EP

United Nations Environment Programme

UNEP(OCA)/MED WG.128/2
8 April 1997

ENGLISH

MEDITERRANEAN ACTION PLAN

Informal Consultation on
Trend Monitoring

Athens, 10-11 April 1997

A PROGRAMME OF TREND MONITORING IN COASTAL AND POLLUTION HOT SPOT AREAS

UNEP
Athens, 1997

TABLE OF CONTENTS

BACKGROUND

Following the first phase of the implementation of the MED POL Programme (MED POL - Phase I) from 1975-1980, the Contracting Parties to the Barcelona Convention approved a ten-year long-term Programme (MED POL - Phase II, 1981-1990) consisting of a monitoring and research component. In 1991 the Contracting Parties extended MED-POL Phase II until 1995 and the Programme was subsequently extended until 1996 to enable its completion and the formulation of the next phase.

In 1992 the Bureau of the Contracting Parties requested the Secretariat to organise the preparation of an in-depth evaluation of the MED POL Programme by experts and scientists external to the MAP office, with the intention to use this evaluation in the drafting of Phase III of MED POL. This evaluation was presented to the Eighth Ordinary Meeting of the Contracting Parties in October 1993 (UNEP, 1993a). During this meeting the Contracting Parties formally agreed to the preparation of MED POL Phase III, covering the period 1996-2005, and set a number of basic objectives and principles for its preparation (UNEP 1993b, Annex IV).

The meeting of experts on the preparation of MED POL Phase III held in Izmir, in June 1994, after reviewing and discussing the achievements and shortcomings of Phases I and II of the MED POL Programme, prepared a draft Programme for MED POL Phase III which was submitted for approval to the Joint Meeting of the Scientific and Technical Committee and the Socio-Economic Committee in April 1995. The document was not considered by the Joint Meeting due to lack of time and consequently the delegations were requested to provide comments to the Secretariat in writing. After reviewing the comments received and taking into account the results of the informal consultation meeting on MED POL III (Athens, December 1995), the document was revised to bring it in line with the Action Plan for the Protection of the Marine Environment and the Sustainable Development of the Coastal Areas of the Mediterranean (MAP Phase II) which was approved by the Contracting Parties in June 1995. The revised document was submitted to the Meeting of MED POL National Coordinators (Athens, March 1996), the Meeting of MAP Focal Points (Athens, May 1996) and finally the Extraordinary Meeting of the Contracting Parties (Montpellier, 1-4 July 1996). where it was adopted (UNEP, 1996).

According to the Annex of the MED POL Phase III Programme, two basic types of monitoring will be organised, compliance and trend monitoring. Trend monitoring will be carried out at four levels, coastal zone trend monitoring, trend monitoring in pollution hot spot areas, trend monitoring of loads and trend monitoring of biological effects.

The present document refers to the first two types of trend monitoring which are coastal zone trend monitoring and trend monitoring in pollution hot spot areas using chemical contaminants.

1. INTRODUCTION

The first phase of MED POL included baseline studies which would generate data which were insufficiently available at the time. However since only very few laboratories were able to perform such types of analyses, during the first phase of MED POL (1975-1980) emphasis was placed on strengthening and upgrading the technical capabilities of national laboratories, mostly in developing countries, so that all countries would be able to participate in the Programme, and on the development of the methodologies needed to implement the Programme. Analytical instruments and laboratory materials were provided while an extensive training programme was pursued. In addition a quality assurance programme was launched which included the development of standard analytical techniques, the maintenance of instruments, intercomparison exercises, etc.

In view of the inexperience of many laboratories, and difficulties inherent to the Programme, the data collected during the first phase of MED POL could not be considered of high quality, largely due to the validity and comparability of the data and the uneven and inadequate geographical coverage of the Mediterranean Sea.

During MED POL Phase II, monitoring was organised on a national level. Each national monitoring programme aimed to cover monitoring of levels of pollutants in the marine environment (nearshore and offshore areas) and monitoring of sources of pollution including inputs through the atmosphere. The areas monitored were usually near the coast, especially those affected by pollution. Assistance to laboratories in the developing countries continued. The QA programme was enhanced with on-job training and split sample exercises. Data were provided to the Secretariat on an annual basis.

MED POL Phase II recommended a number of priority parameters to be monitored, (UNEP, 1986), frequencies and species.

This approach was soon abandoned as it was necessary to redesign each national monitoring programme on the basis of local needs and conditions. For example, in certain cases organohalogens were not monitored as they did not pause a problem in certain countries. The frequency of sampling in most cases changed from seasonal to annual.

Not all national monitoring programmes run smoothly. In many cases there are temporal gaps, despite the effort made to collect samples in such a way as to be able to use the data for the identification of trends, and in certain cases geographical gaps. Furthermore, the data stored in the MEDU data bank were not screened properly, until only very recently. In addition, while an effort was made to use all the data on the Mediterranean level, this proved fruitless as the data originated mostly from polluted areas and did not cover large parts of the Mediterranean. According to the evaluation of the MED POL Programme by external experts and scientists (UNEP, 1993a), this was due to insufficient infrastructure of the participating institutes and limited experience of analytical procedures, or due to limited specification of the work to be done in monitoring agreements. The evaluation concluded that the results of MED POL Phase II could not provide a complete and representative description of the state of the marine environment in the Mediterranean and could not allow an estimate of the balance of inputs.

Statistical analysis for trends of MED POL monitoring data for heavy metals and halogenated hydrocarbons in a MAP/MED POL study carried out in 1992 by Dr Robert Fryer, statistician and member of the ICES Working Group on statistical aspects of monitoring, revealed that due to the inconsistent collection, preparation and chemical analysis of data, objective investigation of between-year variation in contaminant levels is virtually impossible and

that due to an insufficient number of pools on each sampling occasion and/or an insufficient number of years sampled, only very large trends in contaminant levels are likely to be detected (Fryer, 1992).

The results of the above study were not unexpected as the programmes were not designed accordingly.

In the present effort, coastal zone trend monitoring, involves the selection of a number of fixed coastal stations from the national monitoring programmes, to be included in a regional monitoring network for the establishment of trends in the Mediterranean, providing information for the assessment of the overall quality status of the Mediterranean Sea. Trend monitoring in areas under the direct influence of pollution sources includes monitoring in pollution hot spot areas (intensively polluted areas) where control measures have or will been taken;

2. PROGRAMME DESIGN

2.1 Introduction

A monitoring programme of trends in the levels of contaminants over a number of years, should provide a simple description of any observed variation with time in the environmental levels of contaminants, for example as an upward or downward trend.

An obstacle to the above is the occurrence of variation between years caused by other sources of change such as biological variables, i.e. seasonal changes in the physiology and behaviour of organisms, and environmental variables i.e., changes in sediment composition, climatic changes etc. (see Carlberg, 1993; Phillips and Rainbow, 1994 chapters 5, 6 and 8). Furthermore, variations in analytical methods due to variations in sample and data handling may also be a source of between year variation. Large sample-to-sample variations in contaminant concentrations and small differences in mean contaminant levels between samples are found to occur due to sources of random between year variation. Therefore, in order therefore to detect environmental trends, particularly small changes, it is necessary to take into account all other factors contributing to variations, in the design of the monitoring programme and where possible take appropriate measures to ensure that these factors are kept constant or controlled.

A trend monitoring programme should include an indication of the level of change in contaminant concentration that the programme is expected to detect and the desired probability with which the change should be correctly detected. The probability of having a correct answer from the programme must be investigated prior to the start of the programme and if found not to be satisfactory, the programme should be redesigned or the objectives of the programme should be reconsidered. Once a programme design is formulated it must be rigorously followed, including quality control and assessment methods, in particular those necessary to ensure analytical precision and accuracy, since any analytical trend that is not detected and corrected will be ascribed to the trend in environmental contaminant levels.

