ISSN 0142-2499

AQUATIC ENVIRONMENT MONITORING REPORT
Number 37

Analysis and interpretation of benthic community data at sewage-sludge disposal sites

Prepared by the Benthos Task Team for the Marine Pollution Monitoring Management Group Co-ordinating Sea Disposal Monitoring

MAFF
Directorate of Fisheries Research
Lowestoft, 1993

MINISTRY OF AGRICULTURE, FISHERIES AND FOOD
DIRECTORATE OF FISHERIES RESEARCH

AQUATIC ENVIRONMENT MONITORING REPORT
Number 37

Analysis and interpretation of benthic community data at sewage-sludge disposal sites

Prepared by the Benthos Task Team for the
Marine Pollution Monitoring Management Group
Co-ordinating Sea Disposal Monitoring

LOWESTOFT
1993

This report has been compiled by H. L. Rees, MSc PhD (Chairman of the Benthos Task Team) of the MAFF Fisheries Laboratory, Burnham-on-Crouch, Essex CM0 8HA, from whom further copies can be obtained.

Aquat. Environ. Monit. Rep., MAFF Direct. Fish. Res., Lowestoft, (37), 80pp.

CONTENTS Page

/continued

FOREWORD

This report is the second prepared by a UK Benthos Task Team answerable to the MPMMG Group Co-ordinating Sea Disposal Monitoring, and covers work conducted during the period 1990 and 1991. Its main objective is to provide, for the first time, a detailed assessment of the feasibility of deriving 'Environmental Quality Standards' from benthic community data at UK sewage-sludge disposal sites. This was achieved through analyses, by several means, of a series of test data sets, and a number of recommendations are made regarding future implementation of such an approach. This Section also addresses the implications for future monitoring of identification of benthic organisms at higher taxonomic levels.

Later Sections provide information on the outcome of an intercomparison of laboratory performances in the analyses of samples of benthic macrofauna, along with an assessment of the status of different coding systems for archiving of benthic community data, and a review of methodology and future requirements for sampling of the epibenthos. Finally, a set of conclusions covering the range of topics is presented, together with a series of recommendations for future work.

P. W. Greig-Smith
Deputy Director
(Aquatic Environment Protection)
Ministry of Agriculture, Fisheries and Food
Directorate of Fisheries Research

SUMMARY AND RECOMMENDATIONS

The following report is a compilation of work carried out during 1990 and 1991 by members of a Benthos Task Team at the request of a UK Group Co-ordinating Sea Disposal Monitoring. It covers the following topics:

(i) an assessment of the utility of a range of measures for use in the derivation of 'Environmental Quality Standards' (EQSs);

(ii) an evaluation of the consequences for data interpretation of identification of benthic organisms to higher taxonomic levels;

(iii) conduct of a laboratory intercomparison exercise on the efficiency of analysis of benthos samples;

(iv) identification of a flexible taxonomic coding system for use in the archiving of benthos data;

(v) identification of future requirements for improved assessments of the benthic epifauna.

A summary of the outcome of this work is given below, while key topics for future research and development are also listed in Section 7.

(i) EQS derivation

The framework for this task was provided by two Environmental Quality Objectives defined by the above Co-ordinating Group which related to changes in the benthos, namely 'ecosystem maintenance', and 'preservation of the environment (outside the zone of immediate effect)'.

Regarding the setting of associated Environmental Quality Standards for assessing compliance with these Objectives, the first was taken to imply a judgement not only on whether a change had occurred in response to waste disposal well within its known sphere of influence, but also on whether any change was acceptable. The second was taken to imply a judgement on whether any change had occurred in response to waste disposal outside its known sphere of influence, generally without regard to standards of acceptability. This is because, in the latter case, relatively minor departures from the *status quo* will be of interest, even though these would normally be expected to occur well within the boundaries of acceptability defined in accordance with the first Objective.

The primary effect of sewage-sludge disposal at the seabed was considered to be that of **organic enrichment**. An empirical model describing the response of benthic communities along an enrichment gradient provided a means to evaluate the significance of observed changes in selected measures. Such changes (in space or time) are frequently non-linear; moreover, both positive and negative directional trends in some measures may be significant in the identification of effects.

The performance of these measures was assessed on a number of specimen data sets; measures were also scored subjectively for a variety of attributes by individual Task Team members. Finally, the effects on interpretation of data of pooling to higher taxonomic levels were considered. The main conclusions of this work were as follows :

(a) Measures suitable for routine use in EQS derivation included total abundance (A), total taxa (T) and total biomass (B) as ash-free dry weight, along with ratios of these measures (A/T, B/A) which had interpretative value. The Shannon-Wiener diversity index H'($\log_2$) and Multidimensional Scaling Ordination were also identified for complementary use, pending further statistical evaluation of their mode of application to compliance monitoring.

(b) Use of the selected variables in compliance monitoring can best be illustrated by a hypothetical example:

Station:	'Treatment'	'Reference 1'	'Reference 2'
Disposal → site	*	*	*

← Sphere of waste influence at the seabed →

Ideally, all three locations will be subject to the same natural environmental influences. 'Reference 1' is located just outside the known sphere of waste influence, while 'Reference 2' is some distance away. In reality, it may of course be very difficult either to establish the precise boundary for waste influence, or to select for identical environmental conditions at both reference stations. It will therefore be necessary to establish an initial ‘baseline’, derived from pairwise comparisons, against which the significance of any future differences between reference stations may be assessed, beyond those already apparent. (Pairwise comparisons have the potential advantage of allowing for the occurrence of natural synchronous changes in populations with time, assuming proportionality in the response).

(c) Thus guideline values (or ‘Action Points’) designed for the monitoring of compliance with the *status quo* at stations away from the zone of immediate effect, were as follows :

<u>Primary variables</u>
A : baseline value (%) ± 50
T : baseline value (%) ± 20
B : baseline value (%) ± 20

<u>Derived variables</u>
A/T : baseline value (%) ± 50
B/A : baseline value (%) ± 50

The baseline values are calculated from pairwise comparisons of annual measures, as follows:

[(Reference 1/Reference 2)-1] x 100, averaged over at least three years, preferably before discharge commences.

(d) Guideline values (or ‘Action Points’) to test for the <u>acceptability</u> of change at stations within the sphere of waste influence, relative to outside, were as follows :

<u>Primary variables</u>
A : + 200% of reference value
T : + 50% of reference value
B : + 50% of reference value

<u>Derived variables</u>
A/T : + 100% of reference value
B/A : - 50% of reference value

Compliance-testing will involve the same pairwise approach, namely:

[(Treatment/Reference 1 or 2)-1] x 100.

It will be clear that careful selection of stations will be critical to the success of this approach. For both (c) and (d) above, the primary variables are presently considered to be of first importance in compliance monitoring, and action of an appropriate nature would generally follow only when the guideline values of all three were exceeded. In addition, annually sampled stations may be marginally in breach of ‘Action Points’ for up to three successive years, to allow for any anomalous changes attributable to natural causes. This allowance would be overriden during the three-year period if ancillary analyses — which will notably include qualitative assessment of the biological data — provided conclusive evidence of adverse effects attributable to the discharge.

It must be emphasised that these ‘Action Points’ were derived from a consideration of well-defined gradients in <u>quiescent</u> areas, and make allowance for the generally positive responses of the selected variables in the early enrichment phase. Beyond this phase (at which point changes are deemed to be unacceptable), there may be negative changes in some variables, e.g. numbers of taxa. Prior knowledge of the status of benthic communities in relation to the enrichment process will therefore be essential for effective implementation of this approach.

Additionally, some uncertainties remain regarding the applicability of ‘Action Points’ to dispersive localities. They should therefore not be regarded as immutable, but should be set on a site-specific basis to match scientific assessments of the likely local scales of variability. It is also important to emphasise the role of complementary analyses using a variety of other measures of community structure, as an aid to better understanding of the causes of observed trends in the data.

Finally, **it is recommended that high priority be given to further evaluation of statistical methods for compliance testing, and of the related issue of effective sampling design.**

(ii) Identification at higher taxonomic levels

This topic was investigated by means of analysis of test data sets, and appraisals of recent findings in the literature. This is presently an issue of much contention, and concerns included the effects of loss of species-level information on historical comparisons with data at a locality, or from elsewhere; also, the consequences for a developing alliance with ecotoxicologists, whose tests are invariably conducted on named species, and implications for the future maintenance of taxonomic expertise. Positive aspects included the potential time-saving arising from identification to higher taxonomic levels which might, for example, permit the processing of larger numbers of samples from annual surveys.

Analyses of test data sets did not suggest that findings would be seriously compromised as a result of pooling, at least to the level of the family. However, a more rigorous statistical analysis across a wider range of data sets will be advisable.

It is therefore recommended that identification to the level of the family may be adopted for wide-scale descriptive surveys which are conducted in support of EQS programmes, with samples being retained in the eventuality of a future requirement for species-level identification. Identification to species level should continue at sites used for EQS derivation, to allow for further evaluation of any loss in sensitivity, and for any adjustments to be made to the above 'Action Points' which were derived from species-level data.

(iii) Laboratory intercomparison exercise

Twelve laboratories participated in the first exercise, which involved the analysis of a series of samples from a shallow sandy location, and then re-analysis of all samples by a single laboratory [Scottish Office Agriculture and Fisheries Department (SOAFD), Aberdeen]. (A second exercise involving analyses of samples from an offshore muddy sand station is presently in progress). The exercise allowed assessment of the efficiency of identification, enumeration and biomass determination.

There was wide variation in nomenclature, partly reflecting the lack of availability of a standard species directory across all laboratories. Differences could also be attributed to assignments of specimens to a less precise taxonomic level, and to mis-identifications. However, estimates of total numbers of taxa, individuals and diversity showed acceptable variation between laboratories.

Significant time-saving (of about 50%) could accrue when samples were identified to the family level only, though this would be appreciably less once investigators had gained familiarity with the fauna of an area.

Recommendations for future improvements included the conduct of regular laboratory intercomparison exercises, maintenance of in-house checks on sorting efficiency, adoption of a standard methodology for sample processing, circulation of taxonomic references, adoption of a standard species list, maintenance of a reference collection of specimens, establishment of contacts with taxonomic specialists to provide independant verification of specimens and attendance at taxonomic workshops.

It is essential to recognise the resource implications for co-ordinating and — to a lesser extent — participating laboratories, arising from the adoption of these measures. (It is worth noting that devotion of 20% of analytical effort to Analytical Quality Control (AQC) samples is becoming the accepted norm in Water Industry laboratories).

(iv) Taxonomic coding of benthos data

Ideally, a coding system should be flexible to allow for future additions or deletions of taxa and any changes in nomenclature, and hierarchical to permit the archiving and tabulation of named species (or higher groups) in taxonomic order. Several coding systems presently employed in UK laboratories were reviewed, but none appeared to be fully satisfactory.

It is recommended that the US National Oceanographic Data Centre (NODC) coding system be adopted for archiving of UK benthos data since, as well as being hierarchical and flexible, it is also internationally recognised.

It is further recommended that nomenclature for UK surveys follows that of the Species Directory published by the Marine Conservation Society (which is presently being up-dated). One proviso is that there is a need to allocate new NODC codes to several of the listed species, and this is currently the subject of negotiation between US personnel and SOAFD (Aberdeen).

(v) Future requirements for studies of the epibenthos

Quantitative studies of this important — and conspicuous — component of the benthos have traditionally been hindered by sampling problems, especially on coarse deposits. Currently available sampling methods, especially underwater photography, are reviewed and recommendations for future improvements include the need for trials of the comparative efficiency of different methods in carrying out community censuses across a range of substrates.

There is also a need for definitive studies of changes in epibenthic communities along known pollution gradients using standard sampling methods, with the objective of developing an empirical model to aid the interpretation of data from routine monitoring surveys.

1. INTRODUCTION

In 1988, a Benthos Task Team met, at the request of a UK Group Co-ordinating Sea Disposal Monitoring, to develop standard approaches for the sampling and analysis of benthic organisms at UK sewage-sludge disposal sites. The outcome of this work has been reported in MAFF (1989) and Rees *et al*. (1990). The latter account also provided an initial assessment of the feasibility of developing Environmental Quality Standards (EQSs) from the results of benthos surveys, and should be consulted as a background to the work on this topic which is reported below.

Arising from the above activities were a series of proposals for future work, which were considered by the Co-ordinating Group in 1989. As a result, a re-constituted Task Team was set up in 1990 (see Annex 1 for list of members), with the following Terms of Reference:

(i) to develop qualitative and quantitative standards against which the effects of sewage-sludge disposal on the benthos can be judged;

(ii) to assess the effects of pooling data at higher taxonomic levels on the precision and accuracy of measures of benthic community structure;

(iii) to conduct an intercalibration exercise concerned with the sorting and identifying of macrofauna from field surveys;

(iv) to agree on a common coding system for archiving of benthos data;

(v) to develop proposals aimed at improved assessment of the epifauna at sewage-sludge disposal grounds.

In the following report, Section 2 deals with the derivation of Environmental Quality Standards using a variety of numerical measures of community structure applied to sample data sets from the Garroch Head, Forth and Tyne sewage-sludge disposal sites. The performance of chosen measures on time-series data was examined by reference to samples taken in the Tees Bay area as part of a long-standing ICI monitoring programme. Measures were scored against a range of attributes relevant to practical implementation, as an aid to the final selection.

Section 2 also addresses the effects of pooling of benthic data to higher taxonomic levels on the outcome of analyses of test data. This approach has been the subject of much recent debate, and an outline of some of the perceived advantages and limitations arising from its adoption in regular monitoring programmes is givenat Annex 2.

Section 3 reviews the outcome of a laboratory intercomparison exercise and makes proposals for follow-up work. Section 4 is concerned with procedures for taxonomic coding of archived data on marine benthos around the UK coastline, with particular reference to sewage-sludge disposal sites. Section 5 considers the problems associated with sampling of the epifauna at marine disposal sites, especially in areas of hard ground, and makes recommendations for future work aimed at improving assessments of this important component of the benthic biota.

Finally, in Sections 6 and 7 conclusions are drawn from the range of issues addressed by the Task Team, with particular reference to the outcome of EQS derivation, and a series of proposals for future Research and Development are presented.

2. THE DERIVATION OF EQSs FOR THE BENTHOS AT UK SEWAGE-SLUDGE DISPOSAL SITES

2.1 Rationale for development of EQSs

Regarding benthos studies, the relevant Environmental Quality Objectives (EQOs), and associated Standards for the measurement of compliance, are given in MAFF (1989) as follows:

EQO	Criterion	Basis of standard
1. Ecosystem maintenance	Benthic fauna	Deviation from the reference site(s) to be within acceptable limits
2. Preservation of the environment (Outside zone of immediate effect)	Benthic fauna	No deviation from reference site(s)

The requirements of an EQS approach pertaining to the benthos can be summarised as follows:

(i) Within the sphere of influence of sludge disposal, a measure of biological 'quality' must not only be sensitive to anthropogenically-induced change, but also provide a means to assess the **acceptability** of any observed change.

(ii) Outside the sphere of influence, such a measure must be sensitive to anthropogenically-induced change relative to a distant reference site, but this would generally be without regard to standards of acceptability, since relatively minor deviations from the *status quo* will be of interest.

It follows from the above that a good knowledge of the spatial extent of the sphere of influence of sewage sludge at a locality is required, along with an understanding of the likely consequences for benthic communities, in order to design an effective programme for compliance monitoring.

Clearly, variations in environmental conditions between locations — e.g. exposure to wave and tidal action, other anthropogenic influences — will dictate the sampling effort required to meet the above objectives and, in some cases, many sampling stations will be necessary. However, for illustrative purposes, it is helpful to consider an idealised system in which three stations (or groups of stations) are located as in Figures 1(a) and (b).

In Figure 1(a), the 'Treatment' station falls within the sphere of influence of dispersing sludge particles at the seabed. 'Reference' Stations 1 and 2 are not affected by sludge disposal: the former is located just outside the sphere of influence, and the latter at an appreciable distance. Ideally, the three stations will be identical in terms of natural environmental influences. Pairwise comparisons between the 'Treatment' station and 'Reference' Stations 1 and 2 will be concerned both with the identification of change and a judgement as to its acceptability at the former, relative to the latter. The comparison of 'Reference' Stations 1 and 2 will generally require only the measurement of change at one relative to the other, since departures from normality would be expected to occur well within the boundaries of acceptability defined in accordance with Objective 1 above. This pairwise approach to quality assessment also has the potential advantage of allowing for the occurrence of natural synchronous changes which may occur at all sites from year to year (see Sub-Section 2.2.2).

Finally, it should be noted that in areas of low dispersive capacity, it may be necessary to define a 'mixing zone' within which exceedance of an EQS may be permissible. However, in principle, there should be no such requirement at a highly dispersive site.

2.1.1 A description of changes in the benthos in response to sewage-sludge disposal

There are well-documented responses to organic enrichment (e.g. Pearson and Rosenberg, 1978) which permit some generalisations to be made. The following account provides the basis for later recommendations concerning definitions of acceptability of changes in the benthos.

Basis for recognition:
Local alterations in the benthos due to effects of sludge-derived material may take place:

(i) as a result of its motion over the seabed;

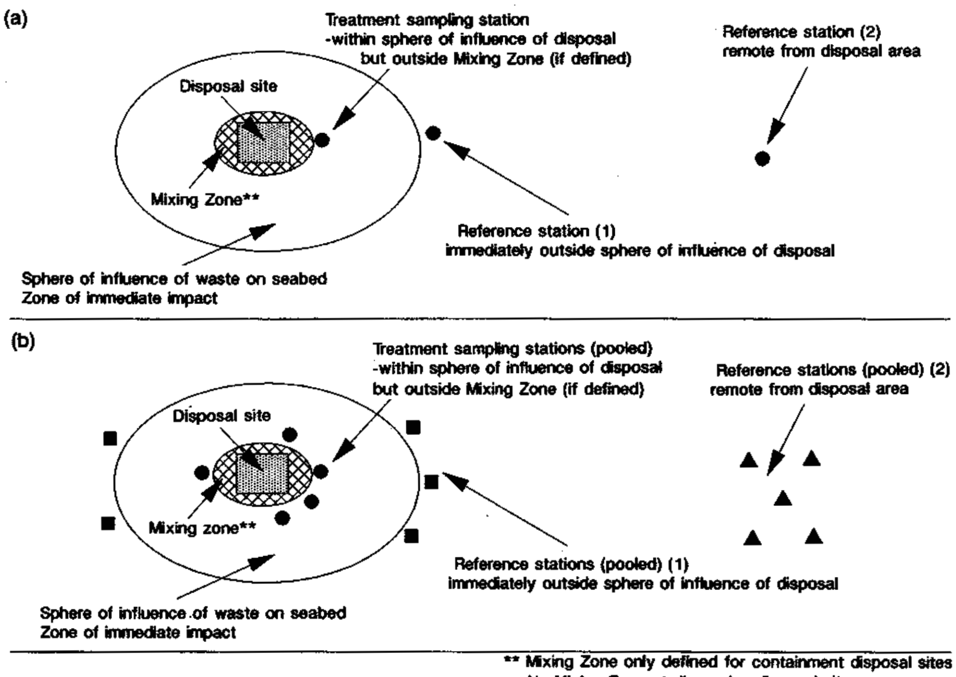


***Figure 1.** Idealised system showing stations for EQS monitoring of sewage-sludge disposal: (a) single stations (with replicates) and (b) pooled stations (with or without replicates)*

(ii) in quiescent areas which promote settling of particulates (these may be remote from the disposal site);

(iii) intermittently as a result of transient accumulations.

These processes may occur at all types of site but the relative importance may vary.

UK experience indicates that the primary effect of sewage-sludge disposal on the benthos is likely to be a consequence of **organic enrichment.** The biological effects of potentially toxic contaminants associated with sewage sludge (other than those generated as by-products of the enrichment process), e.g. certain trace metals, are considered to be secondary in importance at the concentrations generally encountered within sediments.

Nevertheless, the benthos may still have a significant role to play in **bio-accumulation** and then transfer of contaminants to commercial fish via the food chain, even though the concentrations in organisms may be insufficient to induce adverse biological effects.

The structure of a 'normal' benthic community will be determined locally by the prevailing physical and chemical regimes, and hence — in accordance with current ecological theory — may occupy **any point** along a continuum from:

(i) a complex 'climax' community, typically including a variety of long-lived and large-sized organisms distributed at depth within sediments; to

(ii) a simple 'successional' community, with the majority of typically short-lived and small-sized organisms concentrated in the surface sediment layers.

Empirical models suggest that during the progress of organic enrichment, changes in a benthic community may occur in **both directions** along this continuum (see Figure 2). The **extent** of change in community structure, if any, will depend upon its natural successional state, and the contaminant loading.

An important implication of the above model is that changes in the benthos due to organic enrichment may in some respects mirror those that occur naturally in response to physical perturbations at the seabed, for example, those arising from periodically intense wave action. This further highlights the importance of knowledge of the likely scales of natural variability in a waste-receiving area.

It should be noted that while 'maintenance of the benthic community' is in itself an Environmental Quality Objective, studies which are conducted in support of

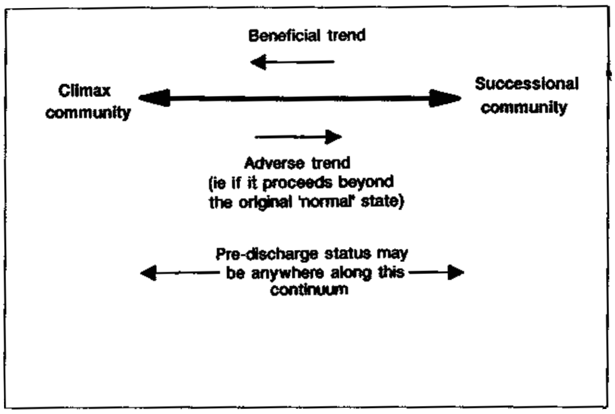


Figure 2. A generic model for change in benthic communities in response to sewage-sludge disposal

this Objective are also useful as indicators of the health of the environment generally.

Examples to illustrate the impact of sewage-sludge disposal on the benthos which - from the outset - occupy extremes of the successional scale are:

(a) Garroch Head

This disposal site is located in a quiescent muddy area off the Scottish west coast, where the 'normal' fauna approaches a climax community characterised by: high diversity, low % dominance and the occurrence of long-lived, deep-burrowing species — such as the heart urchin *Brissopsis* — which are also characteristic of aerobic sediments. (See, for example, Pearson, 1987).

Spatial changes in response to sewage-sludge disposal include an initial <u>increase</u> in biomass and numbers ('enrichment'), followed by a decline. As the centre of the disposal site is approached, and sediments become increasingly anoxic with depth, the trend is towards high % dominance of one or two species at the expense of others ('over-enrichment'). A typical dominant is the deposit-feeding polychaete *Capitella,* occasional specimens of which may penetrate to considerable depths (about 50 cm) in anoxic sediments.

(b) Barrow Deep

This disposal site is located in a 'high energy' area of the outer Thames estuary, south east England and is characterised by well-sorted sand or muddy sand sediments occupying channel systems, especially in an inshore direction. The 'normal' fauna is naturally maintained at an early successional state due to a combination of cyclical and unpredictable disturbances at the seabed (tide and wave action). It is typified by a reduced range of species, high % dominance and erratic fluctuations in near-surface deposit-feeding worms such as *Spiophanes* and *Pectinaria*. (See, for example, Norton *et al.*, 1981; Talbot *et al.*, 1982).

Such effects as have been observed in the area include increased numbers and variety of species at more quiescent locations. Strictly, these effects could be interpreted as **beneficial** since — in accordance with the above model — they represent a shift in the direction of a climax community. Elsewhere, apparently toxic effects were identified at a few stations within the area of sludge impact. Here, the benthos is characterised by a reduction both in the numbers and range of species found.

2.1.2 Measurement of change

The starting point for this assessment was a suggestion of the Co-ordinating Group that consideration be given to the feasibility of detecting a 20% change in selected measures of community structure. This will clearly depend on the ability to achieve acceptable levels of precision for the mean values of each measure.

A typical example of the number of sampling units required in order to achieve specified levels of precision with respect to sample counts is given in Table 1.2 of McIntyre *et al.* (1984). In general, the smaller the mean sample size, the greater the number of sampling units required; it follows that the required sampling effort at sewage-sludge disposal sites is likely to vary both between stations and between sites.

In practice, a limit as restrictive as 20% will be unrealistic for measures such as total abundance, since prohibitively large numbers of samples will usually be required. Further, changes of such magnitude in naturally variable systems may be perceived to have less significance, in terms of pollution responses, for some measures than for others.

As natural changes in sample densities over time are to be expected, it will also be necessary to re-appraise sampling effort on a frequent basis in order to maintain specified levels of precision of mean counts for the purpose of compliance monitoring. Other univariate measures of community structure will require different levels of sampling effort and, again, these will need to be determined on a site-specific basis. These aspects are further addressed in Sub-Section 2.3.

2.1.3 Choice of measures

Initially, a decision was required on which of a range of measures of community structure were likely to be the most suitable for EQS derivation. To aid this judgement, we have analysed three spatial data sets:

(i) Garroch Head: Scottish west coast (1983). Survey design took the form of a transect of stations through the disposal site; one sample was taken from each station using a Van Veen grab, and the macrofauna extracted over a 1 mm mesh sieve. The data show a strong gradient of biological effect in response to sewage-sludge disposal (see, for example, Pearson, 1987).

(ii) Tyne: English east coast (1986). A transect of stations was sampled southwards from the centre of the disposal site; three replicates were taken at each with a Day grab and a 0.5 mm mesh sieve was used to extract the macrofauna. Changes in the benthos and other parallel measures provided circumstantial evidence of a mild enrichment effect (see Rees *et al.*, 1992).

(iii) St Abbs Head: Scottish east coast (1987 and 1988). Survey design took the form of a circular grid of stations. Two samples were taken at each using a Van Veen grab, and the macrofauna extracted using a 0.5 mm mesh sieve. There was a degree of substratum variability between stations, but no evidence for a strong gradient in the benthos data which could be attributed to the effects of sewage-sludge disposal (see, for example, Heap *et al.*, 1991).

Thus the three data sets provide a contrast in terms of sampling design, sampling methods and biological effects at the seabed.

In addition, a time-series data set was selected, in order to examine the sensitivity of measures of data structure to year-on-year changes in benthic populations. The data were from two sandy stations, approximately 500 metres apart and at about 20 m depth, sampled annually off Tees Bay, north east England, between 1974 and 1987. The macrofauna from Smith-McIntyre/Day grab samples were extracted using a 1 mm mesh sieve (see, for example, Shillabeer and Tapp, 1990). In this case, there was no expectation that any directional trends might be attributable to pollution.

Changes in the benthos are frequently non-linear along pollution gradients (e.g. Pearson and Rosenberg, 1978), and this must be allowed for in seeking to quantify the status of communities, both in space and time. (The use of appropriate transformations of data may be helpful in this respect).

In the case of Garroch Head, most traditional descriptors of community structure have been shown to reflect the strong gradient. Indeed, work at this site has the advantage of providing a 'model' for the progress of enrichment effects arising from sewage-sludge disposal, at least in quiescent areas. The identification of effects, if any, at some of the more dispersive sites encountered elsewhere presents more of a challenge. In order to assist judgement on the suitability of different measures of data structure applied to the above data sets, the following criteria were identified as important:

(i) sensitivity;
(ii) amenability to statistical testing;
(iii) effect of pooling (on sensitivity/precision);
(iv) ease of use;
(v) ease of interpretation; and
(vi) range of applicability (NB dispersive versus accumulating sites).

These criteria were employed as an aid to the final recommendations given in Sub-Section 2.3.

Measures selected for further investigation of their performance when applied to the above data sets included:

Multivariate techniques: non-metric Multidimensional Scaling Ordination (MDS).

Univariate and graphical techniques: Species/Abundance/Biomass (SAB) Comparison; Abundance/Biomass Comparison (ABC); Shannon-Wiener diversity ($H'log_2$); Pielou Evenness (E); Coefficient of Pollution.

Full details of these techniques are given in Sub-Section 2.2 below.

It may be noted that 'k-dominance' and 'rarefaction' curves (see Lambshead *et al*., 1983 and Sanders, 1968, respectively) were considered as possible means for EQS derivation in Rees *et al*. (1990). The production of numerical values would require transformation of curves along the lines of that carried out by Clarke (1990). However, the end-products would, in effect, represent new diversity indices with little perceived benefit over H', as all are functions of the apportioning of individuals among species in the community. This line of investigation was therefore not pursued further.

An assessment by the Task Team of the potential for derivation of a Biotic Index for UK benthos led to the initiation of a research project, jointly funded by the Scottish and Northern Ireland Forum for Environmental Research and the National Rivers Authority. Its utility in an EQS context must await the outcome of work presently being conducted by the Water Research Centre.

Finally, we have considered the effects of pooling data to higher taxonomic levels on the ability of different measures to detect changes. This approach — which has received much recent attention in the literature — has the potential advantage of allowing increased numbers of samples to be processed for the same resource commitment, a requirement which is almost certain to arise in the implementation of an EQS approach. A review of some of the implications arising from the adoption of such an approach in routine surveys of the benthos is given at Annex 2.

2.2 Evaluation of measures

2.2.1 Spatial changes

Multidimensional scaling (MDS) ***(R. Warwick)***

A description of this multivariate analytical technique is given at Annex 3.

(i) Garroch Head Disposal Site
A map of sampling stations along the east-west transect is given in Figure 3. The two-dimensional

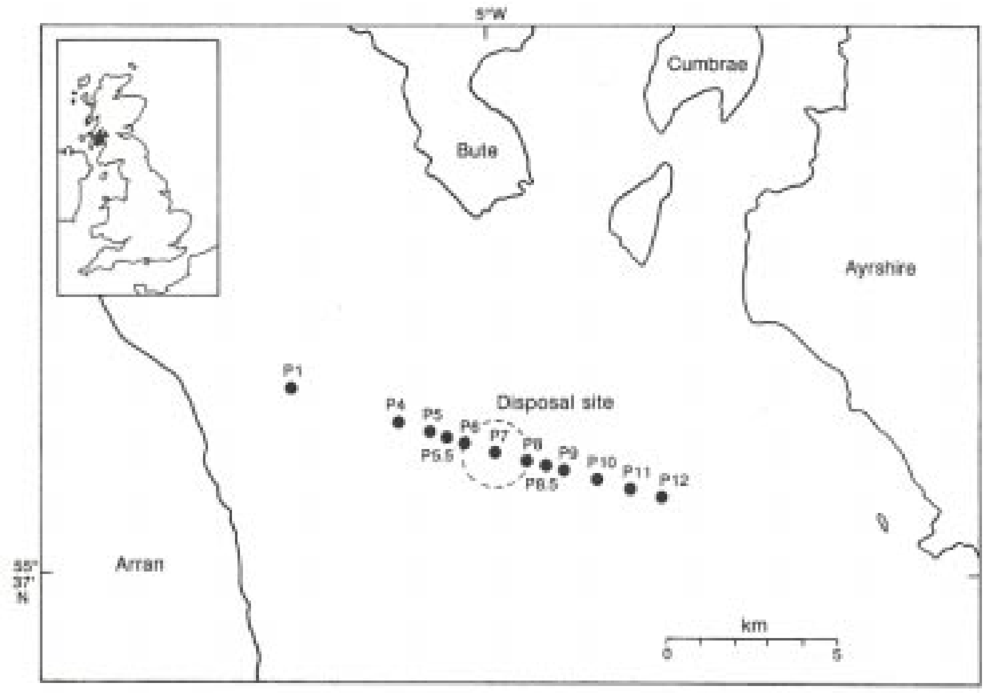


Figure 3. Garroch Head disposal site: station positions

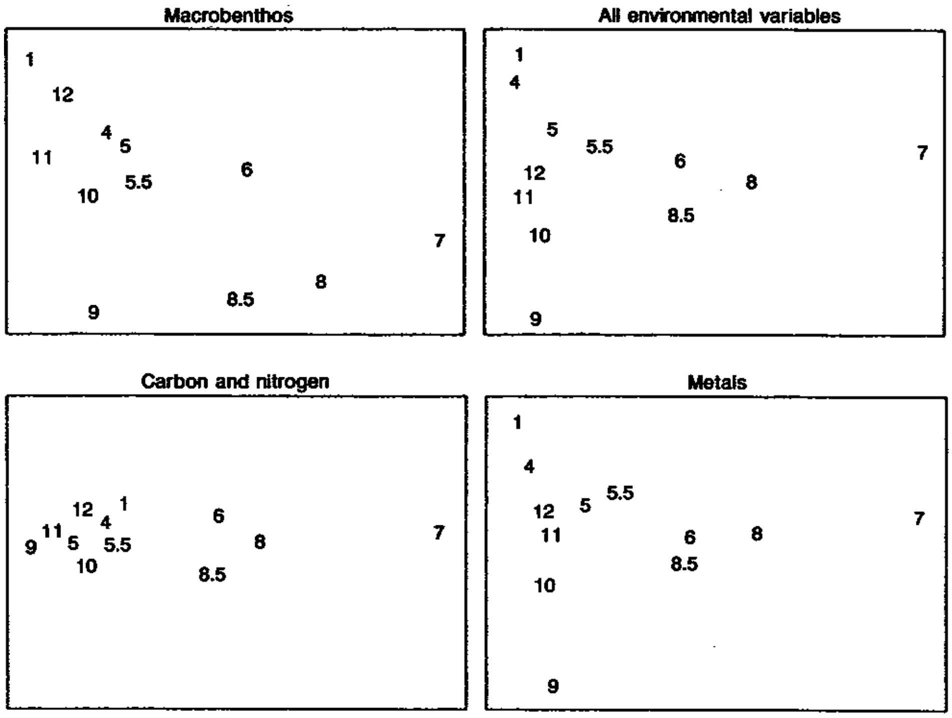


Figure 4. ***Garroch Head disposal site: macrobenthos MDS compared with PCA for different combinations of environmental variables***

MDS configuration for the root-root transformed species *biomass* data is given in Figure 4 (top left). There is a clear gradation of community change from Station 1 at the top left of the configuration to Station 7 (the disposal site centre) at the right, and then back again to Station 12 which is close (i.e. similar in community structure) to Station 1.

A correlation-based Principal Components Analysis (PCA) on the measured environmental variables (carbon, nitrogen and 8 metals) in the sediment closely resembles the structure of the faunistic MDS, suggesting that these are the variables that 'explain' the changes in community structure across the disposal site. The PCA for C and N alone is very one dimensional, whereas the PCA for metals (Cu, Mn, Co, Ni, Zn, Cd, Pb, Cr) more closely reproduces the three dimensional structure of the faunistic MDS. This suggests that, although organic enrichment has been considered the main determinant of benthic community structure at such sites, the toxic effects of heavy metals may also be of importance. (However, it should be noted that many of the metals will be highly correlated with organic carbon and nitrogen values).

All environmental variables along this clear contamination gradient are inter-correlated, and individually all show a close visual correlation with the faunistic MDS. In Figure 5, for example, both lead and carbon have high values to the right and low values to the left of the configuration.

