

This report not to be quoted without prior reference to the Council*

International Council for the
Exploration of the Sea

C.M.1989/E:13
Ref. Statistics Committee
Theme Session T

**REPORT OF THE WORKING GROUP ON
THE STATISTICAL ASPECTS OF TREND MONITORING**

The Hague, 24-27 April 1989

This document is a report of a Working Group of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council. Therefore, it should not be quoted without consultation with the General Secretary.

*General Secretary
ICES
Palægade 2-4
DK-1261 Copenhagen K
DENMARK

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
1 OPENING OF THE MEETING AND ORGANIZATION OF WORK	1
2 ADOPTION OF THE AGENDA	1
3 REVIEW OF THE ACOMP REPORT	1
4 RESULTS OF ANALYSIS OF CMP DATA FOR TRENDS IN CONTAMINANTS IN FISH LIVER TISSUE	1
4.1 Study of the CMP Fish Liver Data	2
4.2 Trend Measurement from Bulk Samples	2
4.3 The Effect of Analytical Accuracy on Tests for Trends	3
5 RESULTS OF ANALYSIS OF CMP DATA FOR TRENDS IN CONTAMINANTS IN BLUE MUSSELS	3
5.1 Evaluation of CMP/JMP Data on Contaminants in Blue Mussels	3
5.2 The Effect of Depuration on Metal Levels in Blue Mussels	4
6 CONSIDERATION OF THE RESULTS OF TREND STUDIES IN CANADIAN ATLANTIC COD FOR 1977-1985 BY UNIVARIATE AND MULTIVARIATE ANALYSES	4
6.1 Principal Components and the Problem of Multicollinearity	6
7 PROGRESS IN THE ANALYSIS OF TRENDS IN CONTAMINANTS IN SEVERAL SPECIES FROM THE SAME AREA	7
8 DRAFT LEAFLETS ON TREND MONITORING TOPICS FOR PUBLICATION IN THE TIMES SERIES	7
9 USE OF SEaweEDS IN MONITORING NON-RADIOACTIVE CONTAMINANTS	8

<u>Section</u>	<u>Page</u>
10 REVIEW OF THE ENGLISH VERSION OF THE PAPER BY P. GROS ON STATISTICAL MANAGEMENT OF MARINE ENVIRONMENTAL DATA . .	8
11 REVIEW OF A PAPER ON THE EFFICIENCY OF POOLING IN MUSSEL WATCH STUDIES	8
12 EFFECT OF LABORATORY BIAS ON TEMPORAL TREND STUDIES . .	9
13 SAMPLE STRUCTURE FOR MONITORING CONTAMINANTS IN BIOTA, SEAWATER, AND SEDIMENTS	9
13.1 Guidelines on the Philosophy, Principles, and Strategy of Monitoring	9
13.2 "An introduction to the study of spatial and temporal trends in contaminant levels in marine biota." by J.F. Uthe et al.	10
13.3 "Sampling strategies for trend monitoring using biota, sediments or seawater" by J.F. Uthe et al.	10
13.4 "Effect of Sampling Structure on Regression Estimates: Why the CMP Sampling Guidelines Should be Followed" by M.D. Nicholson and S.J. Wilson	10
14 OTHER BUSINESS	11
14.1 Papers on Nutrients	11
14.2 Communication from JMG (Dr J. Portmann)	12
14.3 Long-term Trends in Laboratory Analysis	14
14.4 Items from WGEAMS	14
14.4.1 Guidelines	14
14.4.2 Contaminant Matrices Related to the JMP Purposes . .	15
14.4.3 Data Quality and Level of Detectable Change in Contaminant Level	16
14.5 Proposal for Chairman for the period, 1990-1992 . . .	16
15 RECOMMENDATIONS	16
16 CLOSING OF THE MEETING	17
ANNEX 1: List of Participants	18

Section	Page
ANNEX 2: Agenda	20
ANNEX 3: Assessment of Trends in Contaminants in Fish Livers	21
ANNEX 4: Trend Measurement from Bulked Samples	33
ANNEX 5: The Effect of Analytical Accuracy on The Power of Student's Two-Sample t-Test	43
ANNEX 6: The Effect of Depuration on Nussel Analyses	52
ANNEX 7: On Multivariate and Univariate Analyses of Variance	59
ANNEX 8: On Time Trends, 1977-1985, in Canadian Cod Atlantic (<u>gadus morhua</u>)	80
ANNEX 9: Methods for Collecting Benthic Algae for Contaminant Monitoring	85
ANNEX 10: Pooling May Economize a Sampling Program	94
ANNEX 11: Effect of Sampling Structure on Regression Estimates	101
ANNEX 12: Telefaxed Response by WGSATM to WGEAMS Inquires . .	110

1 OPENING OF THE MEETING AND ORGANIZATION OF WORK

The Chairman, Dr J.F. Uthe, opened the meeting at 9:30 a.m. on 24 April 1989. Dr J. van der Meer welcomed participants to The Hague and to the host institute, the Rijkswaterstaat Tidal Waters Division. The Chairman noted that the Group had a large number of items to consider and requested each member introducing an item to prepare a short note for inclusion in the report of the meeting. The list of attendees is attached as Annex 1.

2 ADOPTION OF THE AGENDA

The agenda (Annex 2) was adopted with some modification and the addition of a number of new reports under various items.

3 REVIEW OF THE ACMP REPORT

The appropriate sections of the 1988 ACMP Report were reviewed and felt to accurately reflect the deliberations and conclusions of the 1988 WGSATM meeting. Although not directly related to this item, the Chairman noted that the Working Group on Environmental Assessments and Monitoring Strategies was meeting concurrently, which would cause some difficulty in communicating the results of this year's meeting of WGSATM to WGEAMS. However, a number of items of interest to WGEAMS (see items 13.1 and 14.4) were discussed early in the meeting and the results of these discussions forwarded to WGEAMS immediately by telefax.

4 RESULTS OF ANALYSIS OF CMP DATA FOR TRENDS IN CONTAMINANTS IN FISH LIVER TISSUE

This section and Section 5 of the report describe the ongoing activities of WGSATM concerning the statistical analyses of the Co-ordinated ICES Monitoring Studies Programme (CMP) data on contaminants in fish and shellfish for the determination of temporal trends. With the publication of the statistical analyses of data on contaminants in fish muscle tissue in ICES Cooperative Research Report No. 162, the Group is currently concerned with the evaluation of the data on contaminants in fish liver tissue and shellfish soft body tissue.

These analyses are being conducted to develop the applications of the statistical methods for temporal trend elucidation to data on contaminants in fish liver and shellfish soft body tissues. The results of these statistical analyses will thereafter be transmitted to the ad hoc Working Group on Monitoring (under the Joint Monitoring Group (JMG) of the Oslo and Paris Commissions) for incorporation into their evaluation of the Joint Monitoring Programme (JMP) temporal trend monitoring data.

The statistical analysis of the shellfish data are considered in Section 5 of this report, while the fish liver tissue analyses are covered in Section 4.1, below.

4.1 Study of the CMP Fish Liver Data

A paper, "Evaluating important biological covariables" by M.D. Nicholson, N. Green, and S.J. Wilson, was presented as constituting Part I of the preliminary assessment of data on contaminant concentrations in fish livers.

Data sets from the ICES CMP had been selected where data on individual fish length, liver weight, and fat percent were present. These variables had previously been identified as influencing concentrations of PCB and cadmium in cod livers (see 1988 WGSATM report, ICES C.M.1988/E:27, Section 8.2 and Annex 7).

This paper, included in this report as Annex 3, describes an investigation focussing on the influence of these variables on concentrations of PCB, cadmium, copper, lead, mercury, and zinc in liver tissue of cod, whiting, flounder, and plaice in areas of the North Sea. Not all contaminant/species combinations were present and some data sets were very small (< 20 observations). The importance of variables differed among contaminants and species, although, in some cases, this could have been due to the restricted range of the variables.

Once again, the Working Group commented upon the difficulties of analysing and interpreting incomplete data sets and data of poor quality, and reiterated the importance of following the CMP sampling guidelines.

4.2 Trend Measurement from Bulkied Samples

Tissue from samples of marine biota are sometimes bulkied, either to provide sufficient material for chemical analysis or to reduce the analytical costs. The statistical distribution of contaminant concentrations when analyzed in the bulkied tissue may be different from the distribution of concentrations when analyzed in the tissues on an individual organism basis. This should be taken into account when assessing trends in contaminant concentrations, and has been the subject of continuing work within WGSATM over the past years.

A paper, "Trend measurement from bulkied samples: The effect of pooling on log-normally distributed contaminants" by M.D. Nicholson, N.W. Green, and S.J. Wilson (Annex 4), was presented which examined this problem when concentrations in the tissues of individuals are log-normally distributed. Some simple approximations showed that increasing the degree of bulking would not affect accuracy, but would lead to an apparent increase in geometric mean contaminant levels. These results were extended using computer simulation to monitoring levels of PCB in cod livers. Bulking had the expected effect of not affecting the measured difference between contaminant levels between years, although the probability of finding a significant difference was reduced.

On the basis of the information reported under items 4.1 and 4.2, the group were informed that it is intended that the preliminary assessment of the data on contaminants in fish liver tissue will continue and the results be finalized with a view to publication in the ICES Cooperative Research Report series. It is further intended that this work will be completed in time for the results

of the statistical analyses to be made available to the JMP's ad hoc Working Group on Monitoring, meeting in December 1989.

4.3 The Effect of Analytical Accuracy on Tests for Trends

In practice, chemical analyses of contaminants may be both biased and imprecise. This will have an adverse effect on statistical tests of significance. Imprecision will usually be lost in comparison with the standard deviation of the contaminant in the field. Bias, however, will persist to distort estimated trends. A paper entitled, "The effect of analytical accuracy on the power of Student's two sample t-test", by M. Nicholson and J. van der Meer addressed this topic; it is attached to this report as Annex 5.

Data from the Seventh ICES Intercalibration Exercise for Trace Metals in Biological Tissue were used to estimate precision, variable bias within a year, and persistent bias for nine laboratories. The ability of these laboratories to detect a given trend in either copper or lead in mussels was assessed. As would be expected, the ability to detect trends in lead, where analytical performance was poorer, was more affected than for copper. The Working Group felt this paper to be an important contribution and that the message contained therein should be stressed and brought to the attention of groups such as the ACMP and MCWG.

5 RESULTS OF ANALYSIS OF CMP DATA FOR TRENDS IN CONTAMINANTS IN BLUE MUSSELS

5.1 Evaluation of CMP/JMP Data on Contaminants in Blue Mussels

A draft paper, "Preliminary Evaluation of CMP/JMP Data on Contaminants in Blue Mussels for the Determination of Temporal Trends" by P.B. Nielsen and S.J. Wilson, was submitted to the group for consideration. This paper was introduced as constituting the first stage in the preparation of an analysis of blue mussel (*Mytilus edulis*) data covering the period 1980-1987, which had been submitted to the ICES CMP.

Data submitted by five countries (Belgium, Federal Republic of Germany, Netherlands, Norway and Sweden) were included in the evaluation, comprising 15 sets of mussel samples from 13 locations. Data on seven metals, (cadmium, chromium, copper, mercury, nickel, lead, and zinc) and PCB had been considered for analysis. The paper described the results of two treatments applied to the data.

The first treatment consisted of a basic presentation of "apparent trends" as reflected by plots of the yearly arithmetic means of the reported contaminant concentration values - including both pooled and individually analysed samples. This approach represented a basic means of looking at the data from all available data sets, regardless of data characteristics and ignoring possible biological effects and the effect of pooling.

As a second approach, a more involved analysis was undertaken by investigating the use of shell weight as a possible covariate to adjust contaminant burden estimates for biological variation be-

tween years, using that sub-set of the data where the necessary parameters had been reported.

After considering the results obtained by these two methods, it was agreed that some further work was required in order to assess (i) whether the shell weight relationship constituted a useful approach, and (ii) whether geometric means should be computed as an alternative to arithmetic means in the basic data presentations. The authors agreed to conclude this work as soon as possible after the meeting so as to complete a statistical analyses of the data, by whichever method proved to be most appropriate, for consideration by the JMP and HOC WG on Monitoring.

It was further agreed that a revised manuscript, incorporating this intersessional work would be made available for consideration at the 1990 meeting of WGSATM, for possible publication as a Cooperative Research Report along with the results of the statistical analysis of the fish liver tissue data.

5.2 The Effect of Depuration on Metal Levels in Blue Mussels

During its 1987 meeting, WGSATM had considered possible modifications to the guidelines for sampling shellfish in connection with temporal trend monitoring studies (see Annex 9 of C.M.1987/E:24). As one of the proposals for further work in association with this review, an investigation into the effects of depurating mussels sampled had been suggested.

N. Green presented evidence (Annex 6) that depuration had no effect on concentrations of trace metals (cadmium, copper, mercury, lead, and zinc) found in the soft parts of mussels (Mytilus edulis) collected in waters with low suspended particulate matter (SPM) concentrations. He suggested that, if not already done so, a similar study should be carried out in waters with high SPM concentrations, e.g., by the Netherlands, in accordance with the depuration procedures given by WGSATM in 1987.

The Working Group accepted the results presented in this paper, noting, however, that one cannot, as yet, expand these results to other metals or other types of contaminant. The need for the experiment to be repeated with mussels collected in water with high SPM was supported.

Accepting these comments, N. Green suggested that the draft guidelines for sampling of mussels produced by WGSATM at its 1987 meeting might now be formalized and recommended for adoption by the various monitoring groups. This matter is discussed under Section 14.4.1, below.

6 CONSIDERATION OF THE RESULTS OF TREND STUDIES IN CANADIAN ATLANTIC COD FOR 1977-1985 BY UNIVARIATE AND MULTIVARIATE ANALYSES

In response to a request that WGSATM considers the results of trend studies in Canadian Atlantic cod for 1977-85, by univariate and multivariate methods, Dr Misra presented two papers. The first (Annex 7) compared and contrasted univariate and multivariate results using two years of Belgian flounder data and only two

contaminants to keep the statistical models and analysis as simple as possible for the non-mathematical reader. The second paper (Annex 8), in a similar vein, presented time trend studies on the Canadian Atlantic cod data. Dr Misra noted the following points.

- 1) When criterion variables, Y_i , (contaminants) were uncorrelated, temporal variations examined by several univariate analyses would yield results which are similar to those from a single multivariate analysis.
- 2) However, Y_i variables are generally mutually correlated. For these data, temporal variations which are identified by the multivariate procedure become increasingly more difficult to identify by univariate analysis as the extent of the overlaps in annual ranges of individual Y_i variables increases.
- 3) Separate ANOVAs or ANCOVAs may still be desirable for reasons such as
 - a) when the manager is primarily or exclusively interested in tracking temporal variations of individual or specific contaminants or
 - b) when there are several missing observations in a data set. A good working strategy when employing a series of univariate analyses would be to supplement them with a multivariate analysis for drawing inferences concerning temporal variations and time trends for contaminants.
- 4) For the Canadian Atlantic cod data (4 data years, 10 contaminants) MANCOVA was employed using log (length) as the covariate. The results showed:
 - a) time trend was identified, and
 - b) contaminants which contributed significantly to the time trend were identified individually by MANCOVA. Furthermore these contaminants were the same ones which showed similar (increasing or decreasing) trends in the scatterplots of adjusted yearly-means of Y_i versus years. It should be recalled that these adjusted yearly-mean values are the same for MANCOVA and ANCOVAs.

Questions were asked concerning the interpretation of composites or linear combinations of criteria, i.e. contaminant variables in time-trend investigations by the multivariate procedure. In response Dr Misra said that the MANOVA (and MANCOVA) is very flexible in this respect. It is possible to test the significance of any linear combination of contaminants. The user may, therefore, test sensible biochemical models of contaminants, for example by assigning values for each coefficient of contaminant variables and then testing the significance of this combination. Not all contaminants need be included in an arbitrarily selected linear combination. It is, therefore, possible to test the contribution of one or more contaminants of specific interest to time trends. In particular, contributions (in the multivariate context) of

individual contaminants to time trends can be tested for significance.

The Working Group suggested that the rejection of univariate procedures in favour of multivariate procedures, as stated in the paper, should be modified to note those situations in which the univariate procedure is appropriate; for example, since it is known that missing observations virtually destroy multivariate analysis, its use is thus prohibited for most CMP data sets (WGSATM 1988). The author agreed with this, and a revised text is included as Annex 7 to this report.

The discussion emphasized that results from neither MANCOVA nor ANCOVA were necessarily correct. Simultaneous and individual statistical analysis of metals data does, however, provide different perspectives on what the trend data indicate.

The difference between them is that MANCOVA may indicate trends in a mixture of contaminants and provides a test of whether such a mixture varies with time. Key questions are:

- 1) Are these differences due to competition or synergism between the contaminants? and
- 2) How do they reflect inputs and changes in inputs?

In the former case the multivariate approach is preferable to the simplistic interpretation of trends in levels of a single contaminant as necessarily a direct reflection of inputs and input changes. However, at this stage there are insufficient data to answer these questions.

There was considerable debate regarding confidence limits of adjusted yearly-means calculated by MANCOVA and ANCOVA. It was agreed that both had their uses but that neither should be used without a full understanding of their meaning.

6.1 Principal Components and the Problem of Multicollinearity

Dr Misra presented his paper, "Time trends of chemical contaminant levels in Canadian Atlantic cod (*Gadus morhua*) with several biological variables" which is in press in Marine Pollution Bulletin. He noted that it is recognized that temporal variations and time trends can be studied more efficiently when a covariance procedure employs a multiple linear regression (MLR) on several biological covariables rather than a simple linear regression on one covariable. However, biological variables employed in trend investigations are generally mutually correlated, as shown in the manuscript. As a remedy for this problem, MLR was introduced using principal components (PCs) of the biological variables as covariates. Advantages of the use of this PC-MLR technique in time trend analysis of the Canadian cod data are:

- the multicollinearity problem is circumvented, and

- MLR was much simpler because only one principal component was needed as the covariate in analysis of burden data and two in the case of concentration data. The paper was well received by the Working Group and a number of members expressed interest in trying the procedure.

7 PROGRESS IN THE ANALYSIS OF TRENDS IN CONTAMINANTS IN SEVERAL SPECIES FROM THE SAME AREA

Dr Misra presented the manuscript, "Monitoring of time trends in contaminant levels using a multispecies approach: Contaminant trends in Atlantic cod (*Gadus morhua*) and European flounder (*Platichthys flesus*) on the Belgian coast, 1978-85." The manuscript has been submitted for publication in the open literature. Biological and statistical reasons were given to support the need for studying temporal variations and time trends based upon the analysis of data on all species which are available from an area rather than analysing data on one species at a time. Procedures for the multivariate and univariate analysis of covariance (MANCOVA and ANCOVA) of joint data on contaminants in cod and flounder from ICES rectangle 31F2 (coast of Belgium) were developed and employed in the study of temporal variations and time trends for the individual species and interactions between them. Analysis of all species from an area at one time is more efficient statistically and supplies other information such as the interactions of time trends in contaminant levels between species. The paper was received without comment by the WG.

8 DRAFT LEAFLETS ON TREND MONITORING TOPICS FOR PUBLICATION IN THE TIMES SERIES

The Chairman raised the issue of inordinate delays in the reviewing of manuscripts within ICES Working Groups. This problem was illustrated with the case of the manuscript, "An introduction to the study of spatial and temporal trends in contaminant levels in marine biota." This manuscript, intended for publication as a Techniques in Marine Environmental Sciences (TIMES) document, was first sent to Working Group members for review shortly after the 1987 WGSATM meeting. The author received no written comments although a number of substantial changes were recommended at the 1988 meeting. A revised manuscript was sent out shortly thereafter but, again, no comments were received prior to the 1989 meeting other than a suggestion; this suggestion and the subsequent handling of the TIMES manuscript paper is considered under items 13.2 and 13.3. Various members suggested that the reason for the delay was the workload on group members which resulted in WGSATM items being put off interessionally. WGSATM therefore suggests adoption of the following procedure:

WGSATM is directly involved in the production of two types of ICES publication: the TIMES Series and the Cooperative Research Reports (CRR). To ensure a better and more timely review, the Working Group suggests that the first draft from an author is discussed at the annual Working Group meeting. Before the meeting is adjourned, the Working Group will decide who among them will

take responsibility for reviewing the manuscript. The number of individuals involved will depend upon the condition and importance of the paper. This small group is to work interessionally through the ICES Environment Officer to expedite review of the manuscript and its ultimate recommendation for publication. WGSATM also recommends that ICES, perhaps in cooperation with the Working Group, ensure that the paper is subject to peer review by at least one person external to the Working Group.

9 USE OF SEAWEEDS IN MONITORING NON-RADIOACTIVE CONTAMINANTS

N. Green presented a paper (Annex 9) in which he and M. Munk Hansen described their respective institutes' experiences in the use of seaweeds as contaminant monitors. In some cases seaweeds have reflected contaminant levels in relation to discharges very well, often better than animals such as mussels. The authors suggest that, in order to establish a set of guidelines for the use of benthic algae for monitoring purposes, a review of existing literature is needed, in particular addressing uptake and release characteristics of trace metals and other contaminants in benthic algae, effects of environmental conditions, e.g., salinity, temperature, on these characteristics, seasonality effects, natural variability and the suitability of different species. The Working Group agreed with the usefulness of seaweeds as monitors and the need for this review, but were unable to find or, indeed, recommend a suitable reviewer. WGSATM requests the assistance of WGEAMS and MCWG in the preparation of the review, with the hope that a set of possible guidelines could be prepared for consideration as soon as possible.

10 REVIEW OF THE ENGLISH VERSION OF THE PAPER BY P. GROS ON STATISTICAL MANAGEMENT OF MARINE ENVIRONMENTAL DATA

As nothing was received from the author on whether or not his paper would be available in an English translation for consideration by the Working Group, the matter was not discussed further.

11 REVIEW OF A PAPER ON THE EFFICIENCY OF POOLING IN MUSSEL WATCH STUDIES

Section 4.2 of this report describes work addressing the effects of sample pooling on the statistics of trend determination. One of the reasons why sample pooling is carried out is as a means of economizing the costly resources required in a monitoring study programme.

Jaap van der Meer presented a paper (Annex 10) discussing how pooling may economize a sampling programme. In this paper he developed generalized models incorporating arbitrary cost estimates for some of the main components in a programme (e.g., sampling and analytical costs) and using the power of statistical tests to optimize the models with respect to pooling strategies, i.e., the number and sizes of pools. It was noted that, in the past, relatively little attention has been paid to the final choice of the number of pools and of the number of organisms comprising a pool. Generally pooling has been expressed in terms

of extremes, e.g., one very large pool (i.e., a single sample homogenate) or no pooling at all (i.e., sampling and analysis of individual specimens). Dr van der Meer showed that, in almost all cases, due to the rather flat appearance in the central part of an optimization function as compared to the steep slopes characterising the two extreme situations noted above, a good rule-of-thumb from an economic point of view is to avoid these extremes.

The Working Group was highly complimentary of this paper and it was suggested that an example, including real data (e.g., real costs) with a clear, non-mathematical description of its use, would provide an important contribution for consideration by ACMP and WGEAMS and possibly also for programme managers involved in deciding sampling strategies and resource allocations, etc.

12 EFFECT OF LABORATORY BIAS ON TEMPORAL TREND STUDIES

"The effect of analytical accuracy on the power of Student's two sample t-test" by M. Nicholson and J. van der Meer was discussed under Agenda item 4.3.

