ICES TECHNIQUES IN MARINE ENVIRONMENTAL SCIENCES

No. 20

Temporal trend monitoring: Robust method for analysing contaminant trend monitoring data

M.D. NICHOLSON
The Centre for Environment, Fisheries and Aquaculture Science
Lowestoft Laboratory
Lowestoft, Suffolk, NR33 0HT
United Kingdom

R.J. FRYER
Fisheries Research Services
The Marine Laboratory
PO Box 101
Victoria Road
Aberdeen, AB11 9DB
United Kingdom

J.R. LARSEN
International Council for the Exploration of the Sea
Palægade 2–4
DK-1261 Copenhagen K
Denmark

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Palægade 2–4 DK-1261 Copenhagen K Denmark

September 1998

ISSN 0903-2606

TABLE OF CONTENTS

Temporal trend monitoring:
Robust method for analysing contaminant trend monitoring data

M.D. Nicholson, R.J. Fryer, and J.R. Larsen

Nicholson, M.D., Fryer, R.J., and Larsen, J.R. 1998. Temporal trend monitoring: Robust method for analysing contaminant trend monitoring data. ICES Techniques in Marine Environmental Sciences, No. 20.

ABSTRACT

This document describes a new method of assessing contaminant trends in fish muscle, fish liver, and shellfish.

Previous methods of assessing trend data have often been complicated by the need to respond to unusual features of the data, either in the way they were collected, processed, or in their distribution. These complications are discussed in detail to show the reasoning behind the method described here, which aims to provide a simpler, robust and more complete method of analysing and presenting trends.

Essentially, the method summarizes trends using a smoother, a specific class of smooth curves fitted to median log-concentrations. The theory and methodology of fitting smoothers is new and our knowledge of the performance of the fitted smoothers, particularly with small sample sizes, is only approximate. Although a preliminary application of the new method to the 1993 assessment of the data from the Joint Monitoring Programme of the Oslo and Paris Commissions was promising, more assessment and development of the method will be necessary. A very simple smoother is used here to make the computations and theory easy to follow.

A detailed worked example is provided. Statistical theory and formulae are included as annexes.

Key words: temporal trends, data assessment, smoother

1 INTRODUCTION

This method of analysis is intended for data which have been collected according to the following sampling guidelines for biota (fish and shellfish):

Cooperative ICES Monitoring Studies Programme, Purpose (3):

> *Provision of an analysis of trends over time in pollutant concentrations in selected areas especially in relation to the assessment of the efficacy of control measures* (ICES, 1984),

and Oslo and Paris (OSPAR) Commissions' Joint Monitoring Programme (JMP), Purpose (d):

> *Assessment of the effectiveness of measures taken for the reduction of marine pollution in the framework of the Conventions* (OSPAR, 1994).

JMP Purpose (d) is being replaced in the new OSPAR Joint Assessment and Monitoring Programme (JAMP) by the following:

> *To determine temporal trends, either as a means of assessing the effectiveness of policy measures, or to assess, by the use of suitable indicators, changes and variability in the quality of the marine environment* (OSPAR, 1995).

The main characteristics of 'trend data' collected in this way are:

1) biota are collected annually at the same time within each year;

2) this time should be outside the spawning period;

3) the same size range of the target species should be sampled each year;

4) the sample size should be the same each year.

There are other characteristics which relate to the collection and treatment of individual species. For example, samples of fish should be length stratified; mussels should be homogenized into three equal bulked samples.

The implications of this sampling protocol are that between-year biological variation (e.g., mean length, condition, stock composition) is controlled, and that within-year biological variation (e.g., length) can be removed. This is the basis of the method of analysis described here, and of methods used in the past.

Previous analyses of ICES and JMP temporal trend data (ICES, 1989, 1991; OSPAR, 1992) have followed a method published in ICES (1987): for fish muscle, concentrations of mercury, lead and zinc are assumed to be related to fish length and time, and to be described by a model of the form

$$\textit{log concentration} = \mu_t + \alpha\, \textit{length} + \textit{error}$$

where the *error* term is Normally distributed with constant variance. For copper, chromium, nickel, and PCBs in fish muscle, and all contaminants in fish liver and shellfish, biological variables such as length or shell weight seemed to be unnecessary (Nicholson and Wilson, 1987; ICES, 1991) and a simpler model of the form

$$log\ concentration = \mu_t + error$$

was used.

Temporal trends are assessed by comparing the year-specific intercepts against those of two sub-models:

$\mu_t = \mu$ (the same level in each year),

and

$\mu_t = \mu + \beta t$ (a linear trend).

This relatively simple method has given reasonable results in some cases (Jensen and Cheng, 1987; Nicholson *et al.*, 1991; Rees and Nicholson, 1989). However, in other cases, more complex models (Misra *et al.*, 1990; Warren, 1993a, 1993b) or various *ad hoc* procedures were found to be necessary. Many of the results reported in the ICES/JMP assessments were qualified by comments concerning unexplained outliers and other potentially distorting characteristics of the data.

There are two groups of problems that have been encountered in the past, corresponding to the two main stages of the trend assessment. Firstly, there are problems in the *within-year* analysis. This consists of detection and treatment of outliers, identification of the error distribution, removing the effect of biological covariates, and similar procedures. This stage leads to a series of summary statistics representing the yearly contaminant levels. The problems arising here (Bignert and Nielsen, 1988; Nicholson *et al.*, 1989) are:

1) unexplained outliers;

2) results reported as 'below detection limit';

3) partial bulking, e.g., liver tissues from small fish;

4) variation in length regression coefficients between years.

The effect of problems (1) to (3) is that the summary statistic representing a particular yearly contaminant level could be distorted. Depending on its location in the time series, it will tend either to inflate the between-year variability or to induce a spurious component to an observed trend. Variation in length regression coefficients from year to year implies that trends may be length dependent and trends for small fish could be different from those for large fish.