Figure 6 shows the MDS configuration for the species *abundance* data. Note the close similarity to the *biomass* MDS in Figure 4 (the orientation of such

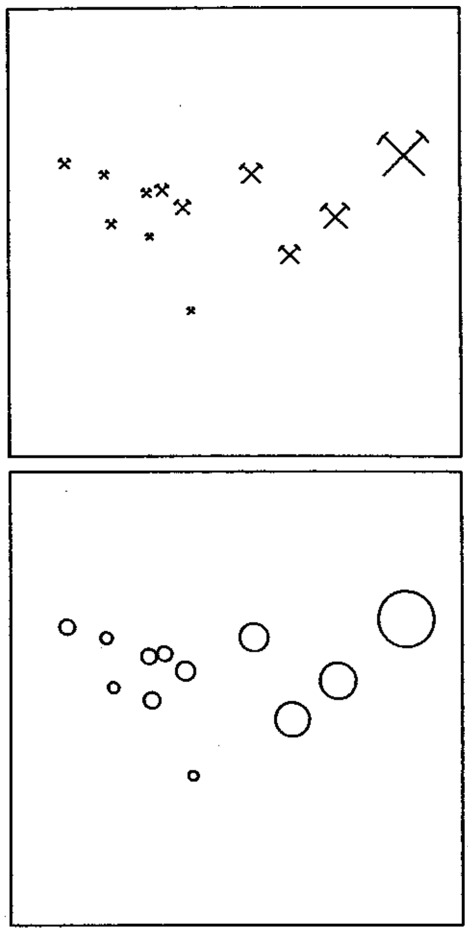


Figure 5. ***Garroch Head disposal site: MDS for macrobenthos species biomass (from Figure 4) with symbols superimposed which are scaled in size according to measured values of lead (pick-axes, top) and carbon (circles, bottom) in the sediment***

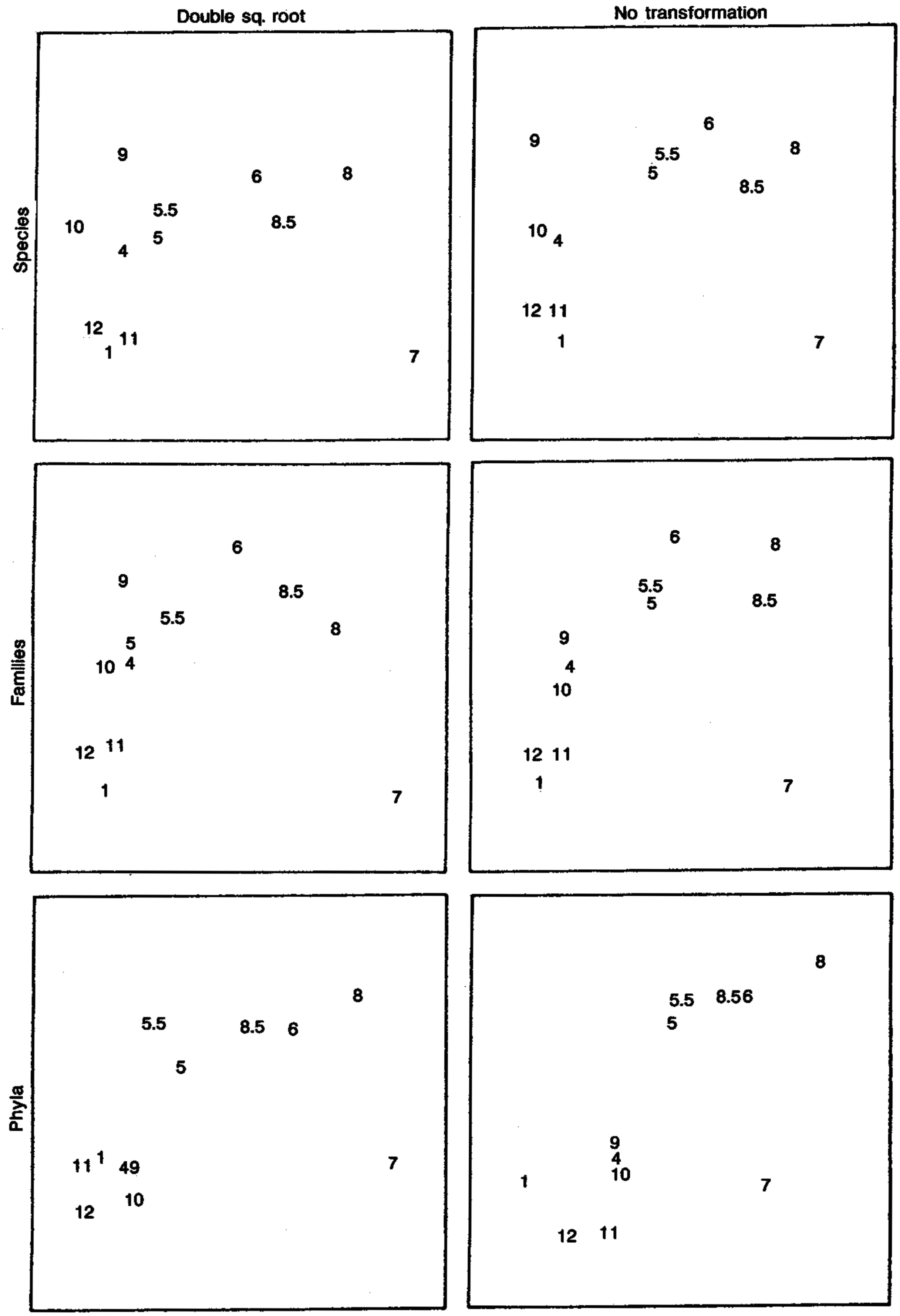


Figure 6. Garroch Head disposal site: MDS ordinations for abundance data at the level of species (top), families (middle) and phyla (bottom)

configurations is abritrary, and the abundance configuration is reflected on the x-axis relative to the biomass configuration). This analysis has been repeated with the species data aggregated into families and also to phyla. Note the close similarity of the configurations at all levels of taxonomic aggregation.

(ii) Tyne Disposal Site

A map of the transect of stations sampled in June 1986 is given in Figure 7. The MDS configuration for the root-transformed species abundance data for individual replicate samples at each station is given in Figure 8(a). Few environmental variables were determined on these individual replicates, but superimposing tomato pip counts onto the MDS configuration (Figure 8(b)) shows a clear correlation, with the most 'polluted' sites to the right and the less polluted to the left.

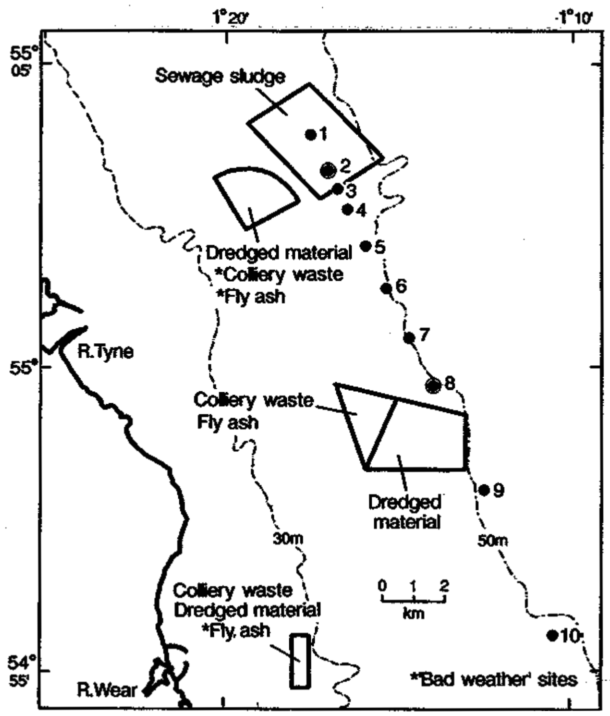


Figure 7. Tyne disposal site: station positions

Non-pollution related variables, e.g. sediment grain size (Figure 8(c)), show no clear relation to the MDS configuration. Thus, community change along the transect is shown to relate to pollution levels. The MDS analyses for the 89 FAMILIES (Figure 8(d)) and 10 PHYLA (Figure 8(e)) show essentially the same configuration of replicates, with the polluted sites 1-4 to the right and the remainder to the left. Superimposing environmental variables onto these configurations therefore gives essentially the same pattern as Figure 8(b), so that the relationship between community change and pollution levels is evident at all levels of taxonomic aggregation, including the phylum level.

Since there is only one determination of most environmental variables for each site, PCA analyses for environmental variables are compared with a faunistic MDS based on the totals for all replicate grab samples taken at each station (Figure 9). The MDS for species, families and phyla (Figures 9(a-c)) show similar patterns, with a more or less horse-shoe shaped arrangement of stations from 3 and 4 on the right to 9 and 10 on the left.

The PCA for all environmental variables shows a similar pattern to the faunistic MDS (Figure 9(d)). Pollution-related environmental variables are mainly responsible for this, as can be seen from Figure 9(e), which is the PCA for these variables (tomato pips, %C, C:N and 6 metals), and Figure 9(f) which is the PCA for non-pollution variables (depth, median grain size, % silt/clay). The former closely resembles the faunistic MDS whereas the latter does not.

(iii) St Abbs Head Disposal Site

The pattern of sampling stations is given in Figure 10(a); the disposal site is at the centre, i.e. Station 13 (see also Figure 20). MDS has been performed on the species-abundance data for June 1988 (the test data) and also for the species-abundance and biomass data from the check monitoring survey of 1987 (Figures 10(b-d). In each case the configuration of sites on the MDS approximately resembles the geographical arrangement of sites in Figure 10(a).

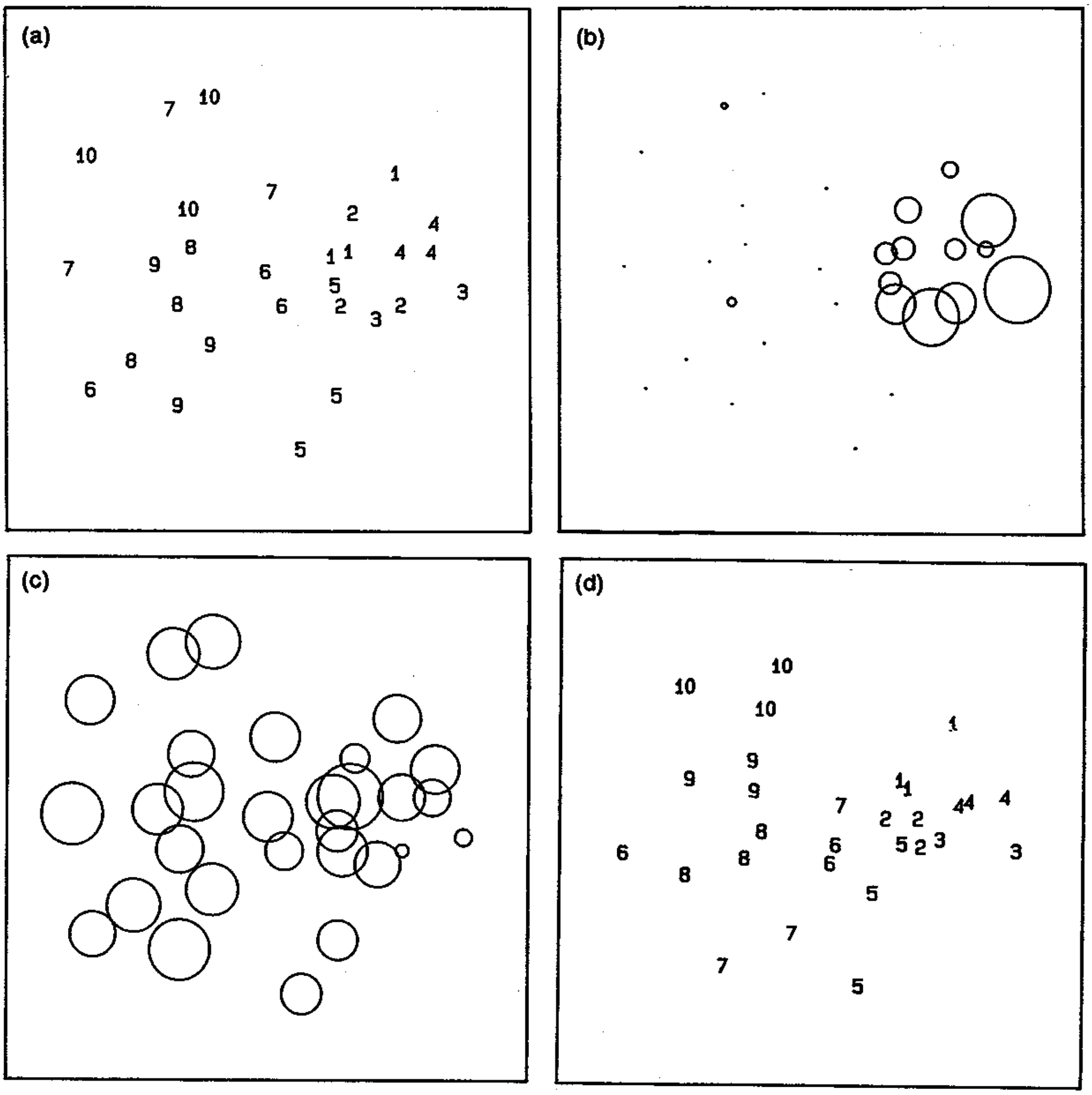
PCA analyses for environmental variables in different combinations (Figure 11) show no such relationship, with geographically distant sites being close together on the PCA, and vice versa. We must therefore conclude that the faunistic pattern is simply due to natural geographical variability across the disposal site, but is not influenced by the disposal activity itself.

In view of the fact that no association between faunistic pattern and pollution could be demonstrated at the species level, the analysis was not repeated at higher taxonomic levels.

General conclusions:

MDS analyses, coupled with methods of relating the configurations to measured values of environmental variables, showed a clear relationship between community structure and pollution levels at the Garroch Head and Tyne disposal sites. This was evidenced at all levels of aggregation of the species data, even to phyla. No community change which could be related to pollution effects was detected at the St Abbs Head ground.

It is important to distinguish between the ability of a technique to identify statistically significant responses to pollution, and the determination of the acceptability of a change or difference which must be based on its magnitude. The ability of a technique to identify small significant changes is nevertheless an important factor in assessing whether the method is sufficiently sensitive for the required EQS application.



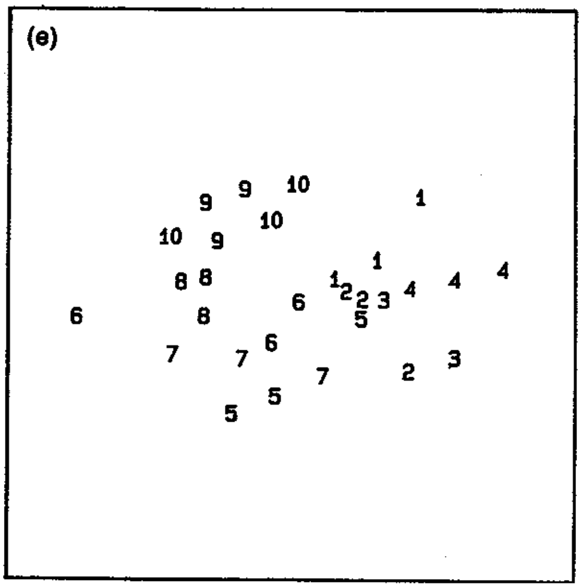


Figure 8. Tyne disposal site: (a) MDS of root-transformed species-abundance data for individual replicates at each station; (b), as (a) but station numbers replaced by circles scaled in size according to counts of tomato pips; (c), as (a) but station numbers replaced by circles scaled in size according to the median particle diameter of the sediment; (d), as (a) with species data aggregated into families and (e), aggregated to phyla

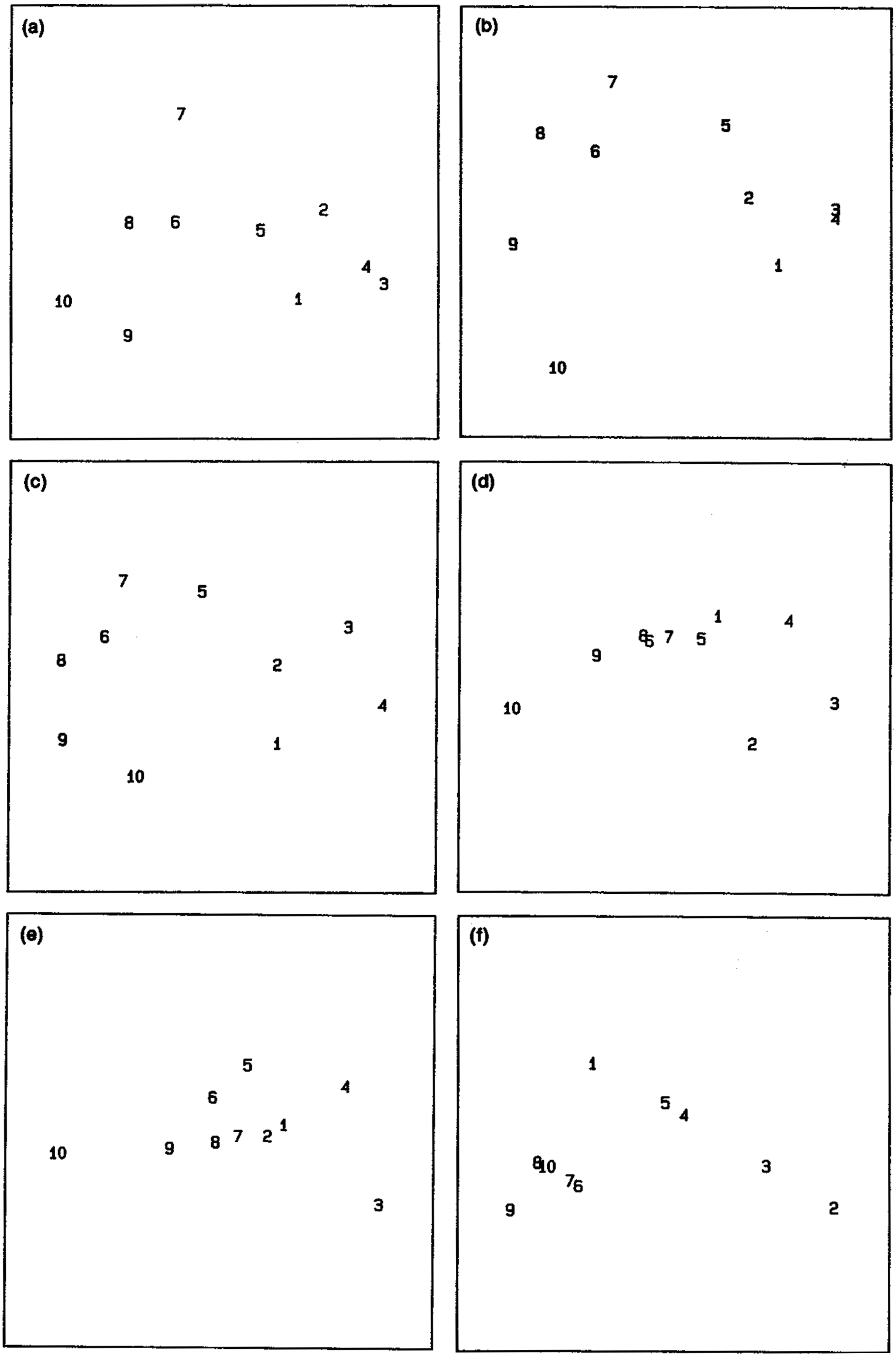


Figure 9. Tyne disposal site: (a-c), MDS for root-transformed species (a), family (b) and phylum (c) abundance data for totals of all replicates at each station; (d-f), PCA for all environmental variables (d), pollution-related variables (e) and natural environmental variables (f). See text for details

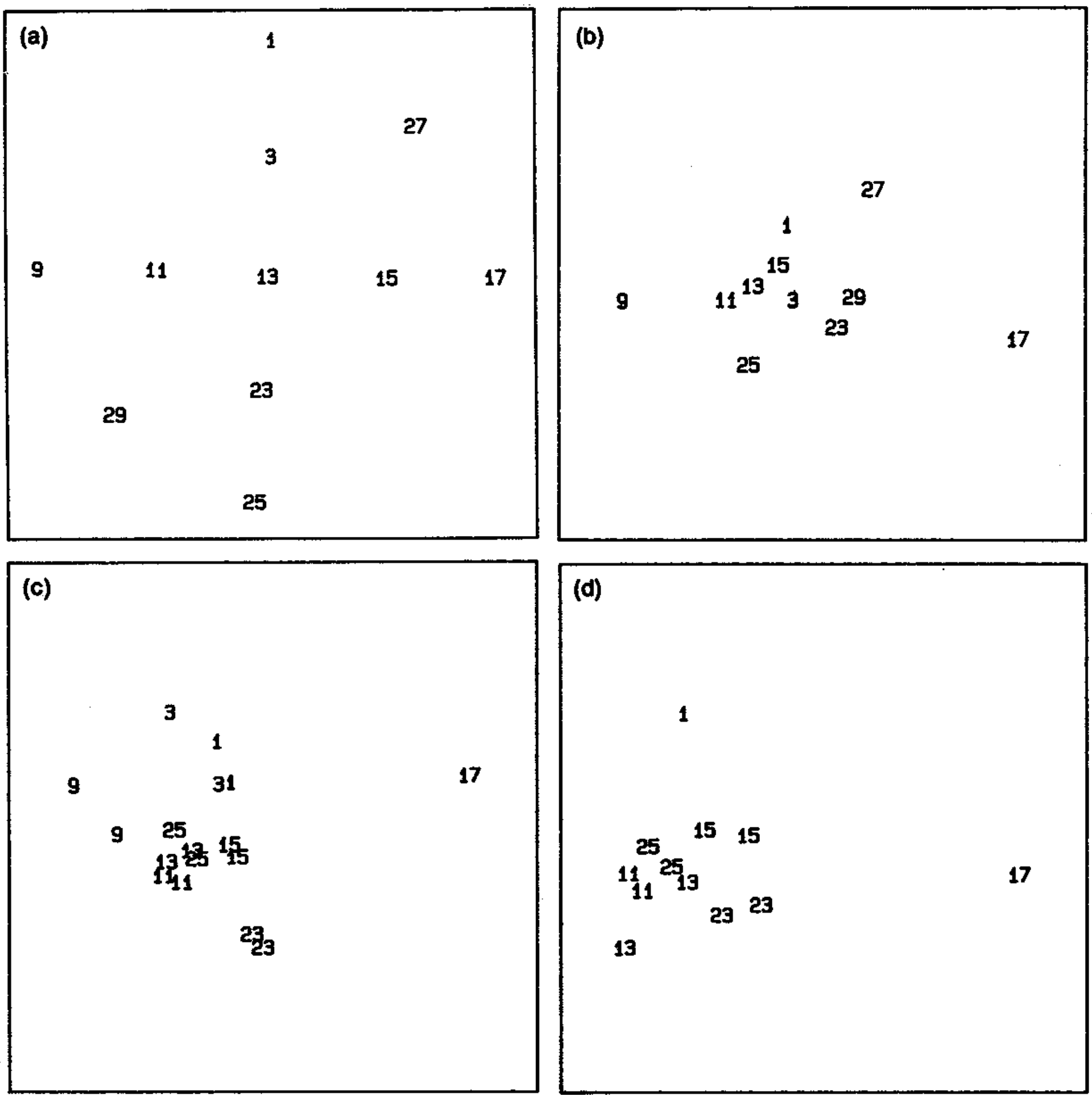


Figure 10. *St Abbs Head disposal site: (a), geographical pattern of sampling stations; (b-d), MDS ordinations of macrobenthos — (b), 1988 abundance; (c) 1987 abundance; (d) 1987 biomass*

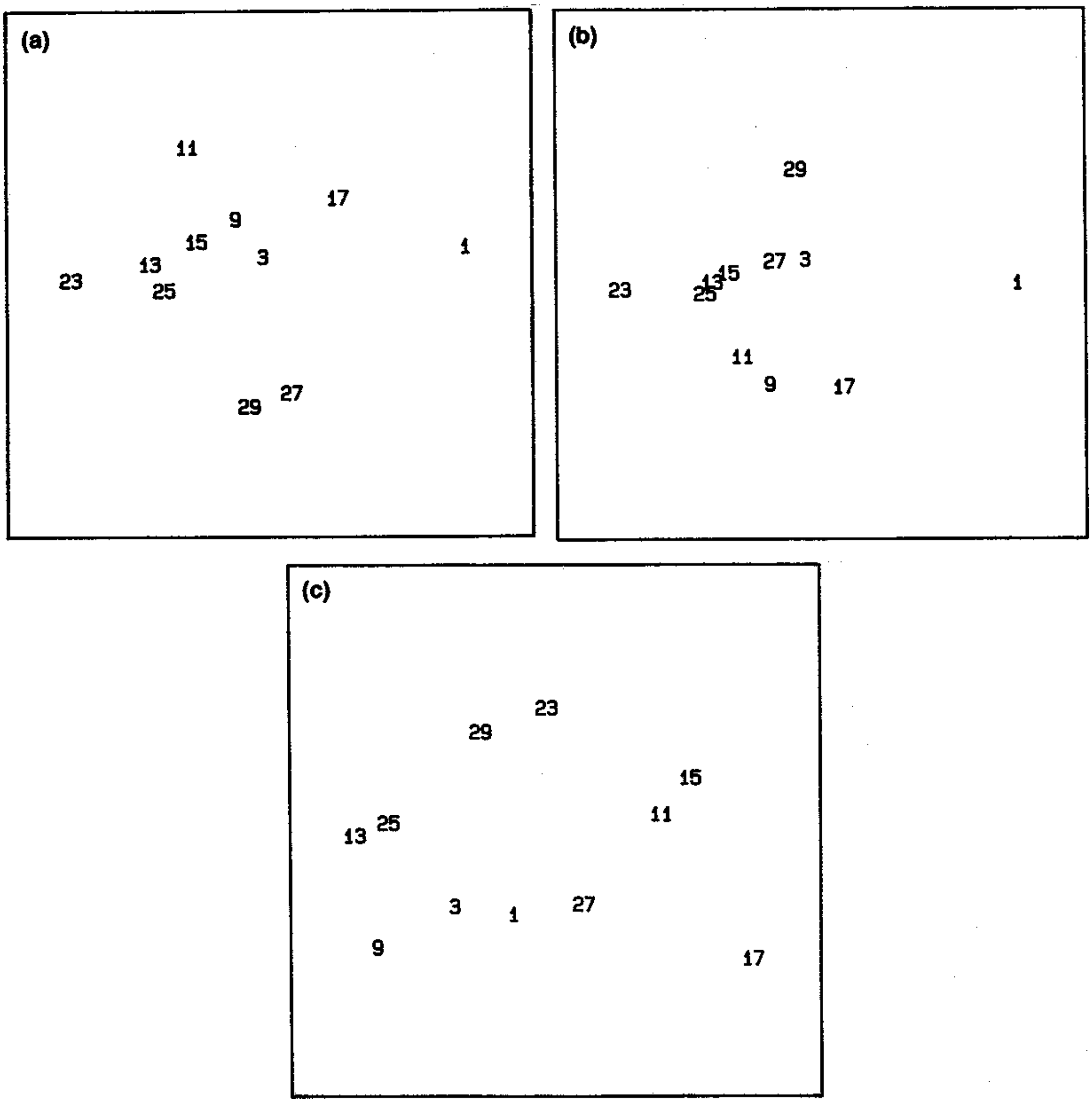


Figure 11. ***St Abbs Head disposal site: PCA for different combinations of environmental variables. (a) all variables; (b) metals in < 63 µm fraction of sediment only (Cu, Zn, Cd, Pb, Cr, Ni, Fe, As, Hg); (c) % silt/clay, median particle diameter of sediment, depth***

Species/Abundance/Biomass (SAB) comparison
(T. Pearson)

The following assessment is divided into two parts. The first part deals with the three spatial data sets (i.e. Garroch Head, St Abbs Head and Tyne) in their original form, that is, with the taxa identified at a high level of taxonomic precision. It describes changes in the selected measures, and then provides a means to judge the acceptability of these changes. The second part assesses the consequences of pooling data to a variety of levels in the taxonomic hierarchy, by reference to changes in the ratio of abundance and taxon number.

A. Analyses of raw data

(i) Garroch Head data, 1983

Measurement of change:
The data relate to a series of stations positioned along an 11 km transect running from east to west across the centre of the site, as shown in Table 1. The changing relationship of the curves representing the total number of species, S, the total abundance of organisms, A, and the total wet weight biomass, B, as distance from the centre of the site increases is diagnostic of the changing response of benthic communities to a decreasing organic gradient.

Table 1. Garroch Head survey faunal statistics, 1983. (Values 0.1 m^{-2})

Station Number	Distance From Disposal site centre (km)	S	% Deviation	A	% Deviation	B	% Deviation	A/S	% Deviation	B/A (x1000)	% Deviation
XP1/ P12	6,5	18	...	57	...	6.5	...	3.2	...	114	...
P4	3	28	+56	183	+221	12.1	+86	6.5	+103	66	-42
P5	2	45	+150	1045	+1733	20.3	+212	23.2	+625	19	-83
P5.5	1.5	42	+133	1459	+2460	13.2	+103	34.7	+984	9	-92
P6	1	27	+50	2872	+4939	29.0	+346	106.4	+3225	10	-91
P7	0	3	-83	330	+479	0.8	-88	110.0	+3337	2	-98
P8	1	11	-39	8327	+14509	26.6	+309	757.0	+23556	3	-97
P8.5	1.5	23	+28	2249	+3846	13.2	+103	97.8	+2956	6	-95
P9	2	29	+61	205	+260	5.2	-20	7.1	+121	25	-78
P10	3	33	+83	165	+189	7.0	+8	5.0	+56	42	-63
P11	4	21	+17	70	+23	3.8	-42	3.3	+3	54	-53

Figures 12(a) and (b) demonstrate the changes observed along the western and eastern arms of the transect, in each case starting from the centre Station P7. At this station, all three statistics are depressed, indicating a highly stressed community with only a few individuals of a very few species surviving. At 1 km from the centre (P6, P8) species numbers had recovered, although were still somewhat low on the eastern side (P8) where, however, the abundance and biomass had risen to extremely high levels.

On the western transect at 1 km from the centre, the abundance and biomass also reached their highest levels. These levels declined precipitously beyond 1 km, whereas species numbers continued to rise until 2 km from the centre on the western transect and 3 km on the eastern arm. There was a secondary, smaller peak in biomass at 2 km from the centre on the western arm. On both ends of the transect the statistics had declined to around background levels at distances beyond 5 km from the centre.

These changes adhere very closely to the idealised model of changing population statistics along an organic gradient proposed by Pearson and Rosenberg (1978) which was based on information drawn from a wide range of studies. Thus the changing values demonstrated across the Garroch Head site may be used as a comparator for changes observed in other areas where the organic gradient is not as obvious.

The changes observed in the population statistics can best be summarised and given some easily appreciated biological meaning by integrating them into the ratios A/S, the abundance ratio, representing the mean number of organisms per taxon in each sample, and B/A, the size ratio, being the mean weight of an individual organism in the sample (Pearson *et al.*, 1982). The changing values of these ratios are also shown in Figures 12(a) and (b).

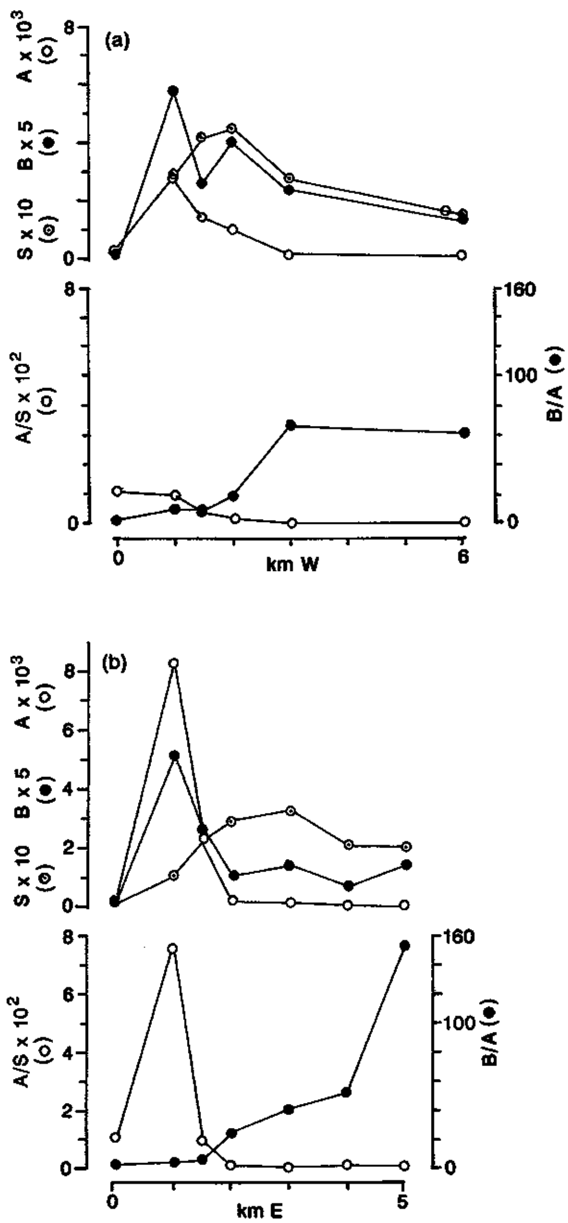


Figure 12. Changes in the benthos along a transect through the Garroch Head disposal site (see text)

It is obvious from reference to Figures 12(a) and (b) that major changes in the benthic communities were taking place at 3 km from the centre of the Garroch Head site. At that distance from the centre the average weight of organisms declined rapidly and the number of individuals per taxon began to rise steeply, indicating a switch in the communities from a relatively stable, diverse system to a system dominated by a lower number of smaller, opportunistic species.

A comparison of the values of the faunal statistics at the stations situated at this distance with those from the stations at either end of the transect representing background levels, will provide some numerical indication of this level of change in the communities. Table 1 lists the SAB statistics and the A/S and B/A ratios for each station along the transect together with their percentage deviation from the mean values of those statistics at the two stations at either end of the transect.

Thus at the stations 3 km from the centre the species numbers were 50-80% higher, the abundances 200% higher, the biomass 10-90% higher, the abundance ratio A/S 60-100% higher and the size ratio B/A 40-60% lower than the mean reference values.

Acceptability of change:
These values might suggest that faunal variability in an area resulting in species numbers increasing by 50%, abundances by 200% and biomass by 50% would represent an unacceptable level of change. Similarly, a change in the value of the abundance ratio of more than 100%, and in the size ratio of more than 50% might be regarded as unacceptable.

Support for these levels is provided by more recent data. Tables 2(a-f) list the data sets for 1985-88 and 1990-91, showing the S, A and A/S statistics and their % deviation from the values recorded at a Reference Station G1, some 10 km north-west of the disposal site centre. The later data sets only include stations up to 3 km from the centre of the site, all of which are affected by enrichment. They all demonstrate that the levels of acceptable change chosen on the basis of the 1983 data set are equally relevant for these data sets, i.e. changes in S greater than 50%, changes in A greater than 200% and changes in A/S greater than 100%, taking place at about 3 km from the centre of the site, represent a switch to lower diversities and higher abundances which may be considered as unacceptable.

Note that this degree of change represents the point in community succession when a transition takes place from an enriched but still highly diverse fauna to an abundant but less diverse community. It does not represent the point when changes are first discernible. This choice of cut-off-point is subjective, but defensible on ecological grounds by detailed reference to what is known about enrichment successions and subsequent faunal recovery times.

(ii) Tyne data, 1986

Measurement of change:
From Table 3 it is apparent that the stations immediately to the south of the disposal site show the greatest increase in the ratio over the assumed normal condition (Station 10 (39)). The two stations on the sludge disposal site have slightly lower ratios, perhaps suggesting that there is a net movement of sludge to the south. Stations 6, 7 and 8 have values very close to that of the control, whereas the station to the south of the fly-ash site is lower, suggesting some depressive effect.

Acceptability of change:
If a quality standard for the abundance ratio was set as any result in excess of 50% higher than the control, then Stations 1-4 would exceed it. If a standard of 100% higher than the control was set, as suggested from the evidence from the Garroch Head site, then no stations would exceed that standard, although Station 3 would come close.

(iii) St Abbs Head data, 1988

Measurement of change:
The St Abbs Head data set consists of information from a series of stations located on a circular grid centred on the sludge release area (see Figure 10(a)). Each of two stations lie to the north, south, east, and west of the central station at intervals of 1 nautical mile, and two additional stations lie 2 nautical miles to the south west and north east of the centre. SAB statistics are given in Table 4. To the west and to the north the abundance ratios decline and the size ratios increase as distance from the centre increases.

Although these changes are relatively small, it is worth commenting that such trends would be expected if enrichment were decreasing in these directions. To the east the abundance ratio declines slightly but the size ratio first increases (Station 15) and then decreases, suggesting some moderate enrichment at 1 nautical mile from the centre in this direction. To the south the abundance ratio increases initially and then declines and the size ratio gradually increases from the centre. This suggests a degree of enrichment 1 nautical mile from the centre in this direction.

These changes are relatively minor compared with those documented on the Garroch Head site. Table 4 compares the values of the statistics at each station with those recorded at Station 1, 2 nautical miles north of the centre, the station showing the least degree of variability.

The deviations from the values found at Station 1 are relatively minor, suggesting that despite the sludge input the faunal characteristics are fairly uniform over the area. Thus the abundance ratio deviates by more

Table 2. Garroch Head survey faunal statistics, 1985-1988 and 1990-91 (values 0.1 m^{-2})

	Station	Distance (km)	S	% Deviation	A	% Deviation	A/S	% Deviation
(a)	1985							
	G1	10	22	-	66	-	3	-
	P7	0	5	-77	6332	>1000	1266	>1000
	P6	1	18	-18	4728	>1000	263	>1000
	P5.5	1.5	45	+104	3693	>1000	82	>1000
	P5	2	50	+127	1064	>1000	21	+600
	P9	2	44	+100	888	>1000	20	+567
	P4	3	43	+95	313	+374	7	+133
	P10	3	39	+77	208	+215	5	+67
	P1	6	14	-36	58	-12	4	+33
(b)	1986							
	G1	10	20	-	53	-	3	-
	P7	0	5	-75	2493	>1000	499	>1000
	P8.5	1.5	10	-50	4792	>1000	479	>1000
	P5	2	41	+105	273	+415	7	+133
	P4	3	28	+40	174	+228	6	+100
	P10	3	36	+80	175	+220	5	+67
(c)	1987							
	G1	10	21	-	72	-	3	-
	P7	0	5	-76	4853	>1000	971	>1000
	P8.5	1.5	12	-43	11780	>1000	982	>1000
	P5	2	41	+95	804	>1000	20	+567
	P4	3	32	+52	402	+458	13	+333
	P10	3	45	+114	349	+385	8	+166
(d)	1988							
	G1	10	18	-	39	-	2	-
	P7	0	6	-67	10	-74	2	0
	P8.5	1.5	35	+94	995	>1000	28	>1000
	P5	2	50	+178	913	>1000	18	+800
	P4	3	32	+78	170	+336	5	+150
	P10	3	30	+67	185	+374	6	+200
(e)	1990							
	G1	10	24	-	82	-	3	-
	P7	0	4	-83	5553	>1000	1388	>1000
	P8.5	1.5	23	-4	11274	>1000	490	>1000
	P5	2	49	+104	979	>1000	20	+567
	P4	3	36	+50	723	+782	20	+567
	P10	3	31	+29	235	+187	8	+167
(f)	1991							
	G1	10	24	-	143	-	6	-
	P7	0	7	-71	1404	+882	201	>1000
	P8.5	1.5	18	-25	9332	>1000	518	>1000
	P5	2	51	+112	895	+526	17	+183
	P4	3	34	+42	404	+182	12	+100
	P10	3	38	+58	361	+152	9	+50

Table 3. Tyne survey faunal statistics, 1986 (values 0.1 m^{-2})

Station	Species		Abundance		A/S	
	No.	% Deviation	No.	% Deviation	Ratio	% Deviation
1(24)	60	+2	892	+61	14.9	+58
2(27)	61	+3	908	+64	14.9	+58
3(29)	61	+3	1100	+98	18.0	+91
4(30)	72	+22	1167	+111	16.2	+72
5(31)	58	-2	762	+37	13.1	+39
6(32)	62	+5	583	+5	9.4	0
7(33)	58	-2	597	+8	10.3	+10
8(34)	58	-2	497	-10	8.6	-9
9(36)	64	+8	516	-7	8.1	-14
10(39)	59	-	554	-	9.4	-

Table 4. St Abbs Head survey faunal statistics, 1988. (Values 0.2 m^{-2})

Station number	Distance and direction from disposal site centre (nm)	S	% deviation	A	% deviation	B	% deviation	A/S	% deviation	B/A	% deviation
1	2N	114	...	1954	...	33528	...	17	...	17	...
3	1N	116	+2	2244	+15	23727	-29	19	+12	11	-35
9	2W	91	-21	1289	-34	12669	-62	14	-18	10	-41
11	1W	118	+3	2188	+12	23058	-31	18	+6	10	-41
13	0	113	-1	2212	+13	15100	-55	20	+18	7	-59
15	1E	109	-4	2227	+14	64170	+91	20	+18	29	+71
17	2E	137	+20	2135	+9	30431	-9	16	-6	14	-18
23	1S	121	+6	4569	+133	47258	+41	38	+123	10	-41
25	2S	108	-5	2235	+14	23271	-31	21	+23	10	-41
27	2NE	117	+3	2829	+45	41472	+24	24	+41	15	-12
29	2SW	113	-1	2933	+50	37441	+12	26	+53	13	-24

than 100% only at Station 23, situated 1 nautical mile south of the centre, and the size ratio differs by more than 50% only at the central station and at the station 1 nautical mile east of the centre.

