The Community Concept and the Phytoplankton

W. T. Williams[1], J. S. Bunt[1], R. D. John[2] and D. J. Abel[3]

[1] Australian Institute of Marine Science, P.M.B. No. 3, Townsville, Queensland 4810, Australia
[2] CSIRO Division of Mathematics and Statistics, Davies Laboratory, Private Bag, Aitkenvale, Qld. 4814, Australia
[3] CSIRO Division of Computing Research, Davies Laboratory, Private Bag, Aitkenvale, Qld. 4814, Australia

ABSTRACT: The community concept of terrestrial ecology is examined with particular reference to its application to phytoplankton distribution. It is suggested that it is more appropriate to regard phytoplankton distribution as entirely probabilistic, comparable to a local situation in a tropical rain forest. The oceans of the world would then be regarded collectively as a single essentially uniform environment inhabitated by a single diverse population. The probability of encountering any given species in any ocean might then be very low but would be nowhere zero. The thesis is examined mathematically with respect to data from the diatom literature, and its further implications are briefly discussed.

INTRODUCTION

All writers on the subject are agreed that the distribution of diatoms is not uniform over the world oceans. Yet difficulties arise from attempting to define the nature of this non-uniformity, so that different systems of description have been proposed for different oceans, resulting in a number of conflicting terminologies. A general outline of the early history of the problem was given by Patrick (1948). Some of the more recent literature has been covered in reviews by Margalef (1969) and Guillard and Kilham (1977). We may begin with Cleve (1897), who recognized associations of phytoplankton species in the North Atlantic Ocean which he termed 'plankton types', and to which he gave descriptive names which have virtually disappeared from the literature. It was pointed out by Gran (1912) that these types, as defined, were mutually exclusive, whereas in reality species distributions overlap. Valentine (1973) uses the terminology of terrestrial plant ecology, defining 'provinces' as 'regions in which communities maintain characteristic composition', and 'communities' as 'recurrent associations' of species 'with the proviso that the association is flexibly defined and need only include a certain percentage of the species in some definitive set of species in order to qualify'. Kozlova and Mukhina (1967) prefer the term 'biocoenoses' for the groups of species they recognized in the Pacific. Braarud et al. (1953) use the term 'communities' for phytoplankton assemblages which characterized particular watermasses in the North Sea, but clearly recognized its shortcomings and used it interchangeably with others, such as 'populations' and 'societies'.

The overlapping nature of such diatom-assemblages is generally accepted, if only as a result of agreement that some species are 'cosmopolitan' – i. e. recorded at some time or another from most of the world oceans. In contrast, there has been a discernible tendency to search for endemism, to establish that certain species are strictly confined to particular areas. Patrick (1948), for example, claimed that endemism could be as high as 10–30 %. It is thus not surprising that attempts to define relatively stable assemblages appear to have lost their earlier wide attraction, so that modern work has often tended to concentrate on the autecology of individual species or groups of commonly-associated species.

Nevertheless, ignoring a problem does not cause it to disappear. Any sample obtained during a phytoplankton survey will contain an assemblage of diatom species; and in this paper we address ourselves to the double question, 'What is the nature of such an assemblage, and what should it be called?' It is difficult to avoid the suspicion that much of the thinking on this subject has been coloured by the pervasive and successful work in temperate terrestrial plant ecology. This has encouraged the widespread use of a term such as 'community'; and Edsbagge (1966a, b) has gone so far as to suggest that the techniques of the Zurich-

Montpelier school of phytosociology should be applied to the study of epiphytic diatoms. We consider that in any such study it is first necessary to examine the historical implications of the term 'community'.