A detailed design of an environmental trend monitoring programme therefore should include the following:

- Description of the objective of the trend monitoring programme
- Determination of the stations to be selected for monitoring
- Determination of the contaminants to be measured
- Selection of the sampling matrices (i.e. biota, suspended matter, sediments, seawater)
- Determination of the species to be utilized
- Selection of tissues for analysis of contaminants in biota
- The timing and frequency of sampling
- The number of samples and size of specimens to be taken for each sample

Methods should subsequently be formulated for quality control and assessment. This document also includes relevant measures for assistance to participants in the regional trend monitoring programme.

2.2 The objective of the Trend Monitoring Programme

As stated above, the general objective of the trend monitoring programme is to provide a simple description of any observed variation with time in the environmental levels of contaminants, for example an upward or downward trend.

More in particular a Trend Monitoring Programme should include an indication of the level of change in contaminant concentration that the programme is expected to detect and the desired probability with which the change should be correctly detected. The objective of a trend monitoring programme may also be formulated on the basis of factors such as the number of samples required to detect an important trend or the minimum allowed level of error to detect an important trend. Power studies allow an estimation of the reliability of the programme in achieving the desired monitoring objective. This process is elaborated in the attached Annex.

2.3 Selection of monitoring stations

A number of fixed coastal stations from the national monitoring programmes will be selected by the MED POL national coordinators in each country, to be used in the trend monitoring programme.

In order to select the location of appropriate stations for the detection of contaminant trends, the knowledge of the ecological dynamics in a specific coastal area as well as of its seasonal and annual patterns is necessary, in particular for evaluating how wide is the area which is (statistically) under, or not, the influence of one specific discharge.

In this context, it could be very useful the support of dynamic information derived from satellite remotely sensed data. As a matter of fact, satellite sensors, could provide spatial and temporal patterns relevant to some sea surface parameters (as temperature, chlorophyll-like pigments, suspended matter) which are directly influenced by river discharges - as well as by plant discharge or coastal runoff in general - sea dynamics, seasons, biology productivity, etc..

The following criteria will determine the sites to be selected for trend monitoring:

- The site will allow the detection of the change in contaminant level that the MED POL trend monitoring programme is expected to detect (as described in section 1.1 on the detailed objectives of the trend monitoring programme and the accompanying Annex), through the selection of a realistic number of samples (see section 1.9 on the number of samples required for trend monitoring).

As already mentioned, the detection of trends in hot spot areas is relatively straightforward, as the change in environmental contaminant levels is expected to be greater than natural variations. At stations where there is no direct input of pollutants and therefore the change in contaminant levels is expected to be lower, the detection of changes in environmental levels of contaminants, with reasonable statistical confidence using realistic sample sizes may be very difficult;

- The site will allow the selection of a sufficient number of biota required for the trend monitoring programme, which fulfil the criteria for the selection of organisms for the purpose of monitoring chemical contaminants;

- The site will be suitable for sediment down-core analysis, particularly as regards sedimentation rates and bioturbation intensity.

2.4 Contaminants to be measured

The selection of contaminants is based upon legislative requirements and analytical capability, as shown by demonstrated accuracy, precision and limit of quantification.

On a Mediterranean scale and on the basis of the past MED POL monitoring data the following contaminants could be selected for measurement in the trend monitoring programme:

a)
- Total mercury in sediment and biota
- Cadmium in sediment and biota

The above may be considered as priority contaminants, for which it could be expected that trend monitoring would be carried out in most, if not all, stations.

b)
- Total arsenic in biota
- Zinc in sediment and biota
- Copper[1] in sediment and biota
- High molecular weight halogenated hydrocarbons in sediment and biota
- Polynuclear aromatic hydrocarbons in biota.

The choice of monitoring the above and other trace elements will be site dependent and will depend upon whether there is an input of sufficient scale.

Total mercury in marine biota will be determined by flameless atomic absorption spectrophotometry (for a detailed analytical methodology of total mercury determination in marine organisms see UNEP/FAO/IAEA/IOC, 1984a).

Total arsenic in marine biota will be determined by hydride generation atomic absorption spectrophotometry (for a detailed analytical methodology of total arsenic determination in marine organisms see UNEP/FAO/IAEA/IOC, 1985).

Total cadmium, total zinc and total copper in marine biota will be determined by atomic absorption spectrophotometry (for a detailed analytical methodology of total cadmium, zinc and copper determination in marine biota see UNEP/FAO/IAEA/IOC, 1984b).

---

[1] Highly unlikely that any trend signal related to anthropogenic influences will be detected

Total mercury in marine sediments will be determined by flameless atomic absorption spectrophotometry (for a detailed analytical methodology of total mercury determination in sediments see UNEP/IAEA, 1985a and UNEP/IOC/IAEA, 1995).

Total cadmium in marine sediments will be determined by flameless atomic absorption spectrophotometry or graphite furnace atomic absorption spectrophotometry when Cd is too low in concentration to be determined by FAAS (for a detailed analytical methodology of total cadmium determination in sediments see UNEP/IAEA, 1985b and UNEP/IOC/IAEA, 1995).

Total copper in marine sediments will be determined by flame atomic absorption spectrophotometry or graphite furnace atomic absorption spectrophotometry when Cu is too low to be determined by FAAS (for a detailed analytical methodology of total copper determination in sediments see UNEP/IAEA, 1985c and UNEP/IOC/IAEA, 1995).

Total zinc in marine sediments will be determined by flame atomic absorption spectrophotometry (for a detailed analytical methodology of total zinc determination in sediments see UNEP/IAEA, 1986 and UNEP/IOC/IAEA, 1995).

High molecular weight halogenated hydrocarbons in sediment and biota and polynuclear aromatic hydrocarbons in biota will be determined by gas chromatography. PCBs and DDTs and other halogenated pesticides will be analysed by packed column gas chromatography or, where available, with more advanced high resolution capillary gas-chromatographs (for a detailed analytical methodology of high molecular weight halogenated hydrocarbons and polynuclear aromatic hydrocarbons see UNEP/FAO/IOC/IAEA. 1986, UNEP/IOC/IAEA, 1988, UNEP/IOC/IAEA, 1992 and UNEP/IAEA/IOC/FAO, in preparation).

2.5 Selection of the sampling matrices

The use of biota to quantify the degree of contamination of aquatic environments has a number of theoretical and practical advantages over the analysis of either natural waters or sediments in monitoring programmes. Most biomonitors exhibit contaminant concentrations which permit relatively simple measurement compared to seawater analysis. The accumulated concentrations of contaminants reflect an average of the short-term temporal fluctuations in contaminant abundance in the ambient water.

Most contaminants of concern in aquatic ecosystems tend to associate preferentially with suspended particulate material, rather than being maintained in solution, although this varies in extent between individual contaminants. Sampling of surficial sediment samples has the advantage of a simple methodology and more general availability. Sediment down-core analysis in order to identify past time trends requires more elaborate sampling material and expertise. Attention should be paid to site selection in relation to sedimentation rates and bioturbation intensity, favoring sites with high sedimentation rates and low bioturbation intensity.

Seawater analysis is not recommended for trend monitoring purposes (except for lindane, or in sites of marked contamination or where appreciable changes in contaminant levels are expected). It is difficult and very costly to obtain high quality data from seawater, particularly for trace metals and in addition it will be difficult to obtain trend information due to the high variability in time and space.

Biota and sediments are therefore considered as the primary matrices for the sampling of contaminants for trend monitoring purposes, presenting the advantage of integrating

contamination over time. Biota and sediments are primary matrices for the measurement of total mercury, cadmium, zinc, copper and high molecular weight halogenated hydrocarbons.

