

ABUNDANCE DISTRIBUTION AND THE EVOLUTION OF COMMUNITY STRUCTURE

Rudolf Harmsen

Biology Department, Queen's University
Kingston, Ontario K7L 3N6 Canada

Received October 12, 1983; December 6, 1983

ABSTRACT: A recent development of considerable importance in the study of community structure is the publication of Sugihara's (1980) niche hierarchy model for resource partitioning. This model assumes minimal community structure, and leads to the prediction of a canonical log-normal abundance distribution of species within any taxocene. I here present a theoretical argument which proposes that more than minimal community structure will develop over evolutionary time in sufficiently stable environments. It is argued that individual species have a tendency to move towards "optimal abundance". This process will result in abundance distributions with a significantly lower value for σ than would be predicted by the niche hierarchy model.

Data from the literature and from some recent collections from temperate and tropical forest communities support the theoretical considerations. The evenness values of the distributions are significantly higher (lower σ values), so that the distributions are no longer canonical. Species abundance distributions from evolutionarily old communities may still fit the log-normal or some other mathematically described distribution, but such a fit is probably merely fortuitous, and has no biological basis.

*

*

*

INTRODUCTION:

Recently, the main focus in the study of community structure has been on the trophic interactions of natural communities, or the interactive dynamics of specific groups of species (e.g. Caswell, 1976; May, 1981; Mound and Waloff, 1978; Paine, 1980).

Where assemblages of organisms within single trophic categories (e.g. guilds or taxocenes) have been studied, the emphasis has been on the statistical methodology of representation rather than on the biological relevance of the taxocene structure (e.g. Bulmer, 1974; Kempton, 1979; Kempton, 1981; May, 1981). A consensus seems to be shaping that the distribution of relative abundances of species from any one taxocene does closely approach the log-normal (May, 1981; Pielou, 1975; Sugihara, 1980; Taylor, 1978; Whittaker, 1975) although agreement is far from universal (Kempton, 1975; 1981). This general finding applies to all gross taxa examined so far. There is, however, no consensus as to whether or not this log-normal distribution or any other mathematically defined distribution has any biological significance (e.g. May, 1981).

In this paper I examine the arguments concerning the biological meaning of the log-normal distribution, and present some data which seem to support the contention that the species abundance distribution is an important indication of the evolutionary state of the community.

SPECIES ABUNDANCE DISTRIBUTION

There is, of course, no reason to assume that a set of values extracted out of nature will always be distributed according to some mathematically describable pattern. However, when such a pattern does emerge, it is important to determine which of three reasons lead to its appearance. First, it may be a rough approximation only,

*

*

*

Evolutionary Theory 6: 283-292 (December, 1983)

The editors thank two referees for help in evaluating this paper.

© 1983, the author

and the pattern is merely in our mind; in the case of abundance distributions many collections fit the log-normal, the negative binomial or the log series equally well; at least two of these "fits" must be completely fortuitous. Second, the pattern may be a statistical property of the natural system being studied, in which case the biological significance may be low. Third, the pattern may reveal the presence of important biological phenomena. Only in the latter case will the study of changes in the observed pattern be a really worthwhile biological activity - one which may tell us something about the nature and dominance of the biological interactions at hand.

Preston (1948; 1962) introduced a format for representing species abundance distribution which not only summarizes the data simply and clearly, but also allows for easy analysis and comparison (see Fig. 1). He also concluded that, for birds and moths at least, the distribution was log-normal. Other distributions have at times

