Detection of pollution-induced disturbance in marine benthic assemblages with special reference to nematodes

K. M. Shaw[1], P. J. D. Lambshead[2] and H. M. Platt[2]

[1] Biometrics and Computing Section, British Museum (Natural History), Cromwell Road, London SW7 5BD, United Kingdom
[2] Nematode Section, British Museum (Natural History), Cromwell Road, London SW7 5BD, United Kingdom

ABSTRACT: Benthic marine nematode assemblages were assessed for possible use in the detection of pollution-induced disturbance. Doubt is cast on a recently proposed method which claims to observe deviation from a log-normal distribution of species abundance. Little evidence is found to substantiate the belief that natural marine nematode assemblages in temperate waters are well described by the log-normal distribution. This conclusion may not be restricted to nematodes. Simple dominance is suggested as a potentially more practical method of assessing changes due to pollution.

INTRODUCTION

For a variety of reasons, marine nematodes have been suggested as possible pollution indicators and to have certain attractions in routine biological monitoring (Ferris and Ferris, 1979; Platt and Warwick, 1980).

To investigate this claim in more detail, both published and hitherto unpublished nematode data were analysed. Invariably, such studies produce a bewildering array of data sets – so a suitable method of data reduction is required to enable comparison between them. Since one of the problems and indeed weaknesses of many pollution studies is the difficulty of finding suitable unpolluted control or reference sites (Gray, 1976) we tried to apply the intrinsic probability-plotting method of detecting pollution-induced changes suggested by Gray and Mirza (1979). This method, which has aroused much interest (Ortner et al., 1982; Soulsby et al., 1982), is based on the frequency distribution of species abundances: deviation from a log-normal distribution of individuals per species it is said may indicate a disturbed or polluted assemblage. However, when we tried to assess our data in this way, we experienced certain difficulties. Attempts to resolve these difficulties led us to a consideration of the basic principles involved and, eventually, to cast doubt on the usefulness of the whole concept.

In this paper, then, we first discuss the probability paper method of estimating goodness-of-fit to the log-

normal. Secondly, we address the question of whether the log-normal is in any case a good description of species abundance distributions for most marine benthic assemblages. Finally, we suggest a simple alternative method which has potential for the detection of pollution-induced changes in these benthic assemblages.

PLOTTING LOG-TRANSFORMED ABUNDANCES ON NORMAL PROBABILITY PAPER

Gray and Mirza (1979) considered that many natural unpolluted assemblages of species have abundance distributions that are well-fitted by the log-normal model. They demonstrated that the effect of pollution is to increase dominance and claimed that this causes a departure from the log-normal model and, further, that such population disturbance is transient, the distribution reverting to log-normal 'after a certain time period'. Gray (1981) further suggested that a heavily polluted assemblage would also have a log-normal distribution, albeit with a larger standard deviation, although Ugland and Gray (1982) subsequently refuted this.

Gray and Mirza (1979) assessed departure from a log-normal distribution by a cumulative plot of log-transformed abundances on normal probability paper. However, a rather primitive method of probability

plotting was used: the log-transformation was achieved by classing into Preston's (1948) octaves, irrespective of the number of species, thereby reducing the data to perhaps as few as 6 points – far too few to assess accurately the shape of the distribution.

Secondly, they chose as plotting points the sample cumulative percentage frequency $i/S \times 100 \%$, where S = total number of species; i = cumulative number of species in each class. Thereby they prevented the highest class from being plotted since the 100 % point is at infinity on the probability scale. Not only does this further reduce the number of points plotted but, if pollution does increase dominance, it is the highest class that will be most affected and hence give the most information. A better choice for plotting position on statistical grounds is $(i - 0.3)/(S + 0.4)$ (Barnett, 1975), which allows the highest class to be plotted.

Any observed species abundance distribution is always truncated on the left to some degree because it is impossible to observe less than 1 complete individual of a species (Preston, 1948). Fig. 1 shows the
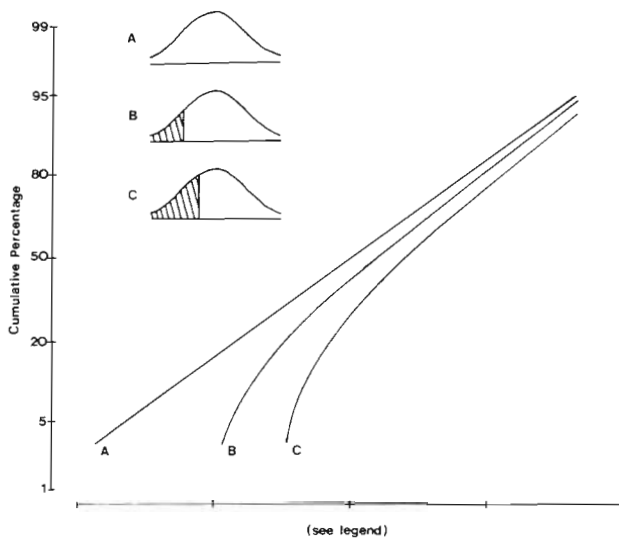

