TECHNIQUES IN MARINE ENVIRONMENTAL SCIENCES

No. 14

Temporal trend monitoring: Introduction to the study of contaminant levels in marine biota

J.F. UTHE, C.L. CHOU, R.K. MISRA, P.A. YEATS, D.H. LORING

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

C.J. MUSIAL

C. Musial Consulting Chemist Ltd
Halifax, Nova Scotia
Canada B3L 4J2

W. COFINO

Free University
Institute for Environmental Studies
De Boelelaan 1115
1081 HV Amsterdam
The Netherlands

INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CONSEIL INTERNATIONAL POUR L'EXPLORATION DE LA MER

Palægade 2-4, DK-1261 Copenhagen K, Denmark

November 1991

ISSN 0903-2606

TABLE OF CONTENTS

Temporal trend monitoring: Introduction to the study of contaminant levels in marine biota

1 INTRODUCTION

The study of contaminants in marine species, sediments, and sea water has been of interest to the International Council for the Exploration of the Sea (ICES) since the early 1970s. The investigation of temporal trends (changes over time in one area) in contaminant levels in fish and shellfish, both as monitors of their environment and from a human health concern, and in sediments and sea water is a topic currently being addressed by the ICES Working Group on the Statistical Aspects of Trend Monitoring (WGSATM).

In this connection, the following two points should be noted:

a) The design of a monitoring programme is a scientific effort, in which expertise in the fields of biology, environmental and analytical chemistry, statistics, and hydrography must be combined, utilizing specific information on the area to be monitored. It is not possible to give a simple blueprint for the design of these programmes. Rather, each study must be specifically designed within the general strategies and guidelines such as those provided by the ICES Advisory Committee on Marine Pollution (ACMP) (ICES, 1989a,b).

b) The objectives of the monitoring programme have to be defined qualitatively and quantitatively. Statistical model calculations based upon real estimates of the overall variance are a prerequisite to programme design, to give insight into the resolving power of the approach. In most cases, conduct of a "pilot project" will be necessary during the design of a programme.

1.1 Background

Contaminant concentrations in marine biota, sediments, and sea water are all characterized by large natural variances. The nature of the relationships between contaminant levels within and between these compartments is complicated, contaminant-specific, and often not yet elucidated. The utility of such measurements in determining temporal trends in contaminant levels in the marine environment, other than in cases where the change is relatively large, remains to be established.

Historically, studies of temporal trends in contaminant concentrations have been shown to be valuable in those cases where the change in contaminant level brought about by, e.g., emission controls, has been much larger than the natural variation. In situations where the change in contaminant levels is less dramatic, the relatively greater natural variance may make it impossible to detect a change in mean contaminant levels with reasonable statistical confidence using realistic sample sizes. In addition to natural variations, contaminant measurement is a source of further variance and the relatively large frequency of erroneous data entries emphasizes the need for careful quality control.

A number of statistical problems were encountered during the first attempts to analyse the data collected under the Cooperative ICES Monitoring Studies Programme (CMP) for temporal trend monitoring purposes. A main source of these problems was the high degree of inconsistency in the application of the sampling guidelines issued by ICES (1984). Thus, a major objective of this paper is to improve the understanding of potential sources of variance and promote a better appreciation of the need for consistent application of guidelines for the study of temporal trends.

This paper is intended to be the first of a series providing specific details for temporal trend studies using different types of species or marine compartments. This first paper is a general one, describing the types of problems encountered in monitoring temporal trends in contaminant levels in finfish.

Although aimed at the technical audience, the need for stringent quality management necessitates its inclusion in this introductory leaflet. There is a need for the analytical laboratory to maintain reliability in its analyses over the period of the trend studies. With decreasing rates of change in contaminant levels over time, such studies may take many years and it is possible that much of the observed change over the study period may reflect analytical change ("drift") in the laboratory. The effects of abrupt changes, e.g., change in methodology, analysts, instruments, etc., must also be controlled. Changes in intralaboratory analytical variance also affect the efficiency of the statistical analysis.

1.1.1 Biota

The Working Group on the Statistical Aspects of Trend Monitoring (WGSATM) and earlier working groups have attempted the analysis of temporal trend monitoring data using a number of different data sets, some collected under the auspices of the Cooperative ICES Monitoring Studies Programme (CMP), and others collected under other international and national programmes. All comparisons of contaminant levels in biota have recognized the need for a stable sampling structure. This is reflected from an early ICES recommendation to focus on six-year-old cod (*Gadus morhua*) through to the most recent (ICES, 1984) recommendation to sample fish within a range of length strata. The WGSATM has recently completed the analysis of data submitted under the CMP since 1978 on contaminants in fish muscle tissue (ICES, 1989c) and fish liver tissue and mussels (*Mytilus edulis*) (ICES, 1991a).

In temporal trend studies, the need to sample the same population/stock during the same period each year (ideally, within the same period of the species' annual cycle) is evident, as is the need for uniform and consistent specimen handling and analysis over the term of the study. It is, therefore, unfortunate that such basic needs are frequently not fulfilled in many practical monitoring programmes. The need for the procedures to be free from significant bias is obvious.

To attempt to account for the influence of natural variations in contaminant concentrations, their relationship with a number of biological covariables has been investigated. It is difficult to provide a definitive list of covariables which are important, as these vary for particular contaminant/species combinations. Thus, a pragmatic solution regarding the selection of covariables is, for the present, to measure all those that are relatively simple to measure or obtained easily as part of other analyses, e.g., the measurement of "fat" during the analysis of organic contaminants. In this way, the information is available for later use in trend analysis, if required, and to assist in detecting errors. For example, one or more of the covariables may be used to estimate another covariable, which can assist in the identification of possible transcription errors or in identifying abnormalities within a specimen. In a cod study (Misra

and Uthe, 1987), fork length, total weight, age (otolith), sex, liver weight, and percent extractable fat in liver and muscle were recorded for each specimen. Each species should have its own set of covariables chosen after discussion among biologists, chemists and the statisticians. The importance of clearly defining the objectives of the study must be emphasized.