The Community Concept in Terrestrial Ecology

The term 'community' has engendered a formidable literature in plant ecology, of which good reviews were given (Shimwell, 1977; Whittaker, 1978). The term is very old; it can be traced back to Humboldt (1805). The concept implied simply that certain assemblages of plant species tended to occur together, and could be so recognized on the ground. It follows that the basic concept is d i s c o n t i n u o u s distribution; different communities could be recognized as such, even though there might be some gradation in composition at inter-community boundaries. That such a model will serve as an adequate working basis for terrestrial temperate plant ecology is shown by the success of the European investigations into phytosociology, notably those associated with the Zurich-Montpelier and Uppsala schools (the literature on these is too extensive to cite here, but can be found in the historical references cited above). Its efficacy is also borne out by the success of modern methods of numerical classification, such as those of Williams and Lambert (1960) and of many subsequent workers.

The suggestion that such discontinuities in plant distribution are the result of discontinuities in the climatic or edaphic environment can be traced back as far as Drude (1896). It is implicit in the terminology of British ecologists such as Moss (1910) and Tansley (1939, 4th impression 1965), whose use of terms such as 'limestone heath', 'chalk grassland' and 'reedswamp' imply that a community is an assemblage of plant species growing in a particular edaphic environment. It follows that a 'community' in this sense is also d e t e r m i n i s t i c ; given a particular environment, a particular group of species will be expected. Even successional changes do not destroy the deterministic concept. Webb et al. (1972), referring to the classic work of Watt (1947), point out that, in a monospecific temperate forest, once the dominant has begun to establish, 'there is only one path open, and such patchiness as exists is composed of different age-classes of one species'.

The assumption of discontinuity has not, however, been universally accepted. In particular, Gleason (1926) has argued that – apart from such major discontinuities as change in geological formations – vegetation is everywhere continuously varying, a view supported later by Goodall (1954). This viewpoint became increasingly relevant as attention turned to multi-species systems such as the tropical rain forest. Such forests characteristically have very high diversity; the number of tree species involved may be of the order of 1 000 to 2 000. These species are not arranged in the tidy blocks of temperate forests; as Aubreville (1938) remarked, they appear to consist of local, apparently random, assortments of species. Nor is there any evidence of regular, ordered succession; as Webb et al. (1972) point out, 'the succession is characterized by departures from orderliness'. Such departures had been stigmatized by Gleason (1926) and Watt (1964) as f o r t u i t o u s , and by Daubenmire (1968) as u n p r e - d i c t a b l e. Aubreville (1938), however, regarded the situation as analagous to linear algebra when the number of equations (environmental parameters) is less than the number of independent variables (plant species), so that an indefinite number of alternative solutions will fulfill the required conditions. An important contribution to the discussion was made by Goodall (1970), who stated: 'The relation of a species to ecotope space should be considered, not in terms of a discrete cell which the species occupies, but of a diffuse region which could be expressed as a nested set of contour shells representing successively greater probabilities of occurrence'.

Finally, Webb et al. (1972), discussing Aubreville's thesis, pointed out that the solutions to his equations were never c o m p l e t e l y unpredictable, in that it was usually possible to define boundary conditions outside which the solution could not lie. They therefore postulated that the tropical rain forest has a two-level structure: a large-scale deterministic framework, largely climatically determined, and a small-scale patchiness which was entirely probabilistic, in that the species-composition of any given patch was determined by which species happened to be flowering at the time the canopy opened to permit regeneration. We may therefore distinguish 2 types of organization: temperate terrestrial vegetation, almost completely deterministic, and tropical rainforest, with a large-scale deterministic framework and a small-scale probabilistic structure. We shall suggest that diatom distribution terminates this sequence, in that it is entirely probabilistic.

The New Model

We suggest that, so far as phytoplankton distribution is concerned, the oceans of the world should be regarded collectively as a single ocean, so that the concept of endemism ceases to be meaningful; we also suggest that, although the probability of collecting any diatom species in an ocean may be extremely low, it is never zero.