Biota only are suitable matrices for the measurement of polynuclear aromatic compounds. Sediment profiles are an unsatisfactory matrix for the measurement of arsenic while no recommendations have yet been made for arsenic measurement in biota.

2.6 Species to be selected for the measurement of contaminants

The trend monitoring programme will carry out measurements for contaminants in species most closely fulfilling the objectives of the programme while at the same time selecting species adhering to the greatest extent possible to the following criteria:

- A simple relationship exists between contaminant concentrations in the species and average concentrations in the surrounding environment;
- The species accumulates the contaminant without being affected by the concentrations encountered in its environment (other than in polluted areas);
- The species is sedentary and thus represents the collection area;
- The species is widespread and abundant in the study region, to allow comparisons among different areas;
- The species lives long enough so that more than one year-class can be sampled, if desired;
- The species is large enough to yield sufficient tissue for analysis;
- The species is easy to collect and hardy enough to survive unfavourable conditions or within the laboratory;
- The species exhibits high bio-accumulation factors, to allow analysis without preconcentration;
- The species tolerates brackish water, to allow comparisons between estuarine and offshore sites;
- The species must be easy to identify with certainty.

The following species were used in the past for MED POL monitoring purposes:

- Bivalves
  Mytilus galloprovincialis, or
  Mytilus edulis, or
  Perna perna, or
  Donax trunculus
  The latter three species were suggested as alternative species if *M. galloprovincialis* did not occur in the area
- Demersal fish
  Mullus barbatus, or

Mullus surmuletus, or
Upeneus mollucensis
The latter two species were suggested as alternative species if *M. barbatus* did not occur in the area.

- Pelagic carnivore fish
Thunnus thynnus, or
Thunnus alalunga, or
Xiphias gladius
The latter two species were suggested as alternative species if *T. thynnus* did not occur in that area.

- Pelagic plankton feeding fish
Sardina pilchardus
Other clupeids were suggested as alternative species if *S. pilchardus* did not occur in the area.

- Crustaceans
Parapenaeus longirostris, or
Nephrops norvegicus, or
Penaeus kerathurus
The latter two species were suggested as alternative species if *P. longirostris* did not occur in the area.

For efficient, cost-effective trend monitoring, particularly in sites which are not directly affected by large pollution discharges or yearly marked changes in discharges due to regulation, it may me useful to focus on one or two species selected as trend monitors. Mollusc species show many of the characteristics involved in the criteria for selection of species for the purpose of contaminant monitoring, and in particular, reflecting the environment's response to changes in inputs. Common mussels are suitable for trend monitoring programmes in temperate coastal waters, reflecting contaminant loads in the water column. Deposit feeders reflect contaminant loads in the sediment.

2.7 The tissues selected for analysis of contaminants in biota

For molluscs, whole soft tissue may be selected for analysis, as it is a relatively simple process providing enough material.

For large crustaceans the digestive gland (hepatopancreas) may be used, which concentrates metallic and organic contaminants.

For fish, muscle may be a suitable tissue for most purposes although it is usually selected for public health concerns, or where liver and kidneys may not provide sufficient tissue for analysis. Most toxic metals accumulate in liver and kidneys. Fatty tissues accumulate hydrocarbons and organochlorines.

2.8 Timing and frequency of sampling

It is a considerable demand on resources to sample and analyse biota and sediments several times every year, where this is not essential. For this reason sampling of biota for trend monitoring of contaminants could generally take place once every year, while sampling of sediments for trend monitoring of contaminants could take place over a larger time frame.

Carrying out sampling of biota during a period in the year when contaminant concentrations are not being significantly affected by changes in biological events, is essential for consistency of sampling. Such periods of minimal change are generally related to periods outside the spawning cycle and when food supply is relatively constant. Food supply and the spawning period are known to cause changes in total body weight, lipid concentration, lipid composition and therefore contaminant levels. In order to avoid such variations it is recommended that sampling take place in the pre-spawning period. In order to obtain comparable data from the various sampling stations it is necessary to establish the pre-spawning period at all these stations in order to ensure that samples are taken at correct occasions.

2.9 The number of specimens to be included in each sample

The number of specimens needed to detect important trends depends on the type of the trend. the magnitude of the trend and the variability in the data. In order therefore to choose an appropriate number of specimens, the statistical power of the monitoring programme should be considered through power studies which examine the types and magnitude of changes that will be detected for a given number of specimens. This process is described in the attached Annex.

The number of specimens within a sample of fish should be sufficient to allow the sample to be collected in a length-stratified manner, i.e., the size of the fish should include as wide a length-range as possible and there should be an equal number of individuals in each length-grouping. The length range of the sample should be such that the individuals in the lower length group yield sufficient tissue for the chemical analyses, and the individuals in the upper length-group should be such that at least 5 individuals can readily be found in the sampled catch. The length range should be divided into at least 5 length intervals of equal size (the log upper length group minus the log lower length group should be equal for each length interval). No length interval should be less than 2-3 cm, otherwise the species is considered as not ideally suited for the analysis. In order to cover the above requirements, at least 25 individuals should comprise a sample of fish. The agreed length-stratification for a particular species should be strictly adhered to each year.

The number of specimens within a sample of mussels should be sufficient to allow as wide a size-range as possible and to provide sufficient material for analysis in groups of the different sizes. Indicatively, a sample of 5-10 individuals for each size range would be obtained. For small crustaceans a sample of 25-50 individuals may be required.

There are as yet no concrete guidelines for the determination of the number of sediment samples to be used in the trend monitoring of contaminants in sediments. Further developments are expected from the ICES Working Group on Environmental Assessment and Monitoring Strategies and the ICES Working Group on Statistical Aspects of Environmental Monitoring.

2.9.1 Pooling of specimens of biota

It may be necessary to pool (bulk) fish tissues, particularly in the case of fish livers and mussel and other shellfish tissues, in order to provide sufficient quantities of material for chemical analysis.

Pooling can distort the statistical analysis of log-transformed data by increasing the yearly mean concentration values and decreasing the power of tests to detect trends (Nicholson *et al.*, 1989).

Nicholson *et al.* (1989) however have shown that in general pooling does not influence trend identification (i.e. differences between years and associated regression coefficients will be unaffected, although trends may be less precisely estimated than from unbulked data), if pooling is consistent between years, i.e. if samples consist of the same number of pools, which contain the same number of specimens.

Keeping the same number of individuals in the pool between years is the most important aspect, i.e. in the pool, for a given length class, the number should be the same each year. It is also important to maintain the same number of pools each year (preferably based on length-stratification of the sample if possible).

3. SUPPORTING MEASURES

The success of the trend monitoring programme will depend largely on the development of adequate supporting measures, consisting of methods for quality assurance and assistance to the participants in the monitoring programme.

3.1 Quality Assurance

Quality assurance of the monitoring programme refers to those procedures which are developed to ensure that analytical results are valid, traceable, reproducible, representative, complete and accurate, i.e. close to the true value; as well as measures developed to assess performance. Methods of quality assurance collectively consist of methods for quality control and quality assessment.

3.1.1 Quality Control Methods

The design of quality control methods involves the development of procedures for each step of the trend monitoring programme which contribute to the eventual production of quality data, and procedures to ensure that each participating laboratory will produce comparable data.

Quality control methods will involve the following:

a) Standard sampling and measurement procedures, including:

- species selection (i.e. methods to ensure the ability to distinguish between two native species belonging to the same genus)

- sample handling (including methods of storage, transportation, preservation, analytical sub-sampling, dissection, homogenization, bone grinding of tissue and sampling, sieving and grinding of sediments)

- biological measurements (methods for measurement of length, total weight, organ weights, methods for the determination of age, sex and fat and water content)

- chemical measurements (methods for the analysis of chemical contaminant residues)

Guidelines and recommended procedures for the storage and pre-treatment of samples following their collection are given in UNEP/FAO/IAEA (1984) and UNEP/FAO/IAEA/IOC (1984c) dealing with heavy metals and halogenated hydrocarbons respectively.