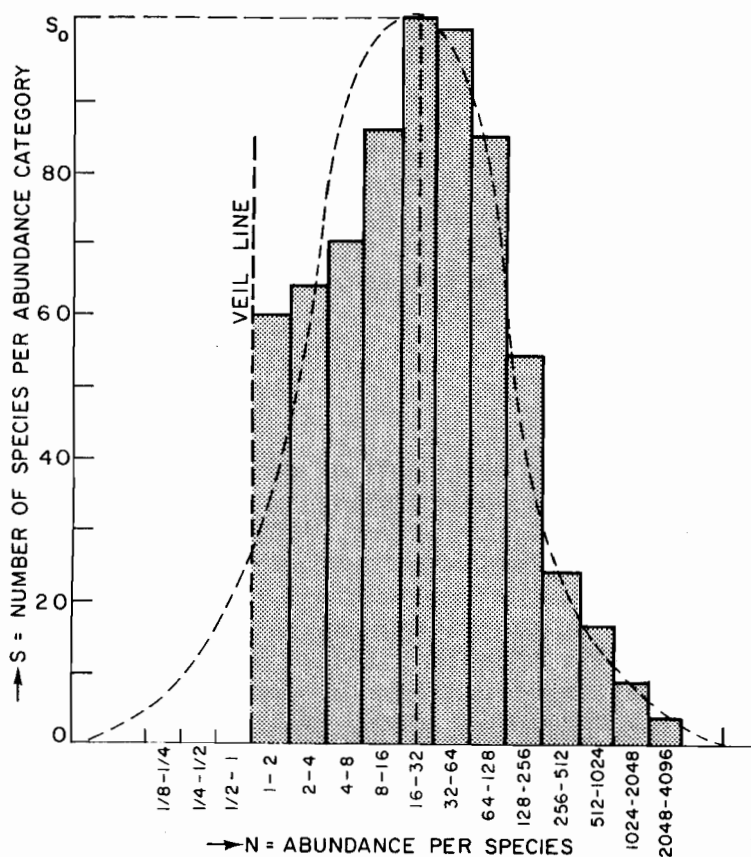


Figure 1. Method of representing quantitative samples (after Preston, (1948)). All species are placed in a bar graph according to their abundance (when the species falls on the border between two abundance categories it is scored half in each category; singletons are scored as a half in the 1-2 category, the other half is not scored, as it would be "lost" in the 1/2-1 category). To the left of the "veil-line" there are no observations, as the species that should be there are so rare that on the average the sample will not include them. For each doubling of the sample size the veil line will shift one abundance category to the left. Only for samples where the total number of species is large enough to obtain a fairly "smooth" bargraph (at least 100) and where the sample size is large enough to have the veil line well to the left of the mode (for moths, at least 4000 individuals) can a log-normal curve be fitted with confidence (Bliss, 1967; Pielou, 1975).

This particular example represents the combined Perth Road collection of 66,484 specimens, consisting of 701 species. N_0 = abundance of modal species; S_0 = the number of species at the mode of the distribution.

been proposed, and some actual field collections appeared to fit the negative binomial or log-series distributions better than the log-normal (Brian, 1953; Kempton and Taylor, 1974), but a further analysis of existing data, and the collection of larger data sets covering more taxa, have led to the general acceptance of the log-normal distribution as the one which fits most existing data sets best (May, 1981; Pielou, 1975; Taylor, 1978; Whittaker, 1975), although there are, at times, compelling practical reasons for the use of the log-series, especially for small samples (Taylor, 1978) and for samples from severely disturbed sites (Patrick, 1973; Kempton, 1981). Indeed, the fit of large samples of species rich communities is so close to a truncated log-normal distribution, that the possibility of a mere superficial likeness must be rejected.

This general conclusion is not at all surprising, as the determination of the abundances of a large number of species within a community is governed by many independent factors, a situation which should be covered by the central limit theorem. Thus, statistically speaking, a log-normal distribution is to be expected (Whittaker, 1972). It is, however, still an open question as to whether any biologically important conclusions can be extracted out of what may be no more than the result of a probability function operating in a diverse system.

