. An important practical consideration has been to automate, as far as possible, all the data extractions and data handling. This has required the evolution of a range of procedures for handling data, development of assessment criteria (both as significant concentrations and significant trends) statistical procedures for analysing the data and forms of data presentation.

The purpose of this manual is to describe how the important phases of the assessment process are carried out, to provide supporting rationale for the procedures that have been adopted, and a record of how some of the assessment criteria were developed. It is a live document and changes may be expected to occur regularly as assessment techniques develop.

1.2 Scope of the manual

The assessment process as currently implemented comprises the following steps:

- Data are extracted from the ICES database according to specifications defined by the data assessors
- Decisions are made on the bases on which the analyses of contaminants and supporting parameters should be expressed and the assessment carried out, and data are converted to the required bases.
- Assessment criteria are agreed. These include background concentrations, background assessment concentrations, and environmental assessment criteria. Significance levels for assessments of temporal trends are established.
- Expressions are derived for the uncertainties in the analytical data for sediment. Procedures for estimating uncertainties are required for the values of contaminants and co-factors, and also for normalized concentrations.
- For biota, performance in internal and external QC exercises is used to derive initial values for weighting factors, and these are adjusted to appropriate statistical weights during subsequent analysis.
- Contaminant concentrations are normalized to agreed co-factors according to matrix and contaminant combinations.
- Time series data are compiled by station and assessed for linear and non-linear trends over selected time intervals, using weighted LOESS smoothers.
- Tabulations and visual representations of trends are supported by comparisons of concentrations with BACs and/or EACs.
- The power of the data series at each station to detect changes in concentration can be calculated.
- Geographically-based presentations of concentrations and trends are drawn.

This manual presents the principles underlying these various stages of the process, together with examples of their application. The assessment methodology continues to evolve on an annual basis, and therefore this manual is a dynamic and changing document. It is intended that this manual will be reviewed annually by MON and kept up to date by amendment as necessary. The text in Chapters 1 – 9 has, in many cases, been derived from original texts developed by ICES or OSPAR Working Groups and Committees with little additional editing. The principles and procedures laid down in these Chapters continue to be followed by OSPAR MON in assessments of CEMP data. However, the numerical values of assessment criteria such as Background Concentrations and Background Assessment Concentrations have been revisited and revised in the subsequent years as new data have become available to the various Groups involved in the process. Chapter 10 tabulates the current values for assessment criteria used in CEMP assessments, which in some cases differ from those in the preceding Chapters.

2. Selection of bases for expressing concentrations

2.1 Introduction

The ICES Environmental Database holds information on a wide range of contaminants in marine matrices. In reporting analytical data ICES, data submitters can choose the basis on which the concentrations are expressed, for example dry weight (dw), wet weight (ww), lipid weight (lw) etc. It is also possible that the basis of the analysis can change within time series from a single station.

In order to create comparability between data within and between stations, and to allow comparison with assessment criteria, it is necessary to choose the bases on which all concentrations must be expressed. The choice of bases aimed to meet several considerations: scientific validity, uniformity for groups of contaminants for particular tissues and a minimum loss of data. As to the latter, the choice of bases might affect the number of data that could be treated in the statistical analyses, depending on available information on dry weights, wet weights and lipid weights. The bases preferred for biota are:

- dry weights for metals, organochlorines and PAHs in bivalve soft body tissues;
- wet weights for metals and organochlorines in fish muscle, including tail muscle (crustaceans), and metals in fish liver; and
- lipid weights for organochlorines in fish liver.

Concentrations in sediment are expressed on a dry weight basis.

2.2 Conversions of bases for field data

Conversions are necessary to ensure that maximum use is made of the field data in the ICES database. Conversion is only done if the contaminant data for the sample are accompanied by the necessary specific conversion information (e.g. a measured value for % dry weight). This decision typically (2005) meant a 0.3-3% loss of data.

2.3 Conversion of assessment criteria to preferred bases

A consequence of the approach described above for the bases used in expressing field data is that assessment criteria (BACs, EACs) also need to be expressed on the same bases. If the assessment criteria are initially expressed on bases which differ from the preferred bases used in this assessment, it necessary to convert between bases, for example from wet weight to dry weight or lipid weight. A comprehensive set of conversion factors were developed during the 1996 assessment of biota data, making use of data in the ICES database. The factors were reassessed in 2005 (Table 2.1) using data from the ICES database, and the conversion factors were confirmed. Factors were also used to convert whole fish EACs to individual tissue based values, where necessary. Examples of conversions are shown in Tables 2.2, 2.3, and 2.4.

Table 2.1 Summary of data available in the ICES Environmental Database (Dec 2005) on % dry weights and % lipid weights for monitored species / tissues and conversion factors

	Number of yearly medians	1 st quartile	Median	3 rd quartile	100 / median	Selected conversion factor
% dry weight						
Blue mussel	1096	15.0	18.3	22.6	5.5	5
% lipid weight in liver						
Cod	552	34.0	45.0	50.9	2.2	2
Whiting	3	41.0	44.0	44.3	2.3	2
Common dab	242	7.7	15.9	23.0	6.3	7 ¹
Flounder	696	6.1	11.0	14.3	9.1	9
Megrim	59	17.4	20.4	27.0	4.9	5
Plaice	91	8.1	11.2	13.9	8.9	9

¹ A conversion factor of 6 would seem more appropriate, but the conversion factor of 7 was used in practice. This is consistent with the conversion factor used in previous assessments.

Details of the conversions applied to the assessment criteria in the 2005 assessment report are shown in Table 2.2. Criteria for bivalves were converted to the preferred base (dry weight). The conversion of EACs for fish fillet or fish liver required first a conversion from fish (whole carcass) to these tissues (Table 2.3). By using the mean lipid weight concentrations in liver for the different species the EACs for these were estimated (Table 2.4).

Table 2.2 Conversion of assessment criteria to the preferred bases in biota as used in 2005 assessment. The bases used in the 2005 assessment by OSPAR MON are highlighted in bold

Contaminant	Species tissue	Conversion type	Factor	ww µg/kg	dw µg/kg
BRC/BAC					
Cd	blue mussel	wt to dw	5x	110	550
Cu	blue mussel	wt to dw	5x	1100	5500
Hg	blue mussel	wt to dw	5x	10	50
Pb	blue mussel	wt to dw	5x	190	950
Zn	blue mussel	wt to dw	5x	30000	150000
CB 153	blue mussel	wt to dw	5x	0.4	2
ΣCB ₇	blue mussel	wt to dw	5x	0.7	3.5
Hg ²	roundfish ¹			50	
Hg ²	flatfish			70⁴	
EAC					
Cd	blue mussel	wt to dw	5x	55.9	279.5
Hg	blue mussel	wt to dw	5x	1.7	8.5
Pb	blue mussel	wt to dw	5x	1690	8450
DDE	blue mussel	wt to dw	5x	10	50
Dieldrin	blue mussel	wt to dw	5x	10	50
γ-HCH	blue mussel	wt to dw	5x	0.29	1.45
TBT	blue mussel	wt to dw	5x	2.4	12
CB 153	blue mussel	wt to dw	5x	2.5	12.5⁵
ΣCB ₇	blue mussel	wt to dw	5x	10	50
Naphthalene	blue mussel	wt to dw	5x	91	445
3 rings PAH (PA+ANT)	blue mussel	wt to dw	5x	1290	6450
4 rings PAH (FLU+PYR+BAA+CHR)	blue mussel	wt to dw	5x	6900	34500
5 rings PAH (BAP+BKF)	blue mussel	wt to dw	5x	1069	5345
6 rings PAH (BGHIP+ICDP)	blue mussel	wt to dw	5x	73	365
Cd	fish ³			7.35	
Hg	fish ³			3.5	
Pb	fish ³			300	
DDE	fish ³			50	
Dieldrin	fish ³			50	
γ-HCH	fish ³			1.1	
CB 153	fish ³			2.5⁵	
ΣCB ₇	fish ³			10	

- 1 Assumed applicable for whiting 2 BRC for Hg only applied to fish fillet
3 Concerns the whole fish 4 Assumed applicable for herring.
5 Assumed to be 25% of Σ CB₇

Table 2.3 *The EACs for organochlorines for fish, and derived EACs for fish liver, flatfish fillet and roundfish fillet as used in the 2005 assessment*

Contaminant	"Fish" µg/kg ww	Fish liver µg/kg ww (="Fish"*10)	Flatfish fillet µg/kg ww (="Fish")	Roundfish fillet µg/kg ww (="Fish"/10)
DDE	50	500	50	5
Dieldrin	50	500	50	5
Lindane	1.1	11	1.1	0.11
CB 153 ¹	2.5	25	2.5	0.25
Σ CB ₇ ²	10	100	10	1

- 1 Assumed to be 25% of Σ CB₇
2 Sum of CB 28, CB 52, CB 101, CB 118, CB 138, CB 153, and CB 180

Table 2.4 *Estimated EACs for fish liver on lipid weight basis as used in the 2005 assessment*

Contaminant	Species	Fish liver µg/kg ww	Conversion factor	Fish liver µg/kg lw
DDE, dieldrin	Cod/whiting	500	2	1000
γ-HCH		11		22
CB 153 ¹		25		50
Σ CB ₇ ²		100		200
DDE, dieldrin	Dab	500	7	3500
γ-HCH		11		77
CB 153 ¹		25		175
Σ CB ₇ ²		100		700
DDE, dieldrin	Flounder/plaice	500	9	4500
γ-HCH		11		99
CB 153 ¹		25		225
Σ CB ₇ ²		100		900
DDE, dieldrin	Megrin	500	5	2500
γ-HCH		11		55
CB 153 ¹		25		125
Σ CB ₇ ²		100		500

- 1 Assumed to be 25% of Σ CB₇
2 Sum of CB 28, CB 52, CB 101, CB 118, CB 138, CB 153, and CB 180

For ΣCB_7 , firm EACs were available in 2005 for mussels and whole fish (OSPAR agreement 1997-15), and provisional EACs are available for fish liver and muscle. These values were recalculated for CB 153 assuming that CB 153 constitutes about 25% of this sum. The uncertainty introduced by this recalculation was considered negligible.

3. Methods used for the determination of Background Concentrations

3.1 Introduction

Background concentrations (BCs) are an established assessment criterion in OSPAR. The definition of background concentrations used by the OSPAR/ICES Workshop on the Evaluation and Update of Background Reference Concentrations (BRCs) and Ecotoxicological Assessment Criteria (EACs) and How these Assessment Tools Should be Used in Assessing Contaminants in Water, Sediment, and Biota that met in The Hague, The Netherlands in February 2004 is:

The Background Concentration is the concentration of a contaminant at a “pristine” or “remote” site based on contemporary or historical data.

BCs are needed to assess progress of contaminant concentrations towards the OSPAR objective of achievement of background/near background concentrations of contaminants, and in assessing the anthropogenic contribution to the observed concentrations of contaminants in the environment. The BC for a man-made substance is zero.

3.2 Background concentrations of contaminants in sediment

In 2004, the ICES Working Group on Marine Sediments in Relation to Pollution (WGMS 2004) and the Working Group on Statistical Aspects of Environmental Monitoring (WGSAM 2004) worked in collaboration to construct draft background concentrations for OSPAR Coordinated Environmental Monitoring Programme (CEMP) metals (Cd, Hg, Pb) and polycyclic aromatic hydrocarbons (PAHs) in sediments. The summary and the outcome of the work of the two WGs are given below.

3.2.1. Background concentrations of metals in sediment

Prior to the meeting of WGMS 2004, the OSPAR/ICES Workshop on the Evaluation and Update of Background Reference Concentrations (BRCs) and Ecotoxicological Assessment Criteria (EACs) and How these Assessment Tools Should be Used in Assessing Contaminants in Water, Sediment, and Biota developed a simple database for the storage and handling of relevant data.

3.2.1.1 METHOD

During WGMS 2004 meeting, the participants gathered much more data, which were processed and added to the database. The data in the database from the BRC/EAC workshop and the new data added at the WGMS meeting were utilised in the following manner for metals:

- Down-core results dated, or expected to be, from approximately 1850 and before were selected;
- Sediments in estuaries were not included because they are affected by the activities and geology in the catchment areas;
- Generally only fine-grained and/or sieved samples were included. Trace metal concentrations in unsieved samples were assessed by taking into account either the lithium content or grain size analyses, as aluminium can be found in large quantities in some coarse samples of glacial origin;
- All data were normalised to a sample containing 50 g kg⁻¹ Al or 50 mg kg⁻¹ Li, following the rules below:
 - The simple ratio method was used for samples sieved on 20 µm and/or the Li content > 40 mg kg⁻¹,
 - In other cases, the OSPAR guidelines were applied taking into account pivot values. However, data resulting from an extrapolation factor of more than 3 were discarded. If the grain size/Al or Li relations showed more variation than analytically expected, the maximum extrapolation was reduced to a factor of 2;

- Data were evaluated on a regional basis, taking only minimum and median values, thereby disregarding any high outliers. The difference between minimum and median values was taken as an indication of variability;
- The new data sets were accepted and the above rules were applied as widely as possible, but only ratio-normalisation was performed. For some of the data sets, only Li or only Al was available.