Acceptability of change:
Thus based on the criteria derived from observations of successional change along the strong gradients at the Garroch Head site the only areas at the St Abbs Head site where faunal conditions might be considered to be approaching an unacceptable level of change are areas 1 nautical mile or less to the east and south of the centre.

B A comparison of abundance ratios calculated on different levels in the taxonomic hierarchy

Introduction:
In the above account, three differing data sets were used to assess the validity of using abundance ratios (total number of individual organisms found in a sample divided by the total number of taxa), i.e. the mean number of organisms per taxon found in the sample. Assessments from these data sets, drawn from surveys of the Garroch Head disposal site in the Firth of Clyde, the St Abbs Head disposal site in the Firth of Forth and the Tyne disposal site were based on ratios calculated from counts of the numbers of species (or lowest identifiable taxon) in the samples. The present analysis extends the assessment to ratios calculated on the basis of higher taxonomic categories, at the family, class and phyla levels, in order to estimate what effect such a reduction in the level of detail might have on the interpretations made from the data bases.

Results:

(i) Garroch Head Sludge Disposal Site survey, 1983 data:
The total abundances of organisms found at each sampling station, together with the numbers in each taxonomic category and the ratios calculated from these data, are listed in Table 5. The variability in these statistics from station to station is compared in Figure 13. The stations are ordered along a 12 km long transect running from west to east across the centre of the disposal site.

Table 5. Garroch Head sludge disposal site data, 1983 (Values 0.1 m^{-2})

Station	A	S	F	C	P	A/S	A/F	A/C	A/P
P1	65	16	15	5	5	4	4	13	13
P4	183	28	20	4	4	6	9	46	46
P5	1045	45	31	8	6	23	34	131	174
P5.5	1459	42	32	11	8	35	46	133	182
P6	2872	27	21	7	5	106	137	410	574
P7	330	3	3	3	3	110	110	110	110
P8	8327	11	9	3	3	757	952	2776	2776
P8.5	2249	23	18	7	6	98	125	321	375
P9	205	29	26	9	7	7	8	23	29
P10	165	33	28	8	6	5	6	21	27
P11	70	21	19	6	5	3	4	12	14
P12	48	21	17	6	6	2	3	8	8

A total abundance of all organisms in sample
S total no. of species recorded in sample
F total no. of families recorded in sample
C total no. of classes recorded in sample
P total no. of phyla recorded in sample
A/S abundance ratio calculated from total species
A/F abundance ratio calculated from total families
A/C abundance ratio calculated from total classes
A/P abundance ratio calculated from total phyla

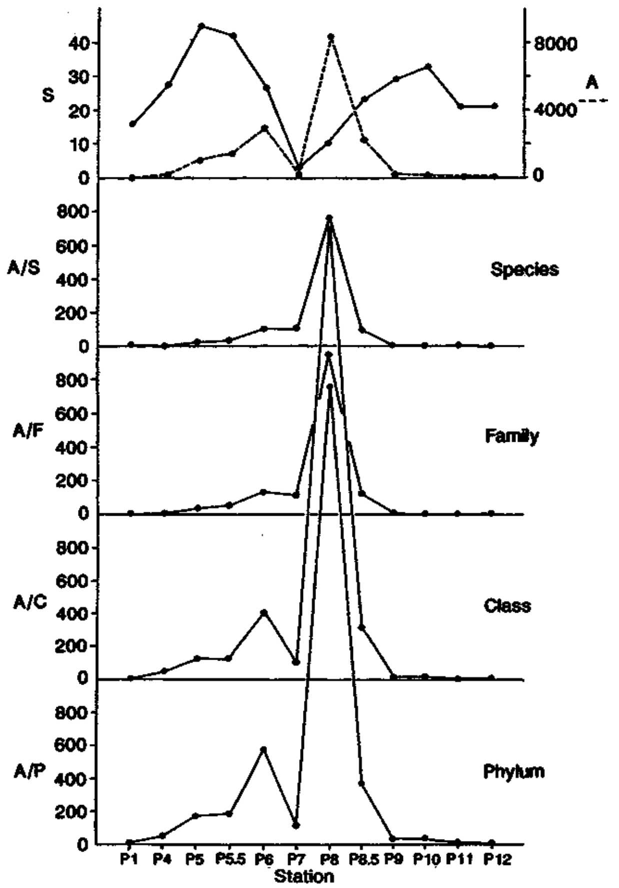


Figure 13. Garroch Head disposal site, 1983: abundance ratio indices at different taxonomic levels

Species numbers peak 2 km either side of the centre of the site and are minimal in the centre, whereas abundances peak 1 km either side of the centre and are low in the centre, although still higher there than in the unenriched sediments at either end of the transect. The abundance ratios at all four taxonomic levels reflect the enormous abundance changes at the central stations rather than the influences of changes in taxon numbers.

All four levels demonstrate the same pattern of increasingly large deviations from background levels with decreasing distance from the centre of the site. All stations within the two at either end of the transect exceed the mean values of the various ratios at the two outer stations by more than 50%.

(ii) St Abbs Head Sludge Disposal Site survey, 1988 data:
The total abundances of organisms found at each sampling station, together with the numbers in each taxonomic category and the ratios calculated from these data, are listed in Table 6. The variability in these statistics from station to station is compared in Figure 14. The stations are sited along radii running N/S, E/W and NE/SW across the centre of the site at Station 13.

Species numbers are highest at Station 17 furthest east and lowest at Station 9 furthest west of the centre. Abundances are highest at Station 23 to the south of the centre and lowest at Station 9. This latter pattern is followed by all the calculated ratios, emphasising the importance of the abundance changes in formulating the ratio values. The values at Station 1, furthest to the north of the centre, have been taken as reference values as in the previous assessment. Only at Station 23 do the ratio values exceed those at the reference station by more than 50%, but they do so whatever the taxonomic level used in the calculation.

(iii) Tyne Sludge Disposal Site survey, 1986 data:
The total abundances of organisms found at each sampling station, together with the numbers in each taxonomic category and the ratios calculated from these data are listed in Table 7. The variability in these statistics from station to station is compared in Figure 15. The stations are situated along a 16 km transect running in a south-easterly direction from the centre of the disposal site (in the direction of the prevailing current).

Table 6. St Abbs Head sludge disposal site data, 1988 (Values 0.2 m^{-2})

Station	Statistic								
	A	S	F	C	P	A/S	A/F	A/C	A/P
1	1954	114	63	22	12	17	31	85	163
3	2244	116	61	21	12	19	33	107	187
9	1289	91	66	20	11	14	19	64	117
11	2388	118	80	20	13	20	30	119	184
13	2212	113	80	20	13	19	28	111	170
15	2227	109	76	18	11	20	29	124	202
17	2135	137	91	21	14	16	23	102	152
23	4369	121	80	20	12	36	55	218	364
25	2235	108	73	21	14	21	31	106	160
27	2829	117	81	23	13	24	35	123	218
29	2933	113	77	19	13	26	38	154	226

A total abundance of all organisms in sample
S total no. of species recorded in sample
F total no. of families recorded in sample
C total no. of classes recorded in sample
P total no. of phyla recorded in sample
A/S abundance ratio calculated from total species
A/F abundance ratio calculated from total families
A/C abundance ratio calculated from total classes
A/P abundance ratio calculated from total phyla

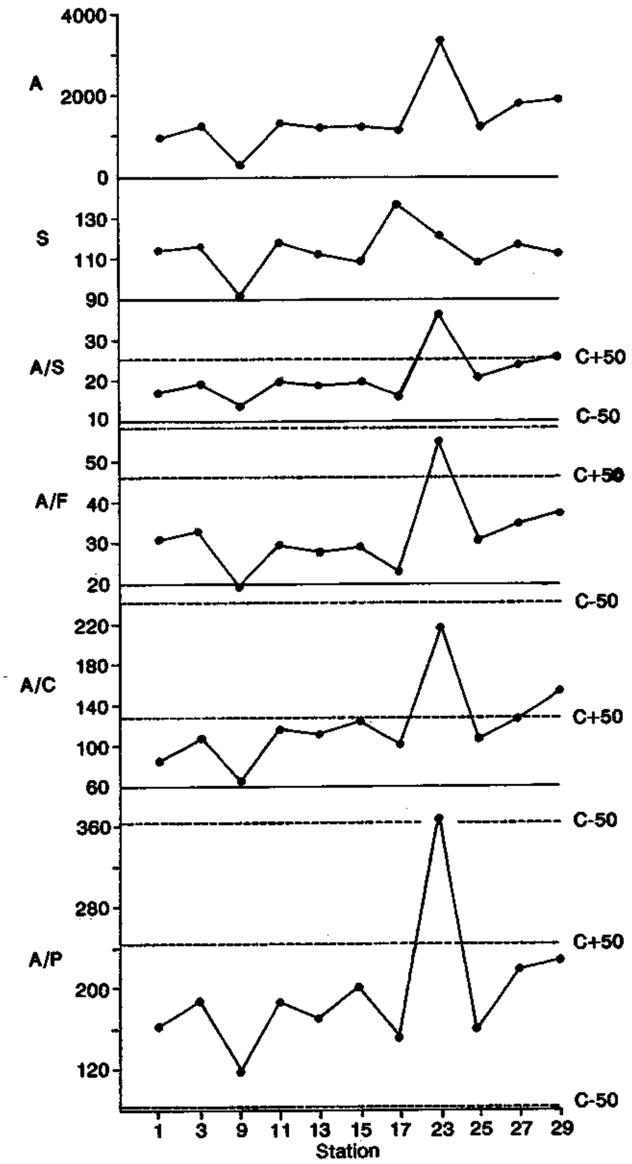


Figure 14. ***St Abbs Head disposal site, 1988: abundance ratios***

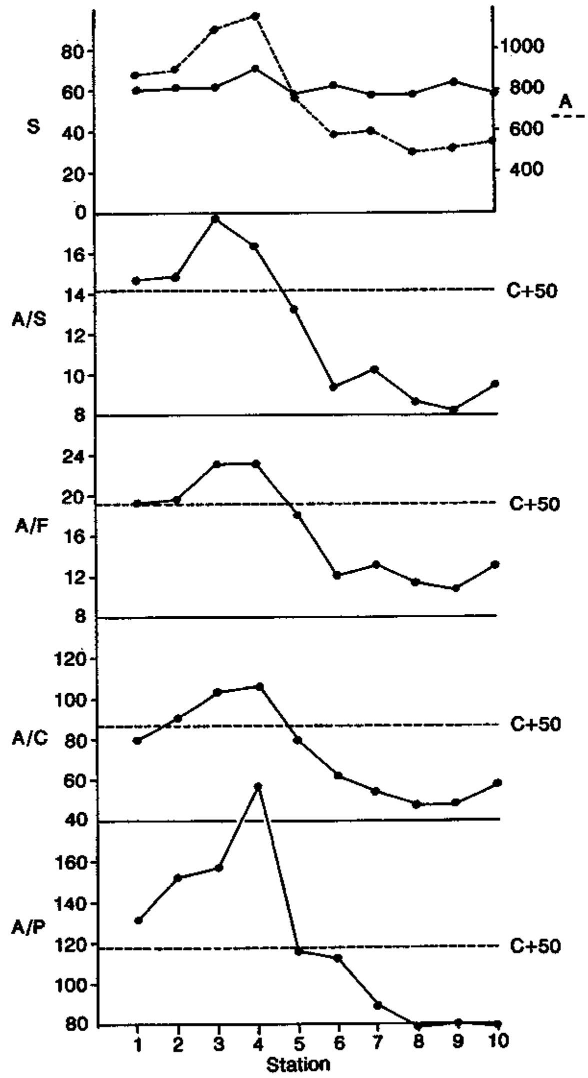


Figure 15. ***Tyne disposal site, 1986: abundance ratios***

Table 7. Tyne disposal site data, 1986 (Values 0.1 m^{-2})

Station	Statistic								
	A	S	F	C	P	A/S	A/F	A/C	A/P
1 (24)	882	60	46	11	7	15	19	80	132
2 (27)	147	61	46	10	6	15	20	91	152
3 (29)	1100	61	47	10	7	18	23	104	157
4 (30)	1167	72	51	11	6	16	23	106	197
5 (31)	762	58	43	10	7	13	17	80	116
6 (32)	583	63	47	9	5	9	12	63	113
7 (33)	597	58	45	11	7	10	13	54	89
8 (34)	499	58	44	11	6	9	11	47	79
9 (36)	516	64	48	11	7	8	11	48	80
10 (39)	554	58	43	10	7	9	13	58	79

A total abundance of all organisms in sample
S total no. of species recorded in sample
F total no. of families recorded in sample
C total no. of classes recorded in sample
P total no. of phyla recorded in sample
A/S abundance ratio calculated from total species
A/F abundance ratio calculated from total families
A/C abundance ratio calculated from total classes
A/P abundance ratio calculated from total phyla

The number of taxa found at each station does not vary greatly along the transect, the only deviation being a slight increase in diversity at Station 4 immediately to the south of the disposal site. Abundances are higher at all stations on and immediately to the south of the disposal site. These variations are reflected in the changes in the abundance ratios which are considerably higher at Stations 1-5 than at the stations more distant from the site. These differences are in excess of 50% higher than the values found at the control station 16 km from the centre of the site at Stations 2-4 when calculated at any level in the taxonomic hierarchy, and also higher than 50% at Stations 1-5 when calculated on the basis of species and families.

Discussion:
It is apparent that the level of taxonomic discrimination makes little difference to the changes in the abundance ratios which can be attributable to the effect of the impact of sludge disposal in these three areas. Even at the coarsest level of discrimination, that of the phyla, the pattern of change in relation to distance from the disposal area is not substantially changed. It could be argued therefore that discrimination to the lower taxonomic levels is unnecessary for the assessment of the overall effects of the disposal operations. However, some basic cautions must be expressed before such an approach can be advocated with any confidence; these are given in Annex 2.

It should also be noted that the abundance ratio is a simple diversity index which, like all such simple indices, is very sample-size dependant. As the level of taxonomic discrimination is reduced the influence of taxon number on the computation of the index declines in relation to abundance, thus exaggerating the sample size dependance.

Abundance/Biomass Comparison (ABC)
(R. Warwick)

An account of the methodology is given at Annex 4.

(i) Garroch Head Disposal Site:
ABC plots for the macrobenthos along a transect of stations across the accumulating sewage-sludge disposal site at Garroch Head, Scotland (Figure 16) are given in Figure 17. Note how the ABC curves behave along the transect, with the peripheral Stations P1 and P12 having unpolluted configurations, those near the disposal site centre at P7 with grossly polluted configurations and intermediate stations showing moderate pollution.

Of course, at the centre of the disposal site itself there are only three species present, so that any method of data analysis would have indicated gross pollution. However, the biomass and abundance curves start to become transposed some distance from the centre at P4 and P11 when species diversity is still high. In Figure 18 the results of a similar analysis are presented, but with the species data aggregated to FAMILIES. Note that these plots are virtually identical to those in Figure 17, so that there would have been no loss of information whatsoever if the fauna had initially been analysed to family level only.

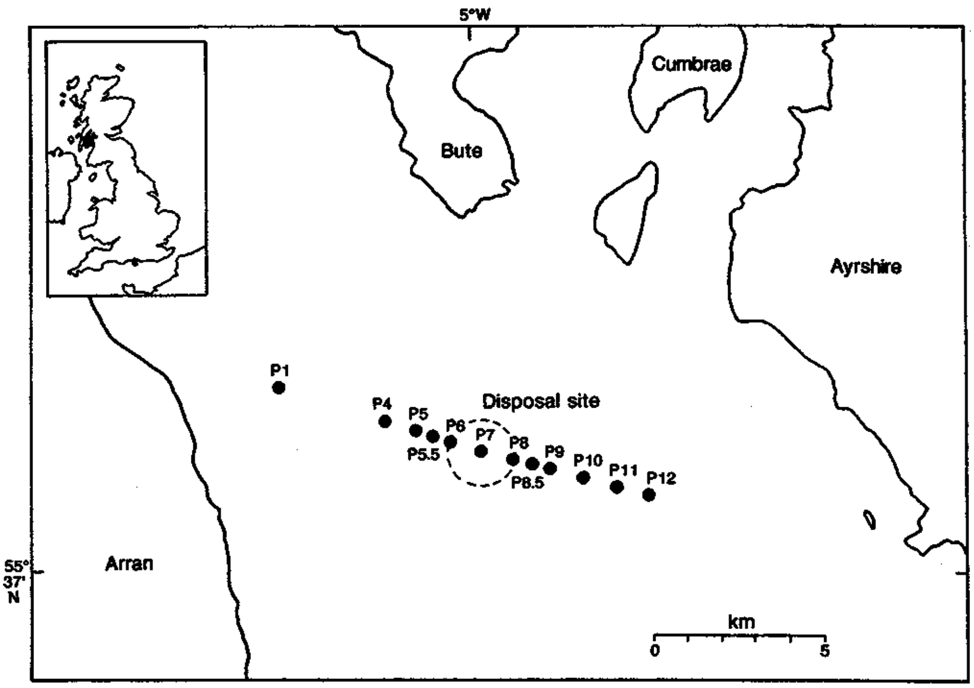


Figure 16. Map showing the location of Garroch Head sewage-sludge disposal site in the Firth of Clyde, Scotland. The centre of the disposal site is denoted by the 'dashed' circle and the positions of sampling stations P1-P12 are identified by dots

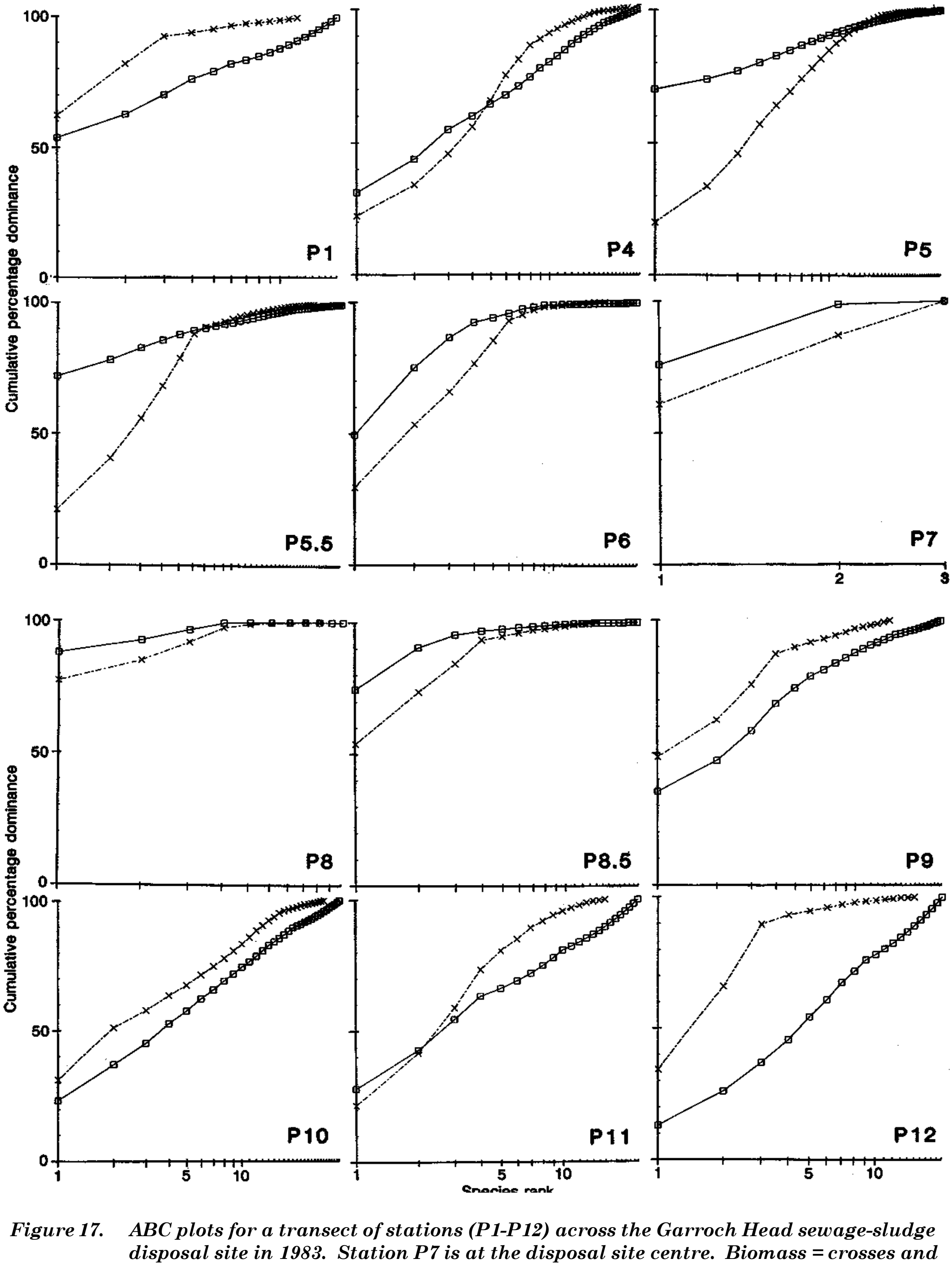


Figure 17. *ABC plots for a transect of stations (P1-P12) across the Garroch Head sewage-sludge disposal site in 1983. Station P7 is at the disposal site centre. Biomass = crosses and dashed lines, abundance = squares and continuous lines. (From: Warwick et al., 1987)*

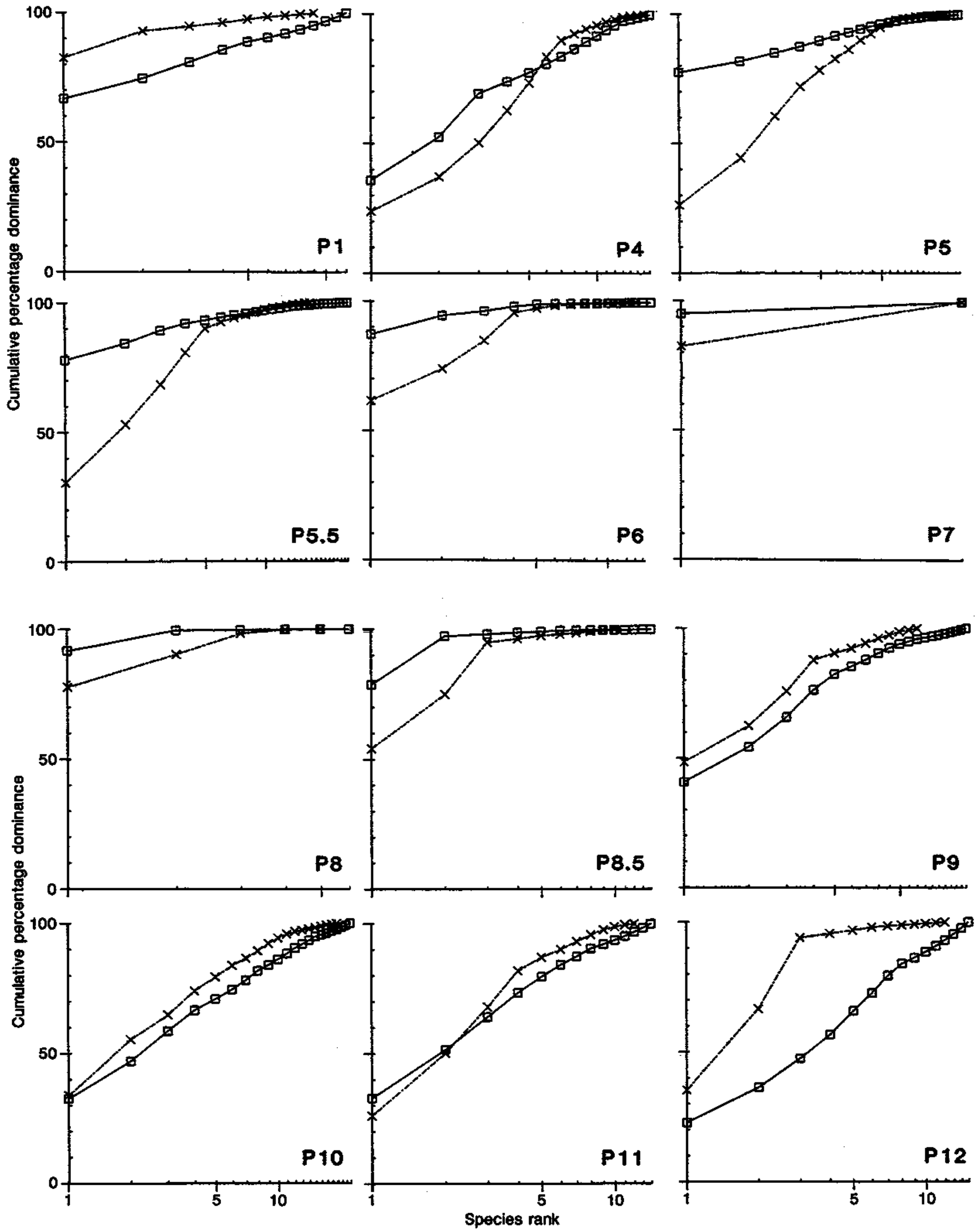


Figure 18. As Figure 17, but with species data aggregated into families

(ii) Tyne Disposal site:
The 1986 test data did not include biomass, and so data collected by the Plymouth Marine Laboratory (PML) on a subsequent survey in 1989 at a limited number of stations along the same transect (see Figure 7) have been used (Figure 19). Note that the ABC plots all show the unpolluted configuration, except for Station 7 which indicates moderate pollution. This must result from some unknown form of disturbance at this site and cannot be related to sludge disposal. Stations 1 and 3

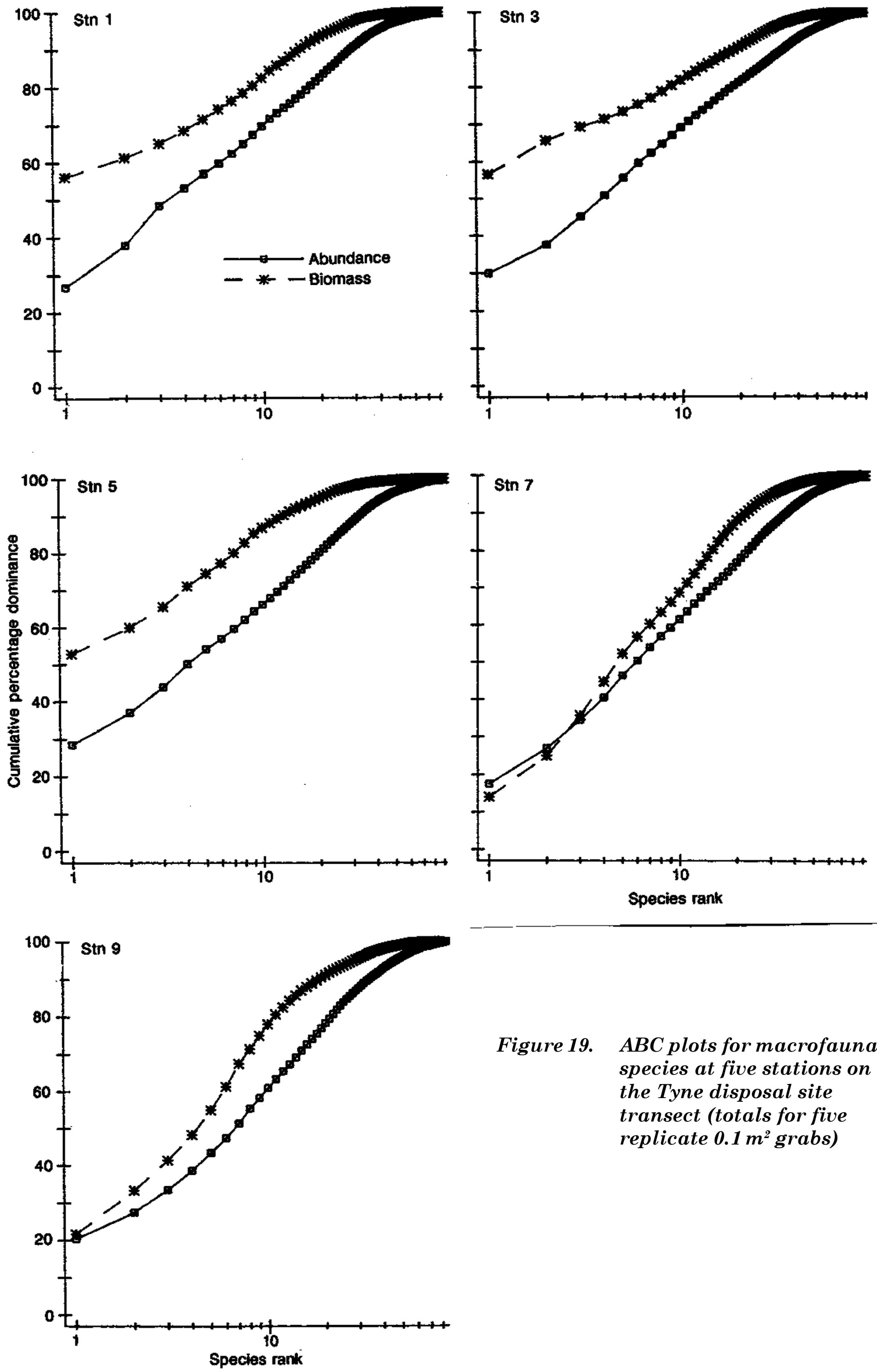


Figure 19. ***ABC plots for macrofauna species at five stations on the Tyne disposal site transect (totals for five replicate 0.1 m² grabs)***

which are likely to be the most impacted both show the unpolluted configuration, so that by this method the level of disposal must be regarded as acceptable. Since there was no species level effect, the data have not been re-analysed at the family level.

(iii) St. Abbs Head Disposal site:
Biomass data were not available for the test data (June 1988), and data from the check monitoring survey of 1987 have been used here. Only seven stations were sampled and two replicates for biomass were only available from Stations 11, 13, 15, 23 and 25 (see Figure 20). This is barely adequate for ABC analysis, so the results (Figure 21) must be viewed with caution.

Stations 15 and 13 appear to be unpolluted: the biomass at Station 15 is overwhelmingly dominated by *Echinocardium cordatum*, and the absence of this species from other stations might well be the result of sampling error. Stations 11 and 25 appear to be moderately polluted, and Station 23 moderately/grossly polluted. Abundance at Station 23 is dominated by the small polychaete *Myriochele oculata*. (This species is a good indicator of intermediate levels of enrichment on many of the oilfields in the central North Sea: see e.g. Gray *et al.*, 1990).

However, taken at face value, the ABC method does indicate unacceptable levels of pollution to the south and west of the disposal site. In view of the uncertainties of interpretation of these data, with so few replicates at each station, analysis at the family level was not considered appropriate.

General conclusions:
Despite the fact that only one replicate grab sample was taken at each station along the Garroch Head transect, a clear trend was evident and detrimental effects were detected 3 km to the east and 4 km to the west of the disposal site. Even with adequate replication (5 grabs per station), no detrimental effect could be associated with the Tyne disposal site. Detrimental effects were noted at some stations on the St Abbs Head ground, but these must remain equivocal because of inadequate replication.

Coefficient of Pollution *(M. Service)*

(i) Method of calculation:
The Coefficient of Pollution (C of P) was described by Bogdanos and Satmadjis (1985) on the basis of a perceived relationship between the infauna and sediment structure.

The calculation involves a number of assumptions:

(a) the number of species increases linearly with the number of individual animals;

(b) sediment structure can be represented by a single value (the sand equivalent, s') based on the relative percentages of sand and silt;

(c) faunal abundance is related to s' and depth.

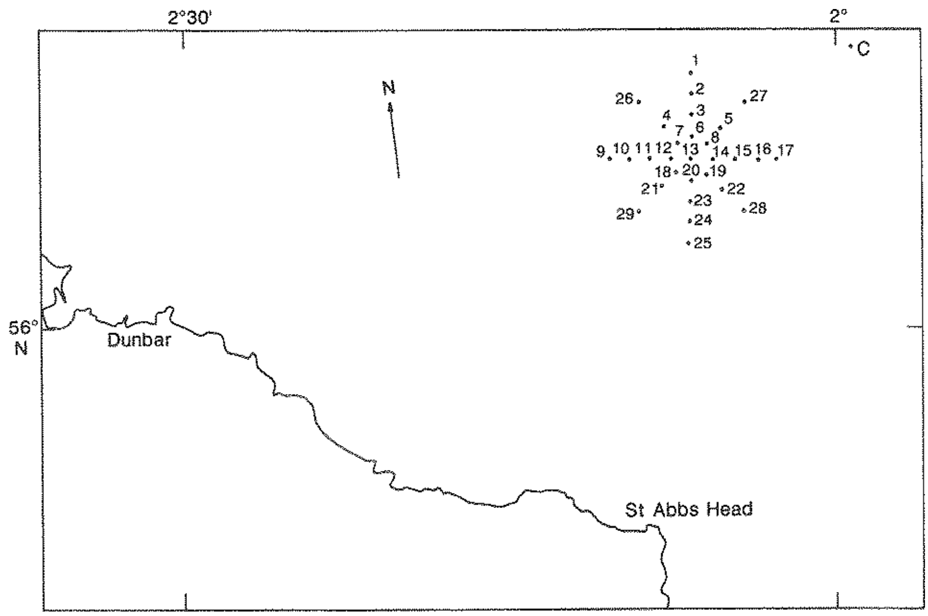
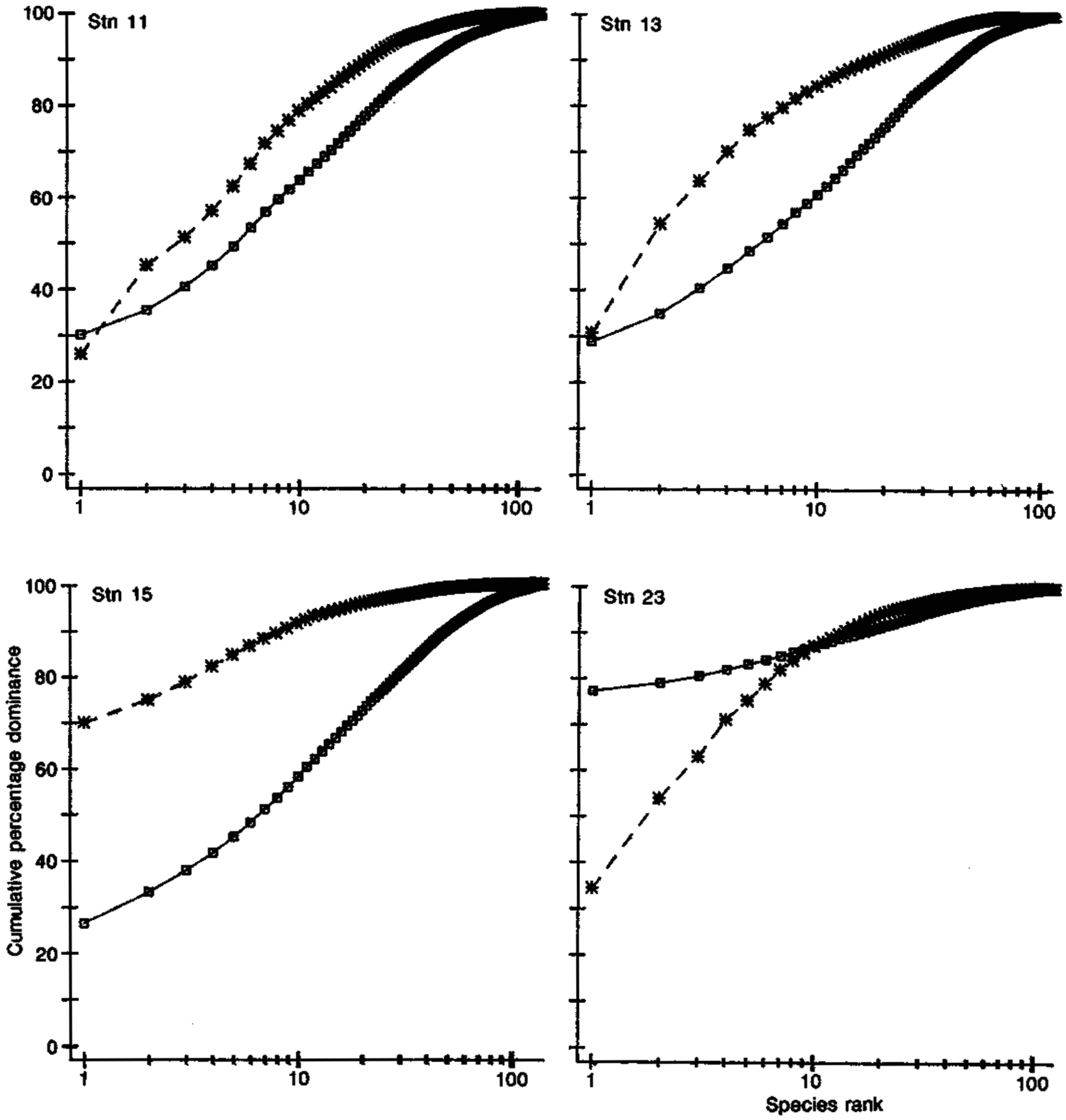


Figure 20. Positions of sampling stations on the St Abbs Head disposal site



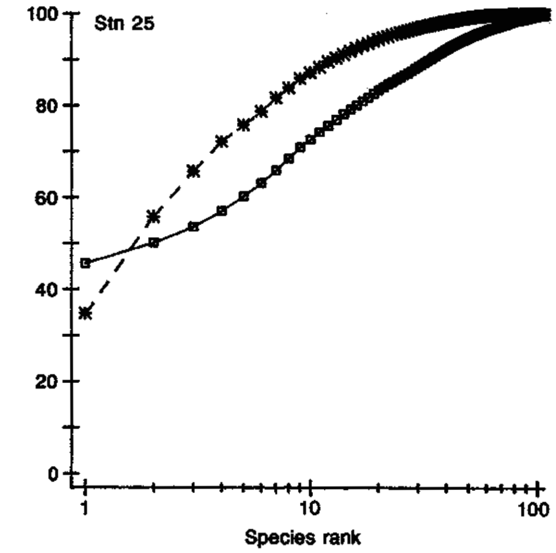


Figure 21. ABC plots for macrofauna species at five stations on the St Abbs Head disposal site (totals of two replicate 0.1 m² grab samples). Squares and solid lines = abundance, asterisks and broken lines = biomass

Following from these assumptions Bogdanos and Satmadjis (1985) generated the following equations:

(1) $s' = s + t(5+0.2s)$
s = %sand
t = %silt

The relationship between faunal abundance, s' and depth is used to calculate the theoretical number of individuals for a given sediment type and depth using the equation:

(2) $i' = (-0.0187s'^2 + 2.63s' - 4)(2.20 - 0.0166h)$
i' = theoretical number of individuals
h = depth(m)

The relationship between numbers of individuals and species is used to calculate the number of species which should correspond to the actual number of individuals assuming an unpolluted environment:

(3) $g' = i/(0.0124i + 1.63)$
g' = theoretical number of species
i = actual number of individuals

The Pollution Coefficient is calculated from the relationship between the calculated values of i' and g' and the actual values, using the following equation:

(4) $P = g'/[g(i/i')^{1/2}]$
P = Pollution Coefficient
g = actual number of species

Coefficients of 1.5-2.0, 2.0-3.0, 3.0-4.0 and 4.0-8.0 are assumed to indicate slight, moderate, heavy and very heavy pollution, respectively. The Coefficient of Pollution has been applied by Maurer and Haydock (1989) and Maurer *et al.* (1991) to data from the California Shelf, and found to compare favourably with other indices.