The second group of problems arises in the stage concerned with the *between-year* analysis. This takes the series of annual contaminant levels and provides a description of the observed pattern of change, together with a statistical test of its significance. The problems that need to be dealt with (Nicholson and Fryer, 1992; Fryer and Nicholson, 1993a) are as follows:

1) the need to detect non-linear trends in mean contaminant levels; and

2) additional random between-year variation in mean contaminant levels.

These problems affect the statistical tests of the observed trend. For example, if random between-year variation is wrongly assumed to be zero, the test described in ICES (1987) will result in too many apparently significant trends where there are none.

These problems have been discussed extensively by the ICES Working Group on the Statistical Aspects of Environmental Monitoring (see, e.g., ICES, 1993). Out of these discussions, a revised method of analysis has been developed, which is presented here. The objectives of the revised analysis are that it should be:

a) simple—be easy to apply and to understand;

b) robust—require no special treatment of outliers, 'less than' values, etc.;

c) correct—provide size-dependent trends, if necessary, estimate the component of random between-year variation, and be valid across as wide a range of conditions as possible.

Section 2, below, describes the within-year component of this new analysis, and Section 3 describes the new between-year component. Section 4 summarizes some of the advantages and disadvantages of this method, while Section 5 gives a worked example. The statistical theory and computational details are presented in Annexes 1 and 2.

2 A ROBUST WITHIN-YEAR ANALYSIS

The relatively simple analyses described in ICES (1989, 1991) were based on the assumption that log concentrations are Normally distributed. Hence, average log concentrations provide a sensible summary of the yearly contaminant level. In practice, some data sets showed evidence of outliers or non-Normality (on a log scale), or contained a large proportion of 'less than' values.

A second assumption is that for those contaminants measured in fish muscle, where log concentration is related to fish size, fish length can be used to improve the precision of the estimated yearly contaminant level by removing this source of variation. This assumption also allows average contaminant levels from different years to be expressed at a common fish length. This is straightforward if the slope of the line relating log concentration to fish length remains the same from year to year. In practice, several data sets showed the slope changing from year to year.

Problems arising from violations of these assumptions may be avoided, or at least reduced, by adopting the simple, non-parametric approach described in the following paragraphs.

Where concentration is not thought to be related to fish length (e.g., Cu, Cr, Ni, and PCBs in fish muscle and all contaminants in fish liver and shellfish), the yearly contaminant level is estimated by the median concentration. The median is not affected by small numbers of outliers, or by 'less than' values (provided that fewer than half of the values are below the limit of detection).

For a series of concentrations $c_1, c_2 \ldots c_n$ within a year, ordered so that $c_1 < c_2 \ldots < c_n$, the median is defined as c_{m+1} [where $m = (n-1)/2$] when n is odd, and $(c_m + c_{m+1})/2$ [where $m = n/2$] when n is even.

If it is necessary to remove the effects of biological covariates, for example, when contaminant concentration is related to fish length (e.g., Hg, Pb, and Zn in fish muscle), the data within a year are first ranked by fish length and then split at the median length into two data sets, labelled *Small* and *Large*. Then the median concentration within each size group is calculated.

Splitting the data in this way will account for some of any relationship between contaminant level and fish length. No assumption about the form of this relationship is necessary. Again, the medians will be less susceptible to the effects of outliers and 'less than' values within each size group. The disadvantage of this approach is that there is no adjustment of contaminant level to a standard fish length. However, this adjustment should not be necessary if the sampling guidelines have been followed. Compared with the previous method, using the median could also reduce the power of the trend tests. However, a preliminary assessment by Fryer and Nicholson (1994a) showed that this is likely to be small.

Where the tissue from a sub-sample of individuals has been pooled, the individual observations should be estimated from the contaminant concentration of the sub-sample. For example, if c_i is the concentration measured in the pooled tissue of a sub-sample of five individuals, to calculate the median we would assume five individual observations: c_i, c_i, c_i, c_i, c_i.

Finally, the medians are transformed to a logarithmic scale. This provides continuity with the results from previous assessments, and implies that a linear trend corresponds to a constant percentage change from year to year. Also, one advantage of working on a logarithmic scale is that predicted contaminant concentrations will not be negative. In practice, trends of the order of 10 % per year will appear linear over a ten-year period on both a logarithmic and an arithmetic scale.

3 A ROBUST BETWEEN-YEAR ANALYSIS

This section considers how to assess variation between the yearly median concentrations and describes tests of any systematic variation (e.g., trend). Where the data are divided into *Small* and *Large* fish, there is also a comparison of the trends in these two groups.

The assessment of temporal effects in the regression analysis used for previous JMP/ICES assessments (ICES, 1987) employed the ratio of the between-year variation to the pooled within-year variation to test the null hypothesis that the average contaminant level in each year was the same. If this null hypothesis was rejected, the between-year variation was then partitioned into one component corresponding to a linear trend, and a second component corresponding to a lack of fit of the linear trend. This second component was intended as a catch-all for any non-linear patterns of between-year variation, such as discrete events or shifts in average level.

Problems arise when both of these components are significant, and the second component actually corresponds to *random* between-year variation. The probability of incorrectly finding a significant linear trend when there is none can be very high using this test (Fryer and Nicholson, 1993a).

If the second component of the between-year signal consists only of random between-year variation, the test can be modified, and the trend component tested against the between-year variance estimated from the random component. However, as discussed by Fryer and Nicholson (1993a), this will be misleading if both random variation and some unknown, non-linear systematic variation are present. Separating these two kinds of information is difficult. The approach they adopted was to fit a robust locally weighted smooth curve (Cleveland, 1979) through the estimated yearly contaminant levels. The scatter around this robust curve was then used to estimate the random between-year variance. The smooth curve is then tested to see whether it is significantly different from a linear trend. An advantage of this method is that even when the non-linear trend component is significant, it may still be possible to infer some tendency to increase or decrease if the essential appearance of the smoother is monotonic

upwards or downwards. This can be inferred from the plotted smoother. This approach is followed here.