3.2.1.2 DISCUSSION

In the discussion, some critical aspects of using and deriving BC values for metals were addressed:

- 1) It was generally accepted that analysis of cores of pre-industrial sediments would be the best choice, as surface sediments cannot be considered as sediments without any influence of anthropogenic contaminants. But even then, there is reason for concern because metals can move within the sediments, as is known for Mn, which could give rise to lower BC values in the cores, as well as naturally increasing values of some metals in the surface sediment.
- 2) Care should be taken especially with diagenetic elements (e.g., Mn, Fe, Cd) when dealing with anoxic sediments. For example, Cd concentrations in sediment can be very high if such sediment was settled under anoxic conditions, i.e., the overlying water was anoxic. The formation of sulphides will function as a collector of metals from the water phase. This is especially notable for Cd and a little less so for Cu, Ni, and other elements which tend to form strong sulphide complexes. Anoxic conditions do not automatically result in higher metal concentrations as it is not expected that sediment from cores originally formed in oxic environments obtain elevated concentrations when they become anoxic at a later stage.
- 3) It should also be noted that in estuaries, sediments could be affected by transport from the catchment areas, which could give a different natural background concentration, albeit one that is difficult to assess.
- 4) Another aspect that had not been taken into account is the possible occurrence of nodules, which contain raised levels of many metals (Borg and Jonsson, 1996). These were probably not present in the cores used for determining BC values. Sediment with recognisable nodules is probably not suitable for assessment using BC values. A suggestion could be to sieve the samples over 20 µm. If the metals present in the nodules are environmentally relevant, i.e., released to the water phase, this will also give higher concentrations in the fine fraction.

While the complications noted above might not be relevant in the underlying data set for the estimation of BC values, deviating conditions should be taken into account when assessing monitoring data.

3.2.1.3 GENERAL APPROACH TO ESTIMATING BCS

For each data set, a range was established by taking median concentrations of each core sample in the region (where more than one individual concentration was available), after the exclusion of data sets felt to be impaired by anoxic conditions or possibly anthropogenic influence (i.e., W-Norway and Biscay Bay and Iberian Coast). Metal concentrations were normalised to both Al and Li concentrations where possible, and a rounded value in the upper range was taken as the BC. Taking the upper value is consistent with earlier work in OSPAR QSR 2000, where only the upper limits for BRCs (now BCs) were taken into account. Likewise, the 2003 meeting of the OSPAR Working Group on Monitoring (MON) decided (before the BACs were developed) to use the upper range values. Consequently, there is no merit in setting ranges and, with regard to a practical use, a single (upper) value as BC is strongly preferred. Some reservations were made with respect to generally applicable values for Ni and As, as consistently lower values were found in the Baltic area.

3.2.1.4 RESULTS

The background concentrations estimated from core data are listed in the upper row of Table 3.1. For comparison, the BRC values from the 1996 workshop on BRCs (Hamburg), when recalculated to 50 g kg⁻¹ aluminium, are given. The table shows that, for most elements, the BC values developed at WGMS 2004 are close to the 1996 upper range values. For Cd, Hg, and Pb, the proposed values are above the former upper values, but less than 10% for Cd and Pb. Turning the 1996 values into round numbers would bring them even closer.

The 1996 report also indicated a global average for shale. The data, after recalculating them to 50 g kg⁻¹ Al, are given for information. In addition, the geologically established values for the content in the earth's crust are given in Table 3.1 (the latter were not normalised to 50 g kg⁻¹ Al).

As the background values were based on core data, it was also roughly investigated whether these values were far from the current concentrations in surface sediments, using results collated from the ICES

database. As it was not possible to normalise those data, only sediments from the fine fraction (20–90 µm) were taken into consideration. Although these samples would also require further normalisation, the compositions are close enough to reference composition to allow a certain degree of comparison. For each metal, the lower 5% percentile was taken as a probable lowest concentration in surface sediments. For all metals, the 5% percentile of ICES data is close to, and often even lower than, the proposed BC value. Comparing BCs with the lower 5% percentile concentrations in the ICES database, it can be seen that Hg and Pb concentrations in current surface sediments are generally well above the suggested background concentration, whereas the other elements are close to or below the suggested background concentration. Especially Ni is below the BC. In addition, the median values are listed, showing that in general the sediments in the OSPAR Convention Area are well above background level, except for Ni where the median concentration is even below the proposed BC. The number of samples for each element is given for information.

The 1996 EAC values are shown, and except for As, the background concentrations are found to fall within the ranges given, with Hg on the lower EAC value. It should be noted that for EAC values no conditions for grain size composition are prescribed for their application.

Table 3.1 Estimated background concentrations and corresponding former BRCs, earth crust values, data held in the ICES database, and OSPAR EACs.

Element	As	Cd	Cr	Cu	Hg	Ni	Pb	Zn
Suggested BC ¹⁾	15 ²⁾	0.2	60	20	0.05	45 ²⁾	25	90
1996 BRC (lower) ³⁾	12	0.04	52	13	0.02	26	8	51
1996 BRC (upper) ³⁾	26	0.17	116	33	0.04	53	23	104
1996 BRC (shale)	6	0.08	51	25	0.02	38	9	65
Earth Crust ⁴⁾	5.5	0.15	51	23		35	11.5	92
ICES ⁵⁾ 5% Percent.	11	0.11	51	14	0.07	18	31	95
Database Median	33	0.50	92	28	0.25	38	94	200
(20–90 µm) (n)	(1304)	(1754)	(1656)	(1858)	(1764)	(1472)	(2162)	(1866)
EAC (low)	1	0.1	10	5	0.05	5	5	50
EAC (high)	10	1	100	50	5	50	50	500

¹⁾The suggested normalised BC values are valid for all regions of the OSPAR Area.

²⁾For these elements, the core data from Region II and the Baltic area suggest that a lower value could be applied for this region.

³⁾The Me/Al ratios were recalculated to a sample composition of 50 g kg⁻¹ Al.

⁴⁾Data for earth crust not normalised.

⁵⁾All sieved data from the ICES database were taken into account, and the lower 5% percentile value and median value were taken as a guideline for the content in surface sediments, with the number of data given parentheses.

Full information on the range of data on metals in sediments used by WGMS and WGSAM to estimate background concentrations and BACs can be found in Annex 3 to the 2004 ICES ACME report. This report contains tabulations and discursive text for each element considered. As an example, the section concerning lead is reproduced below.

3.3 Background Concentrations of lead in sediment

Six cores were available normalised to aluminium, covering Regions I, II, and IV. Normalised to lithium, four cores were available, covering Regions I and IV.

Median lead value for cores normalised to aluminium were between 6.8 mg kg⁻¹ and 9.1 mg kg⁻¹, with the exclusion of Oystergrounds (29 mg kg⁻¹) and the Bay of Biscay and Iberian Coast (19 mg kg⁻¹) (Table 3.2). Generally, the minimum values were within 25% of the median.

For the new data sets, Region I data from Norway was not available, as no Al results were available. Swedish data on lead from Region I (Kattegat/Skagerrak) ranged from 21 mg kg⁻¹ to 29 mg kg⁻¹, with minimum values about 50% of the median values. The 90% percentile was up to 44 mg kg⁻¹ for the Skagerrak. For Region II, a median lead value of 27 mg kg⁻¹ was found, with a maximum of 49 mg kg⁻¹; for Region III, a median of 17 mg kg⁻¹ and a maximum of 20 mg kg⁻¹ were found. For the HELCOM area, median lead values ranged between 14 mg kg⁻¹ and 30 mg kg⁻¹. The 90% percentile was below 46 mg kg⁻¹. The median values from the new data sets around three times higher than those selected from the original data set.

Median lead values for cores normalised to lithium were between 11.2 mg kg⁻¹ and 12.5 mg kg⁻¹, with the exclusion of the Bay of Biscay and Iberian Coast (30 mg kg⁻¹). Generally, the minimum values were within 25% of the median.

For the new data sets, Region I data on lead from Norway had a median value of 20 mg kg⁻¹, with a maximum value 45 mg kg⁻¹. Swedish data from Region 1 (Kattegat/Skagerrak) ranged from 25–35 mg kg⁻¹, with minimum values up to 50% below median values. The 90% percentile was up to 58 mg kg⁻¹ for the Skagerrak. No Li values were available for Region II or Region III data. For the HELCOM area, median values of lead ranged between 27 mg kg⁻¹ and 44 mg kg⁻¹. The 90% percentile was high for the Stockholm Archipelago (76 mg kg⁻¹), but otherwise was below 67 mg kg⁻¹. The median values from the new data sets are about four times above those of the original data set.

The estimated background concentration of lead is set to 25 mg kg⁻¹.

Table 3.2: Median lead value for cores normalised to aluminium					
CONTAMINANT	Pb				
COMPARTMENT	Sediment				
COFACTOR	Al		mg kg ⁻¹		
TYPE	OSPAR Reg.	LOCATION	MEAN	MEDIAN	MIN
CORE	I	Barents Sea	9.6	8.6	6.8
		Spitzbergen	9.7	9.1	8.1
		W-Norway	7.6	7.6	7.3
	II	Oystergrounds	29.0		
	IV	Bay of Biscay and Iberian Coast	25.0	19.0	16.0
MONIT	IV	Biscay Bay	9.0	6.8	10.0
		English Channel	26		32
	V	Bay of Biscay	33.1		30
SURFACE	I	Rockall Bank and Rockall trough	14.0	13.0	11.0
		Laptov Sea, Siberia		15.3	
Added at the 2004 WGMS meeting:					
CORE	I	Skagerrak, 51 cores	31.8	29.7	14.8
		Kattegat, 27 cores	21.4	20.5	14.0
	II	Norway 5 cores	28.8	24.9	15.9
	III	Belfast Lough		16.5	
		Duich 1996		17.5	15.5
	HELCOM	Baltic	29.9	29.8	27.7
		Stockholm Archipelago	27.6	21.2	12.6
		SW Baltic	19.5	18.0	13.0
S Bothnian Sea		15.9	13.8	10.1	

Table 3.2: Median lead value for cores normalised to lithium				
CONTAMINANT	Pb			
COMPARTMENT	Sediment			
COFACTOR	Li	mg kg ⁻¹		
TYPE	OSPAR Reg.	LOCATION	MEDIAN	MIN
CORE	I	Barents Sea	11.2	8.8
		Spitzbergen	12.5	10.5
		W-Norway	11.2	10.7
	IV	Bay of Biscay and Iberian Coast Biscay Bay	30.0	20.0
MONIT	II	English Channel		23.8
	IV	Bay of Biscay		28.2
	V	Rockall Bank and Rockall trough	16.0	11.0
SURFACE	I	Laptov Sea, Siberia	16.7	
Added at the 2004 WGMS meeting:				
CORE	I	Norway, 15 cores	20.9	20.1
		Skagerrak, 51 scores	35.0	17.7
		Kattegat, 27 cores	25.5	19.2
	HELCOM	Baltic	44.2	39.9
		Stockholm Archipelago	38.3	20.5
		SW Baltic	26.4	19.8
		S Bothnian Sea	27.6	19.6

3.4 Background concentrations for PAHs in sediments

For the organic contaminants, BCs were only considered and discussed for the PAHs and specifically those PAH compounds that were part of the OSPAR CEMP in 2004. Two approaches were identified that could lead to appropriate background concentrations, namely, the use of data from deep cores and the use of present-day concentrations in remote (pristine) areas, i.e., distant from known sources and, therefore, probably primarily influenced by atmospheric/diffuse inputs.

The use of data from deep cores would be consistent with the recommendations from the 2004 OSPAR/ICES Workshop on BRCs and EACs that the background concentrations for synthetic substances should be zero. This is also consistent with the OSPAR Strategy for Hazardous Substances, namely, “the ultimate aim of achieving concentrations in the marine environment near background values for naturally occurring substances and close to zero for man-made synthetic substances.”