1.1.2 Sediments

Sediments are an appropriate matrix for monitoring studies because they constitute a contaminant reservoir and contain, either directly or indirectly, food for many organisms. Fine-grained sediments are potentially valuable for trend monitoring studies because, under appropriate conditions, they integrate contaminant inputs over relatively long time periods. Thus, they offer the most suitable matrix in which to examine spatial gradients and temporal trends in contaminant concentrations, particularly in areas where biota or water do not provide appropriate matrices.

In general, the studies involve the determination of contaminants in radiochemically dated (e.g., ^{210}Pb and/or ^{137}Cs) sediment cores. From these data, the fluxes of contaminants and their historical record of deposition can be estimated (e.g., Smith and Ellis, 1982; Smith and Walton, 1980). Problems associated with bioturbation and other factors must also be examined for significant influence. In all sedimentary studies, the natural trace metal variability must be taken into account, i.e., there is a need to remove the effect of grain size through granulometric and/or geochemical normalization, e.g., normalization with regard to lithium or aluminium concentrations (ICES, 1989b).

The value of sediment studies is dependent on the sensitivity of the sediment site for reflecting temporal trends in contaminant levels. For temporal trend studies, sediment accumulation areas with low bioturbation should be used. It is important to estimate the magnitude of change in the flux of a contaminant to the sediment surface that can be identified by a statistically significant change in its concentration in the surface sample. A recent review of this problem by Larsen *et al.* (1989) indicates that it is possible to make estimations of the sensitivity of a given sediment site and the sampling frequency required for trend monitoring using parameters obtained by modelling ^{210}Pb profiles and chemical distributions. Such parameters include estimations of sedimentation rates, depth and intensity of bioturbation, and changes in contaminant fluxes to the sediment surface.

From an analytical viewpoint, recent intercalibration exercises on the analysis of trace metals in sediments indicate that improvements in interlaboratory comparability are required before internationally coordinated programmes concerned with environmentally relevant metals, such as mercury and cadmium, can be undertaken with confidence. Improved comparability can be achieved by the careful use of appropriate sampling and analytical procedures and the use of appropriate reference materials for intra- and interlaboratory quality control. Despite these problems, there is now enough precise analytical data from individual studies to test the use of sediments for temporal trend and spatial distribution monitoring, at least within a certain number of highly qualified laboratories.

1.1.3 Sea water

There has been some consideration given to the use of sea water for temporal trend monitoring of nutrients and trace metals during the last few years. A number of analyses of temporal trends in nutrient concentrations in the North Sea have been conducted, with varying conclusions. There has also been some discussion of potential approaches to temporal trend analysis for trace metals, the general conclusion being that sea water is not a suitable matrix for such studies. Sea water can be used, however, for studies of the geographical distribution of

trace metals and nutrients; an example is the ICES Baseline Study of Trace Metals in Coastal and Shelf Sea Waters (ICES, 1991b).

Monitoring contaminants in sea water differs from monitoring contaminants in sediments (or biota) in that, unlike sediments, sea water does not integrate a signal over time, nor does it retain a record of historical concentrations. As a result, the signal is subject to much greater short-term variability in both time and space. Unless this variability can be understood and quantified, it will manifest itself as noise in the long-term trend signal. The causes of some of this variability are understood and can be accounted for. For example, relationships with salinity can be used to describe some of the observed spatial variability for some dissolved constituents. Diurnal and seasonal variability may be related to climatic influences and biological activity, as well as other sources of natural variability. It will be important to understand and quantify as much as possible of this variability, because the long-term trend signal is likely to be rather small compared to the short-term variability. This reflects the need to determine existing variances before designing a monitoring programme.

The analytical chemistry capability for measuring nutrients and trace metals in sea water is relatively well-established. The collection and analysis of nutrient samples is a long-established procedure in oceanographic studies and a relatively high level of analytical precision and accuracy may be achieved, if appropriate quality assurance procedures are consistently applied. The analytical capability for trace metals has not existed for as long; there are now, however, established procedures for the collection and analysis of a number of trace metals, sea water reference materials, and well-documented quality control procedures.

Sea water has a potential advantage in that it has the capability of more rapidly reflecting changes in contaminant levels than either biota or sediments.

2 GENERAL CONCEPT OF TEMPORAL TREND ANALYSIS

The analysis of temporal trends in contaminant levels involves a number of steps:

1) Selection of a sampling locality;
2) Selection of the appropriate matrix for the contaminants being studied, including appropriate species/populations/stocks of biota;
3) Sampling of appropriate sediments, sea waters or biota;
4) Sediment and sea water sample handling; animal and tissue handling;
5) Measurement of biological and chemical parameters;
6) Data handling;
7) Statistical analysis; and
8) Interpretation.

Problems at any of the earlier steps can lead to difficulties at the later steps, resulting in less accurate, if not erroneous, information. Two points need to be emphasized:

1) Once a protocol is selected, it must be rigorously followed, i.e., the sample structure must be as nearly identical as possible each year; and

2) Analytical quality management and control is absolutely necessary. In the case of temporal trend studies, the analytical laboratory must monitor precision and accuracy with an appreciation that any laboratory trend that is not detected and corrected will be ascribed to the trend in contaminant levels in the environment.

2.1 Quality Management

The importance of quality control and assurance in a monitoring programme cannot be over-emphasized. Without proper care and attention to this, data become suspect and, in many past monitoring programmes, impossible to interpret with any degree of confidence. The consequent waste of resources has been enormous. Therefore, within any monitoring programme, the first step must be to address this question.

Concepts of quality improvement and management have been developed during the last century, originating in industrial environments. Beginning with inspection procedures, the next stage was statistical quality control and, thereafter, quality assurance (QA), incorporating total quality control and the quantification of the costs of QA. Most recently, strategic quality management has been introduced. QA involves (1) the formulation of strategies, specific goals and objectives, (2) the development and implementation of action plans, (3) the use of control systems for monitoring performance, (4) feedback, and (5) taking corrective action.