(ii) Application to test data sets:
In order to test its utility as a tool for measuring the benthic impact of sewage sludge disposal, the Coefficient of Pollution was calculated using data from Garroch Head (1983), St Abbs Head (1988) and the Tyne (1986). Where full particle size data was unavailable, %silt was taken as the percentage less than 63 µm. In the case of Garroch Head data, no particle size data was available, and the relative percentages were assumed to be t=85% and s=15%.

(iii) Results:
Coefficients of Pollution for the three data sets are plotted in Figure 22(a-c), along with % organic carbon at each station. Only one station exceeded the 'slight pollution' category (see above). In fact, all the St Abbs Head and Tyne stations produced low scores.

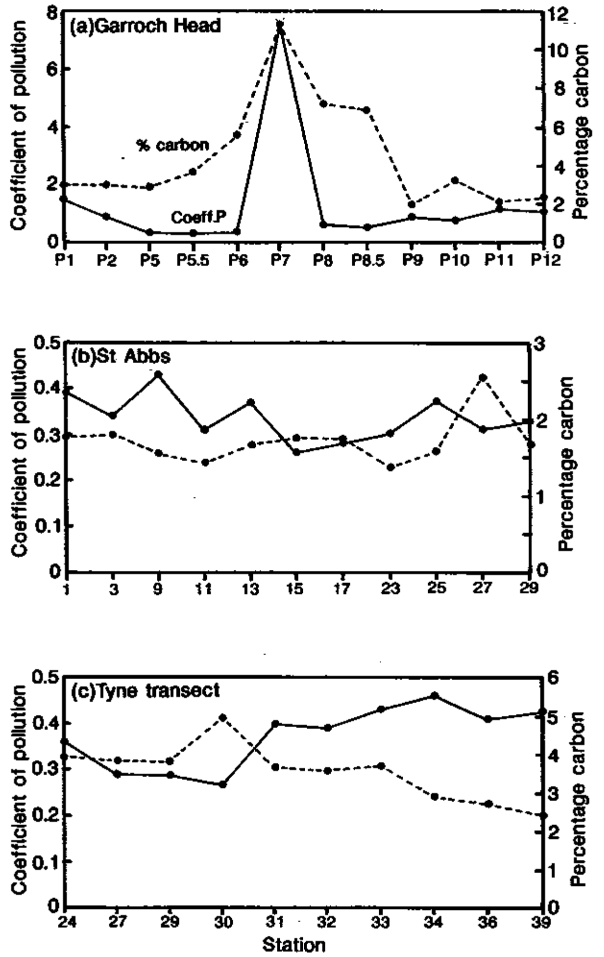


Figure 22. Trends in the Coefficient of Pollution for test data sets

Station P7 at Garroch Head scored 7.6 (= 'very heavy pollution'). This of course agrees with highly elevated organic carbon at this station, along with depressed faunal abundances and numbers of species. However, at the remaining Garroch Head stations, and at the other sites, the Coefficient of Pollution appears to move in the opposite direction to that expected by the organic carbon content (see, for example, trends at Stations P10 and P11 in Figure 22(a), and at Stations 33 and 34 in Figure 22(c)).

(iv) Discussion:
On the basis that the assumptions made regarding particle size were not totally erroneous, the Coefficient of Pollution does not seem to be suitable in its present form for application to UK sewage-sludge disposal sites. The reason appears to be a lack of sensitivity to intermediate pollution levels. In turn, this appears to be due to a general under-estimation of faunal abundance in comparison with natural levels, arising from calculation of equation (2) above. Also, equation (3) above makes no allowance for the fact that faunal abundance will increase under slight to moderate pollution, but numbers of species can either stay constant or show only a slight elevation.

However, it should be noted that both Satsmadjis (1985) and Maurer and Haydock (1989) were able to apply the Coefficient of Pollution with some success. This suggests that further development of this index is needed, before it could be used successfully in UK waters. In particular, the simple regression models used for the prediction of i' and g' need to be re-defined.

Diversity, evenness and other univariate measures ***(M. Elliott and J. Pomfret)***

Introduction:
Between-station variability in the following univariate measures of community structure has been assessed by means of statistical analysis of the test data sets from Garroch Head, St Abbs Head and the Tyne at both the species and family level: number of taxa (T), number of individuals (A), Shannon-Wiener diversity H' $\log_2$ (Shannon and Weaver, 1949), Evenness (E) (Pielou, 1966) and Abundance/Taxa (A/T).

Trends in the primary variables T, A and B and the derivatives A/T and B/A have been examined earlier in this Section. The present assessment is designed to complement the findings from this work: for example, it provides a more detailed insight into relationships between species- versus family-level data; however, some degree of overlap in the coverage was unavoidable.

Rationale for statistical evaluation:
Where replicate samples were available, the data from each location were subjected to an Analysis of Variance (ANOVAR) with the following aims:

(i) to examine for the presence of significant differences between stations;

(ii) to determine the magnitude of variability across the data matrices in their entirety. This was assessed by reference to values of the F-ratio provided by the ANOVAR;

(iii) to assess, from a ranking of F-ratios, which among the above variables gave the strongest signal at individual locations, and to compare their performance between locations;

(iv) to compare variability at the species and family levels, using paired t-tests.

(It was not possible to provide an assessment of variability in biomass, because of insufficient data).

Graphical representation of spatial trends:
It is clear from Figures 23(a-c) that, with few exceptions, trends in selected variables along the transects at the species and family level are identical. (Note that trends in untransformed A/T ratios were discussed earlier in

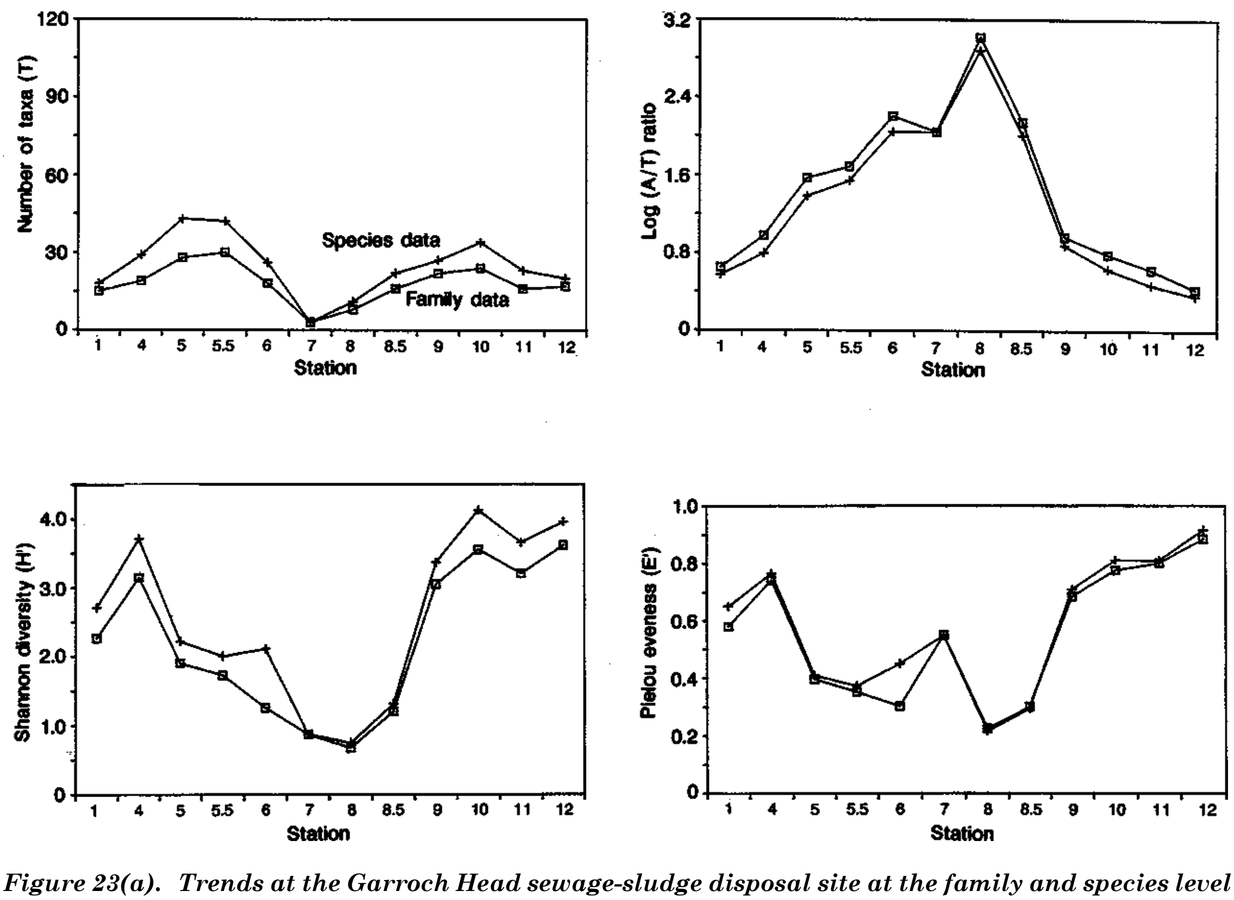


Figure 23(a). Trends at the Garroch Head sewage-sludge disposal site at the family and species level along an east-west transect

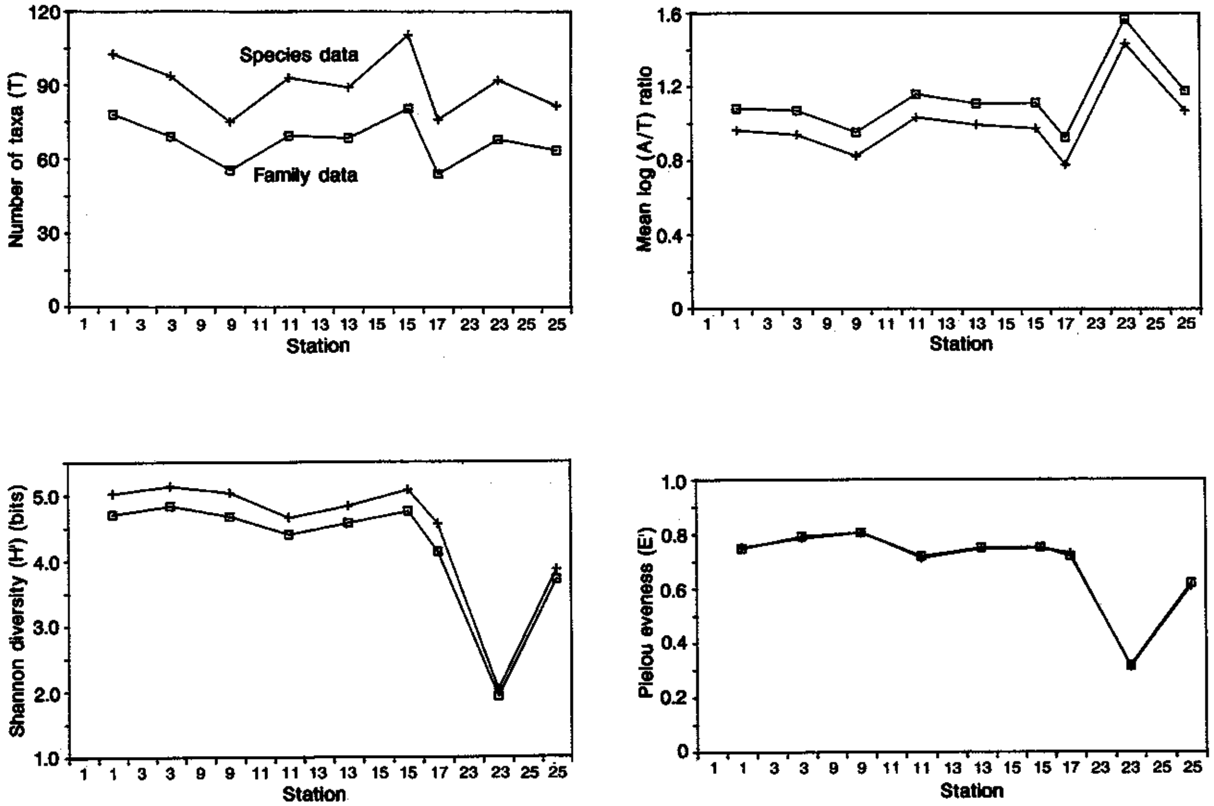


Figure 23(b). Trends at the St Abbs Head sewage-sludge disposal site at the family and species level along a north-south pseudo-transect in 1987

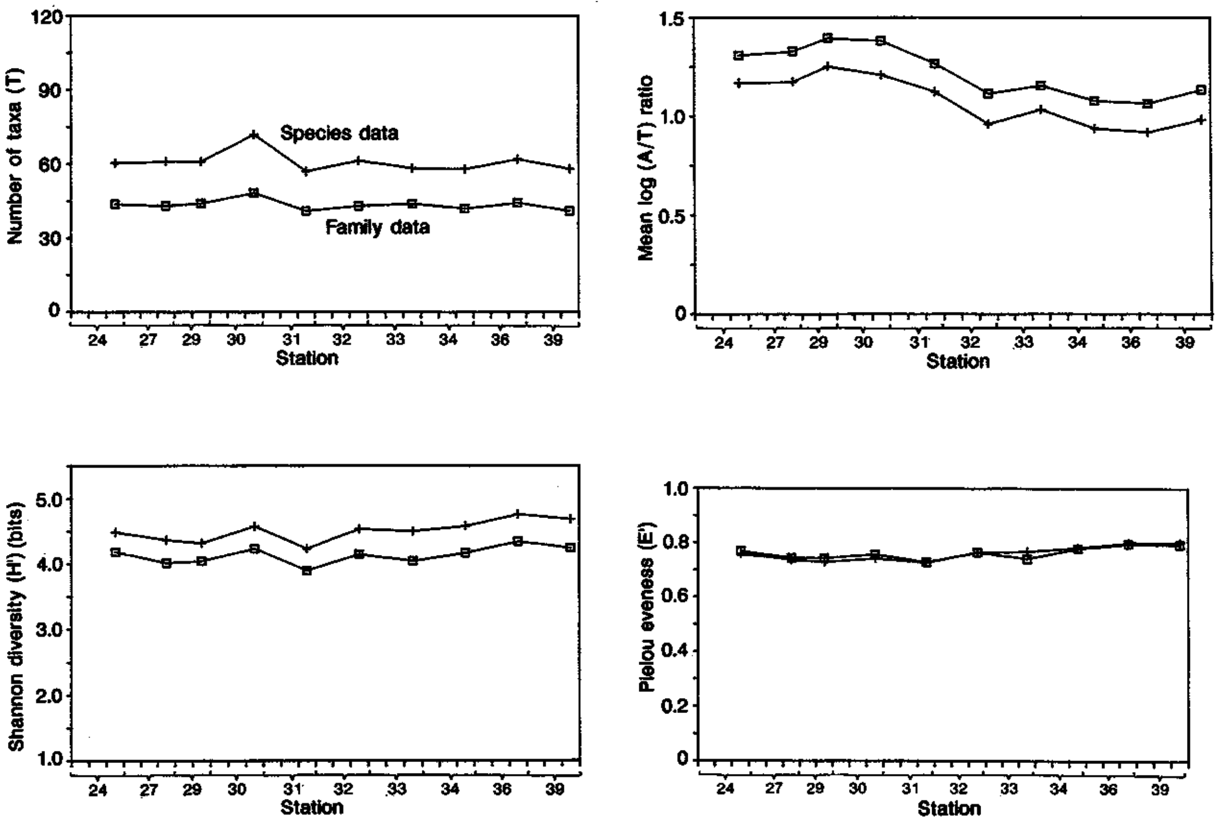


Figure 23(c). Trends of the Tyne sewage-sludge disposal site at the family and species level along a north-south transect in 1986 this Section). The following features are notable:

(i) number of taxa: the same pattern is produced irrespective of the taxonomic level, although this is more accentuated with species data; the absolute values decrease by up to 50% with a change from species to family;

(ii) log A/T: patterns are in good agreement but, as expected, the ratio increased by up to 50% (on an arithmetic scale) for family-level data;

(iii) H': the same pattern is evident, although there is a decrease of up to 10% in absolute values when family data are used; the magnitude of this difference declines with decreasing values of H' so that, for example, at grossly affected stations at the centre of the Garroch Head disposal site (Figure 23(a)), no differences are apparent;

(iv) E: values for St Abbs Head and the Tyne at the species- and family-levels are barely distinguishable; however, there is a less consistent relationship (with cross-over points) at Garroch Head.

It may be concluded that trends in the variables at the species- and family-levels are in good agreement. However, the magnitude of any trend appears dampened at the family-level, and hence it is possible that subtle trends may be obscured. It is also apparent that the patterns of variation in log A/T along the transects are generally in close agreement with the inverse of those for H'. In fact, they are highly correlated at both the species and family levels:

Family data: Pearson $r = 0.376$; $df = 186$; $p<0.001$

Species data: Pearson $r = 0.450$; $df = 197$; $p<0.001$

Given some uncertainty regarding the biological significance of H', it may therefore be more appropriate to use log A/T.

ANOVAR and paired t-test analysis:
This analysis aimed to detect the strength of the trends among the variables at both the species- and family-levels (see Tables 8(a) and (b)).

Table 8(a). Results of ANOVAR for univariate measures

F ratio	Species level data		Family level data	
Dependent variable	St Abbs Head	Tyne	St Abbs Head	Tyne
Abundance (A)	9.559	10.753	9.559	10.753
Log (A)	4.275	8.394	4.275	8.394
Taxon number (T)	4.503	1.301	5.491	0.613
A/T ratio	14.366	12.494	11.001	14.870
Log (A/T)	5.661	9.927	4.548	12.754
Shannon index (H')	30.982	2.626	29.100	2.183
Pielou evenness (E)	23.059	3.935	22.285	1.996
d.f.	8	9	8	9

Significance of F (p values)				
Dependent variable	St Abbs Head	Tyne	St Abbs Head	Tyne
Abundance (A)	0.0022	<0.0001	0.0022	<0.0001
Log (A)	0.0277	0.0001	0.0277	0.0001
Taxon number (T)	0.0239	0.299	0.0133	0.7715
A/T ratio	0.0005	<0.0001	0.0014	<0.0001
Log (A/T)	0.0121	<0.0001	0.0232	<0.0001
Shannon index (H')	<0.0001	0.0366	<0.0001	0.0729
Pielou evenness (E)	0.0001	0.0058	0.0001	0.0980

Table 8(b). Results of paired t-tests on F-ratio values

Results paired are for species- and family-level taxonomy for each of the following variables: Log (A), (T), Log (A/T), H', E

Area	t-value	d.f.	p value for t	Pearson corr. coeff.	p value for corr.
St Abbs	1.14	3	0.335	0.998	0.002
Tyne	0.06	3	0.955	0.981	0.019

(i) St Abbs Head:
The strongest trends were seen with H' and E, followed by A/T; there was little difference in the behaviour of the variables at different taxonomic levels (Table 8(a)). The paired t-test carried out on the F-ratio values (Table 8(b)) indicated no significant difference between the species- and family-levels, and the two data sets were significantly correlated, as would be expected. (The variables A and A/T were excluded from all of the paired t-tests to avoid duplication and undue bias in the analysis).

(ii) Tyne:
The highest F-ratios were for A, A/T and log A/T, although the rank order of the variables differed slightly according to the taxonomic level. For both the species- and family-level data, H', E and T gave the weakest trends. Furthermore, the ANOVAR for T was not significant at either taxonomic level, and those for H' and E were significant at the species level but not at the family level which may, in the latter case, indicate lowered sensitivity. The paired t-test gave the same result as for the St Abbs Head data.

Determination of variability across whole data sets:
This variability has been assessed by computing the means, standard deviations and coefficients of variation for species- and family-level data at Garroch Head, St Abbs Head and the Tyne and is given in Table 9. Variability of all measures at Garroch Head was very high, with coefficients of variation ranging from 40-60%. This was to be expected, given the known strength of the biological signal attributable to sludge disposal. Of 15 stations sampled at St Abbs Head, two were notable for exceptionally high densities especially of the polychaete *Myriochele oculata* arising from patchy recruitment of juveniles. Analyses of the full data set show that coefficients of variation for H' and E were markedly elevated, compared with the outcome of analyses following their exclusion. In pollution

Table 9. Variability in combined test data sets for Garroch Head, St Abbs Head and the Tyne

	Garroch Head	St Abbs*	St Abbs*	Tyne
No. of samples:	12	15	13	29
Log A				
Mean (species)	2.62	2.97	2.90	2.85
Mean (families)	2.62	2.97	2.90	2.85
S.D. (species)	0.74	0.22	0.17	0.15
S.D. (families)	0.74	0.22	0.17	0.15
%cv (species)	28.14	7.41	5.82	5.11
%cv (families)	28.14	7.41	5.82	5.11
T				
Mean (species)	24.83	91.18	92.54	60.90
Mean (families)	18.00	68.18	68.92	43.41
S.D. (species)	11.61	12.30	13.30	6.78
S.D. (families)	7.63	8.88	9.72	4.35
%cv (species)	46.74	13.46	14.37	11.13
%cv (families)	42.38	13.02	14.11	10.01
Log A/T				
Mean (species)	1.30	1.01	0.94	1.07
Mean (families)	1.43	1.14	1.07	1.22
S.D. (species)	0.81	0.19	0.12	0.13
S.D. (families)	0.81	0.20	0.12	0.13
%cv (species)	61.96	19.10	12.51	11.81
%cv (families)	56.65	17.54	11.06	10.68
H'				
Mean (species)	2.56	4.46	4.93	4.51
Mean (families)	2.20	4.19	4.61	4.13
S.D. (species)	1.20	1.01	0.27	0.21
S.D. (families)	1.08	0.94	0.27	0.18
%cv (species)	46.97	22.66	5.44	4.62
%cv (families)	48.91	22.43	5.82	4.34
E				
Mean (species)	0.58	0.67	0.76	0.76
Mean (families)	0.55	0.69	0.76	0.76
S.D. (species)	0.23	0.16	0.05	0.03
S.D. (families)	0.23	0.15	0.05	0.03
%cv (species)	39.67	23.88	6.13	4.27
%cv (families)	41.66	21.74	6.13	4.18

** See text for details*

surveys, such findings may be open to misinterpretation, without recourse to expert biological interpretation of the raw data.

For St Abbs Head (excluding the two abberant stations) and the Tyne, where gradients in the data are much weaker, coefficients of variation for log A, H' and E are approximately 5%, whereas those for T and log A/T are 10-15%. Depths, substrates and current regimes are comparable between these sites, which probably accounts for the similarity. For all these data sets, it is evident that pooling to family level has very little effect on standard deviations or coefficients of variation. The latter were highly correlated (Pearson $r = 0.995$; $df = 10$; $p<0.01$ using arcsin-transformed data), as would be expected.

Treatment/Reference comparison for the Tyne transect: Pairwise comparisons of 'Treatment' and 'Reference' sites (see Figure 1 and Section 2.2.2) were made along the Tyne transect. Stations were pooled into three groups to give a 'Treatment' site (4 northernmost stations in Figure 7), a 'Reference 1' site (4 central stations in Figure 7) and a 'Reference 2' site (2 southernmost stations in Figure 7).

The means with standard deviations of four variables determined from family-level data for each group are given in Table 10, along with the outcome of pairwise comparisons. Data are also given for three single stations corresponding with the locations of the pooled groups. Marginal elevations in log. abundance, taxa and log A/T are evident at the 'Treatment' site, and this is reflected in the Treatment/ Reference comparison. No clear pattern emerged for H'.

It may be inferred from standard deviations that there is very little evidence for statistical differences between stations or station groups. This would be consistent with findings from the Treatment/Reference comparison, which provided no indication of unacceptable effects. It is of some interest to note that, in this instance, the pooling of stations makes little difference to the precision of mean values.

Overall, the results are consistent with mild enrichment near to the disposal site.

Discussion:

1. The use of F-ratios for significance testing of the ANOVAR assumes that the data are normally distributed. Thus p-values associated with F-ratios for A and A/T in particular should be treated with caution.

2. In the comparison between the Tyne and St Abbs Head data sets, it was not possible to identify a single variable which consistently gave the highest F-ratio. Thus at St Abbs Head, the F-ratio for H' and E gave the strongest indication of significant differences between stations, whereas at the Tyne, the ANOVAR for these variables were non-significant at the family level but significant at the species level. The latter may indicate some loss of sensitivity as a result of pooling. However, with few exceptions, the spatial patterns for the data at species level were similar to those at the family level.

3. Similar (opposing) trends in values of H' and log A/T would suggest that the latter might be more suitable for routine use, as it is both easier to calculate and more amenable to biological interpretation.

4. There is clearly scope for more detailed statistical analysis of test data sets, including the use of multiple range tests for identifying stations responsible for the presence of significant differences, and analysis of covariance to link univariate measures to environmental factors. Multiple range testing will also indicate whether the same configuration of stations is achieved at different taxonomic levels.

Table 10. Test application of 'EQS' ratios to Tyne data set (family-level data)

Ratios labelled Trt/Ref % are calculated as ((Treatment/Reference)-1) x 100
Ratios labelled Ref[1]/Ref[2] % are calculated as ((Reference[1]/Reference[2])-1) x 100

	Log(A) Mean	SD	Taxa (T) Mean	SD	Log(A/T) Mean	SD	H' Mean	SD
Trt (4 pooled stations)	2.998	0.069	44.82	4.708	1.35	0.055	4.13	0.160
Ref[1] (4 pooled stations)	2.779	0.112	42.50	3.631	1.15	0.095	4.06	0.183
Ref[2] (4 pooled stations)	2.726	0.057	42.67	5.046	1.10	0.058	4.29	0.119
Trt/Ref[1] %	7.863		5.45		17.04		1.53	
Trt/Ref[2] %	9.977		5.04		22.81		-3.83	
Ref[1]/Ref[2] %	1.959		-0.39		4.94		-5.28	
Trt (Station 1)	2.947	0.032	43.67	3.512	1.31	0.049	4.19	0.235
Ref[1] (Station 5)	2.875	0.089	41.00	5.292	1.26	0.034	3.90	0.138
Ref[2] (Station 10)	2.746	0.079	41.00	3.000	1.13	0.057	4.24	0.099
Trt/Ref[1] %	2.474		6.50		3.35		7.42	
Trt/Ref[2] %	7.318		6.50		15.33		-1.30	
Ref[1]/Ref[2] %	4.727		0.00		11.59		-8.12	

5. It may be concluded that univariate measures provide a useful means to summarise important attributes of benthic community structure, and have the advantage of being amenable to straightforward statistical testing. The primary variables T, A and B are unambiguous descriptors, while the derived variables H' and E, along with ratios of the primary variables, attempt to combine the attributes of species occurrence, and the apportioning of individuals among those species, with varying degrees of sophistication.

6. The limitations attached to single-figure summaries of complex data have been well documented. For example, natural events such as a particularly successful recruitment, or excessive predation, may give rise to changes in diversity indices comparable with those occurring in response to anthropogenic influences. Misinterpretation of the causes of such events can only be prevented by reference to the raw data. Another drawback in the use of univariate measures is that widely different communities can produce similar values. Therefore, employment in assessments of anthropogenic effects should always be accompanied by complementary analyses (especially the use of multivariate techniques) to aid expert interpretation.

2.2.2 Temporal changes

Multidimensional scaling (MDS) ***(R. Warwick)***

Tees Bay (Stations 1 and 2): time-series data

The appropriate questions relating to these data appear to be:

(i) are there any differences in community structure between the two stations? How great are these differences? Are they statistically significant? Is there any trend in the magnitude of the difference over time (increasing or decreasing)?

(ii) at each individual station, are there significant differences in community structure between years? Are these differences statistically significant? Are these changes random or unidirectional over time?

Differences between stations:
MDS has been performed on the individual replicate data (5 replicates at each station) for each year separately, using square-root transformed species abundance data and the Bray-Curtis similarity measure. The ANOSIM test has been used to assess the degree of difference between stations and its statistical significance: see Annex 3 for description of the method.

Stations 1 and 2 are significantly different from each other at the 5% level in the years 1974, 1978, 1979, 1980, 1984 and 1985 (Table 11). The differences can be visualised (separation of 1 and 2 replicates) in Figure 24. The R statistic of the ANOSIM test fluctuates rather randomly from year to year, and there is no clear trend of increase or decrease.

Table 11. Tees Bay: ANOSIM to test for differences between Stations 1 and 2 each year

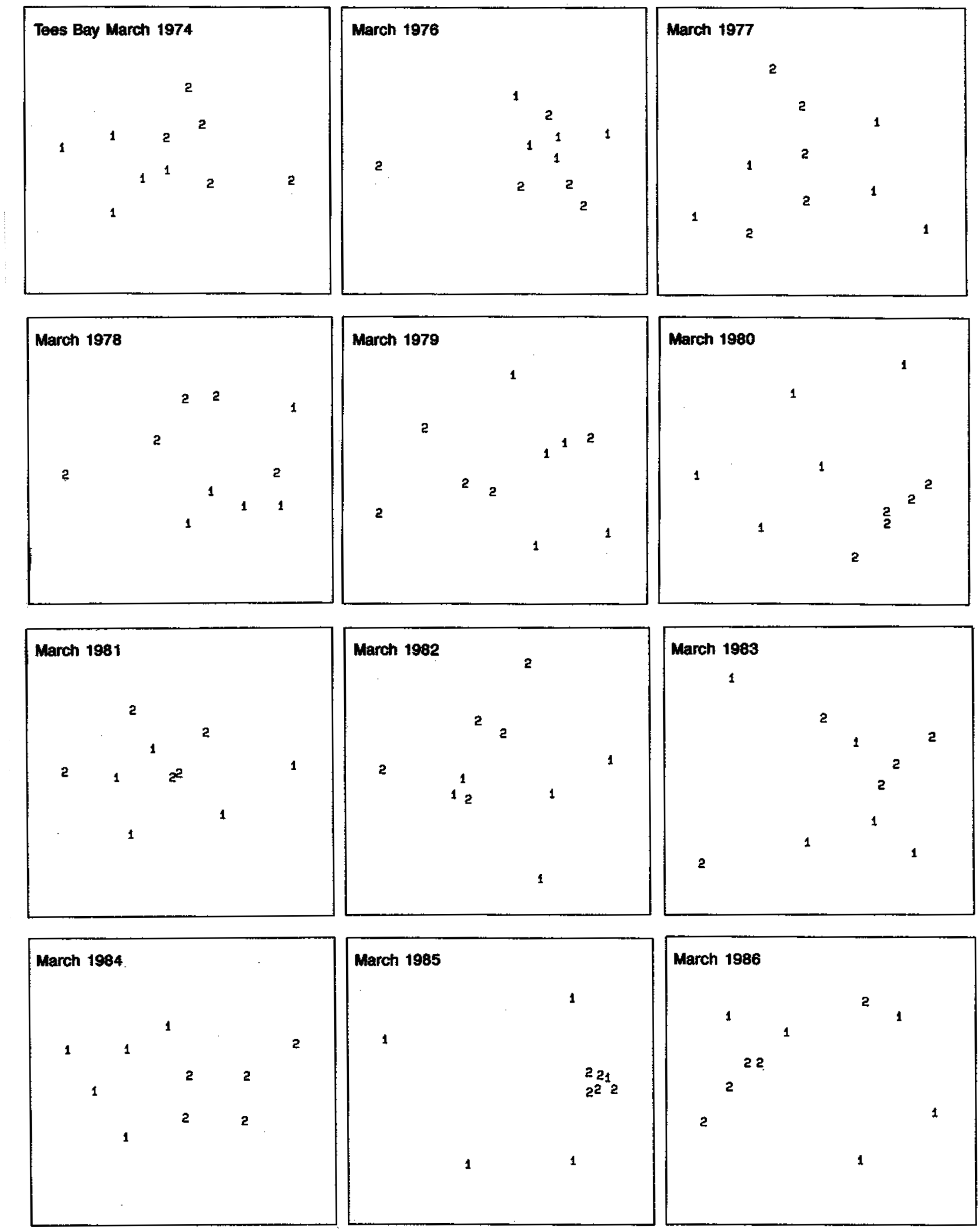
5 replicates at each station, total number of permutations = 126

Year	No. of significant statistics	Statistic value	% Significance level
1974	1	.392	.794
1976	37	.048	29.365
1977	7	.220	5.556
1978	6	.300	4.762
1979	6	.304	4.762
1980	1	.552	.794
1981	64	.004	50.794
1982	27	.108	21.429
1983	13	.144	10.317
1984	1	.660	.794
1985	4	.172	3.175
1986	18	.132	14.286
1987	14	.212	11.111

Community changes at each station over time:
MDS analysis has been performed on the individual replicate data for all years at each individual station, and differences between years tested using ANOSIM. The combined replicates for each station have been analysed together in a single MDS to see whether changes over time are similar at each station.

At each station the individual replicates for a given year tend to cluster together (Figure 25). The differences between every successive year are significant at the 5% level for both stations (Table 12); for this particular data set, it might therefore be argued that the test is <u>over-sensitive</u> to change. The MDS does not reveal a unidirectional trend of change, but rather this is quite random. Variations in the R statistic of the ANOSIM test (= the magnitude of the change) are also quite random. The two stations change in a similar manner over time (i.e. they cluster together on the MDS (Figure 26)).

General conclusions:
Changes in community structure over the years, although significant, appear to be quite random and cannot clearly be associated with any long-term trend of environmental change. The fact that the two stations sometimes differ significantly from each other, and at other times do not, can probably be associated with natural environmental processes. For some unaccountable reason, Station 1 was particularly anomalous in 1985.



