

NOTE

Use and misuse of the log-normal plotting method for detection of effects of pollution – a reply to Shaw et al. (1983)

J. S. Gray

Universitetet i Oslo, Institutt for marinbiologi og limnologi, Avdeling for marin zoologi og marin kjemi, P. B. 1064 Blindern, Oslo 3, Norway

ABSTRACT: In contrast to Shaw et al. (1983) I believe the log-normal distribution of individuals among species is an adequate description of most benthic communities, given that no community should ever be expected to be in complete equilibrium and give a perfect fit. Under organic enrichment departures from a log-normal distribution have been shown to give a consistent pattern and are not highly confusing as Shaw et al. contend. The trends must be treated cautiously however, since pollution is not the only factor leading to a lack-of-fit to a log-normal.

In the preceding paper, Shaw et al. (1983) suggest that (1) there is little evidence to substantiate the belief that benthic marine communities fit the log-normal distribution of individuals among species; (2) the plotting method of detecting deviation from a log-normal distribution is of dubious value; (3) a simple dominance index is a more practical method of assessing pollution induced changes.

In attempting to refute the generality of the log-normal distribution Shaw et al. (1983) use data from Gray and Mirza (1979) and to this plot a log-normal line which they claim has the same standard deviation and mean as the data. The log-normal is not a single-fixed distribution, and for a given number of species and individuals there are many possible log-normal distributions. If one transforms their data (their Fig. 3) to probits and fits a linear regression, the line obtained ($R^2 = 0.934$) is I submit, a good fit to a log-normal as claimed in our original paper. Shaw et al.'s plot clearly has divergent mean (50 % point) and standard deviation (84 % to 50 % points) than our data, and bears little relation to the data. I believe this evidence shows that Station A in Oslofjord does fit the log-normal (Fig. 1).

The original model fitted to the distribution of individuals among species was R. A. Fisher's log-series which is characterized by having the highest number of species in the class with 1 individual per species. Later it was found that more usually the class with 2 to

3 individuals per species was the dominant class and a new model, the log-normal distribution, was derived by Preston (1948). Yet if the sample is not large the mode will often lie in the 1-individual-per-species class (a log series) but as the sample size is increased the mode moves along to the 2 to 3 individual class or higher giving a log-normal. Preston (1962) has an excellent example showing this when sample size was increased from 3 mo data to 6 mo to 12 mo. Thus the

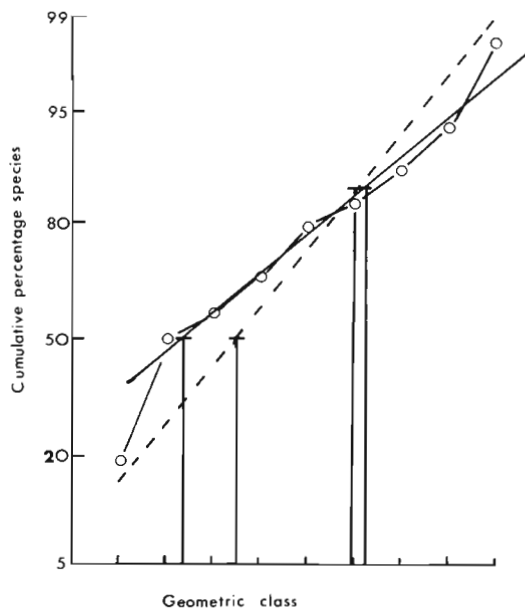


Fig. 1. Oslofjord data from Gray and Mirza (1979). Station A. Solid line: probit plot of cumulative % species against geometric class of individuals per species omitting first point ($y = 4.7172 + 0.2675 \times R^2 = 0.934$). Broken line: Shaw et al.'s (1983) fit to the same data purporting to show the same mean and standard deviation. Vertical line at 50 % point shows the difference in means from the 2 plots, vertical line at 84 % shows the difference in standard deviation (84 to 50 % points) for the 2 plots

distinction between a log-series and log-normal distribution is not always clear and in my experience, within the limits of the chi square goodness-of-fit test, it is usually impossible to distinguish between the 2 models. As the log-normal is the simpler theoretical model (and possibly the more general) Occam's razor dictates use of a log-normal. In my view there is little to be gained by saying that in one data set a log series is a better fit, as done by Shaw et al. (1983) for Williams' lepidopteran data (Fig. 4 in their paper).

The log-normal has been found to be an almost universal model for fits of individuals among species in data as divergent as snakes (in Panama), lizards, breeding birds, phytoplankton and zoobenthos. Plots of many data sets led me to conclude that it was a generally acceptable model for benthic marine communities, provided that the community sampled was heterogeneous, sample size large and the community could be expected to be in some form of equilibrium. Shaw et al.'s (1983) data for nematodes in Strangford Lough (Fig. 6 in their paper) show excellent fits to the log-normal, as does Moore's data from the unpolluted St. Abbs (their Fig. 9) which Gray and Mirza (1979) used. Similarly in their Fig. 5 there are many data sets which give good fits to the log-normal. It seems to me pointless to deny the generality of the log-normal when it has been shown to fit universally so many data sets. The interest is in why some data sets do not fit which was my starting point. Amongst others I cited Moore's data for Newbiggin showing how it diverges from a straight-line log-normal (shown in Shaw et al.'s Fig. 9 and previously in Gray and Mirza, Fig. 4).

The most substantial support for the practicality of the log-normal plotting method as a means of detecting pollution-induced disturbance is Shaw et al.'s Fig. 11 using Pearson's data, which is more complete than Gray and Mirza's (Fig. 3). To quote from Gray and Mirza (1979, p. 143): 'In the prepollution phase the data follow a log-normal distribution almost perfectly and the data span few geometric classes. The transitory phase is characterized by a bend in the straight line log-normal distribution and an increase in the number of geometric classes spanned. The polluted phase

returns to a typical log-normal distribution with a straight line, but the slope of the line is at a less steep angle than in the unpolluted data and furthermore the data now span many more geometric classes (10-11 compared with 6).'

I contend the data show almost perfectly the things we claimed and disagree with Shaw et al.'s (1983) statement that Fig. 11 'reveals a confusing picture.' Shaw et al. claim that the dominance-index method is preferable but their Fig. 10 shows that first in 1969 does this index show an influence of pollution. The log-normal plots showed changes already in 1967 and 1968. Using other methods (Gray and Pearson, 1982) based on the same technique, species sensitive to pollution can be isolated objectively, which is a further rationale for using the log-normal method.

The plotting method is primitive and Gray and Pearson (1982) revert back to simply plot number of species against individuals per species in geometric classes. I agree with Shaw et al. (1983) that the technique should be used carefully and critically and there is a danger of misuse, but I believe there is little evidence presented in their paper to substantiate the claim that the log-normal does not in general fit benthic data and the plotting method is of dubious value.

LITERATURE CITED

- 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
- Gray, J. S., Pearson, T. H. (1982). Sensitive species indicative of pollution-induced change in benthic communities. I. *Comparative Methodology. Mar. Ecol. Prog. Ser.* 9: 111-119
- Preston, F. W. (1948). The commonness and rarity of species. *Ecology* 29: 254-283
- Preston, F. W. (1962). The canonical distribution of commonness and rarity. *Ecology* 43: 185-215
- Shaw, K. M., Lambshead, P. J. D., Platt, H. M. (1983). Detection of pollution-induced disturbance in marine benthic assemblages with special reference to nematodes. *Mar. Ecol. Prog. Ser.* 11: 195-202

Accepted for printing on December 11, 1982