A clear overview of quality management is given by Garvin (1988). Recently, the International Organization for Standardization (ISO) has published a series of standards, ISO 9000-9004, in which guidelines are given on how to select and use quality systems (ISO, 1987). Computer-based programs are also available, e.g., see Keith *et al.* (1988).

The objective is "Total Quality Control"; the whole process from the design of the monitoring programme to the evaluation and presentation of the data should be controlled (Cofino, 1989).

Organizations supervising monitoring programmes can utilize the concepts described above. In line with the ISO standards 9000-9004, it is recommended that the following steps be adhered to:

1) Formulation of a **quality policy**. The quality policy concerns the overall intention and direction regarding quality, as formally expressed by the coordinating organization (ISO, 1986).

2) Establishment of **quality management**. Quality management is defined as the aspect of the overall management function that determines and implements the quality policy. It includes strategic planning, allocation of resources, and other systematic activities for quality, such as quality planning, operations, and evaluations (ISO, 1986).

3) On the basis of the above, development and implementation of a **quality system**. A quality system is defined as the organizational structure, responsibilities, procedures, processes, and resources for implementing quality management.

A QA policy elaborated for multi-laboratory monitoring programmes should specify quality objectives and minimum requirements regarding quality assurance which have to be followed by participating laboratories. The coordinating organization has to take measures for quality assurance of the data submitted. To this end, it is essential to appoint an expert laboratory or body as the "quality manager" specifically for quality control of the data. The appointment of such a laboratory or body should be delineated in the quality policy, along with its specific tasks.

The expert laboratory or body mentioned above should make action plans to achieve interlaboratory comparability and to control and assess the quality of the data. This QA programme should include directives for all technical and administrative procedures, in as much detail as feasible. Certified reference materials should be identified, used, and measured

successfully by all participating laboratories, so that results can be compared on a common basis. A further programme should be devised for external quality assurance of the laboratories, employing, for example, the conduct of intercalibration exercises and the distribution of uncompromised reference materials for analysis along with the actual samples. The expert laboratory or body should prepare an annual "quality report", providing an evaluation of the quality assurance programme and its results. The quality policy and system should be documented in a handbook.

The programme described above combines recent views on quality management with the various approaches aimed at achieving comparability among laboratories (e.g., reports of the ICES Marine Chemistry Working Group; Uriano and Gravett, 1977; Taylor, 1985). It is stressed that each laboratory should maintain a quality assurance programme, carefully tailored to its specific requirements (e.g., Taylor, 1985; Vijverberg and Cofino, 1987).

3 GUIDELINES FOR STUDYING TEMPORAL TRENDS

3.1 General

It is impossible, given the rigorous requirements for temporal trend monitoring, to give a detailed protocol that will suffice for all compartments in all instances. However, there are a few general recommendations which are broadly applicable to sampling biota, sediments, or sea water. They are:

1) Estimate the potential magnitude of the change to be observed. If information is available, estimate the variance associated with the observations. Model the system and determine the number of specimens, or replicate samples of sea water or sediments, required to observe a predetermined level of change at a certain level of probability. This basic information should be used to decide whether trend monitoring of the contaminant(s) using that compartment is a project having a reasonable likelihood of success in relation to the expenditure of resources. Keep in mind that the real situation is not ideal, e.g., it is often impossible to fill the largest and smallest strata when sampling biota within a length-stratified sampling strategy. Estimate the duration of the study and appropriate sampling intervals. There may be a need for potentially long-term monitoring programmes, possibly with high-density sampling, when the detection of changes at levels comparable to the residual variances is desired.

2) Develop, from either historical information or a pilot study, a reasonable and statistically sound sampling protocol.

3) Develop appropriate techniques for quality management, sample handling, and analysis to ensure intralaboratory (important in both temporal trend and geographical distribution monitoring) and interlaboratory (important in geographical distribution monitoring because many laboratories are involved) consistency and comparability. Appropriate controls on data handling must also be included. Some degree of specimen banking may be needed to estimate laboratory drift brought about by sample handling and storage. These factors cannot be checked using reference materials, as they have been prepared to be as stable as possible over time.

4) Develop in all participants an appreciation of the need for adhering to protocols as closely as possible. Once a protocol is selected, it should be rigorously followed, i.e., the sample structure must be as nearly identical as possible each year. This is of particular importance with respect to body length in finfish, the basic criterion on which individual

fish are selected. Obviously, one must ensure that the same population is sampled each year and that, within the sample drawn, specimens are randomly selected within the constraints of the target population and sampling scheme.

5) Identify, if possible, major alterations in contaminant inputs into the study area as well as other factors affecting dynamics within that area.

6) Avoid taking non-representative samples, for example, discard obviously abnormal specimens, such as fish in which post-mortem changes are obvious. However, one should record the incidence and cause of any such discards.

7) The identification of statistically significant trends in contaminant levels is not sufficient to prove that these represent trends in the population at large, the stock within a region, or the environment in the study area. Corroborating evidence is also needed, e.g., results of similar studies in other matrices, information on contaminant inputs, emission controls, and stock size changes.

3.2 The Study of Temporal Trends in Contaminant Concentrations in Biota

3.2.1 Selection of the species/stock for study

Much thought has been given to the selection of appropriate species for use as monitors of temporal trends in chemical contamination. The desired attributes have been listed by Phillips (1980). In summary, the species/stock should be:

1) An accumulator of contaminants reflecting environmental levels in a rational manner;

2) Sedentary;

3) Abundant;

4) Available in multi-year class populations; and

5) Composed of specimens large enough to yield sufficient quantities of tissue for analysis.