13 SAMPLE STRUCTURE FOR MONITORING CONTAMINANTS IN BIOTA, SEAWATER, AND SEDIMENTS

13.1 Guidelines on the Philosophy, Principles, and Strategy of Monitoring

WGSATM reviewed the "Philosophy, Principles and Strategy of Monitoring", that had been adopted by the ACMP in 1988 (1988 ACMP Report, Section 4.1) and noted the following: (i) The paper should emphasize sampling and data handling to the same extent that data quality is emphasized since both of these factors are as important as data quality (analytical chemistry) in monitoring studies. (ii) The paper should note the differences in comparing environmental levels to a fixed standard (human health purposes) and the determination of temporal and spatial trends. With regard to Table 1, the Working Group was informed about the MCWG consideration of this table and it was suggested that the table be modified along the lines described by MCWG, with a discussion added to support a 'common sense' approach to monitoring. All contaminants should be listed as individual chemical compounds and levels of such chemicals judged to be of concern included in the tables. With respect to trend monitoring, the preferred tissue for monitoring purposes should be named (refer also to Section 14.4.2. of this report). The part of Table 2 (produced by MCWG, ref. 1989 MCWG report) addressing seawater should include a section on suspended particulate matter (SPM), identifying for which contaminants SPM is preferred to water itself as the monitoring medium.

13.2 "An introduction to the study of spatial and temporal trends in contaminant levels in marine biota." by J.F. Uthe et al.

This manuscript was intended for publication as the first of a series of TIMES publications addressing trend monitoring in the marine environment. This was the second draft and was accepted by the Working Group with only minor comment on its content. However the Working Group noted paper 13.3 and recommended that this manuscript and the general item sections of 13.3 be combined into one TIMES document. The senior author agreed to this and, with reference to the comments on review of material for publication (cf. Section 8), the Working Group identified Simon Wilson, Norman Green and Mike Nicholson to review the next revision together with the authors.

13.3 "Sampling strategies for trend monitoring using biota, sediments or seawater" by J.F. Uthe et al.

This paper was prepared under the direction of the Chairman as a response to a request for trend monitoring information from the North Sea Task Force. It was considered by the MCWG at its 1989 Meeting, which made certain suggestions and considered the paper a worthwhile introduction to the topic. WGMS and WGEAMS have also considered the manuscript at their 1989 meetings. It had been proposed that the manuscript, when completed with the addition of information on the use of sediments and seawater as monitors, could become a joint publication of the four environmental Working Groups, WGSATM, WGMS, MCWG, and WGEAMS. WGSATM recommends, rather, that the document be split up, with the general section combined with the document mentioned in the Agenda item 13.2. The sections addressing the specific use of compartments or species, e.g., the section dealing with Atlantic cod (*Gadus morhua*), should then be issued as separate TIMES documents. However, the Working Group recognized the need to respond to the request from the North Sea Task Force and recommends therefore that the manuscript in its present form be forwarded to ACMP for the purpose of information on the use of biota in trend monitoring programmes.

13.4 "Effect of Sampling Structure on Regression Estimates: Why the CMP Sampling Guidelines Should be Followed" by M.D. Nicholson and S.J. Wilson

In the ICES analyses of the CMP data on metals in fish muscle tissue (Cooperative Research Report, No. 162), regression coefficients of log (contaminant concentration) on various biological variables tended, with the exception of mercury, to vary significantly between years. This effect could possibly be the result of poor data structure, where the range of the biological variables sampled in some years was very narrow.

In order to re-appraise this apparent ambiguity, to investigate in more detail the mechanisms which could lead to the spurious heterogeneity of regression slopes, and possibly to improve the objective basis for deciding whether it is necessary to conduct the somewhat involved six-step regression procedure in the statistical analysis of all metals in fish muscle tissue, a simple model which might explain the observations was applied to the CMP data (cf. Annex 11). With the exception of mercury, the results

remained inconclusive. However, the results did emphasize the need for the CMP Guidelines on sampling to be followed so that contaminant levels can be adjusted for biological variation.

14 OTHER BUSINESS

14.1 Papers on Nutrients

G. Radach presented results from the analysis of the Helgoland long-term time series, obtained by the Biologische Anstalt Helgoland (Hamburg). Helgoland is situated in the inner German Bight, at the border of central North Sea water and coastal water. Parameters investigated include temperature; salinity; the nutrients: nitrate, nitrite, ammonia, phosphate and silicate; and the biological entities: phytoplankton carbon together with the biomass of diatoms and flagellates (as carbon equivalents). The data represent samples collected at fixed times once a day on working days, starting from 1 January 1962 until the present.

A variety of methods have been applied to investigate the structure of the time series, especially with respect to trend analysis. Linear regression analysis was carried out for the full 23 years as well as for 5-yearly periods; for winter and summer data; for the (more saline) North Sea water and for (fresher) coastal water. Running 5-year-medians together with the 16.6%- and 83.3%-quantiles were calculated, exhibiting the long-term behaviour of the series.

Spectral analysis of variance showed that the parameters investigated have very different variance distributions, some with nearly all of the variance in the annual frequency (e.g., temperature), many others with most of the variance at frequencies of less than one year. This must have consequences for monitoring strategies for the different parameters.

After the variance analysis, variance for a special frequency band was synthesized again, to obtain the long-term variation of the time series, on scales of more than 2 years. When these time series are plotted against each other in a plane (i.e., the phase plane, such as in an T-S diagram), it becomes obvious that the time series of plankton and nutrients in recent years have moved away from the region of those ecosystem states, that were observed in earlier years up until the late 1970s.

Possibilities were discussed as to utilizing the data set for investigating the power of different monitoring strategies.

The results have partly been published as ICES papers (ICES C.M. 1985/L:2 and ICES C.M.1986/C:8). Further publications are being prepared.

In the ensuing discussion, the value of this unique data set and of the analysis performed was stressed and the differences between the statistical approaches applicable to such long time-series and those which have been utilised by the group until now, aimed at detecting temporal trends in short time-series, were remarked upon. It was felt that for some of the outstanding problems to be addressed, it might be helpful to further evaluate

this data set and to repeat some of the analyses using only the winter data (e.g., January data). Further, it was noted that several groups within ICES are dealing with similar or related problems concerning nutrients, and some coordination of their activities by the Marine Environmental Quality and Hydrography Committees was desirable.

To this end the suggestion that ICES hold a mini-symposium at the Statutory Meeting in 1990 was supported, and a suggested title for the mini-symposium was, "Nutrient distribution, transport and trends in the ICES Area."

Jaap van der Meer briefly described the Dutch programme on nutrient monitoring and the handling of these data to date and introduced the publication, "Monitoring the progress of the attempts to reduce nutrient load and inputs of certain compounds in the North Sea by 50%" by R.W.P.M. Laane, J. van der Meer, A. de Vries and A. van der Giessen, Rijkswaterstaat, Tidal Waters Division, P.O. Box 20904, 2500 EX The Hague, Publication GWA0-89.008.

14.2 Communication from JMG (Dr J. Portmann)

With respect to cooperation between WGSATM and the OSPARCOM ad hoc Working Group on Monitoring, WGSATM noted the following:

WGSATM activities will provide the OSPARCOM ad hoc Working Group on Monitoring with statistical analyses of JMP data on Blue mussels for cadmium, chromium, copper, mercury, lead, zinc, and PCB on a formulation basis for the years 1980-1987. As described in Section 5, WGSATM work on these data have not been finalized, however, it appears to suggest that a very simple, basic presentation of the data is all that can be recommended for the data sets considered. Assuming this to be the case, then if the OSPARCOM ad hoc Working Group on Monitoring wish to include the 1988 data this would be a relatively simple procedure for them.

As discussed in Section 4, it is intended that the WGSATM analyses on the fish liver data (including cadmium and PCB on a formulation basis) concerning the data up to 1987 will be completed in time for the OSPARCOM ad hoc Working Group on Monitoring meeting in December 1989. Extending these analyses to include the 1988 data would involve a major work load on the OSPARCOM ad hoc Working Group on Monitoring. The two WGSATM members involved (Nicholson and Green) should inform the ad hoc Working Group of progress by some appropriate point in time to ensure that, if the completion of these analyses is not possible, the ad hoc Working Group is informed of this fact, and the reasons why. Cadmium and PCB should be considered the priority contaminants for analysis by Nicholson and Green.

Concerning the assessment of contaminants in fish muscle tissue, it has already been agreed that the statistical analyses (covering the years 1978-1985) reported in Cooperative Research Report, No. 162 will be used by the ad hoc Working Group on Monitoring as a basis for its assessment work. Extending the statistical analysis for (only) mercury in fish muscle tissue to include the data up to 1988 does represent a significant work load, just to run and interpret the analyses on the samples concerned. Finding someone to conduct this work may present a problem since

Nicholson and Green are already heavily committed to the liver data analyses.

Extending the fish muscle analyses for other contaminants could be contemplated, but if this is considered then it should be restricted to a simple analysis only, i.e., geometric means.

With reference to the questions raised regarding the Dutch (pooled) fish muscle tissue data, the comments stated in the preceeding two paragraphs apply. WGSATM was unable to confirm that it could take on the analysis of the Dutch pooled fish muscle tissue data and considered that the Dutch might be able to identify someone to conduct the analyses of these data, using the six-step regression approach described in Cooperative Research Report, No. 162 for mercury at least.

In relation to the proposal by the Netherlands delegation at the OSPARCOM's Joint Monitoring Group meeting concerning the statistical analysis of seawater data for determination of temporal trends, if the OSPARCOM ad hoc Working Group on Monitoring intends to analyse seawater data, the WGSATM was not able to identify any individual to take on this task. Indeed, given that MCWG questions whether trends in seawater can usefully be addressed using the current JMG monitoring data, it might seem somewhat inconsistent to recommend ICES association with such an activity.

It should also be noted that WGSATM's experience has shown that, in general, the trend monitoring data cannot be properly assessed using a "standard analysis" approach (partially due to data quality considerations). The liver tissue analyses will provide/suggest a possible general statistical approach for looking at these data, but this has to be applied carefully by a statistician who is capable/willing to consider each data analysis individually, that is, analysing trend monitoring data is a labour-intensive, time-consuming activity. The JMG should be made aware of this fact and decide how to deal with it, because responsibility for this activity will eventually fall on the JMG if the WGSATM are to be released to go on to consider other statistical problems.

With respect to the request that the current guidelines (see Coop.Res.Rep. No. 126 (1984)) for sampling and analysis in the use of biota for temporal trend assessment purposes (ICES Objective 3) and spatial distribution purposes (ICES Objective 2) be combined in some way, WGSATM notes the following. Samples collected under Objective 2 are taken only every five years and are designed to give a synoptic geographical overview within the ICES area. Samples are to be selected to be as representative of the area in question as possible with the intent that a large number of sample sites be visited.

Samples collected under Objective 3 are taken annually and are designed to give a suitable data set for quantitative trend analysis. When collected from the same area as samples under Objective 2, it is important that the same, i.e., stable, stock be sampled each year under Objective 3. The importance of fish stock is emphasised under item (e) within the Objective 3 guidelines. While it is recommended that certain geographical areas and species be covered it is important to emphasise the stable stock aspect. There is little point in studying unstable stocks or at-

tempting trend monitoring of the type visualized under Objective 3 with a large number of species from many different areas, particularly in light of the difficulties encountered with the quality of the data sets to date.

Thus, Objectives 2 and 3 are distinctly and importantly different. The only compromise that might be made is to state that a proper sample taken for Objective 3 can be used to estimate a single value for use under Objective 2, an example of the "common sense" application of monitoring guidelines.

However, WGSATM noted that the sampling and analytical guidelines appropriate to Objectives 2 and 3 may need to be modified in light of its investigations carried out on CMP data. Therefore, a small sub-group of WGSATM, comprising Jaap van der Meer (Chairman), Norman Green and Simon Wilson, has been asked to consider these guidelines intersessionally and report to the 1990 meeting of WGSATM. It is envisioned that any recommended changes to the guidelines will be of a nature to aid data analysis, with both past and future data in mind.

14.3 Long-term Trends in Laboratory Analysis

The Chairman reported on discussions of this item within MCWG and noted that WGSATM is also interested in addressing long-term trends in laboratory analytical performance. Dr John Calder of the National Science Foundation in Washington has supplied WGSATM with a set of data (1983-1987) from one of the NOAA laboratories. This information has been passed to Mike Nicholson for consideration within his investigations of the effects of laboratory analytical factors on the determination of trends.

14.4 Items from WGEAMS

A number of items related to WGEAMS were raised during the meeting and were commented on as follows:

14.4.1 Guidelines

In addition to the aspects discussed under Section 14.2 (above), WGSATM reconsidered earlier work which it had carried out in relation to possible modifications to the existing guidelines for monitoring using mussels for the purpose of determining temporal trends (see 1987 WGSATM report, C.M.1987/E:24, Annex 9). This original proposal for a modification to the guidelines had invited comments, requested through ACMP, but none had been received. However, in the light of papers presented at the 1989 WGSATM meeting on the effect of pooling and sample depuration, etc. (cf. Sections 4.2, 5.2 and 11), the group considered that it might be worthwhile to consider the feasibility of finalising the revised guidelines, for resubmission to ACMP as a completed project.

The work reported on the effects of pooling (see Section 4.2 and Annex 4) confirm that, for log-normally distributed contaminants pooling reduces the ability to detect trends. However, this negative consequence has to be considered in relation to the levels of change which the mussel component in a particular monitoring programme is attempting to identify. A second conclusion, that

pooling introduces a bias in the estimated levels, does not affect the ability to detect temporal trends providing pooling is consistent between years, i.e., the number of specimens in a pool is kept constant, as is recommended in the proposed guidelines.

A depuration study (see Section 5.2 and Annex 6) had similarly confirmed that the proposed guidelines were appropriate, i.e., that depuration may be omitted, but that the effect of depuration should be studied in each individual case under consideration.

It was, therefore, considered appropriate that the mussel guidelines be recommended for further consideration by WGEAMS or ACMF for eventual incorporation in the monitoring guidelines, but with a note to the effect that, in cases where temporal trend monitoring using mussels is already underway, it will be necessary for the institute/organization involved to consider whether they should adopt the new guidelines or continue with their existing procedures in order to ensure that their time series is not disturbed.

14.4.2 Contaminant Matrices Related to the JMP Purposes

Further to their discussion on the original matrix tables reported under Section 13.1, WGSATM received, during the course of the meeting, a revised set of three WGEAMS matrix tables. WGSATM noted that the new tables differed somewhat from both the original tables and the suggested modifications produced by MCWG (see 1989 MCWG report, C.K.1989/C:32). It was not clear to WGSATM why a material such as dieldrin had been excluded from the human health purpose table nor why lindane (γ -HCH) was included and not α - or β -HCH in the two other tables. It was suggested that these tables be reconsidered in terms of new contaminants, that all organic listings refer to specific chemical compounds, and that the guidelines be re-examined in relation to the assessment purposes. It is further suggested that the need for identifying "reference" stations in selected areas be assessed along with the need for adequate sample storing ("tissue banking") procedures.

The group further noted that there is some doubt as to the utility of biota as monitors of low-level environmental changes in contaminant levels, i.e., where the influence of such environmentally induced changes in contaminant levels within the biota is of the same magnitude as changes brought about by other environmental and biological factors. In spite of this, the Working Group does not feel that a wholesale shift to water or sediments as the monitors of environmental contamination is warranted since we do not, as yet, have enough good (complete) data upon which to judge the utility of the various abiotic sectors as monitors of environmental change. As an example, the recommendation for normalization in the use of sediments (C.M.1989/E:2, Annex 4) is similar to the requirement for 'normalizing' fish (e.g., the use of a standard-sized fish for adjusting mean contaminant estimates for length effects, first suggested by WGMFNA in 1981); both are procedures that have proven themselves less than simple in application. Obviously, each compartment can be used to monitor changes in contaminant levels within that particular compartment; it is the extrapolation to environmental contaminant levels that is problematical.

Due to the need for rapid response to WGEAMS's request, a telefax (Annex 12) was forwarded. In addition to the above comments, it is noted that:

- 1) Changes imposed on trend monitoring schemes (e.g., a restriction in the number of contaminants monitored in a particular biological tissue) could destroy WGSATM's potential for analysing past-future data sets for trends by multivariate analysis (such as the analysis presented in Annex 8).
- 2) Present trend studies are in their adolescence and properly executed, ongoing studies should continue.
- 3) There is a plea for research into elucidating and modelling processes, a need for identifying "reference stations" for intensive study, and a possible need for appropriate specimen banking.

14.4.3 Data Quality and Level of Detectable Change in Contaminant Level

Using mercury in cod liver and fish muscle as an example, some observations on the data quality needed for detecting a certain level of change are included in Annex 12; the results of this basic example are summarised as a table in Annex 12.

14.5 Proposal for Chairman for the period, 1990-1992

Following nomination by Mike Nicholson, which was unanimously supported, and lacking any other nomination, the group recommended that the present Chairman, Dr. Uthe, be appointed to serve for the next three-year period.

15 RECOMMENDATIONS

- 1) WGSATM recommends that a Mini-symposium be held at the 1990 Statutory Meeting on the theme, "Nutrient distribution, transport and trends in the ICES Area."
- 2) WGSATM recommends that the paper, Annex 5, "The effect of analytical accuracy on the power of student's t-test" by M.D. Nicholson and J. van der Meer be brought to the specific attention of all groups concerned with the monitoring of chemical contaminants in the marine environment.
- 3) WGSATM recommends that the paper, Annex 10, "Pooling may economize a sampling program" by Jaap van der Meer be brought to the specific attention of all groups involved in the management of marine contaminant monitoring programmes.
- 4) Because of the number of items referred to WGSATM by WGEAMS, WGSATM recommends that WGEAMS meet 1-2 weeks prior to the annual meeting of WGSATM.

- 5) WGSATM recommends that a review on the usefulness of marine plants as contaminant monitors be prepared by WGEAMS and MCWG as soon as possible.
- 6) WGSATM recommends that the mussel sampling guidelines be considered without modification for inclusion into the monitoring guidelines with the caveat noted under Section 14.4.1 that ongoing mussel studies should not be altered without careful consideration of the effects of such alterations on the interpretability of the study.
- 7) The Working Group on Statistical Aspects of Trend Monitoring should meet at ICES Headquarters on 23-27 April 1990 to consider the following topics:
 - a) Review of the relevant sections 1989 ACMP report and the report of the 1989 meeting of the Working Group on Environmental Assessments and Monitoring Strategies;
 - b) Completion of the report on the study of trends in contaminant levels in fish liver;
 - c) Completion of the report on the study of CMP data for trends in contaminants in blue mussel;
 - d) Consideration of revisions to sampling and analytical guidelines for ICES Objectives 2 and 3;
 - e) Consideration of papers on the effects of laboratory performance on contaminant trends studies; and
 - f) Consideration of papers addressing trends in other substances, such as nutrients, and other compartments, such as seawater, sediments and suspended particulate matter.

16 CLOSING OF THE MEETING

The Chairman thanked Dr Jaap van der Meer and the Tidal Waters Division of Rijkswaterstaat for hosting the meeting. He then complimented the group on the intensity and productivity of the members both at the meeting and intersessionally. The meeting closed at 16:30 on 27 April 1989.

ANNEX 1

WORKING GROUP ON THE STATISTICAL ASPECTS OF TREND MONITORING
The Hague, 24 - 27 April 1989

List of Participants

Anders Bignert	The Swedish Museum of Natural History P.O. Box 50007 104 05 Stockholm SWEDEN Tel. 46-8-6664115 Fax. 46-8-6664212
Rob Fryer	D.A.F.S. Marine Laboratory P.O. Box 101 Victoria Road Aberdeen AB9 8DB UNITED KINGDOM
Norman W. Green	Norwegian Institute for Water Research P.O. Box 33 Blindern 0883-Oslo-8 NORWAY Tel. 47(2)235280 Fax. 47(2)394189
Jaap van der Meer	Tidal Waters Division Rijkswaterstaat P.O. Box 20907 2500 EX Den Haag THE NETHERLANDS Tel. 070-264101 Fax. 070-241293
Raj K. Misra	Department of Fisheries & Oceans P.O. Box 550 Halifax, N.S. B3J 2S7 CANADA Tel. (902) 426-6208
Mike Nicholson	Fisheries Laboratory Lowestoft NR33 OHT Suffolk UNITED KINGDOM Tel. 0502 562244
Palle Bo Nielsen	Grønlands Fiskeri- og Miljøundersøgelser Tagensvej 135 DK-2200 Copenhagen N DENMARK

Günther Radach

Institut für Meereskunde der
Universität Hamburg
Troplowitzstrasse 7
D-2000 Hamburg 54
FEDERAL REPUBLIC OF GERMANY
Tel. 040-4123-5656
Fax. 040-4123-4644
Telex 272586 IFMH D
Telemail IFM.HAMBURG

Jack Uthe (Chairman)

Department of Fisheries & Oceans
P.O. Box 550,
Halifax, N.S. B3J 2S7
CANADA
Tel. (902) 426-6277

Simon Wilson

ICES
Palægade 2-4
DK-1261 Copenhagen K
DENMARK
Tel. 33 15 42 25
Fax. 01 93 42 15
Telex 22498 ices dk

ANNEX 2

WORKING GROUP ON THE STATISTICAL ASPECTS OF TREND MONITORING
The Hague, 24-27 April 1989

AGENDA

1. Opening of the Meeting.
2. Adoption of the Agenda.
3. Review of the ACMP Report.
4. Result of analysis of CMP data fro trends in contaminants in fish liver tissue.
5. Results of analysis of CMP data for trends in contaminants in blue mussel.
6. Consideration of the results of trend studies in Canadian Atlantic cod for 1977-1985 by univariate and multivariate analyses.
7. Progress in the analysis of trends in contaminants in several species from the same area.
8. Draft leaflets on trend monitoring topics for publication in TIMES.
9. Use of seaweeds in monitoring non-radioactive contaminants.
10. Review of the English version of the paper by P. Gros on statistical management of marine environmental data.
11. Review of a paper on the efficiency of pooling in Mussel Watch Studies.
12. Effect of laboratory bias on temporal trend studies.
13. Sample structure for monitoring contaminants in biota, sea water and sediments.
14. Other business.
15. Recommendations.
16. Closing of the Meeting.

---oo0oo---

ANNEX 3**ASSESSMENT OF TRENDS IN CONTAMINANTS IN FISH LIVERS**
Part I: Evaluating Important Biological Covariables

by

Mike Nicholson, Norman Green, and Simon Wilson

FORWARD

The Working Group on the Statistical Aspects of Trend Monitoring which met in Copenhagen 12-29 April recommended among other points to "complete the study of the CMP data for trends in contaminants in fish liver tissue" (C.M. 1988/E:27, item 11.1). It was requested that this be investigated by the authors of this report.

INTRODUCTION

Fish livers are often used when monitoring trends in heavy metals and chlorinated hydrocarbons for several reasons but mainly because this organ is readily obtainable and tends to accumulate a number of contaminants. However the binding and sequestration capacities of each of these or other contaminants in livers, either alone or in combination with each other is far from known. Phillips (1980) has written an extensive review qualifying the problems related to monitoring contaminants in the organism and stress the importance of measuring lipid in the liver. Further, organochlorines are particularly lipophilic and conversion of concentrations from a wet weight basis to a fat weight basis often greatly reduces the variation amongst samples. However lipid content, let alone lipid type, is often not measured in conjunction with the contaminant. Hence, lipid content is paramount in assessing contaminant found in organism but may vary considerable in nature for example according to: organ tissue, sex, age, trophic level, season, water temperature, etc.. Lipid content is often only crudely measured by weighing the extractable solids as a convenient step in the organochlorine analysis. Some contaminants such as DDT and Endrin even cause lipogenesis and enhance uptake of the these contaminants. He (Phillips, 1985) emphasizes that concentration based solely on a wet weight basis may be only relevant where the change in lipid content is small.

MATERIALS AND METHODS

Selected data

Seven ICES/JMG data sets were selected including five ICES rectangles in the North Sea region 1981-1988. The selection criterion was that measurements of fish length, liver weight and fat percent were reported (Table 1). The data sets included two round fish (Atlantic cod, *Gadus morhua* (=GADU)) and whiting, *Merlangius merlangus* (=MERL)) and two flat fish (European flounder, *Platichthys flesus* (=PLAT), and European plaice, *Pleuronectes platessa* (=PLEU)). There was considerable variation in the biological variables length,

liver (tissue) weight and (liver) fat % and weight (Table 2).