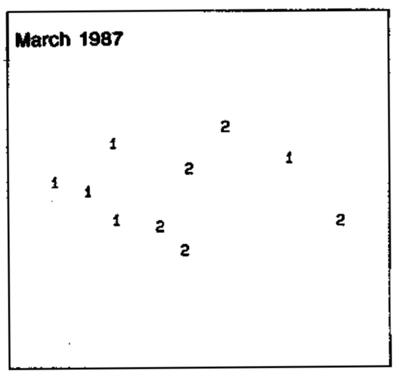


Figure 24. *MDS of root-transformed species abundance data for all replicates in each individual year. 1 = Station 1, 2 = Station 2*

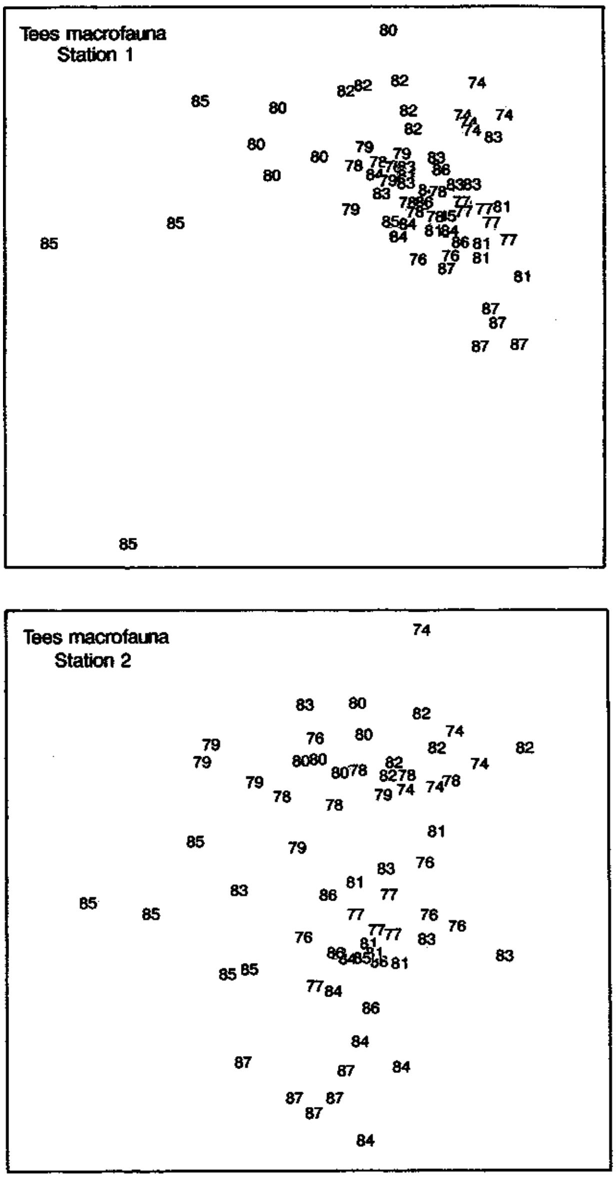


Figure 25. MDS of individual replicate data for all years, for each individual station

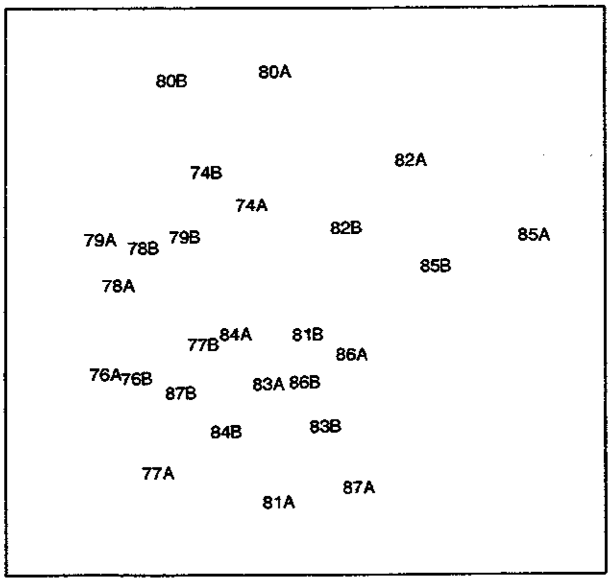


Figure 26. MDS for both stations in all years (all five replicates combined) A = Station 1, B = Station 2

Table 12. Tees Bay: ANOSIM test for differences between years at each station

5 replicates at each station each year

Size of random sample = 1000

Approx. number of permutations =1.238E+054

Station 1

No. of significant statistics = 1

Significance level of sample statistic 0.706 is 0.100 % Approx. 95% C.I. is (0.02%, 0.58%)

Pairwise effects tests (total permutations = 126 in each case)

Years compared	Statistic value	% Significance level
1974-1976	1.00	0.79
1976-1977	0.88	0.79
1977-1978	1.00	0.79
1978-1979	0.45	2.38
1979-1980	0.63	0.79
1980-1981	0.98	0.79
1981-1982	0.98	0.79
1982-1983	0.92	0.79
1983-1984	0.59	0.79
1984-1985	0.40	1.59
1985-1986	0.40	0.79
1986-1987	0.94	0.79

Station 2

No. of significant statistics = 1

Significance level of sample statistic 0.731 is 0.100%

Approx. 95% C.I. is (0.02%, 0.58%)

Pairwise effects tests

Years compared	Statistic value	% Significance level
1974-1976	0.56	1.59
1976-1977	0.34	1.59
1977-1978	0.89	0.79
1978-1979	0.37	3.17
1979-1980	0.57	1.59
1980-1981	0.96	0.79
1981-1982	0.84	0.79
1982-1983	0.65	0.79
1983-1984	0.77	0.79
1984-1985	0.79	0.79
1985-1986	0.76	0.79
1986-1987	0.94	0.79

Species/Abundance/Biomass (SAB) comparison ***(T. Pearson)***

Tees Bay (Stations 1 and 2): time-series data; A/S comparisons

The data from the 5 and 3 replicate compilations show the same trends, thus that from the 5 replicate lists will be discussed (see Table 13).

The thirteen-year means for the two stations are very similar and the variations in the A/S ratios from year to year follow broadly the same patterns at each station, suggesting that both areas are reacting to the same forcing functions.

If it is assumed that a fifty percent deviation from the mean will represent an abnormal deviation from the expected pattern then, at Station 1, lower than expected values were found in five years and higher than expected only in 1987. At Station 2, three years showed lower than expected values and 1984 and 1987 were higher.

At both stations by far the greatest deviation from the means occurred in 1987. If the means for the years 1974-86 are used as the expected values then, at Station 1, only two years, 1980 and 1985 show greater than 50% negative deviations and three years, 1976, 1981 and 1987 are more positive than expected. At Station 2, the years showing greater than expected deviations remain the same.

These analyses suggest some differences between the two stations. Both showed similar trends in 1980 and in 1987, the indices being lower than expected in the former year and very much higher in the latter. In other years, deviations from the standard expected differed at the two stations. Little can be inferred from the patterns of these deviations, other than to suggest that the very high positive deviations in 1987 imply a considerable degradation in the communities in that year, whereas the fairly consistent negative deviations in the period 1978-80 at both stations suggest a general improvement in the communities during that period.

It may be noted that if the data had been examined for deviations greater than 100%, then only Station 2 in 1984, and Stations 1 and 2 in 1987, would have been singled out. This would be consistent with the major departures from an oscillating trend evident in the A/T plots of Figure 28.

Abundance/Biomass Comparison (ABC) ***(R. Warwick)***

For the Tees Bay data set, no biomass data for individual species were available. However, the same procedure would normally be applied to time-series data, as to spatial data (see Annex 4).

Diversity, evenness and other univariate measures ***(H. Rees)***

Tees Bay (Stations 1 and 2): time-series data

(a) Analysis of variance:
The outcome of analysis of variance of data at the species and family levels is shown in Tables 14 and 15, for 5 and 3 replicates, respectively. (Data for 1974 were omitted, so as to provide an unbroken annual time-series). For 5 replicates, there are highly significant differences between years for all variables. With the exception of evenness for species-level data, and H' for family-level data, the between-site comparisons showed no significant differences. There is evidence of significant year x site interaction for all variables other than numbers of taxa.

Table 13. Tees Bay data: A/S comparisons

Station		1 (5 Replicates)		2 (5 Replicates)		
Mean A/S	1974-1987	8.84		8.53		
	1974-1986	5.87		7.06		
	A/S	% of mean 1974-1987	% of mean 1974-1986	A/S	% of mean 1974-1987	% of mean 1974-1986
1974	10.4	+17	+76	6.33	-26	-11
1976	7.3	-17	+25	6.6	-22	-6
1977	8.7	-2	+47	9.3	+10	+32
1978	4.2	**-52**	-29	4.5	-47	-36
1979	3.6	**-59**	-38	3.5	**-59**	**-51**
1980	1.6	**-82**	**-73**	3.8	**-56**	**-53**
1981	10.4	+17	**+77**	9.4	+10	+33
1982	3.3	**-62**	-43	3.9	**-54**	-44
1983	6.0	-32	+2	5.4	-37	-24
1984	5.3	-40	-10	16.5	**+94**	**+134**
1985	2.1	**-76**	**-64**	6.6	-23	-7
1986	7.5	-15	+28	8.8	0	+25
1987	44.5	**+403**	**+658**	26.2	**+207**	**+270**

Table 14. Tees Bay Stations 1 and 2 (1976-87): F-values from two-way analysis of variance; species-level data

(i) 5 Replicates (error df=96)

	Sites df=1	Years df=11	Years x Sites df=11
H'	3.78 ns	25.76 ***	3.16 **
E	4.16 *	31.93 ***	4.59 ***
Taxa (T)	0.02 ns	7.72 ***	1.74 ns
Individuals (A)	0.27 ns	29.78 ***	5.16 ***
A/T	0.24 ns	36.72 ***	6.50 ***

(ii) 3 Replicates (error df=48)

	Sites df=1	Years df=11	Years x Sites df=11
H'	4.73 *	15.99 ***	2.53 *
E	6.89 *	19.96 ***	3.95 ***
Taxa (T)	0.05 ns	3.75 ***	0.98 ns
Individuals (A)	0.01 ns	9.63 ***	2.15 *
A/T	0.00 ns	14.52 ***	3.11 **

ns = not significant
** = significant at 5-1% level*
*** = significant at 1-0.1% level*
**** = significant at <0.1% level*

Table 15. Tees Bay Stations 1 and 2 (1976-87): F-values from two-way analysis of variance; data pooled to family level

(i) 5 Replicates (error df=96)

	Sites df=1	Years df=11	Years x Sites df=11
H'	4.06 *	29.28 ***	3.44 ***
E	3.88 ns	33.63 ***	4.86 ***
Families (F)	0.38 ns	6.99 ***	1.52 ns
A/F	0.30 ns	38.79 ***	6.76 ***

(ii) 3 Replicates (error df=48)

	Sites df=1	Years df=11	Years x Sites df=11
H'	4.59 *	17.58 ***	2.68 **
E	7.24 **	22.26 ***	4.01 ***
Families (F)	0.33 ns	4.25 ***	0.97 ns
A/F	0.01 ns	14.84 ***	3.00 **

ns = not significant
** = significant at 5-1% level*
*** = significant at 1-0.1% level*
**** = significant at <0.1% level*

The results for 3 replicates only are broadly similar; for both species- and family-level data, there are significant between-site differences in H' and E. A comparison of Tables 14 and 15 provides no evidence to suggest that the resolution of the data is significantly impaired as a result of pooling to the level of family. This can be explained by the fact that the complement of taxa at the two sites across all sampling occasions was reduced only by about 13%, on average, as a result of pooling. Analysis of variance of the data from Station 1, derived from 5 replicates at the species and family levels respectively, for each sampling occasion, showed no significant differences between estimates of H', E and the A/T ratio. (For numbers of taxa, F=10.63; df=96; p=0.15%).

It may be concluded from the above analyses that there is a much stronger signal for differences between years, than for differences between sites. Evidence for a measure of co-variability between sites is shown in Figure 27, which gives the fitted regression lines with 95% confidence limits for log/log plots of each of the four variables at Station 1 (x) relative to Station 2 (y). Though the linear relationship is in all cases significant, there is appreciable variability, particularly in the case of numbers of taxa.

Time-series plots of the major variables at each site for species-level data are given in Figure 28. These are expressed as means with 'Least Significant Intervals' (LSIs) at the 95% probability level, and are calculated as follows:

$$LSI = \bar{x} \pm \frac{tS(2/n)^{1/2}}{2}$$

where t = student's t statistic corresponding with the degrees of freedom associated with S.
S = root MSE from the analysis of variance
and n = number of observations in each mean.

In Figure 28, it may be assumed that the means are significantly different if their LSIs do not overlap. Further details of this multiple comparison procedure are given in Andrews *et al*. (1980). It provides a useful — though not infallible — way of determining the significance of trends at 'Treatment' and 'Reference' sites. For individuals, departures from a generally similar trend at both sites are most notable in 1980, 1984 and 1985. In 1987, increases in abundances at both sites are accompanied by depression of diversity and evenness indices, and elevation of the A/T ratios. This suggests a disproportionate increase in counts of one or two species only. There is also a shift from lower to higher mean numbers of taxa at Station 2 relative to Station 1 from 1982, though the differences are generally not significant.

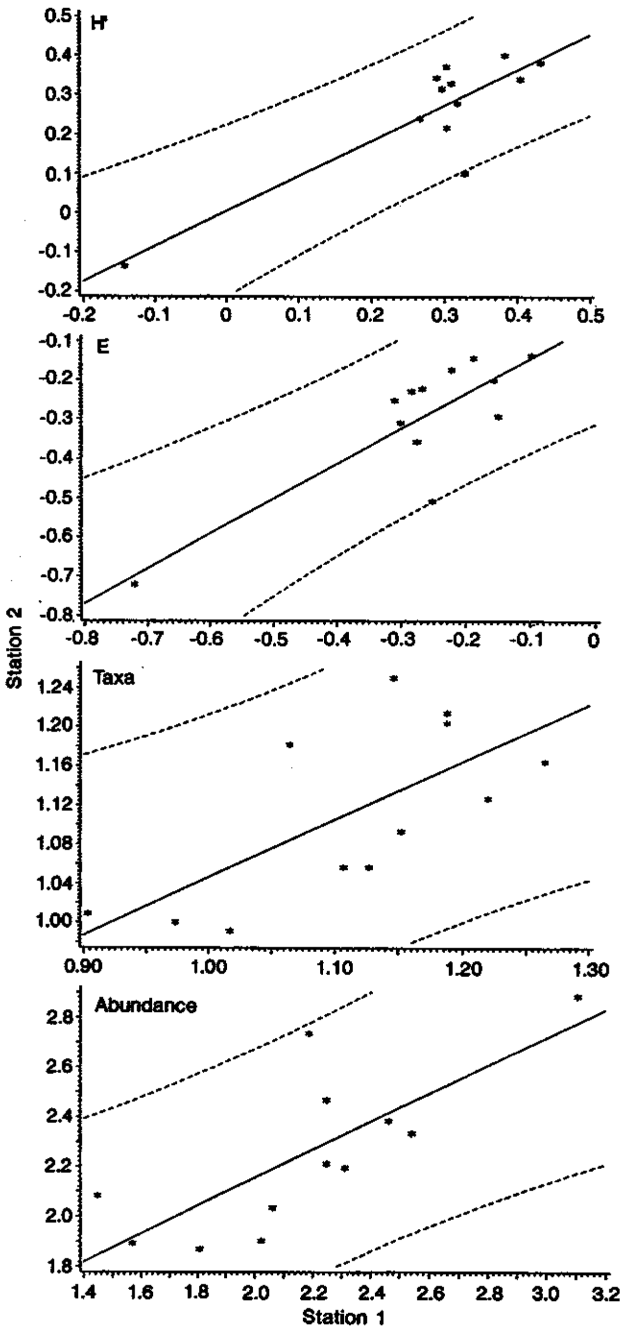


Figure 27. Fitted regression lines, with 95% confidence limits for logarithmic plots of four variables at Tees Bay (Station 1 versus Station 2)

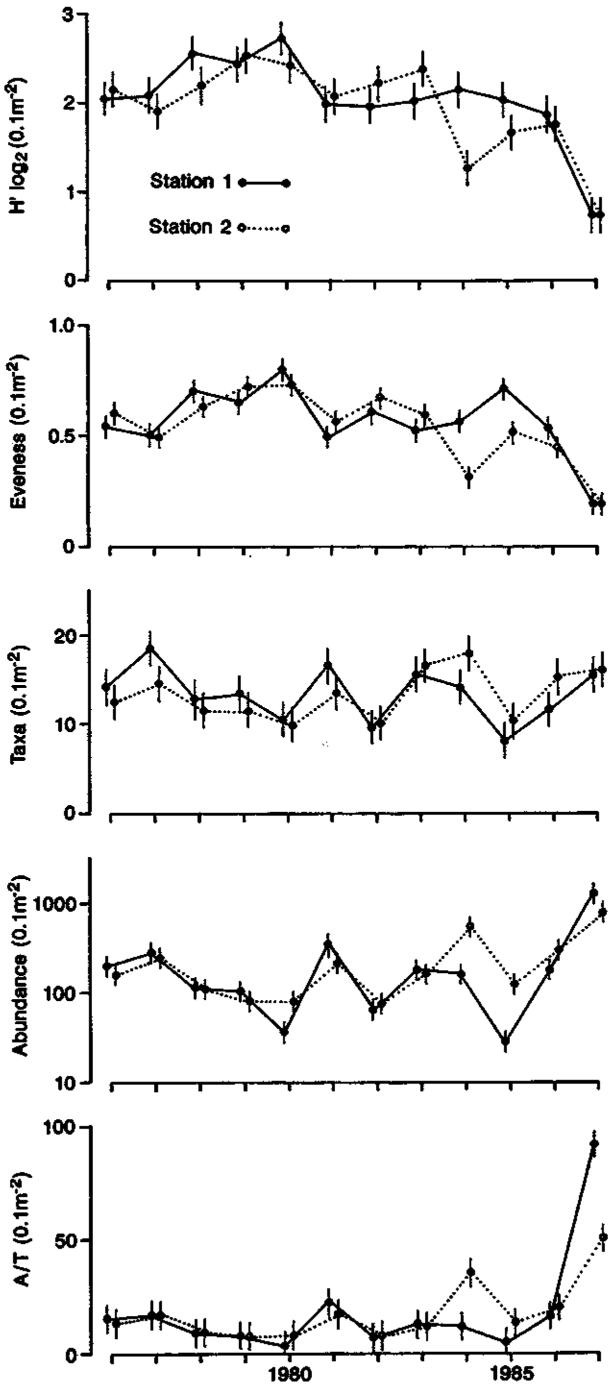


Figure 28. Tees Bay Stations 1 and 2: trends in primary and derived variables at the 'species' level (means ± 95% LSIs)

The fixing of quantitative limits for permissible change at one site relative to another is not straightforward when the data are expressed as in Figure 28, unless it can be ascertained that structural properties of the faunal assemblages at each site are, for practical purposes, identical at the outset of an investigation. Even if this were to be the case, any standard must still allow for the occurrence of natural synchronous changes at treatment and reference sites.

(b) An alternative method for pairwise comparisons and the setting of limits on variability:

This method — which is derived from the 'control/treatment pairing' principle described by Skalski and McKenzie (1983) — involves the calculation of ratios for different measures of data structure, as follows:

(Treatment/Reference - 1) x 100

The 'baseline' value is defined as the arithmetic mean of the Treatment/Reference comparison for the first three years of the data set. In Figure 29, these are represented by solid horizontal lines, and are accompanied by arbitrary boundaries which vary according to the chosen measure. (Station 1 has been arbitrarily designated a 'Treatment' site for the purpose of this comparison).

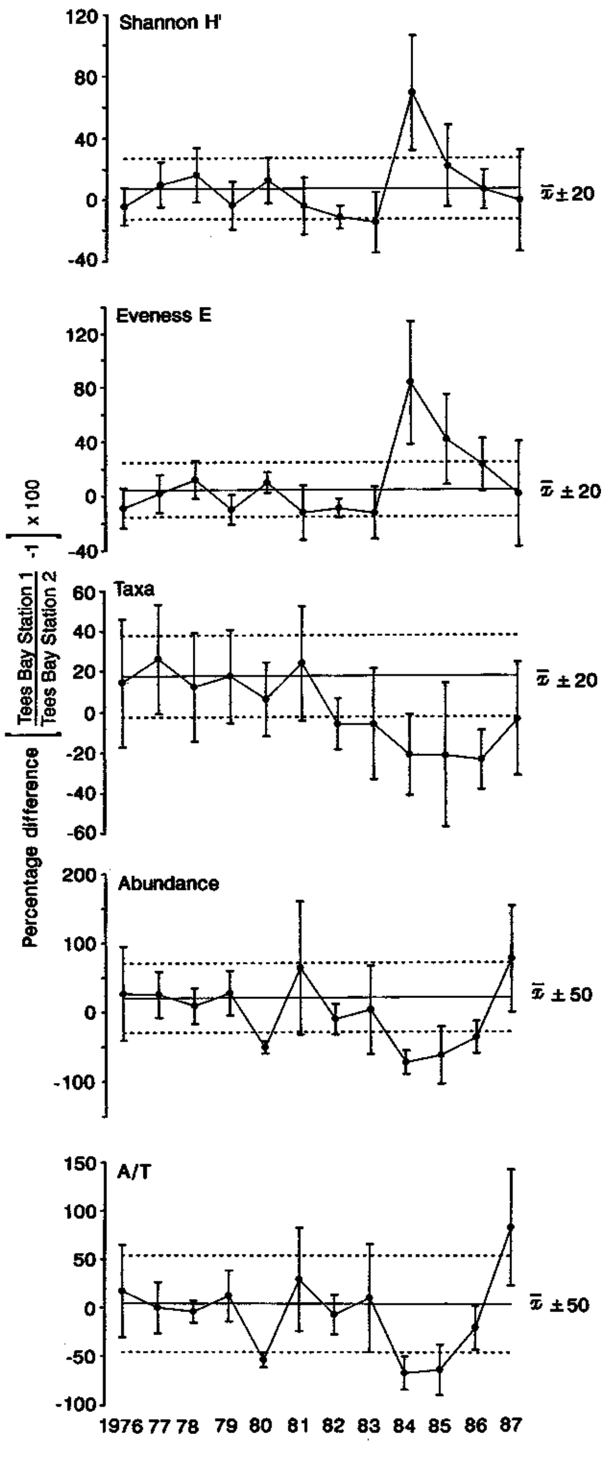


Figure 29. ***Tees Bay Stations 1 and 2 pairwise comparisons (note the differences in scale for the various measures)***

Pairwise comparisons for each measure, and on each sampling occasion, were conducted on the average of 5 replicates for the 'Treatment' and 'Reference' sites. 95% confidence limits were derived by 'Bootstrapping' (see Efron and Tibshirani, 1986), employing a computer program written by R Fryer (SOAFD).

The limit for permissible change (EQS) is defined as 'Annual values for the Treatment/Reference comparison should not significantly exceed ± x of the baseline for a consecutive period of greater than three years.'

(Here, exceedance is not meant to imply that a change is necessarily unacceptable, but simply that the perceived *status quo* has been breached).

Figure 29 shows that variability about the mean values for H' and E is reasonably similar up to 1983, but noticeably higher since then. For numbers of taxa, variability is similar across all years, in contrast to numerical abundance, where there appears to be a 'chance' element to between-sample variability in any one year. In general, variability in the untransformed data would appear to be unacceptably high for the detection of change within the specified limits. Alternative expressions of the data, e.g. log. transformations, may be more appropriate.

The pattern of changes in mean values for all measures is logical, when compared with those occurring at individual sites in Figure 28. For example, negative values for the ratio of taxa occur from 1982 onwards, and there is a marked switch to positive values of diversity and evenness ratios in 1984 and 1985, accompanying significant divergence of numbers of individuals at these times.

A marginal synchronous decline (i.e. at both Treatment and Reference sites) in, e.g. the Shannon diversity index, will result in a systematic change in % difference, if the decline is not proportional between sites. Thus a <u>natural</u> downward trend of this nature over a prolonged time interval could result in a breach of the limit. Possible options to prevent such an occurrence might include periodic 'up-dating' of the baseline value, or use of a running mean, provided there is no major dichotomy in trends at Treatment and Reference sites.

A large synchronous decline which is not in proportion between sites could give rise to a significant change in % difference:

e.g. Trt = 50 taxa; Ref = 40 taxa; (Trt/Ref - 1) x 100 = +25%
Trt = 30 taxa; Ref = 20 taxa; (Trt/Ref - 1) x 100 = +50%

A decline of such magnitude might seem less likely for taxa than for individuals. These hypothetical examples amply illustrate some of the limitations inherent in the use of any single-figure summaries of complex community data, and emphasise the importance of complementary analyses (both physico-chemical and biological) as aids to the identification of causes for observed change. (See also Section 2.2.1).

The relatively low numbers of individuals and taxa found at the Tees Bay stations are characteristic features of exposed, shallow sandy locations. (Contrast the numbers present at the Garroch Head disposal site). In these circumstances, a 100% change in a Treatment/ Reference comparison might signify a change at one site relative to another from, say, 100 to 50 individuals. This raises a question as to whether a % change of such magnitude would be more or less significant than one occurring in a more stable and diverse area which showed a drop from, say, 5000 to 2500 individuals. The example serves to highlight a general concern, relating both to the feasibility of detecting change of a given magnitude, and the attribution of significance, in contrasting community types.

(c) General comments:
The above approach is aimed at defining the *status quo*, and then monitoring for compliance by means of the fixing of arbitrary boundaries. It is perhaps best employed in conjunction with time-series plots of measures at individual sites, as in Figure 28. It should be noted that, in general, the statistical evaluation of data involving the use of ratios of variables should be approached with some caution because, apart from their obvious dependance on the relative rates of change in the numerator and denominator (see above), there is also a risk of generating spurious correlations as a result of the transformations (Jackson *et al*., 1990). When examining for trends in relation to other biotic or environmental data, alternative regression-based approaches, such as the analysis of covariance, may also be useful.

The approach does not directly address the acceptability of changes occurring and, as a result, any breaches of the limits for the chosen measures might best be viewed as **'Action Points'** for further investigations, rather than as formal EQSs.

Guideline values for assessments of acceptability of changes at the Garroch Head sewage-sludge disposal site were suggested in Section 2.2.1. Examples of complementary application of approaches both for the monitoring of compliance with the *status quo*, and for assessing the acceptability of observed changes, are given in Annexes 5-7 for the Tyne, Liverpool Bay and Thames sewage-sludge disposal sites. These provide a good insight into the likely future demands for field sampling and laboratory analysis arising from the adoption of such an approach. Such demands are likely to be particularly significant at the more dispersive areas. At all areas, it is clear that a good knowledge of waste dispersal pathways, accompanied by sound sampling design, will be critical to success.

Table 16. Outcome of attribute scoring exercise (see text)

Score: 1 = bad *2* *3* *4* *5 = good*	Sensitivity	Amenability to statistical testing	Effect of pooling (on sensitivity/ precision)	Ease of use	Ease of interpretation	Range of applicability (NB dispersive v. accumulating sites)
Multidimensional Scaling Ordination (MDS)	5.5.5.	2.2.5.	4.5.5.	3.5.3.	4.5.3.	4.3.5.
	5.5.5.	2.3.2.	4.4.5.	4.3.1.	4.3.3.	5.3.5.
	4.4.4.	3.2.3.	4.4.4.	3.2.3.	3.2.3.	4.5.4.
Species/Abundance/Biomass Comparison (SAB)	4.3.3.	4.3.1.	4.2.3.	4.4.5.	4.3.3.	3.2.3.
	5.4.3.	5.2.1.	4.4.3.	4.3.5.	5.2.4.	5.3.3.
	4.3.3.	2.5.3.	3.4.4.	5.4.5.	4.4.4.	5.4.3.
Abundance/Biomass Comparison (ABC)	3.3.3.	2.3.5.	4.2.5.	3.4.5.	3.3.5.	3.2.5.
	3.4.3.	1.2.1.	4.4.3.	2.3.4.	4.2.3.	4.3.3.
	4.3.3.	2.3.2.	3.3.3.	4.3.3.	4.3.4.	5.4.4.
Coefficient of Pollution (C of P)	1.1.1.	4.1.1.	1.1.1.	2.1.1.	2.1.1.	1.1.1.
	1.1.1.	2.4.1.	1.1.3.	1.1.3.	1.4.1.	1.1.1.
	1.1.1.	3.2.3.	1.2.	1.2.3.	2.2.2.	3.1.2.
Shannon diversity (H')	3.3.3.	5.4.5.	4.2.4.	5.5.5.	3.4.2.	4.3.5.
	3.4.4.	5.5.4.	1.3.4.	4.4.5.	3.1.4.	5.3.3.
	3.3.3.	4.3.3.	4.3.	4.3.3.	4.3.3.	4.3.4.
Pielou evenness (E)	3.3.3.	5.4.5.	4.2.4.	5.5.5.	3.4.2.	4.3.5.
	3.3.2.	5.3.1.	1.3.2.	4.3.5.	2.2.4.	5.2.1.
	2.3.3.	3.3.3.	4.3.	4.3.3.	4.3.3.	3.3.4.
Abundance (A)	3.4.2.	5.4.5.	5.2.5.	5.5.5.	4.5.1.	4.4.5.
	5.2.3.	5.2.4.	1.3.5.	5.4.5.	5.3.2.	5.4.4.
	2.2.2.	4.5.5.	2.	5.5.5.	5.2.4.	4.4.4.
Taxa (S)	4.2.3.	5.2.5.	4.4.4.	5.4.5.	5.2.3	4.2.5.
	5.1.4.	5.3.4.	4.1.3.	5.4.5.	5.2.4.	5.3.4.
	3.1.2.	4.5.5.	4.2.2.	5.5.5.	5.2.4	4.3.4.
Biomass (B)	3.2.3.	5.4.5.	5.5.5.	4.4.5.	3.4.1.	3.2.5.
	5.1.2.	5.3.4.	1.3.5.	3.4.5.	5.2.2.	5.3.3.
	2.2.2.	4.5.5.	2.	5.3.5.	5.2.4.	4.4.4.
Classification techniques (CLASS)	5.4.4.	2.2.1	4.4.4.	3.5.4.	4.5.3.	4.3.5.
	4.4.3.	4.2.1.	3.3.4	2.3.2.	4.3.2.	3.4.4.
	3.5.3.	3.3.3.	3.4.	3.4.4.	3.4.4.	4.5.5.

2.2.3 *Subjective exercise on scoring of analytical methods*

The objective of this exercise was to assess the merit of different analytical approaches to the assessment of benthic community structure (drawing from the experience of members of the Team), as a possible aid to the selection of a suite of methods best suited to the derivation of 'Environmental Quality Standards'. A list of attributes which were considered to be particularly important in such an evaluation was given in Section 2.1.3.

Individual scores for each category are given in Table 16, and a summary of the outcome is given in Figure 30. It must be re-emphasised that this exercise was, by its

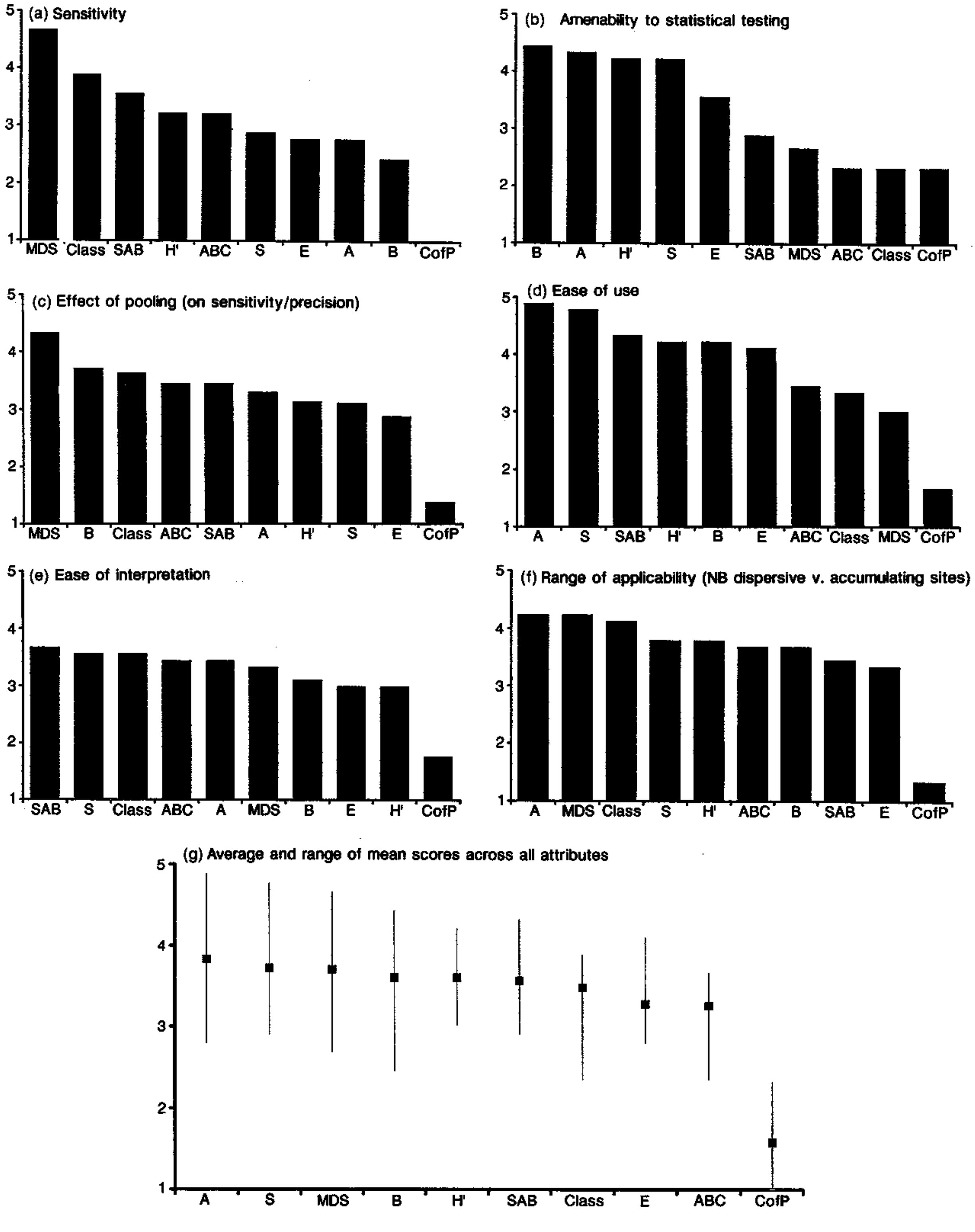


Figure 30. Plots of mean scores for various attributes (see text)

nature, **subjective** and the results should be interpreted in this light. It should also be noted that 'Multidimensional Scaling' was scored as one example of an ordination technique, and hence equivalence with other commonly used methods may be assumed, when considering aspects of its performance.

The Coefficient of Pollution was scored lowest for all attributes (arguably too harshly in some cases), and reflects the conclusions from analyses of test data by this means (see Section 2.2.1). It will be seen (Figure 30(a)) that multivariate techniques (MDS and Classification) were scored highest on sensitivity, as also were methods involving combinations of abundance, taxa and biomass (i.e. SAB and ABC). H' was also considered to have some merit. By contrast, univariate measures all scored highly in terms of amenability to statistical testing (Figure 30(b)). The views expressed in Figures 30(a) and (b) were to be expected, and highlight a need for development and/or greater 'accessibility' of methods for statistical testing of analytical techniques which involve combinations of variables.

The outcome of Figure 30(c) suggests that multivariate techniques (especially MDS), along with ABC, SAB and B are more robust to effects of pooling than the remaining univariate measures. The last, along with SAB, were all scored highly in terms of ease of use (Figure 30(d)), an outcome which doubtless can be linked to Figure 30(b) ('Amenability to statistical testing').

E and H' (and C of P) were perceived to be least easy to interpret (Figure 30(e)). This is of some interest, since the three are more complex in their derivation than simple A/S and B/A ratios characterising the 'SAB' approach. Other univariate/graphical measures consist of — or depend directly upon — the primary variables themselves. However (with the exception of C of P), the mean scores are not markedly different across all methods, which would seem to reflect a view that no single approach provides an easy route to the problem of pollution detection.

The range of scores in Figure 30(f) ('Range of applicability') is also relatively narrow; however, it is significant that methods/measures which are not rooted in formal hypotheses concerning pollution effects score somewhat higher. This no doubt reflects uncertainties about the nature of benthos responses to pollution in dispersive environments.

Figure 30(g) gives averages of scores for each method across all attributes. It will be clear that only C of P (as presently formulated) is considered to be seriously flawed as a means of data analysis. This is not unreasonable as an outcome, since most were perceived to have some merit at the outset of Task Team investigations, and the analysis of test data has largely confirmed these views.

2.3 Recommendations

1. Proposed limits for 'acceptable' change, or those designed to check for maintenance of the *status quo*, are best viewed as 'Action Points' rather than as formal EQSs since the wider environmental consequences (e.g. for commercial fish or man) arising from non-compliance are not rigorously defined in scientific terms, and may in any case vary from one locality to another.

2(a). The following methods are recommended for routine use in EQS derivation :

(i) univariate measures : Abundance (A), Taxa (T), Biomass (B) as ash-free dry weight;

(ii) ratios of univariate measures : A/T, B/A.

2(b). The following methods are recommended for routine use in EQS derivation, subject to further statistical evaluation of their mode of application to compliance monitoring :

(i) univariate measures : Shannon-Wiener diversity index ($H'log_2$);

(ii) multivariate methods : Multidimensional Scaling (MDS).

For the same numbers of samples taken in succeeding years, the precision of the means of the univariate measures will vary due to a combination of natural changes in benthic populations and field sampling error (quite apart from pollution influences). For pairwise comparisons, methods such as 'bootstrapping', may prove to be useful in deriving confidence limits. **In all cases, further evaluation of statistical methods for compliance monitoring is recommended.** (For example, the use of medians, transformations of data, non-parametric and trend-monitoring techniques). **The consequences of this evaluation for future field sampling strategy, and laboratory processing time, must also be addressed.**

In the case of biomass estimates, it may be necessary to adopt agreed procedures for the exclusion of 'large' species, as their occasional presence can greatly increase variability in the data.

3. The exclusion of methods such as clustering techniques and the Abundance/Biomass Comparison (ABC) from the above list is not meant to imply that these are not valued as **interpretative** tools. They should be freely employed in complementary investigations of the consequences of waste discharges; indeed, without such analyses, some of the changes observed in univariate measures may be open to misinterpretation. Future developments may improve their potential for application to standard-setting.