However, there are some difficulties in using PAH data from deep cores to define background concentrations. These include:

- Uncertainties as to the extent of environmental (habitat) change that had occurred during the period of time represented by the core. For example, core data were presented from the Netherlands which suggested that, while the cores had been collected from marine areas, the deeper parts might represent terrestrial or littoral environments;
- Uncertainties as to the rate of degradation of PAHs in sediment cores (oxic or anoxic) over periods of decades to centuries. Significant degradation could lead to observed concentrations being lower than when the sediment was initially deposited;
- Uncertainties arising from the rather incomplete coverage of the OSPAR Convention Area by cores analysed for PAHs;
- The limited occurrence of stable depositional environments from which suitable cores might be obtained for subsequent analysis.

It was therefore concluded that data for present-day concentrations in surface sediments from remote (pristine) areas, i.e., areas distant from known sources and, therefore, probably primarily influenced by

atmospheric/diffuse inputs, would be suitable for the derivation of BCs. The BCs should reflect “consensus” values of these data, and should not be dominated by extreme high or low values.

A database of suitable data had been initiated at the 2004 OSPAR/ICES Workshop on BRCs and EACs, and further data had been made available prior to and during the WGMS meeting. All data were normalised to organic carbon concentrations.

In order to limit the influence of extreme values, the data sets were summarised as median values of normalized concentrations and initially visualised as histograms (e.g. Figure 3.1). The Background Concentrations were then estimated as the medians of the medians. The BC values for PAH compounds obtained by this procedure, normalised to 2.5% organic carbon, are shown in Table 3.4.

Table 3.4 Proposed Background Concentrations for PAH compounds in sediment.

PAH compound	Calculated BC $\mu\text{g kg}^{-1}\text{ OC}$	Proposed BC $\mu\text{g kg}^{-1}\text{ OC}$	Proposed BC $\mu\text{g kg}^{-1}$ normalised to 2.5% OC
Naphthalene	191	190	5
Phenanthrene	667	670	17
Anthracene	109	110	3
Fluoranthene	792	800	20
Pyrene	507	500	13
Benzo[a]anthracene	362	360	9
Chrysene	436	440	11
Benzo[a]pyrene	577	580	15
Indeno[1,2,3,c,d]pyrene	1983	2000	50
Dibenzo[a,h]anthracene	274	270	7
Benzo[g,h,i]perylene	1756	1800	45

Note: The values above for dibenzo[a,h]anthracene are based on only three data points, and therefore have particularly high uncertainty.

Full information on the range of data on PAHs in sediments used by WGMS and WGSAM to estimate background concentrations and BACs can be found in Annex 3 to the 2004 ICES ACME report. This report contains histograms showing the median values of the datasets available for each compound considered. As an example, the Figure concerning benzo[a]pyrene is reproduced here at Figure 3.1.

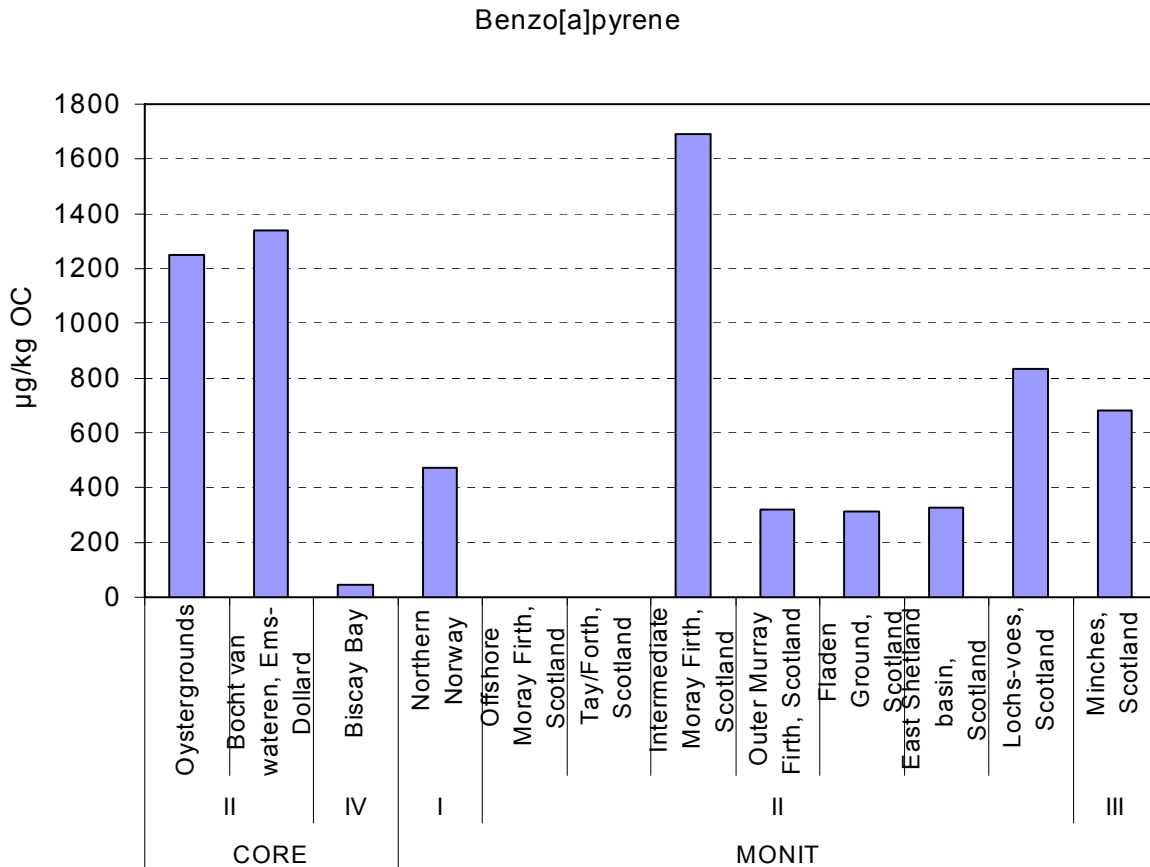


Figure 3.1 Typical dataset used to derive BCs for PAH compounds in sediments

Data were also made available on PAH and organic carbon concentrations in deep sediment cores from the Stockholm archipelago and the southwestern Baltic (south of Skania, Arkona Basin area). Concentrations were normalised to 2.5% organic carbon, and medians were compared to the proposed Background Concentrations (Figure 3.2). The data from the Baltic Sea cores are of the same order of magnitude as the proposed BC values, suggesting that the proposed BC values may also be applicable in the Baltic Sea area.

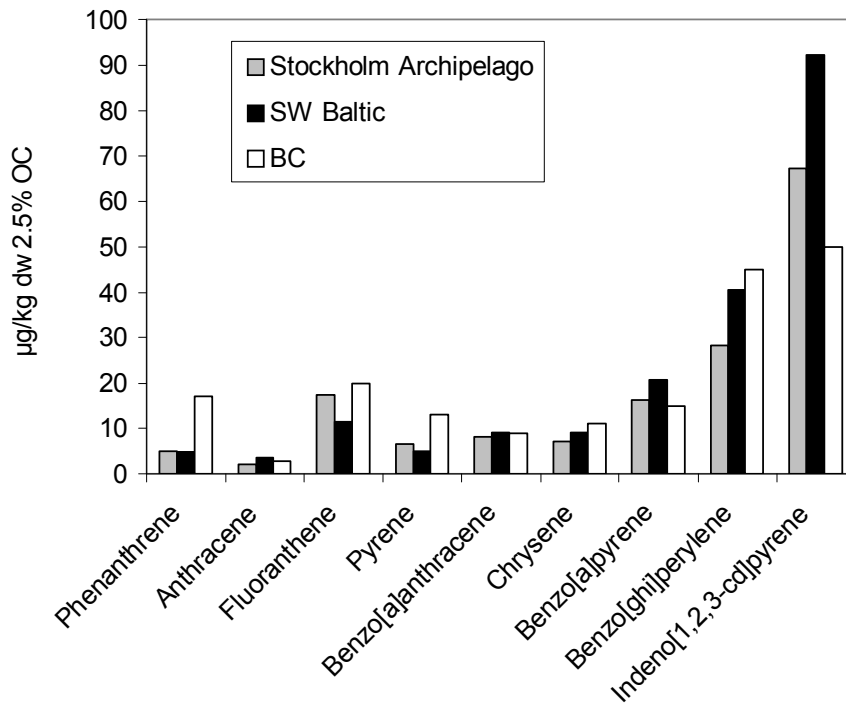


Figure 3.1 Comparison of PAH concentrations in deep cores from the Baltic Sea with proposed Background Concentrations, all normalised to 2.5 % organic carbon.

References

Borg, H., and Jonsson, P. 1996. Large-scale metal distribution in Baltic Sea sediments. *Marine Pollution Bulletin*, 32: pp 8-21.

ICES. 2003. A precautionary approach to testing for compliance with environmental objectives. *In* Report of the ICES Advisory Committee on the Marine Environment, 2003. ICES Cooperative Research Report, 263: pp 156–162.

3.5 Background concentrations of contaminants in biota

Background concentrations and background assessment concentrations for contaminant concentrations in biota have not been established for metals and PAHs. Therefore, the current approach to assessment (2005/2006 CEMP assessment) is to use the Background Reference Concentrations (BRCs) from OSPAR Agreement 1997-4.

4. Application of Background Concentrations i.e. the derivation and use of Background Assessment Concentrations (BACs)

Statistical tests are required to determine whether the concentrations of a contaminant, derived from monitoring data, comply with background concentrations, i.e. will help assess the OSPAR policy objective for hazardous substances namely:

“achieving concentrations in the environment near background values for naturally occurring substances and close to zero for man-made synthetic substances.”

The method used involves the use of the Background Concentrations and it adopts a precautionary statistical approach to the comparison of monitoring data with Background Concentrations. The method adopted requires the establishment of a secondary concentration level, the Background Assessment Concentration (BAC). The BAC is a concentration near to the background and its value for a particular contaminant will depend, for PAHs and metals, on the BC and the residual variance in temporal trend series at the BC. The BC for xenobiotics is zero, and in this case the variance used to derive BACs is the variance at a low concentration that is small but detectable by common analytical methods. The use of BACs is considered as:

- being statistically sound and based on a precautionary approach;
- having a potential for wide applicability covering all contaminants, natural and man-made in all regions of the Convention Area (providing BCs are available);
- being applicable to sediment and biota, and also potentially to water
- having application as a strategic management tool by countries wishing to assess the status of their
- marine environments;
- allowing OSPAR to test its policy objectives.

The use of Background Concentrations in the assessment of CEMP data

4.1 Introduction

The OSPAR policy objective for hazardous substances is

... achieving concentrations in the environment near background values for naturally occurring substances and close to zero for man-made synthetic substances

(OSPAR, 1998). Two questions that follow are:

- How do we quantify *near background* and *close to zero*?
- How do we test whether the objective has been met?

MON 02/4/2 describes a precautionary test for assessing the OSPAR policy objective and shows how it can be incorporated in the current OSPAR methodology for assessing CEMP data. Section 4.2 describes the precautionary test. The test relies on the setting of *Background Assessment Concentrations*¹ that quantify what is meant by *near background* and *close to zero*. Section 4.3 shows how provisional Background Assessment Concentrations (BACs) might be established based on the precision achieved by the CEMP.

4.2 Testing whether concentrations are near background or close to zero

Suppose that we wish to compare the mean concentration [c] of a contaminant in a matrix at a monitoring site to the Background Concentration (BC). The traditional way of testing is to assume that the mean concentration is at background (i.e. [c] ≤ BC) unless there is statistical evidence to show that it is above background (i.e. [c] > BC). Formally, the test involves a null hypothesis H₀ that states that

$$H_0: [c] \leq BC \quad (\text{i.e. concentrations at the site are at background})$$

¹ called *near-background concentrations* in the original document.

The alternative hypothesis H_1 states that

$$H_1: [c] > BC \quad (\text{i.e. concentrations at the site are above background})$$

We conclude that concentrations are at background unless there is sufficient statistical evidence to show that concentrations are above background (i.e. we accept H_0 unless there is sufficient statistical evidence to reject it in favour of H_1).

To make the test, we need to use monitoring data to estimate the mean concentration and put confidence limits around the mean². The test itself uses the lower confidence limit. We conclude that concentrations are at background if the lower confidence limit is below the BC (e.g. site 1 in Figure 4.1). We conclude that concentrations are above background if the lower confidence limit is above the BC (e.g. site 2 in Figure 4.1).