Important procedures for consistent sampling, sample preparation and chemical analysis for trend monitoring of contaminants in biota are described in UNEP/FAO/IOC/IAEA (1993), Uthe, J.F. (1994) and Uthe *et al.* (1991). The role of consistency in the sampling process is described in Fryer (1993).

Guidelines and recommendation procedures for sediment sampling, handling etc., are given in UNEP/IOC/IAEA (1995), Mudroch and Ascue (1995), Mudroch and MacKnight (1994).

b) Data handling procedures. including methods of data translation, data transcription and keeping records of calculations, methods for long-term storage of data in log-books or computer files and methods of data reporting.

Data handling procedures are described in UNEP/FAO/IOC/IAEA (1993), UNEP/IOC/IAEA/FAO (1989), Uthe, J.F. (1994) and Uthe *et al.*, (1991).

c) Use of certified reference materials (CRMs) of identical or similar matrix as the sample analysed and covering the concentration ranges likely to be encountered in the measurements, in order to obtain accurate and precise data following routine measurements. Available reference materials, for the choice of appropriate CRMs are listed in IAEA (1995).

d) Regular analysis of reference materials throughout the monitoring programme in order to ensure that analytical performance is maintained.

e) Regular mandatory participation of the laboratories involved in the trend monitoring programme, in intercomparison exercises in order to ensure the comparability of the data being produced among the participants in the programme.

Intercomparison exercises will involve the analysis, through methods used in their normal monitoring work, by each participating laboratory, of blind samples containing unknown concentrations of a specified analyte or samples of known but undisclosed concentrations of the analyte.

f) Regular calibration, servicing and maintenance of all the equipment

3.1.2 Quality Assessment Methods

Quality assessment methods will be developed for the assessment of the performance of individual laboratories participating in the monitoring programme over time and in relation to the other laboratories participating in the programme

The assessments will be based on the analyses of CRMs and other reference materials. The quality assessment methods will include rules for the selection of reference materials and the frequency of their analyses when carrying out interlaboratory assessments of performance with respect to the quality of trend data.

Methods will be developed for the regular documentation on control charts of the results of analyses of the same reference material over a period of time, in order to obtain

information on whether the results are within the acceptable limits of accuracy and precision and how this information is comparable to the results of other laboratories participating in the programme.

3.2 Other forms of assistance

The necessary training will be provided to ensure that all laboratories are able to participate in the trend monitoring programme and in the relevant quality control and quality assurance activities.

Where problems continue to exist in the sampling process or in the analysis of contaminants, or both, the necessary assistance will be provided on a case-by-case basis for any participant requiring such assistance.

Where continuing analytical problems have been established for a participant, a methodology for split-sampling and analysis of contaminants by another laboratory will be formulated.

4. REFERENCES

Carlberg, S.R. (1993). Correct sampling-What is that? In: Report of the Second FAO/ UNEP/IAEA Training Workshop on the Design of Monitoring Programmes and Management of Data Concerning Chemical Contaminants in Marine Organisms. (Athens, Greece, 22-26 June). FIR/MEDPOL/ATH/3. pp.99-114

Fryer, R.J. (1992). Statistical analysis of the MED POL monitoring data on heavy metals and halogenated hydrocarbons in biota with special reference to temporal trends (1974-1991). UNEP (Restricted circulation), mimeo 147 p.

Fryer, R.J. (1993). The Role of Consistency in the Data Collection Process In: Report of the Second FAO/UNEP/IAEA Training Workshop on the Design of Monitoring Programmes and Management of Data Concerning Chemical Contaminants in Marine Organisms. (Athens, Greece, 22-26 June). FIR/MEDPOL/ATH/3. pp.151-161

Fryer, R.J. and M.D. Nicholson (1993). The power of a contaminant monitoring programme to detect linear trends and incidents. ICES J.mar.Sci., 50:161-168

IAEA (1995). Survey of reference materials, Vol. 1: Biological and environmental reference materials for trace elements, nuclides and microcontaminants. IAEA-TECDOC-854

Mudroch, A. and J.M. Azcue (1995). Manual of aquatic sediment sampling. Lewis Publishers

Mudroch, A. and S.D. MacKnight (1994). Handbook of techniques for aquatic sediment sampling. Lewis publishers

Nicholson, M.D., N. Green and S.J. Wilson (1989). Trend measurement from bulked samples. ICES, Doc. C.M. 1989/E:13 (mimeo)

Phillips, D.J.H. and P.S. Rainbow (1994). Biomonitoring of Trace Aquatic Contaminants. Chapman and Hall. 371 p.

UNEP (1986). Long-Term Programme for Pollution Monitoring and Research in the Mediterranean (MED POL)- Phase II. UNEP Regional Seas Reports and Studies no 28 Rev.1. 26 p.

UNEP (1993a). Evaluation Report of the MED POL Programme. UNEP(OCA)/MED IG.3/Inf.6. 62 p.

UNEP (1993b). Report of the Eight Ordinary Meeting of the Contracting Parties to the Convention for the Protection of the Mediterranean Sea Against Pollution and its Related Protocols. (Antalya, Turkey, 12-15 October) UNEP(OCA)/MED IG.3/5

UNEP (1996). Report of the Extraordinary Meeting of the Contracting Parties to the Convention for the Protection of the Mediterranean Sea Against Pollution and its Protocols. UNEP(OCA)/MED IG.8/7 Annex IV

UNEP/FAO/IAEA (1984). Sampling of selected marine organisms and sample presentation for the analysis of chlorinated hydrocarbons. Reference Methods for Marine Pollution Studies No. 12 Rev.1, 19 p.

UNEP/FAO/IAEA (1985). Determination of total arsenic in selected marine organisms by hydride generation atomic absorption spectrophotometry. Reference Methods for Marine Pollution Studies No. 9, 15 p.

UNEP/FAO/IAEA/IOC (1984a). Determination of total mercury in selected marine organisms by cold vapour atomic absorption spectrophotometry. Reference Methods for Marine Pollution Studies No. 8 Rev. 1, 17 p.

UNEP/FAO/IAEA/IOC (1984b). Determination of total cadmium, zinc, lead and copper in selected marine organisms by flameless atomic absorption spectrophotometry. Reference Methods for Marine Pollution Studies No. 11 Rev. 1. 21 p.

UNEP/FAO/IAEA/IOC (1984c). Sampling of selected marine organisms and sample preparation for trace metal analysis. Reference Methods for Marine Pollution Studies No. 7, Rev. 2, 19 p.

UNEP/FAO/IOC/IAEA (1986). Determination of DDTs and PCBs in selected marine organisms by packed column gas chromatography. Reference Methods for Marine Pollution Studies No. 14, Rev. 1, 20 p.

UNEP/FAO/IOC/IAEA (1993). Guidelines for monitoring chemical contaminants in the sea using marine organisms. Reference Methods for Marine Pollution Studies No. 6, 28 p.

UNEP/IAEA (1985a). Determination of total mercury in marine sediments and suspended solids by cold vapour atomic absorption spectrophotometry. Reference Methods for Marine Pollution Studies No. 26, 15 p.

UNEP/IAEA (1985b). Determination of total cadmium in marine sediments by flameless atomic absorption spectrophotometry. Reference Methods for Marine Pollution Studies No. 27, 13 p.