This model has implications. If we were prepared to make the crude assumption that the factors affecting the abundance of any diatom on any occasion were independent, the likelihood of capture would always take the form $\pi_j\, p_j$, in which the p_j values represent the individual probabilities. We cannot hope to know these; but we can at least consider four areas of action in which they must fall. Since diatoms differ in their rate of growth and reproduction, we can regard each diatom as being associated with a probability p_1 which is unique to that species. However, some diatoms are larger and heavier than others, and some begin their lives attached, though they may later appear in the plankton. The likelihood will therefore involve a second probability p_2, which will be higher in coastal or other shallow waters. A third probability, p_3, can be postulated as subsuming all environmental sensitivities; these will include those many factors of salinity, temperature and the like which bulk large in the current ecological literature. They will, however, also include the effects of predation; and Wroblewski and O'Brien (1974) have investigated the appropriate mathematical models. They conclude that the combined effects of reproduction and predation will result in the formation of 'patches' of finite extent. The situation is thus remarkably similar to that of the tropical rainforest, but seen under tremendous time-acceleration. Finally, there is a 'catch-probability' p_4, according to whether a collector does or does not encounter a 'patch' at a particular time.

It follows that, as in the rainforest, a particular assemblage of species does not represent a temperate terrestrial-type determinate 'community'; it is always a temporary assemblage which may never precisely recur.

It also follows that the probability of collecting any diatom anywhere will rise with the intensity of sampling. Apparent endemism would then be regarded simply as an artefact of inadequate sampling. For example, *Nitzschia kerguelensis*, long believed to be confined to Antarctic waters, is now known to extend to 30 °N (Van der Spoel et al., 1973); Round (1973) provides other cases. Furthermore, if we select any 2 areas in the world, then the more intensively they are both sampled, the more species they may be expected to have in common.

We can offer some circumstantial evidence for this phenomenon from Australian collections. Revelante et al. (1981a) provide a list of 396 taxa of diatoms and dinoflagellates collected over some years from a section of the Great Barrier Reef. Dr. N. Revelante (personal communication) has also collected extensively in the Adriatic Sea, for which she was able to identify unequivocally 122 species of diatoms and dinoflagellates; of these, no less than 101 also appear in the

Australian list. It is even possible to extend such a comparison to a world scale. Guillard and Kilham (1977) provide a comprehensive set of lists of 'prominent' species gathered from the world literature. We were able to extract some 53 lists; these overlapped geographically to a considerable extent, and we have reduced their areas of origin to the following 16: 'Cosmopolitan', Arctic, Antarctic, Western coastal North Atlantic, North Central Atlantic, North-Eastern Atlantic, Central Atlantic, Mediterranean, Atlantic coasts of tropical America, Caribbean Sea and the Gulf of Mexico, Indian Ocean, North Pacific, Central and tropical Pacific, Equatorial, coasts of tropical eastern Pacific, and non-tropical coasts of Eastern Pacific. The lists included a total of 265 species, of which 115 also appear in the Australian diatom list; and these 115 were scattered through all the 16 areas, including even a substantial representation in the Arctic list.

Mathematical Considerations

We think it unlikely that a rigorous test of our hypothesis will ever be possible; the logistic problems involved in the collection and identification of phytoplankton species are such that it is doubtful whether the detailed species-area curves which would be required will ever become available. We can, however, use the Guillard-Kilham areas as a rough test of whether the species distribution is in accordance with expectation. We have postulated that there are no inter-oceanic discontinuities, so that the whole of the world oceans would collectively behave as a 'small area' in the sense of C. B. Williams (1964). Out of the 265 species, we should then expect the number found in 1, 2, 3 etc. of the 16 areas to form a logarithmic sequence of the type αx, $\alpha x^2/2$, $\alpha x^3/3$ where α is the 'index of diversity' and x a parameter dependent on sample size. It must be emphasized that a satisfactory fit would not constitute a proof of our hypothesis. There are other models which would produce the same result; Zipf (1949) has stressed the ubiquity of the logarithmic sequence in a wide range of classificatory situations; and it is conceivable (though we think highly unlikely) that species occurring in only one of the areas might be true endemics. Nevertheless, f a i - l u r e to fit such a sequence would, we suggest, cause misgivings as to the validity of our model. As an extreme counter-example, we would not expect the results to fit a Poisson distribution; this would imply that the occurrence of all diatoms was everywhere substantially equiprobable, which would be biologically absurd.