4. Compliance monitoring will normally require comparisons between locations which, **in principle**, should be subject to identical natural environmental influences. Monitoring for compliance with the *status quo* will involve the following arbitrary boundaries for change as guidelines, based on pairwise (i.e. [(Value at Location x/ Value at Location y)-1] x 100) comparisons, generally at stations outside the known sphere of influence of the waste discharge:

(i) Primary variables
A : 'baseline' value (%) ±50
T : 'baseline' value (%) ±20
B : 'baseline' value (%) ±20

(ii) Derived variables
H' : 'baseline' value (%) ±20
A/T : 'baseline' value (%) ±50
B/A : 'baseline' value (%) ±50

5. For well-placed stations peripheral to the sphere of waste influence, the outcome of the pairwise comparisons should, in principle, be zero, reflecting identical conditions at both. In practice, subtle differences will determine otherwise, and this explains why it will be necessary to establish a 'baseline' value for each variable, against which the significance of any future changes can be assessed. Ideally, this will consist of the mean of pairwise estimates from at least three annual sampling periods prior to commencement of the discharge. This will avoid any uncertainties associated with imprecise definition of the boundaries of waste influence, or with changes in the boundaries with time, e.g. in response to variation in the discharge. As such pre-disposal data are not usually available, a 'baseline' from at least three years of data should still be established, with an additional presumption that the responses to waste inputs at the nearest location, if present, should be stable and minor over this period.

6. The primary variables are of first importance in compliance testing and, as a general rule, a requirement if any for ameliorative action will be indicated only when pairwise comparisons of all three are significantly in breach of the designated limits. (It does not follow that such an event will indicate environmental degradation; indeed the reverse may be true if one of the two locations were to be influenced by a discharge which is decreasing in quantity). The derived variables (especially A/T) are designed to assist expert biological interpretation of changes in the benthic community, but are presently considered to be of secondary importance to compliance testing. (They are in any case integrals of the primary variables).

7. In order to assign limits for acceptable change, judgements will again have to be site-specific, but the following guideline values, derived from experience along organic enrichment gradients in *quiescent areas*, are recommended as a starting point :

(i) Primary variables
A : + 200% of reference value
T : + 50% of reference value
B : + 50% of reference value

(ii) Derived variables
A/T : + 100% of reference value
B/A : - 50% of reference value

Compliance-testing will involve the same pairwise approach as employed above, but in this case one of the locations will be well within the known sphere of waste influence (though generally outside any 'mixing zone') and hence will take the form:

[(Treatment/Reference)-1] x 100

Such an approach to environmental quality assessment emphasises the great importance which must be attached to site-selection. (As a general rule, this will be achieved from an assessment of physico-chemical characteristics).

8. The above values are considered to be indicative of a transition point from 'benign' to structural change in stable benthic communities in response to organic enrichment, and are quite rigorous when compared with the magnitude of changes encountered within the Garroch Head disposal site (see Section 2.2.1). Again, compliance testing will concentrate on the primary variables. Different values (in some cases incorporating negative rather than positive change for the first three measures) would be required if, for example, there was a need to monitor for improvement at locations within a 'mixing zone' which are already grossly affected, and hence in exceedance of levels of acceptable change set for peripheral areas. (On current knowledge, this should not apply to areas other than Garroch Head). This point emphasises the great importance of prior knowledge of the history of contamination at a location.

9. Some 'fine-tuning' of these values — especially those derived from ratios — will be required for local application, assuming that an adequate historical database is available.

10. Marginal breaches of standards of acceptable change (or of those defining the *status quo*) may be permissible for a period of three consecutive years, before action is required which may lead to further controls on the discharge. This is to allow for any anomalous short-term changes due to natural causes, but may be overriden if ancillary analyses provide conclusive evidence of adverse effects arising from the discharge within the three-year period. Accompanying qualitative appraisals of the biological data, along

with, e.g. multivariate techniques, will be particularly important during this period, since a sudden deterioration in conditions could lead to structural change in a community. At this stage, the trend in, e.g. numbers of taxa, may be downward and hence may not, on its own, signify that an unacceptable change had occurred.

11. Whether monitoring for change *per se*, or for the acceptability of change, marked directional trends in ratios of measures between stations may point towards an eventual degradation of benthic community structure, even though the 'Action Points' have not been breached. If sustained over a minimum period of three consecutive years, such trends may be used as triggers for more intensive investigation of possible causes.

12. Establishing criteria for acceptable change in 'high energy' environments is more difficult, since a number of the consequences of natural physical disturbance (e.g. high % dominance by one or two species, low diversity and high temporal variability) are analogous to those observed in stable assemblages in response to an advanced stage of organic enrichment. However, with adequate replication, monitoring for compliance with the *status quo* should be feasible. Wherever possible, stable ('depositional') assemblages in the general vicinity of the discharge should be sought for monitoring of compliance with criteria of acceptable change.

13. Annual monitoring is sufficient; the same sampling and analytical protocols at selected stations must be rigidly adhered to. Sampling must be conducted at the same time of year.

14. Sampling design will be site-specific, but at most sites will require increased replication at a limited number of 'representative' stations, or the pooling of stations (with or without replication). The number of samples required to achieve acceptable levels of precision will vary from site to site, and from time to time, and hence must be locally determined. For example, small-scale patchiness may result in highly variable abundances between samples at a location, which in combination will give rise to high variance and uncertainty. Surveys should nevertheless incorporate an element of redundancy to allow for retrospective analysis of additional samples, as improved precision in any one year may be achievable by this means.

15. It should be recognised that, in addition to modification of existing sampling strategies, a greater commitment of resources will be required for most areas in order to facilitate such an approach.

16. Periodically repeated spatial (grid-type) surveys will be essential to check the continued validity of representative stations.

17. Identification to the level of family may be adopted for such descriptive surveys. However, identification to the level of species should continue for EQS sampling programmes, so as to allow for further evaluation of the effect of taxonomic level on the sensitivity and precision of data, and of its significance for the setting of 'Action Points' (which presently relate to species-level data). More detailed guidance on this issue, along with recommendations, are given in Annex 1.

18. Investigations of effects of sieve mesh size on treatment/reference comparisons may lead to future site-specific refinements of limits for change, and are to be recommended.

19. The above methodology will ideally be complemented by future application of a biotic index, currently under investigation by the Water Research Centre (WRc).

3. A LABORATORY INTERCOMPARISON EXERCISE FOR SAMPLES OF MARINE BENTHOS: SUMMARY OF FINDINGS AND PROPOSALS FOR FUTURE WORK (*M. Elliott, D. Moore and I. Rees*)

3.1 Introduction

The necessity for analytical quality control (AQC) and quality assurance (QA) in field and laboratory benthic techniques was highlighted in Rees *et al*. (1990), along with initial suggestions for achieving these. They are important to ensure the production of valid data (thus leading to laboratory accreditation), and to allow intercomparisons of data either spatially or temporally. The areas identified as being of particular concern are:

(i) standardisation and checking of methods;
(ii) standardisation of taxonomic literature;
(iii) checking of the efficiency of identification;
(iv) inter- and/or intra-laboratory comparison exercises;
(v) intercalibration of data analysis.

In order to address some of these topics, the Task Team organised an inter-laboratory comparison exercise amongst 12 laboratories, many of which carry out the statutory monitoring of disposal grounds. The exercise was conducted in two stages. The first stage required each laboratory to sort, identify and enumerate specimens from one of a series of samples taken at a shallow sandy location off the North Wales coastline.

This provided an indication of the laboratory (i.e. methodological) variability, and the biological variability in the samples. The second stage involved the re-analysis of all samples by one laboratory (SOAFD, Aberdeen), as a means of separating these two sources of variability.

The exercise was designed to assess:

(i) the general capabilities of the laboratories;

(ii) the range of determinand values (times taken for each stage, faunal parameters, etc);

(iii) the value of the data, based on differing levels of taxonomic resolution;

(iv) taxonomic problems (e.g. synonymy, geographical unfamiliarity, lack of literature);

(v) taxonomic precision (the same name assigned each time the species is encountered) and accuracy (the correct name is used);

(vi) the time taken for analyses, and the quality of the data when multiple samples are being processed (i.e. the analytical time-variation with increasing familiarity of the fauna);

(vii) the value of benthic intercalibration exercises, and problems associated with their conduct.

As the Task Team was separately engaged in assessing the implications of carrying out benthic analyses at taxonomic levels higher than that of the species (see Section 2.2), the present exercise gave the opportunity to estimate the saving in processing-time which would result from such an approach.

3.2 Results

The results of the first stage of the exercise are given in Elliott (1990), and a final report is in preparation. One important lesson arising from the first stage was the need of individual laboratories to document the minutiae of analytical procedures, as there was evidently scope for a good deal of variability in approaches but these were not always adequately recorded.

Despite the fact that samples were only moderately diverse, queries could be attached to up to 40% of the taxa in the combined list from all laboratories. These related to the use of synonyms, to the assignment of specimens to a less-precise taxonomic level, or to mis-identifications. Although the use of a standard species directory was advocated as a means of overcoming some of these problems, its restricted availability caused difficulties for some participants.

The average times taken to recover the basic information during the first stage (Table 17) indicates that 207 minutes (± SD 93) were needed to sort and identify the sample to family level, compared to 386 (± 122) minutes to the highest level (usually species). Similarly, where biomass was estimated, the respective times were 325 and 547 minutes, although the latter may be somewhat inflated as they include times from laboratories which do not routinely determine biomass. These represent time-savings of 40-45%. As the sorting time remains the same irrespective of the taxonomic level, the time-saving for the identification to family was 60% in relation to that for species.

However, it would be a mistake to extrapolate from these results to routine laboratory undertakings, because of the unfamiliarity of the sample to most participants, and the fact that only a single sample was analysed. Thus during stage 2, one laboratory re-analysed 11 of the samples. The times taken for the identification and then biomass determination to family and species level were, respectively, 295 and 370 minutes for the first sample. However, this decreased to an average of 96 and 141 minutes for the last 5 samples, a reduction of over 60% in analytical effort. With greater familiarity of the fauna attendant on the analysis of multiple samples, processing to the family level produced a time-saving of 40% in relation to that for species.

It is evident from Table 17 that some stages of the analysis had high coefficients of variation. To some extent, this reflects the differing levels of expertise available, although different techniques were also employed.

The faunal parameters derived from stage 1 of the exercise (which reflect both laboratory and biological variability) varied markedly in the case of biomass, but little for numbers of taxa, abundance, diversity and evenness. However, this was also the case for the stage 2 samples. Thus, although the species lists differed between laboratories, this had little impact on most of the primary and derived faunal parameters. For example, the mean H' for stage 1 is 3.729, whereas that for stage 2 is 3.848, both with a similar coefficient of variation of approximately 8%.

However, multivariate analysis is likely to be more sensitive to taxonomic variability between laboratories, as similarity measures take account of individual assignations in their calculation. The high similarity among the stage 2 samples in relation to the stage 1 samples (Figure 31) indicates the relatively low accuracy obtained.

Table 17. Summary of the outcome of a laboratory intercomparison exercise (see text)

	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6
Times (mins.)						
Staining	180					10
Rough sort	90	105				
Check sort	30	90	190	210	200	
Ident. to family		60	60	40	225	240
Ident. to spp.	150	120	300	180	75	120
Biomass to family						
Biomass to spp.	180			150	70	
Reporting			90			
Total (sort, family)	120	225	250	250	425	240
Total (sort, family, biomass)						
Total (sort, spp.)	270	375	550	430	500	360
Total (sort, spp., biomass)	450			580	570	
Faunal Data						
No. quant. taxa	18	23	29	25	30	24
No. qual. taxa				1	1	2
No. quant. families	17	17	26	20	21	22
Abundance	87	73	114	110	100	111
Biomass (g)	25.75			14.970	68.17	
Sample volume (ml)	100	150	95	77		100

	Lab 7	Lab 8	Lab 9	Lab 10	Lab 11	Lab 12	mean	S.D.	%c.v.
Times (mins.)									
Staining							95	85	89
Rough sort		80		105	120		100	14	14
Check sort		10			30	25	98	83	84
Ident. to family	70	160	150	195			133	73	55
Ident. to spp.	270	430	90	45	300	70	179	114	64
Biomass to family		75					75	0	0
Biomass to spp.	30	10					88	66	76
Reporting	30						60	30	50
Total (sort, family)	70	250	150	300	150	25	207	93	45
Total (sort, family, biomass)		325					325	0	0
Total (sort, spp.)	340	680	240	345	450	95	386	122	32
Total (sort, spp., biomass)	370	765					547	134	25
Faunal Data									
No. quant. taxa	20	18	24	29	24	21	23.8	4.0	16.8
No. qual. taxa	6	2		1			2.2	1.8	81.8
No. quant. families	18	17	21	24	20	14	19.8	2.8	14.3
Abundance	129	73	75	92	110	57	94.3	18.1	19.2
Biomass (g)	31.90						35.2	20.0	56.8
Sample volume (ml)	190	35	50		165		106.9	49.1	45.9

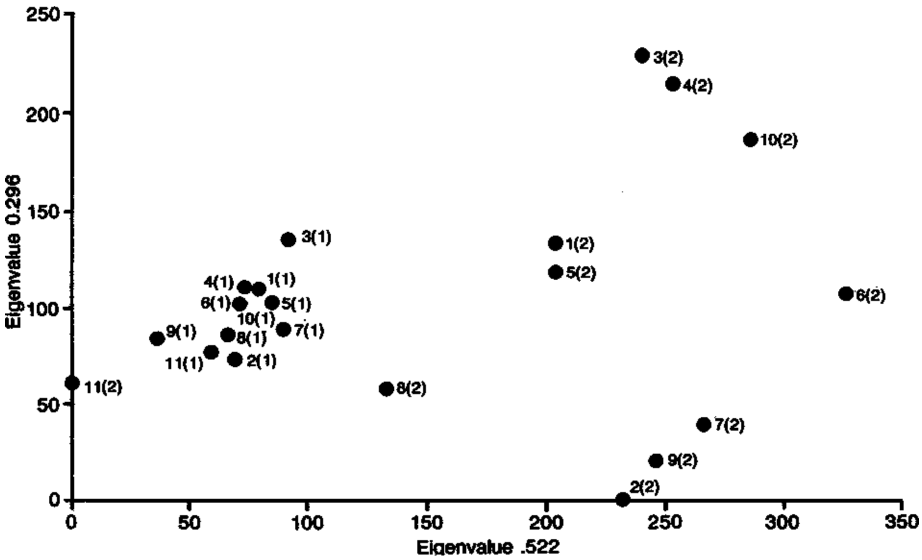


Figure 31. Outcome of the DECORANA ordination of the results from samples analysed for the Biology Task Team intercalibration exercise. [(1) denotes outcome of the second-stage analysis by SOAFD; (2) denotes outcome of first-stage analysis by different laboratories]

3.3 Recommendations

This exercise revealed that although precision amongst benthic laboratories is quite high, accuracy is less so. This clearly has wider repercussions, for example, concerning the use of biotic indices and 'Environmental Quality Standards', and the development of a central data archive. To date, QA/AQC measures have not been widely demanded of benthic laboratories. The following proposals for future improvements in inter-laboratory consistency are therefore made:

(i) maintain in-house checks on sorting efficiency; a routine 10% resorting has been proposed elsewhere (Chapman *et al.*, 1987);
(ii) carry out an independent re-analysis of a given proportion of samples (e.g. 5%) using inter- or intra-laboratory procedures;
(iii) adopt a standard methodology and maintain complete sorting, processing and reporting records for each sample;
(iv) circulate taxonomic references;
(v) adopt a standard species list;
(vi) attend workshops, maintain a specimen collection and employ independent specimen verification;
(vii) participate in ring-tests and other intercalibration exercises concerned with sample analysis and data interpretation.

It may be concluded that benthic intercalibration and laboratory intercomparison exercises are valuable as means of data validation and laboratory accreditation. However, such exercises require that the samples are re-analysed by an independent laboratory and hence these objectives could not be met without an adequate commitment of resources for this purpose. It should be noted that in many areas of work in the Water Industry, devotion of 20% of analytical time to AQC is becoming the accepted norm. Indeed, in some cases, this is a requirement to satisfy regulatory or accreditation organisations.

4. TAXONOMIC CODING OF DATA FROM MARINE BENTHOS SURVEYS AROUND THE UK COAST (*D. Moore and N. Shillabeer*)

4.1 Objectives

Ideally a coding system should fulfil the following criteria:

(i) Flexibility. It should allow easy modification to accept additions and taxonomic changes and should allow the inclusion of unidentifiable groups.

(ii) Information. It should be able to contain a range of information including various taxonomic levels and ecological characteristics.

(iii) It should produce station/species lists in taxonomic order.

(iv) It would be an advantage if the code was recognisable as a particular species, i.e. it had some component derived from the generic and specific names.

4.2 A review of coding systems presently in use at selected UK laboratories

The following organisations were contacted in order to gain a perspective on current practice:

(a) MAFF Burnham-on-Crouch
(b) SOAFD Aberdeen
(c) Plymouth Marine Laboratory (PML)
(d) Institute of Offshore Engineering (IOE), Heriot-Watt University
(e) ICI Brixham Laboratory
(f) NCC, Peterborough
(g) Forth River Purification Board (FRPB)
(h) Clyde River Purification Board (CRPB)
(i) Northumbrian Services Ltd
(j) BP Exploration
(k) Field Studies Council Research Centre (FSC), Fort Popton

The procedures employed were as follows:

(a) MAFF. Historically, a three-letter code was used for each species. This meant that additional species received a coding dissimilar from taxonomically related species. The system was not considered adequate, and they are currently using the MCS coding system.

(b) SOAFD. MCS system directly. Analysis of data in broader taxonomic groups would be a problem.

(c) PML. No system, each data set analysed has the species sequentially numbered.

(d) IOE. MCS is used as a guide to establish the taxonomic location of a species. The species is then coded with a 7 character code:

Phylum	Letter
Class	Letter
Family	2 Numbers
Genus	2 Numbers
Species	Letter

Additional characters would be required if Orders or other information needed to be included in the code.

(e) ICI. Species are coded with a 6 digit number; 2 additional digits are also used to define phyla and feeding type. The original species list was arranged in a taxonomic order based on the Plymouth Marine Fauna; new additions are placed in the MCS order. Originally all species were given a whole number and subsequent intervening species were defined using the numbers after the decimal point e.g.:

089.000 3 *Capitomastus minimus*
089.100 3 *Mediomastus fragilis*
090.000 3 *Capitella capitata*

This system is the basis upon which a package of programs operates. These enable the archiving, retrieval and manipulation of data files. The system cannot cope with analyses at different taxonomic levels: to do this, additional code numbers would need be included for the various levels.

(f) NCC. The MCS system is used as the basis of the coding system. However, because of the problems associated with simply using this coding, the computer system also allows the definition of other 'fields' in which additional coding, referring to information such as taxonomic level, may be included.

(g) FRPB. Species are defined in the data set using the first four letters of the specific and generic names. The computer system relates this listing to a species list containing the full name and the NODC (National Oceanographic Data Centre) code. It is this code that is used to order the species/station matrix. The NODC code is a 12 digit system, two digits being available for each of the taxonomic levels, Phylum, Order, Family, Genus, Species and variety. Further taxonomic levels can be defined. NODC codes are reassigned by the National Oceanographic Data Centre in Washington.

(h) CRPB. The complete list of encountered species has been arranged in taxonomic order and then numbered. Spaces have been left between each species to allow the inclusion of additional species. Code numbers range from 1 to 14805. CRPB had intended to use the NODC system but they received no reply when they asked for codings for species not currently listed.

(i) Northumbrian Services Ltd. Species are coded with an 8 character code:

Phylum	Letter
Class	1 number
Family	2 numbers
Genus	2 numbers
Species	2 numbers

This system allows analysis at taxonomic levels higher than species.

(j) BP. BP have operated the 14 digit Neptune system. However they no longer use any system and depend on their contractors to archive data. They found that the Neptune system was too labour-intensive. If they were to become involved with record keeping again they feel a system based on the MCS, together with coding to allow flexibility and analysis at different taxonomic levels, would be appropriate.

(k) FSC. Species were originally stored under species name and authority. More recently they have used the MCS system directly. They agreed that it needed modification to allow the coding of different taxonomic levels.

4.3 Conclusions

None of the systems currently employed by the laboratories listed above appears to fully satisfy the requirements of an ideal coding system. The FRPB and the IOE systems are the most advanced. The MCS Directory of species should be used to provide the taxonomic structure for any coding system to be employed. The use of a single coding system would probably be unrealistic since any 'global' coding system would be fairly unwieldy and would not lend itself to day-to-day use. We should therefore assume that the use of in-house coding systems which can be related to the global system through computer manipulation will become the norm.

In attempting to conceive a new system, based on the MCS Directory the following procedure was adopted:

(i) an alpha character was generated to represent the phylum. In fact there are more than 26 Phyla and this was the first pitfall;

(ii) a single-digit numeric code was applied at the Class level;

(iii) a two-digit code was applied at the family level.

At this point the coding would not work since there were more than 100 families in some of the Classes. Recoding at the Class/Sub-class level was then attempted, but in the case of the Crustacea this did not resolve the problem since sub-classes such as the *Pericarida* and the *Copepoda* still contained more than the 100 families which could be coded in this new system. The single digit Class/Sub-class code would also need to be upgraded to 2 digits.

Therefore 3 digits would be needed for families if we are to continue along this pathway. The new code would therefore have to have 10 digits to satisfactorily code to species level, or would have to be coded at an even lower level in the hierarchy i.e. Order, or would have to split certain of the Phyla. Since theoretically it should be mandatory to code all taxa using a single protocol this would, in practice, lead to a number of exceptions in the coding rules which might cause problems.

In-house computer analysts at SOAFD were consulted regarding the possible automatic coding of the MCS Directory from floppy disc. They have had long experience of coding systems and on balance they felt that adoption of an existing, proven system would be the route to follow. It was felt that the NODC system was the best presently available, despite its cumbersome format, since it is purely numeric and lends itself well to cross referencing in computer terms.

The following conclusions were therefore reached:

(i) a ten digit code will be needed to code to species satisfactorily;

(ii) this is the same as the ten digits required to code to the same level in the NODC system (the 12 digits mentioned in the original document allows for coding varieties);

(iii) in the long term a 'global', as opposed to national coding system must be used — NODC is furthest advanced in this and is therefore the obvious candidate;

(iv) the MCS Directory on disc will lend itself to being coded through direct computer programming with an indeterminate amount of manual checking, (it is estimated that up to 15% of the British taxa may not exist presently in the NODC system and it can be expected that a further 5-10% may be subject to varying nominations!);

(v) funding must be sought to do this work and time will be required for uncoded species to be coded by NODC.

4.4 Recommendations

It is recommended that the NODC coding system is adopted as a standard format for coding species data in monitoring work in relation to sewage-sludge monitoring, with the proviso that in-house coding systems which can be translated into NODC codes for national and international transmission of data should be permitted. The MCS Directory should still be valid for within-Phylum order and nomenclature but there may be some transposition of Phyla, a point which has been noted in the first edition and may be rectified in the proposed new edition.

5. DEVELOPMENT OF IMPROVED SAMPLING STRATEGIES FOR MONITORING THE EPIBENTHOS AT SEWAGE-SLUDGE DISPOSAL SITES (*H. Rees and M. Service*)

5.1 Background

The epibenthos comprises animals and plants living on — as distinct from within — the seabed, and they may be sedentary e.g. sea anemones, seaweed and hydroid colonies, or motile e.g. crabs, whelks and flatfish. Animals in the former category are typically filter-feeders, while the latter are typically carnivores or omnivorous scavengers. Some groups spend their entire adult life intimately associated with the seabed e.g. hydroids, most crabs and flatfish, while others may be only transiently associated e.g. shrimps and many fish species.

Individual or colonial epibenthic species are frequently the most conspicuous features of seabed sediments, and are certainly the most familiar to non-specialists and the public at large.

Subtidally, the most well-developed epibenthic assemblages normally occur on mixed substrates with a significant coarse component, where the range of micro-habitats can allow colonisation by a wide array of species. On soft substrates, a more restricted range of usually motile species are generally encountered; typically, these are dominated by brittle stars and hermit crabs, which may occur in very large numbers in the presence of a rich food supply.

There are several attributes of the epibenthos of subtidal habitats which can make this group an important target in environmental assessment. For example:

(i) they may have a controlling influence on infauna populations of soft substrates through predation;

(ii) on predominantly rocky areas or tide-swept grounds, they may be the only significant component of the benthos. Such areas may support an exceptionally high diversity and biomass of species, e.g. associated with subtidal mussel beds;

(iii) sedentary epibenthic species provide a direct route for carbon from the water-column to the seabed via filter-feeding; similarly, motile scavengers account for larger particles of settling detritus and other organic matter;

(iv) irrespective of mode of feeding, they can have a significant role in the bio-accumulation and then transfer of contaminants through the food-chain;

(v) many species are preyed upon by fish.

5.2 Sampling approaches

Because of the much wider size-range of organisms encountered compared with the infauna, small (0.1 m^2) grab samplers are generally unsuitable for quantitative assessment. Moreover, on mixed substrates or hard ground, grab sampling devices may operate at low sampling efficiency or not at all. However, heavy-duty grab samplers such as the Hamon grab (see Eleftheriou and Holme, 1984) may be useful for gravel substrates. This device has been employed by French workers in the English Channel, and is currently undergoing trials by MAFF.

A wide range of dredges and trawls have been devised for remote epibenthic sampling, with varying efficiency of organism retention (see e.g. Eleftheriou and Holme, 1984 and below). Epibenthic samplers can be very useful tools in initial exploratory surveys, while a 2-metre beam trawl was recommended for the sampling of the epifauna of soft ground at sewage-sludge disposal sites (see Rees *et al*., 1990). Adaptations to cover use on coarser substrates may be minor, e.g. the use of chafers to limit net damage, or major, e.g. the use of chain-link bags on heavy-duty frames such as scallop and rock dredges. However, as yet no single method can be recommended as a standard for obtaining quantitative data across all substrate types.

Examples of large-scale surveys where epifauna communities have been studied to good effect include Dyer *et al*. (1983), who examined the by-catch from commercial fish trawls throughout the North Sea. Frauenheim *et al*. (1989) also studied the larger epifauna of the North Sea using a 2-metre beam trawl, while Tiews (1983) reported on temporal trends in the epifauna by-catch from shrimp trawls in the German Wadden Sea. Studies of the epifauna at sewage-sludge disposal sites include those of MAFF (1990) for the outer Thames estuary.

Given adequate water clarity, tidal and weather conditions, diving probably provides the best means for quantitative assessment of the epifauna through a combination of direct observation and sampling, but these conditions are not typical for much of the UK coastline. For these and other logistical reasons, diving can be an expensive — and potentially hazardous — option for the conduct of regular offshore monitoring routines, although protocols for survey and sampling of inshore (mainly rocky) habitats by this means are well-defined in the UK, largely through the efforts of conservation interests (e.g. Hiscock, 1990).

Alternative methods for *in-situ* assessment include remotely-deployed underwater video and still photography. These, along with a range of other imaging methods, have recently been reviewed in Rumohr (1991); further detail is provided below. Again, water clarity is an important limiting factor but, in general, this option is likely to be cheaper and less weather-prone than diving surveys. Moreover, they may be operated in areas deeper than those normally accessible to divers. Ideally, a combination of *in-situ* observations by photography and efficient remote sampling of sediments offers the most promising combination for the development of routine monitoring programmes.

5.3 Underwater video techniques for surveying hard ground

Traditionally, the epibenthos of hard substrates has been sampled using a variety of dredges. Quantification of these techniques was usually based on an estimation of the distance travelled by the gear, either by use of the ship's electronic navigator, or from an odometer wheel attached to the gear. However, with this type of equipment the sampling efficiency is reckoned to be low, and much of the time the gear may in fact be 'flying' or bouncing above the seabed.

Remote-control underwater photography has been in use for a number of years to obtain static images of the seabed, and high quality images can be obtained which enable the identification of much of the macro-invertebrate fauna present. These images cover a small area of sea bottom and, while useful for gaining a 'feel' for the bottom conditions, do not give information on the overall distribution of faunal communities.

To allow wider coverage of the seabed, photographic and video cameras have been mounted on a variety of towed skids or sledges. By using a fixed frame, the area in view at any one time can be calculated and this, coupled with a knowledge of the distance covered in any one haul, allows transect-type studies to be made. Normally such sledges are fitted with both video cameras and still photographic cameras. This allows still photographs to be taken at areas which show up on the video camera as 'interesting', or at regular fixed distances.

Still photography is regarded as essential because, despite the improving quality of video, stills provide the sharper definition required for the identification of many species. Additionally, still photographs provide a convenient 'hard copy' for use in reports. Presently, the use of video printers is becoming more widespread; with further improvements in the technology, it is likely that these methods will eventually supersede the use of still cameras.

A spin-off from the development of the off-shore oil industry has been the production of relatively low cost Remotely Operated Vehicles (ROVs), on which a variety of cameras can be mounted. ROVs are self-propelled vehicles controlled by commands from the surface which are relayed down an umbilical cable which also carries the video and other telemetry signals.

The apparent advantage held by ROVs over towed vehicles is their manoeuvrability which offers the freedom to move in three dimensions. This should allow objects to be viewed from a variety of angles, and the vehicle can be stopped or moved back onto an object for further study. In addition, ROVs can be manoeuvred into areas which are otherwise inaccessible except by the use of divers. This is an especially valuable asset when studies are required of conditions beneath fish farms. In practice, however, the small, low-cost (i.e. <£30 000) ROVs are restricted by their limited ability to operate in currents in excess of 1.5 knots.

The area covered by ROVs is generally restricted by the length of umbilical and the water depth. They are not routinely deployed in situations where the surface tender is not anchored. This is not a restriction for towed systems, where tow length is dictated more by operational requirements. Quantification of the image obtained from ROV systems can be difficult as the vehicle will tend to fly at varying heights off the seabed, thus altering the area viewed in frame. This can be overcome by the use of ranging poles, or by keeping objects of a known dimension in the field of view.

A third alternative to the systems listed above comes under the form of the Remotely Operated Towed Vehicle (ROTV). Broadly speaking, this is a towed vehicle whose depth and attitude are controlled by rotors. Such systems allow quite fast towing speeds and the possibility of midwater observations. However the cost of the elaborate control systems required for these devices will tend to limit their use by smaller organisations.

The following standard procedures for the deployment of underwater camera systems can be recommended:

(i) underwater photographic systems should normally comprise of at least one video camera and a suitable quality still camera (preferably 35 mm for ease of processing);

(ii) where towed sledges are used, the field of view of each camera should be known from previous calibration;

(iii) the distance travelled by the sledge should be known, either from the use of the ship's electronic navigator or by the use of a meter wheel attached to the sledge;

(iv) towing should be at a constant speed;

(v) still photographs should be taken at fixed intervals either on a distance or on a time basis. These can be backed up by opportunistic shots taken of 'interesting subject matter' as identified on the video monitor;

(vi) where ROVs are used, the following information should be available or calculable — distance travelled, heading, height above seabed and field of view. Consideration should be given to the laying of fixed transect lines using weighted or leaded rope to act as a fixed reference point.

In addition, the slow towing speeds necessary to obtain high quality images when using towed sledges means that, at most UK disposal sites, transects will be run in the form of controlled drifts along the direction of the prevailing tidal current. It is recommended that a series of parallel transects be run in the centre, upcurrent and downcurrent of the disposal site. The data obtained can be treated at a number of levels which will be partially determined by the quality of the images obtained. Still photographs taken at regular intervals along the transect can be treated as point quadrants, the fauna identified to the appropriate taxonomic level and quantified. Data obtained by 'freezing' the video image at regular intervals can be treated in a similar manner. An alternative treatment is to make counts of prominent species, note faunal adjacencies and treat the data using 'run-like' statistics (see, for example, Caddy and Carter, 1984).

5.4 Future requirements

Significant gaps still remain in our understanding of the role of the epibenthos in marine ecosystems, and of their utility in pollution-monitoring. This is partly an historical consequence of sampling difficulties especially on mixed substrates, but more recent developments (particularly in underwater photography) offer scope for improvements in this area.

Presently, further effort needs to be expended on comparative assessments of the efficiency of currently available sampling methods before standard procedures covering the full range of substrates around the UK coastline can be recommended. There is also a need to assess the sensitivity of individual species and of whole communities to environmental changes, and from this to develop models for the responses of the epifauna to anthropogenic effects, including waste discharges. Subject to a successful outcome of this work, it is likely that the wider adoption of standard sampling and analytical procedures would be less demanding on time than traditional monitoring approaches involving the laboratory processing of the macro-infauna.

Future work requirements may therefore be summarised as follows:

(i) a strategic review of the literature on community studies of the epifauna with special reference to sampling methodology and environmental effects;

(ii) a comparison of the efficiency of remote sampling gear and imaging methods on a range of substrates representative of UK coastal and offshore habitats;

(iii) an assessment of the sensitivity of epifauna communities to waste discharges. This would require the deployment of a range of sampling methods in order to identify changes in community structure, and in the contaminant content of selected species. A key objective of this work would be the construction of an empirical model for anthropogenically-induced changes, which could then provide the framework for assessment of data from routine monitoring programmes;

(iv) an assessment of the implications of changes in the epifauna for (a) commercially exploited species, and (b) other ecosystem components;

(v) the development of standard procedures for cost-effective routine monitoring.

6. CONCLUSIONS

The first report of the Benthos Task Team was largely concerned with issues relating to the harmonisation of monitoring programmes at UK sewage-sludge disposal sites (Rees *et al*., 1990). The present report has concentrated on the identification of measures derived from field surveys of the benthic fauna which are best suited to the interpretation of changes, within a framework of Environmental Quality Objectives and Standards outlined in MAFF (1989).

The principle effects of organic enrichment on the benthos have been well documented, notably by Pearson and Rosenberg (1978), and this knowledge greatly facilitated the establishment of guideline values (or 'Action Points') for permissible change. However, it is important to recognise two difficulties with which the Task Team were confronted at the outset of this work :

(i) classical responses of the benthos along enrichment gradients are generally encountered in <u>quiescent</u> areas, and some uncertainty remains concerning the validity of applying such models to dispersive areas. However, against this it may be noted that there is very little evidence of <u>any</u> significant changes in benthic populations in response to sewage-sludge disposal in such areas, a finding which is of course compatible with the view that this disposal option is environmentally acceptable, on current scientific evidence;

(ii) it became clear that there was a paucity of suitable data against which to rigorously test criteria for permissible change. This was due to a combination of insufficient replication, and limited continuity in sampling over time. (It should be remembered that, to date, survey design at these areas has not been specifically aimed at the derivation of Environmental Quality Standards).

Important conclusions arising from this part of the report therefore include a future need for re-design of surveys to accommodate an EQS approach, a more detailed evaluation of the statistical requirements for compliance-testing, and the requirement for extended time-series data with adequate replication, to allow a more thorough assessment of variability in selected measures of biological change in response to natural and anthropogenic influences.

Presently, the issue of identifying benthic samples at taxonomic levels higher than that of the species is a contentious one for a variety of reasons specified in Annex 2. There was also some unease over the precedent that might be set by adoption of such an approach for applied studies generally, at a time of growing concern in some circles over the decline in UK taxonomic expertise.

On a narrower level, the outcome of analyses of test data sets provided no indication that interpretation of trends was seriously hampered by pooling to higher taxonomic levels, and it is also true that some time-saving will accrue from identification to coarser levels only. The final conclusion, namely that identification to species level should be maintained at sites selected for EQS derivation, reflected the view that a more rigorous statistical analysis of the effects of pooling on precision and accuracy of the data will be advisable.

The outcome of a laboratory intercomparison exercise involving the processing of a series of benthos samples proved to be valuable in highlighting aspects in need of future improvements, and this led to a series of recommendations being made. To date, there has been very little co-ordinated effort in this field, which is in marked contrast to, for example, analytical chemistry. It is therefore most important that the resource implications arising from regular participation are recognised, and this is especially true for those laboratories involved in the central co-ordination of such exercises.

A review of taxonomic coding systems revealed a surprising variety in use among UK laboratories, though none were considered to fully satisfy future needs. The recommendation to adopt the National Oceanographic Data Centre (NODC) system has the distinct advantage of conformity with an internationally recognised code. Further work will be required in the allocation of new codes for a range of UK species, and to this end the establishment of a formal link between NODC and SOAFD (Aberdeen) is being pursued.

Quantitative studies of the epibenthic fauna at marine waste disposal sites have traditionally been limited by sampling problems especially on mixed substrates. This component is frequently the most conspicuous element of the marine benthos, but the significance of its role in the food chain, and its sensitivity to anthropogenic effects are usually ill-defined. Future studies directed at improving sampling efficiency, and at the modelling of responses along pollution gradients, should allow the development of cost-effective monitoring approaches, and will have the further advantage of broadening the scope of biological impact assessments.

7. FUTURE RESEARCH AND DEVELOPMENT NEEDS

(i) Further development and then application to benthic community data of statistical methods for compliance monitoring, and evolution of a tailored statistical package.

(ii) Application of the NODC coding system to benthos data from ongoing surveys of UK sewage-sludge (and other) disposal sites. This will require the establishment of formal links with the US National Oceanographic Data Centre.

(iii) Evaluation of the potential of underwater photographic and other methods for quantitative surveys of the benthic epifauna at waste disposal sites, with particular emphasis on coarse deposits. The objectives of this work will be to assess the sensitivity of the epifauna to anthropogenic influences, to derive an empirical model for such changes (if appropriate) and then to establish protocols for routine monitoring, including compliance-testing against agreed standards.

(iv) Initiation of regular inter-laboratory comparisons of proficiency in sample processing, species identification and data analysis, and the establishment of a formal system of laboratory accreditation. (In this endeavour, it is most important that the resource implications for co-ordinating and — to a lesser extent — participating laboratories are recognised).