Fig. 1. Effect on the probability plot of truncating the log-normal distribution (A) by 15 % (B) and 30 % (C). Scale of x-axis purely arbitrary

effect on the probability plot of truncating a log-normal distribution by 15 % and 30 %. The plots from a truncated distribution might well be misinterpreted as being the result of increased dominance distortion. This could have important repercussions in a biological pollution monitoring programme where an extensive initial base-line survey might be followed by a more restricted series of samples. Gray (1981), who recognized this problem, suggested omitting the first 2 classes. However, this increases the problem of inaccuracy caused by having too few points for fitting the line.

More seriously, deviations from the log-normal are demonstrated by non-linearity at the ends of the curves, the central portion remaining linear for quite large departures from the model. Omitting the lowest points to avoid the effects of truncation and the highest point because of the choice of plotting position considerably reduces the sensitivity of the method.

Gray and Mirza (1979) considered that their data from the unpolluted Station A, Oslofjord, gave a good straight line when plotted on probability paper. Nevertheless, when these data are plotted as log-transformed abundances, the frequency distribution casts doubt on the goodness-of-fit to the log-normal (Fig. 2). Replotting using Barnett's choice of plotting points (Fig. 3) also indicates the lack of fit, both by the non-
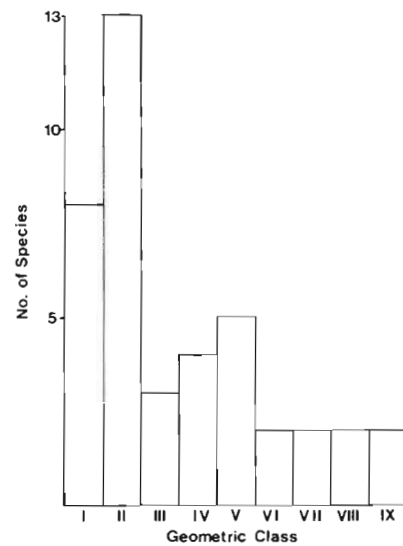

Fig. 2. Log-transformed abundances (x 2 scale) of species from Station A, Oslofjord (data from Gray and Mirza, 1979)
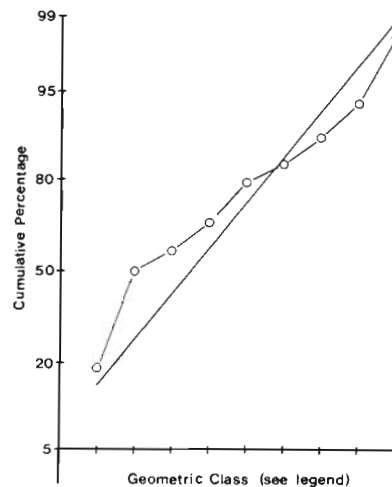

Fig. 3. Oslofjord data of Fig. 2 replotted using Barnett's choice of plotting points. Straight line is theoretical log-normal with same mean and standard deviation as data. X-axis represents x 2 geometric classes I–IX

linearity of the plot and the systematic variation about the theoretical line representing a log-normal distribution with the same mean (3.5) and standard deviation (2.40) as the Oslofjord data.

Probability plotting is, in any case, a relatively insensitive technique which should be used primarily for initial screening. Stenseth (1979) provided evidence that in disturbed assemblages a log-series is a better fit than a log-normal. However, it can be difficult to distinguish these 2 distributions by probability plotting. In particular, the truncated log-normal produced by a moderately sized sample from a log-normal population has a distribution not unlike a log-series. Williams (1964: Table 8) quoted data on macro-lepidopteran captures at Rothamsted Experimental Station (England) in 1935, which includes 6,814 individuals from 197 species and considered these data to fit the log-normal because of the close fit to a straight line as a cumulative plot of log-transformed abundances on probability paper (Fig. 4). Nevertheless, Fisher et al. (1943) considered it to be a good example of the log-series, an opinion echoed by Krebs (1972).

In view of the insensitivity of Gray and Mirza's method, especially in distinguishing between log-series and log-normal models, we question its value in practical pollution analyses.