The statistical assessment of trends for this new method is similar to the previous method in the sense of providing tests of both the linear trend and the lack-of-fit to the linear trend. The difference, as described above, is that the tests now incorporate an estimate of the random variance between years. The tests are described in detail in Annex 1 and consist of comparisons of the residual sums of squares from three models:

Model 1: *log median concentration* = μ

Model 2: *log median concentration* = $\mu + \beta t$, and

Model 3: *log median concentration* = $f(t)$,

where $f(t)$ is some smooth function of time.

The underlying curve $f(t)$ is estimated by a smoother, $f_p(t)$, where p indicates the amount of smoothing. In developing this method, a simple running-line smoother was employed. Initially, its ease of computation and simplicity were attractive, and for typical JMP series of ten years, the choice of smoother did not seem critical. Its smoothing properties were relatively easy to explore and a three-point running mean seemed to offer good properties in terms of minimizing the average mean squared error for an underlying $f(t)$ likely to be of interest. However, with experience and more understanding of this method, a better choice of smoother would be LOESS or cubic spline (Hastie and Tibshirani, 1990). The running-line smoother is still used for demonstration in the worked example in Section 5, below, and in the statistical annexes.

The problem of how much smoothing to do still remains. If the smooth curve is over-fitted, the between-year variance will be underestimated; and if the smooth curve is under-fitted, the between-year variance will be overestimated. Various methods of choosing the degrees of freedom for the smoother are discussed by Hastie and Tibshirani (1990). However, these methods rely on more observations than the five to ten years of typical trend monitoring series. At present, a sensible rule-of-thumb suggested in du Toit *et al.* (1986) is that the smoother should span about 30 % of the data. With ten years of data, this rule implies $p = 1$ for the running-line smoother, and three degrees of freedom for LOESS and cubic spline smoothers.

Where concentration may be related to fish length (e.g., Hg, Pb, and Zn in fish muscle), the analysis is extended as described in Annex 2. The models described above are fitted separately for *Small* and *Large* fish, and then each Model 3 is compared using a series of tests to establish whether the trends for the two groups are different, have the same pattern but at different concentrations, or are identical. If they are identical, there is no evidence of a length effect. If they are parallel, there is a constant (over time) relationship between concentration and length. This was assumed in the previous regression method. If the trends are different for *Small* and *Large* fish, there is not a constant relationship between concentration and length, and the trends are size dependent and should be presented separately.

4 SUMMARY OF STRENGTHS AND WEAKNESSES OF THE NEW METHOD

As with any statistical method, there are both positive and negative aspects of this method. The main positive aspects (strengths) are that:

1) it is robust;

2) it is simple to apply consistently to a large number of data sets;

3) it provides accessible results;

4) it allows for size-dependent trends;

5) it automatically deals with outliers, small amounts of partially pooled data, and 'less than' values; and

6) it allows for random between-year variation.

The main negative aspects (weaknesses) are that:

1) the degree of smoothing is subjective;

2) if the degree of smoothing is wrong, estimates of trend and/or error will be biased;

3) there is no adjustment for length effects; if a species' length range varies from year to year, biological variation may inflate between-year variation;

4) using the median may reduce power; and

5) non-linear trend direction must be inferred graphically.

5 WORKED EXAMPLE

The following example demonstrates the new method using the three-point moving average as the smoother. The data are mercury concentrations in cod muscle for the years 1982–1989, collected by the United Kingdom from ICES areas 34F2, 35F2, and 37F2 (three locations off the east coast of England). Since mercury in fish muscle is expected to vary with fish length, the data have been divided into *Small* and *Large* fish, and the trend assessment applied separately to each size group.

Table 1 shows the mercury concentrations (mg kg^{-1} wet weight) and fish lengths for individual fish in each year. The data have been ranked by fish length, split at the median length (the central row), and then ranked within each size group by mercury concentration. Note that when the number of observations in a year is odd, the median length corresponds to a specific fish with an associated Hg concentration. When the number is even, the median length is the average of the longest *Small* fish and the shortest *Large* fish, and there is no associated concentration.

Table 1. Mercury concentrations (mg kg^{-1} wet weight) in cod muscle (c) and length (l) of the fish by year, separated into *Small* and *Large* fish.

	1982		1983		1984		1985		1986		1987		1988		1989	
	c	l	c	l	c	l	c	l	c	l	c	l	c	l	c	l
	0.05	330	0.08	310	0.05	311	0.08	326	0.04	325	0.05	325	0.06	320	0.06	335
	0.05	340	0.05	312	0.04	340	0.04	344	0.04	330	0.05	330	0.04	330	0.05	340
S	0.06	340	0.06	336	0.04	344	0.03	362	0.04	340	0.06	336	0.05	346	0.07	346
m	0.06	340	0.07	343	0.05	356	0.07	370	0.04	345	0.03	340	0.06	354	0.07	350
a	0.08	360	0.07	353	0.05	363	0.05	376	0.05	364	0.05	348	0.06	364	0.11	361
l	0.09	380	0.05	379	0.04	371	0.06	391	0.11	370	0.04	385	0.06	381	0.13	378
l	0.12	380	0.07	385	0.08	371	0.03	410	0.04	376	0.07	388	0.06	405	0.12	382
	0.06	390	0.09	402	0.06	382	0.04	418	0.04	382	0.06	400	0.06	414	0.12	386
f	0.09	400	0.08	420	0.05	396	0.04	433	0.09	385	0.06	406	0.06	426	0.09	413
i	0.06	420	0.08	422	0.07	432	0.04	435	0.10	430	0.05	410	0.08	435	0.08	426
s	0.05	440	0.08	440	0.05	440	0.06	458	0.07	450	0.04	444	0.08	458	0.06	440
h	0.05	460	0.08	452	0.09	451	0.06	464	0.11	450	0.06	460	0.05	461	0.08	452
	0.08	480														
	0.10	490														
	0.11	500														
median	-	515	0.06	470	-	453	0.08	481	0.05	470	0.05	465	0.06	463	0.08	455
	0.06	530	0.07	492	0.11	455	0.09	486	0.06	493	0.04	466	0.05	464	0.05	478
	0.15	550	0.09	512	0.06	493	0.06	496	0.05	504	0.06	475	0.09	468	0.09	490
L	0.07	560	0.08	526	0.08	525	0.06	585	0.08	536	0.06	554	0.09	534	0.12	545
a	0.08	570	0.09	532	0.09	550	0.08	587	0.11	573	0.14	557	0.07	551	0.09	553
r	0.11	570	0.13	542	0.06	556	0.06	589	0.06	575	0.04	566	0.09	577	0.14	577
g	0.13	620	0.11	551	0.06	610	0.09	604	0.08	597	0.15	572	0.10	582	0.06	588
e	0.13	630	0.12	574	0.11	616	0.08	605	0.10	600	0.12	585	0.07	596	0.11	603
	0.09	640	0.09	622	0.10	634	0.15	620	0.06	635	0.12	630	0.11	622	0.11	625
f	0.09	640	0.19	630	0.09	650	0.14	634	0.09	635	0.10	639	0.16	624	0.19	651
i	0.10	700	0.12	658	0.11	683	0.10	644	0.12	675	0.17	643	0.09	634	0.11	663
s	0.15	740	0.15	693	0.13	728	0.14	669	0.15	680	0.10	648	0.10	666	0.14	676
h	0.14	750	0.14	699	0.15	780	0.14	745	0.14	703	0.10	670	0.12	666	0.17	701
	0.16	760														
	0.17	800														
	0.14	830														