(v) Site-specific assessment of the effects of using different mesh sizes on the efficiency of sampling benthic macro-invertebrates from grab samples, and of the consequences for changes in primary and derived variables and hence compliance-testing.

(vi) Further evaluation of the effects of identification to higher taxonomic levels on compliance-testing measures.

(vii) Site-specific assessments of the utility of a biotic index in appraisals of environmental quality (including compliance-testing), subject to the outcome of present WRc research.

(viii) Assessment of the wider utility of recommended measures for compliance-testing at sewage-sludge disposal sites arising from the present report, e.g. in relation to pipeline discharges, or coastal classification schemes for assessing environmental quality.

(ix) Initiation of collaborative field and laboratory (mesocosm) studies, in order to enhance understanding of benthos responses to contaminants, and to improve predictive capability.

(x) Development of predictive models for benthic community structure using multivariate analyses, as an aid to interpreting anthropogenic influences.

8. REFERENCES

ANDREWS, H. P., SNEE, R. D. AND SARNER, M. H., 1980. Graphical display of means. Amer. Stat., ***34***: 195-199.

BOGDANOS, C., AND SATSMADJIS, J., 1985. Quantitive effect of sediment coarseness and depth on the macrobenthos of an unpolluted and closed Mediterranean Gulf. Rev. Int. Oceanogr. Med., 77-78, 74-85.

CADDY, J. F. AND CARTER, J. A., 1984. Macro-epifauna of the lower Bay of Fundy - observations from a submersible and analysis of faunal adjacencies. Can. Tech. Rep. Fish. Aquat. Sci., No. 1254.

CHAPMAN, P. M., DEXTER, R. N. AND LONG, E. R., 1987. Synoptic measures of sediment contamination, toxicity and infaunal community composition (the Sediment Quality Triad) in San Francisco Bay. Mar. Ecol. Prog. Ser., ***37***: 75-96.

CLARKE, K. R., 1990. Comparisons of dominance curves. J. Exp. Mar. Biol. Ecol., ***138***: 143-157.

DYER, M. F., FRY, W. G., FRY, P. D. AND CRANMER, G. J., 1983. Benthic regions within the North Sea. J. Mar. Biol. Ass. U.K., ***63***: 683-693.

EFRON, B. AND TIBSHIRANI, R., 1986. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. Statist. Sci., ***1***: 54-77.

ELEFTHERIOU, A. AND HOLME, N. A., 1984. Macrofauna techniques. *In*: Holme, N. A. and McIntyre, A. D. (Eds.), 'Methods for the study of marine benthos'. Oxford: Blackwell Scientific Publications (2nd. ed), 140-216.

ELLIOTT, M., 1990. Analytical intercomparison exercises in marine benthic biology. Forth River Purification Board, Report No. TW14/90, 15 pp.

FRAUENHEIM, K., NEUMANN, V., THIEL, H. AND TURKAY, M., 1989. The distribution of the large epifauna during summer and winter in the North Sea and its suitability for environmental monitoring. Senckenbergiana marit., ***20***: 101-118.

GRAY, J. S., CLARKE, K. R., WARWICK, R. M. AND HOBBS, G., 1990. Detection of initial effects of pollution on marine benthos: an example from Ekofisk and Eldfisk oilfields, North Sea. Mar. Ecol. Prog. Ser., ***66***: 285-299.

HEAP, J. M., ELLIOTT, M. AND RHEINALLT, T. ap, 1991. The marine disposal of sewage sludge. Ocean and Shoreline Management, ***16***: 291-312.

HISCOCK, K., 1990. Marine Nature Conservation Review: methods. Nature Conservancy Council, CSD Report No. 1072.

JACKSON, D. A., HARVEY, H. H. AND SOMERS, K. M., 1990. Ratios in aquatic sciences: statistical shortcomings with mean depth and the Morphoedaphic index. Can. J. Fish. Aquat. Sci., ***47***: 1788-1795.

LAMBSHEAD, P. J. D., PLATT, H. M. AND SHAW, K. M., 1983. The detection of differences among assemblages of marine benthic species based on an assessment of dominance and diversity. J. Nat. Hist., ***17***: 859-874.

MCINTYRE, A. D., ELLIOTT, J. M. AND ELLIS, D. V., 1984. Introduction: design of sampling programmes. *In*: Holme, N. A. and McIntyre, A. D. (Eds.), 'Methods for the study of marine benthos'. Oxford: Blackwell Scientific Publications (2nd ed), 1-26.

MAFF, 1989. First report of the Marine Pollution Monitoring Management Group's Co-ordinating Group on Monitoring of Sewage-Sludge Disposal Sites. Compiled by J E Portmann. Aquat. Environ. Monit. Rep., MAFF Direct. Fish. Res., Lowestoft, (***20***): 1-64.

MAFF, 1990. Disposal of sewage sludge from ships: Thames estuary. *In*: Franklin, A. (compiler), Monitoring and surveillance of non-radioactive contaminants in the aquatic environment, 1984-1987. Aquat. Environ. Monit. Rep., MAFF Direct. Fish. Res., Lowestoft, ***(22)***: 33-39.

MAURER, D. AND HAYDOCK, I., 1989. Coefficient of Pollution: Palos Verdes California Shelf 1973 and 1984. Mar. Pollut. Bull., ***20***: 219-222.

MAURER, D., ROBERTSON, G. AND HAYDOCK, I., 1991. Coefficient of Pollution (p): the southern California shelf and some ocean outfalls. Mar. Pollut. Bull., ***22***: 141-148.

NORTON, M. G., EAGLE, R. A., NUNNY, R. S., ROLFE, M. S., HARDIMAN, P. A. AND HAMPSON, B. L., 1981. The field assessment of effects of dumping wastes at sea: 8. Sewage sludge dumping in the outer Thames estuary. Fish. Res. Tech. Rep., MAFF Direct. Fish. Res., Lowestoft, ***(62)***: 1-62.

PEARSON, T. H., 1987. Benthic ecology in an accumulating sludge-disposal site. *In*: Capuzzo, J. M. and Kester, D. R. (Eds.), 'Oceanic processes in marine pollution. Vol. 1. Biological processes and wastes in the ocean'. Florida: R. E. Krieger, 195-200.

PEARSON, T. H. AND ROSENBERG, R., 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanography and Marine Biology: an Annual Review, ***16***: 229-311.

PEARSON, T. H., DUNCAN, G. AND NUTTAL, J., 1982. The Loch Eil Project: population fluctuations in the macrobenthos. J. Exp. Mar. Biol. Ecol., ***56***: 305-321.

PIELOU, E. C., 1966. Species-diversity and pattern diversity in the study of ecological succession. J. Theoret. Biol., ***10***: 370-383.

REES, H. L., MOORE, D. C., PEARSON, T. H., ELLIOTT, M., SERVICE, M., POMFRET, J. AND JOHNSON, D., 1990. Procedures for the monitoring of marine benthic communities at UK sewage sludge disposal sites. Department of Agriculture and Fisheries for Scotland. Scottish Fisheries Information Pamphlet, (***18***): 78pp.

REES, H. L., ROWLATT, S. M., LAMBERT, M. A., LEES, R. G. AND LIMPENNY, D. S., 1992. Spatial and temporal trends in the benthos and sediments in relation to sewage sludge disposal off the north east coast of England. ICES J. Mar. Sci., ***49***: 55-64.

RUMOHR, H. (Ed.), 1991. Imaging methods in benthic ecology. *In*: Keegan, B. (compiler), COST 647 Coastal Benthic Ecology, Activity Report, 1988-1991. Brussels: Commission of the European Communities, 310-351.

SANDERS, H. L., 1968. Marine benthic diversity: a comparative study. Amer. Nat., ***102***: 243-282.

SATMADJIS, J., 1985. Comparison of indicators of pollution in the Mediterranean. Mar. Pollut. Bull., ***16***: 395-400.

SHANNON, C. E. AND WEAVER, W., 1949. The mathematical theory of communication. Urbana: University of Illinois Press, 117 pp.

SHILLABEER, N. AND TAPP, J. F., 1990. Long-term studies of the benthic biology of Tees Bay and the Tees estuary. Hydrobiologia, ***195***: 63-78.

SKALSKI, J. R. AND McKENZIE, D. H., 1983. A design for aquatic monitoring programs. J. Environ. Management, ***14***: 237-251.

TALBOT, J. W., HARVEY, B. R., EAGLE, R. A. AND ROLFE, M. S., 1982. The field assessment of effects of dumping wastes at sea: 9. Dispersal and effects on benthos of sewage sludge dumped in the Thames estuary. Fish. Res. Tech. Rep., MAFF Direct. Fish. Res., Lowestoft, (***63***): 1-42.

TIEWS, K., 1983. Uber die Veranderungen im Auftreten von Fischen und Krebsen im Beifang der deutschen Garnelen-fischerei wahrend der Jahre 1954-1981 - Ein Beitrag zur Okologie des deutschen Wattenmeeres und zum biologischen Monitoring von Okosystemen im Meer. Arch. Fisch-Wiss., ***34***: 1-156.

WARWICK, R. M., PEARSON, T. H. AND RUSWAHYUNI, 1987. Detection of pollution effects on marine macrobenthos: further evaluation of the species abundance/biomass method. Mar. Biol., ***95***: 193-200.

ANNEX 1. Members of GCSDM Benthos Task Team

Mr I Codling
Water Research Centre
Medmenham Laboratory
Henley Road
Medmenham
PO Box 16
Marlow
Buckinghamshire SL7 2HD

Dr M Elliott
Department of Applied Biology
The University
Hull
HU6 7RX

Dr S Hull
Forth River Purification Board
Tidal Waters Section
Heriot-Watt Research Park
Avenue North
Riccarton
Edinburgh
EH14 4AP

Mr D Moore (Secretary)
Scottish Office (Agriculture and Fisheries Department)
Marine Laboratory
PO Box 101
Victoria Road
Aberdeen AB9 8DB
Scotland

Dr T H Pearson
Scottish Environmental Advisory Services
c/o Dunstaffnage Marine Laboratory
PO Box 3
Oban
Argyll PA34 4AD
Scotland

Mr J Pomfret
Analytical and Environmental Services Ltd
Biology Laboratory
Northumberland Yard
East Howdon
Wallsend
Tyne and Wear NE28 OQD

Mr E I S Rees
University College of North Wales
School of Ocean Sciences
Marine Science Laboratories
Menai Bridge
Gwynedd LL59 5EY

Dr H L Rees (Chairman)
Ministry of Agriculture, Fisheries and Food
Fisheries Laboratory
Remembrance Avenue
Burnham-on-Crouch
Essex
CMO 8HA

Dr M Service
Aquatic Sciences Research Division
Dept of Agriculture (Northern Ireland)
Newforge Lane
Belfast
Northern Ireland

Mr N Shillabeer
Zeneca Ltd
Brixham Environmental Laboratory
Freshwater Quarry
Brixham
Devon TQ5 8BA

Dr R M Warwick
Plymouth Marine Laboratory
Prospect Place
The Hoe
Plymouth
Devon PL1 3DH

ANNEX 2. Identification of benthic organisms at higher taxonomic levels: implications for future monitoring

A2.1 Background

Identification of freshwater benthic organisms to the family level — or above — has become a familiar routine in Water Industry surveys of environmental quality (though we understand that there has been a recent trend back towards species-level identification). In the marine environment, it is known that gross changes in the macrofauna in response to anthropogenic effects can also be detected at coarse taxonomic levels (e.g. Rosenberg, 1972; Ferraro and Cole, 1990). However, it has become accepted practice — in both pure and applied investigations — to identify marine organisms to the species level.

The effect on data quality of pooling species of both macrofauna and meiofauna into progressively higher taxonomic levels has recently been studied in the course of a Workshop on biological effects techniques (Warwick, 1988(a); Heip *et al*., 1988). In the case of meiofauna (animals passing through 0.5 mm mesh sieves), identification to higher taxon levels holds the promise of wider use of this group in applied studies, since the advanced taxonomic skills required for identification of individual species, and the time required to do so, is presently a severe constraint (Warwick, 1988(b); Herman and Heip, 1988).

Recent publications (mainly those above) all appear to support the idea of pooling at least to the family level; indeed, Warwick (1988(a),(b)) suggests that taxonomic units higher than the level of species may be more appropriate to studying responses of pollution.

For a variety of reasons specified below, it is <u>essential</u> to establish with confidence that there will be no adverse implications arising from the adoption of such an approach at sewage-sludge disposal sites, since recourse to a full species list in order to check on any ambiguities that may arise from future data analyses will not be possible.

A2.2 Implications for macrobenthos surveys at sewage-sludge disposal sites

Effect on precision/accuracy of results: we have no reason to suppose that the conclusions of the above workers will not apply equally to studies at sewage-sludge disposal sites. However, the opportunity to validate the procedure was afforded by analysis of test data sets from Garroch Head, Forth/St Abbs Head and the Tyne (see Section 2 of main report). Such an assessment was necessary to counter criticism that identification to higher taxonomic levels was not simply an attempt to raise the detection limits for identification of biological effects.

The grouping of species into higher taxonomic levels is inconsistent between major groups of organisms e.g. the levels of discrimination between major crustacean groups, which have innumerable points of surface morphological distinction, are much more detailed and thorough than discrimination within the polychaeta, where distinctions even at the class level are sometimes arguable. This leads to major imbalances in the numbers of major divisions within these important groups which can lead to bias when the relative proportion of such groups changes along gradients of enrichment.

Changes at the higher taxonomic levels are very dependant on the presence or absence of the minor phyla in the samples e.g. the *Platyhelminthes*, *Echiuroidea*, *Sipunculoidea* etc. Such phyla are represented only by occasional and rare individuals and their inclusion in a sample is often a matter of chance at the usually low levels of sample numbers included in the typical benthic survey. The significance of this effect will clearly depend on the taxonomic level at which the analysis is being conducted, e.g. it will have less impact at the family than at the class level.

Effect on historical comparisons: for the analysis of temporal trends, access to the raw data from earlier studies will be required, so as to permit species counts to be pooled to the appropriate taxonomic level. This should not be problematic for previous sewage-sludge disposal surveys done by the same organisation, but difficulties may arise in obtaining access to historical data for surveys done by a different organisation, or for access to relevant data produced in an entirely different connection (e.g. academic studies in the published literature; see below).

Effect on comparisons with the published literature: in the absence of species-level data, it may be very difficult to apply relevant published information concerning individual species to the results of local surveys; in this sense, a substantial body of reference literature could be considered closed. This literature might include ecotoxicological studies, which are conventionally conducted on individual species. Indeed, it seems highly improbable that such methodology could credibly be adapted for tests on (unspecified) species-groups, rather than on named species. Thus the developing alliance of specialists in field ecology and ecotoxicology could be hindered.

Effect on the speed of processing of benthos samples: this is potentially the most attractive aspect of conducting identifications to higher taxon level, since it would facilitate the handling of more samples from disposal site surveys. Information from an inter-laboratory comparison carried out under Task Team auspices (Section 3 of the main report) suggests that significant time saving can accrue from identification to the level of family only. However, it is important to note that the difference in required effort between family- and species-level identification declines in proportion to the number of samples processed from a given locality. This is because familiarity is soon gained with the bulk of species occurring in an area, with the result that most can then be rapidly assigned without further time-consuming recourse to taxonomic keys.

It follows from the above that the time saved during processing of samples from UK sewage-sludge disposal sites is unlikely to be dramatic, in cases where there is continuity in responsible personnel.

Effect on the requirement for benthic expertise at the laboratory bench: is recognition to family level a substantially less skillful undertaking? It is likely that some groups, e.g. polychaete worms and amphipod crustaceans, will become less demanding overall, but in most cases the range of taxa encountered — even at the family level — could still be considerable. For example, the pooling of data to family level for 29 samples taken at the Tyne sewage-sludge disposal site reduced the number of taxonomic units by about 100%, but this still resulted in the presence of 94 taxa.

Job satisfaction: identification to the level of species often carries with it a sense of achievement which can compensate for the more tedious elements in sample processing which are well known to practitioners. An added incentive is that from time to time the finding of a rare or unusual specimen can make a real contribution to scientific knowledge (the public relations value of such findings have not been lost by institutions who have been responsible for such work).

Effect on requirements for taxonomic keys: it is possible that condensed versions concentrating only on families may be all that is required for Water Industry laboratories in the future. (Clearly, this is somewhat at odds with recent advice of the Task Team on taxonomic issues: see Rees *et al*., 1990).

Effect on other pollution monitoring programmes (i.e. other than sewage sludge). Such evidence as is presently available would suggest that the efficiency of pollution detection at higher taxonomic levels would not vary with the nature of the discharge. It may therefore be assumed that a positive recommendation of the Task Team would have wider consequences for the pollution monitoring field. However, it could be argued that such an approach would be less appropriate in highly stressed environments, e.g. estuaries.

Effects on contribution of applied studies to the wider field of benthic ecology. In the narrow sense, this is not a matter of concern for waste disposal managers; however, there are many examples where properly conducted monitoring programmes have provided insights into natural ecological processes, and — given data of the right quality — can readily draw from the results of 'pure' studies. The degree to which such activities would be hindered by a reduction in taxonomic effort is clearly a matter for debate (see above). Retention of specimens to be available for use in future research is recommended to reduce any potential problems in this area.

A2.3 Options

Assuming that the ability to identify effects of waste disposal is not significantly affected by identification of benthic organisms at or above the level of the family (as appears to be the case from the analyses in Section 2 of the main report), then possible options for future work include the following:

(i) identify to an appropriate higher taxonomic level (e.g. to family) from now on. The importance of working to a standard taxon list, at whatever taxonomic level(s) are chosen, should be stressed;

(ii) carry out full species appraisal on initial 'baseline'; thereafter, identify only to (e.g.) family level. However, it may be important to retain samples to allow a more rigorous examination, in the event that such information may become relevant at the 'expert interpretation' stage;

(iii) identify taxa which are most relevant in a pollution context: continue to identify a proportion of these to species as seems appropriate (but do we have adequate knowledge?);

(iv) continue as at present, i.e. identification of all species. This might be at the expense of a facility to handle more samples, and would certainly ignore the growing body of evidence suggesting that identification to higher taxonomic levels is an acceptable compromise.

A2.4 Conclusions

Evidence from the literature, along with Task Team work on test data sets (Section 3 of main report) appears to favour the view that identification at least to the level of family is unlikely to prejudice the ability to detect well-defined anthropogenic effects.

However, it is important to stress that while the option of pooling to higher taxonomic levels may be useful for routine monitoring surveys, this does not imply that data at a more detailed taxonomic level will never be required. Taxonomy to species level may, for example, be more important in estuary monitoring, or in follow-up studies at localities where a decline in environmental quality has been identified from routine surveys. It may also be very important where historical comparisons from the same locality are being drawn, or when comparisons are necessary with the findings of published studies from other regions.

One potential advantage of pooling is that any time saved in sample processing will help to accommodate increases in sample numbers, along with additional requirements such as measurements of biomass and inter-calibration exercises that will be necessary for successful implementation of an EQS approach.

It is therefore recommended that identification to the level of the family may be adopted for descriptive surveys which are periodically conducted in support of EQS sampling programmes at sewage-sludge disposal sites. At sites chosen for EQS derivation, identification should for the time being continue to the level of species, so that further comparative testing of the effects of pooling on the sensitivity of selected measures can be conducted, and to allow for any adjustments to be made to the proposed limits for permissible change which have been developed from species-level data.

A2.5 References

FERRARO, S. P. AND COLE, F. A., 1990. Taxonomic level and sample size sufficient for assessing pollution impacts on the Southern California Bight macrobenthos. Mar. Ecol. Prog. Ser., ***67***: 251-262.

HEIP, C., WARWICK, R. M., CARR, M. R., HERMAN, P. M. J., HUYS, R., SMOL, N. AND VAN HOLSBEKE, K., 1988. Analysis of the comunity attributes of the benthic meiofauna of Frierfjord/Langesundfjord. Mar. Ecol. Prog. Ser., ***46***: 171-180.

HERMAN, P. M. J. AND HEIP, C., 1988. On the use of meiofauna in ecological monitoring: who needs taxonomy? Mar. Pollut. Bull., ***19***: 665-668.

REES, H. L., MOORE, D. C., PEARSON, T. H., ELLIOTT, M., SERVICE, M., POMFRET, J. AND JOHNSON, D., 1990. Procedures for the monitoring of marine benthic communities at UK sewage sludge disposal sites. Department of Agriculture and Fisheries for Scotland. Scottish Fisheries Information Pamphlet, (***18***): 78pp.

ROSENBERG, R., 1972. Benthic faunal recovery in a Swedish fjord following the closure of a sulphite pulp mill. Oikos, ***23***: 92-108.

WARWICK, R. M., 1988(a). Analysis of the community attributes of the macrobenthos of Frierfjord/ Langesundfjord at taxonomic levels higher than the species. Mar. Ecol. Prog. Ser., ***46***: 167-170.

WARWICK, R. M., 1988(b). The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. Mar. Pollut. Bull., ***19***: 259-268.

ANNEX 3. Non-metric Multidimensional Scaling Ordination (MDS): description of method *(R. Warwick)*

A3.1 Background

Multivariate methods of data analysis are known to be much more sensitive in detecting community change than are univariate measures, e.g. diversity indices, or graphical/distributional methods, e.g. dominance curves (Warwick and Clarke, 1991). If the criterion of 'no community change' is to be adopted in EQS evaluation, then the most sensitive techniques should be employed. Non-metric MDS (Kruskal and Wish, 1978) has been widely used and tested in benthic community studies, and has conceptual advantages over other methods in that it does not make the (usually invalid) assumption of multivariate normality in the data. It is recommended as one of the best (perhaps *the* best) ordination technique (e.g. Everitt, 1978; Kenkel and Orloci, 1986).

A3.2 Method

The starting point is a (dis)similarity matrix between samples constructed from the stations/species matrix of appropriately transformed species abundance or biomass data, usually employing the Bray-Curtis similarity measure. The ordination depends only on the RANKS of similarities in the triangular matrix, attempting to construct a sample 'map' (usually in two dimensions) using information in the form 'Sample 1 is closer to Sample 4 (in species composition) than it is to Samples 2 or 3', as in the simple example at Figure A3.1.

The MDS plot is arbitrarily scaled, located, rotated or inverted: it gives the positions of samples *relative* to each other. In the above example it is not difficult to place the points in two dimensions preserving the rank order of dissimilarities exactly. Usually this is not possible and there will be some distortion or 'stress' between the ranked dissimilarities and corresponding distances in the plot.

The MDS algorithm is an iterative process:

(i) specify number of dimensions for MDS plot (=m);

(ii) construct a 'starting map' of n samples (usually a random set of points);

(iii) regress interpoint distances from this map on the corresponding dissimilarities (non-metric MDS uses monotonic (increasing) regression (the Shepard diagram);

Year:	**64**	**68**	**71**	**73**
Sample:	**1**	**2**	**3**	**4**
Species				
Echin.	1.7	0	0	0
Myrio.	2.1	0	0	1.3
Labid.	1.7	2.5	0	1.8
Amaea.	0	1.9	3.5	1.7
Capit.	0	3.4	4.3	1.2
Mytil.	0	0	0	0

Similarities

Sample	**1**	**2**	**3**	**4**
1	-			
2	25.6	-		
3	0.0	67.9	-	
4	52.2	68.1	42.0	-

Rank dissimilarities

Sample	**1**	**2**	**3**	**4**
1	-			
2	5	-		
3	6	2	-	
4	3	1	4	-

MDS plot

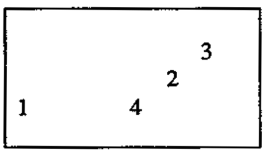


Figure A3.1. ***Summary of procedures involved in generating an MDS plot (the data matrix is a subset of counts from macrofauna surveys of Loch Linnhe)***

(iv) measure goodness-of-fit of the regression ('stress'). Stress = 0 if the distances preserve the rank order of dissimilarities exactly, stress is large if the current map is poorly related to the dissimilarities;

(v) perturb current sample positions on the map, in directions decreasing the stress, using steepest descent algorithm;

(vi) repeat steps (iii) to (v) until no further reduction in stress is possible. Since the algorithm is an iterative procedure it could converge to a local minimum rather than a global minimum of this stress function. Therefore, the procedure is repeated for different random starting configurations to confirm that these give the same solution (with the lowest stress value) several times. Generally, in two dimensions:

Stress <0.05 implies excellent representation,
<0.01 good,
<0.2 still useful, but
>0.3 little better than random points.

A3.3 Significance testing between sample groups

A3.3.1 Requirements

(i) A statistic which is a measure of the difference in benthic community structure between replicate samples taken in a potentially impacted area and a control or reference area.

(ii) Assessment of the statistical significance of the difference between impacted and control sites.

A3.3.2 Approach

Randomisation/permutation tests are appropriate for testing the significance of differences in community structure between groups of sites selected on an *a priori* basis, e.g. between replicate stations at impacted (I) sites and control (C) sites. The computer program ANOSIM (Clarke and Green, 1988) calculates a statistic (R), the value of which is 1 if all I replicates are more similar to each other than to any C replicate and all C replicates are more similar to each other than any R replicate. R has a value of 0 if there are no differences between I and C stations. The I and C labels are then randomly allocated to the sample points and the statistic recomputed a large number of times (say 500 or 1000).

The test will reject the null hypothesis of no differences between sites at the 5% significance level if the observed statistic is greater than its value for 95% of the random labellings. The test operates on the original similarity matrix and the statistic is based on differences in average rank dissimilarities between and within sites. To achieve a significance better than 5% requires a minimum of 4I and 4C replicates (giving 35 permutations).

A3.3.3 Example

None of the test data sets from sewage-sludge disposal areas had adequate replication (see Section 2). However, for illustrative purposes, data from the Tyne transect (1986) were grouped as follows:

Stations 1-4 (11 samples): Close to disposal site (D)
Stations 5-8 (12 samples): Mid-distance from disposal site (M)
Stations 9 and 10 (6 samples): Far from disposal site (F)

The source data file is the triangular Bray-Curtis similarity matrix of root transformed species abundance data.

A3.3.4 Results

Size of random sample = 500.
Approx. possible number of permutations = 6.423E+011
No. of significant statistics = 1
Significance level of sample statistic 0.535 is 0.200 % Approx 95% C.I. is (0.03%, 1.16%).

A3.4 Pairwise effects tests

Station groups	Statistic value	% Significance level	95% Confidence interval	Total possible permutations
(M, D)	0.49	0.20	(0.03-1.16)	1.352E+006
(F, D)	0.83	0.20	(0.03-1.16)	1.238E+004
(F, M)	0.32	0.20	(0.03-1.16)	1.856E+004

These results therefore show that the D and F samples are most dissimilar (R=0.83), followed by the D and M samples (R=0.49) and that the M and F samples are most similar to each other (R=0.32). However, all sample groups are significantly different from each other (significance = 0.2%). If the disposal site becomes more different from the control sites over time, this will be reflected in an increased R value. A two-way version of the ANOSIM test can also be used to assess the significance of changes in both control and impacted sites over time.

At present, there is no statistical testing framework for assessing the significance of differences between R values.

A3.5 Relating MDS configurations to environmental variables

Two approaches are possible:

(i) superimpose symbols scaled in size to represent the measured values of environmental variables onto the faunistic two dimensional MDS configuration (instead of the station numbers). Determine whether there is a visual correlation between the station groupings and the variable under consideration (Field *et al.*, 1982);

(ii) perform a correlation based Principal Components Analysis (PCA) on standardised values for environmental variables in different combinations to see which combination best matches the faunistic MDS.

Both these methods are useful in order to ascertain whether faunistic differences correlate best with measured levels of pollutants, or with naturally occurring environmental differences between stations (e.g. in sediment type or water depth) (see Warwick and Clarke, 1991).

A3.6 Availability of computer software

Computer programs developed or modified for use on PCs, and user-friendly to benthic ecologists, have been recently developed at the Plymouth Marine Laboratory for all the procedures outlined above. If these methods are to be adopted for routine monitoring programmes, the software needs to be made available to all potential users.

A3.7 References

CLARKE, K. R. AND GREEN, R. H., 1988. Statistical design and analysis for a 'biological effects' study. Mar. Ecol. Prog. Ser., ***46***: 213-226.

EVERITT, B. S., 1978. Graphical techniques for multivariate data. Heinemann, London, 128 pp.

FIELD, J. G., CLARKE, K. R. AND WARWICK, R. M., 1982. A practical strategy for analysing multispecies distribution patterns. Mar. Ecol. Prog. Ser., ***8***: 37-52.

KENKEL, N. C. AND ORLOCI, L., 1986. Applying metric and nonmetric multidimensional scaling to some ecological studies: some new results. Ecology ***67***: 919-928.

KRUSKAL, J. B. AND WISH, M., 1978. Multidimensional scaling. Sage Publications, Beverley Hills, California, 96 pp.

WARWICK, R. M. AND CLARKE, K. R., 1991. A comparison of methods for analysing changes in benthic community structure. J. Mar. Biol. Ass. UK., ***71***: 225-244.

ANNEX 4. The Abundance Biomass Comparison (ABC) method for detecting pollution effects on marine macrobenthos: description of method *(R. Warwick)*

A4.1 Background

Under stable conditions of infrequent disturbance the competitive dominants in macrobenthic communities are K-selected or conservative species with the attributes of large body size and long life span: these are rarely dominant numerically, but are dominant in terms of biomass. Also present in these communities are smaller r-selected or opportunistic species with a short life-span which are usually numerically dominant but do not represent a large proportion of the community biomass.

When pollution perturbs a community, conservative species are less favoured and opportunistic species often become the biomass dominants as well as the numerical dominants. Thus, under pollution stress, the distribution of numbers of individuals among species behaves differently from the distribution of biomass among species.

A4.2 Method

The ABC method, as originally described by Warwick (1986) involves the plotting of separate k-dominance curves (Lambshead *et al.*, 1983) for species abundances and species biomasses on the same graph and making a comparison of the forms of these curves. The species are ranked in order of importance in terms of abundance or biomass on the x-axis (logarithmic scale) with percentage dominance on the y-axis (cumulative scale).

In undisturbed communities the biomass is dominated by one or a few large species, each represented by rather few individuals, whilst the numerical dominants are small species with a strong stochastic element in the determination of their abundance. The distribution of numbers of individuals among species is more even than the distribution of biomass, the latter showing strong dominance. Thus, the k-dominance curve for biomass lies above the curve for abundance for its entire length (Figure A4.1(a)). Under moderate pollution, the large competitive dominants are eliminated and the inequality in size between the numerical and biomass dominants is reduced so that the biomass and abundance curves are closely coincident and may cross each other one or more times (Figure A4.1(b)). As pollution becomes more severe, benthic communities become increasingly dominated by one or a few very small species and the abundance curve lies above the biomass curve throughout its length (Figure A4.1(c)).

These three conditions (unpolluted, moderately polluted and grossly polluted) should be recognisable in a community without reference to control samples in time or space, the two curves acting as an 'internal control' against each other.

A4.3 Constraints

Adequate replication of sampling is a prerequisite of the method, since the large biomass dominants are often represented by few individuals, which will be liable to a higher sampling error than the numerical dominants.

A4.4 Advantages

There is less need for extensive control samples in either time or space, although these are of course highly desirable. The method has been shown to be a good indicator of the pollution status of benthic communities in a variety of organic enrichment

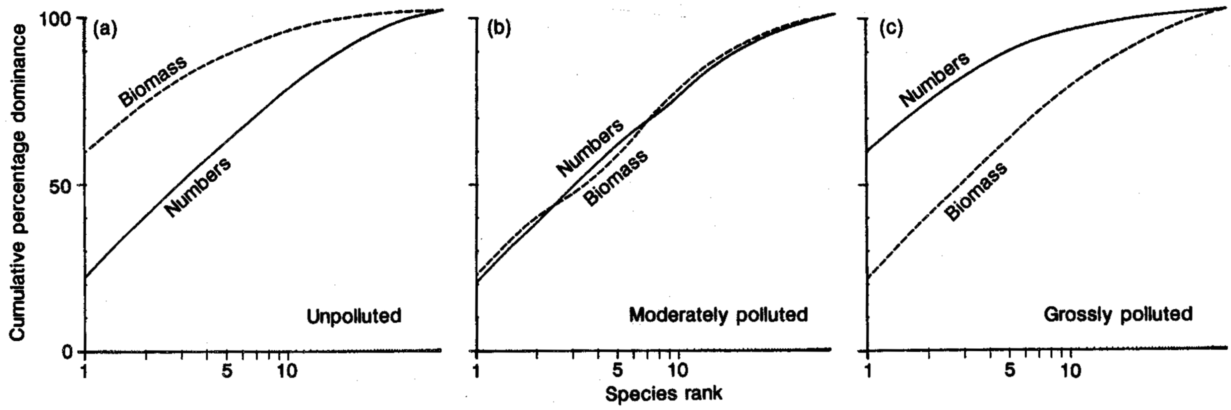


Figure A4.1. Hypothetical k-dominance curves for species biomass and numbers, showing unpolluted, moderately polluted and grossly polluted conditions situations (Warwick *et al*., 1987). The method also appears to be very robust to analysis at the family rather than species level (Warwick, 1988).

A4.5 References

LAMBSHEAD, P. J. D., PLATT, H. M. AND SHAW, K. M., 1983. The detection of differences among assemblages of marine benthic species based on an assessment of dominance and diversity. J. Nat. Hist., ***17***: 859-874.

WARWICK, R. M., 1986. A new method for detecting pollution effects on marine macrobenthic communities. Mar. Biol., ***92***: 557-562.

WARWICK, R. M., 1988. The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. Mar. Pollut. Bull., ***19***: 259-268.

WARWICK, R. M., PEARSON, T. H. AND RUSWAHYUNI., 1987. Detection of pollution effects on marine macrobenthos: further evaluation of the species abundance/biomass method. Mar. Biol., ***95***: 193-200.

ANNEX 5. Limits for change at the Tyne sewage-sludge disposal site *(H. Rees)*

Two stations at and to the south of the disposal site have been sampled annually in May since 1986 (see Figure 7 of the main report). Three 0.1 m^2 Day grab samples (of five) were analysed for the benthic macrofauna using a 0.5 mm mesh sieve, on each sampling occasion. Station 1 (the 'Treatment' station) is known to fall within the sphere of influence of sewage-sludge deposition. Further details are given in Rees *et al.*, 1992; see also Rees and Pearson, 1992).

A5.1 Method of analysis

Summary statistics (Shannon-Wiener diversity (H'$\log_2$), Evenness (E), Abundance (A), Taxa (T), Abundance/ Taxa ratio (A/T)) for samples at 'Treatment' and 'Reference' stations were integrated as follows, for each sampling occasion :

$$(\text{Treatment/Reference} - 1) \times 100.$$

This expression is a convenient way of allowing for the occurrence of natural synchronous changes with time. However, it does assume proportionality between sites in the response of the variable of interest (see Skalski and McKenzie, 1983 and Section 2.2.2 of the main report).

Means and standard intervals (= 95% confidence limits) for each pairwise comparison were generated by 'Bootstrapping' (see Efron and Tibshirani, 1986), using a program written by R. Fryer (SOAFD). These intervals should be treated with some caution, because of the small sample sizes involved, but are useful for illustrative purposes.

A5.2 Acceptability of change (EQS)

This comparison addresses the EQO concerning ecosystem maintenance (see Section 2.1 of the main report), for which the basis of the standard is 'deviation from the reference site(s) to be within acceptable limits'. Using the criteria suggested by T. Pearson derived from observations at Garroch Head (see Section 3.2.1 of the main report), abundances in excess of 200%, taxa (marginally) in excess of 50% and A/T in excess of 100% of values at the reference station may be deemed unacceptable, in that they signal the onset of structural change in communities, beyond initial 'mild' enrichment. The 'Action Points' will be reached when these measures are found to exceed limit values over three successive years. In fact, none of the measures was in breach over the sampling period (Figure A5.1(a-c)) suggesting that, at least when judged against the standards of Garroch Head, annual variability remained within acceptable bounds.

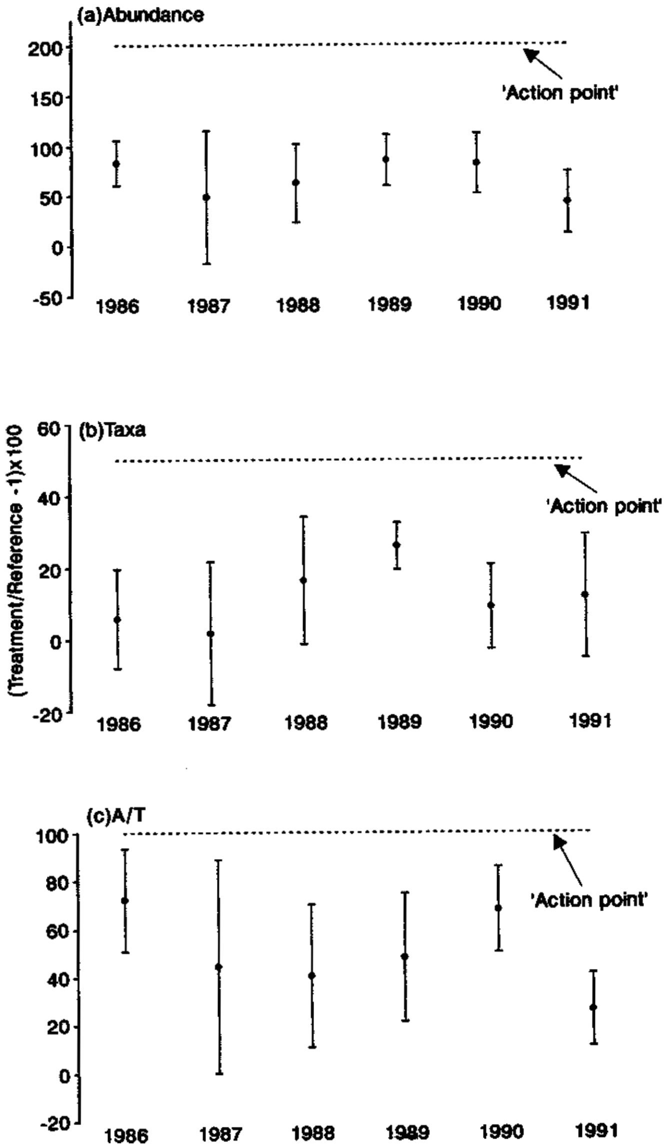


Figure A5.1. Means and 95% confidence intervals for pairwise comparisons of selected variables at the Tyne sewage-sludge disposal site. (Proposed 'Action Points' for acceptable change are superimposed)

A5.3 Monitoring for compliance with the *status quo*

Ideally, this comparison will be made between at least two stations peripheral to the sphere of influence (see Section 2.1 of the main report), in order to satisfy the EQO concerning preservation of the environment outside the zone of immediate effect ('no deviation from reference site(s)'). However, data from a second reference site are not presently available for the Tyne and so, in this case, the comparison is made to satisfy an additional objective, namely to establish that there is no worsening trend in intensity of any impact (within the sphere of influence) with time.