THE LOG-NORMAL MODEL

At this point, rather than considering alternative methods of fitting the log-normal, it is worth investigating in more detail whether natural or undisturbed
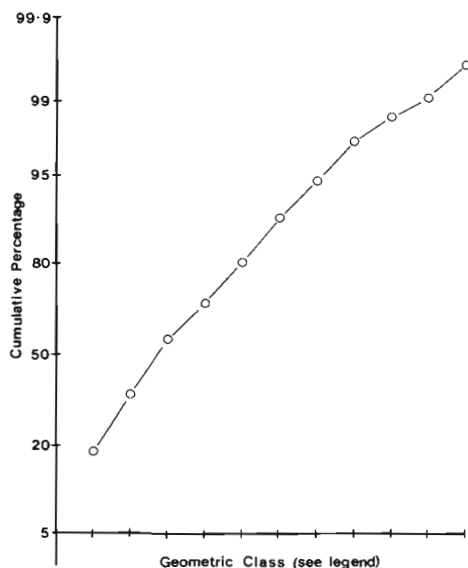
Fig. 4. Probability plot of macro-lepidopteran captures at Rothamsted Experimental Station (Williams, 1964). X-axis represents x 2 geometric classes I–XI

marine benthic assemblages do indeed conform to the log-normal model. Amongst the many attempts to derive the log-normal model, May (1975) provided a purely stochastic argument. Until such an explanation is eliminated, it seems less parsimonious to suggest a more elaborate biological explanation.

May (1975) noted that, for biological assemblages generally, the Central Limit Theorem would suggest a log-normal since (a) organisms tend to increase geometrically and (b) the distribution of relative abundance is liable to be governed by more or less independent factors compounded multiplicatively. For similar reasons, many non-biological multiplicative processes are also log-normal, such as the distribution of wealth in the USA or the international distribution of human populations (May, 1975). It therefore seems logical that a biological assemblage will deviate from the log-normal if it is severely affected by o n e factor or a set of l i n k e d factors, such as often occurs in polluted situations. This is because the distribution of abundances is now overwhelmingly controlled by a single factor rather than by a number of independent factors compounded multiplicatively. It seems reasonable that, by the same argument, a number of natural unpolluted marine environments may be dominated by one or a limited number of biological or environmental factors which could act in such a way as to cause a deviation from the log-normal. It may also be the case that such dominating factors only manifest themselves at certain times of the year.

Fig. 5 shows data from a number of marine nematode assemblages, although Moore's (1973) kelp holdfast fauna includes other taxa in addition to nematodes. Some, such as Platt's (1977) low-tide station (Fig. 5, I) do seem to suggest a log-normal distribution but others are far less convincing. Certainly the distributions in allegedly stressed areas (solid circles in Fig. 5) do not appear any less log-normal than those in unstressed areas. However, it is noticeable that the distributions displaying high dominance are all from areas which might be expected to be undergoing stress, either from pollution or natural environmental factors (e.g. Figs. 5, Q and 5, G respectively).

It is worth keeping in mind the fact that once truncation of a putative log-normal has obscured the mean, accurate estimation of the true mean becomes difficult and it is mathematically hazardous to attempt to fit a log-normal distribution to such data. It would be most unwise to assume that the first geometric class in such a case is the modal class, as for example in Fig. 5, P (or even that it is the second geometric class in Fig. 5, F).