Fitting, of either the Poisson or logarithmic distributions, is not completely straightforward, since the data

are truncated. In fitting the Poisson we have to take into account the fact that, in data such as these, there can be no zero class. If m is the overall mean, and p_i the probability of i occurrences, we fit the truncated Poisson distribution with

$$p_i = \frac{e^{-m}}{1-e^{-m}} \cdot \frac{m^i}{i\,!}$$

If S is the number of species (here 265) and N the total number of records (here 578), the parameters α and x of the logarithmic distribution are conventionally obtained by solving the equations $S = \alpha\,ln(1 + N/\alpha)$ followed by $x = N/(N + \alpha)$. With the present data we obtain $\alpha = 189.4$ and $x = 0.75319$. However, Fisher (1943) has pointed out that this procedure rests on the relationship

$$S = \sum_{n=1}^{\infty} \frac{\alpha}{n}\, x^n = -\alpha\,ln(1 - x)$$

It therefore implies that the sequence is open-ended, whereas in the data we are using, n cannot exceed 16. It would therefore be preferable to fit a truncated distribution, with the summation taken to 16 instead of infinity. The results then obtained are $\alpha = 187.37$, $x = 0.75737$.

The fitted values, down to $n = 8$, for the truncated Poisson, the indefinite logarithmic, and the truncated logarithmic distributions are given in Table 1. It will be obvious that the Poisson fit is impossibly bad; the indefinite logarithmic fit is good, and the truncated logarithmic fit marginally better (X^2 for the two latter, consolidating $n = 13$ to $n = 16$ into a single class, were 12.25 and 11.95 respectively; neither is significant). The results are thus not at variance with the expectations of the model; we wish to claim no more than this.

The Problem of Pattern

Even if we are correct in assuming that phytoplankton distributions are entirely probabilistic, this does

Table 1. Fitting of 3 distributions to the Guillard and Kilham (1977) data (fitting shown down to $n = 8$)

No.	Observed	Truncated Poisson	Indefinite logarithmic	Truncated logarithmic
1	140	92.54	142.65	141.91
2	55	81.89	53.72	53.74
3	27	51.76	26.98	27.13
4	13	23.70	15.24	15.41
5	10	8.69	9.18	9.34
6	11	2.65	5.76	5.89
7	3	0.69	3.72	3.83
8	2	0.16	2.45	2.54
Rest	4	–	–	–

not imply that phytoplankton data are structureless; the probability-patterns will vary in space and time, and may serve to recover some environmental information. Our Australian experience confirms this; Revelante et al. (1981 b, c) have used numerical techniques (classification, ordination and graph-theory approaches) to analyse phytoplankton data from the Great Barrier Reef, with ecologically meaningful results. We have considered that it might be instructive to regard the data we have extracted from the Guillard and Kilham (1977) review as a set of 16 stations defined by the presence or absence of 265 diatom species. We can then define a distance-matrix for all pairs of stations. Since the data-matrix contains many zero entries we have elected to use the 1-complement of the Jaccard measure; this, in the conventional (a, b, c, d) notation of a 2×2 contingency table, is the quantity $(b + c)/(a + b + c)$. We have now used the distance-matrix to compute both a minimum spanning tree (Prim, 1957) and a two-neighbour network (Williams, 1980) to display the configuration of the 16 stations. The results are shown in Fig. 1 and transposed over a world map in Fig. 2. Of the 2 analyses, the network contains the larger amount of information; in particular, revealing several weak affinities which the tree is
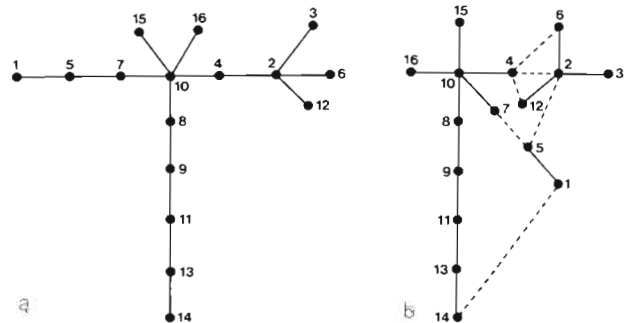