Each species selected for study must be carefully considered and its advantages and problems identified. At the very least, the assumptions that have been made must be described along with the gaps in information on the species. The information base on a particular species may be so deficient that one cannot recommend its use in temporal trend studies, but it can still be suitable for studies associated with human foodstuff wholesomeness.

Because a species is composed of more or less distinct stocks, it is essential to select individual stocks for study, generally ones which are the most reproductively isolated from other stocks. The stocks selected should follow a relatively unchanging pattern of living from year-to-year and feed on a relatively consistent food supply, so that the major effect on contaminant levels in the animal is the level of contaminants in the environment rather than changes in the biological/physiological parameters of the animals. Obviously these requirements are never perfectly met in the real world, but must considered in selecting stocks and areas for study. The choice of tissues to be sampled depends upon the aim of the study, the known physiology of the animal with respect to contaminants, and the availability of sufficient quantities of tissue for analysis.

3.2.2 Sampling Biota for Temporal Trend Studies

It is important that all samples taken for the measurement of temporal trends in contaminant levels have essentially the same structure. For example, in a sampling structure based on length stratification, an appropriate stratification needs to be defined on the basis of a population study. In applying this sampling structure, it is important that all strata defined be filled with the required number of animals, selected randomly within each stratum, and that this scheme be adhered to each year. If significant differences in other biological covariables, such as liver fat content, are present, the roles of these covariables in defining the magnitude of the dependent variable (contaminant) may need to be taken into account. If not, their effect (in a minimum sense) is to add to the unexplained variation in the mean contaminant levels, i.e., differences between means must be larger to be judged statistically significant. Furthermore, true differences between means will be confounded with those associated with the covariables. An effect of this is an interpretation that observed changes resulted from changes in environmental contaminant concentrations.

If one or more of the covariables show significant relationships with the dependent variable, it is necessary to take these effects into account when assessing temporal trends. This requirement is the underlying reason for the rigorous adherence to sampling protocols.

Covariables within a population (and, thus, a sample) differ among individuals and during different time periods within their lifetimes. With few exceptions (e.g., age in adult smelt), individuals, each with their own complex phenotypes, live in a group hierarchy of some sort, demonstrate an age and size (length, weight and organ ratios) structure, are male or female of varying, yet concerted, levels of reproductive status, and demonstrate differences in nutritional adequacy, parasitism and disease incidence. In addition, animals are under the control of environmental factors such as temperature and are thought to be affected by chemical and other types of contamination in their environment. It is obvious that many of these covariables, e.g., age and size, are strongly correlated. In other instances, the relationship between covariables is less obvious, e.g., size and sex, while in other cases the relationship is only suspected, e.g., nutritional status or parasitism and size.

In order to account for the effect of a covariable on contaminant levels, it is necessary to know the mathematical nature of the relationship. Often this is assumed to be linear or log-linear. In many instances, often for pragmatic reasons, only a single covariable (e.g., length, because it is the most convenient measurement) is used, the assumption being that the effects of other variables on contaminant levels are small enough to be ignored, in particular those variables which are strongly correlated to the selected covariable, e.g., a length-stratified sampling scheme is considered appropriate for those species where fish continue to grow in length throughout their lifetimes. It is obviously inappropriate where this does not occur in a simple mathematical manner. However, the use of a single covariable in comparing groups (populations) is highly unsatisfactory if the omitted covariable(s) have different group means. In other cases, e.g., nutritional status, disease and parasitism, little is known about trends in these covariables, although it is suspected that such trends exist due to pollution and the increasing eutrophication of marine coastal areas. It is, in our opinion, advisable to reject immediately all specimens which do not appear to be healthy, as well as rejecting later all specimens with abnormal ratios between biological variables, such as length, weight, etc. The frequency of such abnormalities should be recorded and one must keep in mind that the question of contaminant levels and their trends in "sick" and "healthy" members of a population is important, but separate from the issue of temporal trends of contaminant levels. The best that can be done is to limit the effects of as many as possible of these variables on contaminant levels and correct for the effects of others by stratified (Scott *et al.*, 1983) or fixed level of covariable sampling strategies, e.g., the International Mussel Watch (Phillips, 1980).

3.2.3 Fishing the population/stock

Sampling a marine population involves selection (the use of fishing gear or manual collection), often followed by further selection, e.g., selecting a sample from a multi-tonne haul on a commercial or research vessel. It is assumed that both of these selections are unbiased, i.e., each member of the population being sampled has an equal chance of being selected, e.g., the filling of a stratum on a first-come, first-chosen basis. Due to obvious interactions between individuals and the fishing gear, it is highly probable that the gear shows a degree of selectivity. Even the dumping of a haul on board the vessel may result in a pile of fish that is somewhat sorted according to size as the fish tumble over each other. The selectivity of fishing gear for sick or healthy fish can be debated. It can be argued that sick fish with impaired swimming capability are more readily captured by the gear, but this assumes that sick animals stay within the group. There is no evidence to support this or the possibility that sick fish may actively seek places conducive to avoiding predation (hiding places) or reducing nutritional demands (areas approaching isotonic conditions).

3.2.4 Selection of individual fish

It is important that consistent samples be obtained from year to year (see Section 3.2.2, above). Sample structures can inadvertently be altered by the use of untrained individuals to select specimens. Although much monitoring has been carried out using samples obtained off commercial fishing vessels, often taken at the time of landing, we would caution against such use for trend studies. There are always questions regarding the catch location, the possibility of mixing a number of hauls from different areas on board the vessel, and the problem of the length of time between capture and sampling (e.g., how well does a post-rigor mortis fish reflect the population?). We believe that the best approach is to join an annual fish survey research cruise during which a trawl or two at a selected area is dedicated to the contaminant temporal trend study. This generally means placing a chemist or technician on the vessel for the length of the cruise, but, given the very high costs associated with the study, it is money well spent. This individual is responsible for ensuring the constancy of the sampling procedures and the proper handling of the animals. Selection is based upon length, each of the required strata being filled. The animals should not be weighed or other biological data collected on board the vessel. Accurate weighing on a moving vessel requires very sophisticated balances and there is a very limited amount of time between hauls. Animals should not be cut open or dissected on board due to the possibility of contamination.