These data sets were analyzed for those contaminants which were expressed, or could be expressed, on a wet weight basis. The most abundant data is found for concentrations of PCBs and cadmium, and less so for copper, lead, mercury and zinc (Table 3).

Analysis

Based on previous studies (C.M.1988/E.27) fish length (L), liver weight (W) and fat percent (F) appear to be potentially important covariables for assessing trends in contaminants (C) PCB and cadmium in fish liver. Table 4 presents the results from fitting the model:

$$\ln C = a_{\text{year}} + a_1 \ln L + a_2 \ln W + a_3 \ln F + \text{error}$$

where the error is assumed to be Normally distributed. The table shows the residual sums of squares (RSS) from fitting a null model and models with years and all combinations one, two and three biological variables from all data sets. Residual sums of squares not significantly greater than that from the full model are indicated. Subjectively, this table can be used to identify consistently good sub-models across data sets.

The data were checked for outliers.

RESULTS and DISCUSSION

PCBs

Six data sets were selected. For cod collected from the Oslofjord of Norway (ICES rectangles 47G0 and 48G0) all the biological variables fish length, liver weight and liver fat percent were important for predicting PCBs in cod liver (Table 4, shown by the analysis of covariance that show that the removal of any covariable significantly increases the RSS). Estimates of PCB in cod from the east coast of England indicated that only length was necessary. The differences between the Norwegian and United Kingdom data may in part be due to the differences in lipid content. The Norwegian cod were collected from November to February and the lipid content (weight) of some individuals was low compared to the United Kingdom data which collected fish from July to September (cf., Table 2.4, pers. comm. Simon Wilson, ICES).

The lack of need of biological indicators for whiting and flounder is probably due low values and narrow range of liver - both tissue weight and fat - and of length. The

sample sizes for these species was also small (cf., Table 3).

Cadmium

Of the four data sets selected the two for flounder contained insufficient data. The remaining two data sets for cod (from the Oslofjord) indicated that liver weight was important together with either of length or liver fat.

Copper, Mercury and Lead

Generally, there was no consistent evidence that concentrations of these metals are affected by length, liver weight or liver fat.

Zinc

There is evidence that length may be omitted from the model. Liver weight and liver fat both demonstrate some effect, but this is not consistent.

REFERENCES

Phillips, D.J.H., 1980. Quantitative Aquatic Biological Indicators (dedicated to Charlie Boyden, used by champions) Applied Science Publishers Ltd. London. 480pp..

Nicholson, M.D., Green, N.W. and Wilson, S.J., 1988. Regression models for assessing trends in cadmium and PCBs in cod liver. Report of the 1988 meeting of the Working Group on Statistical Aspects of Trend Monitoring. ICES C.M.1988/E:27, Annex 7. pp.62-68 (mimeo).

Table 1 : No's of obs by years for which tissue weight, %fat and length are present

SAMPLE	YEAR							
	1981	1982	1983	1984	1985	1986	1987	1988
	N	N	N	N	N	N	N	N
GADU34F2	.	20	25	25	25	25	25	.
GADU47G0	10	.	23	.	14	.	25	24
GADU48G0	9	18	.	29	25	25	25	.
MERL35E6	.	.	10	10
PLAT31F3	.	.	.	19	3	15	12	.
PLAT35F5	.	.	.	26	12	10	8	.
PLAT35F6	8	25	.
PLAT36F6	.	.	.	5	.	3	16	.
PLAT37F8	25	25	.
PLAT48G0	8
PLEU34F2	.	13	15	15	15	10	.	.

Table 2 : Overall summary statistics for selected data

Table 2.1 : Length (mm)

SAMPLE	MEAN FISH LENGTH				
	N	MIN	MAX	MEAN	STD
GADU34F2	145	310.000	830.000	506.310	126.962
GADU47G0	96	300.000	1000.000	492.969	110.545
GADU48G0	131	280.000	750.000	420.687	72.025
MERL35E6	20	262.000	352.000	298.150	26.470
PLAT31F3	49	276.000	399.000	326.857	29.229
PLAT35F5	56	231.000	353.000	299.964	33.296
PLAT35F6	33	170.000	380.000	256.848	55.654
PLAT36F6	24	226.000	344.000	281.167	25.139
PLAT37F8	50	150.000	380.000	263.280	56.017
PLAT48G0	8	300.000	630.000	381.250	106.293
PLEU34F2	68	311.000	600.000	419.985	66.225

Table 2.2 : Liver wt (gms)

SAMPLE	LIVER WEIGHT				
	N	MIN	MAX	MEAN	STD
GADU34F2	145	3.700	303.700	52.296	55.944
GADU47G0	96	2.500	163.300	22.628	25.385
GADU48G0	131	2.600	106.800	25.705	22.589
MERL35E6	20	8.400	22.600	13.325	3.440
PLAT31F3	49	1.600	14.400	7.165	3.090
PLAT35F5	56	0.100	8.800	3.629	2.119
PLAT35F6	33	0.460	5.920	2.390	1.352
PLAT36F6	24	0.700	8.900	4.379	2.094
PLAT37F8	50	0.330	62.500	4.906	8.766
PLAT48G0	8	4.000	12.000	7.625	3.021
PLEU34F2	68	1.800	75.400	17.712	12.361

Table 2 (cont'd)

Table 2.3 : %fat in liver

SAMPLE	LIVER FAT %				
	N	MIN	MAX	MEAN	STD
GADU34F2	145	4.700	68.000	34.320	16.618
GADU47G0	96	0.100	67.100	23.523	18.473
GADU48G0	131	0.100	83.200	48.980	20.648
MERL35E6	20	37.000	69.500	58.055	9.763
PLAT31F3	49	2.700	32.000	18.592	7.289
PLAT35F5	56	1.000	34.600	10.116	7.135
PLAT35F6	33	2.110	21.270	8.308	5.097
PLAT36F6	24	6.600	37.500	23.175	7.369
PLAT37F8	50	0.230	25.140	7.029	5.679
PLAT48G0	8	5.500	23.300	13.213	6.727
PLEU34F2	68	0.670	34.000	8.819	7.277

Table 2.4 : Liver fat wt (gms)

SAMPLE	LIVER FAT WEIGHT				
	N	MIN	MAX	MEAN	STD
GADU34F2	145	0.185	206.516	23.737	32.891
GADU47G0	96	0.009	74.200	7.652	12.411
GADU48G0	131	0.005	74.226	14.940	15.584
MERL35E6	20	3.870	13.108	7.845	2.668
PLAT31F3	49	0.084	4.608	1.438	0.992
PLAT35F5	56	0.010	2.284	0.434	0.477
PLAT35F6	33	0.020	0.833	0.198	0.189
PLAT36F6	24	0.073	2.363	1.074	0.659
PLAT37F8	50	0.001	1.956	0.293	0.356
PLAT48G0	8	0.330	2.256	1.006	0.662
PLEU34F2	68	0.015	7.242	1.567	1.464

Table 3 : No's of obs by years expressed on a wet liver wt basis for selected data

Table 3.1 : PCBs (mg/kg wet wt)

	YEAR							
	1981	1982	1983	1984	1985	1986	1987	1988
SAMPLE	N	N	N	N	N	N	N	N
GADU34F2	.	20	25	25	25	25	25	.
GADU47G0	10	.	23	.	14	.	25	24
GADU48G0	9	18	.	29	25	24	25	.
MERL35E6	.	.	10	10
PLAT48G0	8
PLEU34F2	.	13	15	15	15	10	.	.

Table 3.2 : Cadmium (mg/kg wet wt)

	YEAR							
	1981	1982	1983	1984	1985	1986	1987	1988
SAMPLE	N	N	N	N	N	N	N	N
GADU47G0	10	.	23	.	14	.	25	24
GADU48G0	9	18	.	29	25	25	25	.
PLAT31F3	.	.	.	4	3	1	2	.
PLAT35F5	.	.	.	1	4	.	.	.
PLAT36F6	1	1	.
PLAT48G0	8

Table 3.3 : Copper (mg/kg wet wt)

	YEAR		
	1986	1987	1988
SAMPLE	N	N	N
GADU47G0	.	25	24
GADU48G0	25	25	.

Table 3 (cont'd)

Table 3.4 : Lead (mg/kg wet wt)

	YEAR		
	1986	1987	1988
	N	N	N
SAMPLE			
GADU47G0	.	25	24
GADU48G0	25	25	.

Table 3.5 : Mercury (mg/kg wet wt)

	YEAR	
	1981	1982
	N	N
SAMPLE		
GADU47G0	10	.
GADU48G0	5	18

Table 3.6 : Zinc (mg/kg wet wt)

	YEAR		
	1986	1987	1988
	N	N	N
SAMPLE			
GADU47G0	.	25	24
GADU48G0	25	25	.

Table 4 Results of analyses of covariance. Bold type indicates models not significantly worse than the model of log(contaminant) on years, length, log(liver weight) and log(fat percent).

4.1 PCBs

Model	Species/Area						
	Gadu 34F2	Gadu 47G0	Gadu 48G0	Merl 35E6	Plat 31F3	Plat 48G0	Pleu 34F2
-	123	135	135	3.4	-	4.3	111
year	73	96	94	3.1	-	4.3	90
year + l	50	71	91	2.7	-	4.1	90
year + w	52	66	85	3.1	-	4.2	90
year + f	59	41	48	3.0	-	3.5	70
year + l + f	49	32	46	2.5	-	3.5	70
year + l + w	50	65	85	2.7	-	3.7	87
year + f + w	52	40	48	2.8	-	3.5	70
year + l + w + f	48	29	41	2.2	-	3.3	69
residual d.f.	135	88	121	15	-	4	60

4.2 Cadmium

Model	Species/Area						
	Gadu 34F2	Gadu 47G0	Gadu 48G0	Merl 35E6	Plat 31F3	Plat 48G0	Pleu 34F2
-	-	77	133	-	3.1	1.9	-
year	-	51	78	-	1.5	1.9	-
year + l	-	47	78	-	1.1	1.8	-
year + w	-	39	72	-	0.8	1.8	-
year + f	-	39	72	-	1.5	1.7	-
year + l + f	-	38	72	-	1.1	1.6	-
year + l + w	-	36	68	-	0.8	1.8	-
year + f + w	-	35	70	-	0.8	1.7	-
year + l + w + f	-	35	67	-	0.8	1.6	-
residual d.f.	-	88	122	-	3	4	-

4.3 Copper

Model	Species/Area						
	Gadu 34F2	Gadu 47G0	Gadu 48G0	Merl 35E6	Plat 31F3	Plat 48G0	Pleu 34F2
-	-	12	36	-	-	-	-
year	-	12	9.8	-	-	-	-
year + l	-	11	9.8	-	-	-	-
year + w	-	11	8.5	-	-	-	-
year + f	-	11	9.1	-	-	-	-
year + l + f	-	11	9.1	-	-	-	-
year + l + w	-	11	8.3	-	-	-	-
year + f + w	-	11	8.5	-	-	-	-
year + l + w + f	-	11	8.3	-	-	-	-
residual d.f.	-	44	45	-	-	-	-

4.4 Mercury

Model	Species/Area						
	Gadu 34F2	Gadu 47G0	Gadu 48G0	Merl 35E6	Plat 31F3	Plat 48G0	Pleu 34F2
-	-	2.2	11	-	-	-	-
year	-	2.2	10	-	-	-	-
year + l	-	1.9	9.9	-	-	-	-
year + w	-	1.8	10	-	-	-	-
year + f	-	1.5	9.0	-	-	-	-
year + l + f	-	1.3	8.6	-	-	-	-
year + l + w	-	1.1	9.4	-	-	-	-
year + f + w	-	1.5	8.9	-	-	-	-
year + l + w + f	-	1.1	8.5	-	-	-	-
residual d.f. t	-	6	18	-	-	-	-

4.5 Lead

Model	Species/Area						
	Gadu 34F2	Gadu 47G0	Gadu 48G0	Merl 35E6	Plat 31F3	Plat 48G0	Pleu 34F2
-	-	14	27	-	-	-	-
year	-	6.6	27	-	-	-	-
year + l	-	6.1	25	-	-	-	-
year + w	-	5.5	26	-	-	-	-
year + f	-	5.6	26	-	-	-	-
year + l + f	-	5.2	24	-	-	-	-
year + l + w	-	5.5	25	-	-	-	-
year + f + w	-	5.2	26	-	-	-	-
year + l + w + f	-	5.1	24	-	-	-	-
residual d.f.	-	44	45	-	-	-	-

4.6 Zinc

Model	Species/Area						
	Gadu 34F2	Gadu 47G0	Gadu 48G0	Merl 35E6	Plat 31F3	Plat 48G0	Pleu 34F2
-	-	5.4	34	-	-	-	-
year	-	4.7	7.5	-	-	-	-
year + l	-	4.7	7.2	-	-	-	-
year + w	-	4.3	3.5	-	-	-	-
year + f	-	3.5	5.0	-	-	-	-
year + l + f	-	3.5	4.9	-	-	-	-
year + l + w	-	4.2	3.3	-	-	-	-
year + f + w	-	3.5	3.4	-	-	-	-
year + l + w + f	-	3.5	3.3	-	-	-	-
residual d.f.	-	44	45	-	-	-	-

ANNEX 4

TREND MEASUREMENT FROM BULKED SAMPLES:

The Effect of Pooling on Lognormally Distributed Contaminants

by

Mike Nicholson, Norman Green, and Simon Wilson

Summary

The effect of bulking tissue in which contaminant levels are lognormally distributed is shown to be an increase in the mean and a decrease in the power of tests to detect trends. The estimated trends and associated regression coefficients should be unaffected. These results are demonstrated by computer simulation using a model derived from levels of PCBs in cod livers.

1. Introduction.

When measuring contaminant levels in fish tissues (e.g. muscle, liver), the tissue from several fish is sometimes bulked, and the contaminant concentration determined from the homogenised sample. This may be done to provide sufficient material for chemical analysis, or to reduce analytical costs.

However, this procedure could affect the statistical analysis of the data and the estimated mean contaminant concentrations. This note looks at the effect of bulking when concentrations within a tissue type are lognormally distributed.

2. The Effect of Bulking on Lognormally Distributed Concentrations

Suppose x has a lognormal distribution i.e.

$$y = \log(x) \text{ is distributed } N(\mu, \sigma^2).$$

Then, following Aitchison and Brown (1957)

$$E[x] = \exp(\mu + \sigma^2/2)$$

and

$$\begin{aligned} V[x] &= E[x]^2(\exp(\sigma^2) - 1) \\ &= E[x]^2\sigma^2 \quad \text{if } \sigma^2 \ll 1. \end{aligned}$$

Let us assume that we wish to estimate μ . The usual procedure is to analyse the logarithms of the observed x 's; i.e. from a random sample

$$x_1 \ x_2 \ x_3 \ x_4 \ \dots \ x_n$$

we calculate

$$y_i = \log(x_i)$$

and estimate μ by

$$y = \Sigma y_i / n.$$

Suppose the n samples are randomly aggregated into b bulked samples each consisting of r individuals; i.e.

$$n = rb.$$

If x_{ij} is the concentration of the i 'th individual in the j 'th homogenate and w_{ij} is the corresponding weight of tissue, the concentration in the j 'th bulked sample is

$$x_j = \frac{\Sigma w_{ij} x_{ij}}{\Sigma w_{ij}}.$$

Writing y_b for the mean logarithm of the concentrations in bulked samples, then

$$y_b = \Sigma y_j / b$$

where

$$y_j = \log(x_j).$$

The expected value of y_b is given by

$$E[y_b] = \mu + \sigma^2 \sum_j [1 - \frac{\Sigma w_{ij}^2}{(\Sigma w_{ij})^2}] / 2b$$

In the simple case where the w_{ij} are identical, $E[y_b]$ is given by

$$E[y_b] = \mu + \sigma^2(1-r^{-1})/2$$

and the variance by

$$V[y_b] = \frac{\sigma^2}{rb} \\ = \frac{\sigma^2}{n}.$$

Hence y_b is a biased estimator of μ when the sample is bulked. The bias is positive and increases with the degree of pooling. The standard error of y_b is not affected.

Note that when r equals 1 (no bulking) y_b is the logarithm of the geometric mean. When r equals n (one homogenate) y_b is the logarithm of the arithmetic mean.

3. Bulking Within Groups Defined by a Covariate.

Sometimes the concentration is related to some biological variable measured on the sampling unit. For example, suppose $\log(\text{concentration})$ is linearly related to the length of the organism i.e.

$$y = a + \beta l.$$

In this case, the samples are often bulked according to the value of l .

Suppose we have data

$$l_1 \quad x_{11} \dots x_{r1}$$

$$l_2 \quad x_{12} \dots x_{r2}$$

$$\dots \dots \dots$$

$$l_b \quad x_{1b} \dots x_{rb}$$

and the replicates are bulked at each length. Again,

$$y_j = \log(x_j)$$

where

$$x_j = \frac{\sum_r w_{ij} x_{ij}}{\sum_r w_{ij}}.$$

and, if the w_{ij} are identical,

$$E[y_j] = \alpha + \beta l + s^2(1-r^{-1})/2$$

where σ^2 is the residual variance. Hence the intercept is affected by bulking, but not the slope.

These results are demonstrated by a computer simulation study. 25 samples of a lognormally distributed variable with

$$E[\log(x_{ij})] = 1 + l$$

$$V[\log(x_{ij})|l] = 0.5^2$$

and

$$w_{ij} = w$$

were generated in blocks of 5 at each of

$$l = 1, 2, 3, 4 \text{ and } 5.$$

The intercept and slope from a regression of $\log(x)$ on l were estimated both from the individual data and from the "bulked" data. This process was repeated 500 times. The means and standard deviations of the empirical distributions of the estimates were

	intercept	st.dev.	slope	st.dev.
Individual data	1.01	0.22	1.00	0.07
Bulked data	1.10	0.23	1.00	0.07

The empirical standard deviations are very similar, and as predicted, so are the slopes. The intercept from the bulked analysis has increased, and equals the predicted value

$$1 + 0.5^2(1-5^{-1})/2 = 1.1.$$

4. Application to Trends in PCBs Measured in Bulk Cod Livers.

In practice, the w_{ij} will not be identical. Also, bulking will be for fish of similar, but not identical, lengths.

For example, within the ICES Coordinated Monitoring Programme, samples are length stratified and consist of 25 fish, 5 in each of 5 length intervals defined to have equal width on a logarithmic scale. Liver weights vary and tend to increase with fish length. For some contaminants, e.g. PCBs, concentration is also related to the amount of fat in the liver (Nicholson et al, 1988).

Since the range of the biological variables will be reduced by bulking, its effect should be to reduce the precision of an analysis of covariance using these variables.

To assess the effect of bulking on this monitoring programme a second computer simulation study was carried out. Using Norwegian cod data collected in the inner Oslofjord, the following relationships were established for concentrations of PCBs in the cod liver:

$$\log(\text{concentration}) = a + 0.006l - 0.64\log(w) + 0.70\log(\%f) + \epsilon_c$$

ϵ_c distributed $N(0, 0.582)$

$$\log(w) = -2.8 + 0.019l - 0.000013l^2 + \epsilon_w$$

ϵ_w distributed $N(0, 0.602)$

$$\%f = 100(-2.5 + 0.68w + \epsilon_f)/w$$

ϵ_f distributed $N(0, 3.02)$

Where concentration is in mg/kg wet liver weight, l is length in millimetres, w is liver wet weight in grams and $\%f$ is the percentage liver fat.

These relationships were used to simulate a sampling programme to measure the difference in contaminant levels between two years.

The intercept, a , was set to give an average concentration of PCBs of 0.10 mg/kg in the first year, and values in the range 0.10 to 0.27 mg/kg in the second year.

An annual sample consisted of 25 fish, five per length group. The length groups had equal widths on a log-scale, with end points (280, 341, 425, 506, 616, 750). Lengths within a length group were generated from a uniform distribution followed by the corresponding liver weights, percentage liver fats and concentrations of PCBs.

A sample was generated for each of two years and an analysis of covariance of $\log(\text{concentration})$ on l , $\log(w)$, $\log(\%f)$ and years carried out on the individual data and on the data bulked within length groups. The estimated parameters were the intercept in the first year, the difference in the intercepts (trend), and the regression coefficients for l , $\log(w)$ and $\log(\%f)$.

This was repeated 100 times for a range of differences between the mean concentrations in each year.

5. Results of Simulation of Trend Measurements on bulked Data.

Bulking effects were measured in two ways: the first looked at changes in the power of the test for the presence of a year effect; the second looked at bias and precision of the estimated parameters.

Figure 1 shows the power at a 95% significance level for the analyses of the individual and bulked data. An empirically smoothed curve has been drawn through the simulated values.

Figure 2 shows the mean of the estimates of the intercept in the first year banded by 95% tolerance limits.

Figure 3 shows the means of the estimated difference in the intercepts banded by 95% tolerance limits.

Figure 4 shows the means of, for example, the estimated length regression coefficient and 95% tolerance limits.

6. Discussion.

The simulated results of Section 5 agree with the theoretical results of Sections 2 and 3. Trends and regression coefficients have been estimated without bias, but the mean contaminant level has increased due to bulking.

Power is reduced by bulking. As revealed by the widths of the empirical tolerance intervals of the estimated parameters, the reduction due to the decrease in the degrees of freedom is made worse by the effect of bulking the biological variables.

These results agree with an assessment of the effect of bulking on trends in cadmium levels in lobsters (Anon, 1986). This study randomly sub-sampled a large data set to compare results from individual analyses with those from two degrees of bulking. The power of significance tests of year effects and regressions of log(length) were affected, but as here, trends and regression coefficients were estimated without bias. However, because of the reduced power of the tests, the trends and regression coefficients were more likely to be not significantly different from zero.

The implication for trend monitoring is that the method of data collection, the method of statistical analysis and the resulting estimates are inseparable. If there is bulking, then if it is consistent from year to year, the appearance of trends will not be unduly affected, although they may be less precisely estimated than from unbulked data.

If bulking has not been done consistently, the situation is more complicated. A more complex statistical analysis would be necessary to adjust estimates for bulking, but this would require the weights of each contributing tissue.

References

Aitchison J. and Brown J.A.C. (1957) The Lognormal Distribution. University of Cambridge Department of Applied Economics. Monograph 5. Cambridge University Press.

Anon (1986) Pooling strategy as applied to the Belledune lobster contamination problem. Report of the Meeting of the ad hoc Group of Statisticians assisting the Working Group on Marine Pollution and Baseline Monitoring Studies in the North Atlantic on Trend Monitoring. ICES C.M.1986/E:39. Annex 6. (mimeo).

Nicholson M.D. Green N.W. and Wilson S.J. (1988) Regression models for assessing trends in cadmium and PCBs in cod liver. ICES C.M. 1988/E:27. Report of the 1988 meeting of the Working Group on Statistical Aspects of Trend Monitoring Annex 7. (mimeo).