Fig. 1(a). Minimum spanning tree of 16 ocean regions considered as stations based on diatom species distributions. (b) Two-neighbour network of 16 stations, as in (a)

unable to display. The uniformity of the strong affinities is noteworthy. For the weaker affinities, the figure displays only their existence but not their strength. However, it is interesting to note a link between the equatorial and cosmopolitan groups as well as those related to the Arctic group, including predictably strong links with the Northeastern Atlantic and the North Pacific but also with the Antarctic. At the same time, the network indicates only weak affinity between the phytoplankton diatoms of the Central and North Central Atlantic. The Caribbean, Gulf of Mexico group on the other hand, occupies a strong central position. No doubt the reader may find other relationships of interest.
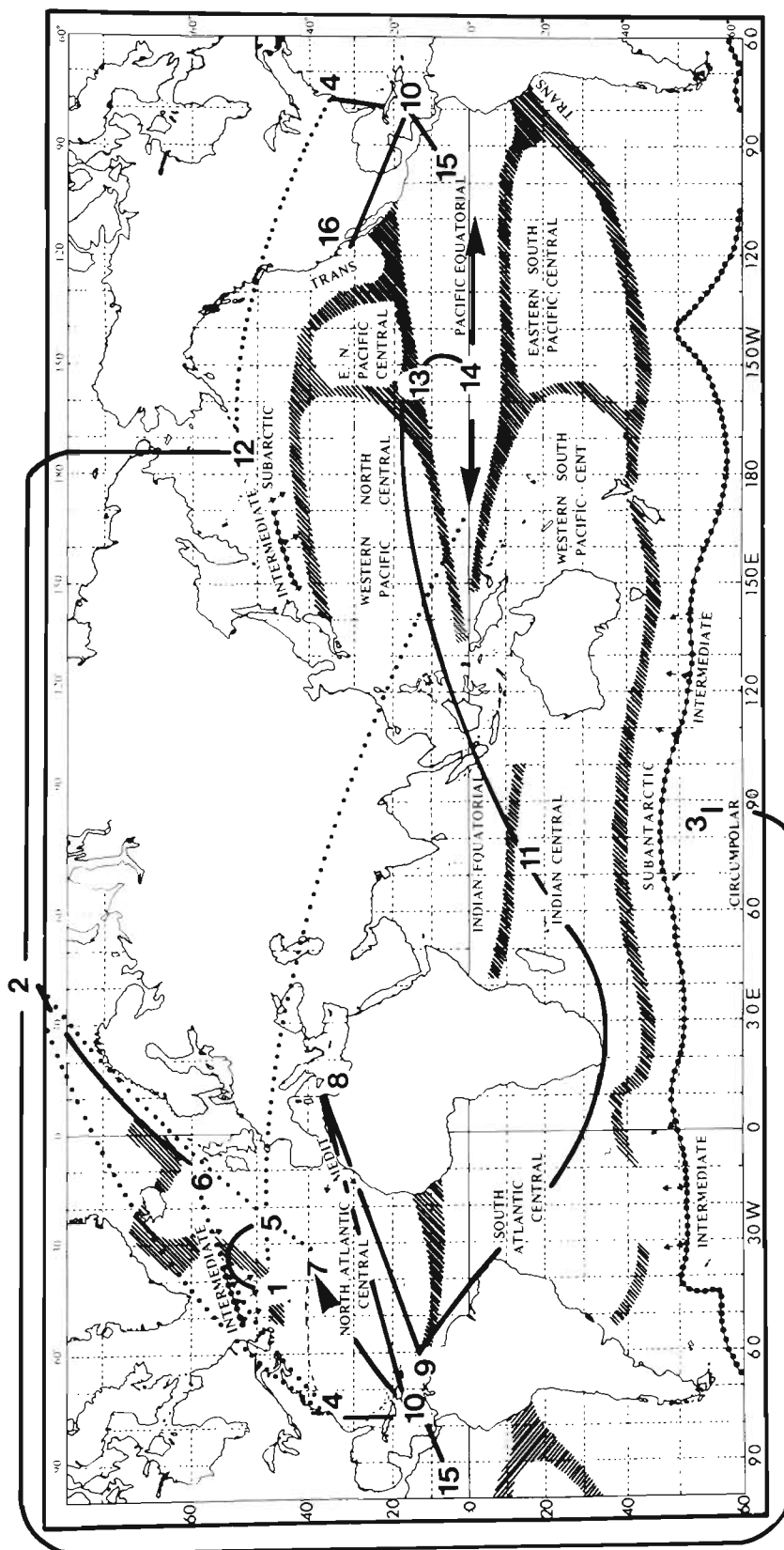
Fig. 2. Map of the world incorporating the network of Fig. 1b