3.2.5 Storage on board the fishing vessel

Fish are selected and frozen whole, i.e., the entire, intact fish, wrapped in plastic, is placed in a single layer directly on the shelves in the freezer to ensure rapid freezing. The proper handling of specimens after capture and selection is very important. While most researchers are aware that specimens can become contaminated by materials from the vessel itself and that the animals have to be stored properly to prevent decomposition, it appears that many investigators are not aware of the speed at which the decomposition of certain tissues occurs. This can have drastic effects on contaminant levels in certain tissues and on inter-tissue contamination ratios (Uthe and Chou, 1987). In the case of cadmium in sea scallops (*Placopecten magellanicus*), there is a high ratio of cadmium in the digestive gland *versus* in the adductor muscle. The digestive gland autolyzes (decomposes) rapidly (in a matter of minutes) after death, releasing cadmium from its protein matrix within the gland, thus allowing it to bind tightly to the protein matrix of the adductor muscle. In the case of studies on cod (Misra and Uthe, 1987), specimens in which the gall bladder had ruptured through autolysis were not processed further.

3.2.6 Measurement of biological parameters

Measurement of the covariables for each specimen, e.g., length, weight, sex, % fat, may pose certain problems. Little information could be found in the literature on operator-to-operator measurement errors. Operator-to-operator error was investigated by Uthe and Chou (1987) in a study of scallop aging techniques and serious problems were identified. This resulted in shell height measurements being preferred to age measurements in relationship to tissue cadmium burdens. Dethlefsen *et al.* (1984) reported significant operator-to-operator error in measurements of fish disease prevalence.

The magnitude of any source of error must also be kept in mind along with its frequency. An error of measurement of 1 mm in shell height determinations in scallops of approximately 100 mm shell height is far less serious than an error of 1 year in 3-6 year-old scallops. When standard least-square procedures are applied in the presence of observational errors in the covariable, (a) estimates of slope and intercept of the simple linear regression equation will no longer be unbiased and consistent, and (b) the accuracy of the prediction of an individual Y (contaminant level) from a data set (sample) will be reduced (Neter *et al.*, 1983).

3.2.7 Tissue sampling and storage in the laboratory

Obtaining tissue samples for chemical analysis from each individual specimen is not without its problems. Obviously, dissection must not affect contaminant levels in the tissue, so some care in the selection of dissection tools is needed along with an investigation of the effect of the tools. We have found steel, not stainless steel, knives quite adequate for use in the study of many contaminants in fish.

Generally, the animal has been frozen and held in this state for some time. In addition to the post-mortem changes noted above, other problems can arise. Obviously, care must be taken to ensure that significant amounts of water are not lost during frozen storage through freeze drying. Frozen tissues cannot be thawed without the tissue expressing fluid which may or may not contain appreciable amounts of the determinand. Substantial amounts (up to 20% of the tissue mass) of fluid may be lost (Uthe and Chou, 1988) and further tissue decomposition may occur. It is necessary to keep the animal frozen to prevent fluid loss and decomposition, yet allow it to soften sufficiently for dissection. Dissection may not be simple due to the fact that many statistical analyses utilize the burden (contaminant concentration times whole tissue weight) of a contaminant in the tissue rather than the concentration. This means that the entire tissue must be removed and weighed. This is relatively simple in the case of an organ like the liver but difficult in the case of the musculature. However, in certain instances, it is feasible to use another estimator of the tissue mass, as demonstrated by Scott *et al.* (1983) for Atlantic cod (*Gadus morhua*) where the musculature mass was estimated by use of the "net" fish weight (total fish weight minus whole liver weight). Contaminant concentrations in muscle were determined in homogenates of both skinned fillets, i.e., containing both red and white muscle tissue. It should be noted that, in our experience, filleting is relatively operator-dependent. Ideally, the same individual should be used to dissect all fish within a single temporal trend study.

3.2.8 Tissue homogenization

It is important to remember that certain tissues, e.g., whole fish musculature, kidney, do not comprise a single type of cell and that a gradient in contaminant concentrations may be present along one axis of the tissue due to a gradient in a cellular constituent, such as fat. It is often assumed that homogenization of a tissue or mixture of tissues is relatively easy to carry out, but any elementary text on sampling for chemical analysis, e.g., Harris (1978), shows the error of

this assumption, especially as applied to the analysis for microconstituents. It is important to check the homogenization technique of the operator by sieving.

Following homogenization, a sub-sample is generally taken and frozen. Ideally, the analytical samples (actual mass to undergo chemical analysis) should also be taken immediately after homogenization. If this is impossible, it is necessary to rehomogenize the frozen homogenate before sampling for analysis. Samples must be stored in cleaned jars with lids. The jars should be of a size to allow convenient rehomogenization prior to sampling for chemical analysis.

3.2.9 Measurement of chemical parameters

Although this document will not address details of the chemical measurement of contaminants and other constituents, the importance of appropriate analytical quality control and the elimination of analytical trends over the years cannot be over-emphasized. It is necessary for the laboratory to ensure that such trends do not occur. Similarly, the laboratory must ensure that its overall analytical precision (this includes all steps carried out within the laboratory) does not change significantly. The use of reference materials has a major role in such quality control, with the caveat that the laboratory manager must ensure that the analysis of reference materials does not differ from routine sample analysis. Sufficient numbers of analyses of reference materials must be carried out to detect changes brought about by batch or chemist changes, as well as to detect analytical changes (analytical drift or trend) within the laboratory (Vijverberg and Cofino, 1987; Taylor, 1985).