It is a matter for some concern that some of the data which best seem to fit a log-normal distribution are amalgamations (stars in Fig. 5) of samples taken from different sites or at the same site but at several separate
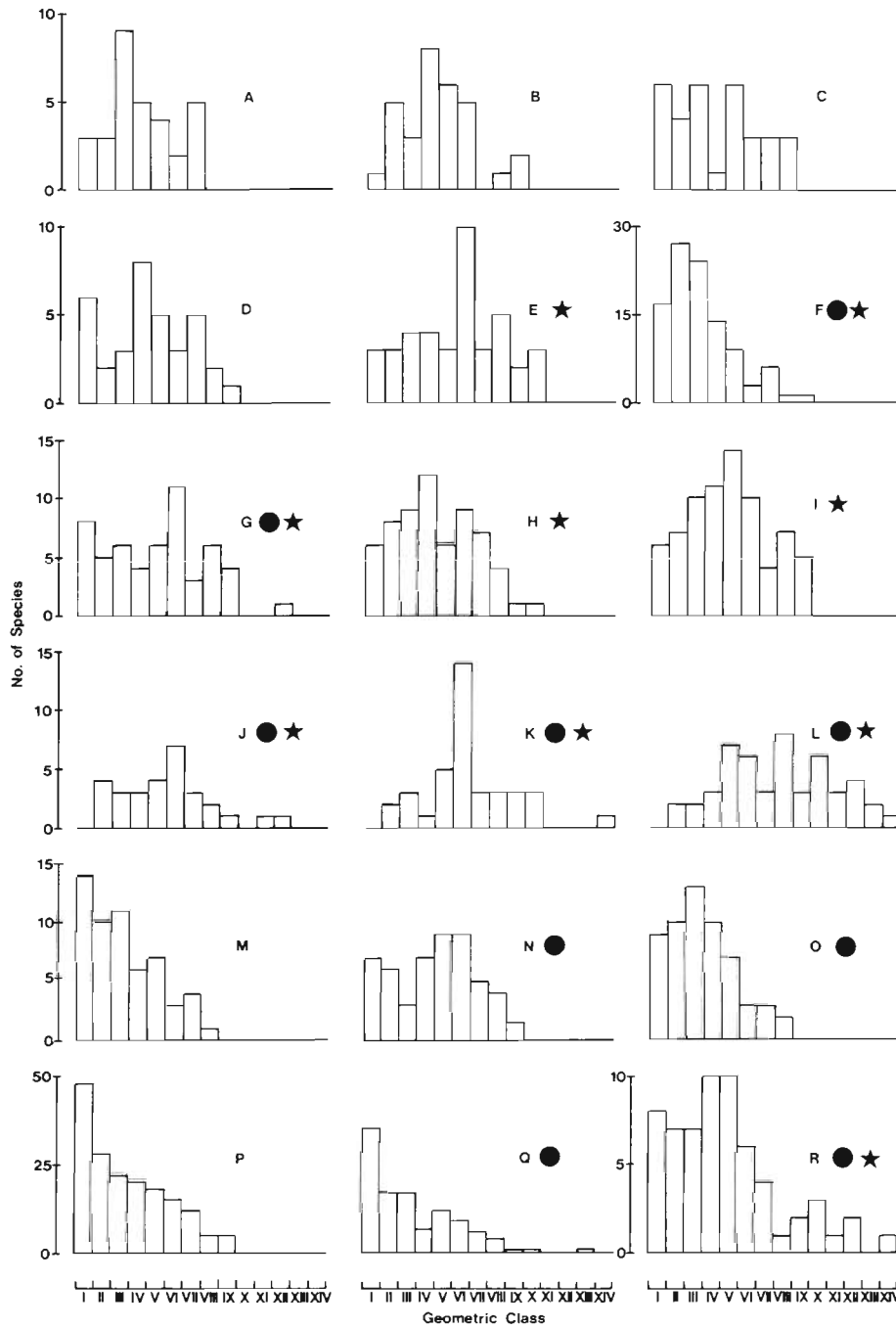
Fig. 5. Log-transformed species abundances from a number of free-living marine nematode assemblages. A–D Lynher estuary, S.W. England: October 1972 (A), January (B), April (C), and July 1973 (D) (Warwick and Price, 1979). E Amalgamated data from Lynher estuary. F Morlaix Bay, N.W. France, Station b (Gourbault, 1981). G–I Strangford Lough, N. Ireland: high (G), mid (H), and low (I) tide stations (Platt, 1977). J–L Irvine Bay, Firth of Clyde, Scotland: high (J), mid (K), and low (L) tide stations (Jayasree, 1976). M–O Firth of Clyde, Scotland: Ettrick Bay (M), Barassie (N), and Prestwick (O) (after Lambshead, unpubl.). P–Q Northeast coast of Britain: St. Abbs (P) and Newbiggin (Q) (Moore, 1973). R amalgamated samples from several localities on the north-east coast of Britain (Moore, 1971). Stars: amalgamated samples taken either from different areas or from the same area over a period of time; solid circles: samples from an area allegedly under stress

occasions throughout the year. One possible explanation for this is that by combining samples one more closely reproduces the a c t u a l log-normal population from which they were drawn.

An alternative explanation is that each individual sample might be governed by a single but different factor or a number of l i n k e d factors which overshadow the rest. But when such samples are amalgamated, these factors now act in a compounded multiplicative manner cancelling out individual effects. In

this way, samples from different non-log-normal populations might, when lumped together, approximate a log-normal distribution.

If this reasoning is correct, then there could be problems in using deviation from the log-normal for detecting pollution. For example, a large baseline survey where many independent populations were sampled and amalgamated could result in a log-normal distribution when analysed. But when smaller samples from the independent populations were monitored for

pollution effects, they might be found to 'deviate' from the log-normal model, thereby falsely giving the impression that the area had recently suffered pollution stress. In any case, the log-normal model should only be applied to large samples of species, although this is by no means always adhered to in practice.

Conversely, by mixing populations from polluted areas the erroneous conclusion might be drawn that the area was not suffering adversely from pollution because the resulting distribution appeared log-normal.

To summarize at this point, it seems to us that there is as yet relatively little evidence to show that, under natural conditions (i.e. unpolluted), marine benthic assemblages in temperate waters conform to the log-normal. However, there are theoretical reasons why data based on pooled samples might artificially simulate such a distribution. Since the data on which this opinion is based come primarily from intertidal or shallow sublittoral northern temperate areas, where 'normal' conditions might be described as relatively unstable, the foregoing observations may not hold true for the more stable environments of polar, tropical or deep seas.