Fig 1

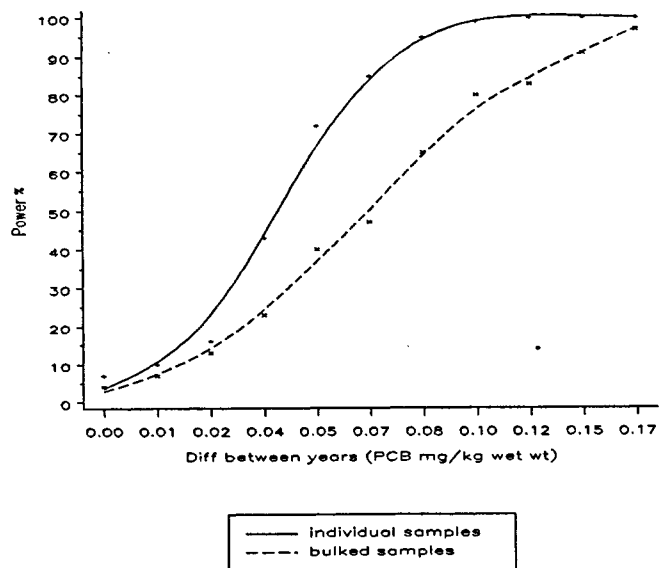
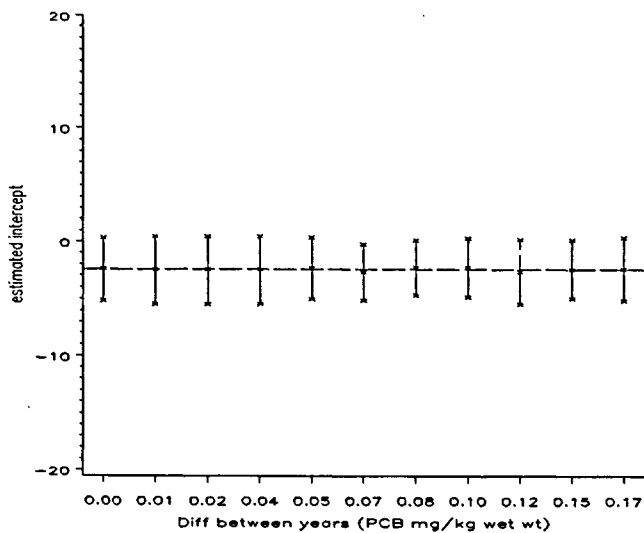


Fig 2
INDIVIDUAL SAMPLES



BULKED SAMPLES

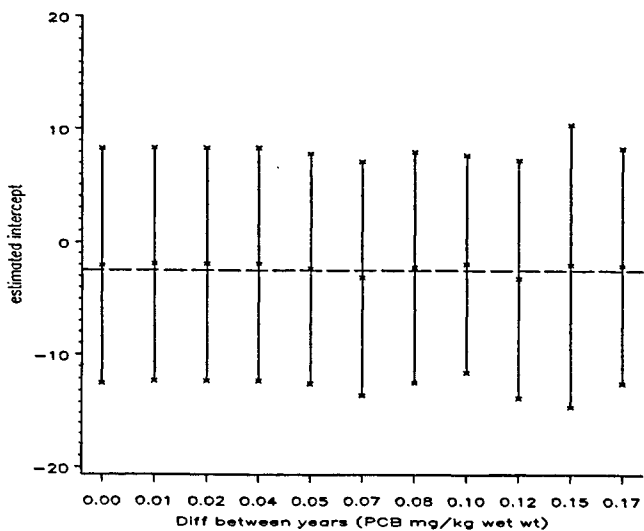
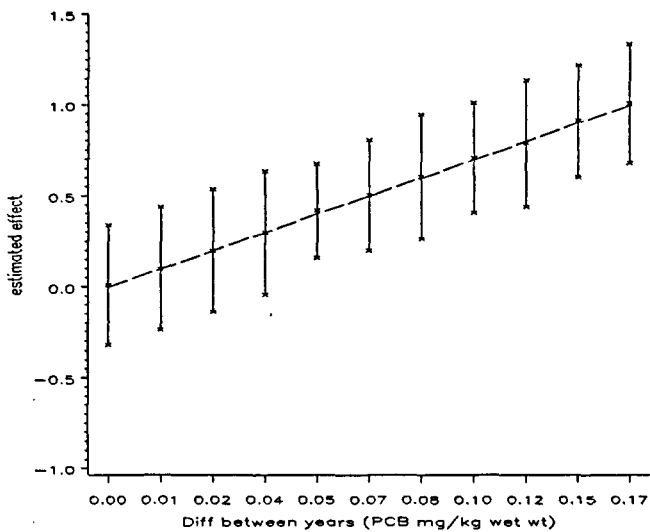


Fig 3

INDIVIDUAL SAMPLES



BULKED SAMPLES

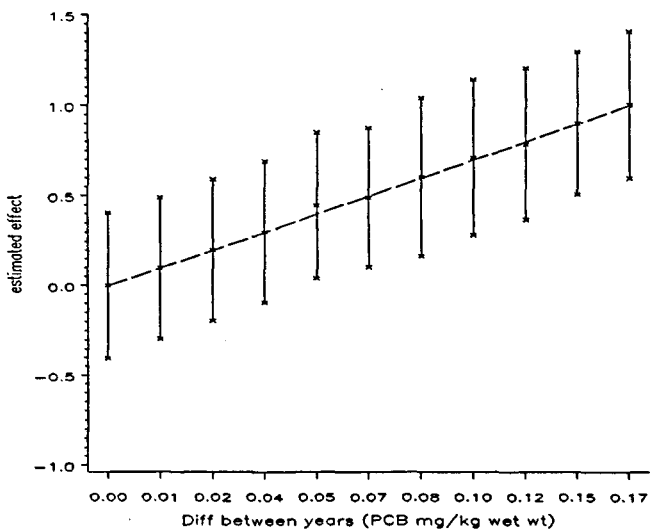
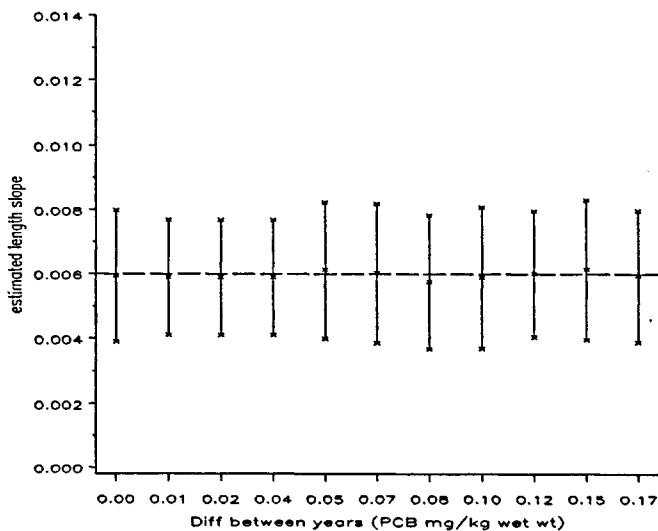
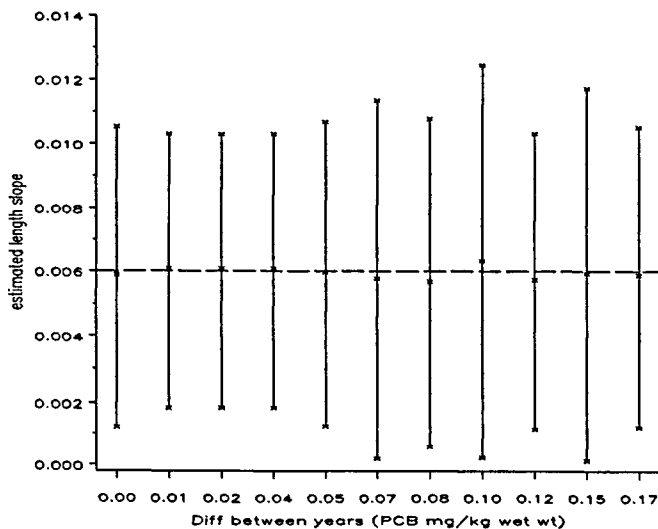


Fig 4

INDIVIDUAL SAMPLES



BULKED SAMPLES



ANNEX 5

THE EFFECT OF ANALYTICAL ACCURACY ON
THE POWER OF STUDENT'S TWO-SAMPLE t-TEST

by

M.D. Nicholson and J. van der Meer

Summary

Analyses of contaminant trend data assume that the level (eg concentration) of contaminant in a sample is observed without measurement error. In practice, the analytical method will have a certain precision, and may also be biased. Precision is defined as the standard deviation applicable to replicated chemical analyses; bias is the true concentration minus the mean concentration of replicate analyses.

Low precision and bias will have an adverse effect on statistical tests of significance. They will tend to change the notional significance level and to decrease the power of the test.

The effects of precision and bias are demonstrated by the change in performance of Student's two-sample t-test. The implications of pooling samples prior to chemical analysis are also considered. These results are illustrated for trend assessments of lead and copper concentrations in Mytilus edulis using data from the ICES cooperative monitoring programme and an international intercalibration exercise for metals in biota.

1. Two-Sample t-Test.

Suppose we have two Normally distributed populations. Using the usual notation, the two-sample t-statistic is

$$t = (x_1 - x_2) / [s(1/n_1 + 1/n_2)^{1/2}]$$

which can be written

$$t = (u+d)(x^2/f)^{-1/2}$$

where

$$d = (\mu_1 - \mu_2) / [\sigma(1/n_1 + 1/n_2)]^{1/2}.$$

Here u is a unit Normal random variate, X^2 is a chi-squared variate with $f = n_1 + n_2 - 2$ degrees of freedom, and d is the non-centrality parameter. (Note that there are different expressions for d ; see Johnson and Kotz, 1970). For convenience, we assume $\mu_1 \geq \mu_2$ and $d \geq 0$.

We write $t(d)$ to signify that, in general, the distribution of the t -statistic depends upon the value of d .

The power of the two-tailed t -test at the 5% level of significance for a given value of d is

$$P[d] = \text{Prob}[|t(d)| > t_{0.975}]$$

where $t_{0.975}$ is defined by

$$\text{Prob}[|t(0)| > t_{0.975}] = 0.05.$$

$P[d]$ increases monotonically with d .

2. Effect of Precision and Bias on Power.

2.1 Effect of Precision on Power.

The variance of an observed value of x will be

$$\begin{aligned}\sigma_o^2 &= \sigma^2 + \sigma_p^2 \\ &= \sigma^2(1+k^2), \text{ say}\end{aligned}$$

where $k = \sigma_p/\sigma$ and σ_p is the precision of the analytical method.

The non-centrality parameter becomes

$$d/(1+k^2)^{1/2}$$

and since

$$d/(1+k^2)^{1/2} < d$$

$$P[d/(1+k^2)^{1/2}] < P[d].$$

Hence analytical variability decreases statistical power. However, since

$$P[0/(1+k^2)^{1/2}] = P[0].$$

the size of the test is unaltered.

Figure 1 shows $P[d/(1+k^2)^{1/2}]$ plotted against $(\mu_1 - \mu_2)/\sigma$ for $k = 0, 0.25, 0.5$ and 1.0 for a t-test with 10 observations from each population and a 5% significance level.

2.2 Effect on Power when Bias is Fixed.

Suppose there are fixed analytical biases of b_1 and b_2 in populations 1 and 2 respectively.

The non-centrality parameter becomes

$$d+db$$

where

$$db = (b_1 - b_2)/[\sigma(1/n_1 + 1/n_2)^{1/2}].$$

The effect of bias will depend upon the values of b_1 and b_2 .

Consider three simple cases:

$$(1) \ b_1 = b_2.$$

In this case, $db=0$ and the bias in each population cancels. The test will be unaffected.

$$\text{ie } P[d+db] = P[d].$$

$$(2) \ b_1 = b \quad b_2 = -b \text{ where } b \text{ is positive.}$$

In this case bias will make the populations appear to be further apart.

$$\text{ie } P[d+db] = P[d+2b/\{\sigma(1/n_1 + 1/n_2)\}]$$

$$> P[d].$$

$$(3) \ b_1 = -b \quad b_2 = b \text{ where } b \text{ is positive.}$$

In this case the power will not increase monotonically with d , or be consistently greater/less than $P[d]$.

ie $P[d] = P[d - 2b / \{\sigma(1/n_1 + 1/n_2)\}]$
 $> P[d]$ when $0 < d < b / \{\sigma(1/n_1 + 1/n_2)\}$
 $< P[d]$ when $d > b / \{\sigma(1/n_1 + 1/n_2)\}$.

Figure 2 shows $P[d+db]$ plotted against $(\mu_1 - \mu_2)/\sigma$ for these three cases with, for example, $b/\sigma = \pm 0.5$.

2.3 Effect when Bias is Variable.

In practice, the biases associated with each population will relate to different laboratories or to the same laboratory operating at different times.

Thus instead of being equal, or equal with opposite signs, the biases are likely to be independent with values drawn from some probability distribution. i.e. the non-centrality parameter

$d - db$

is a random variable, and

$t(d - db)$

is drawn from a compound non-central t-distribution. The form of this distribution will depend upon the distributions of b_1 and b_2 (and hence of db). In the simple case, the two biases might reasonably be Normally distributed and centered on zero.

Figure 3 shows $P(d - db)$, the power when there is random bias, when the standard deviation of the random bias within a population, σ_b , takes the values $\sigma_b/\sigma = 0, 0.25, 0.5$ and 1.0 . As σ_b/σ increases, the probability of incorrectly concluding that there is a significant difference increases, and conversely, as the means become widely separated, the power of the test is reduced.

2.4 Effect of Pooling Samples on Power.

When $n_1 = n_2 = n$, the effect of precision and bias is to change the power of the test from

$P[d]$

to $P[(d + db)/(1 + k^2)^{1/2}]$

where the test has $2(n-1)$ degrees of freedom.

If the n animals are bulked into n_h homogenates each containing n_1 individuals (i.e. $n=n_h n_1$) the power of the test is still

$$P[d]$$

which is changed under the influence of precision and bias to

$$P[(d+db)/((1+n_1 k^2)^{1/2})]$$

with $2(n_h-1)$ degrees of freedom. Thus pooling reduces power both by increasing the effect of σ_p and decreasing the degrees of freedom.

3. Application to Assessment of Trends in *Mytilus edulis*.

3.1. Data.

To assess of the effect of analytical variability on trend measurement data have been selected from the ICES Cooperative Monitoring Programme (ICES, 1984) and the report of the seventh ICES intercalibration for trace metals in biological tissue (Berman and Boyko, 1987).

Concentrations of copper and lead were observed in *Mytilus edulis* collected in ICES area 31F2 during 1986. Each metal was measured four times on a pooled sample of animals. The numbers of animals in each pooled sample ranged from 55 to 170. The means and standard deviations of concentration in an individual were estimated to be

	Copper	Lead
Mean (mg/kg)	5.73	0.920
s.d.	27.9	0.718

Information on possible laboratory precision and bias was obtained from Berman and Boyko (1987). A number of laboratories made replicate analyses of several metals in a variety of sample types in 1985 and again in 1986. The metals included copper and lead, and the sample types included *Mytilus edulis*. For copper and lead the notional true concentrations were similar to the mean levels observed in ICES area 31F2:

	Copper	Lead
True Concentration	7.3	1.7
(mg/kg)		

A measurement made at a particular time is assumed to have a standard deviation σ_b and bias b . Further, b is assumed to vary in time with mean μ_b and standard deviation σ_b .

Columns 1, 2 and 3 of Table 1 gives the estimates of σ_b , σ_b and μ_b for each metal and laboratory for which there were data. Column 4 gives a measure of accuracy, the combined effect of precision and bias (Nicholson, 1989). Here accuracy is defined as

$$\{[|\mu_b| + 1.645(\sigma_b^2 + \sigma_b^2)^{1/2}] / (\text{true concentration})\} 100\%$$

3.3 Simulated Effect of Analytical Accuracy on Trend Assessment.

Column 5 of Table 1 gives the results of a computer simulation study to assess the potential effect of the measured accuracies of the different laboratories.

A monitoring programme for concentrations of copper and lead at levels similar to those observed in *Mytilus* in ICES area 31F2 was assumed to consist of annual samples of 125 individuals bulked into 5 homogenates of 25 animals. For each laboratory the annual mean was assumed to have changed in a direction unfavourable to its estimated fixed bias. The magnitude of the change was chosen so that the power of the test would be 90% if bias and precision were zero.

4. Discussion.

The targets for analytical accuracy are often defined only in terms of the replicate variability observed within a laboratory over a short time interval. However, bias also makes a contribution to analytical noise, and may in fact have a more adverse effect than poor precision (Nicholson, 1989). Targets should therefore be defined for both precision and bias, either separately or as some compound quantity such as accuracy.

As shown in Table 1, bias is not a fixed quantity associated with a laboratory, but may fluctuate with time.

This implies that the analytical accuracy applicable to a laboratory should not be measured from the results of a single experiment or from data collected over a short time interval.

It can be difficult to choose a target for accuracy. One criterion is that analytical accuracy should be sufficiently good not to influence the interpretation of data from monitoring programmes.

The results of our simulation study show that the permitted accuracy depends very much on the level of variability of the contaminant in the population. Thus for copper, for which variability in the population was large, even very poor accuracy had little effect on the power of the test. For lead, where variability was low, power was severely affected; either power was low, or appeared to be high but in fact a decrease in concentration was being reported as a significant increase (marked * in Table 1).

References

- Berman S.S. and Boyko V.J. (1987). ICES seventh round intercalibration for trace metals in biological tissue. ICES-7/TM/BT (part 2) Preliminary report.
- Johnson N.L. and Kotz S. (1970) Distributions in statistics: Continuous univariate distributions - 2. John Wiley and Sons, New York.
- Nicholson M.D. (1989) Analytical results: How accurate are they? How accurate should they be? Mar. Pollut. Bull. 20, 33-40.

Table 1. Summary of analytical characteristics by laboratory and metal.

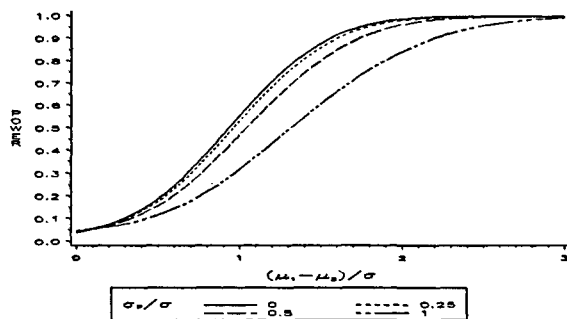
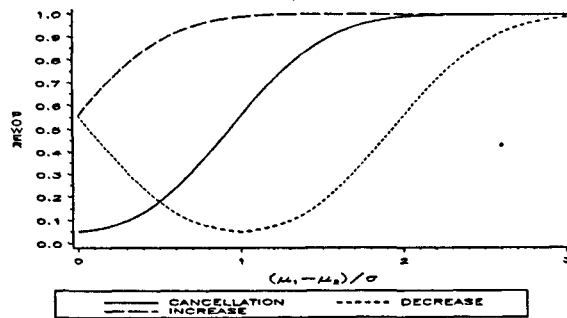
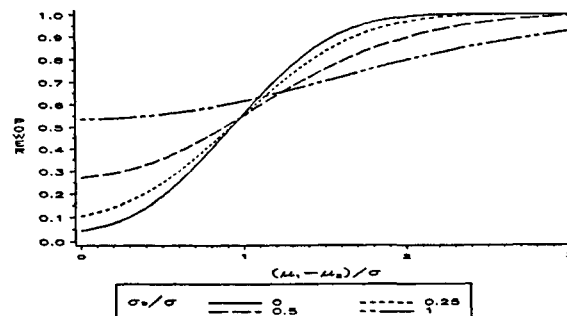
Copper

Laboratory	σ_p	σ_b	μ_b	%accuracy	Power
1	0.33	0.01	-0.27	13	89
2	0.49	0.00	0.33	18	88
3	0.40	0.56	-0.60	27	87
4	0.14	0.08	-0.23	8	89
5	0.74	0.26	-0.67	31	85
6	0.28	0.19	0.28	13	88
7	0.18	0.27	-0.37	14	88
8	0.74	1.42	1.19	60	80
9	0.20	0.10	0.24	9	88

Lead

Laboratory	σ_p	σ_b	μ_b	%accuracy	Power
1	0.15	0.36	-0.02	43	63
2	0.20	0.66	-0.01	74	74*
3	0.10	0.56	-0.58	90	78*
4	-	-	-	-	-
5	0.16	0.25	-0.25	44	43
6	0.79	0.47	1.52	176	53*
7	-	-	-	-	-
8	-	-	-	-	-
9	0.39	0.47	1.75	157	89*

POWER OF TWO SAMPLE T-TEST FOR SAMPLES OF SIZE 10

FIGURE 1
ADDITIONAL ANALYTICAL VARIANCEFIGURE 2
FIXED BIAS OF ARBITRARY SIGNFIGURE 3
VARIABLE ANALYTICAL BIAS

ANNEX 6

THE EFFECT OF DEPURATION ON MUSSEL ANALYSES

by

Norman W. Green

1. Abstract

The procedure agreed upon by JMP for pretreatment of mussels calls for 12-24 hour depuration period to allow for discharge of pseudo-faeces (ICES, 1986). This paper presents results that indicate that at least for one of the three JMP-regions in Norway this step is not necessary.

2. Introduction and Purpose

Depuration is considered important in pretreatment of mussel samples presumably because contaminant bound particles found in the gut may result in artificially high determinations. Therefore, depuration is particularly important for mussels collected in waters with high total suspended matter (TSM). The TSM that is not ingested in the mussels (i.e. pseudo feces) is not relevant when assessing trends. Depuration in some cases significantly increases the pretreatment work load; requiring: intensive cleaning of the aquariums and related material, transportation of this equipment to the field, collection of ambient seawater and the availability electricity and cooling facilities.

3. Materials and Methods

To test the effect of depuration on mussel (*Mytilus edulis*) samples were collected from the Sørfjorden and the Hardangerfjord (JMP areas 63 and 62, respectively) on the West coast of Norway (Figure 1). These areas are largely influenced by heavy metal discharges primarily originating at the head of the Sørfjord, in particular cadmium, mercury, lead and zinc. The effect of some discharge metals can be traced over 100km away to the outer regions of the Hardangerfjord.

During October 1988 mussels were collected from 0-2m depth at two stations: St.6 Kvalnes in the inner Sørfjorden and St.15 Vikingneset in the Hardangerfjorden (Fig.1). About 100 individuals were collected for each of three size intervals: 2-3, 3-4 and 4-5cm. Fifty individuals in each size interval were depurated and the remaining fifty were kept cool and dry until the soft parts were removed. Each subsample of fifty was split in two equal parts as parallel samples.

Two twenty liter aquariums were used for depuration. These and perforated polyethylene trays and airng rods were cleaned before hand

in nitric acid and rinsed in turn by distilled water, acetone and cyclohexane, respectively. The aquariums were placed in the cooling room ($\approx 5^{\circ}\text{C}$). For each station the aquariums were filled with surface seawater collected in the vicinity of the mussel bed and the three sizes of mussel were placed on the tray about 5 cm below the water level. The mussels were left in the aquariums 12-15 hours being aerated continuously. After this time the mussels (including the non depurated ones) were kept alive under cool and dry conditions until dissection of the soft tissue within 5-6 days after collection. The samples were then frozen.

The soft tissue samples were analyzed for cadmium, copper, mercury, lead and zinc. The method of chemical analysis has been described elsewhere (Green 1987).

Statistical analysis employed the principle of reduction of the residual sum of squares (Weisberg 1985). A regression model was fitted for the log contaminant concentrations and indicator variables for station (2), size (3) and depuration (2).

4. Results and Discussion

The concentrations of metals found in the mussels tissues are shown in Appendix Table A. The results from the statistical analysis (Table 1) indicated that the effect of sampling site is significant for all metals, i.e., mercury, cadmium, copper, lead and zinc (higher concentrations are found in Sørkjorden) and for lead and zinc the effect of shell length was significant (concentrations generally increased with length). The effect of depuration was not found significant in any case.

Total suspended matter (TSM) in the surface waters of these fjord regions is relatively low. The highest values are found at the head of the Sørkjord and varied from 0.2 to 0.9 mg/liter during June and September 1987 and February 1988 (Skei, pers. medd. 1989). This roughly compares to another JMP area in Norway, the Oslofjord region, which has 0.3-1.7 mg/liter August 1988 (Skei, pers. medd. 1989). These ranges are considerably lower than areas near the mouth of considerable river runoff such as near the mouth of the Rhine or Ems along the Netherlands coast where depuration may have significant influence on the contaminant content in the mussel. From 0 to 10 km from the Netherlands coast the mean TSMs were found to be 48-91 mg/liter. But 70 km from the coast the mean TSMs were much lower (<3 mg/liter) and comparable to Norwegian conditions (JMG, 1989).

