THEORY AND MODELS IN VEGETATION SCIENCE

Advances in vegetation science   8

Edited by
EDDY VAN DER MAAREL

Theory and models in vegetation science

Proceedings of a Symposium, Uppsala, July 8–13, 1985

Edited by
I.C. PRENTICE and
E. VAN DER MAAREL

Reprinted from Vegetatio, volume 69

CONTENTS

Preface

July 8–13, 1985, an international group of scientists met in Uppsala for a symposium on the subject 'Theory and models in Vegetation science'. A volume of over 70 extended abstracts had already been published in time for the symposium (Leemans *et al.,* 1985). That volume included contributions from nearly all of those who gave talks or presented posters at the symposium. The present volume represents the fully-refereed proceedings of the symposium and features articles by a majority of speakers, plus a handful by poster authors, and two that were sent independently to Vegetatio and seemed timely and relevant to the symposium's theme.

As organizers, we tried to bring together for the symposium people whose interests covered several key aspects of modern vegetation science: vegetation dynamics, on shorter or longer time scales; the analysis of community data, and of vegetation-environment relationships in both time and space; and the functional basis of vegetation in terms of the individual plants and plant populations that it comprises. We encouraged contributors to focus on theory and models – not necessarily mathematical models, but also conceptual models that might contribute to the development of theory and mathematical models.

Reviewing these articles reinforced two somewhat opposing impressions. First, vegetation science (in our broad interpretation of the phrase) is not unified and has no established theoretical basis. Many of the more quantitative contributions draw on theory developed in cognate fields – for example population biology, climatology or ecosystems theory – while the theoretical papers dealing with the middle ground of vegetation science tend either to be very general and conceptual, or to be specifically concerned with data analysis rather than with functional aspects of the vegetation itself. The whole collection therefore seems extremely heterogeneous in approach and language. But on the other hand, we have an impression of great competence and vitality, and of a potential for rapid progress in new theoretical developments if the necessary cross-fertilizations can be made. The various conceptual papers should encourage us to combine our field studies and advanced data-analytical resources creatively with our understanding of the processes that make vegetation work. The theoretical advances that one foresees may owe little to 'theoretical ecology' as it is often (narrowly) defined, but the theory will be applicable, testable and indigenous to our subject. This is the task of the 'Working Group for Theoretical Vegetation Science' of the International Association for Vegetation Science, a working group newly established at the Uppsala symposium as the successor to the 'Working Group for Data Processing' who originally sponsored the symposium. The change of name reflects a shift in what the members of the old working group consider as the growing-point in the application of quantitative methods to the study of vegetation.

I. Colin Prentice
Eddy van der Maarel
Institute of Ecological Botany
Box 559 – S 751 22 Uppsala
Sweden

Relations between community theory and community analysis in vegetation science: some historical perspectives*

Imanuel Noy-Meir[1] & Eddy van der Maarel[2]*
[1]*Department of Botany, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel*
[2]*Institute of Ecological Botany, Uppsala University, Box 559, S 751 22 Uppsala, Sweden*

Abstract

Relations between community theory and community analysis are elucidated from a historical viewpoint. Four groups of relevant aspects are distinguished. This paper concentrates on aspects of spatial variation in species composition and species distribution relations.

Three major elements of community theory are discussed: the response of species to habitat conditions, interactions between species (competition, facilitation), and chance and random processes, both in the environment (disturbance) and in populations (dispersal). The debate island biogeography is used as a starting point for reviewing some theories about vegetation in terms of the elements they involve.

Early community theories are discussed with emphasis on Clements ('competition', 'facilitation') versus Gleason ('chance'), and Braun-Blanquet ('habitat') versus Du Rietz ('habitat', 'competition', 'chance'). Du Rietz's theory of vegetation is manysided and well ahead of its time. Then the development and significance of quantitative plant ecology for community theory is treated with emphasis on Whittaker ('habitat', 'competition') versus Goodall ('habitat', 'facilitation'). The potential use of objective and quantitative methods of analysis for hypothesis testing should be explored much further.

New theoretical ideas are summarized, including the role of the regeneration cycle of individual plants and neighbour interactions. Both interaction experiments and models for community dynamics will contribute to the general theory of plant communities. Perspectives for the development and testing of null hypotheses in vegetation science are discussed in view of the mutual relationships between community theory and community analysis, and the relations with dominance – diversity studies and multi-species pattern analysis indicated. The simplest hypothesis, i.e. that species combinations occur in sites independently and at random, is of little value. The next step would be that observed co-occurrences are determined by the overlap of species distributions over habitats. Then hypotheses on interaction and facilitation become relevant.

Introduction

Ever since ecology started as a science of its own,

plant ecologists have not been content with studying only one or few plant populations and species. Many have also tried to deal with 'vegetation', with entire plant communities consisting of many species. The main aspects of vegetation that have been of interest can be grouped as follows:

---

* The first author did part of the work for this paper during two stays at Uppsala. We thank Colin Prentice for critical comments on the manuscript.

1. structure, physiognomy, life forms; 2. diversity, dominance; 3. spatial variation in species composition, species distribution relations, scale of pattern; 4. temporal variation, stability.

Each of these aspects involves also the relations to the environment. Here we will focus on the third aspect, which has been of central interest in vegetation science for a long time. It concerns the 'phytosociological' structure of plant communities, the pattern or order that may be found in the co-occurrence and co-abundance of the various plant species and its variation in space.

We will adopt the following simple definitions for use in this contribution:
- species co-occurrence is the simultaneous presence of two or more plant species in the same unit of observation, irrespective of any expectation value.
- species association is any co-occurrence deviating from random expectation, either positively or negatively ('exclusion')
- species co-abundance is the simultaneous occurrence of two or more species with a certain abundance (however measured)
- species correlation is any co-abundance deviating from expectation.

Concerning each of these aspects of vegetation, ecologists have tried to derive statements at the community level empirically, by detecting patterns in the analysis of field data; and theoretically, by deducing them from processes and relations observed at the species-population level.

The two approaches, the theoretical and the empirical, are logically complementary and might have been expected to develop in close interaction. In this essay we would like to examine to what extent they have actually interacted in the history of vegetation science, in the specific context of phytosociology. The empirical analysis of species co-variation in space is usually based on observations on species occurrence and abundance in spatial units, sites or quadrats. The identification of patterns in a set of such observations can be done by a variety of methods: by informal judgement, by ordering of data tables, by calculation of similarity and association coefficients, by classification and ordination.

The theory of plant (and animal) communities is based mainly on three major elements:
a. the response of each species to habitat conditions; b. interactions between species: positive (facilitation), as well as negative (competition) and asymmetric (predation); c. random processes, both in the environment (fluctuation, disturbance) and in populations (dispersal, establishment, death).

In the following tabulations we will use the following codes to indic these elements: 'habitat' will refer to species responses; 'competition' and 'facilitation' to the types of interaction; 'disturbance' will be the code for all random processes in the environment; 'chance' will stand for all random biological processes. Theories of plant and animal communities differ mainly in the emphasis they put on each of these elements, in the importance they assume each of them to have in nature.

In the context of this paper two questions arise: what do different theories or hypotheses predict about observable patterns of species distribution, co-occurrence and co-abundance? And conversely: what can be inferred from such patterns about the value of different theories or hypotheses?

A digression into animal ecology

The above questions have been approached by animal ecologists in the last 20 years. Much attention has been given to island faunas. Two opposite views on the main processes structuring animal communities have been presented. MacArthur (e.g., 1972) has extensively and persuasively proposed that they are structured mainly by competition, leading either to exclusion or to niche- and character-displacement. Among many studies claiming to support this theory, of special interest for our theme is that by Diamond (1975) who presented an intensive analysis of data on the distribution and co-occurrence of birds in Melanesian islands, as evidence that the island faunas obey 'assembly rules' that result from competitive exclusion or displacement.

Simberloff (1978) and others severely criticized the evidence for the role of competition in structuring communities. Among other things they have reanalyzed Diamond's bird data, claiming that they can be explained by random colonization of islands (Connor & Simberloff, 1979). They argued that this is a simpler 'null hypothesis' of community structure (i.e. based on chance and random processes) which needs to be rejected before other hypotheses are considered (Strong, 1980).

One way to compare different theories and views about communities may be by comparing the relative importance they explicitly or implicitly assign to each of the main elements of community theory, as listed above. Schematically, these views can be represented as follows (+, + +, + + + suggest emphasis):

	habitat	competi-tion	facilita-tion	chance	disturb-ance
MacArthur-Diamond	+ +	+ + +			+
Simberloff-Connor	+ +			+ + +	

– number of + suggest strength of emphasis –

The debate that developed (e.g., Diamond & Gilpin, 1982; Connor & Simberloff, 1983; and the entire November 1983 issue of the American Naturalist) may have seemed irrelevant to most plant ecologists. However, it raised questions which need to be reconsidered in our field too.

Early plant community theory

In plant ecology there has not always been close interaction between community theory and analysis of observations. Different groups or even generations of plant ecologists have been much more interested in the one than in the other, and have developed it ahead more or less independently of the other. A brief historical review may be useful here. The first quarter of this century was the age of the great general theories of vegetation. It was also a period of development of systematic and more or less repeatable methods for collecting and analyzing data on the distribution of species in sites.

If we attempt to rate the leading plant ecologists of this period in terms of the relative emphasis they put on the main factors in community theory, we find that all of them considered species response to habitat to be dominant in structuring vegetation. The effect of spatial variation of the habitat on plant distribution is obviously ubiquitous and striking. However, Clements (1904) stressed positive interactions between plants as a second major force both in succession and in the maintenance and integration of stable communities. In his theory of the integrated community he was preceded by J. Paczoski, the founder of phytosociology (see Trass, 1975), and with some reservations, supported by Tansley (1920).

Gleason (1917, 1926), in his 'individualistic' approach, denied the importance of facilitation between species. He proposed that species combinations were created by the individual responses of species to the environment and by chance events of dispersal following environmental fluctuation and disturbance were more important in creating species combinations. Gleason stressed the continuous nature of variation in species composition, which was consistent with his theory. Both Clements and Gleason recognized that competition had a role in structuring vegetation (Clements, Weaver & Hanson, 1929; Gleason, 1926).