Table 2 shows the median concentrations for the *Small* fish and the *Large* fish by year.

Table 2. Median concentrations of mercury (in mg kg^{-1} wet weight) for *Small* fish and *Large* fish by year.

	1982	**1983**	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**
Small fish	0.060	0.075	0.050	0.045	0.045	0.050	0.060	0.080
Large fish	0.130	0.115	0.095	0.090	0.085	0.100	0.090	0.110

Applying the computations described in Annex 1 separately for *Small* and *Large* fish to log median concentrations, the residual sums of squares (*RSS*) from the three trend models, with associated degrees of freedom (*df*), are:

			RSS	
		df	***Small* fish**	***Large* fish**
Model 1	Mean	7	0.3415	0.1480
Model 2	Linear regression	6	0.3367	0.1166
Model 3	Smoother	5	0.0914	0.0402

The analyses of variance for *Small* and *Large* fish are shown in Table 3.

Table 3. Results of analyses of variance for *Small* fish and *Large* fish (*SSQ* denotes sum of squares, and *MS* denotes mean square).

	df	***SSQ***	***MS***	**F-ratio**	**% Probability**
Small fish					
Systematic year effects	2	0.2501	0.1250	6.84	3.7
Non-linearity	1	0.2453	0.2453	13.42	1.5
Linearity	1	0.0048	0.0048	0.26	77.9
Error	5	0.0914	0.01828		
Large fish					
Systematic year effects	2	0.1078	0.0539	6.70	3.9
Non-linearity	1	0.0764	0.0764	9.50	2.7
Linearity	1	0.0314	0.0314	3.91	10.5
Error	5	0.0402	0.00804		

The log median concentrations for both size groups have a significant systematic between-year effect with a significant non-linear component and a non-significant linear component.

The trend signals for *Small* and *Large* fish are compared using the analysis described in Annex 2. The residual sums of squares from the three models are:

		df	*RSS*
Model 3	Separate lines	10	0.1316
Model 2	Parallel lines	12	0.1824
Model 1	Same line	13	1.4984

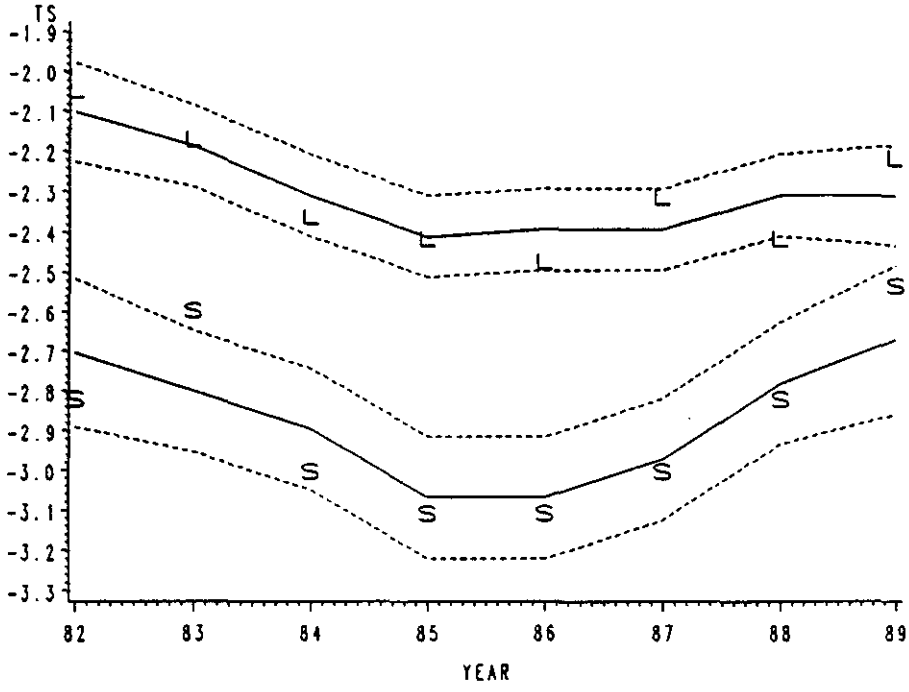
and the analysis of variance table is

Effect	*df*	*SSQ*	*MS*	**F-ratio**	**% Probability**
Different slopes	2	0.0505	0.0252	1.92	19.7
Different intercepts	1	1.3160	1.3160	100	0.0
Error	10	0.1316	0.01316		

The fitted smoothers are not coincident (confirming that mercury concentration depends on fish length), but there is no evidence that they are not parallel.

Figure 1 shows the median log-contaminant concentrations for *Small* and *Large* fish plotted against year together with the fitted smoothers and their 95 % confidence limits. These have been calculated using $\pm t\sqrt{\frac{s^2}{2}}$ for the endpoints (1982, 1989) and $\pm t\sqrt{\frac{s^2}{3}}$ for the other years (1983–1988), where t is the 97.5 percentile for Student's t-distribution with (in this case) ten degrees of freedom and s^2 is the estimated residual variance.

