INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA

MARINE ENVIRONMENTAL QUALITY COMMITTEE
ICES CM 1993/E:16 (REF L)

UNIVARIATE APPROACHES TO THE DERIVATION OF QUALITY STANDARDS FOR MARINE BENTHOS

A.J.R.Cotter[1] & H.L.Rees[2]

1. Directorate of Fisheries Research, Ministry of Agriculture Fisheries and Food, Lowestoft, England NR33 0HT.

2. Directorate of Fisheries Research, Ministry of Agriculture Fisheries and Food, Burnham on Crouch, England CM0 8HA.

Abstract:

Total abundance, numbers of taxa, abundance/taxa, and Shannon-Wiener diversity have been put forward as measures which may be suitable for monitoring and control as environmental quality standards for benthic ecosystems. This paper addresses statistical aspects of these measures by analysis of three time-series of benthic data: i) at two coastal sites off the River Tees, NE England, ii) near a discharge of particulate waste at Boulby, Cleveland, and iii) in the Firth of Forth estuary where sewage and industrial waste are discharged. In each case, 'treatment' and 'reference' stations were nominated and the four measures compared as ratios between the sites. It was found that logarithmic transformation of the ratios for all measures improved homogeneity of variance and normality, as well as making charts of the ratios more informative. The analyses lead to suggestions for sampling strategy and the layout of treatment and reference sites. It is noted that it is not necessary to identify species to calculate the four measures, merely to distinguish them. This could ease the taxonomic analysis of grab-samples, but it also implies that the measures are not the most powerful for assessing the health of benthic communities.

Introduction

In the United Kingdom, protection of marine ecosytems is based on authorisation of effluent discharges based on stipulated limits for certain chemical determinands, and in some cases these are based on ecological quality objectives defined for the receiving area. Questions repeatedly arise on whether the concentrations and levels laid down as acceptable will in fact protect the vital components of the ecosystem, and, from the other perspective, whether there is over-protection and therefore a waste of resources needed elsewhere. Ideally, environmental quality standards (EQSs) framed in terms of ecological measures which relate directly to the health and vigour of the ecosystem would be available. Benthic ecosystems are especially relevant because they contain many non-mobile species likely to be affected by local disturbances, and because the sea bed is the collecting place for the sediment which is a feature of most of man's activities, including waste disposal, mineral extraction and construction works.

Rees & Pearson (1992) and MAFF (in press) reviewed recent work in the United Kingdom directed towards the formulation of benthic EQSs. Three basic univariate measures were identified as suitable for routine use: total abundance [*A*] (i.e. the total number of individuals of all species found in a grab-sample), total taxa [*T*] (i.e. the total number of the lowest taxonomic level identified in the grab-sample, eg. species, families), and total biomass [*B*] (as ash-free dry weight). These can also be combined as the ratios *A*/*T* and *B*/*A* and supplemented with the Shannon-Wiener diversity index [*H'*] for interpretive purposes. The current intention is that such measures be compared as ratios between impacted or 'treatment' sites, a reference site (1) just outside the area of influence, and another reference site (2) some distance away. It is expected that multi-variate analysis and physico-chemical determinations would also be made to aid interpretation of *A*, *T* and *B* .

The present paper considers whether logarithmic transformation of results for *A* and *T* could be beneficial for their application in EQS's, this being an obvious choice of transformation given that ratios between two sites are of interest. Sampling and related matters are also discussed. Three data sets are used: i) at a coastal location off the R. Tees, NE England (Shillabeer & Tapp, 1990), ii) at Boulby, Cleveland, where large quantities of particulate waste were discharged from a pipeline (Craig *et al.*, 1993), and iii) in the Firth of Forth, near Grangemouth where sewage and industrial wastes were present. Unfortunately, results for *B* were not available in these sets. The following work was conducted under the auspices of the United Kingdom Group Co-ordinating Sea Disposal Monitoring (see e.g. MAFF, 1992).

1) Tees coastal site

The data consisted of measures made individually on 5 grab-samples taken at two sandy sites separated by some 500 m at about 20 m depth on 13 occasions, once in 1974 and then annually from 1976 to 1987. Taxa were available at both the species and family level, thus permitting a comparison of how the selected measures performed when calculated with each level. Station 1 was arbitrarily taken as the 'treated' site and station 2 as a 'reference'. Each of the five grab-samples from one site was arbitrarily paired with one from the other, giving five independent ratios to calculate an annual mean ratio and a variance. The variances were subjected to Bartlett's χ^2 test to assess homogeneity from year to year. The individual ratios were then standardised for year to year variation by subtracting the annual mean before calculating the other statistics shown in table 1. [Statistical note: the standard deviations were calculated from the annual variances weighted by their d.o.f.. Skewness and kurtosis statistics were not corrected for annual d.o.f. but the approximate estimates suffice to examine the effects of transformation.]