Clements and Gleason formulated their opposing theories on the basis of wide personal field experience and careful observation. However, neither of them appeared to be concerned with collecting and analyzing field data for the specific purpose of testing their respective hypotheses (McIntosh, 1975). In summary:

	habitat	competi-tion	facilita-tion	chance	disturb-ance
Clements	+ + +	+ +	+ +		+
Gleason	+ + +	+		+ +	+ +

On the European side, the two main theoreticians to be discussed are Braun-Blanquet and Du Rietz. The 'phytosociological' approach by Braun-Blanquet (1913, 1921, 1925) can be summarized as follows (see Westhoff & Van der Maarel, 1973; Van der Maarel, 1975):

1. Plant communities are conceived as types of vegetation recognized by their floristic composition, which better express their relationships to one another and to the environment than any other characteristic.

2. Amongst the species that make up the floristic composition of a plant community some are more sensitive indicators of environmental relationships; these diagnostic species include character-species with an occurrence largely confined to one community type, differential-species characterizing a few types towards related other types, and constant companions.

Although Braun-Blanquet (e.g. 1964) started his textbook with a chapter on species interactions he only vaguely and usually implicitly related his discrete community concept to the idea that the

responses of individual species to the habitat are reinforced and modified by strong positive dependences between plant species.

Du Rietz (1921; Du Rietz et al., 1918) is known as the main leader of the Uppsala School with its own methodology of recording and classifying vegetation, and it is appropriate to pay special tribute to him during this Uppsala symposium. What is less remembered today is that Du Rietz was one of the few ecologists of that time who tried to create a link between his empirical methods and results and a coherent causal theory of plant communities. His theory was to a large extent speculative, but in some aspects it is surprisingly modern even today. This seems to be an appropriate time and place to review his ideas in some detail. The main empirical patterns found by Du Rietz are:

1. Some species combinations are found much more frequently than others; these were called associations. Note that this observation was also the basis of Braun-Blanquet's work.*

2. Associations defined by dominants are usually sharply bounded in the field (or separated by narrow transition zones), even where the habitat changes continuously.

3. Sometimes different associations are found in patches of the same habitat.

4. Many species that are not dominant occur constantly in some associations but not in others.

5. When vegetation belonging to an association is sampled by quadrats of a certain size ('minimal area'), then the frequency of species in quadrats has a U-shaped distribution: most species are either very frequent (80−100% 'constants') or very rare (0−20%) ('the law of constants', or 'the law of frequencies', Raunkiaer (1918, 1934).

These findings were the basis of his method of recording and classifying vegetation. Du Rietz (1921) suggested a theory to explain these patterns. It was taken for granted that, in general, the distribution of plant species reflected the relative

suitability of habitats for them. However, Du Rietz postulated that where the habitat tolerances of two potentially dominant species overlap, strong competition between them results in dominance of the one and exclusion of the other (rather than equilibrium). This was to explain the sharp boundaries between associations even where no habitat discontinuity could be observed. Which species won in each location depended not only on the suitability of the habitat for the two species but also 'substantially on the number of propagules or shoots that each species can throw into the struggle'. This in turn, depended on their relative abundances in 'the surrounding vegetation' and on 'the situation at the beginning of the struggle (which species arrived first at the locality)'.

The decisive role of these two factors, the historical and the topographical, could explain according to Du Rietz, why in patches of the same habitat different associations occur and persist for long periods. In effect, Du Rietz, well ahead of his time, was postulating that competition between plant species could result in alternative stable dominance states, depending on priority or initial composition (Gause, 1936) and on a spatial mass effect (Shmida & Ellner, 1984; note that the term vicinism, coined by Nordhagen 1940, covers essentially the same phenomenon). Indeed Gause (1936) quoted Du Rietz' work and was probably influenced by it.

The distribution of non-dominant species, and their constancy in certain associations, Du Rietz explained by tolerance of the competitive conditions created by the dominants. He also speculated that they might contribute to the competitive power of the association by some kind of diffuse competition between 'vital species combinations'. These two factors were to explain why many constants change together at association boundaries.

Both Du Rietz's theory of vegetation and his empirical methodology were severely criticized by his contemporaries (e.g. Nordhagen, 1922, 1928; Braun-Blanquet, 1921, 1925). Moreover, the 'law of constants' (the U-shaped distribution of frequencies), which Du Rietz regarded as proof for the existence of associations in his sense, was shown to be derivable statistically from simple assumptions ('null hypotheses'?) involving random distribution of plants (Romell, 1920; Kylin, 1926). Du Rietz (1928, 1932) defended his ideas for some time but did not succeed in convincing his contemporary critics.

---

* In fact, the two leaders, although rivals in the beginning, reconciliated rather soon after a period of vivid discussions; Braun even dedicated a copy of the 1964 edition of his Pflanzensoziologie to Du Rietz with the words 'Seinem lieben Freund und Aufbaugenossen'. And indeed, companions they became, in the building phase of vegetation science.

The ideas of Braun and Du Rietz may be characterized as follows:

	habitat	competi-tion	facilita-tion	chance	disturb-ance
Braun-Blanquet	+ + +	+	+		+
Du Rietz	+ + +	+ + +		+ +	+

Plant community theory and quantitative plant ecology

In the second quarter of this century methods for describing and classifying vegetation were further developed and applied to studying the vegetation in many parts of the world. But all this work seemed to contribute little to the solution of the dilemma of the contrasting theories about plant communities (Whittaker, 1962). The arguments continued to be mainly scholastic. One reason for this appeared to be that the methods of sampling vegetation and analyzing the vegetation data were not sufficiently quantitative and objective.

Among the pioneers of modern quantitative methods of vegetation analysis around 1950, Whittaker (1952, 1956) and Curtis & McIntosh (1951) were clearly motivated by the unresolved conflict between Clements and Gleason (see Westman & Peet, 1982), while Goodall (1953, 1954) referred to the controversy between Braun-Blanquet and Du Rietz. The overall result from their studies, and many subsequent ones, was that unbiased samples from real vegetation showed largely continuous variation which could be demonstrated by ordination. At the same time some clustering of sites and species in the continuum usually occurred, which could be used for classification. (See Greig-Smith, 1964, for a critical evaluation of this period).

It became clear that 'the pattern of variation shown by the distribution of species among quadrats of the earth's surface chosen at random hovers in a tantalizing manner between the continuous and the discontinuous' (Webb, 1954, see also Williams, Lambert & Lance, 1966). Thus both methods, ordination and classification, were recognized as useful alternatives or rather as com-plementary and were further developed together (Greig-Smith, Austin & Whitmore, 1967; Orlóci, 1967; Van der Maarel, 1969; Whittaker, 1972).

Patterns of co-occurrence of species were used to seek patterns of resemblance between sites, in classification, for instance in association analysis and its derivates (Williams & Lambert, 1959; Lambert & Williams, 1962; Hill, Bunce & Shaw, 1975) and in phytosociological table sorting methods (Van der Maarel, Janssen & Louppen, 1978), as well as in ordination, notably in reciprocal averaging and its derivates (Hill, 1973; Hill & Gauch, 1980).

The concept of nodum was introduced to recognize the occurrence of relative discontinuities in the vegetational continuum (Poore, 1956) and some numerical methods attempted to formalize this concept (Williams & Lambert, 1961; Noy-Meir, 1971; Noy-Meir & Whittaker, 1977). The problem of continuity also became a matter of discussion within classical phytosociology and especially Tüxen (1955), in a reply to Ellenberg (1954), made clear that it are the nuclei ('foci') as 'correlation concentrates' which characterize types, not their peripheries (See Westhoff & Van der Maarel, 1973, for a discussion).

It has often been implied and sometimes been stated explicitly (McIntosh, 1967; Whittaker, 1962, 1967) that the continuous variation usually found in the results of quantitative vegetation analysis has discredited the Clementsian view of vegetation and has supported the Gleasonian view (for a contrary opinion, see Langford & Buell, 1969). Goodall (1963) showed that the relations between the individualistic concept and the unclassifiable continuum (both Gleason's ideas) were not so simple. In a subsequent paper, he suggested a hypothesis of his own (Goodall, 1966), in which positive interactions and evolutionary adaptation of subsidiary species to dominants play an important role. Whittaker (1967) and Goodall (1966) fit in our scheme as follows:

	habitat	competi-tion	facilita-tion	chance	disturb-ance
Whittaker	+ + +	+ +		+	+
Goodall	+ + +	+	+ +	+	

Since the 1950s methods of ordination and classification have been greatly improved (Greig-Smith, 1983) and (as this symposium shows) still are being developed. Their main use has been the detection and description of specific patterns in particular communities and their interpretation by specific hypotheses involving environmental factors, and sometimes also biological interactions and history. In this context of 'hypothesis generation' (Lambert & Dale, 1964) the contribution of the methods to plant ecological field research has been immense.

One might have expected that the development of an objective methodology for collecting and analyzing data on plant communities would also lead to a more rigorous evaluation and to a sharper formulation of general hypotheses about plant communities. This potential use of the new methodology was not explored much further, however. Occasionally, simulated vegetation data generated by theoretical models were analyzed by multivariate methods (Swan, 1970; Austin & Noy-Meir, 1971; Gauch & Whittaker, 1972) but the aim was to test the methods, not the models. A rare exception is Dale's (1977) reference to Goodall's hypothesis in his application of graph theory methods to species associations.

The loss of interest by plant ecologists in generalizations about plant communities is perhaps not accidental. It may reflect the realization that the classical theories were inadequate, and the feeling that the processes in plant communities are probably too intricate and complex to expect any general pattern to be observable at the community level. This intricacy has been brought home by many recent empirical studies on plant population and community dynamics and by new theoretical insights into the processes operating in plant communities.

New theoretical ideas on plant communities

Whittaker & Levin (1977) developed the idea that plant communities can only be understood as a dynamic mosaic of patches. The patches are generated in part by the habitat template but to a large extent also by the spatial aspects of population processes and interactions, in particular dispersion.

Grubb (1977) analyzed the role of the various stages of the population regeneration cycle, and in particular dispersal, germination and establishment, in determining species distribution and community structure. The effects of habitat and other species may often operate not through resource availability to adult plants, but through less obvious modifications of the conditions for regeneration of the species. There has been a rapid accumulation of empirical evidence that species distributions and co-occurrences are often determined largely by the availability of certain kinds of gaps or micro-sites for germination and establishment.

Other mechanisms involved in co-occurrence of plant species have been discussed by Braakhekke (1980) and Shmida & Ellner (1984), in particular environmental fluctuations, patch dynamics and spatial mass effects by which input of seeds from neighbouring habitats can maintain a population in habitats where it would otherwise be non-viable.