Many laboratories report "less than" values for a variety of contaminants in certain matrices. Such results cause severe difficulties from a statistical analytical point of view. Although suggestions have been made on how to handle a low (1-10%) frequency of such reports, they are not particularly satisfactory from a trend point of view, because inevitably the frequency of "less than" values must change significantly over time when real trends are present. Analysts should modify their methodology to avoid reporting "less than" values. If this is impossible, there is little point in attempting temporal trend studies with that particular species/tissue/ analysis combination and a more appropriate combination should be selected.

4 DATA HANDLING

Manual data handling errors, both transcription and translation entry errors, have been shown to be a major problem in temporal trend studies, together with errors of misinterpretation (entering information other than that required, e.g., weight of matter taken for chemical analysis rather than organ or tissue weight). In addition to order of magnitude errors, e.g., entering 1.987 as 19.87 and transposition errors, e.g., entering 1.987 as 1.897, there is a frequency of other errors not readily explained. Unfortunately, experience has shown that returning print-outs of computer entries to the data originators for verification may still not result in the detection and correction of such errors unless extreme care and dedication are applied. The effects of such errors range from essentially negligible to those of sufficient magnitude to affect the statistical analysis and interpretation of the data. Large outliers (in a statistical sense) can be identified and eliminated, but there is an intermediate level which can still affect the analysis.

4.1 Types of Entry Errors

Entry errors of the types described above are referred to as operator errors by Juran and Cook (1974) and comprise three types:

1) **Inadvertent**: errors characterized by the fact that the operator is not aware that the error

has been committed. Such errors are generally randomly distributed over operators and error types;

2) **Technical**: errors resulting from a lack of "skill" by identifiable operators within one operation, i.e., a single error type. An error-prone operator with errors in all operations is another type of technique error specific to an individual; and

3) **Willful**: errors in which the operator knows that the error has been made. The error is a result of conflict arising from management instructions and operator attitudes. Reduction of willful errors lies within the individual laboratories, if such errors are a problem, and will not be discussed further here.

4.2 Avoiding Data Handling Errors

4.2.1 Data forms

A reduction of data handling errors is possible through education of the operators, and the use of well-designed data forms containing well-prepared, unambiguous instructions (Rigby and Swain, 1975), including an example of a correctly completed form. These authors note the following from their study of data forms:

1) Most instructions do not follow design principles established by research;

2) Capital letters alone are very difficult to read;

3) Spaces between columns should be less than one eye fixation, i.e., one-half inch of white space;

4) Groups of five elements (rows) should be used;

5) Elements should be ordered in a logical way; and

6) Forms should be prepared for the convenience of the user, not the originator.

4.2.2 Other remedies

A variety of other remedies are available for the prevention of inadvertent errors (Juran and Cook, 1974) including:

1) **Foolproofing**: the use of certain methods for the detection and elimination of errors;

2) **Automation**: the use of electronic data processing;

3) **Conversion to comparison**: the operator should be able to compare his results directly with a standard rather than with his memory of a standard; and

4) **Traceability**: the ability to identify the data originator and sample history.

Foolproofing

Foolproofing methods, in this case, involve the use of redundancy in coded language to minimize errors, i.e., each code word must differ from all others by at least two letters, the probability of such a double error occurring being substantially less than that for a single error.

Single letter or number codes are the most error prone. Active proofreading, e.g., reading aloud, using an independent proofreading crew, should be employed as much as possible.

Automation

Automation, i.e., electronic data processing, can eliminate translation and most transcription errors providing the programs are capable of carrying out the manipulations in the necessary manner. Machine errors can be significant and in many cases not noticed without investigation, e.g., the rounding errors imposed by the machine's capabilities when large independent variables are entered in statistical calculations (Neter *et al.*, 1983). Required information can be lost during data manipulation, such as the loss of internal standard correction factors in certain analyses where the loss occurred only in real samples and not the standards (pers. comm., D.E. Wells, Aberdeen). However, computers are invaluable for reducing error frequency and aiding error detection through studies of the ratios of certain data entries and the identification of suspect numbers. Instrument automation, however, can cause problems if the operator does not have a thorough knowledge of how the system works, e.g., draws baselines and integrates chromatograms.

Comparison

Comparison is a powerful method of preventing errors, particularly blunders (defined as gross errors by Mikhail and Ackerman, 1976). The detection of such errors is facilitated by the operator having at hand standards with which to compare his results rather than using his memory or having no information. Thus, in addition to copies of completed data forms, the analyst should be supplied with information on expected ranges of results for all measurements.

Traceability

Traceability is also effective in reducing errors. In addition to being able to show the history of each sample, each measurement should be identifiable as to the individual who made the determination. Data forms should be signed by both the individual who entered the data and the individual who checked the entries.

5 STATISTICAL ANALYSIS

Although this document does not provide detailed statistical procedures, some notes are appropriate. It is important to recognize the need for high quality data sets for the intensive statistical analyses required with temporal trend data sets. Many statistical analyses require normalization of the data, elimination of outlying observations, and very few, if any, missing observations. Chemists must be made to realize the importance of this and not expect the statistical analysis to correct for deficiencies in the data.

Statistical analysis for temporal trends has been discussed at length by the WGSATM (Anon., 1987, 1986, 1985). A step-wise approach using the overall fish length as a primary covariate for certain contaminants in fish muscle tissue has been recommended by this group in the first instance (Anon., 1986). At its 1987 meeting, WGSATM considered the progress which it had made to date in analyzing temporal trends in contaminant levels in biota within the data sets held by ICES as part of the Cooperative ICES Monitoring Studies Programme (CMP) using a two-stage analytical procedure (ICES, 1987; Nicholson, 1985). The first stage investigates the effect of covariables (length, age, sex, etc.) on contaminant levels and establishes whether or not the relationships are the same between years. The second stage is the comparison of contaminant levels adjusted for common values of biological variables to investigate trends.

Modifications to the step-wise procedure were presented later by Misra (Anon., 1988).