CONCLUSIONS

In this paper we have argued the thesis that the distribution of planktonic diatoms in the seas and oceans of the world is that to be expected from a single population in an essentially uniform environment. We suspect that this conclusion is likely to apply to groups other than the diatoms and that phytoplankton distributions are fundamentally and generally probabilistic as in the case locally in terrestrial rainforests. This view is not completely novel; it could be regarded as a logical extension of the views of workers such as Margalef (1969).

As we have explained, it does not follow that the existence of patterns in phytoplankton species distributions is denied. It is quite obvious from the literature that such patterns are clearly recognizable. Their causes, nonetheless, are most unlikely to be determinate except perhaps at much grosser taxonomic scales. We consider it inappropriate, therefore, to apply the term 'communities' to what are, in effect, more or less transient assemblages or associations having substantial, if variable, affinities between them.

The model we offer transcends issues which otherwise might be regarded as merely semantic and, we believe, has ecological substance. Practically, the model may aid phytoplankton researchers who must come to terms not only with the limitations of available methods, but with environmental and biological circumstances which are so basically distinct from those on land where the meaning of the term 'community' can largely be rid of ambiguity. We recommend that a neutral term such as 'species assemblage' should be used in future, without any implication of other than stochastic association unless the nature of that association can be specified.

LITERATURE CITED

Aubreville, A. (1938). La foret coloniale: les forets de l'Afrique occidentale francaise. Annls Acad. Sci. colon. 9: 1–245

Braarud, T., Gaarder, K. R., Grontued, J. (1953). The phytoplankton of the North Sea and adjacent waters in May, 1948. Rapp. P.-v. Reun. Cons. Perm. Int. Explor. Mer 133: 1–87

Cleve, P. T (1897). Phytoplankton of the Atlantic and its tributaries, P. T. Cleve, Uppsala, Sweden

Daubenmire, R. (1968). Plant communities, Harper & Row, New York

Drude, O. (1896). Deutschlands Pflanzengeographie, Engelhorn, Stuttgart

Edsbagge, H. (1966a). Distribution notes on some diatoms not earlier recorded from the Swedish west coast. Botanica Mar. 11: 54–63

Edsbagge, H. (1966b). Zur Ökologie der marinen angehefteten Diatomeen. Botanica gothoburg 6: 9–139