Harper (1982) pointed out that since almost all interactions in plant populations and communities are between neighbouring plants, both empirical study and theoretical understanding must begin at this level. The classical work by Watt (1947) on small-scale processes of individual growth and competition, and the resulting patterns within the plant community has been instrumental for the development of these ideas. The power of this approach at least for relatively simply structured vegetation such as grassland, has been amply demonstrated by Harper. A theory of plant communities based directly on neighbour interactions is a challenge yet to be faced. It will mainly be focussed on plant communities with one layer. For complicated structures, notably in forests, the concept of 'neighbour' has to be expanded considerably.

This is the place to mention the development of models for community dynamics based on Watt-type processes at the individual level. Especially models of forest dynamics based on processes of gap formation, recolonization, growth and death of trees (see Shugart, 1984 for a synthesis) have con-

tributed to community theory. An example of a straightforward application of such a model to a simple, one-layer heathland community focussing on spatial interactions between individuals is found in Van Tongeren & Prentice (1986).

There has been a renewed interest and an increasing research effort on the role of interspecific competition as well as other forms of interference in plant communities. This development has probably been influenced by the central role that competition has obtained in the theory of animal communities. It has also benefited from the use of techniques measuring competition between plant species (De Wit, 1960) for evaluation of competition and coexistence at the community level (Braakhekke, 1980). Tilman (1982) made the first attempt to formulate a coherent theory of plant communities structured mainly by competition for various resources, while considering also the effects of disturbance.

The theory of disturbance and succession in plant communities has been considerably refined and different kinds of disturbance and different causes, mechanisms and pathways of vegetation dynamics (both fluctuation and succession) have been distinguished (Connell & Slatyer, 1977; Noble & Slatyer, 1980; Pickett, Collins & Armesto, 1987).

Several hypotheses have been suggested which assign to herbivores an important role in structuring plant communities, but we cannot discuss them here.

Community analysis and hypothesis testing

In view of what we have learned already about plant populations it seems unlikely that any simple general theory of the community will ever explain adequately the variety of phenomena in vegetation. Does this mean that we must despair of any generalizations about plant communities? It may yet be possible to formulate and test hypotheses about plant communities that are general and realistic, if not precise (Levins, 1967). They are also unlikely to be simple. They are likely to have many distinctions and qualifications, many 'if's and 'but's. Important ingredients, in addition to the elements listed at the beginning of this paper, are likely to be the pattern of spatial and temporal variation in the environment, the nature of the potentially dominant species and, in some cases, the role of herbivores.

In the process of testing and developing general hypotheses about vegetation, experimental and dynamical studies of the processes involving plant populations and environmental factors will play a major role. The analysis of patterns of species distribution and co-occurrence, as for instance in Turkington & Harper's (1979) study of a *Trifolium repens* pasture, can contribute useful preliminary evidence and guidance in the context of such studies. The combined analysis of spatial and temporal variation as in series of permanent quadrats should be particularly rewarding. Clearly, the more closely relevant the methods of vegetation analysis are to explicit hypotheses, the more useful they could be.

It has been stated that a main function of the analysis of species-in-sites data is to generate hypotheses (Lambert & Dale, 1964). However, it can also be used to test hypotheses and discriminate between alternative ones.

It seems logical that hypothesis testing should start from simple 'null hypotheses' (Strong, 1980). (Note that Greig-Smith, (1964), made the same point in his treatment of association between species!). Only when such null-hypotheses are rejected need we proceed to more elaborate alternatives. However, Strong and Simberloff's hypothesis that species combinations occur in sites independently and at random, is likely to be of limited value in terrestrial plant communities. This is due to the differential responses of plant species to the inevitable spatial heterogeneity of the environment at almost all scales. Significantly non-random associations or correlations between species, both positive and negative, are found in most samples of vegetation and at many different scales of observation, except in unusually homogeneous environments or in very early stages of colonization.

A suitable next hypothesis in the spirit of Strong and Simberloff, but also in the spirit of Gleason and Whittaker, would be that the observed species

co-occurrences and co-abundances are determined only by the overlap, or segregation in the species distributions over habitats, and that the distributions are all independent of each other. This 'independent habitat response' hypothesis may still require more precise operational formulations, and suitable tests applicable to field data need to be developed. Any significant deviation from this hypothesis would then be evidence that the observed co-ocurrences and co-abundances can not entirely be accounted for by the individualistic responses to habitat, but most likely to interactions between species, be it competition of facilitation. If no significant deviations are found, this does not prove that such interactions do not occur, but only that there is no clear evidence for them in the observed patterns of species distribution and co-occurrence. To disprove the 'no-interactions' hypothesis experimental studies will then be necessary.

Several papers in this symposium show developments in this context. Improved methods for the determination of species responses to environmental gradients (Austin, 1987; Ter Braak & Gremmen, 1987) will certainly facilitate testing of independent habitat response as a null hypothesis. Also, residual variation not explained by the habitat response could then be subject of further study (see also Fresco, 1982). The tests for the randomness of distribution of species modes (Austin, 1987) and lower and upper boundaries (Shipley & Keddy, 1987) are in fact tests of this hypothesis, or at least parts of it.

Once the response of species to habitat has been accounted for, a next step could be to test hypotheses derived from the theory of competition for resources and coexistence by differentiation. One partial test is for the regularity of distribution of species modes along environmental gradients (Austin, 1987). However, the scope for this may prove to be limited when only observations on spatial co-variation at one time are analyzed. The testing of hypotheses relating to competition, as well as those relating to disturbance, will require observations on species co-variation in time as well as in space, both in the natural state of communities and in experimentally manipulated communities. Since such

data will be multivariate, suitable methods of community analysis will be required.

Such studies will link up with new approaches in the understanding of dominance – diversity relations (Whittaker, 1965; McNaughton & Wolf, 1970; Hubbell, 1979; Gray & Mirza, 1979; Braakhekke, 1980; Sugihara, 1980; Tilman, 1982; Shmida & Ellner, 1984).

Community theory and scale of observation

As discussed by Greig-Smith (1964, 1983) species association is scale-dependent. If the quadrat size corresponds to the average size of the plant units (individuals, shoots) involved, negative associations between species will be found by nature of the fact that two plant units cannot usually occupy the same place. Positive associations between species may occur if one other species has larger individuals and the quadrat size used is the same as the size of these large individuals. Absence of species association will arise if we deal with frequent species in quadrats of the same habitat type. (This will be recognized as the 'stopping rule' in the original association analysis of Williams & Lambert, 1959, to arrive at homogeneous clusters). As was demonstrated by Greig-Smith (1964) such an absence of association may turn into a negative correlation if species presence is replaced by some quantitative measure of species performance. Clearly, by changing the size of the unit of observation, or/and by changing the size of the total universe sampled, i.e. its beta-diversity, species co-occurrences will disappear or appear. This points to two interrelated ideas on hierarchical relations by T. H. F. Allen as expressed in Allen & Starr (1982), Allen, 1987) viz. (1) the relation between scale of observation and the hierarchy of communities (ecosystems) at different levels of complexity, and (2) data transformation as a scaling operation.

Evidently, we first have to define the desired level of abstraction and scale of observation most akin to the phenomenon to be studied and then interpret the results of our hypothesis testing within the framework chosen. Some scales may be more rewarding than others as far as community theory

is concerned. One may compare the relatively superficial observations on entire island faunas with the fine-scale plant-to-plant approaches mentioned above.

Clearly, the scale-dependence of species association can also be approached by the analysis of multi-species pattern, an approach that is becoming integrated with community analysis (Noy-Meir & Anderson, 1971; Whittaker, Gilbert & Connell, 1979; Greig-Smith, 1979; Bouxin & Gautier, 1982).

Conclusion

It remains to be seen whether in the future development of vegetation science there will be a closer relation between community analysis and community theory than there has been in the past. We hope to have demonstrated that there are perspectives for such a development. Occasions like the Uppsala symposium can at least facilitate communication between plant ecologists working on those two aspects of our discipline.

References

Allen, T. F. H. & Starr, T. B., 1982. Hierarchy: perspectives for ecological complexity. Univ. Chicago Press, Chicago.

Allen, T. F. H., 1987. Hierarchical complexity in ecology: a noneuclidean conception of the data space. Vegetatio 69: 17−25.

Austin, M. P., 1987. Models for the analysis of species response to environmental gradients. Vegetatio 69: 35−45.

Austin, M. P. & Noy-Meir, I., 1971. The problem of nonlinearity in ordination: experiments with two-gradient models. J. Ecol. 59: 763−773.

Bouxin, G. & Gautier, N., 1982. Pattern analysis in Belgian limestone grasslands. Vegetatio 49: 65−83.

Braakhekke, W. G., 1980. On coexistence: a causal approach to diversity and stability in grassland vegetation. Agric. Res. Rep. 502. Pudoc, Wageningen.

Braun-Blanquet, J., 1913. Die Vegetationsverhältnisse der Schneestufe in den Rätisch-Lepontischen Alpen. Neue Denkschr. Schweiz. Naturforsch. Ges. 48: 1−347.

Braun-Blanquet, J., 1921. Prinzipien einer Systematik der Pflanzengesellschaften auf floristischen Grundlage. Jahrb. St. Gallen Naturwiss. Ges. 57-II: 305−351.

Braun-Blanquet, J., 1925. Zur Wertung der Gesellschaftstreue in der Pflanzensoziologie. Vierteljahrschr. Naturforsch. Ges. Zürich 52: 122−149.

Braun-Blanquet, J., 1964. Pflanzensoziologie. 3. Aufl. Springer, Wien.

Clements, F. E., 1904. The development and structure of vegetation. Bot. Survey of Nebraska, Univ. Nebraska, Lincoln.

Clements, F. E., Weaver, J. E. & Hanson, H. C., 1929. Plant competition. An analysis of community functions. Carnegie Inst. Washington.

Connell, J. H. & Slatyer, R. O., 1977. Mechanisms of succession in natural communities and their role in community stability and organization. Am. Nat. 111: 1119−1144.

Connor, E. F. & Simberloff, D., 1979. Assembly of species communities: chance or competition? Ecology 60: 1132−1140.

Connor, E. F. & Simberloff, D., 1983. Neutral models of species co-occurrence patterns. In: D. R. Strong Jr, D. Simberloff, L. G. Abele & A. B. Thistle (eds), Ecological communities: conceptual issues and the evidence. Princeton Univ. Press Princeton.

Curtis, J. T. & McIntosh, R. P., 1951. An upland forest continuum in the prairie-forest border region of Wisconsin. Ecology 32: 476−496.

Dale, M., 1977. Graph theoretical analysis of the phytosociological structure of plant communities: the theoretical basis. Vegetatio 34: 137−154.

De Wit, C. T., 1960. On competition. Versl. Landbouwk. Onderz. 66(8): 1−82.