By comparing the results of Bartlett's test in the first and second parts of table 1 it can be seen that logarithmic transformation appreciably reduced the heterogeneity of the variances of the annual ratios of all measures. All skewness statistics were reduced in magnitude by the transformation, implying that the distribution of the log ratios (as deviations from the annual mean) was closer to Gaussian (normal) symmetry than untransformed ratios. Only abundance ratios were significantly different from normality, as gauged from the standard errors of the skewness statistics, approximately $\sqrt{(6/52)}$ = 0.34 for all measures. Kurtosis statistics were not generally reduced by log transformation, but abundance ratios were again exceptional, the large values seen in the table implying that the tails of the distributions were heavier than in a normal distribution. The standard error of these statistics was approximately $\sqrt{(24/52)} = 0.68$ for all measures, and abundance ratios were significantly different from normality. Presumably this also influenced the significant kurtosis values seen in the table for *A/T* for both species and family level identification. A further point worth noticing in the table is that species and family based measures tended to behave in unison.

The conclusions to be drawn are that logarithmic transformation of the ratios of these measures is likely to be beneficial for charting trends and for statistical tests, both of which are aided by homogeneous, normally distributed variability. Identification of all individuals to the species level does not appear to be necessary for these purposes.

2) Boulby

The data were in the form of annual summary statistics for abundance, taxa and diversity at 3 stations over the period 1970 to 1990 (except 1975). Station 33, close to the outfall, may be regarded as the treated site, station 56 as reference site 1, and station 83 as reference site 2.

Charts of the ratios and log ratios of abundance for sites 33:56 and sites 56:83 (fig. 1) are notable in that the plot of (ratio-1) over time gives a very different impression of variability than the plot of log ratio below. When viewed as a pure ratio, site 33 compared with site 56 showed relatively low but constant abundance from 1972 onwards; on the other hand, when viewed as logarithms, there was a severe decline from 1976, then recovery, then decline again from 1986. The difference has arisen in part from the single, unusually high abundance observed at station 56 in 1971 which has caused elongation of the computer drawn vertical axis for (ratio-1), and a consequent scaling down of subsequent, smaller variations. However, more important deceptions can be seen in the periods 1976 - 1980, and 1985 -1991 when the chart of untransformed ratios does not reveal the reducing fractional values which the log chart clearly does. Since these signified low abundances at the treated site when the reference sites were showing high and/or increasing abundances, serious environmental disturbance at the treated site could easily be overlooked without the log chart. There was however one notable weakness in the log chart: missing values arising when abundance at the treated site was zero, e.g. 1981. This could easily be overcome by transforming zeros to 0.1. Charts of taxa and other measures (not shown) made at Boulby also demonstrated the generally superior visual informativeness of the logarithmic plots. It is notable from the charts that sites 56 and 83 served well as reference sites, maintaining reasonably constancy in their log ratios so that comparisons with the treatment:reference pair were possible.

3) Firth of Forth, Grangemouth

This set of data consisted of annual species×station matrices for the period 1972 to 1990. Sampling did not take place in the same month in every year, and taxonomic categories were not always at the species level. Nevertheless, the set is unusually long and complete. Fig 2 shows the location of the main sampling transects. Charts of *A*, *T*, *A/T*, and *H'* on each transect (figs 3A-D) give contrasting impressions of the state of the benthic ecosystem along the estuary and over time. Data for *A* suggest an impoverished fauna at transects 3, 4 and 5. Data for *A/T* suggest the same but also that transects 7 and 9 could be only sparsely settled too. *T* for all transects is on the other hand quite uniform with transect 9 somewhat richer in species than the others; *H'*, indicates comparable diversity among all the transects.

Given the widespread presence of sewage and industrial effluents in this part of the Firth of Forth, as well as a salinity gradient, treated and reference sites are difficult to identify. For illustrative purposes, transect 5 was taken as the 'treated' site, and transect 9 as a 'reference' site. The stations within each transect (up to 5/year) were deemed to be replicates to permit analysis as for the Tees site, above. Table 2 shows the results of forming ratios of *A*, *T*, *A/T*, and *H'* between the two transects.

It can be seen that in all cases Bartlett's χ^2 was reduced by the logarithmic transformation. (There were 16 degrees of freedom), although significant heterogeneity of variance was still in evidence. Skewness and kurtosis also responded favourably implying that transformation improved normality. (Significance levels have not been calculated since there were too few data in some years.) Comparing table 2 with table 1, it is striking that variability of the treatment/reference ratios for all measures was substantially greater in the Firth of Forth than offshore near the R. Tees.

Another consequence of logarithmic transformation is worth noting. In order to give effect to EQSs it is necessary to agree 'action points', in this context, maximum values for the ratios between treated and reference sites which, if exceeded, call for action to improve the environmental situation. For example, for *A*, +200% has been suggested as a value indicative of unacceptable change in response to organic enrichment (Rees &

Pearson, 1992). The raw abundance figures for the 5 grab-samples taken on transects 5 and 9 in 1979 and shown in table 3, were selected and paired arbitrarily to illustrate a property of the transformation. When averaged as percentage change ratios (column 3 of the table) the result, +872 %, clearly called for action. When averaged as log ratios the result, -0.604, clearly did not, since +200% is equivalent to a ratio of 3:1, and ln 3 = 1.099 > -0.604.