Fisher, R. A. (1943). A theoretical distribution for the apparent abundance of different species. In: Fisher, R. A., Corbet, A. S., Williams, C. B. (eds.) 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

Gleason, H. A. (1926). The individualist concept of the plant association. Bull. Torrey Bot. Club 53: 7–26

Goodall, D. W. (1954). Vegetational classification und vegetational continua. Angew. PflSoziol. 1: 168–182

Goodall, D. W. (1970). Statistical plant ecology. Ann. Rev. Ecol. Syst. 1: 99–124

Gran, W. H. (1912). Pelagic plant life. In: Murray, J., Hjart, J. (eds.) The depths of the ocean. Macmillan, London, pp. 307–386

Guillard, R. R. L., Kilham, P. (1977). The ecology of marine planktonic diatoms. In: Werner, D. (ed.) The biology of diatoms. Blackwell, London, pp. 372–469

Humboldt, A. von (1805). Essai sur la Geographie des Plantes, Levrault, Paris

Kozlova, O. G., Mukhina, V. V. (1967). Diatoms & silicoflagellates in suspension and floor sediments of the Pacific Ocean. Int. Geol. Rev. 9: 1322–1342

Margalef, R. (1969). Crupos de especies asociadas en el fitophlancton del mar Caribe (N. E. de Venezuela). Invest. Presq. 33: 287–312

Moss, C. E. (1910). The fundamental units of vegetation. New Phytol. 9: 18–53

Patrick, R. (1948). Factors affecting the distribution of diatoms. Bot. Rev. 14: 473–524

Revelante, N., Williams, W. T., Bunt, J. S. (1981a). Phytoplankton in the waters of the Great Barrier Reef. I. Floristic composition and preliminary remarks. Mar. Ecol. Prog. Ser. (Submitted)

Revelante, N., Williams, W. T., Bunt, J. S. (1981b). Phytoplankton in the waters of the Great Barrier Reef. II. Species occurrences in time and space. Mar. Ecol. Prog. Ser. (Submitted)

Revelante, N., Williams, W T., Bunt, J. S. (1981c). Phytoplankton in the waters of the Great Barrier Reef. III. Seasonal changes, and the effect of Trichodesmium (Oscillatoria) spp.). Mar. Ecol. Prog. Ser. (Submitted)

Round, F. E. (1973). The biology of the algae, St Martin's Press, New York

Shimwell, D. W. (1977). The description and classification of vegetation, Sidgwick & Jackson, London

Tansley, A. G. (1939). The British Islands and their vegetation, University Press, Cambridge

Valentine, J. W. (1973). Evolutionary palaeoecology of the marine biosphere, Prentice-Hall, New Jersey

Van der Spoel, S. Hallegraeff, G. M., Van Soest, R. W M. (1973). Notes on variation of diatoms and silicoflagellates in the South Atlantic Ocean. Neth. J. Sea Res. 6: 518–541

Watt, A. S. (1964). The community and the individual. J. Ecol. 52 (Suppl.): 203–211

Webb, L. J., Tracey, J. G., Williams, W. T. (1972). Regeneration and pattern in the subtropical rain forest. J. Ecol. 60: 675–695

Whittaker, R. W (ed.) (1978). Classification of plant communities, Dr. W. Junk b. v., The Hague

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

Williams, W. T. (1980) Twonet: A new program for the computation of a two-neighbour network. Aust. Comput. J. 12: 70

Williams, W. T., Lambert, J. M. (1960). Multivariate methods in plant ecology. II. The use of an electronic digital computer for association-analysis. J. Ecol. 48: 689–710

Wroblewski, J. S., O'Brien, J. J. (1974). On the physical and biological scales of phytoplankton patchiness in the oceans. Proc. Sixth Liege Colloquium an Ocean Hydrody-namics. Mem. Soc. Ray. Sci. Liege. (Draft only sighted)

Zipf, G. K. (1949). Human behavior, Addison-Wesley Press, Cambridge, Massachusetts