Diamond, J., 1975. Assembly of species communities. In: M. Cody & J. Diamond (eds), Ecology and evolution of communities, pp. 342−444. Harvard Univ. Press, Cambridge, Mass.

Diamond, J. & Gilpin, M., 1982. Examination of the 'null' model of Connor and Simberloff for species co-occurrences on islands. Oecologia 52: 64−74.

Du Rietz, G. E., 1928. Kritik an pflanzensoziologischen Kritikern. Bot. Not. Lund 1928: 1−30.

Du Rietz, G. E., 1928. Kritik an planzensoziologschen Kritikern. Bot. Not. Lund 1928: 1−30.

Du Rietz, G. E., 1932. Vegetationsforschung auf soziationsanalytischer Grundlage. Abderhalden Handb. Biol. Arbeitsmeth. 11: 293−480.

Du Rietz, G. E., Fries, T. C. E. & Tengwall, T. A., 1918. Vorschlag zur Nomenklatur der soziologischen Pflanzengeographie. Svensk Bot. Tidskr. 12: 145−170.

Ellenberg, H., 1954. Zur Entwicklung der Vegetationssystematik in Mitteleuropa. Angew. Pflanzensoz. Festschr. Aichinger 1: 134−143.

Fresco, L. F. M., 1982. An analysis of species response curves and of competition from field data: some results from heath vegetation. Vegetatio 48: 175−185.

Gauch Jr, H. G. & Whittaker, R. H., 1972. Coenocline simulation. Ecology 53: 446−451.

Gause, G. F., 1936. The principles of biocoenology. Quart. Rev. Biol. 11: 320−333.

Gleason, H. A., 1917. The structure and development of the plant association. Bull Torrey Bot. Club 44: 463.

Gleason, H. A., 1926. The individualistic concept of the plant association. Bull. Torrey Bot. Club 53: 7−26.

Goodall, D. W., 1953. Objective methods for the classification of vegetation. I. The use of positive interspecific correlation. Austr. J. Bot. 1: 39−63.

14

Goodall, D. W., 1954. Objective methods for the classification of vegetation. III. An essay in the use of factor analysis. Austr. J. Bot. 2: 304–324.

Goodall, D. W., 1963. The continuum and the individualistic association. Vegetatio 11: 297–316.

Goodall, D. W., 1966. The nature of the mixed community. Proc. Ecol. Soc. Austr. 1: 84–96.

Gray, J. S. & Mirza, F. B., 1979. A possible method for the detection of pollution-induced disturbance on marine benthic communities. Mar. Poll. Bull. 10: 142–146.

Greig-Smith, P., 1964. Quantitative plant ecology, 2nd ed. Butterworths, London.

Greig-Smith, P., 1979. Pattern in vegetation. J. Ecol. 67: 755–779.

Greig-Smith, P., 1983. Quantitative plant ecology, 3rd ed. Blackwell, Oxford.

Greig-Smith, P., Austin, M. P. & Whitmore, T. C., 1967. The application of quantitative methods to vegetation survey. I. Association analysis and principal component ordination of rain forest. J. Ecol. 55: 483–503.

Grubb, P. J., 1977. The maintenance of species richness in plant communities: the importance of the regeneration niche. Biol. Rev. 52: 107–145.

Harper, J. L., 1982. After description. In: E. I. Newman (ed.), The plant community as a working mechanism, pp. 11–25. Blackwell, Oxford.

Hill, M. O., 1973. Reciprocal averaging: an eigenvector method of ordination. J. Ecol. 61: 239–249.

Hill, M. O., Bunce, R. G. H. & Shaw, M. W., 1975. Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland. J. Ecol. 63: 597–613.

Hill, M. O. & Gauch Jr, H. G., 1980. Detrended correspondence analysis: an improved ordination technique. Vegetatio 42: 47–58.

Hubbell, S. P., 1979. Tree dispersion, abundance, and diversity in a tropical dry forest. Science 203: 1299–1309.

Kylin, H., 1926. Begriffsbildung und Statistik in der Pflanzensoziologie. Bot. Not. Lund 1926: 81–180.

Lambert, J. M. & Dale, M. B., 1964. The use of statistics in phytosociology. Adv. Ecol. Res. 2: 59–99.

Lambert, J. M. & Williams, W. T., 1962. Multivariate methods in plant ecology. IV. Nodal analysis. J. Ecol. 50: 775–802.

Levins, R., 1968. Evolution in changing environments. Princeton Univ. Press, Princeton.

Langford, A. N. & Buell, M. F., 1969. Integration, identity and stability in the plant community. Adv. Ecol. Res. 6: 83–135.

MacArthur, R. H., 1972. Geographical ecology: patterns in the distribution of species. Harper and Row, New York.

McIntosh, R. P., 1967. The continuum concept of vegetation. Bot. Rev. 33: 130–187.

McIntosh, R. P., 1975. H. A. Gleason – 'individualistic ecologist' 1882–1975: his contributions to ecological theory. Bull. Torrey Bot. Club 102: 253–273.

Noble, I. R. & Slatyer, R. O., 1980. The use of vital attributes to predict successional changes in plant communities subjected to recurrent disturbances. Vegetatio 43: 5–21.

Nordhagen, R., 1922. Vegetationsstudien auf der Insel Utsira im westlichen Norwegen. Bergens Mus. Årb. 1920–21 Naturv. Rekke 1: 1–149.

Nordhagen, R., 1928. Die Vegetation und Flora des Sylenegebiet. I. Die Vegetation. Skr. Norsk. Vidensk. Akad. Mat. Naturv. Kl. 1927(1): 1–612.

Nordhagen, R., 1940. Studien über die maritime Vegetation Norwegens. I. Die Pflanzengesellschaften der Tangwälle. Bergens Mus. Årb. 1939/40 Naturv. Rek. 2.

Noy-Meir, I., 1971. Multivariate analysis of the semi-arid vegetation in southeastern Australia: nodal ordination by component analysis. Proc. Ecol. Soc. Austr. 6: 159–193.

Noy-Meir, I. & Anderson, D. J., 1970. Multiple pattern analysis, or multiscale ordination, towards a vegetation hologram? In: G. P. Patil (ed.), Statistical Ecology 3: 207–225. Penn. State Univ. Press.

Noy-Meir, I. & Whittaker, R. H., 1977. Continuous multivariate methods in community analysis: some problems and developments. Vegetatio 33: 89–98.

Orlóci, L., 1967. An agglomerative method for classification of plant communities. J. Ecol. 55: 193–206.

Pickett, S. T. A., Collins, S. L. & Armesto, J. J., 1987. A hierarchical consideration of causes and mechanisms of succession. Vegetatio 69: 109–114.

Poore, M. E. D., 1956. The use of phytosociological methods in ecological investigations. IV. General discussion of phytosociological problems. J. Ecol. 44: 28–50.

Raunkiaer, C., 1918. Recherches statistiques sur les formations végétales. Kgl. Dansk. Vidensk. Selsk. Bio. Medd. 1: 1–80.

Raunkiaer, C., 1934. The life forms of plants and statistical plant geography. Oxford.

Romell, L. G., 1920. Sur la règle de distribution des fréquences. Svensk Bot. Tidskr. 14: 1–20.

Shmida, A. & Ellner, S., 1984. Coexistence of plant species with similar niches. Vegetatio 58: 29–55.

Shipley, B. & Keddy, P. A., 1987. The individualistic and community- unit concepts as falsifiable hypotheses. Vegetatio 69: 47–55.

Shugart, H. H., 1984. A theory of forest dynamics. Springer, New York.

Simberloff, D., 1978. Using biogeographic data to determine if colonization is stochastic. Am. Nat. 112: 723–726.

Strong Jr, D. R. 1980. Null hypotheses in ecology. Synthese 43: 271–285.

Sugihara, G., 1980. Minimal community structure: an explanation of species abundance patterns. Am. Nat. 116: 770–787.

Swan, J. M. A., 1970. An examination of some ordination problems by use of simulated vegetation data. Ecology 51: 89–102.

Tansley, A. G., 1920. The classification of vegetation and the concept of development. J. Ecol. 8: 118–149.

Ter Braak, C. J. F. & Gremmen, N. J. M., 1987. Ecological amplitudes of plant species and the internal consistency of Ellenberg's indicator values for moisture. Vegetatio 69: 79–87.

Tilman, D., 1982. Resource competition and community structure. Princeton University Press, Princeton.

Trass, H., 1975. Some classification problems in Scandinavian phytocoenology. In: Some aspects of botanical research in the Estonian SSR, pp. 188–224. Academy of Science, Tartu.

Turkington, R. & Harper, J. L., 1979. The growth, distribution and neighbour relationships of *Trifolium repens* in a permanent pasture. I. Ordination, pattern and contact. J. Ecol. 67: 201–218.

Tüxen, R., 1955. Das System der nordwestdeutschen Pflanzengesellschaften. Mitt. Flor.-soz. Arbeitsgem. N.F. 5: 155–176.

Van der Maarel, E., 1969. On the use of ordination methods in phytosociology. Vegetatio 19: 21–46.

Van der Maarel, E., 1975. The Braun-Blanquet approach in perspective. Vegetatio 30: 213–219.

Van der Maarel, E., Janssen, J. G. M. & Louppen, J. M. W., 1978. Tabord, a program for structuring phytosociological tables. Vegetatio 38: 143–156.

Van Tongeren, O. & Prentice, I. C., 1986. A spatial model for vegetation dynamics. Vegetatio 65: 163–173.

Watt, A. S., 1947. Pattern and process in the plant community. J. Ecol. 35: 1–22.

Webb, D. A., 1954. Is the classification of plant communities either possible or desirable? Bot. Tidsskr. 51: 362–370.

Westhoff, V. & Van der Maarel, E., 1973. The Braun-Blanquet approach. In: R. H. Whittaker (ed.), Ordination and classification of communities, pp. 617–726, Junk, The Hague.

Westman, W. E. & Peet, R. K., 1982. Robert H. Whittaker 1920–1980: the man and his work. Vegetatio 48: 97–122.

Whittaker, R. H., 1952. A study of summer foliage insect communities in the Great Smoky Mountains. Ecol. Monogr. 22: 1–44.

Whittaker, R. H., 1956. Vegetation of the Great Smoky Mountains. Ecol. Monogr. 26: 1–80.

Whittaker, R. H., 1962. Classification of natural communities. Bot. Rev. 28: 1–239.

Whittaker, R. H., 1965. Dominance and diversity in land plant communities. Science 147: 250–260.

Whittaker, R. H., 1967. Gradient analysis of vegetation. Biol. Rev. 42: 207–264.