Figure 1. Plot of median log-contaminant concentrations for *Small* (S) fish and *Large* (L) fish against year, with the fitted smoothers (solid line) and their 95 % confidence limits (dashed lines).



6 REFERENCES

Bignert, A., and Nielsen, P.B. 1988. Some tests with artificially pooled samples from *Mytilus edulis*. *In* Report of the Working Group on Statistical Aspects of Trend Monitoring, Annex 4. ICES CM 1988/E:27.

Cleveland, W.S. 1979. Robust locally-weighted regression and smoothing scatterplots. Journal of the American Statistical Association, 74: 829–836.

du Toit, S.H.C., Steyn, A.G.W., and Stumpf, R.H. 1986. Graphical Exploratory Data Analysis. Springer Verlag Inc., New York.

Fryer, R.J., and Nicholson, M.D. 1993a. The power of a contaminant monitoring programme to detect linear trends and incidents. ICES Journal of Marine Science, 50: 161–168.

Fryer, R.J., and Nicholson, M.D. 1993b. Predicting future contaminant levels in relation to EQOs using temporal trend monitoring data. ICES CM 1993/E:18.

Fryer, R.G., and Nicholson, M.D. 1994a. The power of the new robust method of assessing contaminant trend monitoring data. *In* Report of the Working Group on Statistical Aspects of Environmental Monitoring, pp. 165–170. ICES CM 1994/E:6.

Fryer, R.G., and Nicholson, M.D. 1994b. Analysis of variance tables for the new robust method of assessing contaminant trend monitoring data. *In* Report of the Working Group on Statistical Aspects of Environmental Monitoring, pp. 155–164. ICES CM 1994/E:6.

Hastie, T.J., and Tibshirani, R.J. 1990. Generalized additive models. Chapman and Hall, London.

ICES. 1984. The ICES Coordinated Monitoring Programme for Contaminants in Fish and Shellfish, 1978 and 1979, and Six-Year Review of ICES Coordinated Monitoring Programmes. ICES Cooperative Research Report No. 126, p. 96.

ICES. 1987. Guidelines for temporal trend analysis of data on contaminants in fish. *In* Report of the ICES Advisory Committee on Marine Pollution, 1986, pp. 115–128. Cooperative Research Report No. 142.

ICES. 1989. Statistical analysis of the ICES Cooperative Monitoring Programme data on contaminants in fish muscle tissue (1978–1985) for determination of temporal trends. Cooperative Research Report No 162. 147 pp.

ICES. 1991. Statistical analysis of the ICES Cooperative Monitoring Programme data on contaminants in fish liver tissue and *Mytilus edulis* (1978–1988) for determination of temporal trends. ICES Cooperative Research Report No 176. 189 pp.

ICES. 1993. Report of the Working Group on Statistical Aspects of Environmental Monitoring. ICES CM 1993/Env:6.

Jensen, A., and Cheng, Z. 1987. Statistical analysis of trend monitoring data of heavy metals in flounder (*Platichthys flesus*). Marine Pollution Bulletin, 18: 230–238.

Misra, R.K., Uthe, J.F., and Wyncke, W. 1990. A weighted procedure of analysing for time trends in contaminant levels in CMP data: applications to cod and flounder data of the Belgian coast (1978–1985). Journal du Conseil, Conseil International Pour l'Exploration de la Mer, 47: 65–75.

Nicholson, M.D., and Fryer, R.J. 1992. The statistical power of monitoring programmes. Marine Pollution Bulletin, 24: 146–149.

Nicholson, M.D., and Wilson, S.J. 1987. Identification of trends in levels of metals in fish muscle: appraisal of the statistical analysis and data quality. *In* Report of the Working Group on Statistical Aspects of Trend Monitoring, Annex 3.2. ICES CM 1987/E:24.

Nicholson, M.D., Green, N.W., and Wilson, S.J. 1989. Trend measurement from bulked samples. *In* Report of the Working Group on Statistical Aspects of Trend Monitoring, Annex 4. ICES CM 1989/E:13.

Nicholson, M.D., Green N.W., and Wilson, S.J. 1991. Regression models for assessing trends in cadmium and PCBs in cod livers from the Oslofjord. Marine Pollution Bulletin, 22: 77–81.

OSPAR. 1992. Assessment of temporal trend monitoring data for 1983–1988. *In* Monitoring and Assessment, Part C, pp. 145–205. Oslo and Paris Commissions, London.

OSPAR. 1994. Principles and Methodology of the Joint Monitoring Programme. Oslo and Paris Commissions, London.

OSPAR. 1995. The Joint Assessment and Monitoring Programme. Oslo and Paris Commissions, London. 45 pp.

Rees, H.L., and Nicholson, M.D. 1989. Trends in lead levels in the horse-mussel from the Western North Sea. Marine Pollution Bulletin, 20: 86–89.

Warren, W.G. 1993a. Modelling temporal trends in marine populations: pollutant concentrations as a function of fish length. *In* Report of the Working Group on Statistical Aspects of Environmental Monitoring, Annex 3. ICES CM 1993/Env:6.

Warren, W.G. 1993b. Some applications of the Kalman Filter in fisheries reseach. ICES CM 1993/D:57.

ANNEX 1

LINEAR SMOOTHERS AND STATISTICAL TESTS OF TRENDS

Consider a series of observations c_t made in years y_t, where $t = 1, \ldots T$ and T is the total number of yearly observations. (This notation allows for missing years.) Underlying the c_t is an assumed true curve $f(y_t)$ such that

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

and ε_t has variance σ^2. The underlying curve is estimated by a smoother, $f_p(y_t)$, where p indicates the amount of smoothing.

Linear smoothers can be written

$$f_p(y_t) = S\underline{c}$$

where S is a $T \times T$ matrix determined by the choice of smoother and $\underline{c}$ is the T-vector of contaminant levels.