5. Conclusion

This study indicated that depuration has no statistical significance on the concentrations of mercury, cadmium, copper lead and zinc in mussel collected in Sør fjord and Hardangerfjord. The level of TSM in Sør fjord is low. It is suggested that depuration may be excluded from pretreatment steps in mussel monitoring in cases where there is reasonable evidence that it has no bearing in the results.

6. References

- Green, N.W., 1987. Joint Monitoring Program. Overview of analytical methods employed by JMP in Norway 1981-1987. 32pp..
- JMG, 1989. 1987 National comment of the Netherlands on the JMP data submitted. Fourteenth Meeting of the Joint Monitoring Group, Vigo 24-27 January 1989. JMG 14/3/8. 46pp..
- ICES, 1986. Interim reporting format for contaminants in fish and shellfish. JMP-version. Revised May 1989. 46pp..
- Weisberg, S., 1985. Applied Linear Regression. Second edition. John Wiley & Sons, 324pp.

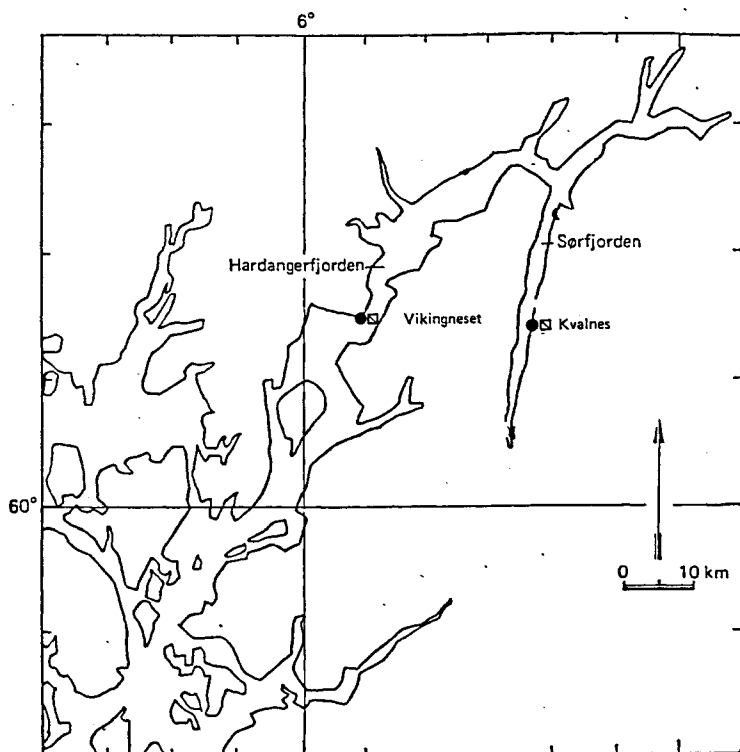


Figure 1. Map of mussel (*Mytilus edulis*) sampling stations in the Sør fjorden and Hardangerfjorden of Norway.

Table 1. The effect of collection site (2 stations), mussel length (three size categories: 2-3, 3-4, and 4-5cm), and depuration using the principle of the reduction of the residual sum of squares (RSS) based on log-e (metal in ppm dry weight)

	df	Hg	Cd	Cu	Pb	Zn
<hr/>						
RSS						
Whole model	19	0.22092	0.47834	0.44574	1.38950	0.36543
Effect of:						
station (st)	20	5.3410	29.322	1.72895	17.8786	5.9895
length (lg)	21	0.26684	0.62617	0.57816	2.4914	0.56571
deputation (dp)	20	0.23321	0.49443	0.44583	1.39136	0.36828
lg + dp	22	0.27914	0.64225	0.57825	2.4932	0.56856
st + dp	21	5.3533	29.338	1.72903	17.8805	5.9923
<hr/>						
F-value						
station (st)	1	440.347	1145.690	54.698	225.472	292.415
length (lg)	2	1.975	2.936	2.822	7.534	5.207
deputation (dp)	1	1.057	0.639	0.004	0.025	0.148
lg + dp	3	1.669	2.190	1.883	5.031	3.520
st + dp	2	220.703	573.163	27.351	112.749	146.280
<hr/>						
Statistical significance						
station (st)	1	***	***	***	***	***
length (lg)	2	ns	ns	ns	**	**
deputation (dp)	1	ns	ns	ns	ns	ns
lg + dp	3	ns	ns	ns	**	*
st + dp	2	***	***	***	***	***

Table A. Concentrations (ppm dry weight) of mercury (Hg), cadmium (Cd), copper (Cu), lead (Pb) and zinc (Zn) in blue mussel *Mytilus edulis* collected at St.6 Kvalnes in the inner Sør fjord and St.15 Vikingneset in the Hardangerfjorden, October 1988. Three shell length intervals were collected at each station: 2-3, 3-4, and 4-5cm. Two samples of approximately 25 individuals were depurated (D) and two samples of approximately 25 individuals were not depurated (N).

Station	Test		Hg	Cd	Cu	Pb	Zn
Size	Count						

St.6, Kvalnes, Sør fjorden							
2-3 cm	D	25	0.42	56.1	6.71	17.2	346
		25	0.38	59.9	8.54	19.7	482
	N	25	0.37	55.4	8.18	16.0	369
		25	0.34	51.0	8.29	15.3	337
3-4 cm	D	25	0.34	53.4	6.56	19.2	423
		25	0.41	55.0	7.57	24.0	437
	N	25	0.37	53.1	6.69	22.7	478
		25	0.34	50.6	6.98	17.7	368
4-5 cm	D	25	0.32	53.4	8.73	12.4	343
		25	0.37	56.3	7.00	21.7	398
	N	25	0.34	52.7	8.61	25.7	452
		25	0.45	59.3	11.5	33.0	505
St.15, Vikingneset, Hardangerfjorden							
2-3 cm	D	25	0.18	5.14	7.06	2.71	133
		25	0.16	5.46	4.95	2.64	125
	N	25	0.16	4.82	4.97	2.39	127
		25	0.15	4.34	4.75	2.03	129
3-4 cm	D	25	0.15	6.51	4.44	4.48	166
		25	0.12	5.30	4.08	3.13	162
	N	25	0.13	6.54	4.93	3.81	177
		25	0.15	7.22	4.73	6.14	182
4-5 cm	D	25	0.15	9.06	5.17	6.49	196
		25	0.16	7.14	5.84	5.52	150
	N	25	0.14	8.59	5.36	5.63	202
		25	0.12	4.96	3.76	3.50	135

ANNEX 7

ON MULTIVARIATE AND UNIVARIATE ANALYSES OF VARIANCE

by

R.K. Misra, J.F. Uthe, and W. Vynke*

Marine Chemistry Division
Physical and Chemical Sciences Branch
Department of Fisheries and Oceans
P.O. Box 550, Halifax, Nova Scotia, Canada B3J 2S7

Summary

A series of univariate analyses of variance (ANOVA) performed on data on several dependent variables, Y_i , is of severely restricted scope as it will ignore correlations among Y_i . Multivariate analysis of variance (MANOVA) is the appropriate analysis of data where Y_i are mutually correlated. Comparison of ANOVA and MANOVA techniques is presented through (a) a brief review of ANOVA and MANOVA models and (b) their application to data on trace metal (Cu and Zn) levels in individual flounder caught in 1981 and 1983 off the Belgian coast (ICES rectangle 31F2). Individual values of Y_1 and Y_2 were shuffled to achieve the following three values of their coefficient of correlation, (1) high and positive, (2) near zero, and (3) high and negative. In case (2), ANOVA and MANOVA outcomes were consistent with each other, neither analysis identifying differences between year-means. In cases (1) and (3) this difference was identified by MANOVA as highly significant whereas the ANOVA outcome was exactly the same as case (2).

* Station de Pêche, Ostende, Belgium.

Introduction

Several books on multivariate statistics are now available. See, e.g. Kshirsagar (1972), Bock (1975), Harris (1975), Kendall and Stewart (1976), Morrison (1976), Mardia et al. (1979), Srivastava and Carter (1983), Bray and Maxwell (1986), and Johnson and Wichern (1988). Harris (1975) and Bray and Maxwell (1986) were used particularly heavily in the preparation of this presentation. An assortment of descriptive and inferential techniques, known as multivariate statistics, has been developed to analyse data sets consisting of m predictor (also called independent, class or classification, indicator, group, etc.) variables and p outcome (also called dependent, response, criterion, etc.) variables. Various statistical techniques are distinguished primarily on the sizes of m and p . However, explicit labeling of these as "univariate" or "multivariate" is problematic since common usage is not consistent with logic (Harris 1975). For example, if we imply that multivariate techniques are applied to situations involving two or more dependent measures (such as contaminants) then multiple linear regression (MLR) is not a multivariate procedure, which is counter to common practice, (e.g. Lassen 1983). Again, considering that measures such as age, sex, and year are predictor variables analysis of variance (ANOVA) is a special case of MLR. However, ANOVA is almost invariably labeled as a univariate analysis, probably because researchers who conduct laboratory experiments gave this label a long time ago (Harris 1975).

Analysis of covariance (ANCOVA), model 2, as recommended by the International Council for the Exploration of the Sea (ICES 1987a) as the statistical procedure of analysing for trends over time in contaminant concentrations in fish stocks from selected areas, is only an ANOVA where predictor variables are a mixture of class (year) and measured (length) variables. In this presentation we label analysis of variance as univariate

ANOVA when it employs only one criterion variable, Y (i.e. $p = 1$) and as multivariate ANOVA or MANOVA when it employs two or more criterion variables, Y_i , $i = 1, \dots, p$ ($p \geq 2$).

The Need for Multivariate Analysis

Complexity of a scientific inquiry frequently requires an investigator to collect observations on several variables, Y_i , taken on individual subjects. Numerous data sets collected on contaminant levels in food fish, including (a) sixty-two (53 metal and 9 organochlorine) data sets on contaminants in fish muscle tissue, analysed (ICES 1987b, 1989) under the ICES Cooperative Monitoring Programme (CMP) and (b) contaminant concentration data on Canadian Atlantic cod (*Gadus morhua*) of Scott et al. (1981), are common examples of data employed in time trend monitoring programmes (ICES 1982) developed by the ICES Advisory Committee on Marine Pollution (ACMP) for fish and shellfish. On each fish several criterion variables (contaminants) are measured simultaneously. These variables are, generally, correlated mutually. For data of this type, multivariate statistical approaches are favoured as reflected in the following statements

A series of univariate statistical analyses carried out separately for each of the variables is, in general, not adequate as it ignores the correlations among the variables. It may even be misleading sometimes. On the contrary, multivariate analysis can throw light on the relationships, interdependence, and relative importance of the characteristics involved and yield more meaningful information, in general. (Kshirsagar 1972)

.....researchers in all of the sciences -
 behavioral, biological, or physical - have long since
 abandoned sole reliance on the classic univariate
 design for very excellent reasons. (Harris 1975)

Bock (1975) notes that (a) Multivariate statistical methods are emphasized because they make it possible to encompass all data from an investigation in one analysis. (b) This approach results in a clearer, better organized account of the investigations than do piece meal analyses of portions of the data. Concerning the application of multivariate statistics in the behavioral sciences, Bray and Maxwell (1982) note that it appears that this will be the, "dominant method of analysis in the future".

The example on flounder presented later in the text would show that when criterion variables (contaminants) are correlated, temporal variations are identifiable by MANOVA but not by ANOVA. In this sense univariate analyses can be misleading. Nevertheless, total rejection of univariate procedures in favour of the multivariate procedure may be neither sensible nor desirable in the time trend investigations which involve comparisons of mean contaminant values. Missing observations (particularly when there are several of them, as is the case in many of the CMP data sets) virtually destroy multivariate analysis. Univariate analysis may be employed when the manager is interested in , primarily or exclusively, tracking temporal variations of individual contaminants or of a specific contaminant. Supplementing univariate analysis with a multivariate analysis would make a good strategy. The two procedures provide different perspectives on temporal variations. In cases where the outcomes of the two analyses yield contradicting inferences, the investigator may (1) exercise caution in making conclusions from univariate analyses and (2) supplement these analyses by an extension to the multivariate level (a) to

examine the significance of individual contaminants in the multivariate complex and (b) test for sensible biochemical models based upon composites of contaminants

ANOVA and MANOVA

The MANOVA concept is a straightforward extension of ANOVA where the number (p) of dependent variables is two or more. Multivariate analysis of covariance (MANCOVA) is an extended form of ANCOVA and consists of the MANOVA of Y_i , $i = 1, \dots, p$ values corrected for variations in the covariate (e.g. length) X based on regressions of Y_i on X . Step 1 of ANOVA tests the overall null hypothesis (H_0) of equality of means of populations (years). This may be supplemented by step 2 which tests the significance of specific comparisons (contrasts) among group means such as the difference between means of two years, time trends, second degree terms of deviations from the time trend. MANOVA also tests the overall H_0 and evaluates mean differences among groups but in doing so it considers correlations among Y_i variables by analysing these simultaneously rather than ignoring these correlations by analysing one Y_i at a time which is how ANOVA will handle $p (\geq 2)$ variables Y_i . As an added bonus in MANOVA, contrasts of step 2 can be defined not only for group means of Y_i but also for group means of linear combinations of Y_i , e.g. time trends of the difference, $Y_1 - Y_2$, in addition to time trends of each. An exposure of mathematical models of ANOVA and MANOVA and their applications on data will explain these features more clearly. In order to keep the mathematical tedium at its bare minimum, we present these models for the case of two populations (years) based upon two criterion variables, Y_1 and Y_2 . For applications of these procedures we consider data on trace metal levels in individual fish provided by ICES (see format in ICES 1986). A detailed account of the screening of data on Cu, Zn, Hg, and length observed in individual cod and

flounder (ICES rectangle 31F2) is given in Misra et al. (1988). Individual fish specimens were sorted by the qualifying items, (i) tissue type, "muscle", (ii) analysis type, "individual", and (iii) for contaminant concentrations in an individual; qualifier code was "=" (as opposed to less than or greater than), basis = 0 (wet) and unit = 0 (mg·kg⁻¹). In this presentation we used a part of this data set, viz. observations on Cu (Y₁) and Zn (Y₂) of flounder caught in 1981 and 1983. Data were transformed to common logarithms and are shown in Table 1.

In the single classification ANOVA the groups are formed by one classification variable with K levels (for the flounder data of this presentation the classification variable is year with K = 2). For a criterion variable, Y_i, the ANOVA model is given as

$$Y_{ijc} = \mu_{ij} + e_{ijc}, \quad (1)$$

$$j = 1, \dots, K \text{ and } c = 1, \dots, n_j$$

where

n_j = the number of individuals in group j and

μ_{ij} = population mean of Y_i for year j.

It is stipulated that Y_i value for an individual is represented by the mean of the group it belongs to. Any discrepancy between an individual observation and the group mean is attributed as "error". Mean μ_{ij} is usually unknown and is estimated by the sample mean \bar{Y}_{ij} for year j. Estimate \hat{e}_{ijc} (we denote it as V_{ijc} for typing convenience) of error for individual c in year j on variable Y_i is given as

$$V_{ijc} = Y_{ijc} - \bar{Y}_{ij} \quad (2)$$

For example, for variable Y₁ of individual 1 of year 1

$$V_{111} = Y_{111} - \bar{Y}_{11}$$

which = 2.602 - 2.524 or 0.078 (Table 1) and, similarly for its Y_2 ,

$$V_{211} = 4.120 - 4.011 \text{ or } 0.109 \quad (3)$$

It is obvious that these V_{ijc} terms, when summed for all individuals of a group on a Y_i will add to zero, i.e.

$$\sum_{c=1}^{n_j} V_{ijc} = 0$$

for a given i and j . For Y_i sum of square (SS) errors,

$$SSW_i = \sum_{j,c} V_{ijc}^2 \quad \text{or} \quad \sum_{j,c} (Y_{ijc} - \bar{Y}_{ij})^2 \quad (4)$$

provides an index of the magnitude of error and is referred to as SS within groups. Estimate s_i^2 of variance of random variable e_{ijc} is given by

$$SSW_i / (n - K), \text{ where } n = \sum_{j=1}^K n_j$$

is the total number of observations. For the data of Table 1,

$$n_1 = 19, n_2 = 25, n = 44, K = 2, \\ SSW_1 = 0.31818, SSW_2 = 1.33176, s_1^2 = 0.0076 \text{ and } s_2^2 = 0.0317 \quad (5)$$

Overall H_0 for ANOVA model (1) on Y_i is

$$H_0: a_{i1} = a_{i2} (= \dots = a_{ik}) \quad (6)$$

as opposed to the alternative hypothesis H_1 which says that at least two group (year) means are not equal. If H_0 is true then there is obviously no need for using separate parameters a_{ij} for years. So we reduce model (1) to the model where all a_{ij} are assigned a common value (a_i) for variable Y_i . This a_i is estimated as the grand mean \bar{Y}_i of an n observations regardless of group. Error for individual c in group j is then given as

$$V_{ijc} = Y_{ijc} - \bar{Y}_i \quad (7)$$

Sum of squared value of these errors is referred to as total sum of squares

(SST_i) and the difference, $SST_i - SSW_i$ as between group sum of squares (SSB_i). SSW_i and SSB_i are the "unexplained" (by the model) and "explained" components of the total variation, SST_i on Y_i . F test of H_0 is based on the comparison of SSB_i with SSW_i and is given by

$$F_i = \frac{SSB_i / (K-1)}{SSW_i / (n-K)}$$

In the MANOVA model variables Y_i are analysed simultaneously, so that overall H_0 is

$$a_{11} = a_{12} (= \dots = a_{1K}), \text{ and}$$

$$a_{21} = a_{22} (= \dots = a_{2K})$$

or, in vector notation,

$$\underline{a}_1 = \underline{a}_2 (= \dots = \underline{a}_K) \quad (9)$$

where \underline{a}_i is a column vector of p elements. H_1 of (9) is that group means are not all equal on at least one Y_i variable. Note that a test of H_0 of (9), based on separate ANOVAs of Y_1 and Y_2 , will employ sum of squared errors, $\sum_{j=1}^2 V_{1jc}$ and $\sum_{j=1}^2 V_{2jc}$ on Y_1 and Y_2 , but not their sum of products (SP_{12}) which is $\sum_{j=1}^2 V_{1jc} \cdot V_{2jc}$. In MANOVA, SSW_i , SST_i and SSB_i of separate ANOVAs are replaced by matrices W , T and B , respectively, where

$$W = \begin{bmatrix} SSW_1 & SPW_{12} \\ SPW_{12} & SSW_2 \end{bmatrix}, \quad T = \begin{bmatrix} SST_1 & SPT_{12} \\ SPT_{12} & SST_2 \end{bmatrix}$$

$$\text{and } B = T - W \text{ or } \begin{bmatrix} SST_1 - SSW_1 & SPT_{12} - SPW_{12} \\ SPT_{12} - SPW_{12} & SST_2 - SSW_2 \end{bmatrix}$$

The matrix analog of "Between"/"Within" is $B \cdot W^{-1}$ where W^{-1} is the inverse of W matrix (named equation (10)). The sum of products is closely related to the correlation between Y_1 and Y_2 as can be seen, e.g. from the following computation for the estimate of the pooled within group coefficient of correlation (r_{12}) between Y_1 and Y_2 :

$$r_{12} = SPW_{12} / (SSW_1 \cdot SSW_2)^{1/2} \text{ which also equals}$$

$$\text{Covariance}(Y_1, Y_2) / [\text{Variance}(Y_1) \cdot \text{Variance}(Y_2)]^{1/2} \quad (11)$$

From (10) and (11) it is obvious that when Y_i variables are highly correlated, SP_{12} will have a large influence on the test of H_0 . To demonstrate this on the flounder data of Table 1 we analysed these by MANOVA and two separate ANOVAs three times. Each time (a) we did not change the data set itself, i.e. we used the same Y_1 and Y_2 values as shown in Table 1 so that means, ranges and variances of Y_1 and Y_2 remained the same, but (b) structure of the data was changed by shuffling values of Y_1 and Y_2 between individuals so that correlation, r_{12} , changed. Values of pooled within group r_{12} generated were +0.98 (+ve and high), -0.10 (as close to zero as we could get with these data) and -0.97 (-ve and high). Probability (P) levels of significance in tests of univariate H_0 were 0.86 and 0.22 for Y_1 and Y_2 , respectively. There was thus no evidence against H_0 (recall that the convention is to reject H_0 when $P \leq 0.05$). As expected, these P values for separate ANOVAs stayed the same for all three data sets. The findings of MANOVA, however, were consistent with those of separate ANOVAs only in case 2 (where r_{12} was near zero) with $P = 0.46$. Multivariate H_0 was rejected overwhelmingly ($P = 0+$, i.e. close to zero) in cases where r_{12} were 0.98 and -0.97.

To sense how MANOVA attains greater power of discriminating between group (years) when Y_1 and Y_2 are correlated, we present scatterplot

(Figure 1) of Y_1 and Y_2 when $r_{12} = 0.98$. Projections of the plotted points on the Y_1 -axis or the Y_2 -axis overlap considerably (also see the % overlap values of Table 1). One would, therefore, expect a nonsignificant F for testing univariate H_0 of ANOVA. We also discern (Figure 1), rather clearly, (i.e. considering how messy these flounder data are), a straight line separating the two groups. Even though drawn arbitrarily (by freehand) in the figure, it has separated data of 1981 (A) from 1983 data (B) well, by keeping most As on side and most Bs on the other.

A brief account through figures underlying the distributions shown in Figure 1 is provided as follows: In a complete bivariate normal population, to every pair of measurements (Y_1, Y_2) there is a probability (frequency), $f(Y_1, Y_2)$. This is shown in Figure 2 (from Seal 1964), with the probability (frequency) axis drawn perpendicular to the plane of the Y_1 and Y_2 axes. Notice that for a given value of Y_1 there is a normal distribution of Y_2 values and vice versa. By looking down from the top in Figure 2 and working outwards from the origin we can draw a contour (an ellipse) such that the central 95% (or any other specified percentage) of observations lie inside it. Figure 3 (from Seal 1964) shows concentric ellipses for some percentage values. About a contour we note that (a) it is an ellipse because the variances of Y_1 and Y_2 will, in general, be different. In Figure 3 variability of Y_1 is larger than that of Y_2 . In the special case when these variances are equal, the ellipse is replaced by a circle. (b) Variables Y_1 and Y_2 are uncorrelated. Figure 4 gives the scatterplot of a hypothetical data set when Y_1 and Y_2 are uncorrelated, this ellipse will tilt as in Figure 5 where r_{12} is assumed to be positive. Also, when r_{12} is large, the ellipse will be long and thin. Figures 6a and b (from Johnson and Wichern 1988) give additional views of the bivariate normal distribution. For Figure 6 (a) $r_{12} = 0$ and for 6 (b) $r_{12} = + 0.75$. Note that due to the presence of correlation

the probability tends to concentrate along a line. Figure 5 is extended to show bivariate frequency distributions for one population (Figure 7) and two populations (Figure 8, from Bray and Maxwell 1986).

Observations made earlier in the text pertaining to Figure 1 are reiterated based upon Figure 8. Univariate distributions for the two populations on Y_1 and Y_2 overlap so extensively that separate ANOVAs will likely not identify differences between the group means. On the other hand, at the two dimensional level, populations overlap considerably less, thereby making the separation between groups by MANOVA possible. The extent of overlap between populations will reduce with the increase of the correlation between Y_1 and Y_2 . Although computations for the MANOVA are rather involved mathematically, these need not deter a practitioner from using it. This is because complete programs for MANOVA are readily available. Consequently, the end user of these programs will have hardly any mathematics to deal with. Therefore, here we give just a brief outline of the algorithm of MANOVA computations.