A SIMPLE DOMINANCE INDEX

Gray and Mirza's (1979) method largely depends on a change in dominance to indicate pollution. However, there are simpler and more sensitive methods of detecting dominance in an assemblage of organisms: one is the rank species abundance (RSA) curve. Species are ranked in order of abundance, and species abundance is plotted against species rank. In practice, the rarer species may safely be omitted from the plot.

Platt (1977) investigated the nematodes at 3 stations, high, mid and low tide, on an unpolluted intertidal sandflat in Strangford Lough, Northern Ireland. He speculated that there was a cline governed by increasing physical stress (or disturbance) from low to high tide. It might be expected therefore that the high tide station should more clearly exhibit the qualities of a disturbed assemblage. If Platt's data are plotted as log-transformed abundances (according to Barnett, 1975) on probability paper (Fig. 6) then the plots obtained are remarkably similar for all 3 stations. However, dominance is detectable in the high tide plot by the increase in the number of geometric classes spanned and the levelling of the plot from Classes IX to XI caused by the most dominant species outstripping the rest by three geometric classes. If Platt's (1977) data are plotted as RSA curves (Fig. 7) the progressive increase in dominance from low to high tide is more readily apparent.
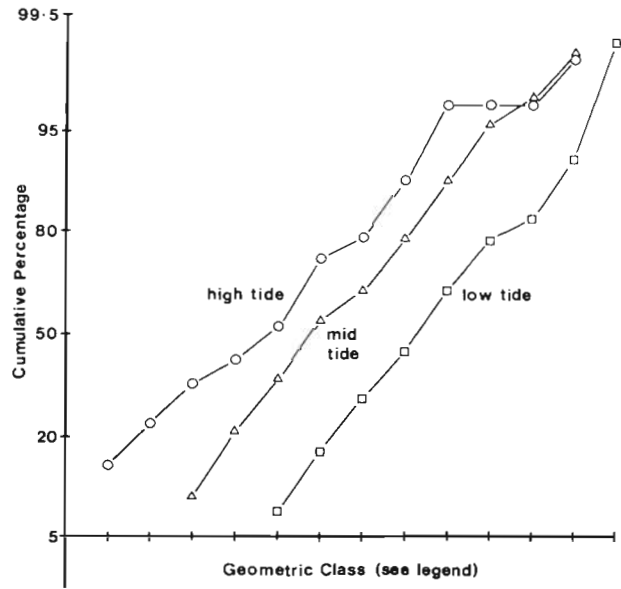

Fig. 6. Probability plot of data from Strangford Lough, N. Ireland (after Platt, 1977). X-axis indicates x 2 geometric class; first point on each curve represents Class I (i.e. the curves are staggered)
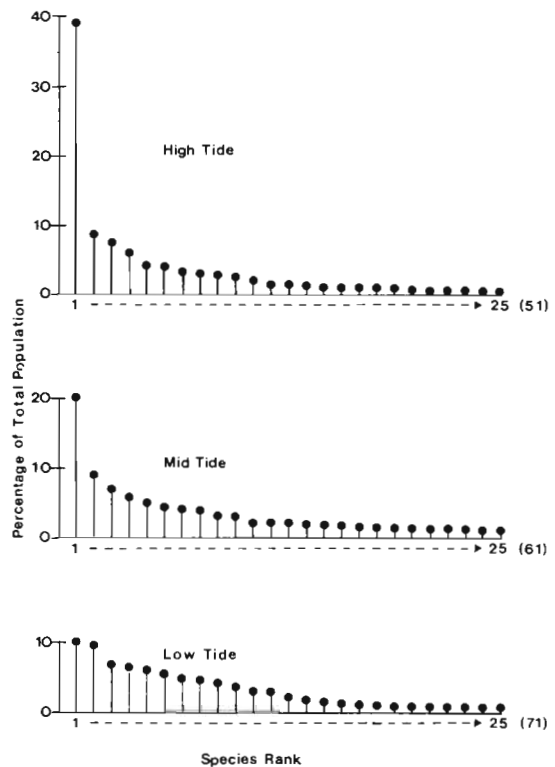

Fig. 7. RSA curves of the 25 most common species from Strangford Lough (after Platt, 1977) data with the total number of species shown in parentheses. Figure after Platt and Warwick (1980) but note that in the original figure the y-axis was wrongly labelled