This test is sometimes called a *brown* test because the benefit of doubt is against the environment. To illustrate, if the mean concentration is estimated with poor precision there can be a large chance of concluding that concentrations are at background even when the mean concentration is actually much greater than the BC (e.g. site 3 in Figure 4.1). As such, brown tests are counter to the precautionary principle advocated by OSPAR (OSPAR 1998).

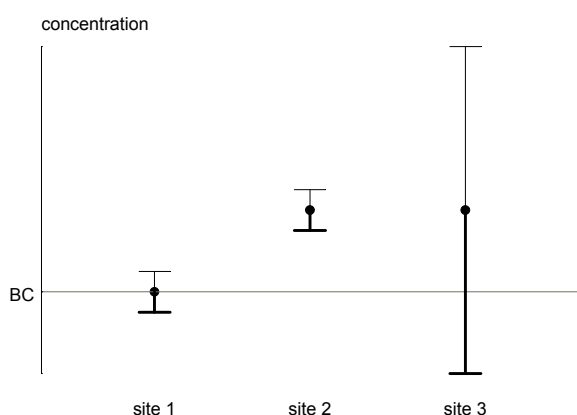


Figure 4.1 Illustration of the brown test; the dots indicate the estimated mean concentration and the bold bars indicate the lower confidence limit. We conclude that concentrations at site 2 are above background and that concentrations at site 1 and 3 are at background.

A precautionary or *green* test is one where the burden of demonstrating that concentrations are at background rests with the monitoring programme. The test is obtained by reversing the order of the hypotheses in the brown test. In the green test, we assume that the mean concentration is above background unless there is statistical evidence to show that it is at background. The null and alternative hypotheses are:

$$H_0: [c] > BC \quad (\text{i.e. concentrations above background})$$

$$H_1: [c] \leq BC \quad (\text{i.e. concentrations at background})$$

The test now uses the upper confidence limit. We conclude that concentrations are above background if the upper confidence limit is above the BC. We conclude that concentrations are at background if the upper confidence limit is below the BC. The precautionary nature of the test is illustrated in Figure 4.2 where we conclude that concentrations are above background at site 3. However, the green test also concludes that concentrations are above background at site 1 (Figure 4.2) so the test is too precautionary.

² The current OSPAR methodology for assessing the CEMP data does this by fitting a smoother to annual contaminant indices and using the pointwise confidence bands around the fitted smoother.

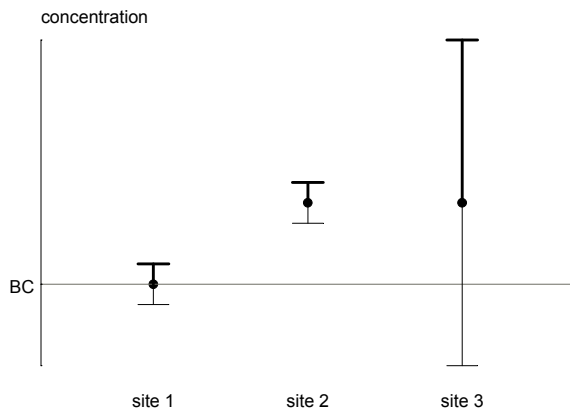


Figure 4.2 Illustration of the green test; the dots indicate the estimated mean concentration and the bold bars indicate the upper confidence limit. We conclude that concentrations are above background at all three sites.

A solution is to use a modified green test in which some environmental protection is sacrificed for a more effective test. The modified green test assumes we can establish a Background Assessment Concentration (BAC) below which concentrations can be considered near background. We now assume that the mean concentration is above background³ (i.e. $[c] > BAC$) unless there is statistical evidence to show that it is near background (i.e. $[c] \leq BAC$). The null and alternative hypotheses are now:

$$H_0: [c] > BAC \quad (\text{i.e. concentrations above background})$$

$$H_1: [c] \leq BAC \quad (\text{i.e. concentrations near background})$$

The test again uses the upper confidence limit. We conclude that concentrations are above background if the upper confidence limit is above the BAC. We conclude that concentrations are near background if the upper confidence limit is below the BAC. To illustrate, we now conclude that concentrations at site 1 are near background, but that concentrations at sites 2 and 3 are above background (Figure 4.3).

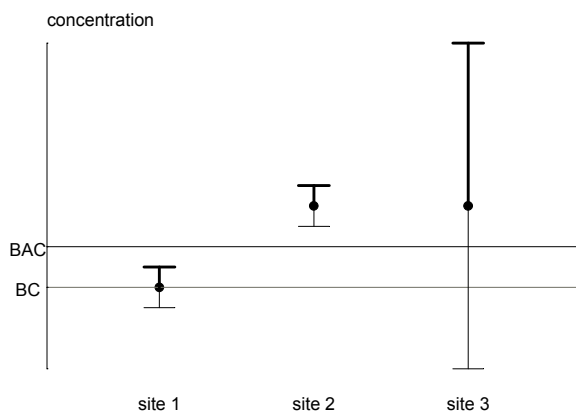


Figure 4.3 Illustration of the modified green test; the dots indicate the estimated mean concentration and the bold bars indicate the upper confidence limit. We conclude that concentrations are above background at sites 2 and 3 and near background at site 1.

4.3 Setting Background Assessment Concentrations

The main practical difficulty with the modified green test is setting a suitable Background Assessment Concentration. The BAC should be both:

- relevant – low enough to reflect near background concentrations
- effective – high enough that we are likely to conclude that concentrations are near background when $[c] = BC$.

To illustrate effectiveness, consider site 1 (in Figures 4.1-4.4) where concentrations are at background (i.e. $[c] = BC$). If the BAC is set too close to the BC relative to the precision of the monitoring programme, we

³ strictly, above near-background.

are likely to conclude, incorrectly, that concentrations at site 1 are above background (Figure 4.4). On the other hand, if the BAC is compatible with the precision of the monitoring programme, we are likely to conclude, correctly this time, that concentrations at site 1 are near background (Figure 4.3).

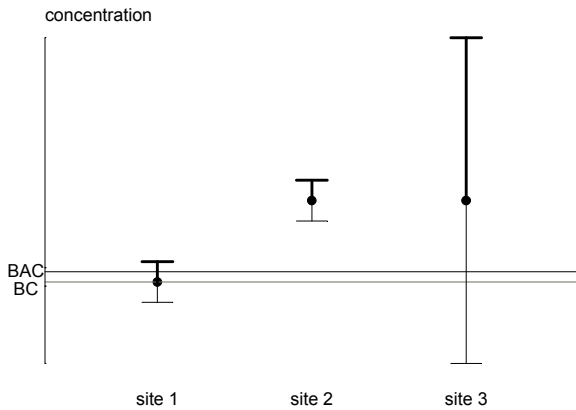


Figure 4.4 If the BAC is set too close to the BC relative to the precision of the monitoring programme, we are unlikely to ever conclude that concentrations are near background (see site 1).

In an ideal world, the BAC would be set on environmental / ecotoxicological grounds. The monitoring programme would then be designed with sufficient precision to give a high probability of concluding that concentrations are near background when $[c] = BC$. However, in practice, there is often no firm environmental / ecotoxicological basis for setting the BAC and monitoring budgets will limit the achievable precision.

A pragmatic way forward is to consider what is achievable under the CEMP. CEMP data can be assessed to evaluate the precision of the programme. Provisional BACs can then be set to give a high probability of concluding that concentrations are near background when $[c] = BC$. However, these BACs should only be agreed if they are relevant. One way of considering relevance would be to compare them to lower EACs.

To illustrate the process, we considered temporal monitoring data from the UK National Marine Monitoring Programme. The table gives the precision of the programme⁴ summarised by contaminant group and matrix.

	Sediment	Shellfish	Fish	Water
Metals	11%	14%	21%	11%
CBs	32%	30%	36%	
PAHs	21%	27%		

Figure 4.5 shows the corresponding probability⁵ (power) of concluding that concentrations are near background when $[c] = BC$ as the BAC increases relative to the BC. Thus, for metals, setting the BAC to be twice the BC would give at least 90% power of concluding that concentrations are near background when $[c] = BC$. Comparable power for PAHs and CBs would be achieved by setting the BAC to be 2.5 times and 3.5 times the BC respectively. Of course, different multipliers could be used for contaminant group / matrix combination, or indeed for each contaminant / matrix combination, if appropriate.

⁴ expressed as the % coefficient of variation on the estimated mean concentration in the final monitoring year; it is assumed that there are annual data over a ten-year period and that the data are modelled following the current OSPAR methodology using median log concentrations as the annual index, a LOESS smoother, and a seven-year fixed window; the % coefficient of variation is then $100\psi\sqrt{(SS')_{TT}}$ where S is the smoothing matrix, the subscript TT denotes the elements corresponding to the final monitoring year; and ψ is the residual standard deviation about the fitted smoother.

⁵ assuming a one-tailed t-test at the 5% significance level (i.e. using the upper 95% confidence limit).

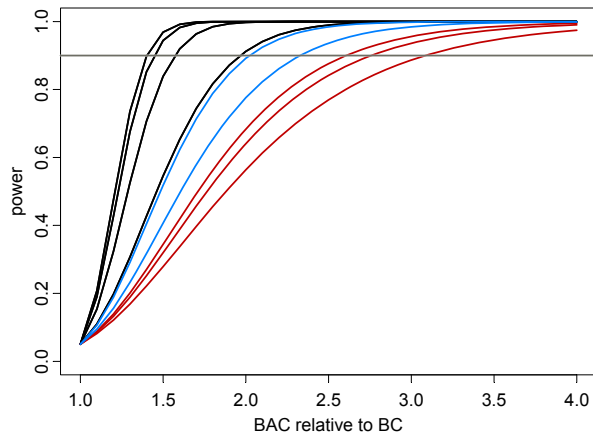


Figure 4.5 The power of concluding that concentrations are near background when $[c] = BC$ for different values of BAC based on UK monitoring data. Running from left to right, the black lines are for metals in sediment, water, shellfish and fish; the blue lines are for PAHs in sediment and shellfish; and the red lines are for CBs in shellfish, sediment and fish.

4.4 Technical method for deriving of BACs

BACs are used to make precautionary tests of whether observed concentrations are near background. The BAC is a concentration greater than the BC that quantifies what is meant by near background or close to zero. The test assumes that the mean concentration $[c]$ is above background⁶ (i.e., $[c] > BAC$) unless there is statistical evidence to show that it is near background (i.e., $[c] \leq BAC$). Formally, the null and alternative hypotheses are:

$$H_0: [c] > BAC \quad (\text{i.e., concentrations above background})$$

$$H_1: [c] \leq BAC \quad (\text{i.e., concentrations near background})$$

and H_0 is rejected in favour of H_1 if the upper confidence limit on $[c]$ is below the BAC.

BACs should be both low enough to reflect near background concentrations and high enough that we are likely to conclude that concentrations are near background when $[c] = BC$. In the absence of other objective means of setting the BAC, the current precision of the CEMP data can be used to set a provisional BAC. Specifically, the BAC can be set to give a 90% probability of concluding that concentrations are near background when $[c] = BC$.

Technically, the BAC can be constructed as follows. Under the current OSPAR methodology⁷, the mean log-concentration in the final year of a time series with at least ten years of data and residual standard deviation ψ will be estimated with a standard deviation

$$\text{s.d.} = \psi \sqrt{(SS')_{TT}} = 0.727\psi$$

where S is the smoothing matrix and the subscript TT denotes the elements corresponding to the final monitoring year. Given this precision, setting the BAC to satisfy

$$\log BAC - \log BC = 3.18 \text{ s.d.} = 2.31\psi$$

will ensure that a one-tailed t-test of size 5% will have 90% probability of concluding that the mean log-concentration is below $\log BAC$ when the true mean log-concentration is $\log BC$. Ignoring the philosophical difficulties of moving from mean log-concentrations to mean concentrations, this suggests that if one can find a value of ψ that is typical of data collected under the CEMP, the provisional BAC should be

$$BAC = BC \exp(3.18 \text{ s.d.}) = BC \exp(2.31\psi).$$

Temporal monitoring data from the UK National Marine Monitoring Programme were used to establish typical levels of variability (field and analytical combined) at concentrations near background. For each contaminant/matrix combination, the residual standard deviation ψ was estimated for all available time series. This was then used to calculate the % coefficient of variation (% CV) of the estimated mean concentration in the final year of a ten years time series:

⁶ strictly, above near-background

⁷ time series of annual median log-concentrations, LOESS smoother, and a seven-year fixed window.

$$\% \text{ CV} = 72.7\psi$$

Since residual standard deviations, or equivalently the % CVs, might be greater at lower concentrations where analytical measurements are more challenging, the relationship between variability and mean concentration was summarised by a robust LOESS smoother. The smoother was then used to estimate a “typical” % CV at the BC, and from this, the provisional BAC was calculated as

$$\text{BAC} = \text{BC} \exp(3.18 \text{ CV})$$

Modifications were needed to deal with CBs where the BC is zero. Specifically, the BAC for individual CBs was calculated as

$$\text{BAC} = 0.1 \exp(3.18 \text{ CV})$$

The factor 0.1 was adopted because it is twice the QUASIMEME constant error in estimating CBs (expressed as $\mu\text{g kg}^{-1}$) and thus represents low concentrations that should be measurable. The CV is again taken from the robust LOESS smoother and represents typical variability in monitoring data at concentrations around $0.1 \mu\text{g kg}^{-1}$. The BAC for $\Sigma 7\text{CB}$ was calculated as

$$\text{BAC} = 0.4 \exp(3.18 \text{ CV})$$

on the premise that 0.4 represents a concentration of $0.1 \mu\text{g kg}^{-1}$ in the dominant CB and nominal concentrations of $0.05 \mu\text{g kg}^{-1}$ in the other six CBs.