The two groups (Figures 1 and 8) can be separated by a line which is a linear combination (Z) of the original variables, Y_1 and Y_2 , and is given as,

$$Z = b_1Y_1 + b_2Y_2$$

The coefficients, b_1 and b_2 , are assigned values which will make the linear combination Z discriminate between the two groups more than any other linear combination. Searching for such b_i ($i = 1, 2$) is done by maximizing SSB/SSW on Z . This Z is called, "discriminant function" and is, in fact, eigen vector of the $B \cdot W^{-1}$ matrix of (10). Also, eigen value of this matrix is the ratio SSB/SSW on Z . We could use individuals' values of the single composite Z as univariate observations to test H_0 of MANOVA by the univariate t test or F test performed on Z values.

As stated earlier in the text, the test of overall H_0 is only the first of the two steps of analysis a practitioner would usually be interested in. Rejection of H_0 merely indicates that at least two groups differ on at least one Y_i variable. This knowledge is not sufficient for a practitioner who needs to probe into specifics of differences among groups, e.g. general time trend, responses, e.g. contribution of an individual Y_1 or of Y_2 to the general time trend, and combinations of groups and responses, e.g. time trend for $Y_1 - Y_2$. A number of simultaneous test procedures for multivariate multiple contrasts are known (see, e.g. Ramsey 1982, Bird and Hadzi-Pavlovic 1983). In spite of the weakness associated with it, the Roy union-intersection approach may appeal to a practitioner most because (a) It is the most general method of examining such contrasts, and (b) It provides the convenience of leading directly to simultaneous confidence bounds on linear compounds of groups and/or variables.

Additional Notes

In the following we give only a few notes concerning the MANOVA, much more could have been said about it.

1. When dependent variables, Y_i , are correlated, not only will performing separate ANOVAs on Y_i ignore their correlations, but also their F tests will not be statistically independent.
2. For the flounder example given here, p was 2. Matrix $B \cdot W^{-1}$ consisted of 4 ($=p^2$) elements. MANOVA simplifies the complexity of handling these many numbers to just one, i.e. Z . In one sense the handling of the MANOVA procedure presented earlier is done by reducing the suite of p measures on each individual to a single Z score and then performing a univariate F-ratio on Z variable.

3. MANOVA provides a unique medium for examining variations among years (including time trends) by employing relationships among original variables, Y_i , rather than by looking at each Y_i in isolation. For example, the linear combination of Y_i provided in Z discriminates between groups (years) best (of all possible linear combinations). Therefore, magnitude of weights assigned to Y_i in Z could be used in extending interpretation on MANOVA results. With the simultaneous use of Y_i variables (as in MANOVA) it is possible to examine temporal variations and time trends for specific combinations of patterns of variables which are of interest to a practitioner.
4. For the flounder example there was only one discriminant function. In the general case of $p \geq K$ and $K \geq 2$, the number of discriminant functions is s where s is the smaller of p and $K-1$. Discriminant functions, Z_1, \dots, Z_s , have the following characteristics: (a) They are uncorrelated with one another. (b) The q -th discriminant function, Z_q , is associated with the q -th largest eigen value (also called the characteristic root) of the $B \cdot W^{-1}$ matrix, and is given by the q -th eigen vector, and (c) A Z_q is a relevant source of differences among groups only if it is statistically significant.
5. Unlike the ANOVA there are four procedures for testing the multivariate H_0 . These procedures, which differ in the way the information in the s eigen values is combined, are: (i) Wilk's lambda, (ii) Pillai-Bartlett trace, (iii) Roy union-intersection procedure, and (iv) Hotelling-Lawley trace. It is not clear which of these procedures is most robust, has the greatest statistical power (power is the probability that H_0 is rejected when it is false) and should be most preferred. Wilk's lambda is historically the most widely used procedure and may be preferred when characteristic roots of $B \cdot W^{-1}$ are nearly equal. Roy union-intersection

approach which employs greatest characteristic root (gcr) of $B \cdot W^{-1}$ may appeal to the practitioner "because of its greater heuristic and didactic values" (Harris 1975) and for other reasons such as (Harris 1975): (a) It leads directly to multiple comparison procedures for groups and variables and (b) It clarifies the connection between discriminant analysis and MANOVA.

6. Issues of distributions of Y_i (or better, of error terms), homogeneity of their variabilities and statistical power of tests are relevant for MANOVA, as they are for ANOVA. ANOVA model requires that Y has a normal distribution within each group and that K groups do not differ in their variance (Y) values. These assumptions are paralleled in MANOVA which requires multivariate normal distribution of Y_i variables and equality of their covariance matrices. The following are noted (Morrison 1976, Stevens 1980, Ramsey 1982, Muller and Peterson 1984, Bray and Maxwell 1986): (i) A priori estimation of power when planning a MANOVA requires estimation of many unknown parameters, which is a difficult task. (ii) The size of within group correlations among Y_i variables influences power of the MANOVA test. Depending on these correlations, MANOVA can at times be less powerful than separate ANOVA tests on individual Y_i variables. For example, Table 4.7 of Morrison (1976) gives a comparison of univariate and multivariate powers and an example with two variables, Y_1 and Y_2 which show that (a) when $r_{12} = 0$, the multivariate test has less power than the univariate one, and (b) As this correlation increases, multivariate power eventually surpasses the univariate power.
7. General MANOVA is available in several packages. We have developed our own package (which is currently being extended) which gives extended analyses, such as examination of contributions of individual characters to

time trends, time trends of differences among contaminants, quadratic discriminators with covariance, multivariate analysis of attribute data, MANCOVA with principal component covariates.

TABLE 1. 1981 and 1983 flounder data on Y_1 (Cu) and Y_2 (Zn), their means, ranges, and percent overlaps between distributions for the two years. All measurements were transformed to common logarithms.

1981			1983		
Individual No	Y_1	Y_2	Individual No	Y_1	Y_2
1	2.602	4.120	1	2.672	4.382
2	2.556	4.057	2	2.580	4.167
3	2.653	4.053	3	2.491	4.045
4	2.690	4.208	4	2.568	4.114
5	2.462	3.989	5	2.518	3.987
6	2.568	4.333	6	2.556	4.127
7	2.591	4.345	7	2.505	3.959
8	2.415	3.903	8	2.663	4.057
9	2.447	4.099	9	2.613	3.954
10	2.431	3.843	10	2.380	3.863
11	2.431	4.176	11	2.431	4.053
12	2.462	3.783	12	2.568	3.845
13	2.672	3.741	13	2.544	3.820
14	2.491	4.013	14	2.477	3.771
15	2.477	4.022	15	2.602	3.982
16	2.342	3.850	16	2.462	3.880
17	2.556	4.074	17	2.447	4.029
18	2.580	3.926	18	2.643	3.753
19	2.532	3.681	19	2.505	3.790
20			20	2.447	3.840
21			21	2.505	4.176
22			22	2.491	3.679
23			23	2.491	3.884
24			24	2.447	3.724
25			25	2.380	3.713
Mean	2.524	4.011		2.520	3.944
Range	2.342-	3.681-		2.380-	3.679-
	2.690	4.345		2.672	4.382
% Overlap	84	100		100	94

References

- Bird, K.D. and D. Hadzi-Pavlovic. 1983. Simultaneous test procedures and the choice of a test statistic in MANOVA. *Psych. Bull.* 93: 167-178.
- Bock, R.D. 1975. *Multivariate Statistical Methods in Behavioral Research.* McGraw-Hill Book Co., New York, NY.
- Bray, J.H. and S.E. Maxwell. 1982. Analysing and interpreting significant MANOVAs. *Rev. Educational Res.* 52: 340-367.
- Bray, J.H. and S.E. Maxwell. 1986. *Multivariate Analysis of Variance.* Sage University Papers Series on Quantitative Applications in the Social Sciences 07 - 054. Beverly Hills, CA.
- Harris, R.A. 1975. *A Primer of Multivariate Statistics.* Academic Press, New York, NY.
- International Council for the Exploration of the Sea. 1982. Report of the Advisory Committee on Marine Pollution, 1981. ICES Coop. Res. Rep. No. 112.
- International Council for the Exploration of the Sea. 1986. Interim reporting format for contaminants in fish. Revision May 1986.
- International Council for the Exploration of the Sea. 1987a. Report of the Advisory Committee on Marine Pollution, 1986. ICES Coop. Res. Rep. No. 142.
- International Council for the Exploration of the Sea. 1987b. Report of the 1987 meeting of the Working Group on Statistical Aspects of Trend Monitoring. ICES C.M. 1987/E:24.
- International Council for the Exploration of the Sea. 1989. Statistical analysis of the ICES CMP data on contaminants in fish muscle tissue for determination of temporal trends (Purpose 3). ICES Coop. Res. Rep. No. -in press.
- Johnson, J.R. and D.W. Wichern. 1988. *Applied Multivariate Statistical Analysis.* Prentice-Hall, Englewood Cliffs, NJ.
- Kendall, M and A. Stewart. 1976. *The Advanced Theory of Statistics, Vol 3.* Hafner Press, New York, NY.
- Kshirsagar, A.M. 1972. *Multivariate Analysis.* Marcel Dekker Inc. New York, NY.
- Lassen, H. 1983. Multivariate analysis applied to monitoring time trends in contaminant levels in fish. ICES Statutory Meeting C.M. 1983/D:4.
- Mardia, K.V., J.T. Kent and J.M. Bibby. 1979. *Multivariate Analysis.* Academic Press, New York, NY.

- Misra, R.K. and J.F. Uthe. 1987. Methods of time trend analysis applied to contaminant levels in Canadian Atlantic cod (*Gadus morhua*). Can. J. Fish. Aquat. Sci. 44: 859-865.
- Misra, R.K., J.F. Uthe and D.B. Swetnam. 1988. Examination of ICES contaminant data sets: Sweden-Kattegat (43G1) and Belgium Coast (31F2) for time trends. Annex 3. ICES Statutory Meeting, Paper C.M. 1988/E:27.
- Morrison, D.F. 1976. Multivariate Statistical Methods. McGraw-Hill Book Co. New York, NY.
- Muller, K.E. and B.L. Peterson. 1984. Power analysis for multivariate linear models: New results and SAS makes it practical. In: 1984 SUGI Proceedings. Cary, NC: SAS Institute.
- Ramsey, P.H. 1982. Empirical power of procedures for comparing two groups on p variables. J. Educational Statistics. 7: 139-156.
- Scott, D.P., J.F. Uthe and C.L. Chou. 1981. Further considerations of time trend determinations of contaminant levels in Canadian cod. ICES Statutory Meeting, Paper C.M.1981/E:47.
- Seal, H. 1964. Multivariate Statistical Analysis for Biologists. John Wiley & Sons Inc., New York, NY.
- Srivastava, M.S. and E.M. Carter. 1983. An Introduction to Applied Multivariate Statistics. North-Holland, New York, NY.
- Stevens, J.P. 1980. Power of the multivariate analysis of variance tests. Psych. Bull. 88: 728-737.

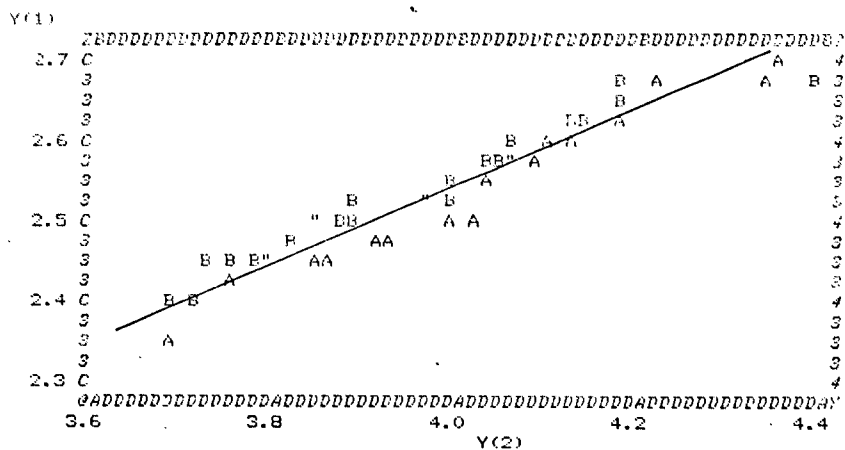


Fig. 1, Scatterplot of Y_1 versus Y_2 ($r_{12} = 0.98$) - flounder data.
Legend: 'A' for 1981 data and 'B' for 1983 data.

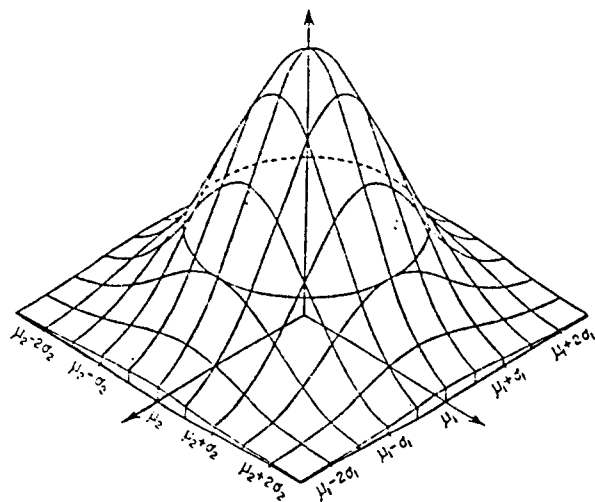


Fig. 2; A bivariate normal distribution.

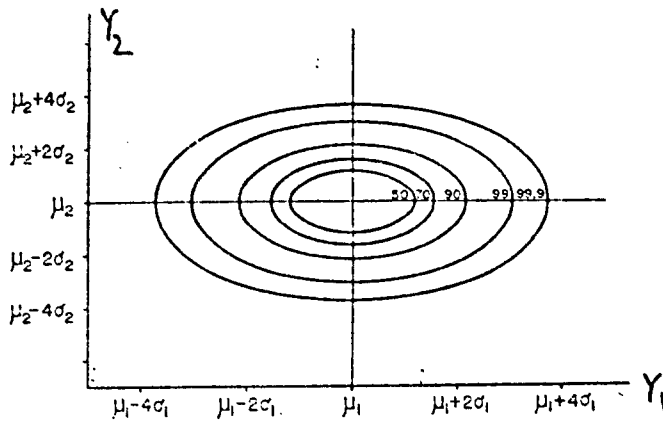


Fig. 3; Concentric ellipses in the (Y_1, Y_2) plane.

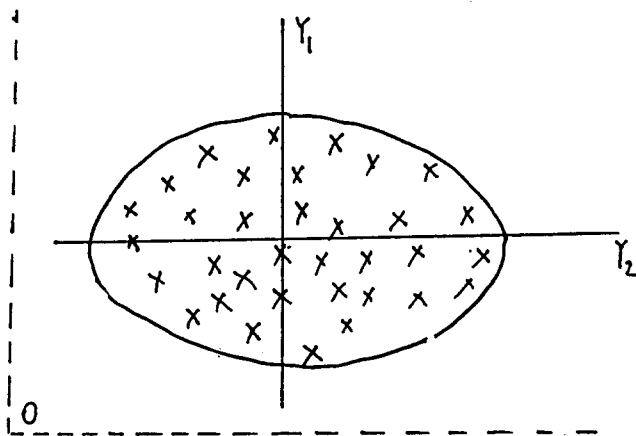


Fig. 4 ; Scatterplot when Y_1 and Y_2 are uncorrelated.

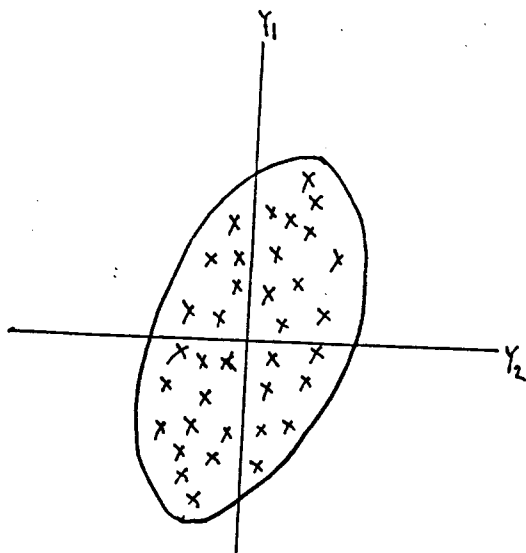


Fig. 5; Scatterplot when Y_1 and Y_2 are correlated positively.

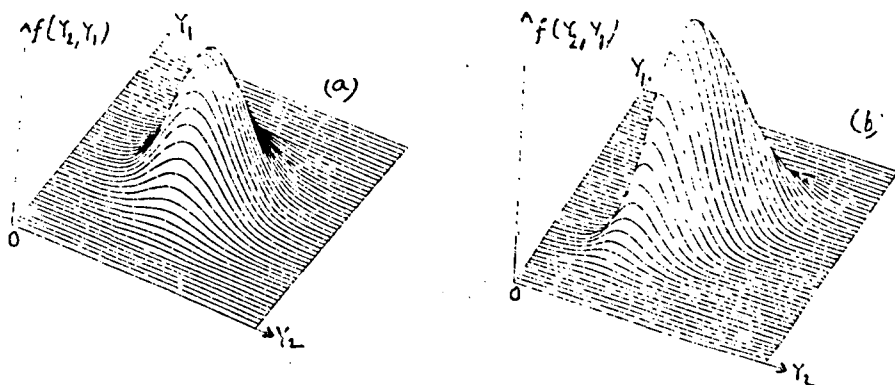


Fig. 6; Additional view of a bivariate normal distribution.
(a) $r_{12} = 0$ and (b) $r_{12} = 0.75$

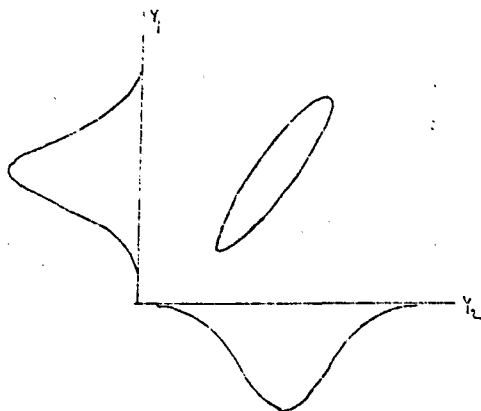


Fig. 7; A bivariate frequency distribution.

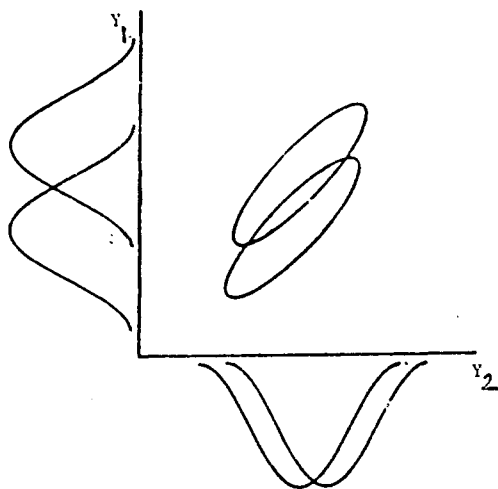


Fig. 8; Two bivariate frequency distributions.

ANNEX 8

ON TIME TRENDS, 1977-1985,
IN CANADIAN COD ATLANTIC (GADUS MORHUA)

by

R.K. Misra, J.F. Uthe, and C.L. Chou

Marine Chemistry Division
Department of Fisheries and Oceans
P.O. Box 550, Halifax, N.S., Canada B3J 2S7

The need for multivariate analysis of variance (MANOVA) and covariance (MANCOVA) in preference to the univariate analysis of variance (ANOVA) or analysis of covariance (ANCOVA) in time trend investigations was clearly shown in the study by Misra and Uthe (1989-this meeting). However, in the example given the application to time trend analysis was not shown. In this paper we do so using the Canadian Atlantic cod data.

Our investigation of time trends in the Canadian coastal environment started in 1977. A number of contaminants in Canadian Atlantic cod (*Gadus morhua*) from the southern Gulf of St. Lawrence were assessed employing a length-stratified sampling strategy spread over as wide a range as practicable (Scott et al. 1978). Misra and Uthe (1987) analysed 3-year (1977, 1978, and 1979) cod data by multivariate analysis of covariance (MANCOVA). In the *ad hoc* statisticians' meeting in 1984 (Uthe et al. 1984) continuation of this investigation beyond the three-year period was recommended, to determine if this species could be used to follow long-term trends in coastal chemical contamination. The need for one or two additional samples, with longer than one-year intervals between samplings, was identified. Following this recommendation an additional sample was collected in 1985. Misra et al. (1988) reported to WGSATM the results of their study of time trends over the 8-year period of 1977-85 based upon MANCOVA of data for 1977, 1978, 1979, and 1985.

Materials and Methods

For information on the basic data and materials and methods see Scott et al. (1978, 1983) and Misra and Uthe (1987). Concentration data analysed by Misra et al. (1988) consisted of measurements on each fish of ten contaminant-tissue combinations designated as Y_i , $i = 1, \dots, 10$ (after transformation to their common logarithms), for MANCOVA. These were: Y_1 (Zn-M), Y_2 (As-L), Y_3 (Cd-L), Y_4 (Cu-L), Y_5 (Hg-L), Y_6 (Se-L), Y_7 (Zn-L), Y_8 (PCB-L), Y_9 (α -HCH-L), and Y_{10} (HCB-L), where -L is liver and -M is muscle. In addition to these, six biological characters, including length (cm), were measured (Misra and Uthe 1987). The "best" multiple linear regression (MLR) has not been established, therefore Misra et al. (1988) employed one covariate at a time in their MANCOVAs on Y_i . In the following, we present a part of that MANCOVA for time trends where the covariate, X , was log length.

Table 1 gives sample sizes, means, and ranges of X and Y_i for concentration data. Vector \underline{b} of coefficients b_i , $i = 1, \dots, 10$, of regression of Y_i on X was computed from "within year" SS and SP. Null hypothesis $\underline{\beta} = 0$ (where $\underline{\beta}$ is the parameter vector for \underline{b}) was rejected (probability level, P , of significance = 0+, i.e.. close to zero), showing thereby that MANCOVA is warranted, i.e. years should be compared based on their mean Y_i after adjusting these for variations in the values of the covariate X among individuals. Adjusted year-means, Y_i , are shown in Table 2. Note that these adjusted means would be the same for ANCOVA and MANCOVA. Overall MANCOVA null hypothesis, H_0 , (step 1, referred earlier) that years do not differ in their vectors of adjusted means of Y_i , was rejected ($P = 0+$). This indicated that significant linear combinations of groups (years) and/or variables (Y_i) existed. Following this, examination of the linear time trend vector of adjusted year-means was one of the analyses done by Misra et al. (1988) in the Step 2 MANCOVA. This was significant ($P = 0+$). Contributions

of individual Y_i to the time trend vector was examined on the basis of 95% simultaneous confidence intervals. Table 3 shows the lower and upper confidence limits for each Y_i . A Y variable would contribute significantly to the linear time trend if its two limits were of the same sign (and thus did not enclose zero value). Also, increasing or decreasing trends would be indicated if the two limits are both positive or negative, respectively. Using these guidelines, contaminants with significant contributions to the time trend were identified as follows: (a) With both limits positive: Cd-L (Y_3), Cu-L (Y_4), and Hg-L (Y_5), (b) With both limits negative: Zn-L (Y_1), PCB-L (Y_8), and HCB-L (Y_{10}). These findings are corroborated by visual examination of increasing and decreasing trends of adjusted year-means of individual contaminants (Figure 1).

In conclusion, this paper has been prepared in response to the comment made at the 1988 meeting of WGSATM that MANCOVA identifies too many sources of variation as significant. This has not been true in this study where contaminants contributing significantly to time trend in MANCOVA were the same as identified by visual observation for univariate analysis.