The log-normal distribution can be expressed (after Preston, 1948; 1962) as follows:

$$S_{(R)} = S_o \underline{a}^2 R^2 \quad (1)$$

where S_o is the number of species at the mode of the distribution (see Figure 1), R is the logarithm to base 2 of N/N_o , and \underline{a} is an expression of the width, or of the slope of the gaussian curve, defined as follows:

$$\underline{a} = 1/(2^{\frac{1}{2}} \sigma) \quad (2)$$

It must be clear from the above that the numerical relationships among the species of a taxocene as encountered in a natural community are defined by S_o and \underline{a} as long as we assume a log-normal distribution of relative abundances. May (1976) develops this point in more detail. If we assume (not necessarily correctly) that the truncated left tail of the distribution also follows a log-normal distribution, the total number of species of the taxocene under study, S^* , can be calculated. S^* is a measure of the species richness of the taxocene.

The measure \underline{a} is mathematically, if not biologically, independent of S^* , and represents the evenness or equitability of species abundances in the taxocene. Any set of measures of species abundances within a defined group of organisms, therefore, will provide us with two biological measures: the species richness (S^*), and the community evenness (\underline{a}). The accuracy and reliability of these measures depends on 1) the degree to which truncation has been eliminated (i.e. large sample size), 2) there being enough species to produce a smooth approximation to a gaussian curve and 3) both the time period and/or the space over which the collection is made being homogeneous.

MacArthur (1957) first attempted to arrive at some definable abundance distribution through an actual mechanism of resource partitioning among the constituent species of a community (e.g. the broken stick model), but the resultant distributions do not match reality well at all. A series of subsequent workers in the area have introduced more biological realism into the mechanism of resource partitioning (e.g. Bulmer, 1974). Some of these later models lead to a log-normal distribution. The latest proposed model, which I will refer to as the "niche hierarchy" model (Sugihara, 1980), assumes sequential random breakage of a multidimensional "stick", and leads to not only a log-normal distribution, but the particular one parameter family of log-normal distributions which appears to coincide with most, if not all distributions observed in nature.

Preston (1948; 1962) already noted that the measure of the slope (\underline{a}) of the log-normal curve, was consistent among various communities, with a value of $\underline{a} \approx 0.2$. As \underline{a} is a measure of the equitability or evenness of species abundance, it became a matter of speculation that this constancy in value was of biological significance (Hutchinson, 1953). May (1975; 1976) has pointed out that the relative constancy of \underline{a} could be merely a mathematical property of the log-normal distribution, based on a general mathematical insensitivity to even fairly major structural differences in biological community composition. Whereas May's explanation will undoubtedly contribute towards low variability in values of \underline{a} , Sugihara's explanation is much more attractive, in that it is based not only on a biologically acceptable mechanism of resource partitioning, it also says something of biological importance: if the value of \underline{a} falls within a certain range, the community which was sampled has "minimal community structure" only. That is, the resources were partitioned according to a random pattern, similar to a random breaking of a stick within a minimal set of rules. Figure 2 is an adaptation of Sugihara's figure, and shows that all reported abundance distributions have evenness values (here represented by σ) that fall within one standard deviation of the value predicted by the niche hierarchy model of resource partitioning.