Discussion

Some form of charting as time series is likely to form an important part of any control programme for EQSs framed in terms of ecological measures. The use of logarithmic ratios to compare treated and reference sites seems preferable to the use of untransformed ratios as it produces smooth series with reasonably normal errors, and is not too prone to erratic false alarms. The log transformation was better at revealing when the reference site was becoming enriched ecologically even though the treated site was almost barren. However, using logarithmic transformation, it is necessary to deal separately with zero values when they arise, and to be aware that ambiguity over the interpretation of results can arise when averaging is involved. In the case of the Firth of Forth which was used to illustrate the latter point, there was much variability, and the validity of transect 9 as a practical reference site is doubtful because of very low abundances occuring there on occasions. This points to the problem of finding adequate reference sites in a location like this which is both confined and subject to a salinity gradient, but better sites would not guarantee that similar ambiguities might not sometimes arise.

The arbitrary pairing of data from different sites was done here to create illustrative data and is not recommended for standard practice involving EQSs since different pairings would give different results. An objective sampling scheme would involve the collection of grab-samples from the treatment and reference sites on the same occasion. Replicate sets could be collected on different days within the selected season in order to minimise statistical dependence among the replicates caused by weather conditions, navigational errors and other factors in common. Apart from the statistical benefits of blocking comparisons in this way, the need to average logarithms probably would not then arise.

The four measures tested here, abundance, taxa, abundance/taxa, and diversity are all readily available from a standard taxa×stations data matrix and therefore may as well be computed. Charted results for the Firth of Forth suggested that they provide independent descriptors of the benthic environment, as is desirable for regulated measures (Cotter, in press). However, the picture painted by these four measures is not always an easy one to interpret in terms of the health of the ecosystem, particularly when different areas, e.g. off Tees and Firth of Forth, show diverse ranges for ratios of the measures between treated and reference sites. It is likely that action levels for ecologically based EQSs would have to be site-specific.

Another impediment to interpretation can be the presence of environmental gradients between treated and reference sites. In the Firth of Forth, there was an obvious salinity gradient between transects 9 and 5 which could render them invalid for EQS purposes. However it is difficult to be sure that influential ecological gradients are not present even when the sites have been well studied before selecting them for EQS comparisons. Where possible, it may be advisable to have more than one reference site of each type (1 or 2) and locate them on opposite sides of the treatment site, or even in a ring around it. The pooled grab-sample results would then be corrected for linear trends in abundance or taxa across the area.

The results from the Tees analysis indicated that identification down to species level is not necessarily adding much for the type of environmental quality controls tested here. Since the identification of benthic samples is a time-consuming and skilled task, the costs of compliance monitoring work might be reduced by restricting taxonomic identification to the family level, at least for those groups whose species are hard to distinguish or whose relevance to a biological quality assessment is thought to be marginal. Indeed, one might go further, for none of the four measures, A, T, A/T, or H, requires benthic individuals to be actually named; they need only be distinguished. The same is true of other benthic ecological assessment methods e.g. the abundance/biomass comparison method (Warwick *et al.* 1987), and the log-normal distribution method (Gray, 1989).

This potential for economising can alternatively be viewed as a weakness, since few benthic surveys would be regarded as satisfactory if they did not involve a reasonable taxonomic effort, yet the methods considered do not utilise the information laborously gained. To take extreme examples, they would not distinguish two sites with comparable A, T and H but different species compositions, nor would differing A, T or H signal when species compositions were similar at two sites. This underlines the importance of supplementary multivariate and physico-chemical measures, but even using these, it would be preferable if, given full taxonomic information, ecologically based EQSs could be derived in terms of measures which utilise it. One possibility is a trophic index, now under development for UK coastal waters (Codling & Ashley, 1992). This should allow sites to be compared in terms of species present and their trophic functions. Another is some sort of dissimilarity measure. Faith *et al.* (1991) have tested various types for their power to detect differences between treated and reference areas, pointing out that they are affected by changes in the taxonomic constitution of the communities. Statistical properties of such measures appear to merit further research.

Acknowledgements

The authors are grateful to Mr N. Shillabeer of the Brixham Environmental Laboratory, Zeneca Ltd, for provision of the data used in the above analyses.

References

Codling, I.D. & Ashley, S.J. (1992). Development of a biotic index for the assessment of pollution status of marine benthic communities. Report NR 3102/1, WRc plc, PO Box 16, Marlow, SL7 2HD, England.

Cotter, A.J.R. (in press). Compliance testing of two or more water quality determinands using quantiles. *Environmetrics.*

Craig, N.C.D., Shillabeer, N., & Parr, J.C.W. (1993). Long-term monitoring of the effects of a dense effluent discharged to North Sea waters. *J. Instit. Wat. Engin. & Manag.* 7(3), 262-269.

Faith, D.P. (1991). Statistical power and BACI designs in biological monitoring: comparative evaluation of measures of community dissimilarity based on benthic macroinvertebrate communities in Rockhole Mine Creek, Northern Territory. Australia. *Aust. J. mar. freshwater Res.* 42, 589-602.