The residual sum of squares from the fitted smoother is given by

$$RSS = \sum_{t=1}^{T} [c_t - f_p(y_t)]^2$$

with residual degrees of freedom given by

$$df = T - tr(2S - SS').$$

If $f_p(y_t)$ is an unbiased estimator of $f(y_t)$, then an unbiased estimator of σ^2 is given by

$$s^2 = \frac{RSS}{df}$$

and tests of significance about the fitted model can be based on the fact that

$$\frac{df \times s^2}{\sigma^2} \approx \chi^2_{df}.$$

There is a wide choice of potential smoothers (see Hastie and Tibshirani, 1990). For example, a running-line smoother calculates $f_p(y_t)$ from the local linear regression calculated from the $2p + 1$ points centred on y_t. This can be extended by using a weighted regression, where the weights decrease with distance from y_t. The $2p + 1$ points can be taken from the symmetric nearest neighbourhood as above, or from the neighbourhood containing the $2p + 1$ points closest to y_t. The LOESS smoother used by Fryer and Nicholson (1993a, 1993b) uses an iterative procedure, first weighting by distance, and then down-weighting points with large residuals. Smoothers are available directly in statistics packages such as GENSTAT and S-Plus, or indirectly as macros in SAS (see, e.g., du Toit *et al.*, 1986).

The statistical assessment of trends when a smoother is used is similar to that reported in ICES (1989, 1991) in the sense of providing a test of the linear trend and of the non-linear trend. The major difference is that the procedure now incorporates an estimate of the random between-year variance, against which the non-random components are tested.

Writing c_t for the median log concentration in year y_t, the tests consist of comparing the residual sums of squares from the three fitted models

Model 1: $c_{1t} = \bar{c} = T^{-1}\sum_{t=1}^{T} c_t$

Model 2: $c_{2t} = a + by_t$

where a and b are the usual least-squares estimators, and

Model 3: $c_{3t} = f_p(y_t) = S\underline{c}$.

The residual sums of squares for these models are

$$RSS_1 = \sum_{t=1}^{T}(c_t - c_{1t})^2$$

$$RSS_2 = \sum_{t=1}^{T}(c_t - c_{2t})^2$$

$$RSS_3 = \sum_{t=1}^{T}(c_t - c_{3t})^2$$

with degrees of freedom

$df_1 = T - 1$

$df_2 = T - 2$

$df_3 = T - tr(2S - SS')$

respectively. The analysis of variance is given by

Effect	***df***	**Sum of Squares**	**F-ratio**	**% Probability**
Systematic year effect	$df_1 - df_3$	$RSS_1 - RSS_3$	$\frac{[RSS_1 - RSS_3]df_3}{RSS_3[df_1 - df_3]}$	$100[1 - P(F, df_1 - df_3, df_3)]$
Non-linearity	$df_2 - df_3$	$RSS_2 - RSS_3$	$\frac{[RSS_2 - RSS_3]df_3}{RSS_3[df_2 - df_3]}$	$100[1 - P(F, df_2 - df_3, df_3)]$
Linearity	$df_1 - df_2$	$RSS_1 - RSS_2$	$\frac{[RSS_1 - RSS_2]df_3}{RSS_3[df_1 - df_2]}$	$100[1 - P(F, df_1 - df_2, df_3)]$
Error	df_3	RSS_3	$s^2 = \frac{RSS_3}{df_3}$	

where $P(F, f_1, f_2)$ is the cumulative probability for an observed F-ratio from an F-distribution with f_1 and f_2 degrees of freedom.

To demonstrate, consider a running-mean smoother, consisting of the average of the p points to the left and p points to the right of c_t. For example, with $p = 1$ we have

$$S = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & . & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & . & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & . & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & . & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & . & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & . & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

with residual degrees of freedom

$$df_3 = \frac{2T-1}{3}.$$

The following example demonstrates the computational details of this analysis using imaginary data from 1982 to 1988, with a missing year in 1985. Note that y_t has been scaled to $y_t - 1980$. Although the data span seven years, the missing year implies $T = 6$.

t	y_t	c_t	$c_t - \bar{c}$	$a + by_t$	$c_t - (a + by_t)$	$f_1(y_t)$	$c_t - f_1(y_t)$
1	2	3	–1	3.893	–0.893	2 = (3 + 1)/2	1
2	3	1	–3	3.929	–2.929	4 = (3 + 1 + 8)/3	–3
3	4	8	4	3.964	4.036	5 = (1 + 8 + 6)/3	3
4	6	6	2	4.036	1.964	6 = (8 + 6 + 4)/3	0
5	7	4	0	4.071	–0.071	4 = (6 + 4 + 2)/3	0
6	8	2	–2	4.107	–2.107	3 = (4 + 2)/2	–1
		$\bar{c} = 4$	$RSS_1 = 34$		$RSS_2 = 33.964$		$RSS_3 = 20$
			$df_1 = 5$		$df_2 = 4$		$df_3 = 3.67$

Note that non-integer degrees of freedom are possible for fitted smoothers.

The computations for Model 1 (mean) and Model 2 (regression of c_t on y_t) are

$$RSS_1 = \sum_{t=1}^{T} (c_t - \bar{c})^2$$

$$b = \frac{\sum_{i=1}^{T} (y_t - \bar{y})(c_t - \bar{c})}{\sum_{t=1}^{T} (y_t - \bar{y})^2}$$

$$a = \bar{c} - b\bar{y}$$

$$RSS_2 = \sum_{t=1}^{T} (c_t - \bar{c})^2 - \frac{(\sum_{i=1}^{T} (y_t - \bar{y})(c_t - \bar{c}))^2}{\sum_{t=1}^{T} (y_t - \bar{y})^2}$$

The analysis of variance gives

Effect	***df***	***SSQ***	***MS***	**F-ratio**	**% Probability**
Systematic year effect	1.33	14.00	10.50	1.99	25.5
Non-linearity	0.33	13.964	41.90	7.68	6.12
Linearity	1.00	0.036	0.036	0.01	94.0
Error	3.67	20.00	$s^2 = 5.46$		

REFERENCES

du Toit, S.H.C., Steyn, A.G.W., and Stumpf, R.H. 1986. Graphical exploratory data analysis. Springer Verlag, New York.