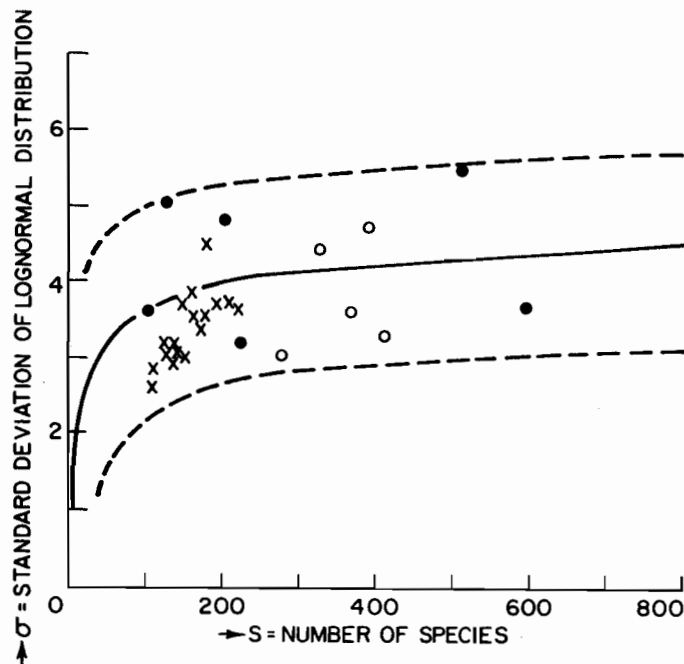


Figure 2. The relationship between the breadth (σ) of the lognormal distribution curve and the number of species (S) (after Sugihara, 1980), as predicted by the niche hierarchy resource partitioning model. The solid line is the mean prediction, the dotted lines are one standard deviation from the mean. Real-world collections are represented by ● birds, x diatoms, and ○ moths. Only those taxocenes with samples greater than 100 species are represented. It is clear from the figure that the Sugihara hypothesis is supported by all real-world collections.

*

*

*

EVOLUTION OF COMMUNITY STRUCTURE

It is unsettling that all reported abundance distributions that have reliably fitted log-normal curves (i.e. those based on taxocenes with more than 100 species) have values indicating minimal community structure only. Such a situation leads to the conclusion that once established, a community does not respond to the evolution-

ary changes in the qualitative parameters of its constituent species or, put differently, that the numerical relationships of communities represent a stable, interactive system, which does not change over time. If this were indeed true, it begs an explanation. However, it is not necessarily true. Certainly, in the case of all reported moth collections, the communities that were sampled were either naturally young and/or disturbed. Clearly, a minimally disturbed community, in a naturally rich and old region should be sampled and analysed in order to establish if community evolution does take place, and how it affects community structure.

In developing a theoretical argument for some form of community evolution, one would perhaps best start from a point of view directly derived from the niche-hierarchy model of resource allocation, and assume that rare species are rare because they have a narrowly defined resource base, whereas common species have a broad resource base. This is an important implication, since the efficiency of resource utilization is dependent on the resource component utilized. At very narrowly defined resources, the suitable resource is so rare that much efficiency is lost due to high search time investments. At the other end of the spectrum, very common species rely on such a common set of resources that their resource utilization efficiency suffers from having to depend on variable foraging methods, broad enzyme spectra, and the like. Two opposing selective forces must at all times affect the fitness of resource specificity (or niche breadth?), in much the same way as migratory and

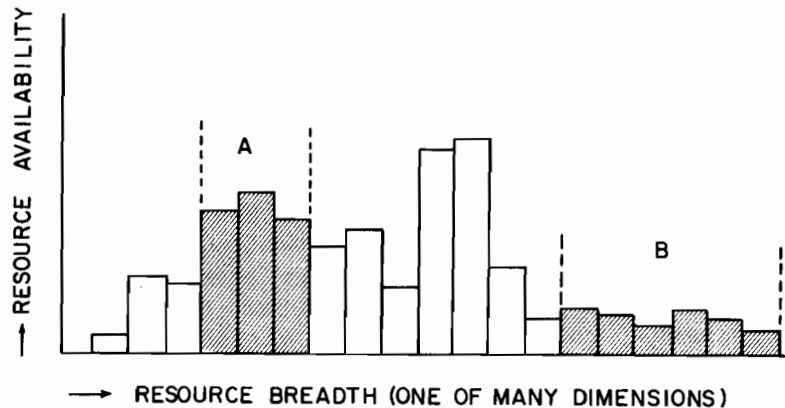


Figure 3. A two dimensional representation of an hypothetical resource spectrum, aimed at defining some terms. Resource availability refers to the density of useable resource of any one type. Resource breadth refers to the diversity of resource types available to a taxocene, e.g. types of foliage or types of prey, and how different they are. "A" indicates a high volume of a more or less homogeneous resource; "B" indicates a part of the resource spectrum where the availability is insufficient to provide a species with an adequate volume or quantity of homogeneous resource.