Gray, J.S. (1981). Detecting pollution induced changes in communities using the log-normal distribution of individuals among species. *Marine Poll. Bulletin* 12, 173-176.

MAFF (1992). Fourth report of the Group Co-ordinating Sea Disposal Monitoring. *Aquatic Environ. Monit. Rep.* No 31, 38 pp.

MAFF (in press). Analysis and interpretation of benthic community data at UK sewage-sludge disposal sites. *Aquatic Environ. Monit. Rep.* No 37, Directorate of Fisheries Research, Lowestoft, England.

Rees, H.L. & Pearson, T.H. (1992). An approach to the setting of environmental quality standards at marine waste disposal sites. ICES CM 1992/E:33.

Shillabeer, N. & Tapp, J.F. (1990). Long-term studies of the benthic biology of Tees Bay and the Tees estuary. *Hydrobiologia*, 195, 63-78.

Warwick, R.M., Pearson, T.H. & Ruswahyuni (1987). Detection of pollution effects on marine macrobenthos: further evaluation of the species abundance/biomass method. *Mar. Biol.* 95, 193-200.

Captions to tables

Table 1: Summary statistics for 'treated' and 'reference' sites at the offshore Tees dump-site as (1) percentage change ratios and (2) logarithmic ratios; sampling conducted 13 times between 1974-1987 (5 grab-samples) and data corrected for variation of the annual mean. * signifies $p<0.05$, ** $p<0.01$.

Table 2: Summary statistics for 'treated' and 'reference' sites in the Firth of Forth near Grangemouth as (1) percentage change ratios and (2) logarithmic ratios; sampling

conducted 18 times between 1972-1990 (up to 5 grab-samples) and data corrected for variation of the annual mean. * signifies $p<0.05$, ** $p<0.01$

Table 3: Abundance data for grab-samples taken from transects 5 and 9 in 1979, Firth of Forth; column 3 shows percentage change ratios, and column 4 the log transformed ratios after pairing. The 'action level' is taken from Rees & Pearson (1992).

Table 1: Summary statistics for 'treated' and 'reference' sites at the offshore Tees dump-site as (1) percentage change ratios and (2) logarithmic ratios; sampling conducted 13 times between 1974-1987 (5 grab-samples) and data corrected for variation of the annual mean. * signifies $p<0.05$, ** $p<0.01$.

1. (treated/reference-1)/100

Measure	Standard deviation	Minimum	Maximum	Skewness	Kurtosis	Bartlett's χ^2
A	72.9	-146.6	234.6	1.12**	2.77**	54.4**
T (species)	30.9	-48.4	70.3	0.55	-0.37	11.7
T (families)	30.7	-55.8	66.4	0.33	-0.40	17.0
H' (species)	25.3	-56.2	67.0	0.33	0.48	21.6*
H' (families)	25.3	-56.7	73.5	0.45	1.16	28.6**
A/T (species)	45.8	-110.5	116.5	0.64	1.49*	37.4**
A/T (families)	46.7	-115.1	112.2	0.58	1.33	31.9**

2. Ln(treated/reference)

Measure	Standard deviation	Minimum	Maximum	Skewness	Kurtosis	Bartlett's χ^2
A	0.57	-1.07	2.04	0.78*	3.06**	29.7**
T (species)	0.29	-0.57	0.64	0.32	-0.39	10.0
T (families)	0.28	-0.55	0.55	0.069	-0.53	12.0
H' (species)	0.23	-0.52	0.42	-0.21	-0.36	16.1
H' (families)	0.23	-0.53	0.41	-0.16	-0.15	20.6
A/T (species)	0.43	-0.80	1.40	0.47	1.66*	26.1*
A/T (families)	0.46	-0.84	1.48	0.46	1.63*	23.6*

Table 2: Summary statistics for 'treated' and 'reference' sites in the Firth of Forth near Grangemouth as (1) percentage change ratios and (2) logarithmic ratios; sampling conducted 18 times between 1972-1990 (up to 5 grab-samples) and data corrected for variation of the annual mean. * signifies $p<0.05$, ** $p<0.01$

1. (treated/reference-1)/100

Measure	Standard deviation	Minimum	Maximum	Skewness	Kurtosis	Bartlett's χ^2
A	1790	-3470	10153	4.03	28.2	266.8**
T	151.4	-267.9	797.0	3.29	19.7	81.8**
H'	79.0	-125.1	356.8	2.07	9.2	58.7**
A/T	1257	-2598	7609	4.63	36.6	283.4**

2. Ln(treated/reference)

Measure	Standard deviation	Minimum	Maximum	Skewness	Kurtosis	Bartlett's χ^2
A	1.96	-4.44	4.70	0.26	1.26	27.0*
T	0.94	-1.71	2.03	0.19	-0.08	15.4
H'	0.65	-2.20	1.27	-0.54	2.58	28.1*
A/T	3.19	-3.07	4.08	0.52	2.92	31.7*

Table 3: Abundance data for grab-samples taken from transects 5 and 9 in 1979, Firth of Forth; column 3 shows percentage change ratios, and column 4 the log transformed ratios after pairing. The 'action level' is taken from Rees & Pearson (1992).