5. Methods for normalisation of contaminant concentrations in sediments

5.1 Introduction

In addition to pollution levels, contaminant concentrations in sediments also depend on the composition of individual sediment samples. In trend assessments, trends should be detected irrespective of changes in, for example, particle size distribution, or organic carbon content. The procedures for normalisation are described in Technical Annex 5 to the JAMP Guidelines for Monitoring Contaminants in Sediments (reference number 2002-16), and were applied to the data in the current assessment.

5.2 Normalisation and required parameters

Contaminant concentrations in sediments are normalised as followed:

$$C_{SS} = (C_M - C_X) \frac{N_{SS} - N_X}{N_M - N_X} + C_X \quad (1)$$

where:

- C_{SS} Normalised concentration
- C_M Measured concentration of contaminant
- C_X Pivot value for the contaminant
- N_X The pivot value for the cofactor
- N_M The measured concentration (C_M) of the cofactor
- N_{SS} Reference composition of the sediment as represented by cofactor content

The constants C_X and N_X (i.e. the pivot values for contaminant and cofactor) have been discussed in Annexes 8 and 9 to the 2002 report of the ICES Working Group on Marine Sediments in Relation to Pollution (WGMS). The report describes how to estimate pivot values for contaminants as well as co-factors. The report proposed a procedure for normalising the measured values taking into consideration the pivot values. Also a proposal was made for a reference composition (N_{SS}).

For normalisation (ICES WGMS 2002 report) of sediment data, it is necessary to establish pivot values for the different contaminants and also the errors associated with normalisation. Pivot values, cofactors and measures of their uncertainty are available for all of the contaminants involved in this assessment (Table 4.1). In addition, a kind of state of the art standard analytical error is listed; the fixed (s) and variable (v) component. The pivot values vary slightly with the analytical method used. The pivot values were derived from data where a strong partial digestion method was used. Total digestion methods may free more metals, especially aluminium, from the sediment. A detailed analysis of the influence of digestion procedure on pivot values is given in Appendix 4B of the 2005 OSPAR CEMP data Assessment.

Table 4.1 Sediment assessment – summary information for pivot values, where the pivot values (N_x) used for different determinands measured by different extraction/ digestion methods. The table also lists the values of the constant error (S_{Cm}) and variable error (V_{Cm}) components of the variance of each determinand/method combination, and the concentrations (N_{ss}) of cofactors to which the contaminant concentrations are normalised (From Annexes 8 and 9 of the ICES WGMS2002 report). Note that $S_{Nm} = S_{Nmco}$

Parameter	Parameter Description	Unit	Digestion*	C_x or N_x	S_{C_x} or S_{N_x}	S_{Cm} or S_{Nm}	V_{Cm}	N_{ss}
CB101	2.2'.4.5.5'-pentachlorobiphenyl	µg/kg	nn	0	0.05	0.05	0.1	
CB118	2.3'.4.4'.5'-pentachlorobiphenyl	µg/kg	nn	0	0.05	0.05	0.1	
CB138	2.2'.3.4.4'.5'-hexachlorobiphenyl	µg/kg	nn	0	0.05	0.05	0.1	
CB153	2.2'.4.4'.5.5'-hexachlorobiphenyl	µg/kg	nn	0	0.05	0.05	0.1	
CB180	2.2'.3.4.4'.5.5'-heptachlorobiphenyl	µg/kg	nn	0	0.05	0.05	0.1	
CB28	2.4.4'-trichlorobiphenyl	µg/kg	nn	0	0.05	0.05	0.1	
CB52	2.2'.5.5'-tetrachlorobiphenyl	µg/kg	nn	0	0.05	0.05	0.1	
SCB7	sum of CBs.-Sum of (CB28. CB52. CB101.	µg/kg	nn	0	0.05	0.2	0.1	
DIELD	dieldrin	µg/kg	nn	0	0.05	0.05	0.1	
DDEPP	DDE (p.p')	µg/kg	nn	0	0.05	0.05	0.1	
HCHG	gamma-HCH (gamma-	µg/kg	nn	0	0.05	0.05	0.1	
CORG	organic carbon	%	nn	0	0.02	0.02	0.05	2.5
AL	aluminium	g/kg	Ps	4	4	0.5	0.03	50
LI	lithium	mg/kg	Ps	4	5	2	0.05	52
AS	arsenic	mg/kg	Ps	3	1.5	0.5	0.06	
CD	cadmium	mg/kg	Ps	0.03	0.06	0.02	0.08	
CR	chromium	mg/kg	Ps	13	6	3	0.05	
CU	copper	mg/kg	Ps	1	1	0.5	0.06	
HG	mercury	mg/kg	Ps	0	0.04	0.02	0.05	
NI	nickel	mg/kg	Ps	2.5	1.1	3	0.08	
PB	lead	mg/kg	Ps	2	2.2	3	0.08	
ZN	zinc	mg/kg	Ps	8	9	3	0.03	
HCB	hexachlorobenzene	µg/kg	nn	0	0.05	0.05	0.1	
TBTIN	tributyltin (TBT)	µg/kg	nn	0	0.05	0.05	0.1	
ANT	anthracene	µg/kg	nn	0	5	3	0.1	
BAA	benzo[a]anthracene	µg/kg	nn	0	5	3	0.1	
BAP	benzo[a]pyrene	µg/kg	nn	0	5	3	0.1	
BGHIP	benzo[ghi]perylene	µg/kg	nn	0	5	3	0.1	

Table 4.1 Sediment assessment – summary information for pivot values, where the pivot values (N_x) used for different determinands measured by different extraction/ digestion methods. The table also lists the values of the constant error (S_{Cm}) and variable error (V_{Cm}) components of the variance of each determinand/method combination, and the concentrations (N_{ss}) of cofactors to which the contaminant concentrations are normalised (From Annexes 8 and 9 of the ICES WGMS2002 report). Note that S_{Nm} = S_{Nmco}

Parameter	Parameter Description	Unit	Digestion*	C _x or N _x	S _{Cx} or S _{Nx}	S _{Cm} or S _{Nm}	V _{Cm}	N _{ss}
CHR	chrysene	µg/kg	nn	0	5	3	0.1	
FLU	fluoranthene	µg/kg	nn	0	5	3	0.1	
ICDP	indeno[1.2.3-cd]pyrene	µg/kg	nn	0	5	3	0.1	
NAP	naphthalene	µg/kg	nn	0	3	3	0.1	
PA	phenanthrene	µg/kg	nn	0	5	3	0.1	
PYR	pyrene	µg/kg	nn	0	3	3	0.1	
TRI	triphenylene	µg/kg	nn	0	3	3	0.1	
MF63	Fraction <63 µm	%	nn	0	0.5	0.5	0.03	115
MF20	Fraction <20 µm	%	nn	0	0.5	0.5	0.03	85
MF16min	Fraction <16 µm after mineralisation	%	nn	0	0.5	0.5	0.03	55
AL	aluminium	g/kg	Tot	14	6	0.5	0.03	58
AS	arsenic	mg/kg	Tot	5	3	0.5	0.06	
CD	cadmium	mg/kg	Tot	0.03	0.06	0.02	0.08	
CR	chromium	mg/kg	Tot	13	6	3	0.05	
CU	copper	mg/kg	Tot	3	1	0.5	0.06	
HG	mercury	mg/kg	Tot	0	0.04	0.02	0.05	
LI	lithium	mg/kg	Tot	7	5	2	0.05	52
NI	nickel	mg/kg	Tot	4	2	3	0.08	
PB	lead	mg/kg	Tot	9	3	3	0.08	
ZN	zinc	mg/kg	Tot	13	5	3	0.03	
AL	aluminium	g/kg	Pw	3	2	0.5	0.03	40
AS	arsenic	mg/kg	Pw	1.5	1.5	0.5	0.06	
CD	cadmium	mg/kg	Pw	0.03	0.06	0.02	0.08	
CR	chromium	mg/kg	Pw	10	6	3	0.05	
CU	copper	mg/kg	Pw	1	1	0.5	0.06	
HG	mercury	mg/kg	Pw	0	0.04	0.02	0.05	
LI	lithium	mg/kg	Pw	3	2	2	0.05	40
NI	nickel	mg/kg	Pw	2.5	1.1	3	0.08	
PB	lead	mg/kg	Pw	2	2.2	3	0.08	

Table 4.1 Sediment assessment – summary information for pivot values, where the pivot values (N_x) used for different determinands measured by different extraction/ digestion methods. The table also lists the values of the constant error (S_{Cm}) and variable error (V_{Cm}) components of the variance of each determinand/method combination, and the concentrations (N_{ss}) of cofactors to which the contaminant concentrations are normalised (From Annexes 8 and 9 of the ICES WGMS2002 report). Note that $S_{Nm} = S_{Nmco}$

Parameter	Parameter Description	Unit	Digestion*	C_x or N_x	S_{C_x} or S_{N_x}	S_{C_m} or S_{N_m}	V_{C_m}	N_{ss}
ZN	zinc	mg/kg	Pw	8	9	3	0.03	
AL	aluminium	g/kg	nn	4	4	0.5	0.03	50
AS	arsenic	mg/kg	nn	3	1.5	0.5	0.06	
CD	cadmium	mg/kg	nn	0.03	0.06	0.02	0.08	
CR	chromium	mg/kg	nn	13	6	3	0.05	
CU	copper	mg/kg	nn	1	1	0.5	0.06	
HG	mercury	mg/kg	nn	0	0.04	0.02	0.05	
LI	lithium	mg/kg	nn	4	5	2	0.05	52
NI	nickel	mg/kg	nn	2.5	1.1	3	0.08	
PB	lead	mg/kg	nn	2	2.2	3	0.08	
ZN	zinc	mg/kg	nn	8	9	3	0.03	
HG	mercury	mg/kg	Pe	0	0.04	0.02	0.05	
NONE	no cofactor for normalization	g/g	nn	0	0	0	0	1
SPAH3r	Sum of 3r PAH. 2 compounds	µg/kg	nn	0	6	6	0.1	
SPAH4r3	Sum of 4r PAH. 3 compounds	µg/kg	nn	0	9	9	0.1	
SPAH4r4	Sum of 4r PAH. 4 compounds	µg/kg	nn	0	12	12	0.1	
SPAH6r	Sum of 6 PAH. 2 compounds	µg/kg	nn	0	6	6	0.1	

* Digestion codes: nn (not required); Pw (Partial weak), Ps (Partial strong) Pe (Partial mercury only), Tot (Total)

6. Trend analysis

6.1 Statistical analyses prior to trend analysis

For the use in an assessment, it is important that the data are of acceptable quality. Several types of quantifiable QA information can be available to the data assessors:

Results of reference materials reported to the ICES database.

Z-scores obtained in the QUASIMEME laboratory performance studies and reported to the ICES database

Information on the whole suite of Z-scores collected by QUASIMEME on CDs and supplied to ICES via the participating laboratories

In early assessments, data screening for analytical quality resulted in two classes of data: 'acceptable' and 'unacceptable'. Only data with 'acceptable' QA were taken forward and assessed for trends using smoothers, with each observation being given equal statistical weight. However, many data were rejected as 'unacceptable' and this led to the shortening or loss of many time series. Nicholson et al. (2001) argued that the QA acceptance criteria were too stringent and that some data, previously rejected as unacceptable, could be used in future assessments if they were appropriately down-weighted in the statistical analysis.