*

*

*

congregatory behaviours are balanced in the Δ -model of Taylor and Taylor (1977, 1979), but at an evolutionary level rather than a behavioural one. This line of argument would suggest the existence of something like an "optimal abundance": very rare species tend to be under strong selection to broaden their resource base (resulting in increased abundance), very common ones are under the opposite selective pressure, namely to narrow their resource base (resulting in decreased abundance). A variety of complicating factors will, undoubtedly prevent the existence of one single optimal abundance for all species of one taxocene, the main one being the unevenness, or "lumpiness" of the resource base: some relatively homogeneous resources can have very high "volumes", whereas in other regions of the resource spectrum the

availability may be very small (see Figure 3). Even if optimal abundances do exist, one would not necessarily expect species to reach such values, as during the gradual approach to optimality, the environment will change, partly independently of the utilizing species, but partly as a direct result of competitive and exploitative co-evolutionary interactions with other species.

All in all, however, one would expect the variance, or breadth of the distribution curve (e.g. σ) to reduce over time (or a to increase). This would be especially true in regions of predictable climate, where stable interactions and reliable resources will allow fine evolutionary adaptations in the absence of major, short term ecological fluctuations. The theory sketched out above, therefore, predicts values of σ significantly lower than those predicted by the niche hierarchy model for those communities that are old, in predictable environments, and undisturbed. A number of quantitative collections, falling within the above specifications have recently been made, and I will deal with their analysis.

THE DATA BASE

My analysis is restricted to nocturnal macrolepidoptera, i.e. species belonging to the following superfamilies (as used by Common, 1970): Geometroidea, Bombycoidea, Sphingoidea, Notodontoidea, and Noctuoidea. This is a taxonomically arbitrary, but ecologically meaningful assemblage as most species have phytophagous larvae, one or two generations a year, and individual energy requirements of the same order of magnitude. The group forms a taxocene in the traditional sense of the word (Hutchinson, 1967). The collections were made in 1) a slightly disturbed forest locality near Perth Road, 2) a somewhat more disturbed locality near Chaffey's Locks, and 3) an agricultural site near Glenburnie, all within 30 km of one another in southeastern Ontario (for detailed data sets and collecting conditions, see Ward *et al.*, 1974). For comparison, additional data from disturbed sites were obtained from the literature (Dirks, 1937; La France, 1968; Preston, 1948 and pers. comm.; Williams, 1939). Some very recent collections were made in low montane tropical rainforest in Papua New Guinea, one such a collection (Bulolo Gorge) is sufficiently analysed to be included here. Unfortunately, the various collections referred to in Table 1 were not all made with exactly the same collecting method. This may add a hidden variable to the data, however, I do not believe this to be a serious problem.

RESULTS OF DATA ANALYSIS

For each of the collections available in sufficient quantitative detail, parameters of the distributions were estimated using the continuous approximation method of Bliss (Bliss, 1967; Pielou, 1975). Although strictly speaking the Poisson log-normal is more appropriate, the error introduced with large samples (with modal abundance at least one octave to the right of the origin) is less serious than with smaller samples. Using the 1934 Rothamstead data, Bulmer (1975) found little difference between the grouped log-normal fitted by Bliss and the Poisson log-normal fitted by maximum likelihood. Kempton and Taylor (1974) provide an example in which the two methods gave discrepant results, but this was based on a small sample with modal abundance hovering around the veil line.

All currently available moth data are summarized in Table 1, which gives the number of species collected, the number of individuals in the sample, and the evenness parameters σ and a , as well as a habitat characterization for each collection. The combined Perth Road collection is graphically represented in Figure 1.