The suggestion that multivariate statistical methods should be used rather than univariate methods has also been made within WGSATM for a number of years (Misra, 1985; Misra and Uthe, 1986; Misra *et al.*, 1985; Misra and Uthe, 1987). It has been argued that currently utilized univariate linear models are only particular cases of the more general multivariate linear model and that, in situations such as temporal trend studies of contaminants in fish, where dependent variables (contaminants) are mutually correlated, the multivariate, and not univariate, models offer more appropriate methods of analysis. However, some believe that multivariate results are difficult to interpret and present to managers in simple, meaningful ways (Anon., 1987). This is still under discussion within the group. Multivariate analysis of covariance has been applied by Misra to contaminant concentrations in cod liver and muscle tissues with trends in levels clearly identified (Misra and Uthe, 1987). Also, WGSATM is investigating the usefulness of conducting multivariate analysis on data for all species from one area, rather than studying each species separately.

These considerations clearly show the need for adequate statistical consultation at all stages of the programme with a statistician who will be responsible for the development and application of appropriate statistical analyses of the data.

6 INTERPRETATION

Although statistical analysis can identify trends and judge the probability of the trends being real, it must be emphasized that good science cannot rest on statistics alone. Trends should be accepted as real only when there is corroborating evidence supporting them, e.g., a trend of decreasing contaminant levels in fish should be corroborated by decreased inputs into the area. Given the complex biochemical/physiological nature of a living organism and the complex nature of the processes influencing a contaminant in the organism's environment, it is difficult to determine how much of any observed trend in a particular fish population reflects environmental change. The study of both a variety of species and the abiotic compartments (sediments, suspended particulate material, and sea water) will be needed to throw further light on this relationship.

7 DISCUSSION

Given the above, it is obvious that a concentration of quality control procedures only at one or more steps of the analysis is not sufficient for temporal trend studies. It is necessary that quality control be an integral part of every step of the study, from sampling through final data analysis and interpretation, i.e., quality management. Essentially, this means reducing the magnitude of errors at each step down to the random error level characterizing that step and rigorously adhering to the analytical protocols. The precision and the operator-to-operator error must be determined for each measurement. Some estimate of the accuracy of each determination, chemical and otherwise, is also needed. Manual data handling must be reduced to the minimum amount required using appropriate hard- and software.

It is unfortunate that we do not have materials equivalent to reference materials for checking errors in the non-chemical measurement steps of the procedure. At the very least, it is necessary to specify measurements that have shown themselves to be relatively immune from operator error, e.g., the use of shell height rather than age in estimating an animal's exposure time. If it is necessary to employ measurements which are characterized by a large operator-to-operator error, it will be essential to use the same operator each time or to train operators to

reduce the error to an acceptable level. Obviously, the role of training in trend studies cannot be neglected. The measurements should be regarded with suspicion if the procedure is being used for the first time. It is equally important that operators do not become complacent about their ability to measure and become careless, because the error level is certain to increase.

The need to generate comparable results, the problems of interlaboratory biases, and the presence of operator-to-operator errors within laboratories has led to the adoption of standard methods of analysis by such chemical agencies as the Association of Official Analytical Chemists. Such recommended methods have been thoroughly tested, levels of intra- and interlaboratory error determined and controlled, and expected levels of precision set, i.e., analysts have standards with which to compare their results. Methods are described in exact detail and analysts diverging from recommended procedures are expected to demonstrate comparability and, hopefully, improvements by their modifications. The use of standard chemical determination methods in fish contamination studies has already been recommended by Holden *et al.* (1983) based on the results of intercomparison exercises on the determination of trace metals and organochlorine residues in tissues of fish recommended by ICES for monitoring purposes.

A similar approach is needed for temporal trend studies. The first step is to determine which species are suitable for which types of trend monitoring. The second step is to describe, in detail, procedures for the investigation of the suitability of selected marine organisms as monitors of temporal trends of contaminant levels in the marine environment. The utility of sea water and sediment contaminant measurements for trend studies must be similarly considered.

8 ACKNOWLEDGEMENT

The authors wish to recognize the valuable contributions from J.M. Bewers, N.W. Green, M.D. Nicholson, J.F. Pawlak and S.J. Wilson, who reviewed the manuscript.

9 REFERENCES

Anon. 1988. Report of the 1988 Meeting of the Working Group on the Statistical Aspects of Trend Monitoring. Doc. ICES C.M.1988/E:27. 85 pp.

Anon. 1987. Report of the 1987 Meeting of the Working Group on the Statistical Aspects of Trend Monitoring. Doc. ICES C.M.1987/E:24. 75 pp.

Anon. 1986. 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. Doc. ICES C.M.1986/E:39. 45 pp.

Anon. 1985. Report of the Meeting of the *ad Hoc* Group of Statisticians Assisting the WGMPNA on Trend Monitoring Issues. Doc. ICES C.M.1985/E:10. 34 pp.

Cofino, W.P. 1989. Methodology of chemical monitoring in the marine environment. Hegol. Meersunters., **43**: 295-308.

Dethlefsen, V., Watermann, B., and Hoppenheit, M. 1984. Sources of variance in data from fish disease surveys. Arch. Fish. Wiss., **34**: 155-173.

Garvin, D.A. 1988. Managing quality. The Free Press, MacMillan Inc. New York, NY.

Harris, W.E. 1978. Sampling, manipulative, observational and evaluative errors. American Laboratory, **10**: 31-39.

Holden, A.V., Topping, G., and Uthe, J.F. 1983. Use and relevance of analytical intercomparison studies in monitoring the marine environment. Can. J. Fish. Aquat. Sci. **40** (Suppl. 2): 100-110.

ICES. 1991a. Statistical Analysis of the ICES Cooperative Monitoring Programme Data on Contaminants in Fish Liver Tissue and *Mytilus edulis* (1978-1988) for the Determination of Temporal Trends. *ICES Coop. Res. Rep.* No. 176. 189 pp.

ICES. 1991b. A Review of Measurements of Trace Metals in Coastal and Shelf Sea Water Samples Collected by ICES and JMP Laboratories During 1985-1987. *ICES Coop. Res. Rep.* No. 178. 64 pp.