For the Treatment/Reference comparison (Figure A5.2(a-e)), values of abundance and A/T significantly in excess of zero, and evenness significantly (though not markedly) less than zero, are consistent with mild enrichment at the treatment station. The *status quo* in each case is defined by the solid line, which represents the mean of the first three years of sampling. The boundaries for permissible change were fixed on a purely pragmatic basis according to observed variability in the particular measure. (See also the time-series analysis of Tees Bay data in Section 2.2.2 of the main report).

The 'Action Points' are again defined as measures in breach of specified levels, sustained for a period of at least three consecutive years. It will be clear that in no case is there evidence for a trend either towards deterioration or improvement.

As expected, there is much greater variance associated with estimates employing A and hence A/T; logarithmic transformation may be more appropriate (see Section 2.2.1 of the main report). It will be understood that 'Action Points' will be intrinsically less significant for this comparison, since they do not indicate unacceptable effects. Action required following breaches (assuming that changes do not also exceed the boundaries for acceptability as defined above) might include more intensive investigations to check on the continued validity of the sampling stations.

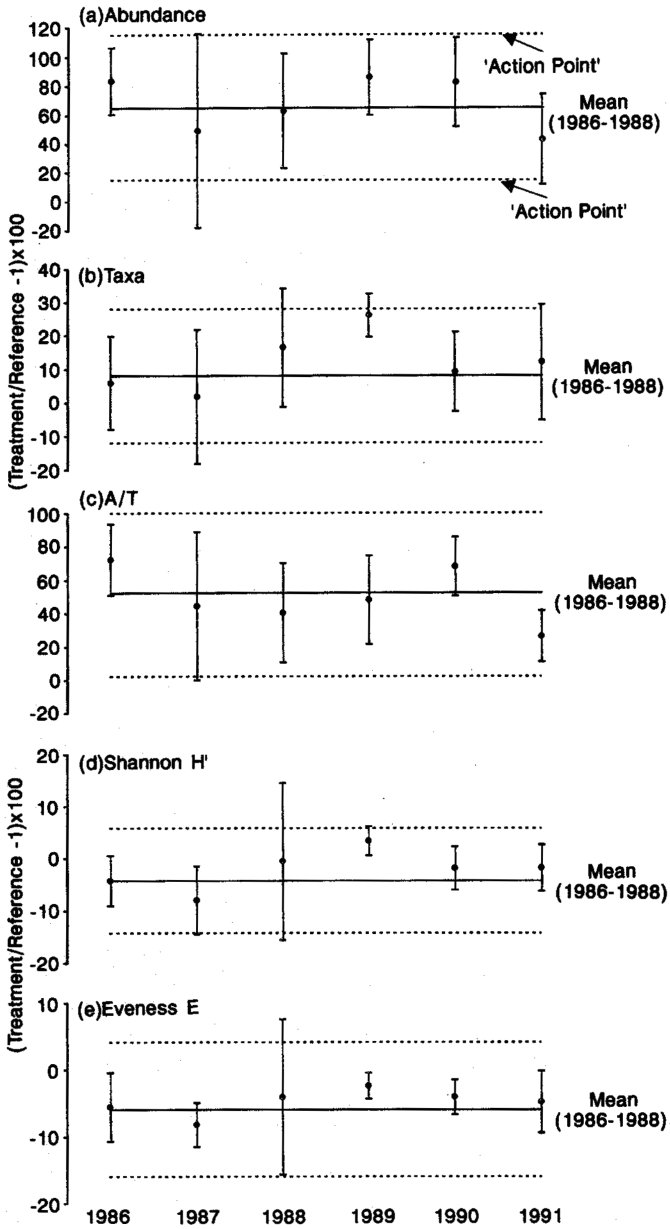


Figure A5.2. ***Means and 95% confidence intervals for pairwise comparisons of selected variables at the Tyne sewage-sludge disposal site. (The 'baseline' values and proposed 'Action Points' for monitoring of compliance with the status quo are superimposed)***

A5.4 Concluding remarks

For a given measure of data structure, precision can vary between years — sometimes appreciably — despite the fact that the same number of samples were analysed. Field sampling error — especially associated with weather conditions at the time of sampling — is likely to be important, as well as natural variability in benthic populations. It is therefore advisable to take 'back-up' samples in anticipation of a need for extra analyses, especially where measures are known to approach 'Action Points'.

Additional Treatment and Reference sites, along with greater replication, are required for the Tyne site in order to fulfil the requirements of this approach to quality assessment. 'Action Points' may in any case need to be adjusted in line with the capacity of the data to allow statistically significant breaches to be detected. Finally, there is a general need to examine the effects of mesh size on Treatment/Reference comparisons, and hence on the setting of realistic 'Action Points'.

A5.5 References

EFRON, B. AND TIBSHIRANI, R., 1986. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. Statist. Sci., ***1***: 54-77.

REES, H. L. AND PEARSON, T. H., 1992. An approach to the setting of Environmental Quality Standards at marine waste disposal sites. ICES CM1992/E:33, 10 pp. (mimeo).

REES, H. L., ROWLATT, S. M., LAMBERT, M. A., LEES, R. G. AND LIMPENNY, D. S., 1992. Spatial and temporal trends in the benthos and sediments in relation to sewage sludge disposal off the north east coast of England. ICES J. Mar. Sci., ***49***: 55-64.

SKALSKI, J. R. AND MCKENZIE, D. H., 1983. A design for aquatic monitoring programmes. J. Environ. Management, ***14***: 237-251.

ANNEX 6. Limits for change at the Liverpool Bay sewage-sludge disposal site *(I. Rees)*

A6.1 Introduction

Suggested 'Action Points' for assessing the acceptability of changes in the benthos were derived largely from experience at the Garroch Head sewage-sludge disposal site, where quiescent conditions allow accumulating organic detritus to modify the species - abundance relationship (see Section 2.2.1 of the main report). In such conditions, locations at or very near disposal sites are likely to be influenced in an obvious way. By contrast, in Liverpool Bay, conditions at the sludge disposal site are strongly dispersive, and the area licensed for sludge disposal covers over 29 square miles (Figure A6.1(a)). Effects, if detectable in such a locality, are likely to be either widespread or widely patchy. On a trial basis, the 'Treatment' versus 'Reference' methods have been applied to monitoring data from wide areas in Liverpool Bay in order to explore the implications of implementing an EQS approach in dispersive situations.

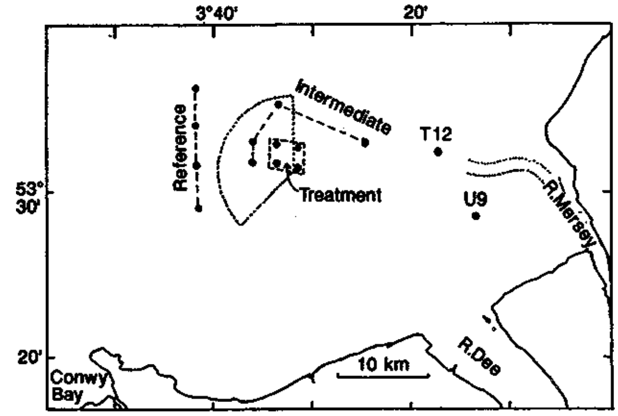


Figure A6.1(a). Location of Liverpool Bay disposal site, and of stations grouped for ecological quality comparisons (see text)

Within the potential range of influence of the sludge disposal site, four soft-bottom community types are recognisable. North and west from the point of the disposal site, the sediment is mainly heterogeneous muddy and gravelly sand supporting an abundant, rich and diverse benthic fauna. To the south, there tends to be more mobile sand, but in places it only partly covers fields of lag stones and boulders. This more unstable site carries a fauna less than a quarter as abundant as that in the more stable muddy and gravelly sands. Inshore, the contrast is between very sparse faunas of the sand banks and the muddy sands of accumulative pockets which, at times, carry huge populations of a limited range of species typical of an '*Abra*' community.

Sludge has been disposed of to sea in Liverpool Bay since the late nineteenth century. Since the late 1970s, licensed amounts have been steady at 80 000 dry tonnes and human populations in the catchment have remained fairly constant during the 20 years since the major multi-disciplinary study of 1970 (DOE, 1972). However, there have been significant changes in the balance of industrial contributions as closures and plant reconstructions reduced discharges of most heavy metals by more than an order of magnitude. Virtually all the sludge discharged to Liverpool Bay has been subject to both aerobic and anaerobic digestion processes.

A6.2 Benthos data sets for comparison and reference

Following an initial spatially-extensive survey of Liverpool Bay in 1970 (Rees *et al.*, 1972), a monitoring scheme was initiated in 1972. This has been kept going, almost annually, with only minor variations since then. The monitoring strategy has been based on the premise that, since conditions were strongly dispersive, the sampling effort needed to be quite widely dispersed as well. In particular, it was necessary to take account of the direction of the near-bed residual dispersal and to include inshore pockets, at a distance from the disposal site, into which repeated resuspension-cycles sorted refractory sludge-residues. Figure A6.1(b) shows the dispersal pattern of sludge-residues as indicated by tomato pip counts per m^2, averaged over the 1972-1990 period. To cover such a heterogeneous environment with a reasonable economy of effort, only two replicates were normally taken at 25 to 30 stations each year.

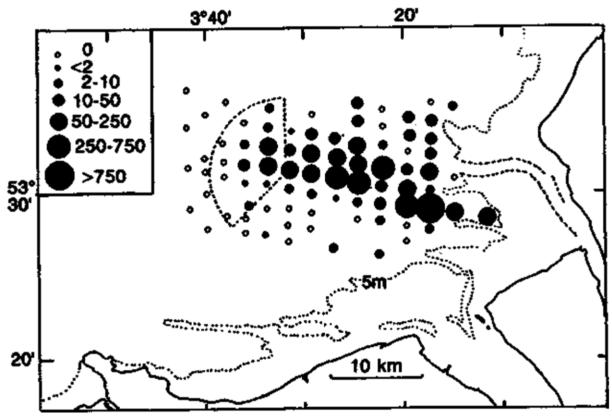


Figure A6.1(b). Distribution of tomato pips averaged over 1972-90 period (nos m^{-2})

The monitoring regime established as long ago as 1972 does not coincide with the likely demands of an EQS approach, as presently proposed. Such an approach is most suitable in situations with localised impacts, and calls for sufficient replication to put statistical confidence limits on the comparisons between single stations. Of necessity, with the Liverpool Bay data, clusters of stations have had to be linked as quasi-replicates to artificially meet confidence-limit criteria. In such a heterogeneous mixed-sediment environment, and with a large disposal site, pooling appropriate quasi-replicates may be more valid than relying on single-station data (see Figure A6.1(a)).

A cluster of 4 stations with similar sediments, spaced in a nearly square formation about 1.2 miles apart and lying on the eastern boundary of the disposal site, has been sampled very regularly since the 1970s. The stations, coded L10, L11, M10 and M11, are in the area most obviously receiving freshly-deposited sludge, as evidenced by samples often having the chemical smell characteristic of sludge from the Davyhulme works, and having more tomato pips than at any other stations in the immediate vicinity of the disposal site (average seed numbers 108 m^{-2}). This cluster of stations is well suited to being regarded as the 'Treatment' pseudo-station for this disposal site.

The choice of stations to cluster as a 'Reference' pseudo-station to compare with the disposal site is more difficult. To the east of the disposal site, stations would be unsuitable as this is the predominant direction for residual sludge dispersal. To the south, across the tidal axis, the sediment becomes more sandy and mobile while, to the north, the sediment becomes more consistently muddy. Even though, to the west, there is an environmental gradation away from the plumes of the Mersey and Dee estuaries, this is the only realistic direction in which to pick 'Reference' locations.

For this exercise, a line of stations running N-S and lying well to the west of the sludge disposal site was used as the 'Reference'. The four, coded G7, G9, G11 and G13, are about 6 miles west of the 'Treatment' cluster and well outside the western boundary of the sludge disposal site. Although it is theoretically possible for sludge discharged on ebbing tides to reach there, virtually no tomato pips have ever been found at any of these stations. The 'Reference' stations are not as closely spaced as the 'Treatment' ones, from G7 to G13 being about 7.5 miles. They have not all been sampled as often as the 'Treatment' ones, and inevitably there are some variations in sediment, but all have gravelly sand with enough mud to be partially cohesive.

No obvious cluster of stations could be found to regard as an 'Intermediate' pseudo-station. Instead, stations that were spaced within the disposal site, but away from the part where the discharges are concentrated, were grouped with stations north-east of the site. Although widely spaced, these 'Intermediate' stations all had sediment and expected faunas fairly similar to the other two grouped pseudo-stations. The stations chosen as 'Intermediates' for this exercise were coded K10, K11, L13 and Q12.

It has long been recognised that the most organically-stressed benthic communities of Liverpool Bay are to be found in localised pockets, where advective and sorting processes accumulate particulate organic detritus. A notable accumulative pocket, identified as likely to collect sludge residues, lies in the Burbo area between the Mersey and Dee estuaries. Though about 10 miles ESE of the preferred sludge disposal location, the Burbo pocket is directly in the path of residual near-bed dispersal, and collects tomato pips at densities exceeding 750 m^{-2}. The proportionate contribution of sludge to the total loading of degradable organic matter is difficult to estimate.

Another inshore site that has been regularly monitored is near the Mersey Bar, where the main ebb outflow from the estuary, channelled by training banks, meets a landward near-bed residual transport path from the bay. Although organic loading here may be high, obvious sewage contamination as evidenced by tomato pips is fairly small (average seed count c25 m^{-2}).

Given the likely interplay of effects from Mersey outflow and from sludge disposal, there is nowhere in the inner part of the bay that could be regarded as a 'Reference' location for the above inshore pockets. However, samples have been taken for other purposes covering roughly similar inshore locations off an estuary in Conwy Bay. For this exercise, a cluster of six muddy *Abra* community stations, taken in the same season and year, were chosen as a 'Reference' for the Burbo and Mersey Bar localities.

A6.3 Comparisons of community attributes

In line with Section 2 of the main report, the same measures were employed, namely Taxa (at the species level, but excluding non-quantifiable colonial forms), Total Abundance (excluding colonial species), Abundance/Taxa Ratio, Shannon Diversity H' and Pielou's Evenness Index E. Preliminary consideration was also given to the use of various other derived community attributes that might help overcome variability arising from having to pool spaced-out samples. As the statistical validity of these has yet to be confirmed, it would be premature to report results.

Table A6.1 shows arithmetic means and 95% confidence limits for the five selected measures, corresponding with the pseudo-station clusters in the offshore parts of Liverpool Bay in and around the sludge disposal site. It is apparent that, in all cases, the 95% confidence limits overlap, so that any statistically significant differences are unlikely between the 'Treatment' (disposal site), 'Reference' and 'Intermediate' clusters. The Shannon and Pielou Index values were clearly all within the ranges that are regarded as normal for offshore sediments.

Table A6.1. Mean values for basic community attributes at offshore localities in Liverpool Bay (1990 data). 'Reference' = pooled data from 4 stations west of the sludge disposal site; 'Treatment' = pooled data from 4 stations where initial sludge deposition normally occurs; 'Intermediate' = 4 stations around and within the licensed disposal area but not obviously contaminated. Figures in brackets are 95% Confidence Limits of the means

	'Reference'	'Treatment'	'Intermediate'
No. of samples:	8	8	8
Taxa	40.3 (36.6-43.9)	39.1 (34.4-43.9)	39.1 (35.7-42.5)
Abundance	233.0 (125.1-340.8)	179.6 (125.1-234.1)	230.9 (161.6-300.1)
A/T ratio	5.53 (3.30-7.75)	4.48 (3.34-5.62)	5.87 (4.30-7.44)
Shannon H'	4.12 (3.81-4.43)	4.25 (4.05-4.45)	4.00 (3.72-4.28)
Pielou E	0.779 (0.706-0.851)	0.809 (0.761-0.856)	0.757 (0.710-0.803)

There are several different ways of using the 'Environmental Quality' (EQ) formula (i.e. [Treatment/Reference -1]x100). For the purpose of this exercise, three were tried, and shown to give slightly different answers. This is assumed to be because the data were not distributed normally in all cases. Firstly, the simple expedient of using mean values for the comparisons was tried. However, this approach fails to give confidence limits. Secondly, the basic data sets were subjected to bootstrap statistical techniques to derive confidence limits (see Section 2.2.2 of the main report). Thirdly, the EQ formula was applied to the pairings of individual samples in 'Treatment' v 'Reference' matrices . In the case of the offshore Liverpool Bay pseudo-stations, where there were 8 samples in each, these became 64-cell matrices. Means and 95% confidence limits were then calculated for the 64 resulting comparisons.

In Table A6.2, the figures produced by the different methods of calculation are set out for both the 'Treatment' v 'Reference' comparison, and the 'Intermediate' v 'Reference' one. It will be noted that, although the confidence limits overlap, there were substantial differences, depending on whether the comparison was based on individual samples or their averages.

Table A6.2. Values for the 'Ecological Quality Comparison Index', derived from three different mathematical treatments. Comparisons were made between the 'Reference', 'Treatment' and 'Intermediate' pseudo-sites using 1990 data. Averages were based on means for the 8 samples in each pseudo-site. Bootstrap estimates were derived from a SOAFD computer program (see text). Matrix figures were calculated from Treatment/Reference or Intermediate/Reference comparisons for all 64 pairs in the 8x8 matrices

	Averages	Bootstrap	Matrix
Taxa			
(Treatment)	-2.8	-2.8 (-16.8-11.2)	-1.1 (-6.4-4.2)
(Intermediate)	-2.8	-2.8 (-14.6-9.0)	-1.1 (-5.6-3.3)
Abundance			
(Treatment)	-22.9	-22.9 (-67.0-21.2)	-18.8 (-4.3-41.8)
(Intermediate)	-0.91	-0.9 (-59.1-57.4)	52.7 (23.2-82.2)
A/T			
(Treatment)	-19.0	-19.0 (-59.1-21.1)	12.0 (-6.4-30.3)
(Intermediate)	6.2	6.1 (-46.9-59.1)	46.6 (22.2-71.1)
Shannon			
(Treatment)	3.2	3.0 (-5.7-11.8)	4.1 (1.0-7.3)
(Intermediate)	-2.9	-3.0 (-12.3-6.3)	-1.9 (-5.4-1.5)
Pielou			
(Treatment)	3.9	3.8 (-6.9-14.5)	5.3 (1.2-20.9)
(Intermediate)	-2.8	-2.9 (-13.2-7.5)	-1.4 (-5.3-2.5)

Table A6.3 'Ecological Quality Comparison Index' values for inner 'accumulative' parts of Liverpool Bay, along with values derived from the same community attributes for the sludge disposal site. The Conwy Bay 'Reference' consisted of 6 muddy stations, from which an average value for each variable was computed. Underlined values may exceed 'quality criteria'

	Disposal site (Offshore Ref.)	Burbo (Conwy Bay Ref.)	Mersey bar (Conwy Bay Ref.)
Taxa	-2.8	<u>-35.1</u>	-24.6
Abundance	-22.9	<u>74.3</u>	198.1
A/T	-19.0	<u>221.1</u>	<u>346.4</u>
Shannon	3.2	<u>-53.7</u>	<u>-45.3</u>
Pielou	3.9	<u>-49.4</u>	<u>-42.7</u>

Ecological quality comparison between the inshore locations in Liverpool and Conwy Bays are given in Table A6.3. For many of the measures, the inshore sites of Liverpool Bay differ from the chosen reference by more than the suggested 'Action Points'.

A6.4 Time-series comparisons

To assess temporal variability in quality-index values, comparisons were made using data from stations that have been sampled regularly since the mid-1970s. Only the A/T measure was used, values for which were derived from combined pairs of samples taken at each station on each survey. Thus the data are not directly comparable with the above analyses. For example, grouped pairs of replicate samples gave values for A/T and diversity that were about 3 to 5% higher than single samples. However, comparing like-with-like in the EQ formula should not affect the detection of any trends.

In Figure A6.2, histograms show the variability of the 'Treatment' v 'Reference' comparison between a station at the sludge-disposal site (M10) and a 'Reference' Station (G9). The greatest difference was in 1979, when elements of the communities normally associated with inshore mud pockets temporarily dominated the disposal site. Similar comparisons are made in Figure A6.2 between another station at the disposal site (M11) and one that could be regarded as 'Intermediate' rather than a true 'Reference'. The range of differences between these two stations (which are more closely spaced) was not as great.

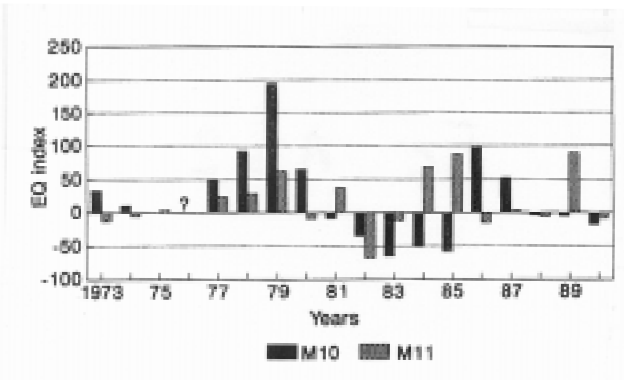


Figure A6.2. EQ Index from A/T Ratios

A6.5 Conclusions

In this preliminary assessment, quality comparisons would seem to show that, in 1990, the state of the benthos at the locations on the disposal site most likely to be impacted, could not be distinguished from the available reference locations. This was in spite of there being obvious, though not accumulative, contamination of the sediments on the disposal site. By contrast, accumulative locations remote from the actual disposal site (but which nevertheless receive refractory residues) do exceed proposed 'Action Points', when paired with equivalent reference locations.

It may therefore be concluded that, in the absence of organic-matter accumulation at the disposal site, such modifications to the fauna as may occur are not reflected in consistent modifications of the species-abundance relationship. Nevertheless, accumulations remote from an actual disposal site can influence such relationships enough for 'Action Points' to be exceeded.

The Liverpool Bay case suggests that the 'Treatment' versus 'Reference' approach can be applied to heterogeneous and dispersive disposal-sites. Some adaptation of sampling approaches may be needed, particularly with regard to the spacing of 'quasi-replicates' at suitable 'Reference' locations. Moreover, reliance should not be placed solely on measures based on the gross species-abundance relationship, if sensitivity is required.

At quiescent disposal sites, it could be argued that relatively severe influences on the benthos may be accepted, because the degraded areas will be both small and circumscribed. By definition, at dispersive locations such as Liverpool Bay, the area which could be influenced would neither be small nor circumscribed.

Thus tracers of sewage-sludge may be detectable over large areas, and biological 'Action Points' may be breached at distance from the disposal site. In the case of Liverpool Bay, it is therefore essential that the spatial dimension to changes at selected sites is taken into account in deriving criteria for acceptable change.

A6.6 References

DEPARTMENT OF THE ENVIRONMENT, 1972. Out of Sight, Out of Mind. Report of a Working Party on sludge disposal in Liverpool Bay. Vols ***1*** and ***2***. HMSO. London.

REES, E. I. S., WALKER, A. J. M., AND WARD, A. R., 1972. Benthic Fauna in relation to Sludge Disposal. Vol ***2***, Appendix 14, pp. 297-343, *In* 'Out of Sight, Out of Mind. Report of a working Party on sludge disposal in Liverpool Bay'. HMSO. London

ANNEX 7. Limits for change at the Barrow Deep sewage-sludge disposal site *(I. Codling)*

A7.1 Introduction

Several surveys of the outer Thames estuary have been carried out with the objective of determining some of the features of the impact of sewage-sludge disposal at the Barrow Deep site. Perhaps the most extensive series of surveys have been carried out by WRc between 1985 and 1990 in collaboration with, and in 1990 for, Thames Water Utilities (TWU). During this time both the physico-chemical and the biological status of the Outer Thames estuary sediments have been assessed.

A7.2 Definition of the zone of immediate effect

A key objective of the WRc/TWU collaborative research project (ap Rheinallt *et al.*, 1990) was to define the areal extent of sewage sludge impact. This was achieved by dividing the study area up into a grid using the DECCA navigation chains. The results from determinations of the following variables were combined for each cell in the grid:

- faecal bacteria densities;
- proportion of organic carbon in the <90 µm fraction;
- concentrations of metals, persistent organic compounds and coprostanol in the <90 µm fraction.

For each non-metal determinand, only the top 10% of the observed range was included. For the metals, only results that were significantly greater than the median were included. A correction/weighting factor was applied to compensate for the variable number of data points in each cell of the grid. The resultant ‘index’ was plotted and arbitrary values of >5 and >20 were taken to represent areas that showed values of the ‘index’ that were ‘slightly elevated’ and ‘elevated’.

The ‘elevated’ cells in the Barrow Deep, including the disposal site itself, and in the Barrow Swatchway can be identified, with supporting evidence from surveys tracking radiolabelled sludge, as those most likely to receive greatest contributions of the above determinands from sewage-sludge disposal.

The zone of immediate effect can therefore be defined as the Barrow Deep and the Barrow Swatchway (Figure A7.1).

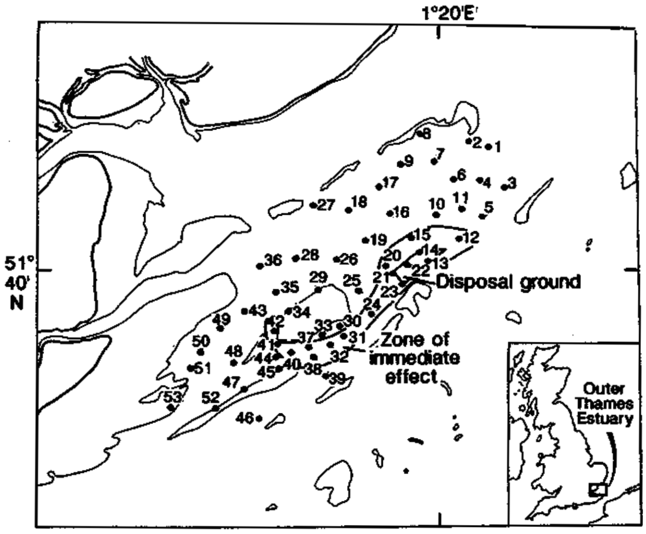


Figure A7.1. Thames EQS: zone of immediate effect

A7.3 Definition of sites for EQS purposes

The approach suggested in Section 2.1 of the main report requires the identification of at least 3 sites, or groups of sites, located within, just outside and distant from the zone of immediate effect of sewage sludge, respectively. Furthermore these sites should be similar with respect to depth and sediment type.

It is also important that the data to be used should have been collected using standard methods. This is particularly pertinent to the Barrow Deep disposal site where surveys have been undertaken by a number of organisations. For this reason, only data collected by WRc have been used in this preliminary assessment.

The geographical location of the selected groups of sites are illustrated in Figure A7.2, an indication of their environmental characteristics based on 1990 data are summarised in Table A7.1 and of their pollution characteristics based on 1990 data in Table A7.2.

It is apparent that the groups of sites are similar in terms of the mean values of both median particle size and depth.

The ranges, however, are somewhat different and cover approximately 80% of the range of values encountered in the whole survey.

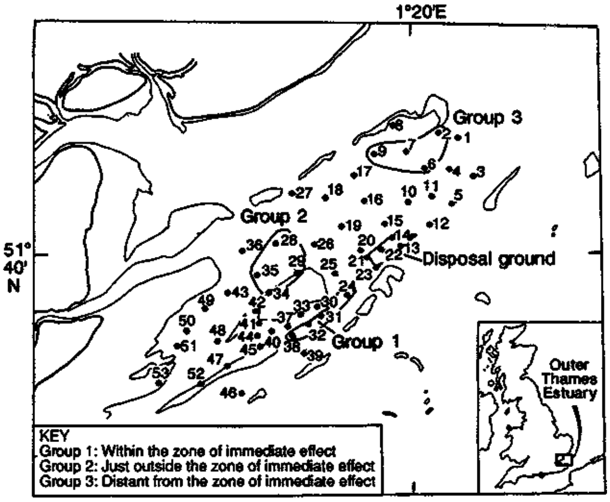


Figure A7.2. Thames EQS: station groupings in relation to the disposal site

Table A7.1. Physical characteristics of Groups 1-3 (1990 data only)

Group	Site	Median grain size	Depth
1	24	0.18	20
	31	0.23	16
	32	0.31	11
	38	0.25	6
Mean		0.24	13.25
2	29	0.3	10
	34	0.27	10
	35	0.2	11
	28	0.15	6
Mean		0.23	9.25
3	2	0.21	16
	6	0.14	18
	7	0.21	10
	9	0.14	7

Table A7.2. Pollution status of Groups 1-3. (1990 data only)

Group	Site No.	Cd	Hg	Org. C.	S-RC spores	Microtox
1	24	0.18	0.16	1.23	220	2.815
	31	0.15	0.35	1.45	1600	0.330
	32	0.22	0.42	1.84	2	0.695
	38	0.2	0.12	1.00	2	6.855
Mean		0.187	0.262	1.38	456	2.67
2	28	0.14	0.11	1.67	26	3.09
	34	0.23	0.44	1.37	280	2.18
	35	0.14	0.23	2.12	350	6.565
	29	0.14	0.11	0.68	240	12.455
Mean		0.16	0.22	1.46	224	6.07
3	2	0.15	0.09	2.19	2	1.995
	6	0.09	0.06	1.46	2	9.27
	7	0.13	0.09	0.1	2	41.315
	9	0.13	0.09	0.82	2	7.195
Mean		0.125	0.082	1.14	2	14.94

A7.4 Method of analysis

Sites within these groupings have been sampled in 1985, 1988 and 1990. For each sampling occasion the following summary statistics have been calculated:

Shannon - Weiner diversity ($H'log_2$);
Abundance (A);
Taxa (T);
Abundance/Taxa ratio (A/T).

The site groupings have been termed as follows:

within the zone of immediate effect - 'Treatment'
just outside the zone of immediate effect - 'Reference 1'
distant from the zone of immediate effect - 'Reference 2'

The following general formula was used to integrate each of the summary statistics for each of the years (where possible) to test compliance with the two EQOs:

(Treatment/Reference - 1) x 100.

A7.5 Acceptability of change (EQS)

This comparison addresses the EQO concerning ecosystem maintenance for which the basis of the standard is 'deviation from the reference site(s) to be within acceptable limits'. Such limits have been suggested by T. Pearson from observations at Garroch Head (Section 2.2.1 of the main report). These limits are marked on Figure A7.3(a-c) as 'Action Points'.

Values for this comparison for 1985, 1988 and 1990 are given in Table A7.3, and confidence intervals for the 1990 values have been generated using the 'Bootstrapping' routine (see Section 2.2.2 of the main report).

The mean values for pairwise comparisons of abundance, taxa and A/T (Figure A7.3(a-c), show that the Thames 'Treatment' sites are within the suggested limits of acceptability. Figure A7.3(d) illustrates the ranges of the expression for H', for which no 'Action Point' is presently available.

A7.6 Monitoring for compliance with the *status quo*

A fundamental requirement for the derivation of this EQS is the definition of the *status quo*. A suggestion for the definition of the *status quo* has been the mean of three years data for each of the summary statistics. At present there is insufficient data at the two reference site groups to allow the *status quo* to be defined.

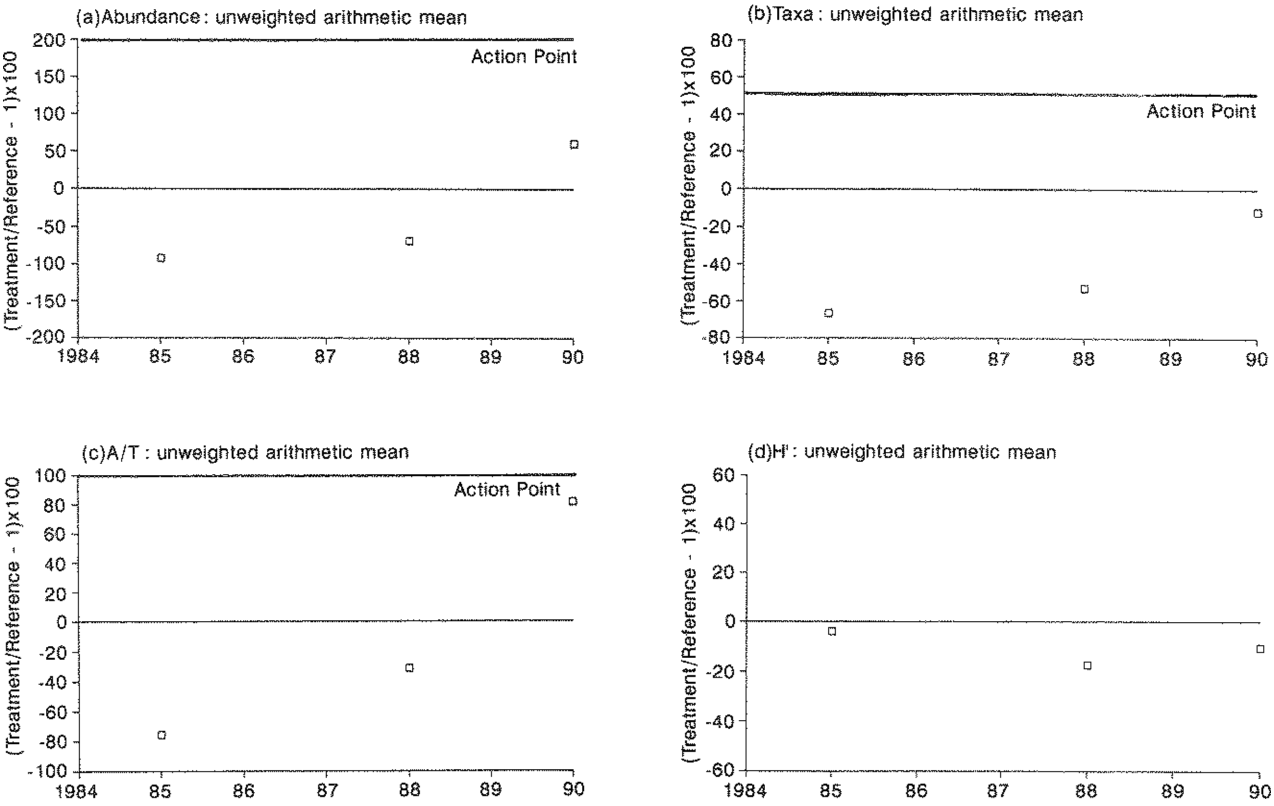


Figure A7.3. 'Treatment/Reference' comparisons for the outer Thames Estuary

Table A7.3. Means and confidence intervals (where appropriate) for the 'Acceptability of change' comparison

Summary statistic	Year	Mean	Lower limit	Upper limit
A	1985	-93.12		
	1988	-69.65		
	1990	60.146	-78.122	198.414
T	1985	-66.67		
	1988	-53.25		
	1990	-11.732	-29.447	5.983
A/T	1985	-76.21		
	1988	-31.72		
	1990	80.92	-67.390	229.354
H'	1985	-4.10		
	1988	-17.67		
	1990	-10.603	-17.104	-4.102

A7.7 Qualifying remarks

The validity of the above statements about acceptability is dependent on the quality and quantity of the data used, the method of calculation of the means and the relevance of the stated 'Action Points'.

The quality of the data used has, at least, been standardised by the use of only WRc data. The quantity of data used has been constrained by the sampling regime of the previous surveys. It has been possible to locate 4 sites for each of the 3 site groupings but sampling at these sites was not replicated. Mean values of the summary statistics have therefore been calculated from, at best, four data values. A design of at least three replicates at each site has been identified as being advantageous and this is clearly not met in this case.

The method of calculation of the mean values for this assessment has been simple unweighted arithmetic means, i.e. no correction has been applied for differing sample sizes.

The values of the 'Action Points' were derived from observations at the Garroch Head disposal site which is a containment site. The Barrow Deep disposal site is a dispersive site and as such differs fundamentally from Garroch Head. The 'Action Points' may be set too high because, for example, a 200% enhancement of abundance may never be achievable on this ground whatever the disposal regime. At present, no attempt has been made to define different values for the 'Action Points' for specific use on the Barrow Deep disposal site.

It is apparent from this preliminary assessment that further monitoring data is required for the successful application of the benthic biological EQS. Future surveys should include replication of those sites in the identified 'Treatment' and 'Reference' groups.

A7.8 References

ap Rheinallt, T., Nixon, S. C., Codling, I. D. and Orr, J., 1990. Thames estuary - environmental status with respect to sludge disposal - Final Report. WRc Report no. FR 0074.

Ministry of Agriculture, Fisheries and Food
Directorate of Fisheries Research
Fisheries Laboratory
Lowestoft
Suffolk
NR33 OHT
England

AQUATIC ENVIRONMENT MONITORING REPORT Number 37