Of the Ward *et al.* (1974) collections, the one from the agricultural site falls well within the range of evenness parameters of previously published collections. Its value (3.30) is lower than the value predicted (4.36) by the niche hierarchy model, but not significantly so, (see Figure 4). The other collections, however, do not conform to the previously set pattern. All the values are much lower than those

Table 1. Estimated log-normal parameters for large (> 10,000 individuals) collections of macrolepidoptera.

Site and data source	Sample Size		σ	\underline{a}	Habitat
	Individually	Species			
Saskatoon (Preston 1948 and pers. comm.)	87,110	279	4.22	.167	Grassland - probably disturbed
Lethbridge, Alberta (Preston, 1948 and pers. comm.)	303,251	291	4.17	.170	Agricultural Station
Orono, Maine (Dirks, 1937)	43,909	342	3.42	.207	Open field in foreground, timbered lot
Glenburnie, Ontario (Ward et al., 1977)	17,226	427	3.30	.214	Old field- hawthorn scrub with remnants of original vegetation
Rothamsted, England (Williams 1939)	15,595	241	2.85	.248	Cultivated area, agricultural station
Chaffey's Locks, Ontario (Ward et al., 1974)	18,146	571	2.72	.260	Beech-maple forest with tracts of pine - hemlock and some open pasture land nearby
Perth Road, Ontario Collection 1 (Ward et al., 1974)	10,818	469	2.53	.280	Beech-maple and pine-oak-hickory forest
Perth Road, Coll. 2	12,966	488	2.48	.286	" " " "
Perth Road, Coll. 3	21,684	612	2.50	.283	" " " "
Perth Road, Coll. 4	21,016	595	2.55	.277	" " " "
Perth Road Combined	66,484	701	2.61	.271	" " " "
Bulolo Gorge, Papua New Guinea (Herbert, pers. comm.)	3,843	632	1.94	.360	Albizia - Araucaria forest, in forest reserve, appr. 1000m altitude.

of previously published moth collections, and fall significantly below the values predicted by the niche hierarchy model (see Figure 4). This is especially true for the New Guinea collection. Note, however, that the New Guinea collection is fairly small, so that the precision of parameter estimation is low.

DISCUSSION

It seems that much of the discussion that has centered upon the apparently universal canonical nature of log-normal species abundance distributions (e.g. Hutchinson, 1953; May, 1975; 1976; 1981; Sugihara, 1980) has been led astray by the fact that older, undisturbed communities are not so easily sampled as agricultural stations or road sides. It is obvious now, that the question itself, i.e. "Why are all species abundance distributions more or less canonical?" is in error. The forest lepidopteran taxocenes reported here, have abundance distributions

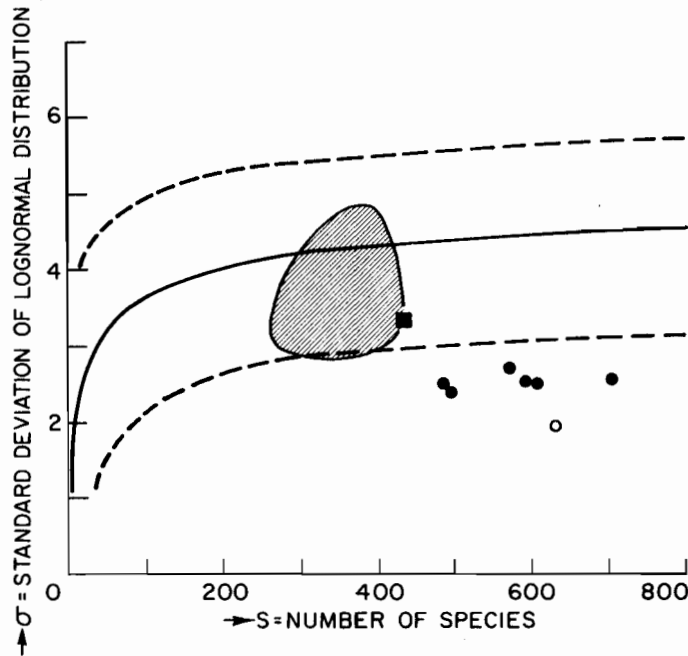


Figure 4. The Sugihara prediction values of σ , as in Figure 2. Only the moth data are retained, as the shaded area. The Glenburnie (disturbed) collection is added (■), as well as all the Ontario forest collections (●), and the Bulolo Gorge collection (○). Clearly, the niche hierarchy model does not adequately cover the undisturbed forest community collections. More than minimal community structure exists in the communities.