Whittaker, R. H., 1972. Convergences of ordination and classification. In: E. Van der Maarel & R. Tüxen (eds), Grundfragen und Methoden der Pflanzensoziologie, pp. 39–57, Junk, Den Haag.

Whittaker, R. H., Gilbert, L. E. & Connell, J. H., 1979. Analysis of two-phase pattern in a mesquite grassland. J. Ecol. 67: 935–952.

Whittaker, R. H. & Levin, S. A., 1977. The role of mosaic phenomena in natural communities. Theor. Pop. Biol. 12: 117–139.

Williams, W. T. & Lambert, J. M., 1959. Multivariate methods in plant ecology. I. Association analysis in plant communities. J. Ecol. 47: 83–101.

Williams, W. T., Lambert, J. M. & Lance, G. N., 1966. Multivariate analysis in plant ecology. V. Similarity analyses and information analysis. J. Ecol. 54: 427–445.

Hierarchical complexity in ecology: a noneuclidean conception of the data space*

T. F. H. Allen**
Botany Department, University of Wisconsin, Madison, WI 53706, USA

Abstract

Vegetation is the consequence of the interaction of a series of widely differing processes, each uniquely scaled. Extensive slow processes pertain to high levels of organization, while fast local processes pertain to lower levels. Curvature in ordination gradients is often not artefact, but the result of interference between different levels. As straight gradients are lengthened by the inclusion of more heterogeneity in the data, the nature of relationships change between species and their environment and each other at distant places in environmental space. With change in these relationships, movement down the gradient does not always mean the same thing, and this causes curvature. In plotting a noneuclidean space onto a euclidean reference, the change in metrics causes apparent curvature. The technical causes of curvature (bimodality, double zeros, beta diversity) fit this model. Data transformations scale the analyses so that different levels are reflected in results. Between levels, when the processes of the lower level are not local enough to be trivial, the pattern from new upper level processes cannot assert a new straight gradient with coarser grained criteria. Thus transformation and the emergence of curvature followed eventually by new straight gradients allow the linking of different levels in an orderly fashion.

Introduction

Vegetation may be seen as the outcome of a large number of processes operating at different rates over regions of different size. Larger, slower, ubiquitous processes form the context for those which are faster and more local. All of these processes can be significant, although they play very different roles. For example, there is no tree unless there was a microsite with a seed to germinate in it, nor is there a tree unless the landscape takes the form that it does after, say, a glaciation. Thus there are various levels of explanation for vegetation and it is the

job of the vegetation scientist to tease them apart. This paper presents a conception of the data space that leads to a protocol for achieving multilevel description of vegetation with the potential for relating different levels.

Hierarchy theory uses the concept of level as one of its organizing principles. Levels are not asserted to be especially real. Rather, levels are populated by entities of characteristic size behaving within particular time frames. These entities emerge with a given observation protocol involving a particular scale of perception. It is misguided to assert that some levels, usually those populated by tangibles like organisms, are less arbitrary than others. While upper levels are ordered above lower levels, there is no reason to assert that any particular level is the one which is above the reference level. Allen *et al.* (1984) have shown that the levels which emerge above or below are dependent on the criteria that are

used for observation, several levels being possibilities for the next one up or down. When the word level is used here it is in this general sense.

Teasing the levels apart is achieved by looking at the system in different ways through changes in the level at which measurements are taken. The critical changes in observation involve using a different resolving power or a different extent in the universe of observation. Sometimes a change is also made in observational criteria, say from measurements of species abundance to records of species presence; one measures something different. Each observation criterion has at least an implied resolution and universe. Hierarchy theory, which as much as anything is a theory of observation, offers three important principles that facilitate the analysis of community data: (1) Levels of organization are populated by entities whose attendant processes behave at characteristic frequencies, that is they have characteristic periods of time between returns in state. The frequencies order a hierarchy of levels so that large slow entities are found at high levels of organization; (2) Big is not just more, it is different, because there can be change in observational criteria and phenomena across levels. For example, physiology is rarely an adept criterion for considering social systems, even though the members of the population in question do have physiologies; (3) Complexity in ecological systems comes from the interaction of several levels of organization (Allen et al., 1984). Complex patterns appear because there is interference between the frequency characteristics of different levels. For example, species compositions of a complex forest mosaic may be fixed for centuries by the tendency of tree to persist once established, but the various dominants may be selected by the accidental coincidence of a dry afternoon of fire interacting with a seed bed that is the consequence of a mast year. Had the evening rain not come to put out the fire, the landscape boundaries might be very different. Centuries of landscape may arise from the happenstance of a few hours of weather interacting with a cyclical reproductive pattern of a few years.

Only processes operating over a commensurate range and speed can be related simply (Levins, 1974). Very fast local processes are unaffected by very slow large scale processes except in a contextual fashion because, over a short time frame pertinent to the fast processes, the slow processes do not have time enough to proceed. Conversely, over the long time periods associated with the slow processes, the fast processes are generally perceived as fixed at some average level (Simon, 1962). Exceptions to this situation are uncommon and involve amplification of the consequences of a small scale accident such that the upper level is radically changed (supersaturated, or supercooled physical systems or forests with a history of fire suppression are susceptible to the amplification of small accidents). Nevertheless, for most of the time in stable systems, only processes with frequency characteristics of the same order of magnitude can exhibit behavior generally relevant to each others' functioning.

Ordination, transformation and curvature

One of the tools for the task of teasing apart ecological explanations is multivariate reduction of vegetation data, and its role is to simplify the pattern so that sets of simply related processes can be identified and offered as interrelated explanatory principles. The pattern is rate independent but arises from rate dependent processes pressing against structural constraints. The processes belong to a close lower level. The strategy of data reduction is to filter out the effects of all process behavior whose natural frequencies are not of the same order of magnitude as the phenomenon of interest. All other faster processes are for the moment considered as generating noise, while slower processes generate a background hum (noise in the vernacular sense). Transformation changes the dominant frequency characteristics of the data set so as to emphasize different constellations of processes (Allen et al., 1984a). For example, transformation of vegetation cover data to presence data ignores the details of local species dynamics and emphasizes the coarser grained signal of local species extinction and colonization (Allen & Wileyto, 1983). Averaging across cells of a serial plankton data matrix, Allen et al. (1984b) removed daily fluctuation from the ordination pattern and allowed the emergence of seasonal regions of stability and eventually annual cycles. Conversely, differencing a plankton matrix allowed Allen et al. (1977) to elicit local signals of week to week environmental control of the plankton from a data set that had until that

analysis only revealed patterns interpretable as general seasonal control. Allen *et al.* (1984b) moved the observer up the hierarchy of processes, while Allen *et al.* (1977) moved him down a level. Followed by data reduction, transformation allows access to levels one at a time, so achieving a simplification and offering explanation. Comparison of differently scaled analyses of the same equivalent field collections allows a clear picture of both the level to be explained and the level to which the explanatory underlying processes belong.

More coarse grained transformations (e.g. expressing datum values as binary presence/absence as opposed to quantitative records) involve less resolving power, for they lump together as one state, conditions which are recognized as different in more fine grained expressions of the field data. Coarse grained transformation of extensive data sets reveal upper levels that reflect slow extensive processes (Allen *et al.*, 1984b), while more finely grained transformations of smaller sets reveal lower levels governed by high frequency processes of local influence in time and space (Allen *et al.*, 1977; Allen & Wileyto, 1983).

It is worthwhile to view the differences between levels and their attendant frequencies in geometric terms. Small scale processes that are homogeneous in a local circumstance may, at some distant point in the global space, change or disappear. This change can be considered as a change across the global space in the criteria for significant ecological difference. This translates easily to a noneuclidean model of the vegetation space. Noneuclidean spaces are locally flat enough to be reasonably approximated by a plane, but are globally curved as the local departures from a plane accumulate. Relative to the local reference space, the metric changes across the larger space. In commonplace ecological experience this refers to the way that in some ecological contexts a small change matters, while in other circumstances quite a large change makes no difference. For example, plants read a small change in water availability in the same terms throughout the entire change. To be completely accurate the first half of the small water loss is different from the second half, but not so that it is worth considering, for it may not be measurable, and anyway there is always some point at which one must set something as an invariant frame. However, over a large change in water regime, local differences in the sig-

nificance of water availability have accumulated so that the difference matters. In a mesic situation a lowering of water one unit makes little difference, but in a drought a similar change can make all the difference between life and death. Consider fires with much as opposed to little fuel as distant points in a noneuclidean space. Not only will the fires be of different sizes, but the smaller fire may be of greater consequence. A big fire in chaparral is a normal process of recycling in a constant fire cycle mosaic, but a second small fire coming too soon can devastate the regeneration process. In ecological spaces the measures of difference themselves change with position in the space.

As indicated above, different data transformations use different grains, and the size and heterogeneity of the data matrix determine the extent of the analysis. Change in grain causes a change in the level perceived just as does the extent of the study. It is easier to see how the noneuclidean model above applies to changes in extent, but since both grain and extent determine the level perceived, the model also applies to changes in grain. A larger extent can include new ecological circumstances in which formerly constant criteria become variable.

In that context, consider one of the cases included in Austin's (1980) 'Searching for a model for use in vegetation analysis'. The model distinguishes between a species grown in monoculture exhibiting its 'physiological response' to a gradient, as opposed to its 'ecological response' under natural conditions when competition suppresses the optimal response (Ellenberg, 1953; Mueller-Dombois & Ellenberg, 1974) (Fig. 3). In monoculture the universe of discourse is local and represents a low level of organization, while natural conditions that involve many species at a site include are more inclusive and so represent a higher level that requires a wider discourse. For species to have more or less equivalent competitive ability, they would all have to retain the form of the response curve they each have in monoculture, even when growing in mixed stands. The only difference would be a lowering of density and an attendant loss of biomass according to a dilution factor proportional to the number of species in the mixed stand. In the unlikely event of all species having the same competitive ability, vegetation, the upper level of organization, would be simply the sum of physiological response curves. Different environments would be the context in which com-

petition would have a specific outcome; change the context by moving down the gradient and the outcome of competition changes accordingly. However, competition is asymmetric in favor of winners, reward for superior competitive ability often involving the virtual exclusion of the loser. As a result, poor competitors may be completely excluded from what are their physiologically optimum environments. At environmental extremes, such losers of competition will be away from their physiological optimum but may be able to reach the vigor of their physiological response at that point on the gradient. These situations may represent the mode or modes of their ecological response curves. At environmental extremes, the species have a new type of relationship to each other and a closer relationship to the limits imposed by the physical environment.