Fryer, R.J., and Nicholson, M.D. 1993a. The power of a contaminant monitoring programme to detect linear trends and incidents. ICES Journal of Marine Science, 50: 161–168.

Fryer, R.J., and Nicholson, M.D. 1993b. Predicting future contaminant levels in relation to EQOs using temporal trend monitoring data. ICES CM 1993/E:18.

Hastie, T.J., and Tibshirani, R.J. 1990. Generalized additive models. Chapman and Hall, London.

ICES. 1989. Statistical analysis of the ICES Cooperative Monitoring Programme data on contaminants in fish muscle tissue (1978–1985) for determination of temporal trends. Cooperative Research Report No 162. 147 pp.

ICES. 1991. Statistical analysis of the ICES Cooperative Monitoring Programme data on contaminants in fish liver tissue and *Mytilus edulis* (1978–1988) for determination of temporal trends. ICES Cooperative Research Report No 176. 189 pp.

ANNEX 2

TESTING BETWEEN TRENDS FOR *SMALL* AND *LARGE* FISH

Where contaminant concentration is related to fish length, smoothers can be fitted separately to the time series for *Small* and *Large* fish and compared to see whether they are the same, parallel, or different.

Writing $s_1\ s_2\ ....\ l_1\ l_2\ ....$ for the time series of median log concentrations in *Small* and *Large* fish, respectively, and $\underline{c} = (s_1\ l_1\ s_2\ l_2\ ....\ )'$ for the augmented vector of these observations, then the smoothers are given by $S_i\underline{c}$ for Model i where

Model 1 = same smoother,

Model 2 = parallel smoothers, and

Model 3 = different smoothers.

Again, we will use the three-point running mean smoother with $p = 1$ to demonstrate.

For Model 1, a common smoother for both *Small* and *Large* fish, S_1 is given by

$$S_1 = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & . & . \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & . & . \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & . & . \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & . & . \\ 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & . & . \\ 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \end{bmatrix}$$

Intuitively, this is equivalent to averaging the individual smoothers for *Small* and *Large* fish.

The residual degrees of freedom for S_1 are given by

$$df_1 = 2T - tr(2S_1 - S_1S_1') = \frac{5T - 1}{3}.$$

For Model 2, where the trends are parallel, S_2 is given by

$$S_2 = \begin{bmatrix} \frac{1}{4}+\frac{1}{2T} & \frac{1}{4}-\frac{1}{2T} & \frac{1}{4}+\frac{1}{2T} & \frac{1}{4}-\frac{1}{2T} & \frac{1}{2T} & -\frac{1}{2T} & \frac{1}{2T} & -\frac{1}{2T} & \cdot & \cdot \\ \frac{1}{4}-\frac{1}{2T} & \frac{1}{4}+\frac{1}{2T} & \frac{1}{4}-\frac{1}{2T} & \frac{1}{4}+\frac{1}{2T} & -\frac{1}{2T} & \frac{1}{2T} & -\frac{1}{2T} & \frac{1}{2T} & \cdot & \cdot \\ \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{2T} & -\frac{1}{2T} & \cdot & \cdot \\ \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & -\frac{1}{2T} & \frac{1}{2T} & \cdot & \cdot \\ \frac{1}{2T} & -\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \cdot & \cdot \\ -\frac{1}{2T} & \frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \frac{1}{6}-\frac{1}{2T} & \frac{1}{6}+\frac{1}{2T} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Intuitively, this is equivalent to Model 1 with the smoothers centred on the corresponding averages for *Small* and *Large* fish.

The residual degrees of freedom for S_2 are given by

$$df_2 = 2T - tr(2S_2 - S_2S_2') = \frac{5T-4}{3}.$$

For Model 3, S_3 is given by

$$S_3 = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & \cdot & \cdot \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \cdot & \cdot \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \cdot & \cdot \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & \cdot & \cdot \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \cdot & \cdot \\ 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

This corresponds to the individual smoothers fitted separately to *Small* and *Large* fish.

The residual degrees of freedom for S_3 are given by

$$df_3 = 2T - tr(2S_3 - S_3S_3') = \frac{4T-2}{3}.$$

Writing RSS_i for the residual sum of squares from the i'th model, and df_i for the corresponding residual degrees of freedom, the analysis of variance is given by

Effect	*df*	Sum of Squares	F-ratio	% Probability
Different slopes	$df_2 - df_3$	$RSS_2 - RSS_3$	$\frac{[RSS_2 - RSS_3]df_3}{RSS_3[df_2 - df_3]}$	$100[1 - P(F, df_2 - df_3, df_3)]$
Different intercepts	$df_1 - df_2$	$RSS_1 - RSS_2$	$\frac{[RSS_1 - RSS_2]df_3}{RSS_3[df_1 - df_2]}$	$100[1 - P(F, df_1 - df_2, df_3)]$
Error	df_3	RSS_3	$s^2 = \frac{RSS_3}{df_3}$	

This is demonstrated using an extended version of the artificial example introduced in Annex 1.