Fig. 8 shows data on the fauna from kelp holdfasts recorded by Moore (1973) from 2 areas, Newbiggin and St. Abbs. Newbiggin is supposedly polluted and there is considerable dominance; roughly two-thirds of the individuals belong to a single species. In contrast, the
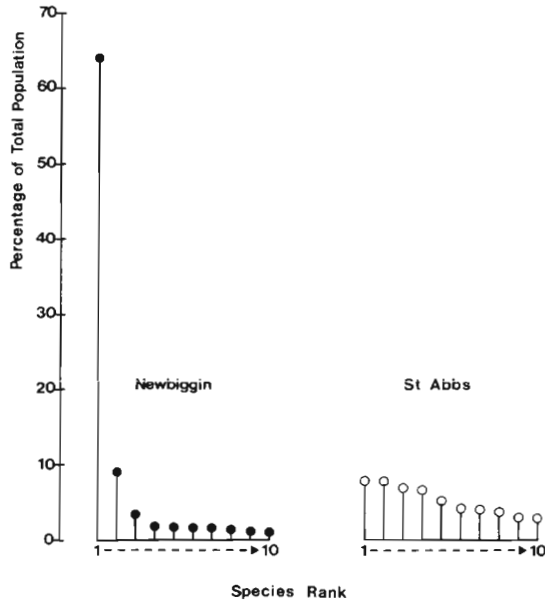

Fig. 8. RSA curves of kelp holdfast fauna (after Moore, 1973) with only the 10 most common species plotted

unpolluted St. Abbs fauna exhibits little dominance, there being little difference in abundance among the more common species. The RSA curves clearly show the enormous difference between the 2 samples. But when these same data are plotted on probability paper (Fig. 9) the 2 plots are remarkably similar. Although increased dominance in the Newbiggin fauna can be detected by the greater number of geometric classes
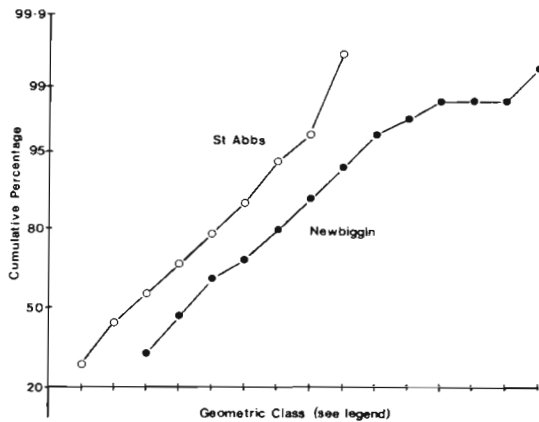

Fig. 9. Probability plot of data from kelp holdfasts (after Moore, 1973). X-axis as in Fig. 6

spanned and the levelling-off caused by the larger difference in abundance between the commonest and the next most abundant species, the effect is far less clear than in the RSA plot.

RSA curves have the advantage of being both very sensitive to dominance changes and mathematically simple and easy to apply. Furthermore, they ease the problem of species identification since, unlike attempts to generate data for complicated diversity indices, only the most common species need be determined.

But it would obviously be convenient to be able to express dominance as a single number. The simplest way of achieving this is to express the abundance of the most common species as a percentage of the total sample. Fig. 10 shows this simple dominance index
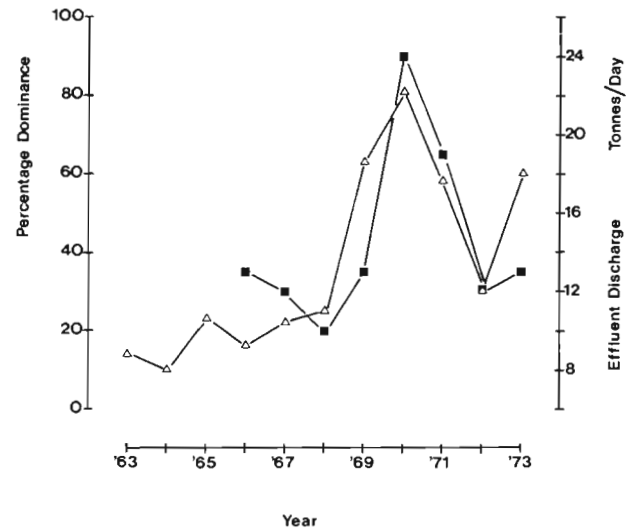