In the extensive ecological space discussed above, the differences in species relationships arise not from a new outcome to competition in a new place but from different organism tolerances that remove some of the would be competitors. Not just the interactions between species change, but the criteria for interaction change in the space that includes environmental extremes. The competition mediated selection of species in the middle of the space gives way to a more direct physical environmental selection of species at the periphery. This does not imply that there is no competition at environmental extremes, but that it may be in different terms than the competition in the middle range. One might expect interference to be the dominant process in the middle range, while competition for limited resources might be the principal interaction at the periphery. This change in explanatory criteria can be translated easily into the noneuclidean geometric model for vegetation space. In a euclidean model, the reference frame for the interaction between species and between species and environment is constant because the metric is constant across the entire space. In a noneuclidean model, the metric (the criterion for distance) changes because the criteria for species-species and species-environment themselves relations are different in various parts of the ecological space. When spaces with different criteria for distance are projected onto each other, curvature appears. A noneuclidean space is not curved by its own standards; note that the environment space is ordered by constantly scaled environ-

mental gradients, it is just that environmental difference means different things for the explanation of vegetation in different regions. Curvature only appears when the environmental space is projected onto a euclidean vegetation space.

If curvature occurs in a gradient on an ordination, sometimes it is artifact and distortion with no biological meaning, but often it is the interaction between two biologically valid criteria, one more local than the other. Curvature can come from the interaction of different levels of organization. Since explanation is the expression of a level in terms of interactions at one lower level, the localization of the analysis to just one level is helpful. Curvature, however, shows the interference of explanation at one level by that at another. This creates problems for explanation by simple reduction to a single lower level but, nevertheless, identifies a valid meta-explanation addressing the issue of complexity in the data. Let us consider the reasons more usually offered for curvature and see if they are not perhaps special cases of the noneuclidean model offered above.

Sources of curvature and the noneuclidean model

Unimodal species curves do cause gradient curvature because of like scores on opposite sides of the mode (Beals, 1973). Only the species mode has an unambiguous relationship to environmental gradients. All other values for a species, performance share the same value with some other distant part of the space and therefore tend to pull those distant points towards each other, so causing curvature. Despite this well known curvature arising in unimodal situations, bimodal species curves are more often cited as a cause because their curvature is commonly more severe and the equivalence of distant modes is obvious. Bimodality is generally considered uncommon and something that almost completely disappears if one has the 'right' axis orientation. However, the Curtis (1959) data that underlie 'The Vegetation of Wisconsin' show over 90 unequivocally bimodal species. If such a significant data set as that shows bimodal species with such regularity, bimodal curvature cannot be dismissed as unusual pathology (Fig. 1).

The problem of curvature is not given adequate account if it is dismissed as mere distortion. Curva-
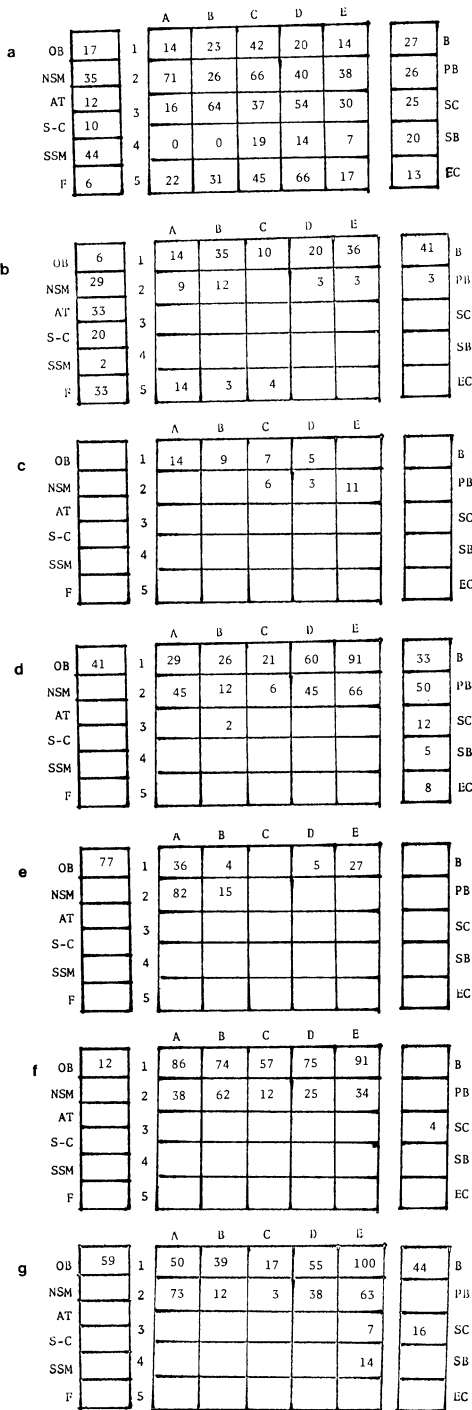
a

		A	B	C	D	E		
OB	17	14	23	42	20	14	27	B
NSM	35	71	26	66	40	38	26	PB
AT	12	16	64	37	54	30	25	SC
S-C	10	0	0	19	14	7	20	SB
SSM	44	22	31	45	66	17	13	EC
F	6							

b

		A	B	C	D	E		
OB	6	14	35	10	20	36	41	B
NSM	29	9	12		3	3	3	PB
AT	33							SC
S-C	20							SB
SSM	2							EC
F	33	14	3	4				

c

	A	B	C	D	E	
OB	14	9	7	5		B
NSM			6	3	11	PB
AT						SC
S-C						SB
SSM						EC
F						

d

		A	B	C	D	E		
OB	41	29	26	21	60	91	33	B
NSM		45	12	6	45	66	50	PB
AT				2			12	SC
S-C							5	SB
SSM							8	EC
F								

e

		A	B	C	D	E	
OB	77	36	4		5	27	B
NSM		82	15				PB
AT							SC
S-C							SB
SSM							EC
F							

f

		A	B	C	D	E		
OB	12	86	74	57	75	91		B
NSM		38	62	12	25	34		PB
AT							4	SC
S-C								SB
SSM								EC
F								

g

		A	B	C	D	E		
OB	59	50	39	17	55	100	44	B
NSM		73	12	3	38	63		PB
AT						7	16	SC
S-C								SB
SSM						14		SB
F								EC

Fig. 1. A reproduction of Curtis' unpublished summary data for selection of 90 species showing bimodality. The central grid represents community-types from southern and warm environments at the bottom to cool and northern at the top. Wet is left and dry is right. Row 1 = Boreal Forest, 2 = Northern Hardwood, 3 = Southern Hardwood, 4 = Savanna, 5 = Prairie. The extreme columns are occupied by rare and peculiar communities which are extremely wet to the left and extremely dry to the right. They are similarly arranged warm to cool. OB = Open Bog, NSM = Northern Sedge Meadow, AT = Alder Thickett, S-C = Shrub-Carr, SSM = Southern Sedge Meadow, F = Fern; A = wet, B = wet mesic, C = mesic, D = dry mesic, E = dry. B = Bracken, PB = Pine Barrens, SC = Shaded Cliff, SB = Sand Barrens, EC = Exposed Cliff. Some special communities mentioned in Curtis' appendices occur in the central grid. 1A = cedar swamp, 4C = oak opening, 4D = oak barrens, 4E = cedar glades. There are no communities corresponding to 4A and B. The text of Curtis (1959) distinguishes between wet mesic and mesic oak openings, but the tabular summaries aggregate these two types. The members of stands in the boreal forest are from Maycock (1957). (a) Total numbers of stands included in each community type. In cases where bimodality may be an artifact of small or uneven numbers, the species concerned were excluded. (b)–(g) Distribution of six species with striking bimodality. Numbers in the grids represent percentages of the total number of stands of a community type in which the species occurs. (b) *Aster umbellatus* (c) *Corallorhiza striata* (d) *Gaultheria procumbens* (e) *Ledum groenlandicum* (f) *Linnaea borealis* (g) *Vaccinium myrtilloides*.

ture can take many forms, and in extreme cases the entire ordination can be circular. A case in point is the circular ordination in Fig. 2, a hitherto unpublished ordination of the vegetation of the Wasatch Mountains performed by Grant Cottam. If the curvature is only artifact, then there arises the problem of how come the ordering makes sense to the field ecologist all the way round (G. Cottam, pers. comm.). One might expect most of the sequence on the circle to make sense even if the curvature were artifact, but the critical point in Cottam's ordination is that there is not a beginning that mismatches with the other end. The whole circle fits.

Allen & Starr (1982) have suggested that a well known bimodality in Black Spruce (*Picea mariana*) is the interference of an upper level of organization into the processes of a straight gradient over the mesic range. This is a special case of the difference between 'physiological response' as opposed to 'ecological response' (Ellenberg, 1953; Mueller-Dombois & Ellenberg, 1974). With paired ordinations in both species and environmental space,
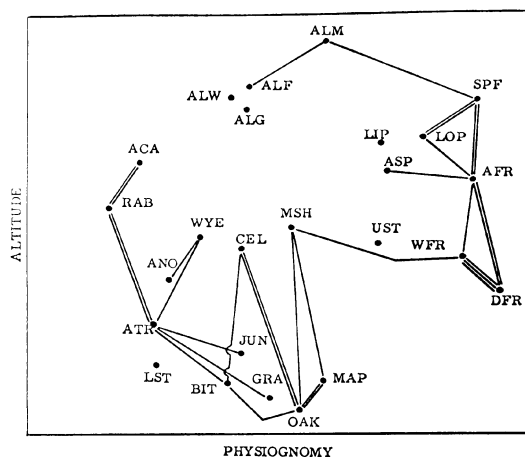
Fig. 2. A Bray & Curtis (1957) ordination in vegetation space of stands in the Wasatch mountains from an unpublished manuscript by Cottam, Ream & Cottam (nd). Working clockwise around the circle, the communities are Alpine wet meadow ALW, Alpine grassland ALG, Alpine Fell-field ALF, Upland herb meadows ALM, Spruce fir SPF, Lodgepole pine LOP, Alpine fir, Aspen ASP, Douglas fir DFR, White fir WFR, Upper streamside UST, Mixed shrub MSH, Maple brush MAP, Oak brush OAK, *Cercocarpus ledifolius* CEL, Juniper JUN, Grassland GRA, *Cercocarpus-Purshia* BIT, Lower streamside LST, *Artemisia tridentata* ATR, *Wyethia* WYE, *Artemisia nova* ANO, Rabbit brush RAB, *Artemisia cana* ACA. Community relationships are identified by the index of similarity (Curtis, 1959) three lines 50, two lines 40, one line 30. The circle can be closed by merely lowering similarity criteria to 20 whereupon ACA joins ALF. A 45° rotation was required to produce the axes presented.