UNEP/IAEA (1985c). Determination of total copper in marine sediments by flameless atomic absorption spectrophotometry. Reference Methods for Marine Pollution Studies No. 33, 13 p.

UNEP/IAEA (1986). Determination of total zinc in marine sediments by flame atomic absorption spectrophotometry. Reference Methods for Marine Pollution Studies No. 39, 10 p.

UNEP/IAEA/IOC/FAO (in preparation). Sample work-up for the analysis of chlorinated hydrocarbons in the marine environment. Reference Methods for Marine Pollution Studies No. 71

UNEP/IOC/IAEA (1988). Determination of DDTs and PCBs by capillary gas chromatography and electron capture detection. Reference Methods for Marine Pollution Studies No. 40, 18 p.

UNEP/IOC/IAEA (1992). Monitoring of petroleum hydrocarbons in sediments. Reference Methods for Marine Pollution Studies No. 20, 72 p.

UNEP/IOC/IAEA (1995). Manual for the geochemical analyses of marine sediments and suspended particulate matter. Reference Methods for Marine Pollution Studies No. 63, 74 p.

UNEP/IOC/IAEA/FAO (1989). Contaminant monitoring programmes using marine organisms: Quality Assurance and Good Laboratory Practice. Reference Methods for Marine Pollution Studies No. 57, 23 p.

Uthe. J.F. (1994). Sample Handling-Samples. In: Report of the Second FAO / UNEP Sub-Regional Workshop on the Monitoring of Chemical Contaminants in Marine Biota for Trends. (Lesvos Island, Greece, 19-22 October). FIR/MEDPOL/LES/4. pp.48-53

Uthe. J.F., C.L. Chou, R.K. Misra, P.A. Yeats, D.H. Loring, C.J. Musial and W. Cofino (1991). Temporal trend monitoring: Introduction to the study of contaminant levels in marine biota. ICES Techniques in Marine Environmental Sciences, No 14. 18 p.

ANNEX

Power studies of a programme design for temporal trend monitoring in biota

Table of contents

INTRODUCTION

The present annex aims to explain how to carry out a power study for the design of a monitoring programme. The explanation follows the applied statistics which have been developed for the planning of programmes using biota to detect temporal trends of contaminants in the environment as well as following general guideline for UNEP' s Regional Seas programmes. The presentation is relevant to the statistical methodology itself, therefore any environmental issues for the programme design (such as site and species selection) are out of the scope of the present annex and are reported in the main part of the document.

In a broad sense, power studies encompass activities which link together the specification, the optimization and the effectiveness (i.e. how good the programme is) of a monitoring programme through the use of statistics. A power study may be directed either at estimating the effectiveness of a programme which follows a specific *ad-hoc* protocol as well as having a certain laboratory performance, or the identification of the number of samples that should be collected to detect an important trend, or the estimation of which intensity (i.e. how strong the variation of contaminant concentration is over time) in trend can be detected by a programme etc.

The exercise can be roughly summarized in the following steps:

a) make clear the purpose of the study and "translate" it into a statement (monitoring objective) that can be handled through the use of statistical methods;

b) gather all the information which is required for the study and perform the actual calculation;

c) compare the result of the study with any expected (or required) values which may have been set prior to the beginning of the study. If the result is not satisfactory, decide on any possible revision of the protocol of the programme. An example may explain the issue: the investigator wants to decide the number of samples to be collected to achieve a target value in the effectiveness of the programme and reaches the conclusion that the adopted protocol cannot provide this value regardless of the number of samples, so the programme' s protocol should be revised to search for any optimization.

The annex is organized as follows:

- the first section explains how to formulate the monitoring question about the trend in a statistically sound way;

- the second section deals with monitoring objectives, in particular it presents how the objective should be formulated and which information it should include;

- the third section explains the power calculation: it presents the relevant quantities and emphasizes such issues as how many samples should be collected and what is the maximum allowed level of error to detect an important trend;

- the fourth section provides examples of power studies for objectives relevant to some monitoring questions.

1. FORMULATION OF MONITORING QUESTIONS ABOUT TRENDS

This section is relevant to an important part of applied statistics: the test of hypothesis, which can be found in any text book (such as Zar, 1984 and Sokal and Rohlf, 1995).

The investigator has in mind a question such as : *is there a trend in contaminant level over time?* To search evidence about the answer from the monitoring data this person formulates two mutually exclusive (i.e. if one is true the other is false and vice versa) hypotheses:

- the first denotes "no difference" in contaminant level over time and is called null hypothesis (H_0);

- the second denotes "difference" and is called alternative hypothesis (H_1),

subsequently the investigator tests H_0 against H_1 through the use of a proper statistical test. If H_0 is rejected in favor of H_1 then the investigator concludes that a trend in contaminant level over time exists. It is important to note that the non-rejection of the null hypothesis does not necessarily mean that the hypothesis is true: it denotes that there is not enough evidence to conclude that it is false.

Example: how to formulate the hypothesis

The investigator searches evidence for a linear trend of contaminant level over time. This trend is written as: $y = a + bt$, where b is the variation per time unit (or slope of y versus t).

$b=0$ denotes no trend of y over t (i.e. a flat trend) whereas $b \neq 0$ denotes trend.

The hypotheses are :
H_0.: $b=0$ and H_1.: $b \neq 0$.

1.1 Error in hypothesis testing and power

The investigator should realize that a true null hypothesis occasionally will be rejected, which means that a trend is erroneously identified when in fact it does not exist. This error occurs with probability α and its value is chosen by the investigator itself. A common value (but not mandatory) is $\alpha=0.05$ which means that the investigator is keen to accept to wrongly reject a true null hypothesis 5% of the time. The probability of accepting a true null hypothesis is $1-\alpha$. On the other hand if H_0 is in fact false the test may not detect it. This error occurs with probability β. **The power of the test is the probability of rejecting the null hypothesis when it is false and therefore should be rejected, and is defined as $1-\beta$** (Zar, 1984). In trend monitoring $\beta=10$ % (i.e. power= 90%) may be considered an acceptable value, but this is not mandatory.

The following table summarizes the probabilities of having correct or wrong answers from the test for trend (showing in parenthesis the above-mentioned typical values for probabilities):

Decision taken	Null Hypothesis true	Null Hypothesis false
Accept Null Hypothesis	1-α (95%)	β (10%)
Reject Null Hypothesis	α (5%)	power=1-β (90 %)

It is important to remark that both errors (α, β) cannot be minimized for a given programme and intensity for the trend: the smaller α is chosen the larger β will be. Should both errors be small, then only a strong trend is likely to be detected. Table 1 shows, as an example, the minimum detectable trend for different combinations of α and β. The values refer to a sensible set of specifications (i.e. the values are those reported for medium variability in example 4.1 which is described later in the text) and therefore should not be generalized).

Table 1

Test for a linear trend: minimum detectable trend (b) for different α and β values

α	β	relative variation of contaminant concentration per year
.05	.05	.13
.1	.1	.10
.2	.2	.07
.05	.10	.12
.05	.20	.10
.10	.05	.12
.20	.05	.10
.10	.20	.09
.20	.10	.09

2. OBJECTIVES OF MONITORING PROGRAMMES

The power study of the design of a programme is based on a monitoring objective which is:

a. sufficiently detailed;

b. expressed in a measurable way;

c. formulated in a way to allow handling through statistical techniques (statistically sound formulation).

In particular the objective must account for:

(i) the trend itself that should be detected (magnitude and shape over time);

(ii) the design information such as programme duration, number of specimens to be collected and pooling of specimens into a composite sample;

(iii) the managerial aim relevant to the effectiveness of the programme i.e., how much the assurance of a correct answer from the programme is desirable;

(iv) information on the uncertainty which affects the estimated value for the level of contaminant in the environment. This uncertainty arises from the analytical performance of chemical analysis (precision and accuracy) and variability in the population of monitoring organisms.