Fig. 10. Dominance index (△) and effluent discharge (■) for the years 1963–1973, from Loch Eil macrobenthos data (after Pearson, 1975). Pollution began in 1966, there being no discharge before this date

plotted against effluent discharge taken from the results obtained by Pearson (1975: Station 2) in his study of the long-term effects of pulp mill discharge on the macrobenthos of Loch Eil. There is remarkable conformity between the 2 plots, except for the final 1973 point where the species richness had been severely reduced. Fig. 11 shows the same data plotted as log-transformed abundances. These complete results reveal a confusing picture but it is clear that the curves are being strongly influenced by dominance as indicated by the flat-spots. It is interesting that Gray and Mirza (1979) interpret the 1967 and 1968 curves as a 'transitory phase' and draw a broken line: we would question whether these curves are any less deserving of a straight line being fitted than those either before or
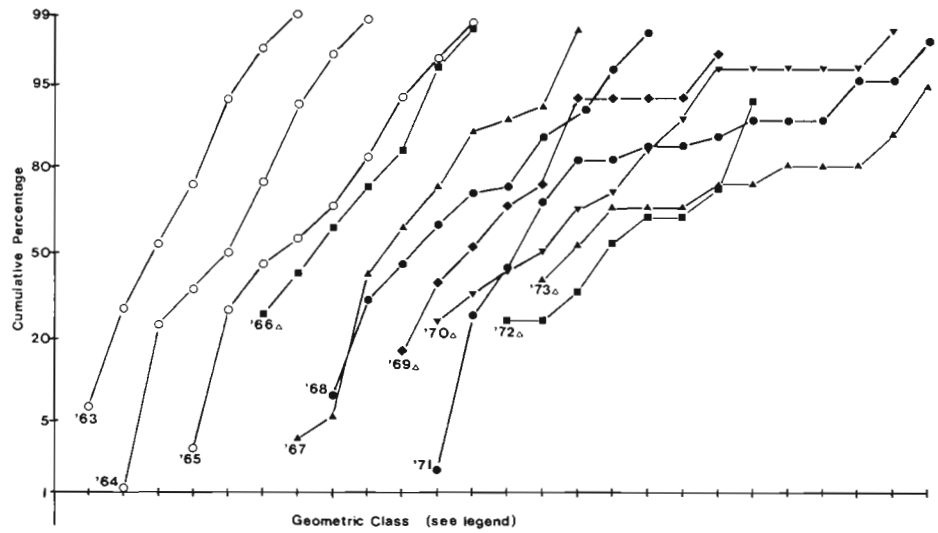
Fig. 11. Probability plot of Pearson's (1975) data. X-axis represents x 2 geometric class; first point on each curve represents lowest class observed (Class II if indicated by △, Class I otherwise). Open circles (1963–1965): data prior to effluent release; curves with solid symbols: data obtained after discharge began

after. Surprisingly, however, Gray and Mirza (1979) do fit a straight line to Pearson's 1975 data.

The dominance index for freeliving nematodes from a number of different environments is shown in Table 1. Just as the RSA curves vary considerably in nature, so do the indices vary in magnitude. But for generally similar habitats, dominance does appear to be higher in areas which are alleged to be polluted or otherwise stressed.

However, in the light of our current state of knowledge of marine nematode ecology, a dominance index would only be of use in pollution monitoring where data from reference areas or base-line surveys were available for comparison. But it is to be hoped that when more data are available it might prove possible to predict the expected dominance index for the ne-

matode assemblage typical of any given natural environment.

We make no claim of sample-size independence for this index. From the nematode data collected in the Firth of Clyde, it would appear empirically that a minimum sample size of about 150 individuals is required. As macrofauna is less diverse than meiofauna, possibly a larger sample size would be required to generate equivalent macrofauna data.

Similarly, the index is not independent of species richness. Theoretically, minimum dominance occurs when all the species have equal abundances of N/S, whereupon the dominance index takes its minimum value of $1/S$. However, this value is not obtained in nature and the effect appears unimportant for S greater than about 30 species.

Table 1. Dominance index for freeliving marine nematode assemblages from several localities

Habitat	Locality	Allegedly polluted	Dominance index %	Source
Kelp holdfast	St. Abbs, North Sea	No	7.6	Moore (1973)*
Kelp holdfast	Newbiggin, North Sea	Yes	64.1	Moore (1973)*
Kelp holdfast	72 samples, NE coast	Mostly	43.9	Moore (1971)
Mid-tide mudflat	Lynher estuary, Cornwall	No	14.8	Warwick and Price (1979)
Low-tide fine sand	Strangford Lough, N. Ireland	No	10.3	Platt (1977)
Mid-tide fine sand	Strangford Lough, N. Ireland	No	19.9	Platt (1977)
High-tide fine sand	Strangford Lough, N. Ireland	No	38.8	Platt (1977)
Low-tide fine sand	Irvine Bay, Firth of Clyde	Yes	19.8	Jayasree (1976)
Mid-tide fine sand	Irvine Bay, Firth of Clyde	Yes	34.6	Jayasree (1976)
High-tide fine sand	Irvine Bay, Firth of Clyde	Yes	48.0	Jayasree (1976)
Sublittoral mud	Morlaix Bay, Brittany	Yes	21.0	Gourbault (1981)
Low-tide fine sand	Prestwick, Firth of Clyde	Yes	16.4	Lambshead (unpubl.)
Low-tide fine sand	Barassie, Firth of Clyde	Yes	17.1	Lambshead (unpubl.)
Low-tide fine sand	Ettrick, Firth of Clyde	No	19.9	Lambshead (unpubl.)