*

*

*

that are very different from the canonical. This illustrates two points. Firstly, May's (1976) argument that the apparent canonical nature of the observed abundance distributions is merely the result of the mathematical properties of the log-normal is false. Undoubtedly, the sluggish response of parameters such as σ to biological changes in community structure reduces the chance of finding communities that deviate from the canonical, but the fact that such communities do exist proves that biological factors can modify community structure sufficiently to obtain σ values significantly different from the originally established values.

The biological suitability of the niche hierarchy model, the extreme closeness of its predictions to those of the canonical lognormal (Sugihara, 1980), and the observation that most taxocenes approximate the structure predicted by the model make the model a very appealing one. Do the currently reported data reject the model? I don't think so; they do, however, limit the model to one which describes only the

EVOLUTION OF COMMUNITY STRUCTURE

initial establishment of community structure. I suggest that a totally different biological mechanism takes over during the long period of community evolution. The theoretical arguments presented earlier cover the kind of mechanism that must be responsible for the changes that occur in the species abundance distribution of a taxocene while it constitutes a part of a stable community. If such a community becomes disturbed (naturally, or by man), a partial or complete return to a structural state such as the one predicted by the niche hierarchy model will take place.

It is somewhat ironical, at this stage, to once again return to the log-normal distribution. Sugihara (1980) has convincingly presented the case for minimal community structure being based upon taxocene abundance distributions that are not only log-normal, but of a very specific type of log-normal. The apparent shift away from the originally established community structure through complex evolutionary processes such as the ones presented above, leaves us with communities of higher levels of organization. The taxocenes in such communities display species abundance distributions that have shifted considerably away from some assumed original pattern, yet, these observed distributions appear to fit the truncated log-normal pattern reasonably well (see Fig. 1). Such fairly close fits, however, may have no theoretical basis, especially since all collections still show a good deal of truncation, hiding possible major skews. As suggested above, the evolutionary mechanisms affecting the high abundance and low abundance tail ends of the distribution curves are different in nature. One would predict, therefore, that as communities evolve further and further away from their original structure, the true log-normal base will be more and more obscured. It leaves the researcher with a dilemma: we can only study community evolution by comparing parameters that assume a log-normal or other mathematically defined distribution, while we are fairly certain that, in highly structured communities, at least, such defined distributions no longer exist.

ACKNOWLEDGEMENTS

I am especially indebted to Paul Hebert and Philip Ward, my co-workers in this entire project, and to the members of the Queen's University EEB Group for extensive discussions and comments on the earlier drafts of the manuscript. Further thanks are extended to George Sugihara and to an anonymous reviewer for helpful suggestions for improvement. Financial support from the Queen's University Advisory Research Committee is gratefully acknowledged.

REFERENCES

- Bliss, C.I. (1976) Statistics in Biology. Vol. 1. McGraw-Hill, New York.
- Brian, M.V. (1953) Species frequencies in random samples from animal populations. *J. Anim. Ecol.* 22: 57-64.
- Bulmer, M.G. (1974) On fitting the Poisson log-normal distribution to species-abundance data. *Biometrics* 30: 101-110.
- Caswell, H. (1976) Community structure: A neutral model analysis. *Ecol. Monogr.* 46: 327-354.
- Common, I.F.B. (1970) Lepidoptera. In: The Insects of Australia (ed). Waterhouse, D.F. Melbourne U.P. Melbourne.
- Dirks, C.O. (1937) Biological studies of Maine moths by light trap methods. *Maine Agric. Exper. Stn. Bulletin* 389.
- Hutchinson, G.E. (1953) The concept of pattern in ecology. *Proc. Acad. Nat. Sci. Philadelphia* 105: 1-12.
- Hutchinson, G.E. (1967) A Treatise on Limnology. Wiley, New York.
- Kempton, R.A. (1975) A generalized form of Fisher's logarithmic series. *Biometrika* 62: 29-38.