Loucks (1962) showed the curvature of environmental gradients superimposed on the ordination of species space. That curvature was related to bimodality for Black Spruce (*P. mariana*) when it was in turn superimposed on the environmental ordination. Over the mesic part of the vegetation in Northern Wisconsin, there is much interference between plants with respect to light. The physical environment mediates the outcome at a given site by favoring one species over another during the initial stages of resource capture before interference fixes the stand composition. However, at both dry and wet environmental extremes, water and nutrients are respectively limiting resources. The role of light interference is diminished by the physical environment. This makes a fundamental change in the context of species interaction. Black Spruce (*P. mari-*

ana) is only able to survive in the ameliorated light competition regimes at the ends of the gradient. Curtis (1959) notes 'regardless of substrate, successful establishment of spruce seedlings requires abundant light'. The level of interference with regard to light is high, and constantly so, across the local mesic universe, but it is variable across the larger universe that includes various densities of canopy. There is a basic difference between above and below ground competition in terms of symmetry (J. Weiner, pers. comm.).Competition for light is asymmetric (shorter plants do not shade taller plants at all) while competition for nutrients and water is the division of a limited resource, where the resources captured by the less successful party are unavailable to the winner. The notion of a noneuclidean space is helpful here because the measure of difference itself changes across the global space as the very architecture of the competitive forces takes a new form.

High beta diversity is also considered a cause of curvature, although it must be noted that high beta diversity does not invariably lead to curvature (Van der Maarel, 1979). Here again the same notion of changes in units of measurement is appropriate. For a certain species, one end of the gradient is beyond its tolerance. It could occur in the mesic range but is restricted to one end of the gradient. This localization increases beta diversity and contributes to curvature. Note how this situation has much in common with bimodality caused by competition suppressing some species in the mesic range. Beta diversity may be raised by competition truncating the range of some species. In the case where only beta diversity is raised and bimodality does not occur, the truncated species is of too narrow a tolerance to manifest a second mode on the far side of the fierce competition; the difference from the bimodal case is trivial. Given this similarity to bimodality, the same noneuclidean model pertains (Fig. 3). Thus saturation of a distance measure, one of the effects of high beta diversity, can have ecological significance. The 'double zero problem' which gives opposite ends of gradients similarity, relates to bimodality as its simple obverse. The absence of the mesic species, for perfectly good ecological reasons, from extreme environments is exactly the reason for convergence in the competition regime at opposite ends of the gradient. There is much ecological significance to be found in curva-
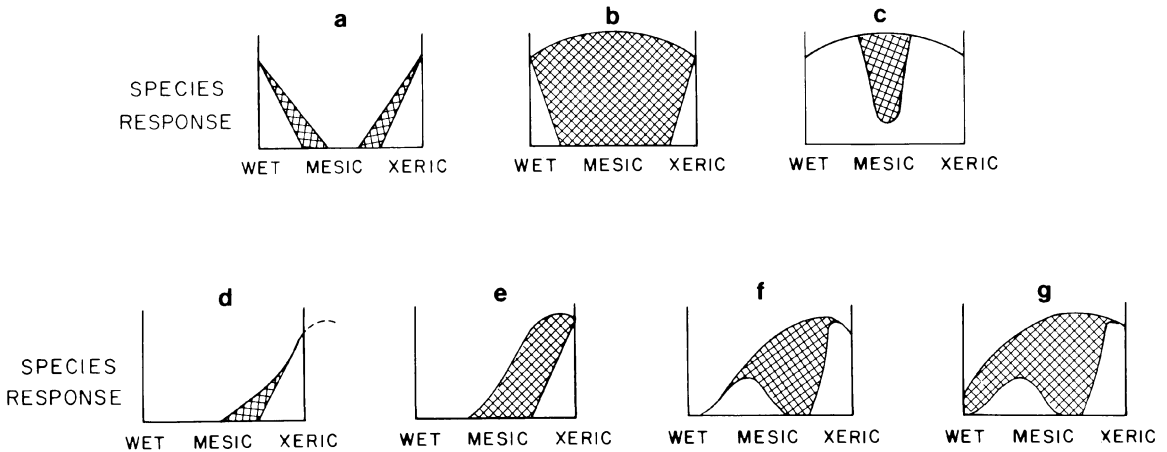
Fig. 3. Examples of various lengths of the mesic range applied to different species response curves along a moisture gradient. Shaded regions indicate the difference between the response curve of the fundamental niche and the realizable niche of suppression. (a) Ecotypic differentiation and specialization. This curve could also be the result of a common status on a secondary gradient. (b) Same realizable niche as 'a' but from wide ranging suppression of a broad modal fundamental response. (c) Same fundamental response more narrowly suppressed giving two modes abutting the mesic condition. (d) Minimal suppression of a xeric specialist. (e) Some unimodal response as 'd' but from wide ranging suppression of a potentially broad tolerence species. (f) Same fundamental response curve as 'e' realizes a skewed bimodal curve under narrowly mesic suppression. (g) Same bimodal response curve as realized in 'f' but by broad suppression of a potentially wide ranging species with a fundamental mode only slightly on the xeric side of mesic.

ture if the noneuclidean model is used.

There are some explicitly noneuclidean methods for applying curvature, albeit constant, in the psychometric literature, but the means whereby we might apply them to ecological situations is less than transparent (Lindman & Caelli, 1978). Ecological gradient straightening techniques such as detrended correspondence analysis (Hill & Gauch, 1980) or step across (Williamson, 1978) imply a correction of a noneuclidean space. Both techniques explicitly normalize around local parts of the ordination space and apply a correction to each part of the space. Those ordinations essentially change the implied metric so that the whole space appears flat when projected onto a euclidean space. However, each part of that euclidean space has its own special relationship to the global vegetation space. This is not necessarily bad, but it does obscure the differences between the local spaces. Sometimes the ecologist may be more interested in the changes in organizing principles rather than in maintaining a standard relationship of vegetation space to simple physical gradients. The change in the importance of water may be more significant for some questions that mere changes in the water re-

gime per se. Loucks' (1962) environmental scalars include a nonlinear reference frame and begin to address these questions for specific situations for water, nutrients and heat. By introducing changes in criteria as a general principle, the noneuclidean conception of the data space might offer a more general model that applies to many different ordination procedures.

Conclusion

As we apply successive data transformations, different biological and ecological principles dominate the data structure. Not only should ecologists be interested in all these principles, but they should try to work out the relationships between them. Ordination curvature is a tool for that job. Data transformations and changes in the size of the data matrix change the scale of the dominant processes that are reflected in the data structure. When this happens what is a straight gradient at one scale becomes curved as the implicit size of the global space increases with use of more coarse grained transformations or as the explicit size of the global

space increases with inclusion of more stands and species. Curvature is often a reflection of a change in the criteria that apply in the local spaces. As the vegetation space implied by the transformation becomes larger, eventually the original set of processes become so local as to be irrelevant. At that point the original gradient based on local considerations becomes lost in the curvature, and a new gradient, reflecting coarser grained processes, replaces it. The lost gradient might be one from an ecology explained by the quantitative outcome of competition, and the one that replaces it might be a gradient where tolerance to the physical environment takes over as the dominant organizing principle.

The image of a hierarchy of processes is helpful here. All the way from an acorn germinating in a bear's footprint right up to the glaciation that moved the mountain on which the bear lived, there are explanations for an oak tree at a site. Only adjacently scaled processes associated with those explanations can be related. Most of the intervening processes between footprint and glaciation are independent of each other: historical accident occurs below, while the fixed context sits above (McCune & Allen, 1985a, b). As the scale of perception is changed through data transformation, the ecologist's eye moves up or down through the hierarchy. The movement is not from one discrete level of explanation to the next but is continuous. Constellations of related processes gradually move through a fish-eye lens; at one level of perception they are reflected as a straight gradient only to be lost in curvature and replaced by another set of explanatory processes and their gradient. The reductionist program searches for the level which is the true explanation, dismissing other levels as mere epiphenomena (Levins & Lewontin, 1980): that program is implicit in the dismissal of curved gradients as inferior explanations. That is not to say that ordinations with straight gradients have been a waste of time. Much of what we know about community structure and dynamics has come from such analyses. Nevertheless, an insistence on straight gradients limits the ecologist's perceptions and isolates his explanations from their context of higher levels and their mechanism from lower levels. The context emerges when grain is coarsened and extent is widened, although the change introduces curvature in the formerly straight gradient. Mechanisms emerge expressed on new curved gradients when resolving power is increased and extent is narrowed.

By formalizing scale considerations we can unravel complexity. Changes in criteria for assessment of ecological explanation become expected and stop taking us by surprise. Curvature of spaces, until now seen as a problem, becomes a friend. We can expect to achieve some translation between different modelling considerations that have heretofore been seen as largely unrelated. As changing grain and extent become a regular part of data analysis, models that imply multiple stable states may be seen as a differentiation of the environment that is left structurally undefined in models of simple local stability (Holling, 1973). Models which are seen as equilibrium based will be seen easily as the description of the context of things that yield to non-equilibrium models (G. Harris, pers. comm.). When linked to larger scaled processes, disturbance can be incorporated as the onset of regeneration in a larger system (C. S. Holling, pers. comm.).

References

Allen, T. F. H., Bartell, S. M. & Koonce, J. F., 1977. Multiple stable configurations in ordination of phytoplankton community change rates. Ecology 58: 1076–1084.

Allen, T. F. H., O'Neill, R. V. & Hoekstra, T. W., 1984a. Interlevel relations in ecological research and management: Some working principles from hierarchy theory. USDA Forest Service Gen. Techn. Rep. RM-110, Fort Collins.

Allen, T. F. H., Sadowsky, D. A. & Woodhead, N., 1984b. Data transformation as a scaling operation in ordination of plankton. Vegetatio 56: 147–160.

Allen, T. F. H. & Starr, T. B., 1982. Hierarchy: Perspectives for ecological complexity. University of Chicago Press, Chicago.

Allen, T. F. H. & Wileyto, E. P., 1983. A hierarchical model for the complexity of plant communities. J. Theor. Biol. 101: 529–540.

Austin, M. P., 1980. Searching for a model for use in vegetation analysis. Vegetatio 42: 11–21.

Beals, E. W., 1973. Ordination: Mathematical elegance and ecological naiveté. J. Ecol. 61: 23–36.