Point (i) above describes the trend in terms of the magnitude and shape of variation of contaminant level over time. The former refers to the maximum difference in contaminant level in the time-series. The latter accounts for how this variation occurs over time. For example variation may be easily described by an upward or downward shape such as a straight line or an exponential variation. Other shapes, such as random variation, may indeed occur.

An objective which is formulated in the proper way allows to link monitoring questions, programme specification and effectiveness and to focus the attention on the problem to be solved. Those designing the programme will formulate the objective according to their questions. Typical questions may be *how many samples are required to detect an important trend?* or *what is the maximum allowed level in error to detect an important trend* or *what is the minimum trend which can be detected given the level in error and the number of samples.*

2.1 Examples of monitoring objectives

This section provides examples of monitoring objectives. The calculation is described in section 4.

Objectives may include statements which can be arbitrary. These are relevant to the duration of the programme and the intensity of the trend. The former may be either the minimum number of years which is required to have a reliable answer from the programme or the time-span in which any answer should be provided. The latter refers to the intensity which is expected. If there is no interest in any of these quantities they must be intended to be statistical requirements, i.e., they must be components of the power study.

OBJECTIVE 1

Suppose that the manager of an international monitoring programme has gathered information on components of errors from several laboratories and different areas, such as those reported in Table 3. This person wants to see the power that the local programmes are expected to have. Values in Table 3 have been used for demonstration.

The monitoring objective is formulated as follows:

Identify the power of a monitoring programme to detect a trend of 10 % per year in a 10 year programme, at different levels of analytical and sampling variability and number of specimens collected

Comment. The choice of the magnitude for the trend is arbitrary. The choice of the duration of the programme may follow managerial constrains.

The calculation is presented in example 4.1

OBJECTIVE 2

Suppose that the investigator of one laboratory has reliable estimates for components of error. The investigator intends to reduce the cost for analysis, therefore the effect of pooling specimens into a composite sample is explored. Values in Table 3 have been used for demonstration.

The monitoring objective is formulated as follows:

Identify the power of a monitoring programme to detect a trend of 10 % per year in a 10 year programme at different levels of pooling and number of specimens collected and one level of analytical and sampling variability.

Comment. The choice of the magnitude for the trend is arbitrary. The choice of the duration of the programme may follow managerial constrains.

The calculation is presented in example 4.2.

OBJECTIVE 3

Section 1.1 presented the topic relevant to errors in hypothesis testing. This example explores the trade-off between α-error and minimum detectable trend. Suppose that the investigator intends to identify which magnitude in trend is likely to be detected at different values of α (i.e. different level of risk in false rejection of H_0). Values in Table 3 have been used for demonstration. The example has been tailored around the number of specimens collected.

The monitoring objective is formulated as follows:

Identify the minimum detectable trend to identify a linear trend during 10 years at different levels of α and number of specimens collected, with 90 % power.

Comment. The choice of the duration of the programme may follows managerial constrains.

The calculation is presented in example 4.3.

OBJECTIVE 4

Suppose that the investigator intends to identify the target value for ψ required to detect important trends. The monitoring objective is formulated as follows:

Identify the maximum value for total error to detect a trend of 10 % per year in a 10 year programme, with α=0.05 and 70 % power.

Comment. The choice of the magnitude for the trend is arbitrary. The choice of the duration of the programme may follow managerial constrains.

The calculation is presented in example 4.3.

3. COMPUTATION OF STATISTICAL POWER

This section explains how to calculate the statistical power. Part 3.1 describes which quantities are involved; part 3.2 reports on the calculation of power for the test for a linear trend. Appendix 1 describes the statistical methods employed in part 3.2. Appendix 2 explains how values for error can be obtained from monitoring data.

3.1 Which quantities are involved in the calculation of power

The probability of having a correct answer from the programme (i.e. 1-α or 1-β as described in section 1) depend on:

- the signal-to-noise ratio of contaminant variation over time;

- the specifications of the programme such as: duration, number of specimens collected and degree of pooling into a composite sample.

In particular its value increases with:

- duration (hereafter denoted as T in years);

- number of specimens (R),

whereas its value decreases with:

- noise

- degree of pooling, i.e., how many specimens are included in a composite sample (I) prior to its analysis.

The magnitude of the trend as well as its shape over time should be accounted for when the value for intensity of the signal is calculated. In fact theory (Fryer and Nicholson, 1993) shows that certain shapes of trend are detected easier than others.

Noise denotes the uncertainty between the measured level of contaminant in the environment and its actual value (which is unknown). For programmes following UNEP guidelines (UNEP/FAO/IOC/IAEA, 1993) noise is conveniently defined as the total residual variance (ψ^2) about the mean:

$$\psi^2 = \sigma_y^2 + \sigma_w^2/R + \tau_y^2 + I\,\tau_w^2/R \quad (1)$$

where R is the number of specimens collected and I is the number of specimens which may be pooled in each composite sample. The components for error are:

σ_y : Random between year sampling variability. This component accounts for the error between the contaminant level measured in the monitoring population and its actual level in the environment. It arises from uncontrolled sources which affect the entire population. Simply, this component attempts to quantify the "accuracy" of the monitoring population to estimate the level of contaminant in the environment;

σ_w. Random sampling variability within a year. This is the variability which is left in the annual data after the effect of any co-variables (such as size) has been accounted for. This component arises from individual variability in the accumulation features within the monitoring population;

τ_y. Accuracy of chemical analysis;

τ_w. Precision of chemical analysis.

The method of calculation of the components of sampling error is provided in Appendix 2.

3.2 Power of the test for a linear trend

The previous section (3.1) shows that the signal must be defined also with respect to the shape of variation of contaminant level over time. Therefore the shape of the trend should be decided prior to planning any test for a trend.

It is sensible to aim to detect a linear trend as this shape is among the easiest to be identified (Nicholson and Fryer, 1992). Therefore a programme which is not likely to detect a linear trend will not identify other more difficult shapes.

Power is a function of:

- α-value;
- duration of the programme (T);
- signal-to-noise ratio ($|b|/\Psi$).

Power values, programme duration and signal-to-noise ratio (computed for $\alpha = 0.05$) have been combined in Table 2. Provided that the statistical requirements of the test are satisfied (see Appendix 1), Table 2 can be used to solve monitoring objectives (see example 4.4).

Table 2 may be insufficient (for example if $|b|/\Psi$ is out of the Table or $\alpha \neq 0.05$ is chosen) or tedious to use on a routine basis, thus the computation can be carried out with the proper probability distribution. Power can be calculated from a non-central F distribution on 1 and T-2 degrees of freedom, with non-centrality parameter δ (see Zar, 1984).

The quantity δ accounts for the signal-to-noise ratio as well as for the number of years. For a linear trend δ takes the form (Nicholson *et al.*, 1996):

$$\delta = b^2(T-1)T(T+1)/12\psi^2 \quad (2)$$

$$power = 1 - \text{prob } F(F1-\alpha, 1, T-2, \delta), \quad (3)$$

where F1-α is the 100(1-α)th percentile of a central F-distribution on 1 and T-2 degrees of freedom. These probabilities can be obtained from statistical tables (fpr a full description of the procedure see Cohen, 1977), or they can be computed with packages such as SAS, SPLUS, STATABLE (and perhaps others). The manuals of the packages are sufficiently detailed to allow the actual computer function to be written easily.