* Moore's (1973) data includes other taxa in addition to nematodes

CONCLUSIONS

A number of conclusions may be drawn from this study, all of which have to be more fully tested before any degree of universality can be assigned.

(1) Normal probability paper is a poor tool for the unequivocal detection of the goodness-of-fit of biological assemblages to a log-normal distribution of species abundances. It is also insensitive in detecting changes in cumulative species distribution patterns of such assemblages.

(2) Apparent changes in cumulative species distribution patterns may be the result of differences in the size or nature of samples.

(3) Natural assemblages of benthic marine organisms in shallow temperate environments do not usually conform to the log-normal distribution.

(4) A simple assessment of dominance might prove of use in detecting disturbed or polluted faunal assemblages.

LITERATURE CITED

Barnett, V (1975). Probability plotting methods and order statistics. Appl. Statist. 24: 95–108

Ferris, V R., Ferris, J. M. (1979). Thread worms (Nematoda). In: Hart, C. W., Fuller, S. L. H. (eds.) Pollution ecology of estuarine invertebrates. Academic Press, London. p. 1–33

Fisher, R. A., Corbet, A. S., Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. J. Anim. Ecol. 12: 42–58

Gourbault, N. (1981). Les peuplements de nématodes du Chenal de la Baie de Morlaix (premières données). Cah. Biol. mar 22: 65–82

Gray, J. S. (1976). Are marine base-line surveys worthwhile? New Scient. 70: 219–221

Gray, J. S. (1981). Detecting pollution induced changes in communities using the log-normal distribution of individuals among species. Mar. Pollut. Bull. 12: 173–176

Gray, J. S., Mirza, F. B. (1979). A possible method for detecting pollution-induced disturbances in marine benthic communities. Mar. Pollut. Bull. 10: 142–146

Jayasree, K. (1976). Free-living marine nematodes from polluted intertidal sand in Scotland. Unpublished Ph. D. thesis, Aberdeen University

Krebs, C. J. (1972). Ecology. Harper and Row, New York

May, R. M. (1975). Patterns of species abundance and diversity. In: Cody, M. L., Diamond, J. M. (eds.) Ecology and evolution of communities. Belknap Press, Cambridge, Mass. p. 81–120

Moore, P. G. (1971). The nematode fauna associated with holdfasts of kelp *(Laminaria hyperborea)* in north-east Britain. J. mar. biol. Ass. U.K. 51: 589–604

Moore, P. G. (1973). The kelp fauna of northeast Britain. II. Multivariate classification: turbidity as an ecological factor. J. exp. mar. Biol. Ecol. 13: 127–163

Ortner, P. B., Hill, L. C., Cummings, S. R. (1982). Variation in copepod species assemblage distributions. The utility of the log-normal approach. Mar. Pollut. Bull. 13: 195–199

Pearson, T. H. (1975). The benthic ecology of Loch Linnhe and Loch Eil, a sea-loch system on the west coast of Scotland. IV. Changes in the benthic fauna attributable to organic enrichment. J. exp. mar. Biol. Ecol. 20: 1–41

Platt, H. M. (1977). Ecology of free-living marine nematodes from an intertidal sandflat in Strangford Lough, Northern Ireland. Cah. Biol. Mar. 18: 261–273

Platt, H. M., Warwick, R. M. (1980). The significance of free-living nematodes to the littoral ecosystem. In: Price, J. H., Irvine, D. E. G., Farnham, W. F. (eds.) The shore environment, Vol. 2. Ecosystems. Academic Press, London, p. 729–759

Preston, F. W. (1948). The commonness and rarity of species. Ecology 29: 254–283

Soulsby, P. G., Lowthion, D., Houston, M. (1982). Effects of macroalgal mats on the ecology of intertidal mudflats. Mar. Pollut. Bull. 13: 162–166

Stenseth, N. C. (1979). Where have all the species gone? On the nature of extinction and the Red Queen hypothesis. Oikos 33: 196–227

Ugland, K. I., Gray, J. S. (1982). Log-normal distributions and the concept of community equilibrium. Oikos 39: 171–178

Warwick, R. M., Price, R. (1979). Ecological and metabolic studies on free-living nematodes from an estuarine mud flat. Estuar. coast. mar. Sci. 9: 257–271

Williams, C. B. (1964). Patterns in the balance of nature. Academic Press, London

Erratum

Re: Shaw, Lambshead and Platt, *Mar. Ecol. Prog. Ser. 11: 195–202, 1983*

● On p. 201, 'Pearson's 1975 data' should read 'the 1970 data of Pearson (1975); see Fig. 11'.