T5 (treated)	T9 (reference)	(T5/T9-1)*100%	ln(T5/T9)
4	216	-98	-3.989
474	10	+4640	+3.859
3	9	-67	-1.099

Figure 1.

BOULBY: abundance comparisons for sites 33, 56 & 83. a)ratio−1 b)log ratio

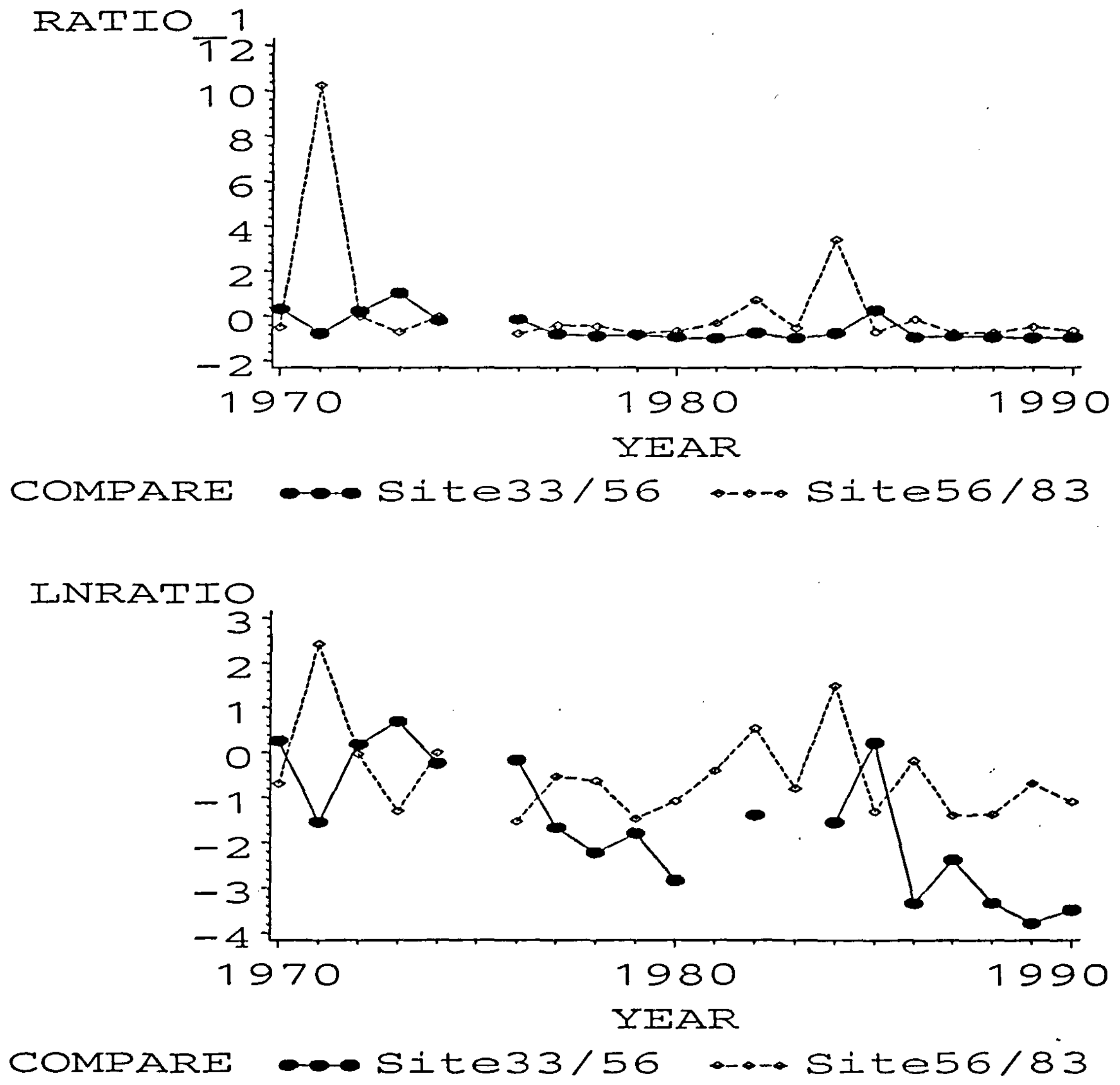

Figure 2. Location of Main Sampling Transects, Grangemouth Main Channel

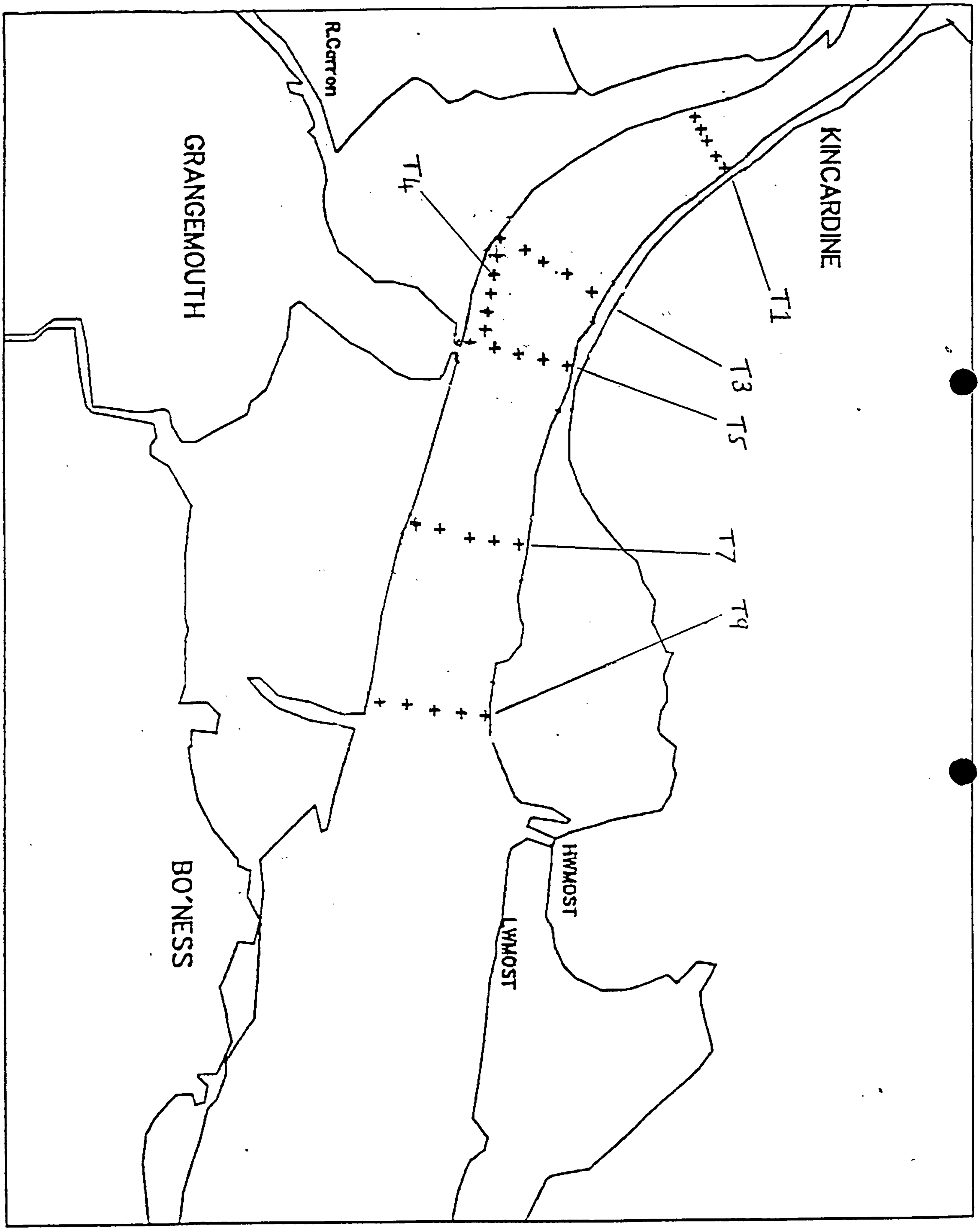

Figure 3a.