Table 2

Test for a linear trend. Values of $|b|/\psi$ corresponding to different powers (rows) and numbers of years (columns). α=0.05 (from Nicholson *et al.*, 1996, modified)

	0.50	0.60	0.70	0.80	0.90	0.95	0.99
5	0.906	1.035	1.176	1.344	1.584	1.786	2.175
6	0.616	0.700	0.791	0.899	1.051	1.178	1.421
7	0.459	0.520	0.586	0.664	0.773	0.864	1.037
8	0.360	0.408	0.459	0.520	0.604	0.674	0.806
9	0.293	0.332	0.373	0.422	0.490	0.546	0.653
10	0.245	0.278	0.312	0.352	0.409	0.455	0.544
11	0.209	0.237	0.266	0.300	0.348	0.388	0.462
12	0.181	0.205	0.230	0.260	0.301	0.335	0.400
13	0.159	0.180	0.202	0.228	0.264	0.294	0.350
14	0.141	0.160	0.179	0.202	0.234	0.261	0.311
15	0.127	0.143	0.161	0.181	0.210	0.233	0.278
16	0.114	0.129	0.145	0.163	0.189	0.211	0.251
17	0.104	0.117	0.132	0.148	0.172	0.191	0.228
18	0.095	0.107	0.120	0.136	0.157	0.175	0.208
19	0.087	0.098	0.110	0.125	0.144	0.160	0.191
20	0.080	0.091	0.102	0.115	0.133	0.148	0.176
21	0.078	0.084	0.094	0.106	0.123	0.137	0.163
22	0.069	0.078	0.088	0.099	0.115	0.127	0.152
23	0.065	0.073	0.082	0.092	0.107	0.119	0.141
24	0.060	0.068	0.077	0.086	0.100	0.111	0.132
25	0.057	0.064	0.072	0.081	0.094	0.104	0.124

Table 3

Values for components of error (logarithmic scale) relevant to mercury contamination in fish from an international monitoring programme for three levels of variability (Anon, 1995)

variability	σ_w	σ_y	T_y	T_w
Low	0.08	0.22	0.09	0.04
Medium	0.26	0.28	0.13	0.05
High	0.52	0.42	0.24	0.10

4. EXAMPLES OF POWER STUDIES FOR CERTAIN MONITORING OBJECTIVES

This section describes the calculation relevant to the examples of objectives provided in section 2.1. Because of the routine calculation which is required to prepare figures 1 to 3 those interested in repeating the example as an exercise should perform at least part of the calculation with a calculator:

a. If Table 2 is used, calculate ψ;

b. If the power is calculated from a non-central F distribution (either using software or statistical tables), calculate ψ and δ.

Example 4.1 (objective 1)

For convenience the calculation has been made only for three combinations of errors in Table 3 (along the rows), for I=1 (no pooling) and α=0.05.

a. calculate δ for different values of R (number of specimens), by substituting values accordingly in equations (1) and (2);

b. calculate power from non-central F distribution on 1, T-2 degrees of freedom and non-centrality parameter δ.

The result is shown in figure 1.

Remark. The study aims to explore extreme scenarios of monitoring performance, so it is assumed that low sampling variability occurs with low analytical variability and high sampling variability occurs with high analytical variability. In the first case (fig. 1 top) power is very high whereas in the second instance the level for power is unacceptable (fig. 1, bottom). A more realistic picture in given by a medium level of variability (Table 3, center). In this case the programme is likely to have a power ≈ 90 % (fig.1, middle)

Example 4.2 (objective 2)

The computation follows example 4.1 except for the calculation of δ which takes into account different combinations of I and R.

The result is shown in figure 2.

Remark. Following a strict statistical consideration, the study suggests that pooling does not affect power, provided that a reasonably large number of specimens are collected (i.e., R>20). The reason for this depends on the large value for σ_y which dominates ψ (Table 3). σ_y does not change with R or with I (see formula 1), therefore any optimization of the programme should be directed to reducing σy.

Example 4.3 (objective 3)

a. Calculate from the non-central F distribution the value of δ at α, and 1, T-2 degrees of freedom which leads to power=90 %;

b. calculate b from δ for different values of R (by substituting values accordingly in equations (1) and (2)).

The result is shown in figure 3.

Remark. The example explores the trade-off between α, β and minimum detectable trend. For simplicity β has been held constant. The choice of α and target value for β depends on the managerial aim of the programme and the risk (in managerial terms) of an α or a β error. A programme for general trend surveillance (i.e. no trend direction is expected) where the programme itself is not meant to be a regulatory tool, may have α and β of the same magnitude (in fig. 3 α=β=10 %). In a programme to monitor a recovery measure, an α-error is less problematic than a β-error, so a larger α allows the detection of smaller trends.

Example 4.4 (objective 4)

The following example describes the use of Table 2.

The monitoring objective is satisfied (for T=10 and power 70 %) if:

$|b|/\psi \geq 0.312$; then the maximum error is:

$\psi \leq 0.06$. This value is compared with those that arise from the components in Table 3 (for R=25 and I=1):

- low variability ψ^2=0.0165, ψ=0.128
- medium variability ψ^2=0.087 ψ=0.296,

therefore medium variability will lead to the target value being satisfied.

Suppose now a more strict value for power, say 90 %. From Table 2:

$|b|/\psi \geq 0.409$; then

$\psi \leq 0.244$. Therefore only low variability (Table 3) will satisfy the target value for ψ.

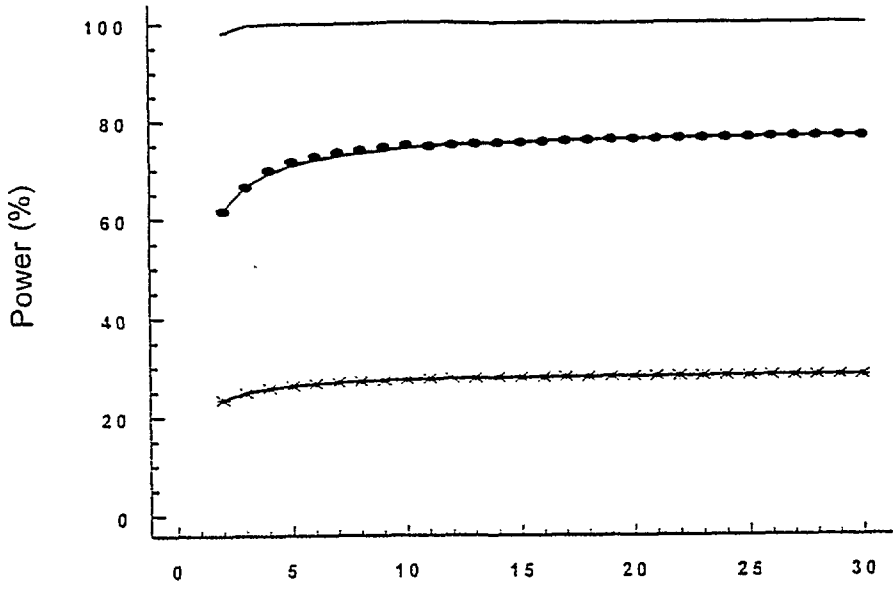

Figure 1. Objective 1 (example 1). Power (%) for the test for a linear trend versus number of samples per year for different levels of error. α=0.05. No pooling (l=1). Values for *b*, *T* and errors, as in the text.