The main purpose of statistical analysis prior to trend analysis is to develop appropriate weightings to be applied to individual data points in the subsequent weighted LOESS smoother and trend analysis. The procedures used for biota differ from those applied to sediment. In both cases, use is made of a range of quality assurance information, such as that obtained from Laboratory Performance Studies (QUASIMEME) and internal quality control procedures, such as the analysis of reference materials. The objective for field data for biota and sediment is to obtain estimates of the uncertainty in each data point.

6.2 Uncertainty in biota analysis

6.2.1 Treatment of data of varying analytical quality

Since the 2005 CEMP data assessment, data are given analytical weights according to the available QA. An iterative procedure (Fryer, 2004) is then used to convert these analytical weights into statistical weights that account for the relative magnitudes of the environmental and analytical variances. A weighted loess smoother (Nicholson & Fryer, 2001; Uhlig, 2001) is then fitted to the median log-concentrations, leading to the same tests of significance as before.

Three types of quantitative QA information are potentially available for each data point:

- QUASIMEME z-scores, as supplied to ICES by analytical laboratories, through the collations of laboratory performance study results distributed to participating laboratories on CD by QUASIMEME in summer 2004, and provided by laboratories for subsequent years
- QUASIMEME z-scores held in the ICES 2.2 database – these are taken to be acceptable if between -2 and +2
- CRM values – these are taken to be acceptable if the CRM lab concentration is within 25% of the CRM true concentration

Inspection of the QUASIMEME data already held in the ICES database strongly suggested that they were not very reliable, because a) they were incomplete and b) it was not clear how the multiple z-scores obtained each year for each determinand-matrix combination were reported to ICES. Procedures were therefore developed to summarise the QUASIMEME CD data to give a single annual expression of performance for each determinand-matrix combination. Specifically, the individual z-scores in each year for each determinand-matrix were squared, summed and compared to the critical value of a χ^2 distribution. The data were considered to indicate acceptable performance (pass) if:

$$\sum_{i=1}^n Z_i^2 < \chi_n^2(0.95)$$

where n is the number of z-scores and $\chi_n^2(0.95)$ is the upper 95 percentile of a χ^2 distribution on n degrees of freedom. In practice, the critical value was computed as $3.84 \times n^{0.66}$.

Analytical weights were then assigned to the annual contaminant indices according to the Table below:

Table 6.1. Assignment of analytical weights to annual contaminant indices

QUASIMEME 2004 CD and for subsequent years	QUASIMEME data held in ICES 2.2 database	CRM performance	Analytical weight
Pass	Pass	Pass	1.0
Pass	Pass	Fail or absent	1.0
Pass	Fail or absent	Pass	1.0
Pass	Fail or absent	Fail or absent	1.0
Fail or absent	Pass	Pass	1.0
Fail or absent	Pass	Fail or absent	0.7
Fail or absent	Fail or absent	Pass	0.7
Fail or absent	Fail or absent	Fail or absent	0.2

6.2.2 Persistent bias

A problem with many trend analysis procedures is that the conclusions can be affected by the presence of persistent bias in the data. Both detection of, and correction for, persistent bias are technically difficult. To partially address this problem, the OSPAR MON intersessional group, MIG, considered that step changes in concentrations reported in field data coinciding with step changes in QA performance might indicate some forms of persistent bias. Visual examination of some time series suggested shifts in field concentrations had occurred when QA practices had changed. More formally, a process was developed to compare concentrations in field samples in periods before and after step changes in QA performance, as indicated by step changes in the analytical weights applied to data points. The summary tables provided to assessors flagged those time series in which such changes may have occurred. Time series with less than three years of data with good QA (an analytical weight of unity) were also flagged with a view to them being excluded from the assessment.

6.3 Uncertainties in sediment analysis

*Developed from ICES CM 2005/Z:04, "Pragmatic estimation of uncertainty in normalized concentrations of contaminants in sediments", F Smedes, *I M Davies, R J Fryer*

6.3.1 Introduction

The OSPAR Working Group on Monitoring (MON) met at ICES HQ in December 2004 to undertake a temporal trend assessment of contaminant concentrations in sediment and biota, using data collected under the OSPAR Co-ordinated Environmental Monitoring Programme (CEMP) held in the ICES Environmental Database. Whilst temporal trends in biota had been periodically assessed over the preceding 14 years, this was the first attempt to assess temporal trends in sediment.

Much time was spent in the two years before the assessment establishing the principles behind the assessment and the statistical approaches to be used. The preparation was carried out by OSPAR MON and its intersessional group MIG, building on a series of tasks carried out by the ICES Working Groups on Statistical Aspects of Environmental Monitoring, Marine Sediments in Relation to Pollution, and Marine Chemistry. Dr Mike Nicholson was closely involved in the fundamentals of this work, and his influence will be felt for many years to come.

An important issue that had arisen in previous assessments concerned the quality of the data held in the ICES database. The assessors recognised that poor quality data could obscure real trends, or could suggest trends erroneously. To minimise this concern, and to provide clear evidence of analytical quality, ICES has required, since about 1990, that field data submitted to the database must be accompanied by supporting analytical quality control (AQC) information. This can take various forms, including analyses of certified reference materials (CRMs) and results from intercalibration exercises and laboratory performance studies, such as those run by QUASIMEME.

Previous assessment groups used this information to establish minimum performance standards for each combination of contaminant and matrix. Field data that were not supported by AQC data of sufficient quality were eliminated from assessments. Whilst this positively selected data of known good quality, it eliminated other data which might have been of equally good quality (e.g. field data submitted before the submission of AQC data became mandatory), thus reducing the length of some time series and the geographical coverage of the assessment.

Nicholson et al (2001) analysed the sources of variance in temporal trend data (for biota) held at ICES and found that the analytical variance was often a relatively small component of the overall variance. They concluded that the AQC criteria should be relaxed or that the AQC information should be used in a different way. Based on this analysis, OSPAR MON undertook, as a principle of the 2004 assessment, to use an alternative form of data analysis that did not eliminate data. The approach was to weight each data point according to its analytical quality, using the AQC information held by (or absent from) ICES. A weighted smoother was then fitted to the data and tests of significance were applied to the fitted trend lines.

This paper has two purposes. First, it describes how, in the assessment of CEMP data, MON 2004 used the AQC information held by ICES to estimate the uncertainties in normalised concentrations of contaminants in sediment, and how these uncertainties were then used as analytical weights in the trend assessment. Second, it describes developments in the methodology that will be considered for use by MON 2005 in a further assessment of CEMP data.

6.3.2 Procedures used by MON in assessing CEMP data

6.3.2.1 UNCERTAINTY IN NORMALISED CONCENTRATIONS

Contaminant concentrations in sediment were normalised to account for changes in the bulk physical composition of the sediment, such as changes in particle size distribution or organic carbon content. Normalisation requires pivot values, estimates of the concentrations of contaminants and normalisers in pure sand (Table 6.2). Normalised concentrations were given by:

$$C_{SS} = (C_M - C_X) \frac{N_{SS} - N_X}{N_M - N_X} + C_X \quad (1)$$

where:

- C_{SS} Normalised concentration of the contaminant
- C_M Measured concentration of the contaminant
- C_X Pivot value for the contaminant
- N_{SS} Reference concentration of the normaliser
- N_M Measured concentration of the normaliser
- N_X Pivot value for the normaliser

The procedures for normalisation are described in OSPAR (2002). The derivation of pivot values and their uncertainties for contaminants and normalisers are described in WGMS (2002). Some normalised concentrations were environmentally inadmissible (i.e. negative) and were therefore excluded from the assessment.

Variance/uncertainty of the normalised concentrations, is related to the uncertainties in the contaminant and normaliser analyses and the natural variabilities in the pivot values. Estimation of this uncertainty can be approached through an approximation based on the propagation of errors through the formulae:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

$$\frac{\text{Var}(XY)}{(XY)^2} \approx \frac{\text{Var}(X)}{X^2} + \frac{\text{Var}(Y)}{Y^2}$$

$$\frac{\text{Var}(X/Y)}{(X/Y)^2} \approx \frac{\text{Var}(X)}{X^2} + \frac{\text{Var}(Y)}{Y^2}$$

where X and Y are any positive random variables (WGMS, 2002). The approximation can be improved by noting that when X and Y are correlated:

Table 6.2 Sediment assessment – summary information for pivot values, where the pivot values (N_x) used for different determinands measured by different extraction/ digestion methods. The table also lists the values of the constant error (s_{Cm}) and variable error (V_{Cm}) components of the variance of each determinand/method combination, and the concentrations (N_{ss}) of cofactors to which the contaminant concentrations are normalised (From Annexes 8 and 9 of the ICES WGMS2002 report). Note that $S_{Nm} = S_{Nmcof}$

Par	ParDescription	Unit	Digestion*	C_x or N_x	S_{C_x} or S_{N_x}	s_{Cm} or s_{Nm}	V_{Cm}	N_{ss}
CB101	2,2',4,5,5'-pentachlorobiphenyl	µg/kg	nn	0	0,05	0,05	0,1	
CB118	2,3',4,4',5-pentachlorobiphenyl	µg/kg	nn	0	0,05	0,05	0,1	
CB138	2,2',3,4,4',5'-hexachlorobiphenyl	µg/kg	nn	0	0,05	0,05	0,1	
CB153	2,2',4,4',5,5'-hexachlorobiphenyl	µg/kg	nn	0	0,05	0,05	0,1	
CB180	2,2',3,4,4',5,5'-heptachlorobiphenyl	µg/kg	nn	0	0,05	0,05	0,1	
CB28	2,4,4'-trichlorobiphenyl	µg/kg	nn	0	0,05	0,05	0,1	
CB52	2,2',5,5'-tetrachlorobiphenyl	µg/kg	nn	0	0,05	0,05	0,1	
SCB7	sum of CBs,-Sum of (CB28, CB52, CB101,	µg/kg	nn	0	0,05	0,2	0,1	
DIELD	dieldrin	µg/kg	nn	0	0,05	0,05	0,1	
DDEPP	DDE (p,p')	µg/kg	nn	0	0,05	0,05	0,1	
HCHG	gamma-HCH (gamma-hexachlorocyclohexane)	µg/kg	nn	0	0,05	0,05	0,1	
CORG	organic carbon	%	nn	0	0,02	0,02	0,05	2,5
AL	aluminium	g/kg	Ps	4	4	0,5	0,03	50
LI	lithium	mg/kg	Ps	4	5	2	0,05	52
AS	arsenic	mg/kg	Ps	3	1,5	0,5	0,06	
CD	cadmium	mg/kg	Ps	0,03	0,06	0,02	0,08	
CR	chromium	mg/kg	Ps	13	6	3	0,05	
CU	copper	mg/kg	Ps	1	1	0,5	0,06	
HG	mercury	mg/kg	Ps	0	0,04	0,02	0,05	
NI	nickel	mg/kg	Ps	2,5	1,1	3	0,08	
PB	lead	mg/kg	Ps	2	2,2	3	0,08	
ZN	zinc	mg/kg	Ps	8	9	3	0,03	
HCB	hexachlorobenzene	µg/kg	nn	0	0,05	0,05	0,1	
TBTIN	tributyltin (TBT)	µg/kg	nn	0	0,05	0,05	0,1	
ANT	anthracene	µg/kg	nn	0	5	3	0,1	
BAA	benzo[a]anthracene	µg/kg	nn	0	5	3	0,1	
BAP	benzo[a]pyrene	µg/kg	nn	0	5	3	0,1	

Table 6.2 Sediment assessment – summary information for pivot values, where the pivot values (N_x) used for different determinands measured by different extraction/ digestion methods. The table also lists the values of the constant error (s_{Cm}) and variable error (V_{Cm}) components of the variance of each determinand/method combination, and the concentrations (N_{ss}) of cofactors to which the contaminant concentrations are normalised (From Annexes 8 and 9 of the ICES WGMS2002 report). Note that $S_{Nm} = S_{Nmcof}$