ICES. 1989a. Report of the ICES Advisory Committee on Marine Pollution, 1988. *ICES Coop. Res. Rep.* No. 160. pp. 18-23.

ICES. 1989b. Report of the ICES Advisory Committee on Marine Pollution, 1989. *ICES Coop. Res. Rep.* No. 167. pp. 22-36 and pp. 68-76.

ICES. 1989c. Statistical Analysis of the ICES Cooperative Monitoring Programme Data on Contaminants in Fish Muscle Tissue (1978-1985) for Determination of Temporal Trends. *ICES Coop. Res. Rep.* No. 162. 147 pp.

ICES. 1987. Report of the ICES Advisory Committee on Marine Pollution, 1986. *ICES Coop. Res. Rep.* No. 142. pp. 15-17 and pp. 115-128.

ICES. 1984. Details to be followed for sample collection, preparation, and analysis in the conduct of cooperative monitoring, pp. 96-100. *In* The ICES Coordinated Monitoring Programme for Contaminants in Fish and Shellfish, 1978 and 1979, and Six-Year Review of ICES Coordinated Monitoring Programmes. *ICES Coop. Res. Rep.* No. 126.

ISO. 1986. International Standard 8402. Quality Vocabulary. International Organization for Standardization. Paris.

ISO. 1987. International Organization for Standardization. Paris.
International Standard 9000. Quality management and quality assurance standards. Guidelines for selection and use.
International Standard 9001. Quality systems. Model for quality assurance in design/development, production, installation and servicing.
International Standard 9002. Quality systems. Model for quality assurance in production and installation.
International Standard 9003. Quality systems. Model for quality assurance in final inspection and test.
International Standard 9004. Quality management and quality system elements. Guidelines.

Juran, J.M., and Cook, H.M. jr. 1974. Inspection and test. *In* Quality control handbook. Ed. by J.M. Juran, F.M. Gryna jr., and R.S. Bingham jr. McGraw-Hill Co., New York.

Keith, L.H., Johnston, M.T., and Lewis, D.L. 1988. Defining quality assurance and quality control sampling requirements, pp. 85-87. *In* Principles of environmental sampling. Ed. by L.H. Keith. American Chemical Society, Washington.

Larsen, B., Pheiffer Madsen, P., and Jensen, A. 1989. Sensitivity of sediments in pollution monitoring, pp. 109-115. *In* Report of the ICES Advisory Committee on Marine Pollution, 1988. *ICES Coop. Res. Rep.* No. 160.

Mikhail, E.M., and Ackerman, F. 1976. Observations and least squares. IEP--A Dun-Donnelly Publisher, New York, NY.

Misra, R.K. 1985. The analysis of time trends in contaminant levels in Canadian Atlantic cod (*Gadus morhua*). 1. Choice of a multiple linear regression model. Doc. ICES C.M.1985/E:34. 15 pp.

Misra, R.K., and Uthe, J.F. 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., and Uthe, J.F. 1986. The analysis of time trends in contaminant levels in Canadian Atlantic cod (*Gadus morhua*). 4. Time trends for interactions of chemical contaminants. Doc. ICES C.M.1986/E:30. 10 pp.

Misra, R.K., Uthe, J.F., Chou, C.L., and Gellately, B. 1985. The analysis of time trends in contaminant levels in Canadian Atlantic cod (*Gadus morhua*). Identification of outlying observations in the multiple linear regression (MLR) analysis. Doc. ICES C.M. 1985/E:35. 37 pp.

Neter, J., Wasserman, W., and Kutner, M.H. 1983. Applied linear regression models. Richard D. Irwin, Inc., Homeward, IL. 547 pp.

Nicholson, M.D. 1985. The treatment of time-effects in the statistical analysis of contaminant monitoring data. Doc. ICES C.M.1985/E:31. 10 pp.

Phillips, D.J.H. 1980. Quantitative aquatic biological indicators. Applied Science Publishers Inc., London. 488 pp.

Rigby, L.V., and Swain, A.D. 1975. Some human-factor applications to quality control in a high technology industry, pp. 201-216. *In* Human Reliability in Quality Control. Ed. by C.G. Drury and J.G. Fox. Taylor & Francis Ltd., London. Halsted Press, New York.

Taylor, J.K. 1985. Standard reference materials: Handbook for SRM users. National Bureau of Standards, Gaithersburg, MD., USA. 85 pp.

Scott, D.J., Uthe, J.F., and Chou, C.L. 1983. Choice of variables in estimation of time trends in contaminant levels in fish. Working Doc. ICES Working Group on Marine Pollution Baseline and Monitoring Studies in the North Atlantic Meeting, Copenhagen, February, 1983.

Smith, J.N., and Ellis, K.M. 1982. Transport mechanism for Pb-210, Cs-137 and Pu fallout radionuclides through fluvial-marine systems. Geochim. Cosmochim. Acta, **46**: 941-954.

Smith, J.N., and Walton, A. 1980. Sediment accumulation rates and geochronology

measured in the Saguenay fjord using Pb-210 dating method. Geochim. Cosmochim. Acta, **44**: 225-240.

Uriano, G.A., and Gravett, C.C. 1977. The role of reference materials and reference methods in chemical analysis. CRC Crit. Rev. Anal. Chem. 361.

Uthe, J.F., and Chou, C.L. 1988. Factors affecting the measurement of trace metals in marine biological tissues. Inter. J. Environ. Anal. Chem., **71**: 67-84.

Uthe, J.F., and Chou, C.L. 1987. Cadmium in sea scallop (*Placopecten magellanicus*) from clean and contaminated areas. Can. J. Fish. Aquat. Sci., **44**: 91-98.

Vijverberg, F.A.J.M., and Cofino, W.P. 1987. Control procedures: Good laboratory practice and quality assurance. *ICES Techniques in Marine Environmental Sciences* No. 6. 15 pp.