- Kempton, R.A. (1979) The structure of species abundance and measurement of diversity. *Biometrics* 35: 307-321.
- Kempton, R.A. (1981) The stability of size ordinations in ecological surveys. In: Hiorns, R.W. and Cooke, D. (ed.) The Mathematical Theory of the Dynamics of Biological Populations II. pp. 217-230. Academic Press, London/New York.
- Kempton, R.A. and Taylor, L.R. (1974) Log-series and log-normal parameters as diversity discriminants for the Lepidoptera. *J. Anim. Ecol.* 43: 381-399.
- La France, J. (1968) The nocturnal insect catches at predetermined time intervals in the organic soil district of Ste. Clotilde, Southwestern Quebec. *Ann. ent. Soc. Quebec* 13: 32-54.
- MacArthur, R.H. (1957) On the relative abundance of bird species. *Proc. Nat. Acad. Sci. Wash.* 43: 293-295.
- May, R.M. (1975) Patterns of species abundance and diversity. In: Ecology and Evolution of Communities (eds.) Cody, M.L. and Diamond, J.M. Harvard U.P. Cambridge, Mass.
- May, R.M. (1976) Patterns in multi-species communities. In: Theoretical Ecology, Principles and Applications (ed.) May, R.M. Blackwell, Oxford.
- May, R.M. (1981) Patterns in multi-species communities. In: Theoretical Ecology, Principles and Applications (ed.) May, R.M. (second edition). Blackwell, Oxford.
- Mound, L.A. and Waloff, N. (eds.) (1978) Diversity of Insect Faunas Blackwell, Oxford.
- Paine, R.T. (1980) Food webs: interaction strength, linkage, and community infrastructure. *J. Anim. Ecol.* 49: 667-685.
- Patrick, R. (1973) Use of algae, especially diatoms, in the assessment of water quality. American Society for Testing and Materials, Special Tech. Publ. 528: 76-95.
- Pielou, E.C. (1975) Ecological Diversity. Wiley, New York.
- 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 and 410-432.
- Sugihara, G. (1980) Minimal community structure: an explanation of species abundance patterns. *Am. Nat.* 116: 770-787.
- Taylor, L.R. (1978) Bates, Williams, Hutchinson - a variety of diversities. In: Diversity of Insect Faunas (eds.) Mound, L.A. and Waloff, N. Blackwell, Oxford.
- Taylor, L.R. and Taylor, R.A.J. (1977) Aggregation, migration and population mechanics. *Nature, London* 265: 415-421.
- Taylor, R.A.J. and Taylor, L.R. (1979) A behavioural model for the evolution of spatial dynamics. In: Population Dynamics. (eds.) Anderson, R.M., Turner, B.D. and Taylor, L.R. Blackwell, Oxford.
- Ward, P.S., Harmsen, R. and Hebert, P.D.N. (1974) An annotated checklist of macroheterocera of south-eastern Ontario. *J. Res. Lepidoptera* 13: 23-42.
- Whittaker, R.H. (1972) Evolution and measurement of species diversity. *Taxon* 21: 213-251.
- Whittaker, R.H. (1975) Communities and Ecosystems. Macmillan, New York.
- Williams, C.B. (1939) An analysis of four years captures of insects in a light trap. Part I. General survey; sex proportion; phenology; and time of flight. *Trans. r. ent. Soc. Lond.* 89: 79-130.

*

*

*