References

- Misra, R.K. and J.F. Uthe. 1987. Methods of time trend analysis applied to contaminant levels in Canadian Atlantic cod (*Gadus morhua*). Can. J. Fish. Aquat. Sci. 44: 859-865.
- Misra, R.K., J.F. Uthe, C.J. Musial and C.L. Chou. 1988. The analysis of time trends in contaminant levels in Canadian Atlantic cod (*Gadus morhua*). 5. Time trends, 1977-1985, employing a multivariate linear model. ICES Statutory Meeting, C.M.1988/E:3.
- Scott, D.P., J.F. Uthe and C.L. Chou. 1978. Regression analysis of heavy metal and organochlorine residue concentrations in a statistically sampled population of Atlantic cod (*Gadus morhua*). ICES Statutory Meeting, C.M. 1978/E:16.
- Scott, D.P., Uthe, J.F. & Chou, C.L. 1981. Further considerations of time trend determinations in contaminant levels in Canadian cod. ICES Statutory Meeting C.M. 1981/E:47.
- Scott, D.P., J.F. Uthe and C.L. Chou. 1983. Choice of variables in estimation of time trends in contaminant levels in fish. Paper presented to the ICES Working Group on Marine Pollution Baseline and Monitoring Studies in the North Atlantic, Copenhagen, 1-4 February, 1983.
- Uthe, J.F., C.L. Chou, M.M. Hansen, H. Lassen, M.D. Nicholson, D.J. Scarrett, D.P. Scott and A. Sreedharan. 1984. Report on the *ad hoc* meeting of statistician's workshop on contaminant trend monitoring in fish stocks. ICES Statutory Meeting C.M. 1984/E:46.

Table 1. Mean, Minimum, and Maximum values for X and Y Variables for Concentration Data.

Year	1977 (n = 34)			1978 (n = 44)			1979 (n = 45)			1985 (n = 35)		
Variable	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Y ₁ Zn-M ($\mu\text{g}/\text{kg}$)	3.624	3.526	3.746	3.615	3.546	3.743	3.669	3.555	3.892	3.556	3.389	3.644
Y ₂ As-L ($\mu\text{g}/\text{kg}$)	0.291	0.076	0.599	0.243	-0.004	0.656	0.378	0.097	0.658	0.349	-0.032	0.728
Y ₃ Cd-L ($\mu\text{g}/\text{kg}$)	1.776	1.342	2.423	2.179	1.763	2.746	1.844	1.415	2.412	2.396	2.093	2.831
Y ₄ Cu-L ($\mu\text{g}/\text{kg}$)	0.602	0.255	1.194	0.619	0.057	0.978	0.806	0.037	1.438	0.845	0.598	1.239
Y ₅ Hg-L ($\mu\text{g}/\text{kg}$)	1.358	0.903	1.690	1.426	1.000	2.000	1.325	1.000	1.914	1.781	1.447	2.117
Y ₆ Se-L ($\mu\text{g}/\text{kg}$)	2.915	2.763	3.117	3.222	2.699	3.705	3.129	2.724	3.597	2.973	2.472	3.179
Y ₇ Zn-L ($\mu\text{g}/\text{kg}$)	1.185	1.013	1.365	1.182	0.898	1.387	1.265	1.014	1.511	1.224	0.969	1.452
Y ₉ PCB-L ($\mu\text{g}/\text{kg}$)	0.274	0.004	0.719	0.463	0.124	0.809	0.424	0.143	0.872	0.162	-0.194	0.873
Y ₉ α -HCH-L ($\mu\text{g}/\text{kg}$)	1.365	1.491	2.076	1.315	1.342	1.997	1.809	1.255	2.025	1.752	1.431	2.013
Y ₁₀ HCB-L ($\mu\text{g}/\text{kg}$)	1.631	1.491	2.067	1.658	0.845	1.903	1.624	1.114	1.826	1.330	0.845	1.556

Table 2. Adjusted Year-Means of Y Variables.

Variable	Year			
	1977	1978	1979	1985
Y ₁	3.700	3.695	3.748	3.636
Y ₂	-0.269	-0.343	-0.206	-0.231
Y ₃	0.545	0.891	0.558	1.119
Y ₄	-0.069	-0.081	0.107	0.152
Y ₅	-0.133	-0.131	-0.229	0.237
Y ₆	2.172	2.246	2.355	2.203
Y ₇	0.991	0.978	1.062	1.023
Y ₈	-1.840	-1.747	-1.781	-2.027
Y ₉	2.513	2.491	2.484	2.462
Y ₁₀	1.550	1.574	1.540	1.247

Table 3. Lower and Upper Confidence Limits for Individual Contaminants in the Time Trend Contrast Based on 95% Simultaneous Confidence Intervals.

Variable	Lower Limit	Upper Limit
Y ₁	-2.5429	-0.3829
Y ₂	-1.7934	4.0283
Y ₃	5.4114	13.7729
Y ₄	0.3253	8.0500
Y ₅	4.1077	11.9895
Y ₆	-3.2559	2.4278
Y ₇	-1.3144	2.4038
Y ₈	-7.5194	-2.0105
Y ₉	-3.2309	1.5976
Y ₁₀	-9.7174	-3.2946

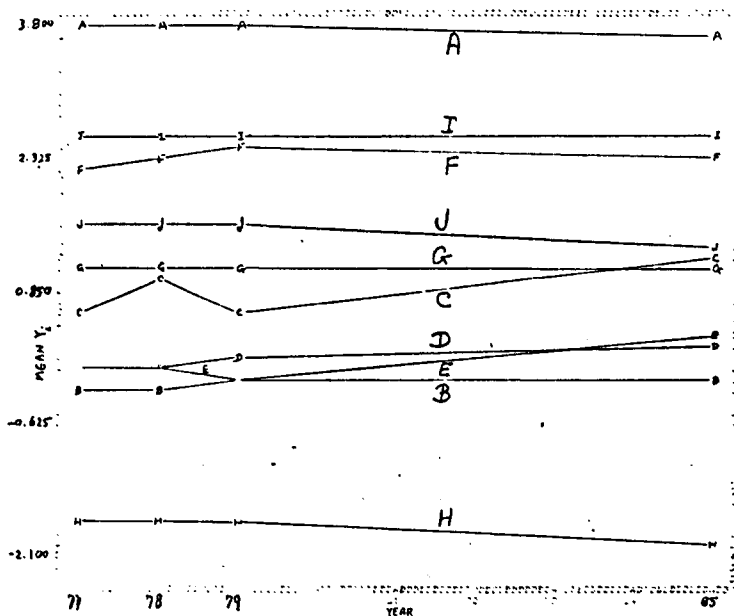


Figure 1. Plot of adjusted year-means of Y_i versus sampling year. ("A" is Y_1 (Zn-M), "B" is Y_2 (As-L), "C" is Y_3 (Cd-L), "D" is Y_4 (Cu-L), "E" is Y_5 (Hg-L), "F" is Y_6 (Se-L), "G" is Y_7 (Zn-L), "H" is Y_8 (PCB-L), "I" is Y_9 (α -HCH-L) and "J" is Y_{10} (HCB-L))

ANNEX 9**METHODS FOR COLLECTING BENTHIC ALGAE FOR CONTAMINANT MONITORING**

by

Norman Green and Martin Munk Hansen

1. Background

The meeting of the ICES Working Group on Statistical Aspects of Trend Monitoring (WGSATM) was held 12-20. April 1988 in Copenhagen. WGSATM considered a request by another ICES working group to describe the methodology for using seaweeds in connection with contaminant monitoring. WGSATM noted that there was not resources to make a complete overview but that Munk Hansen from the Greenland Environment Research Institute (GERI) and Green from the Norwegian Institute for Water Research (NIVA) would undertake to describe what methods were applied at their respective institutes as well as their experience in using seaweeds as monitoring organisms.

2. Biological properties of benthic algae related to their use as monitoring organisms

Benthic algae are plants located at water depths where light intensity permits photo synthesis. Due to practical reasons the species normally used for monitoring purposes are the ones in the intertidal zone. The algae are fixed to the bottom, normally to hard substrates as rock or stones. Some species can become rather old. For the knotted wrack (*Ascophyllum nodosum*), where age can be determined easily, shoots of 10-15 years are common.

Benthic algae including species in the intertidal zone occur in the northern hemisphere in the temperate zone and up to the high arctic. Some species have a large geographical distribution.

Benthic algae take up all nutrients from the surrounding water. Thus no nutrients are absorbed from the substrate to which the plant is anchored. The shoots grow at the tip for the indicator organisms mentioned below. However growth also occurs along the length of the shoots. The shoot tips therefore are constituted of newly produced material and the amount of older material increases towards the base of the shoot.

3. Species used for monitoring

3.1 At NIVA two brown algae have been commonly used to monitor spatial and temporal changes in contaminants (mostly metals, in a few cases phosphorus, nitrogen and PAH):

egg or knotted wrack	<u>Ascophyllum nodosum</u>
bladder wrack	<u>Fucus vesiculosus</u>

To a minor extent also serrated wrack (Fucus serratus) and the green algae Cladophora spp. and Enteromorpha spp. have been utilized for metal monitoring.

3.2 At GERI brown algae have been used for monitoring of spatial and temporal trends in metal pollution from mining operations. The species are:

bladder wrack	<u>Fucus vesiculosus</u>
(no English name)	<u>Fucus distichus</u>

Furthermore the above mentioned species and knotted wrack (Ascophyllum nodosum) have been collected over three years at four unpolluted localities in order to study the natural variations in the concentrations of elements.

4. Sources of variability of monitoring data

The uptake and release of metals in algae has not been adequately investigated in regards to season and the plant's age. The same applies to individual variations (in review by Knutzen 1985, (in Norwegian) includes discussions on background levels, biomagnification and the effect of: seasonal variation, tissue age, salinity, individual differences, place of growth on the seabed, metal speciation and metal binding properties.)

Knutzen (1985) identified the following sources of variation when using algae for the purpose of monitoring metals:

- The uptake and release of metals can be extremely slower during winter than during summer in temperate regions.
- The concentration of metals can increase with age.
- The standard deviation of individual variation is on the order of 10-30 %.

Studies in Danish waters (Meyer and Kjær 1988, in Danish) have shown that the speed of growth is very variable over the year and between stations. Thus length growth has been shown to be two times larger at a locality at the north coast of Zealand than at localities in the Sound near Copenhagen in the period from April to September, while the growth September-October was up to five times larger at the Copenhagen sites. Since metal concentrations in brown algae are dependent on age of the analyzed part of the shoot such variations in growth rate may affect the results when using algae for monitoring spatial trends. However the monitoring of temporal trends would be less influenced by locality differences in growth rate.

5. Practical guidelines for monitoring using benthic algae

5.1 Sampling period

Collect samples at the same time of year for best results in regards to studies on temporal variation or regional comparisons. Collection during late autumn to early spring should be avoided if the purpose is to monitor short term changes (<1 year).

5.2 Sampling localities

Depending on the purpose of the monitoring, samples can be collected on small islands or at the tip of peninsulas (to monitor pollutants from a distant source) or close to the point source of pollutant, e.g. a river mouth.

NIVA collects samples within a zone homogenously distributed with wrack and for a horizontal distance of 10 - 30 m.

GERI collects each sample within a horizontal distance of c. 5 meter.

5.3 Number of samples per locality

To determine the significance of differences in the concentrations of pollutants spatially or temporally it is necessary to know the variance in the measured quantities. To do that more than one sample per locality is necessary. GERI collects 2 - 3 samples at all localities, each sample being collected within a horizontal distance of c. 5 meter.

5.4 Selection of samples

Bulk each sample from several individual plants preferably 10-20 plants and 10 shoots from each plant.

The sample should be free of epiphytes. If epiphytes are unavoidable these have to be removed and a note should be made to this respect.

Old or deformed portions of the plant should not be collected.

Fertile portions (receptacles) should not be collected.

5.5 Preparation of sample

Select particular section of the alga depending on the purpose of monitoring. NIVA routinely selects the upper portion of knotted wrack (Ascophyllum nodosum); from the tip to just over the second bladder. This portion corresponds to the time period 0 to <2 years past, usually up to 18 months for samples collected during the autumn. Where short term changes are anticipated the algae are partitioned into year classes; current years growth corresponds to the shoot above the first (uppermost) bladder. One bladder is formed each spring. For bladder wrack (Fucus vesiculosus) NIVA uses the upper 5-10(15) cm which corresponds to the shoots of recent growth (1-2 years). For this species GERI selects the upper 4-8 cm. This length probably covers the same time span as

the part collected by NIVA due to slower growth in the arctic.

NIVA rinse the samples in ambient seawater while GERI washes the samples twice in deionized water to remove as much salt water as possible. This is done because the ions in the salt water upon drying of the seaweed will be concentrated to levels that may interfere with some methods of analysis. Before storage the samples are drip dried.

5.6 Storage of samples

Samples are placed in plastic bags which do not contain contaminating substances. Polyethylene is preferred. The bags should be appropriately marked with station identification, date of collection and species (it is often difficult to identify species later in frozen condition). The samples should be frozen as soon as possible, within 24 hours provided the samples are kept cool and out of the sun.

5.7 Analysis of samples

Wet weight determination of samples of brown algae is very inaccurate because the algae produces excessive amounts of mucus while being drained. Therefore analyses are expressed on a dry weight basis. GERI and NIVA freeze dry the samples and homogenization is performed in an agate ball mill.

6. Comparison of monitoring properties of benthic algae and mussels

Mussels have been the most commonly used monitoring organisms in the intertidal zone. Using benthic algae as monitoring organisms is another way of monitoring the same marine environment. To decide whether a monitoring program should be based on mussels or algae or both it is necessary to evaluate the qualities of both types as quantitative indicators.

6.1 Integration time

When mussels are used for trend monitoring it is recommended to use small individuals. For practical reasons the soft parts of mussels smaller than c. 2 cm shell length are difficult to sample in sufficient quantities for analysis. In southern Europe a length of 2 cm may be reached in less than a year, but in

the arctic 2 cm mussels are 6-8 years old. Some metals are bound very firmly in mussels, and the use of individuals several years old may reflect the metal levels in the environment integrated over several years. The shoot tips of benthic algae, which is the material the present guidelines recommend, can everywhere be selected to reflect a short time period thus enabling the monitoring of short term changes.

6.2 Experiences with different elements.

At GERI mussels and brown algae have been used simultaneously on a number of occasions. This has given some insights into how the two types of indicators react to different elements.

6.2.1 Zinc and copper

For two elements zinc and copper, brown algae have been found to be superior to mussels as monitoring organisms. It seems that mussels very poorly reflect the environmental concentrations of these elements, and the reason is probably that their internal levels are physiologically controlled.

6.2.2 Lead

Brown algae as well as mussels are useful indicators for lead. The results obtained by the two types, however, differs somewhat.

A difference in spatial distribution is observed several places. The most clearcut is at Ivigtut where a point source of dissolved lead is found. The enrichment factors at the source compared to background levels are about the same for mussels and algae. However the concentrations in the algae decrease close to the source while the concentrations in the mussels decrease more slowly with distance. This may be interpreted as the result of different uptake characteristics: The algae only accumulating dissolved lead and mussels accumulating dissolved as well as particulate lead. The dissolved lead in the seawater close to the source quickly may be connected to organic and inorganic particles thereby leaving it unavailable for the algae. However it would still be available for the mussels feeding on particles suspended in the water.

A difference between temporal trends in mussels and algae has been observed at Marmorilik where lead pollution has been decreasing during the last 10 years. This decrease is very pronounced in algae but smaller in the mussels. The reasons behind the difference may be the long integration time for the mussels as described in section 6.1, but differences in time of the proportions of dissolved and particulate lead can not be excluded.

6.2.3 Various rare elements

At naturally occurring point sources of a number of rare elements in southern Greenland the accumulation in mussels and algae has been studied. Mussels reflect expected geographical distributions of elements as rare earth elements (lanthanum, cerium, samarium, neodymium and others) and hafnium. The geographical trends are not observed in algae.

7. Future work

The intention of this paper has been to describe NIVA and GERI methods and experiences when using benthic algae as monitoring organisms. In order to prepare a published set of guidelines for the use of benthic algae for monitoring purposes a review of existing literature should be made. A review should address subjects as:

- uptake and release characteristics of elements
- effects of environmental conditions (salinity, temperature, nutrients)
- seasonality
- natural variability
- suitability of different species

We would recommend that the working group initiate the preparation of a review as outlined. If that could be finished before the 1990 meeting of the group a finalized set of guidelines probably could be established in 1990.

8. Literature cited

Knutzen, J., 1985. "Bakgrunnsnivåer" av utvalgte metaller og andre grunnstoffer i tang. Øvre grense for "normalinnhold", naturbetingede variasjoner, opptaks- og utskillelsesmekanismer.

NIVA-rapport 0-83091, 121 sider.

Meyer, K.S. and A. Kjær, 1988. Årstidsvariation i vækst, morfologi og reproduktion samt unge planters vækst hos *Fucus vesiculosus* L.

Final exam paper, University of Copenhagen.

ANNEX 10

POOLING MAY ECONOMIZE A SAMPLING PROGRAM

by

Jaap van der Meer

Abstract

When estimating concentrations of contaminants in marine organisms, the confidence of these estimates can be improved considerably by the pooling of organisms before chemical analysis. Minimization of a loss function with constraints, given in this paper, will give the optimal number of pools and the optimal number of organisms within a pool. Due to the rather flat appearance of this function, the following rule of thumb might work well: avoid either the strategy of no pooling or the use of only a few large pools.

Introduction

When considering a programme of sampling and chemical analysis of contaminants in marine organisms the question is often raised: should individual specimens have to be pooled before analysis?

Statistical theory can contribute to answer this question and help the chemist to economize his sampling study. In the recent past the WGSATM has given considerable attention to this problem (Lassen 1983, 1984, Anonymous 1986a, 1986b, Misra 1987, Uthe & Chou 1987, Bignert & Nielsen 1988), mainly in connection with the use of biological covariates. It was concluded that pooling is an attractive strategy, if only one covariate, e.g. length is involved. In such case length stratification with small intervals should be applied, the number of organisms in a pool should be constant and the weights of the organisms within a pool must have small variance. Due to the interrelationship between length and weight the latter will probably be true.

However, little attention was paid to the final choice of the number of pools and of the number of organisms within a pool. This study will show on what grounds this choice can be made and that statistics can be helpful in economizing a sampling programme and thus saving costly resources.

A model

In the following the effect of log-transformation of concentrations (c.f. Nicholson et al. 1989) and the presence of outliers will be ignored. The use of covariates and variation in the weight of organisms will be dealt with in the next section.

Consider the following model:

$$\hat{y} = \mu + \epsilon_1 + \epsilon_2$$

$$\epsilon_1 \sim \text{IND} (0, \sigma_1^2/nm)$$

$$\epsilon_2 \sim \text{IND} (0, \sigma_2^2/m)$$

where

\hat{y}	estimated concentration
μ	population parameter
n	number of specimens in a pool
m	number of pools
σ_1^2	variance among specimens
σ_2^2	variance among chemical analyses

The objective of each sampling programme should be to maximize the power of some planned statistical test given certain estimated costs or to minimize the costs in achieving a desired power.

Let us consider the first objective. For reasons of convenience and without loss of generality the objective of maximizing the power will be restated as minimizing the confidence interval of the mean.

Define c_1 (in some monetary or arbitrary unit) as the costs of using an extra organism and c_2 as the costs of conducting an extra chemical analysis. So the total marginal costs equal:

$$c = n.m.c_1 + m.c_2$$

The confidence interval of \hat{y} equals:

$$2.t_{\alpha, m-1} \cdot s_{\hat{y}}$$

So the objective can be written as follows

$$\text{minimize } t_{\alpha, m-1} \cdot \sqrt{\sigma_1^2/nm + \sigma_2^2/m}$$

with respect to n and m , given the constraint

$$n.m.c_1 + m.c_2 < c$$

(some particular estimate s_y is replaced by σ_y).

If we know c , c_1 and c_2 and assume values for σ_1 and σ_2 the problem can be solved iteratively.

Covariate included

Given the model

$$\hat{y} = \mu + \beta(x - \bar{x}) + \delta_1 + \epsilon_2$$

where

$$\delta_1 \sim \text{IND} (0, (1-R^2) \cdot \sigma_1^2)$$

This results in

$$\sigma_{\hat{y}}^2 (\text{at } \bar{x}) = \sqrt{(1-R^2) \cdot \sigma_1^2 / nm + \sigma_2^2 / m}$$

where

$$R^2 = \text{explained fraction of the variance}$$

If there are more covariates n will in general be equal to 1. So we might include the possibility of using covariates in our optimization problem. Define the costs of including a covariate as $c_3 \cdot n \cdot m$

Weight differences within a pool

Lassen (1984) showed that

$$\sigma_{\text{pool}}^2 = \Sigma w_i^2 / (\Sigma w_i)^2 \cdot \sigma_1^2$$

which, if $w_i = w_j$ for all i and j , reduces to

$$\sigma_{\text{pool}}^2 = \sigma_1^2 / n$$

rewriting gives

$$\sigma_{\text{pool}}^2 = [(\sigma_w / \mu_w)^2 + 1] \cdot \sigma_1^2 / n$$

So if the coefficient of variation of the weights within a pool is known, the possibility of non-stratified sampling can be included in the optimization problem. In such cases we could define the costs of stratified sampling as $c_4 \cdot n \cdot m$

Resources as discrete units

Until now c_1 and c_2 have been assumed to be constant. However, in one survey one is only able to catch, as an example q fish or mussels. If one wishes to catch more one must organize another sampling cruise (ship day). The same might be true for the chemical analysis. If you want to analyse more samples than a single chemist can handle (r analyses), then you must pay another chemist full time (figure 1).

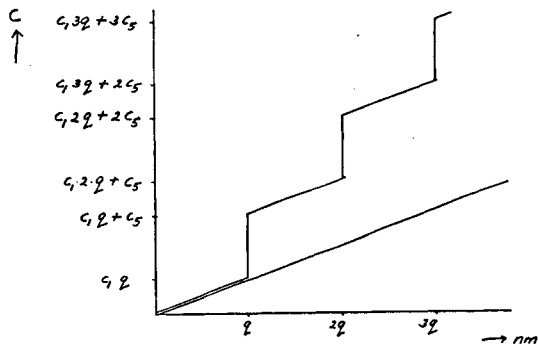


Figure 1. Two examples of a cost function.

$$\text{So } c = c_1 \cdot n \cdot m + c_5 \cdot \text{INTEGER}(n \cdot m / q) + c_2 \cdot m + c_6 \cdot \text{INTEGER}(m / r)$$

However, if

$$c_1 \cdot q \ll c_5$$

$$c_2 \cdot r \ll c_6$$

then the cost function simplifies again, to

$$c = k \cdot c_5 + l \cdot c_6$$

where

- k number of ship days
- l number of chemists

General solution

Given $c_1 - c_6$, σ_1 , σ_2 , R^2 and CV, we wish to minimize

$$t_{\alpha, m-1} \cdot \sqrt{(1-R^2) \cdot (CV^2+1) \cdot \sigma_1^2/nm + \sigma_2^2/m}$$

under the constraint

$$n \cdot m \cdot c_1 + m \cdot c_2 + g \cdot n \cdot m \cdot c_3 + h \cdot n \cdot m \cdot c_4 < c$$

where g is equal to the number of covariates (if $g=0$ then $R^2=0$) and h equals 1 in the case of stratified sampling ($CV=0$) and otherwise equals 0. Alternatively one can use the more complicated cost function. The calculations can easily be done in a spread-sheet program.

Examples

The first example uses the simplest cost function, with the values $c=200$, $c_1=1$, $c_2=10$ and $\sigma_1=10$.

Figure 2 shows that not-pooling ($m=19$) or using large pools ($m=2$) are both inefficient in terms of variance estimates. A rule of thumb might be: keep away from these extremes.