			Model 3	**Model 1**	**Model 2**
t	y_t	*Small* c_t	$f_S(y_t)$	$(f_S(y_t) + f_L(y_t))/2$	$(f_S(y_t) + f_L(y_t))/2 + (\bar{c}_S - \bar{c}_L)/2$
1	2	3	(3 + 1)/2 = 2	(2 + 7)/2 = 4.5	4.5 – 2.5 = 2.0
2	3	1	(3 + 1 + 8)/3 = 4	(4 + 7)/2 = 5.5	5.5 – 2.5 = 3.0
3	4	8	(1 + 8 + 6)/3 = 5	(5 + 8)/2 = 6.5	6.5 – 2.5 = 4.0
4	6	6	(8 + 6 + 4)/3 = 6	(6 + 10)/2 = 8.0	8.0 – 2.5 = 5.5
5	7	4	(6 + 4 + 2)/3 = 4	(4 + 11)/2 = 7.5	7.5 – 2.5 = 5.0
6	8	2	(4 + 2)/2 = 3	(3 + 12)/2 = 7.5	7.5 – 2.5 = 5.0
		$\bar{c}_S = 4$			
t	y_t	*Large* c_t	$f_L(y_t)$	$(f_S(y_t) + f_L(y_t))/2$	$(f_S(y_t) + f_L(y_t))/2 + (\bar{c}_S - \bar{c}_L)/2$
1	2	6	(6 + 8)/2 = 7	(2 + 7)/2 = 4.5	4.5 + 2.5 = 7.0
2	3	8	(6 + 8 + 7)/3 = 7	(4 + 7)/2 = 5.5	5.5 + 2.5 = 8.0
3	4	7	(8 + 7 + 9)/3 = 8	(5 + 8)/2 = 6.5	6.5 + 2.5 = 9.0
4	6	9	(7 + 9 + 14)/3 = 10	(6 + 10)/2 = 8.0	8.0 + 2.5 = 10.5
5	7	14	(9 + 14 + 10)/3 = 11	(4 + 11)/2 = 7.5	7.5 + 2.5 = 10.0
6	8	10	(14 + 10)/2 = 12	(3 + 12)/2 = 7.5	7.5 + 2.5 = 10.0
		$\bar{c}_L = 9$	$RSS_3 = 37$	$RSS_1 = 129.5$	$RSS_2 = 54.5$
			$df_3 = 7.3$	$df_1 = 9.7$	$df_2 = 8.7$

Effect	*df*	Sum of Squares	*MS*	F-ratio	% Probability
Different slopes	1.4	17.5	12.5	2.47	15.7
Different intercepts	1	75.0	75.0	14.80	0.6
Error	7.3	37.0	$s^2 = 5.07$		

Note that this analysis assumes that the errors around the smoothers for *Small* and *Large* fish are independent. One way in which this assumption might be violated is if contaminant levels for *Small* and *Large* fish are given by

$$c_{St} = f_S(y_t) + \varepsilon_{St} + \delta_t$$

$$c_{Lt} = f_L(y_t) + \varepsilon_{Lt} + \delta_t$$

where $V[\varepsilon_{St}] = V[\varepsilon_{Lt}] = \sigma_\varepsilon^2 \quad V[\delta_t] = \sigma_\delta^2$

and the ε's are independent between size groups and times, but the δ's are independent between times and common to each size group. This more general model allows for both independent and correlated errors.

It is easy to show that neither (RSS_1–RSS_2) nor (RSS_2–RSS_3) depend on the value of σ_δ^2, but that

$$E[RSS_3] = \frac{4T-2}{3}\left(\sigma_\varepsilon^2 + \sigma_\delta^2\right)$$

and that df_3 varies between $\frac{2T-1}{3}$ and $\frac{4T-2}{3}$ as $\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\delta^2}$ varies from 0 to 1.

Hence, the effect of σ_δ^2 is to inflate the residual sum of squares and to overestimate the residual degrees of freedom. Thus, the overall effect is unclear. However, simulations of tests for different values of σ_δ^2 suggest that the effect is to make the tests of parallel and coincident smoothers conservative.

A more comprehensive analysis of variance which attempts to take account of non-zero values of σ_δ^2 is presented by Fryer and Nicholson (1994). This also incorporates corrections to the degrees of freedom proposed by Hastie and Tibshirani (1990) to make the tests approximate more closely to an F-distribution.

REFERENCES

Fryer, R.G., and Nicholson, M.D. 1994. Analysis of variance tables for the new robust method of assessing contaminant trend monitoring data. *In* Report of the Working Group on Statistical Aspects of Environmental Monitoring, pp. 155–164. ICES CM 1994/E:6.

Hastie, T.J., and Tibshirani, R.J. 1990. Generalized additive models. Chapman and Hall, London.

ICES Techniques in Marine Environmental Sciences

No. 1	Cadmium and lead: Determination in organic matrices with electrothermal furnace atomic absorption spectrophotometry
No. 2	Trace metals in sea water: Sampling and storage methods
No. 3	Cadmium in marine sediments: Determination by graphite furnace atomic absorption spectroscopy
No. 4	Lipophilic organic material: An apparatus for extracting solids used for their concentration from sea water
No. 5	Primary production: Guidelines for measurement by ^{14}C incorporation
No. 6	Control procedures: Good laboratory practice and quality assurance
No. 7	Suspended particulate matter: Collection methods for gravimetric and trace metal analysis
No. 8	Soft bottom macrofauna: Collection and treatment of samples
No. 9	Sediments and suspended particulate matter: Total and partial methods of digestion (*videotape available*)
No. 10	Organic halogens: Determination in marine media of adsorbable, volatile, or extractable compound totals
No. 11	Biological effects of contaminants: Oyster (*Crassostrea gigas*) embryo bioassay
No. 12	Hydrocarbons: Review of methods for analysis in sea water, biota, and sediments
No. 13	Biological effects of contaminants: Microplate method for measurement of ethoxyresorufin-O-deethylase (EROD) in fish
No. 14	Temporal trend monitoring: Introduction to the study of contaminant levels in marine biota
No. 15	Temporal trend monitoring: Contaminant levels in tissues of Atlantic cod
No. 16	Benthic communities: Use in monitoring point-source discharges
No. 17	Nutrients: Practical notes on their determination in sea water
No. 18	Contaminants in marine organisms: Pooling strategies for monitoring mean concentrations
No. 19	Common diseases and parasites of fish in the North Atlantic: Training guide for identification
No. 20	Temporal trend monitoring: Robust method for analysing contaminant trend monitoring data
No. 21	Chlorobiphenyls in marine sediments: Guidelines for determination
No. 22	Biological effects of contaminants: Cholinesterase inhibition by organophosphate and carbamate compounds
No. 23	Biological effects of contaminants: Determination of CYP1A-dependent mono-oxygenase activity in dab by fluorimetric measurement of EROD activity