ABUNDANCE from April 1972; Transects 1 3 4 5 7 9, GRANGEMOUTH

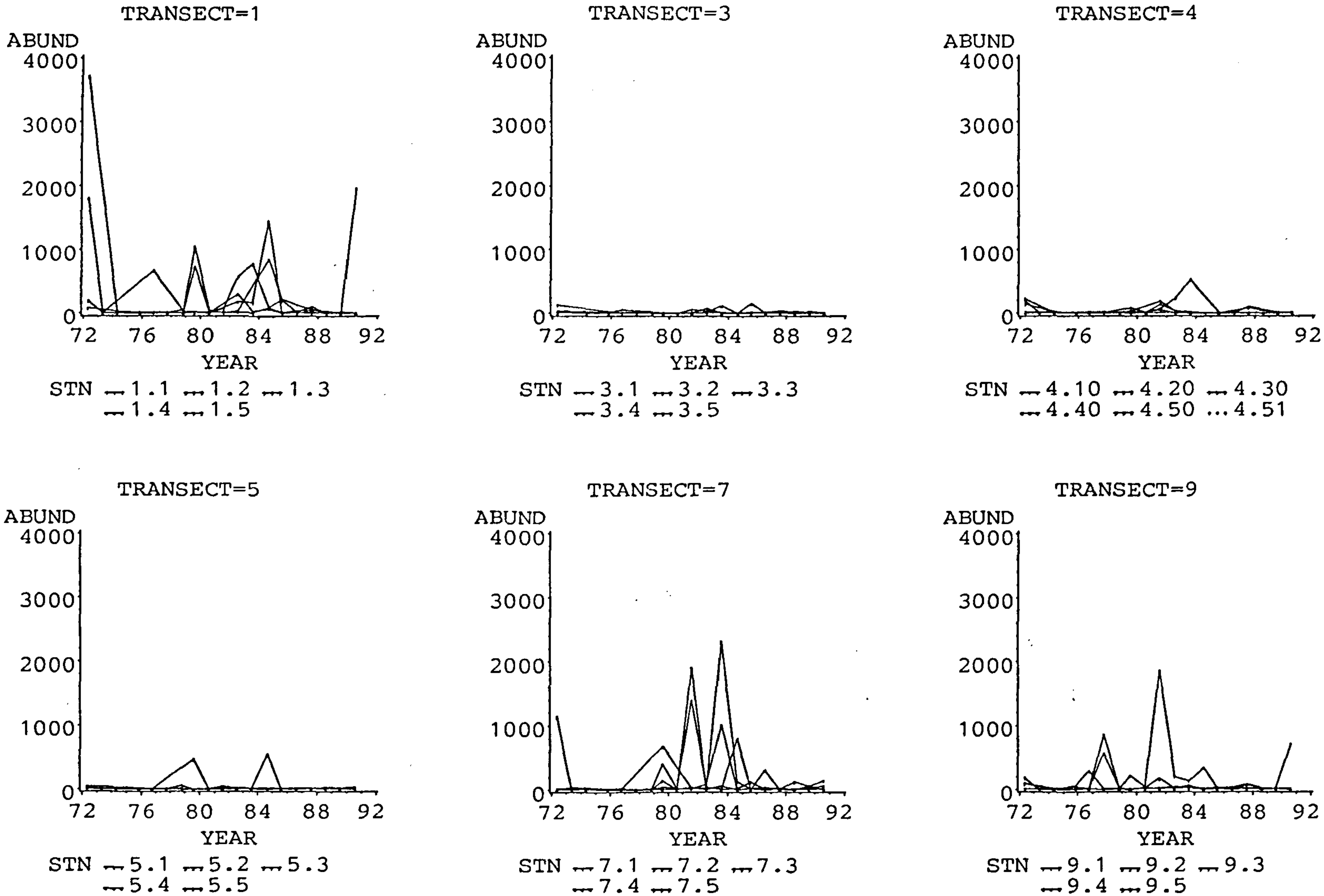

Figure 3b.