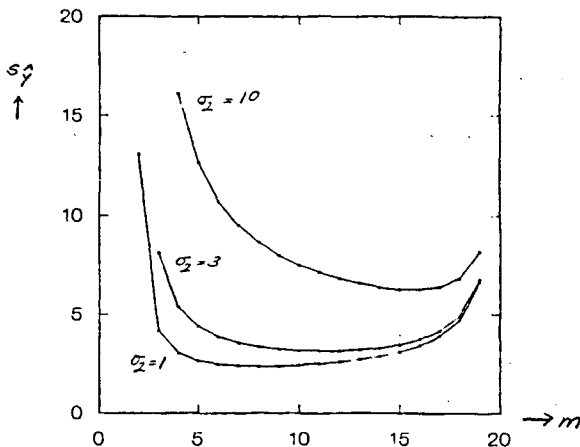


Figure 2. Standard error of the estimated concentration (s_y) related to the number of pools (m) for three values of the variance among chemical analyses (σ_2^2). See text for further explanation.

The second example includes one covariate. In this particular case ($\sigma_2=3$, $c_3=1$ and other parameters as in the first example) there is no use in measuring the covariate if the explained fraction of the variance is less than .5.

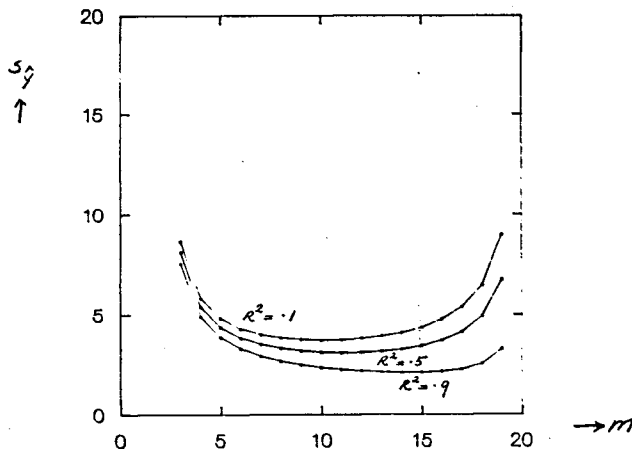


Figure 3. Standard error of the estimated concentration (s_y) related to the number of pools (m) for three values of the variance explained by the covariate (R^2). The curve with $R^2=.5$ equals the one without using a covariate. See text for further explanation.

Literature

- Anon., 1986a. Pooling strategy as applied to the Belledune lobster contamination problem. In: Annex 6 of the Report of the Meeting of the ad hoc Group of Statisticians assisting the Working Group on Marine Pollution Baseline and Monitoring Studies in the North Atlantic on Trend Monitoring. ICES, Doc. C.M.1986/E:39.
- Anon., 1986b. Pooling of samples. In: Section 3.2 of the Report of the Meeting of the ad hoc Group of Statisticians assisting the Working Group on Marine Pollution Baseline and Monitoring Studies in the North Atlantic on Trend Monitoring. ICES, Doc. C.M.1986/E:39.
- Bignert, A. and P.B. Nielsen. 1988. Some tests with artificially pooled samples from M. edulis. In: Annex 6 of the Report of the 1988 Meeting of the Working Group on Statistical Aspects of Trend Monitoring. ICES, Doc. C.M.1988/E:27.
- Lassen, H. 1983. Multivariate analysis applied to monitoring time trends in chemical contaminant levels in fish. ICES, Doc. C.M.1983/D:4.
- Lassen, H. 1984. In: Uthe, J.F. et al. Report of the ad hoc Statisticians' Workshop on Contaminant Trend Monitoring in Fish Stocks. ICES, Doc. C.M.1984/E:46.
- Misra, R.M. 1987. Some implications of statistical analysis for time trend when data are pooled. In: Annex 6 of the Report of the 1987 Meeting of the Working Group on Statistical Aspects of Trend Monitoring. ICES, Doc. C.M.1987/E:24.
- Nicholson, M., N. Green and S. Wilson. 1989. Trend measurements from bulked samples. The effect of pooling on lognormally distributed contaminants. In: Annex 4 of this Report.
- Uthe, J. and C.L. Chou. 1987. The use of pooled samples in measuring time trends in contaminant levels in Marine biota. In: Annex 7 of the Report of the 1987 Meeting of the Working Group on Statistical Aspects of Trend Monitoring. ICES, Doc. C.M.1987/E:24.

ANNEX 11

EFFECT OF SAMPLING STRUCTURE ON REGRESSION ESTIMATES:

Why the CMP Sampling Scheme Should be Followed

by

M.D. Nicholson and S.J. Wilson

Introduction

The ICES Cooperative Monitoring Programme (CMP) for fish and shellfish outlined a sampling scheme for using these organisms to monitor trends in contaminant levels (Anon., 1983). The scheme recommended that for fish, a length-stratified sample of at least 25 individuals should be collected annually at the same time and place.

A procedure for analysing the data collected under the CMP scheme (Anon., 1986) assumed:

- (1) log(contaminant concentration) is linearly related to fish length.
 - (2) homogeneity of regression slopes
 - (3) Normal noise with constant variance.
- and
- (4) Additional covariates such as weight, age or sex could be omitted.
- The procedure included steps to establish whether or not the assumptions were met.

When this procedure was applied, there was surprising inconsistency in the adequacy of the proposed statistical model (Nicholson and Wilson, 1987). Although the length effect was linear when present, it was not consistently present, even for subsets of species or contaminants. Further, about half of the data sets showed the regression coefficient of length fluctuating from year to year.

This effect could be a consequence of poor data quality. For example, as described below, it could arise if the sampled length range varied between years. Nicholson and Wilson (1987) examined variation in length sampling, and attempted to relate this to the results of the analyses of the CMP data.

This note looks in more detail at mechanisms that could lead to spurious heterogeneity of regression slopes, and re-examines the CMP data for evidence of these mechanisms.

2. Regression Models For Trend Assessment

2.1 Simple Regression Model

Writing $y = \log(\text{contaminant concentration})$
 $x = \text{length}$

then

$y = a + \beta x + \text{error}$
 and $\text{variance}(\text{error}) = \sigma_e^2$.

If this model is correct, then y can be predicted at a particular value of x - usually the overall mean length observed in the data. We can then look at the way the length-adjusted predicted values change from year to year.

This simple, and empirical, model becomes more difficult to apply if the regression coefficient, β , varies from year to year. There are really two problems. The first is what does it mean if β truly varies? The second is that β may appear to vary because of the distribution of the data.

2.2. Suppose β truly varies

What does it mean if, for example, contaminant concentration in a species decreases with fish length in 1988, but increases in 1989? This effect must surely be due to some complex chemical or physical phenomenon; eg interaction between contaminants or mixing of various sub-stocks, each having its own history of exposure. Levels going up or down in the fish population may have no simple or direct correspondence with levels in the environment.

Further, of what value are the measured trends? The predicted length-adjusted concentrations will, by definition, be different except at the length at which the lines intersect. Testing for year effects adjusted to the overall mean length is a very specific test, implying that there is something particularly interesting about fish at that length. Yet this value may have arisen by chance, a consequence of vague sampling objectives in different years.

2.3. It is possible that β will appear to vary.

The second problem is that if the sampled length distribution in each year varies, the estimated β may vary accordingly. There are various ways that this can happen. Here is one: because length is usually measured to the nearest centimetre, the recorded length contains some measurement noise. The effect is to reduce the slope for the observed data to

$$\beta' = \beta(1 + \sigma_n^2 / \sigma_x^2)^{-1}$$

where σ_n and σ_x are the standard deviations of the noise and x respectively. Usually σ_n is small relative to σ_x , but consider Figure 1.

Here there are two years, and the length range in the second year is restricted to only large fish such that $\sigma_n/\sigma_x=1$. It is easy to show that the slope in the second year is $\beta/2$ and the ratio of the predicted concentrations at the overall mean length is given by

$$\exp[\beta/4(x_1-x_2)].$$

With β typically having a value near 0.1, and a distance between the two mean lengths of 40 cm, this gives a spurious trend of a 270% increase - when in fact there was none.

You could argue that this is a very extreme example, involving large extrapolations. True. But the point is that there are mechanisms that contrive with badly structured data to generate fluctuating regression slopes. Analysing the data as if the slopes were really different can produce alarmingly wrong trends.

3. A General Regression Model for Trend Data

The errors-in-variables model described above can be embedded in a more general model. Consider Figure 2. In this model, data within a single year consist of clusters of observations. Two clusters have been shown. The important feature is that there are now two relationships, the relationship between points within a cluster, and the relationship between clusters. Each has its own slope. i.e.

β_w = slope within a cluster

β_b = slope between clusters.

The overall relationship can be defined in terms of the within- and between-cluster covariance matrices between y and x . i.e.

$$\Sigma_w = \begin{vmatrix} \sigma_e^2 + \beta_w^2 \sigma_w^2 & \beta_w \sigma_w^2 \\ \beta_w \sigma_w^2 & \sigma_w^2 \end{vmatrix}$$

$$\Sigma_b = \begin{vmatrix} \beta_b^2 \sigma_b^2 & \beta_b \sigma_b^2 \\ \beta_b \sigma_b^2 & \sigma_b^2 \end{vmatrix}$$

giving

$$v[y] = \beta_b^2 \sigma_b^2 + \beta_w^2 \sigma_w^2 + \sigma_e^2$$

$$V[x] = \sigma_b^2 + \sigma_w^2$$

$$\text{and } \text{Cov}[yx] = \beta_b \sigma_b^2 + \beta_w \sigma_w^2.$$

The overall slope is thus a weighted average of the within- and between-cluster slopes i.e.

$$\begin{aligned} \beta &= \text{Cov}[yx]/V[x] \\ &= (\beta_b \sigma_b^2 + \beta_w \sigma_w^2) / (\sigma_b^2 + \sigma_w^2). \end{aligned}$$

We see that the ordinary regression model has σ_w^2 equal to zero. The length-stratified sampling scheme proposed for the CMP constrains non-zero values of σ_w^2 to be zero. The errors-in-variables model of the previous section has β_w equal to zero.

The general form could arise, for example, if the within-cluster effect was due to the biological interaction between contaminant level and fish length; the between cluster effect could be due to the history of exposure, for example if fish in the larger clusters had migrated across areas of high contaminant levels not visited by fish in the smaller clusters. These are only suggested as possibilities.

The observed regression coefficient now depends on the length distribution through σ_b^2 . If this is large, β will be dominated by β_b ; if small, by β_w .

4. Application to CMP data

We have

$$V[x] = \sigma_b^2 + \sigma_w^2.$$

or

$$\sigma_b^2 = V[x] - \sigma_w^2.$$

Then

$$\text{Cov}[yx] = \beta_b V[x] + (\beta_w - \beta_b) \sigma_w^2.$$

Hence, if β_b , β_w and σ_w^2 are constant within a data set but σ_b^2 has varied from year to year, a plot of observed values of $\text{Cov}[yx]$ against $V[x]$ should fall along a straight line with slope β_b and intercept $(\beta_w - \beta_b) \sigma_w^2$.

If the simple (and desirable) model of Section 2.1 is correct, the plot should have a well defined slope passing through the origin. This plot will also arise if $\beta_b = \beta_w$.

If there is no slope and the points are scattered around zero, there is no length effect at all.

If there is no slope but the points are scattered around some non-zero value, there is only a within-cluster length effect.

If there is a slope and a non-zero intercept, there is both a within- and between-cluster length effect but with different slopes.

The CMP data consist of 6 metals measured in 7 species by different countries monitoring different areas, although not in all metal-species combinations. To provide sufficient observations, data were analysed by contaminant and species, combining data from different areas and countries. This gave a total of 26 data sets.

The observed values of $\text{Cov}[y_x]$ and $V[x]$ were not available. However, from an earlier analysis, the quantities

$$b_{ij} \quad \text{se}(b_{ij}) \quad \text{and} \quad n_{ij}$$

were available for year i within area j where b_{ij} is the estimate of β_{ij} . Since

$$\beta = \text{Cov}[xy]/V[x]$$

$$\text{and} \quad V[b] = \sigma_e^2(n-1)/V[x],$$

then defining

$$\text{and} \quad Y = [b_{ij}/\text{se}^2(b_{ij})]/(n_{ij}-1)$$

$$X = [1/\text{se}^2(b_{ij})]/(n_{ij}-1)$$

a plot of Y against X should have a slope of β_b and an intercept of $(\sigma_w^2/\sigma_e^2)(\beta_w - \beta_b)$.

The results are given in Table 1. The first half of the table gives the slope (upper) and intercept (lower) from the least squares fit of Y on X . Numbers in bold type are significantly different from zero. The second half of the table gives the mean Y . Numbers in bold type are significantly different from zero, but the numbers in this half of the table were only tested if the slope in the first half of the table was not significant (otherwise enclosed in brackets).

5. Discussion

The only unambiguous results seen in Table 1 are for mercury, for which all species (except herring) show a

significant slope (i.e. $\beta_D \neq 0$) and a non-significant intercept (i.e. $\beta_D = \beta_W$ or $\sigma_W = 0$). This reflects the reported analyses of mercury in the CMP data (Wilson and Nicholson, 1987).

For the remaining 5 metals, there are no consistent results. Copper has some suggestion that either $\beta_D \neq \beta_W$ or $\sigma_W \neq 0$; zinc has some suggestion that $\beta_D = \beta_W$ or $\sigma_W = 0$. However, for the most part, metals other than mercury do not have a convincing, consistent length effect.

In some ways, these results are not surprising. For a particular contaminant-species combination data have been aggregated from different countries and different years. The assumption that σ_W is constant may not be realistic.

Also, inconsistency of size effect on trace metals other than mercury is already well documented (Phillips, 1980).

Clearly there is a difficulty. For metals other than mercury, we are trying to identify a wayward and intermittent size effect using data which may be poorly suited for the purpose. Therefore it is most important to follow the sampling guidelines with respect to length stratification. Keeping the sampled length range wide within years and constant between years will increase the likelihood of establishing the size effect without ambiguity.

Alternatively, for these metals it may be necessary to accept that there is no simple, robust, regression-based method of analysis that can be applied. It may be desirable to avoid the need for model based statistical analysis by adopting an appropriate sampling scheme whereby possible influential variables are held constant. Length stratified sampling would be one possible scheme. However, again it would be essential that the sampling scheme was followed.

References

- Anon. (1983) Report of the Working Group on Marine Pollution Baseline and Monitoring Studies in the North Atlantic. ICES CM 1983/E:3. 43pp. (mimeo).
- Anon. (1986) Report of the ad hoc group of statisticians assisting the Working Group on Marine Pollution Baseline and Monitoring Studies in the North Atlantic. ICES CM 1986/E:39. 45pp. (mimeo).
- Nicholson M.D. and Wilson S.J. (1987) Identification of trends in levels of metals in fish muscle: appraisal of the statistical analysis and of data quality. Appendix 3.2.

Report of the Working Group on the Statistical Aspects of Trend Monitoring. ICES CM 1987/E:24. 75pp. (mimeo).

Phillips D.J.H. (1980) Quantitative Aquatic Biological Indicators. Applied Science Publishers, London.

Wilson S.J. and Nicholson M.D. (1987) Statistical analysis of the ICES CMP data on contaminants in fish muscle tissue - for trend determination (purpose 3). Appendix 3.1. Report of the Working Group on the Statistical Aspects of Trend Monitoring. ICES CM 1987/E:24. 75pp. (mimeo).

Table 1. Summary of results. See text for explanation.

Metal	Species						
	Flounder	Sole	Cod	Plaice	Whiting	Dab	Herring
Mercury	2.98 0.014	3.60 -0.036	2.70 -0.006	1.40 0.021	5.19 0.052	10.44 -0.004	1.17 0.000
Copper	-1.37 0.037	-2.18 0.029	-0.45 -0.023	0.56 -0.125	1.86 -0.219		
Zinc	-3.86 0.036	-0.30 -0.008	-0.06 0.121	-0.73 -0.063	-0.82 -0.185		
Lead	-2.67 -0.012	-9.08 -0.001	0.70 -0.045				
Chromium	-8.97 0.085	1.94 -0.076	0.10 -0.109				
Nickel	-1.35 0.023	1.29 -0.015	-0.31 -0.131				
Mercury	(0.048)	(0.019)	(0.342)	(0.069)	(0.372)	(0.016)	0.001
Copper	-0.047	-0.082	-0.167	(-0.075)	-0.129		
Zinc	(-0.053)	-0.053	-0.069	(-0.246)	-0.489		
Lead	-0.047	(-0.057)	0.034				
Chromium	(-0.046)	-0.025	-0.096				
Nickel	-0.051	0.015	-0.259				

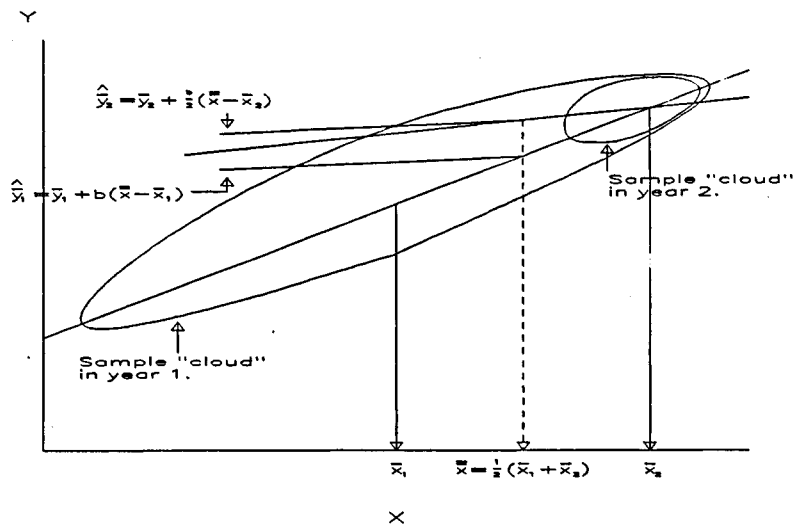
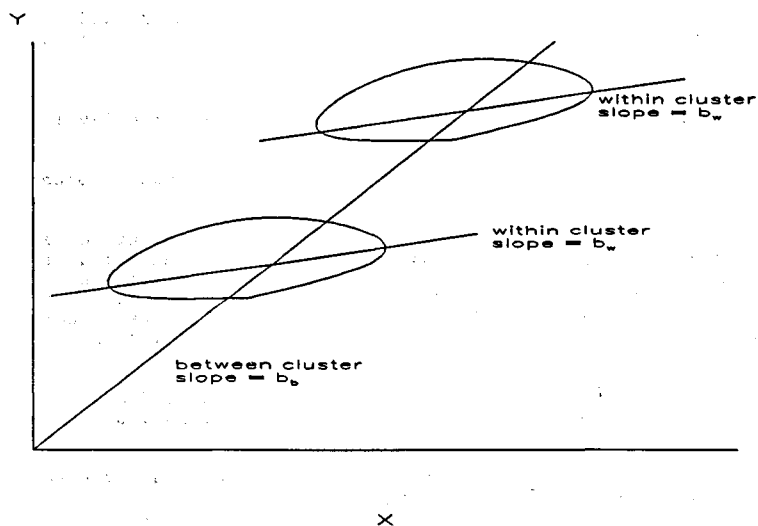


FIGURE 2



ANNEX 12

TELEFAXED RESPONSE BY WGSATM TO WGEAMS INQUIRIES

MATRIX 1 - Matrix table for monitoring in relation to the protection of human health.

- a) It was noted that the statistics for 'looking at' compliance with tolerance standards and for investigating trends are very different, and that the former had not so far been considered within WGSATM. (Had the availability of human health 'standards' been a consideration in the construction of this matrix?).
- b) For 'fatty' fish such as mackerel and herring, it was considered that an entry under fish muscle tissue for PCBs might be appropriate.
- c) The issue of the origin and intended application of the table was raised, i.e. was the table intentionally concerned only with the JMP list of contaminants?, otherwise it was questioned why, for example, dieldrin had not been included in the list of contaminants. (The 1985 baseline study had suggested dieldrin as a contaminant with levels in some areas approaching tolerance levels, and therefore a possible candidate for inclusion in the table). Similarly, is this why, e.g. lindane is in the table but not the other HCH compounds, etc?
- d) Should the table include PCB or CBS?, i.e., should individual chemical compounds be specified (i.e., if individual CBS, then which?).

MATRIX 2 - Matrix table for monitoring to assess existing levels of pollution.

It was not felt that WGSATM could comment on this table, other than in terms of some very general remarks:

It was noted that a continual revision of this table might be required, depending upon the interpretation of what 'monitoring to assess existing levels of pollution' constituted (a similar point could be made in relation to the other tables as well). It was, however, also noted that all 'primary' entries in the table were in abiotic media (with the one exception of TBT in shellfish!).

MATRIX 3 - Matrix table for monitoring to assess the effectiveness of measures to reduce pollution (trend monitoring).

As a basic statement in relation to a response to WGEAM's request, it was considered that, in the absence of high quality data (in all senses of the phrase), it was very difficult for WGSATM to prepare a comprehensive answer for WGEAMS.

No 'seawater' data had yet been considered by WGSATM, sediment data had only been considered in relation to a review of the results of an intercalibration exercise. Thus, for all practical purposes no comments on the non-biota components in the table could be made.

It was, however, considered that the group had a VERY LIMITED amount of data which would possibly allow it to identify which compartments in biota might be most useful or appropriate for certain contaminants. At the same time it should be appreciated that, despite completion of a series of statistical analyses on the ICES CMP trend sets for contaminants in biota, investigations of the complex biological influences on levels of contaminants in biota are only just beginning.

Given this situation, reservations were expressed as to the ability or utility of interpreting biota trend monitoring data in relation to changes in environmental levels in response to 'measures taken to reduce marine pollution', in all but those cases where very marked changes in levels were involved (e.g., near strong point sources such as the Belledune lobster example).

To consider the possible identification of appropriate biological compartments, an exercise was completed involving a review of the results obtained from the analyses of the CMP data for mercury in fish muscle (cod, flounder, sole, plaice and whiting) and liver tissue (cod):

NOTE - mercury was selected as a 'best case', in the sense that it was the only contaminant for which analyses over a relatively wide range of species/area data sets provided consistent results, suitable for comparison.

ALSO - the results represent an average across the species mentioned above.

The following table identified whether it is possible (probability > 0.95) to detect a certain percentage change in mercury concentration over a ten year period by monitoring using fish muscle and liver tissue.

Change	Muscle	Liver
10%	n	n
20%	n	n
30%	y	n
40%	y	n
50%	y	y
60%	y	y

It is IMPORTANT to note that the quantity/quality of the data, and the assumptions used to generate this comparison mean that it is not really safe to conclude that for Hg, monitoring in fish muscle is superior to monitoring in fish liver. Rather, the purpose of the table is to show (roughly), whether, given the quality of the data available, it would be possible to detect some arbitrarily defined levels of change.

Furthermore, it is IMPORTANT to note that the levels of change referred to are levels of change WITHIN the organism, which cannot necessarily be assumed to reflect levels of ENVIRONMENTAL change.

The WGSATM decided that it would be appropriate to send WGEAMS the following extract from the recommendations in their 1985 report:

"..... more information is needed on biological processes influencing the uptake, metabolism, etc. of contaminants and the transfer of contaminants through the food chain. The suggestion was made that each country should establish a long-term, high intensity monitoring station to examine the cyclical and seasonal changes in the animals and their uptake/retention of contaminants as a background to the analysis of time trends of contaminants on an annual basis."

Whether WGEAMS is now considering supporting this aspect, e.g., initiating 'research activities' aimed at resolving some of these questions, (possibly even setting up international reference stations, establishing sample banks, etc.) is considered to be very relevant not only to the activities of WGSATM, but also to all the programmes related to pollution monitoring.

(Last note - nutrients not included in any of the matrix tables?).