Par	ParDescription	Unit	Digestion*	C_x or N_x	S_{C_x} or S_{N_x}	s_{Cm} or s_{Nm}	V_{Cm}	N_{ss}
BGHIP	benzo[ghi]perylene	µg/kg	nn	0	5	3	0,1	
CHR	chrysene	µg/kg	nn	0	5	3	0,1	
FLU	fluoranthene	µg/kg	nn	0	5	3	0,1	
ICDP	indeno[1,2,3-cd]pyrene	µg/kg	nn	0	5	3	0,1	
NAP	naphthalene	µg/kg	nn	0	3	3	0,1	
PA	phenanthrene	µg/kg	nn	0	5	3	0,1	
PYR	pyrene	µg/kg	nn	0	3	3	0,1	
TRI	triphenylene	µg/kg	nn	0	3	3	0,1	
MF63	Fraction <63 µm	%	nn	0	0,5	0,5	0,03	115
MF20	Fraction <20 µm	%	nn	0	0,5	0,5	0,03	85
MF16min	Fraction <16 µm after mineralisation	%	nn	0	0,5	0,5	0,03	55
AL	aluminium	g/kg	Tot	14	6	0,5	0,03	58
AS	arsenic	mg/kg	Tot	5	3	0,5	0,06	
CD	cadmium	mg/kg	Tot	0,03	0,06	0,02	0,08	
CR	chromium	mg/kg	Tot	13	6	3	0,05	
CU	copper	mg/kg	Tot	3	1	0,5	0,06	
HG	mercury	mg/kg	Tot	0	0,04	0,02	0,05	
LI	lithium	mg/kg	Tot	7	5	2	0,05	52
NI	nickel	mg/kg	Tot	4	2	3	0,08	
PB	lead	mg/kg	Tot	9	3	3	0,08	
ZN	zinc	mg/kg	Tot	13	5	3	0,03	
AL	aluminium	g/kg	Pw	3	2	0,5	0,03	40
AS	arsenic	mg/kg	Pw	1,5	1,5	0,5	0,06	
CD	cadmium	mg/kg	Pw	0,03	0,06	0,02	0,08	
CR	chromium	mg/kg	Pw	10	6	3	0,05	
CU	copper	mg/kg	Pw	1	1	0,5	0,06	
HG	mercury	mg/kg	Pw	0	0,04	0,02	0,05	
LI	lithium	mg/kg	Pw	3	2	2	0,05	40

Table 6.2 Sediment assessment – summary information for pivot values, where the pivot values (N_x) used for different determinands measured by different extraction/ digestion methods. The table also lists the values of the constant error (s_{Cm}) and variable error (V_{Cm}) components of the variance of each determinand/method combination, and the concentrations (N_{ss}) of cofactors to which the contaminant concentrations are normalised (From Annexes 8 and 9 of the ICES WGMS2002 report). Note that $S_{Nm} = S_{Nmcof}$

Par	ParDescription	Unit	Digestion*	C_x or N_x	S_{C_x} or S_{N_x}	s_{Cm} or s_{Nm}	V_{Cm}	N_{ss}
NI	nickel	mg/kg	Pw	2,5	1,1	3	0,08	
PB	lead	mg/kg	Pw	2	2,2	3	0,08	
ZN	zinc	mg/kg	Pw	8	9	3	0,03	
AL	aluminium	g/kg	nn	4	4	0,5	0,03	50
AS	arsenic	mg/kg	nn	3	1,5	0,5	0,06	
CD	cadmium	mg/kg	nn	0,03	0,06	0,02	0,08	
CR	chromium	mg/kg	nn	13	6	3	0,05	
CU	copper	mg/kg	nn	1	1	0,5	0,06	
HG	mercury	mg/kg	nn	0	0,04	0,02	0,05	
LI	lithium	mg/kg	nn	4	5	2	0,05	52
NI	nickel	mg/kg	nn	2,5	1,1	3	0,08	
PB	lead	mg/kg	nn	2	2,2	3	0,08	
ZN	zinc	mg/kg	nn	8	9	3	0,03	
HG	mercury	mg/kg	Pe	0	0,04	0,02	0,05	
NONE	no cofactor for normalization	g/g	nn	0	0	0	0	1
SPA3r	Sum of 3r PAH, 2 compounds	µg/kg	nn	0	6	6	0,1	
SPA4r3	Sum of 4r PAH, 3 compounds	µg/kg	nn	0	9	9	0,1	
SPA4r4	Sum of 4r PAH, 4 compounds	µg/kg	nn	0	12	12	0,1	
SPA6r	Sum of 6 PAH, 2 compounds	µg/kg	nn	0	6	6	0,1	

* Digestion codes: nn (not required); Pw (Partial weak), Ps (Partial strong) Pe (Partial mercury only), Tot (Total)

$$\begin{aligned}\text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) \\ \frac{\text{Var}(XY)}{(XY)^2} &\approx \frac{\text{Var}(X)}{X^2} + \frac{\text{Var}(Y)}{Y^2} + \frac{2\text{Cov}(X, Y)}{XY} \\ \frac{\text{Var}(X/Y)}{(X/Y)^2} &\approx \frac{\text{Var}(X)}{X^2} + \frac{\text{Var}(Y)}{Y^2} - \frac{2\text{Cov}(X, Y)}{XY}\end{aligned}$$

where Cov denotes covariance. This leads to an estimate of the uncertainty in normalised concentrations:

$$\begin{aligned}\text{Var}(C_{SS}) &= \left(\frac{N_{SS} - N_X}{N_M - N_X} \right)^2 \left\{ \text{Var}(C_M) + \text{Var}(N_M) \left(\frac{C_M - C_X}{N_M - N_X} \right)^2 \right\} \\ &+ \left(\frac{N_{SS} - N_M}{N_M - N_X} \right)^2 \left\{ \text{Var}(C_X) + \text{Var}(N_X) \left(\frac{C_M - C_X}{N_M - N_X} \right)^2 \right\}\end{aligned}\quad (2)$$

This derivation assumes that N_X , C_X , N_M , C_M are independent. The uncertainties in the pivot values contribute little to the uncertainty in the normalised concentration when N_M is close to N_{SS} .

The natural variabilities of the pivot values are given in Table 6.2. These were used for all data, including those normalisers (e.g. organic carbon) for which the pivot value is essentially zero. The uncertainties in the contaminant and normaliser analyses were modelled as a combination of a fixed component and a variable component

$$\begin{aligned}\text{Var}(C_M) &= s_{C_M}^2 + f_{C_M}^2 v_{C_M}^2 C_M^2 \\ \text{Var}(N_M) &= s_{N_M}^2 + f_{N_M}^2 v_{N_M}^2 N_M^2\end{aligned}$$

where

s_{C_M} , s_{N_M} are the absolute standard errors in C_M and N_M respectively; the fixed error component

v_{C_M} , v_{N_M} are the coefficients of variation in C_M and N_M respectively; the relative error component

f_{C_M} , f_{N_M} are inflation factors based on AQC information held at ICES for C_M and N_M respectively.

The fixed error components are given in Table 6.2 and were applied to all data unless the detection limit of the analysis was available, when they were set to one third of the detection limit. The relative error components are also given in Table 6.2 and represent state of the art (i.e. good) analytical performance. The inflation factors were based on the AQC information held by ICES and increased the relative error component if analytical performance appeared poor (see below). Some normalised concentrations had a relative uncertainty of more than 100% and were excluded from the assessment.

6.3.2.2 INFLATION FACTORS: INCORPORATING AQC INFORMATION IN ESTIMATES OF UNCERTAINTY

The AQC information at ICES consisted of:

- the analyses of CRMs reported to ICES and held in the ICES database
- Z-scores from QUASIMEME laboratory performance studies in the ICES database
- complete information on QUASIMEME Z-scores, covering all exercises in the last 10 years, collated by QUASIMEME onto CDs and supplied to ICES, by the participating laboratories, in 2004.

The AQC data in the ICES database were very heterogeneous. For example, many CRMs had been used and the QUASIMEME Z-scores were incomplete and unreliable. Reasons for this included the difficulty of accommodating the results from more than one QUASIMEME Round per year in the reporting format, and inconsistencies between the Z-scores in the database and those on the CD provided by QUASIMEME. There were also clear examples of gross reporting errors affecting both the QUASIMEME and CRM data. The AQC data were therefore “cleansed” before the assessment (Appendix 1).

The CRM data were converted to Z-scores, assuming a coefficient of variation of 12.5%, by:

$$Z = \frac{M - A}{0.125A}$$

where M is the measured concentration and A is the certified reference value. Thus, each measured concentration C_M or normaliser N_M potentially had several Z-scores which could be used to judge its analytical quality: one from the CRM analysis, one from QUASIMEME held on the ICES database, and perhaps several from the QUASIMEME CD. The inflation factor for C_M was calculated as

$$f_{C_M}^2 = \frac{\sum Z^2}{n}$$

where the summation runs over all the Z-scores for the contaminant and n is the number of Z-scores. A similar calculation was used for the inflation factor for N_M . A consequence of this procedure was that, when many Z-scores were available from the QUASIMEME CDs, they tended to dominate the estimation of the inflation factors.

In practice, it was found that large Z-scores, possibly affected by e.g. misreporting of units, could unduly influence the estimation of the inflation factors. To overcome this, high Z-scores were truncated at a value of 3, leading to a maximum inflation factor of 3. The inflation factors were also truncated below at 1, since this corresponds to state of the art analytical performance. An inflation factor of 3 was used for data with no supporting AQC information.

6.3.2.3 TEMPORAL TREND ASSESSMENT

For each time series, the normalised concentrations and their uncertainties were used to construct annual contaminant indices, which were then assessed for trends.

The annual contaminant index was taken to be the weighted average of the log normalised concentrations, where the weights were a suitable combination of the analytical variation (i.e. the uncertainties computed above) and, where possible, an estimate of the within-year environmental (field) variation. Specifically, let c_{yi} be the normalised concentration of sample i in year y and let σ_{yi} be the analytical standard deviation associated with this concentration (i.e. the uncertainty calculated above). The analytical standard deviation of the normalised log-concentration $\log(c_{yi})$ is then approximately σ_{yi}/c_{yi} . Let τ be the within-year environmental standard deviation (assumed known for now). The joint analytical and within-year field variance of $\log(c_{yi})$ is then

$$v_{yi} = \tau^2 + \sigma_{yi}^2 / c_{yi}^2,$$

and the annual contaminant index z_y is taken to be the weighted average of the $\log(c_{yi})$:

$$z_y = \frac{\sum_i w_{yi} \log(c_{yi})}{\sum_i w_{yi}}$$

where $w_{yi} = 1 / v_{yi}$. The joint analytical and within-year environmental variance of the annual contaminant index z_y is

$$1 / \sum_i w_{yi}.$$

When there were multiple samples in at least one year of the time series, the within-year environmental standard deviation τ was estimated by restricted maximum likelihood with the analytical standard deviations assumed known and equal to σ_{yi} . When there was only one sample each year, τ was taken to be zero, and the variance of the annual contaminant index is only a measure of the analytical variability.

To assess the annual contaminant indices for trends, an iterative procedure (Fryer, 2004) was first used to construct statistical weights that were a suitable balance of the variance of the annual indices and any additional environmental variation (e.g. between-year environmental variation and, if only one sample per year, the within-year environmental variation). A weighted loess smoother (Nicholson & Fryer, 2001; Uhlig, 2001) was then fitted to the data and tests of significance were applied to the fitted trend lines.

The procedures were used to assess 3506 time series of contaminant concentrations in sediment. Some time series were assessed several times using different normalisers. The time series varied in length within

the period 1978-2003. The monitoring sites ranged from Spain to the North of Norway and Iceland, but most were within Region II (the Greater North Sea) of the OSPAR Convention area. Standard outputs included:

- Graphical presentations of time series, with the fitted smoothers and confidence limits around the fitted smoothers
- Statistical significance of linear and non-linear trends in each time series
- Comparisons of the upper confidence limits of the final fitted values of the time series with Background Assessment Concentrations (BACs), and Environmental Assessment Concentrations (EACs) (Nicholson and Fryer, 2003; OSPAR, 2004)
- Statistical assessment of the power of the monitoring data to detect changes in concentrations, expressed as the minimum detectable annual percentage change (Nicholson et al, 1997).

The full report of the assessment is available on the OSPAR website at www.ospar.org.

6.3.3 Conclusion

At the time that this report was written, it was unclear what impact using equation (3) to estimate uncertainty would have on trend assessment. The uncertainties would be reduced, but if all they reduce in a similar way then the statistical weights that are finally applied to each data point in the loess smoother might not change much, leading to a trend line and tests of significance that are similar to before. However, the partitioning of the variation in the data into analytical and environmental components would differ, with implications for monitoring design. The revised estimation of errors was used in the subsequent 2005 and 2006 MON assessment of CEMP data.