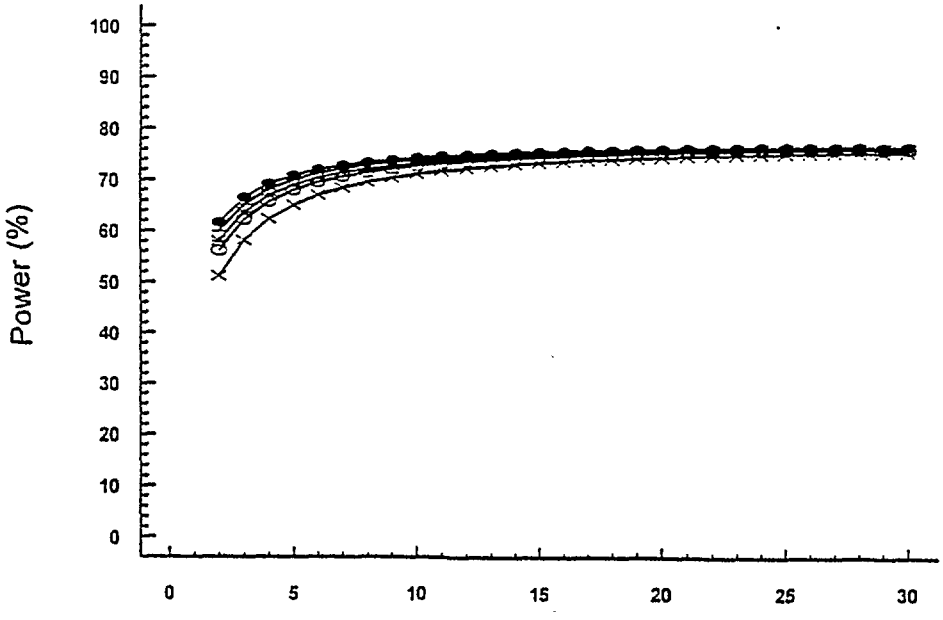

Figure 2. Objective 2 (example 2). Power (%) for the test for linear trend versus number of samples per year for different degree of pooling (●: l=1, +: l= 5, o: l=15, x: l=30). -α value=.05. Values for *b*, *T* and errors, as in the text.

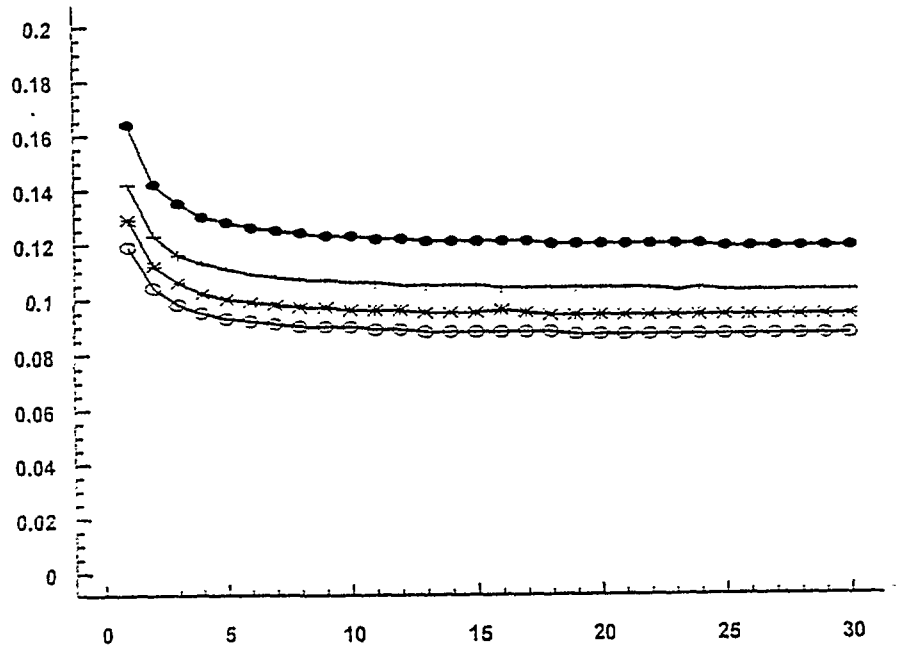

Figure 3. Objective 3 (example 3). Minimum detectable trend (year $^{-1}$) versus number of samples per year for different α values (●: =.05, +: = .1, x: =.15, o: =.20). Target power=90%. No pooling (l=1). Values for *T* and errors, as in the text.

Appendix 1. Description of the statistical method

Let y_t be the mean value of contaminant concentration in the year *t*. A linear trend for the expected value of y_t is written as:

$$E[y_t] = a + bt.$$

Assume that y_t are independent and normally distributed (Zar, 1984) about the linear trend with constant variance:

$$Var[y_t] = \psi^2.$$

Then the evidence for a linear trend is given by the value for *b* (see example 1) which is calculated by regressing y_t on years. The null hypothesis:

$$H_0: b=0$$

is tested against the alternative hypotheses:

$$H_1: b \neq 0$$

by F-test on 1 and *T*-2 degrees of freedom (Zar, 1984).

Appendix 2. How values for components of error can be obtained

Components of error may be estimated from monitoring data, provided they have been documented. In particular accuracy (τ_y) and precision (τ_w) of chemical analysis must be known.

Here it is assumed that all specimens of one year are collected in one occasion (i.e. same time of the year) and analyzed in one batch.

The process goes through the following steps:

1. Calculate ψ^2. This is the residual variance of mean yearly values regressed on years. Unless a certain pattern is the best fit (such as linear, exponential), it is not appropriate to impose a "formula" to the data. Therefore a generic function (such as locally weighted smooth curves), which fits the mean yearly values regardless of any desired pattern being imposed, should be used to calculate ψ^2.

2. Calculate the within year sampling variance (σ_w^2) from the within year variance by correcting it for τ_w (by difference, i.e., subtract τ_w^2 from the within year variance).

3. Calculate the between year sampling variance (σ_y^2) from ψ^2 by correcting it for σ_w^2, τ_y and τ_w (i.e., $\sigma_y^2 = \psi^2 - \sigma_w^2 - \tau_y^2 - \tau_w^2$).

Note on data-transformation. Error values in Table 3 were calculated on a logarithmic scale. This type of transformation has been used (Anon, 1989; Anon, 1991 and Zangrandi, 1996) with monitoring data (contaminants in marine organisms) to satisfy the assumption required by the method (see appendix 1).

5. REFERENCES

Anon (1989). Statistical Analysis of the ICES Cooperative Monitoring Programme Data on Contaminant in Fish Muscle Tissue (1975-1985) for the determination of temporal trends. ICES Cop. Rep. no. 162

Anon (1991). Statistical Analysis of the ICES Cooperative Monitoring Programme Data on Contaminant in Fish Liver Tissue and *Mytilus edulis* (1978-1988) for the determination of temporal trends. ICES Cop. Rep. no. 176

Anon (1995). Report of the ICES Advisory Committee on the Marine Environment. ICES Cop. Res. Rep. no. 212

Cohen, J (1977). Statistical power analysis for the behavioural sciences. Academic press, New York

Fryer, R. and M. Nicholson (1993). The power of a contaminant monitoring programme to detect linear trend and incidents. ICES J. Mar. Sci. 50:161-168

Nicholson, M. and R. Fryer (1992). The statistical power of monitoring programmes. Mar.Poll.Bull. 24(3):146-149

Nicholson, M., R. Fryer and C.A. Ross (1996). Designing monitoring programmes for detecting temporal trends in contaminants in fish and shellfish. Mar.Poll.Bull. (in press)

Sokal, R.R. and F.J. Rohlf (1995). *Biometry 3d ed.* Edited by W.H. Freeman, New York. 887 p.

UNEP/FAO/IOC/IAEA (1993). Guidelines for monitoring chemical contaminants in the sea using marine organisms. Reference Methods for Marine Pollution Studies No. 6, UNEP

Zangrandi, M. (1996). Monitoring the level of contaminant in marine organisms for temporal trend: objectives and effectiveness of a monitoring programme. In: Report of the Working Group on statistical aspects of environmental monitoring. ICES CM 1996/D:1 pp.75-88

Zar, J.H. (1984). *Biostatistical Analysis.* 2nd ed. Prentice-Hall, Englewood Cliffs, N.J., 718 p.