TAXA from April 1972; Transects 1 3 4 5 7 9, GRANGEMOUTH

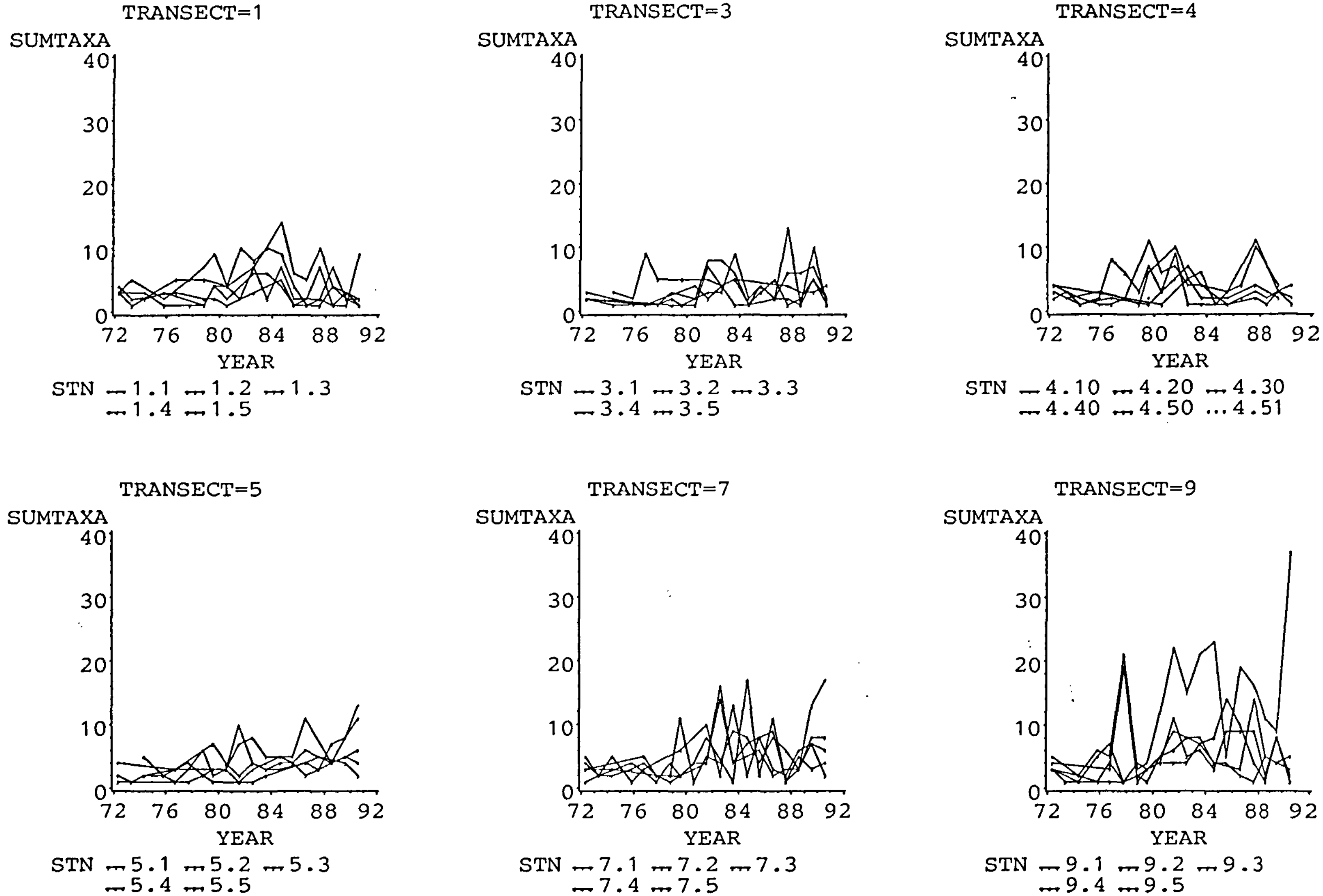

Figure 3c.

ABUNDANCE/TAXA from April 1972; Transects 1 3 4 5 7 9, GRANGEMOUTH

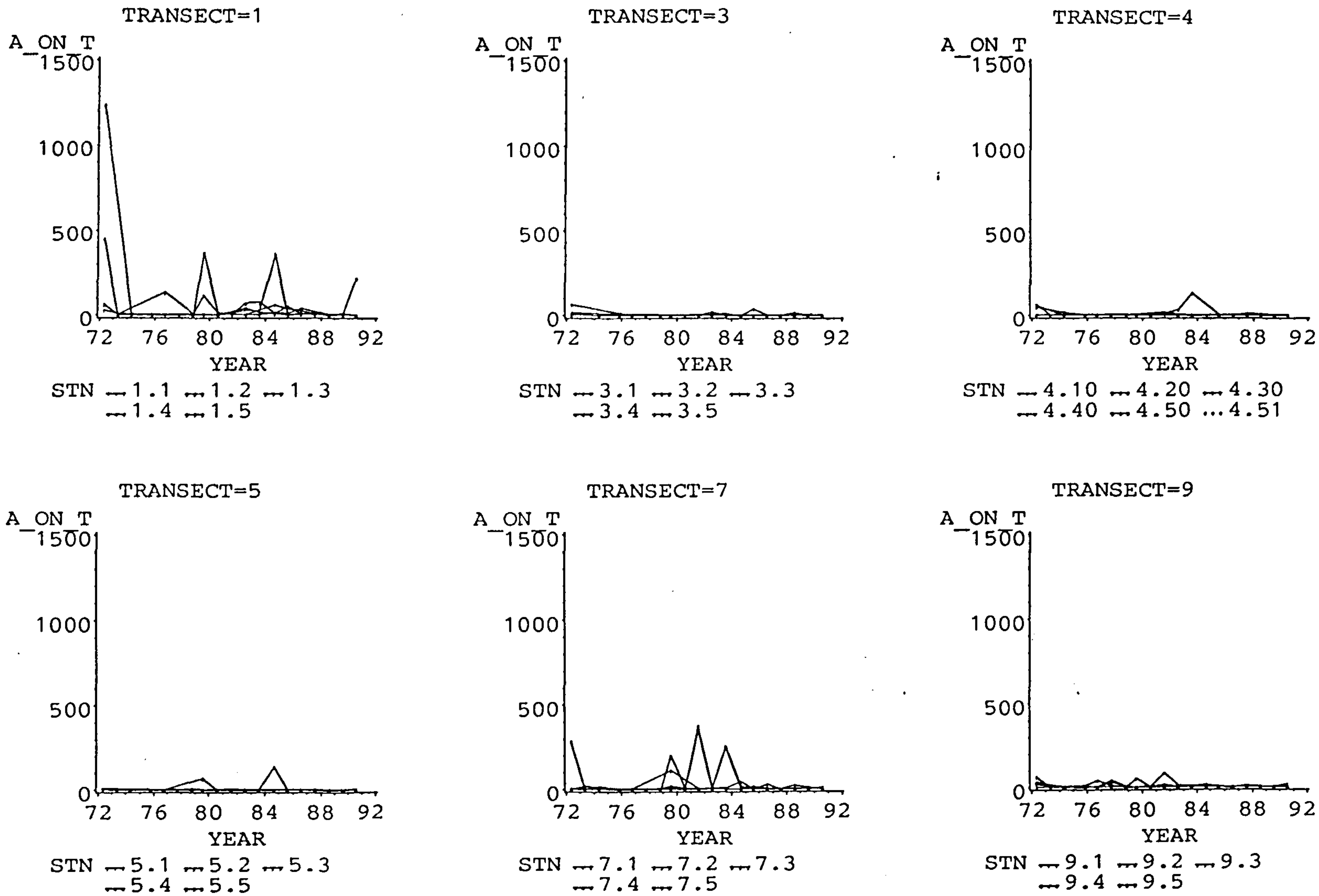

Figure 3d.

DIVERSITY from April 1972; Transects 1 3 4 5 7 9, GRANGEMOUTH