7. Method used for trend analysis of time series

7.1 Introduction

At the last OSPAR assessment of temporal trends in contaminants in biota (OSPAR, 1999), data screening for analytical quality resulted in two classes of data: 'acceptable' and 'unacceptable'. Only data with 'acceptable' QA were used and were assessed for trends using smoothers with each observation given equal statistical weight. However, many data were rejected as 'unacceptable' and this led to the shortening or loss of many time series. It has since been argued (Nicholson et al, 2001) that the QA acceptance criteria were too stringent and that some data, previously rejected as unacceptable, could be used in future assessments if they were appropriately down-weighted in the statistical analysis.

Fitting a weighted smoother is straightforward if the statistical weights are known beforehand (e.g. Hastie & Tibshirani, 1990; Nicholson & Fryer, 2001; Uhlig, 2001). The statistical weights should be inversely related to the total environmental and analytical variance each year. Appropriate methods for estimating them will depend on the QA information available. For example, Nicholson & Fryer (2001) show how the EM algorithm can be used to estimate the environmental variance and hence the statistical weights when the analytical variance is known each year (e.g. from control chart information). However, most time series in the ICES databank do not have a complete record of analytical variances over time. Nicholson & Fryer (2002) extend the methodology to estimate missing analytical variances. However, they recognise that, for routine trend assessments, the method is too complex and makes too many assumptions that are difficult to substantiate.

An alternative (Nicholson & Fryer, 2002) is to use the available QA information to categorise the analytical quality of data as *Good*, *Poor*, *Unknown* and *Unacceptable* and allocate statistical weights $1 > W_{poor} > W_{unknown} > 0$ accordingly. This approach is simple and intuitively appealing. However, the choice of statistical weights is arbitrary and takes no account of the relative importance of the analytical variance to the total environmental and analytical variance. For example, all data with 'poor' analytical quality will have the same statistical weight, even though such data should be down-weighted less when the environmental variance dominates the analytical variance (when poor analytical quality doesn't matter so much).

This document presents a third approach that provides a compromise between the two methods described above. It is assumed that available QA information can be used to construct an *analytical weight* for each datum, ranging from 0 (totally unacceptable) through to 1 (totally acceptable). An iterative procedure is then used to convert these analytical weights into statistical weights that account for the relative magnitudes of the environmental and analytical variances. The approach should be easy to apply routinely to data in the ICES databank.

7.2 Theory

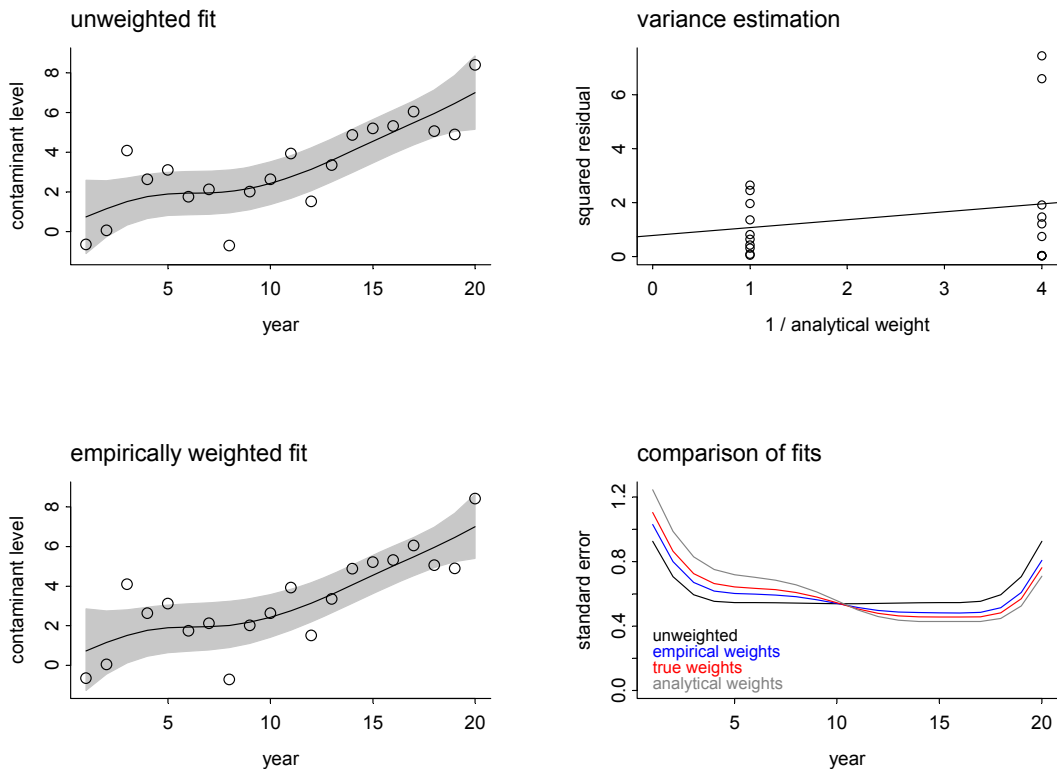
Assume that a contaminant time series can be described by the model:

$$y_t = f(t) + \varepsilon_t$$

where y_t is the annual contaminant index in year t , $f(t)$ is a smooth function of time describing the underlying trend in contaminant levels, and ε_t is the 'noise' in year t from both environmental (i.e. field) and analytical variation. Further, assume that the noise can be decomposed into two terms:

$$\varepsilon_t = \delta_t + \tau_t$$

where δ_t is the noise due to environmental variation (both between- and within-years) and τ_t is the noise due to analytical variation. Finally, assume that the noise terms are mutually independent and normally distributed:



References

- Hastie TJ, Tibshirani RJ, 1990. Generalized additive models. Chapman and Hall, London.
- Nicholson MD, Fryer RJ, 2001. Weighting procedures for assessing trend data of variable quality. ICES Working Group on Statistical Aspects of Environmental Monitoring 2001.
- Nicholson MD, Fryer RJ, Law R, Davies IM, 2001. Criterion for screening data for analytical accuracy in contaminant trend studies. ICES Marine Chemistry Working Group 2001.
- Nicholson MD, Fryer RJ, 2002. Weighted smoothers for assessing trend data of variable analytical quality. ICES Working Group on Statistical Aspects of Environmental Monitoring 2002.
- OSPAR, 1999. Report of an assessment of trends in the concentrations of certain mussels, PAHs and other organic compounds in the tissues of various fish species and blue mussels: OSPAR Ad Hoc Working Group on Monitoring 1998.

Uhlig S, 2001. The LOESS smoother: incorporation of uncertainty data and the behaviour with missing values. ICES Working Group on Statistical Aspects of Environmental Monitoring 2001.

8. Presentation of temporal trend assessments

In the initial output from the trend analysis software, the assessments of each parameter are summarised in two pages, one each for sediment and biota, preceded by an explanatory text. Each page contains regionalized tabulations and graphical representations of detected trends.

Firstly, contaminant data are tabulated by OSPAR region and, for biota, by species group. The number of time series, the number of significant linear trends and the number of time series where the mean concentration in the final year is significantly below the BAC (or BRC) are graphically presented.

Secondly, mean concentrations in the final year of each time series are tabulated, by OSPAR region. Time series for which the mean concentration is significantly below the BAC (BRC) are shown by filled circles; open circles indicate mean concentrations that are not significantly below the BAC (BRC).

Finally, selected time series are graphically illustrated. Time series can be selected where they showed a significant change in concentration over the full period of the time series, or over some subset of the data, for example for the preceding ten years or where the upper confidence limit on the mean concentration in the final year exceeded the BAC (BRC). When it is not possible to plot all time series, a selection can be made, for example to include only those which were most significant or showed the highest levels were given.

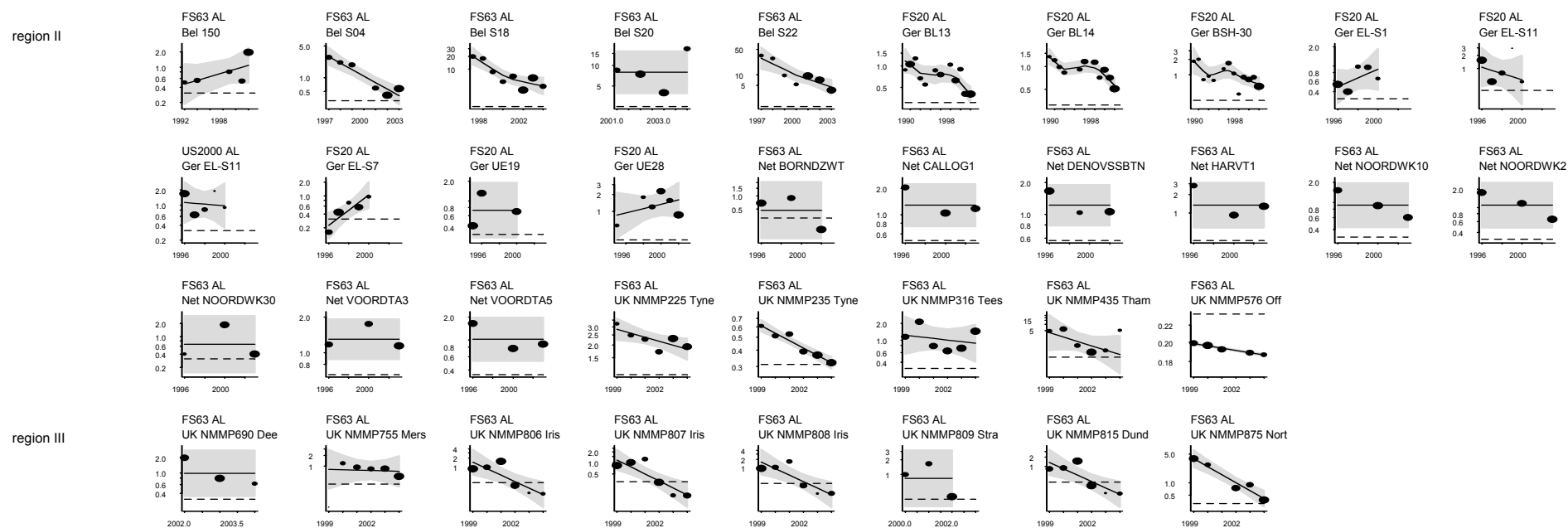
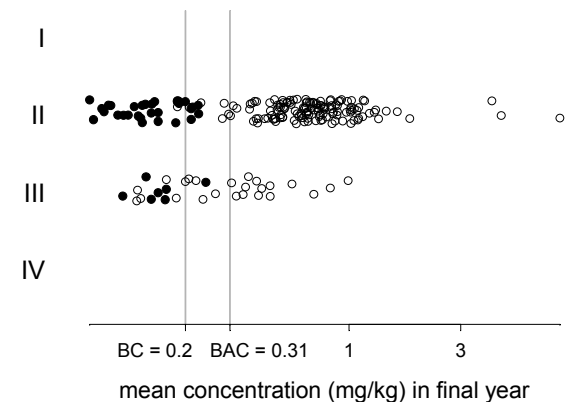
In the explanatory texts on the assessment, the following phrases have been used to explain statistical results:

- a. “*trends*” refer to linear trends, significant at the 5% level,
- b. “*mean concentrations are at background*” or “*mean concentrations are close to zero*” means that the upper confidence limit on the fitted mean concentration in the last year of monitoring is below the BAC.

An example of the approach is given overleaf

Cadmium in sediment

Normaliser	OSPAR Region	Number of time series				Trends		UCL <
		3-4	5-6	7+	total	up	down	BAC
Aluminium	II	98	48	24	170	1	13	31
	III	13	18	0	31	0	5	7
		111	66	24	201	1	18	38



9. Power of temporal trend programmes to detect changes in concentrations of contaminants

The between year residual variance about fitted time trends determines the statistical power of the monitoring programmes to detect changes with time. As the residual variance differs between stations, so will the power. Various approaches can be used to describe the statistical power of the monitoring to detect trends.

The approach used here considers the minimum rate of change that could be detected by annual monitoring over a period of 10 years. This expression of the capability of the programme is not dependent on the length of the data series. However, it is dependent on the assumption that the variance observed in the available time series is a reasonable reflection of the variance that would be observed over longer periods, i.e. the residual variance about fitted time trends is not strongly dependent on the length of the time series. In order to reduce the uncertainty arising from this assumption, this approach has been applied only to time series which contain 5 or more years of data.

An example of the result of these assessments for time series of contaminants in biota is given in the following histogram concerning cadmium in biota, created during the 2004 meeting of MON. The results of this analysis could be used by Contracting Parties to identify those time series with relatively high and low power to detect changes, and therefore used in the optimisation of monitoring programmes.