Bray, J. R. & Curtis, J. T., 1957. An ordination of the upland forest communities of Southern Wisconsin. Ecol. Monogr. 27: 325–349.

Curtis, J. T., 1959. The vegetation of Wisconsin. Univ. of Wisconsin Press, Madison.

Ellenberg, H., 1953. Physiologisches und ökologisches Verhalten derselben Pflanzenarten. Ber. Deutsch. Bot. Ges. 65: 351–362.

Hill, M. O. & Gauch Jr., H. G., 1980. Detrended correspondence analysis: an improved ordination technique. Vegetatio 42: 47–58.

Holling, C. S., 1973. Resilience and stability of ecological systems. Ann. Rev. Ecol. Syst. 4: 1–24.

Levins, R., 1974. The qualitative analysis of partially specified systems. Ann. New York Acad. Sci. 231: 123–138.

Levins, R. & Lewontin, R., 1980. Dialectics and reductionism in ecology. Synthese 43: 47–78.

Lindman, H. & Caelli, T., 1978. Constant curvative Riemannian scaling. J. Math. Psychol. 17(2): 89–109.

Loucks, O. L., 1962. Ordinating forest communities by means of environmental scalars and phytosociological indices. Ecol. Monogr. 32: 137–166.

Maycock, P. F., 1957. The phytosociology of boreal conifer-hardwood forests of the great lakes region. Ph.D. thesis University of Wisconsin, Madison.

McCune, B. & Allen, T. F. H., 1984. Will similar forests develop on similar sites? Can. J. Bot. 63: 367–376.

McCune, B. & Allen, T. F. H., 1984. Forest dynamics in the Bitterroot Canyons, Montana. Can. J. Bot. 63: 377–383.

Mueller-Dombois, D. & Ellenberg, H., 1974. Aims and methods of vegetation ecology. Wiley, New York.

Simon, H. A., 1962. The architecture of complexity. Proc. Am. Phil. Soc. 106: 467–482.

Van der Maarel, E., 1979. Transformation of cover-abundance values in phytosociology and its effects on community similarity. Vegetatio 39: 97–114.

Williamson, M. H., 1978. The ordination of incidence data. J. Ecol. 66: 911–920.

A dynamical systems perspective on vegetation theory

David W. Roberts*
Departments of Botany and Forestry, University of Wisconsin, Madison, USA
Current address: Department of Forest Resources and Ecology Center, Utah State University, Logan UT 84322-5215, USA

Abstract

A dynamical systems perspective is employed to develop a simple conceptual model of vegetation and environment as coupled dynamical systems. The conceptual model characterizes the influence of environment on vegetation, the effect of vegetation on the environment, and the subsequent response of vegetation to the modified environment. Vegetation and environmental dynamics are modeled as trajectories in complementary state spaces, with the trajectories jointly determined by the position of a given site in both spaces. The vegetation and environment state spaces are coupled by the physiological requirements of the component species and the modification of environment by vegetation. From a dynamical systems perspective, current vegetation theory and analyses overemphasize environmental determination of vegetation composition and neglect the effects of vegetation on environment. A dynamical systems perspective is capable of synthesizing previous concepts of vegetation; the continuum and community type concepts are possible consequences of site specific differences in vegetation metabolism and environmental plasticity.

Introduction

Austin (1980) argued that the critical factor delaying the development of better analyses of vegetation was the lack of a well-developed theory or model of vegetation. Austin (1980, 1985) and Gauch (1982) suggest that current theory is based primarily on the direct gradient analyses of Whittaker (1967, 1978), and artificial data sets derived from these analyses. Direct gradient analysis itself, however, is a data analysis based on Whittaker's climax pattern hypothesis (Whittaker, 1953, 1974), which is derived from Gleason's individualistic concept (Gleason, 1926, 1939), with additional emphasis on competition.

In this paper I develop a simple conceptual model of vegetation and environment which I propose as an alternative basis for the design and interpretation of vegetation analyses. The model employs a dynamical systems perspective, emphasizing the mechanism which underlies vegetation dynamics, and the consequences of this mechanism on the relation of vegetation composition and environment. The dynamical systems perspective owes a strong debt to earlier vegetation theories, and I view it as an evolutionary development. Accordingly, I develop the concepts of the model historically, and then provide some detail about the structure and behavior of the model. Finally, I will contrast a dynamical systems perspective of vegetation with other vegetation concepts, and emphasize the implications of this perspective for vegetation analysis.

Historical elements of vegetation models

Clements (1916) strongly emphasized the dynamic nature of vegetation, demonstrating that the interpretation of the distribution of vegetation in space required consideration of vegetation processes in time. Additionally, Clements described the capacity of vegetation to modify the environment, and termed this process 'reaction.' Clements was criticized for overestimating the role of reaction in vegetation development, but the concurrent consideration of time and space, and an appreciation for the capacity of vegetation to modify the environment are critical to the development of a dynamical systems-based vegetation model.

Gleason (1926, 1939) proposed that the presence of a plant species in a community requires the successful migration of propagules to the site, and that all site factors fall within the ecological amplitude of the species, which differs for each species. Gleason considered the other species in the community to be part of the environment when considering a species' amplitude. Modern vegetation ecology owes to Gleason the concepts that the initial composition and subsequent invasion of a community are stochastic processes, and the explicit consideration of the individual ecological amplitude of each species.

Cooper (1926) argued that vegetation dynamics is universal and occurs on broadly varying time scales; the long-term stability envisioned by Clements is an illusion. He also stated: 'The organisms produce vegetation change through their effects on the environment and *upon each other through the environment*' (Cooper, 1926, my emphasis). With few exceptions, such as true symbiosis, root grafts, and some forms of allelopathy, every plant-to-plant interaction is mediated by the environment.

According to Tansley (1935) vegetation cannot be considered apart from the environment or animal communities as the fundamental entity is the 'ecosystem, ... a particular category among the physical systems that make up the universe'. Accordingly, Tansley argued that succession is a continuous, sequential change driven by an internal mechanism of the system and he combined Cooper's (1926) observation that plants act through the environment with the concept of reaction.

This dynamical systems concept of Tansley was developed further primarily by Russian vegetation scientists who employed a similar concept of 'biogeocoenosis,' and who were empathetic with a systems point of view.

Khil'mi (1962) clarified a number of concepts that Tansley left unspecified. He argued that the modification of the environment by plants is in part passive, resulting from their presence as physical bodies. A greater effect is due to the metabolic processes of plants, affecting the temperature and humidity of the air, and the physical and chemical properties of the soil. Plants are limited by their metabolism to a certain range of environments. However, once established they modify the environment to such an extent that they must be considered to largely create their own environment.

Khil'mi (1962) reasoned that not all types of vegetation are equally effective at altering their environment, but rather the capacity to modify the environment is roughly proportional to vegetation biomass and the rate of metabolic activity of the component species. Consequently, the transformatory capacity of forests is especially great. While each species modifies the environment in a manner specific to its metabolism, in practice it is not possible to isolate the effect of each species. Rather, we observe the net, or synthesized action of the vegetation taken as a whole. The basic processes of vegetation are biophysical phenomena. Consequently their study demands a combined biological and physical approach.

Khil'mi's work was followed by Sukachev & Dylis (1964), who argued that within a vegetation type species vary markedly in their capacity to modify the environment. The capacity is roughly proportional to biomass, stature, and rate of metabolic activity, so that most of the influence is exerted by the canopy dominants. Because the vegetation both responds to and strongly modifies the immediate environment, the ecosystem (or biogeocoenosis) is a self-modifying system, and can be modeled as such.

A synthesis of the views of the referent ecologists is that the distribution of vegetation in time and space must be considered concurrently. The composition and subsequent development of vegetation is partly stochastic and partly determined by the self-modifying properties of a dynamical system. Plants interact through the environment, and the net or synthesized interactions determine the properties of the system.

A dynamical systems vegetation model

Goodall (1963) laid the foundation for a dynamical systems model of vegetation by characterizing two complementary *n*-dimensional spaces. The first space is a vegetation composition space, where each species defines an axis, and each axis is scaled according to some measure of the species' abundance. The second space is an environmental space where every axis corresponds to some physical site factor, and where each axis is scaled according to the appropriate measurement for that factor. The ecological amplitude of each species occupies a sub-volume of environment space where all site factors are within the metabolic requirements of the species. The points within this ecological amplitude have varying favorability to the species, generally with an optimum somewhere in the interior of the space and less favorable conditions near the perimeter of this space.

As suggested by Gleason (1926, 1939), the exact shape and position of this subspace is different for each species, and depends on how we measure the environment. If we measure the environment in a gross sense, so that all species in a community experience the same environment, then the shape of each species' subspace will depend in part on which other species occur in the community, corresponding to the species' synecological amplitude. If we measure the environment as each plant experiences it, the subspaces will remain constant (corresponding to the species autecological or physiological amplitude), but each species will experience a somewhat different environment. These concepts are treated in more detail by Mueller-Dombois & Ellenberg (1974), and Wuenscher (1969, 1974).

A site in a forest which has recently experienced a catastrophic forest fire will have a position in environment space reflecting the current environmental conditions, and a position in vegetation space corresponding to the abundance of those few plant species which survived the fire or migrated onto the site after the fire. The initial vegetation which develops on this site will be determined partly by the chance immigration of propagules, and partly by on-site buried seeds, rhizomes, or sprouts. As the vegetation develops, the site's position in vegetation space will move, following a trajectory determined by the increase in biomass and increasing numbers or changing proportions of species. Simultaneously, the position of the site in environment space will move also, following a trajectory reflecting decreased light availability, changes in soil moisture, and other site changes induced by the vegetation. The trajectories of the site in the two spaces are coupled by the physiological requirements of the species and their effect on the site.

At first the trajectory in vegetation space will be controlled largely by the different growth rates and life spans of the pioneer species, reflecting the life history characteristics of these species. Soon, however, the site's position in environment space will migrate into, or out of, the ecological amplitudes or subspaces of some of the available species. Site favorability will increase for some species, and decrease for others, as a result of vegetation-induced changes in environment. For many of the established species, environmental conditions will remain within the subspace in which they can persist, but will move out of the smaller subspace within which they can reproduce, limiting such species to one life span on the particular site.

All species affect environmental conditions and each other through the environment. Conceptually, each species exerts a vector (a directed force) on the position of the site in environment space, with the force of the vector proportional to the product of the species, biomass and metabolic activity rate. The orientation of the vector is determined by the particular effect of the species on each environmental factor. The net or synthesized effect of the whole plant community is simply the vector sum of all species vectors.

The capacity of vegetation to modify the environment is limited. Under the influence of vegetation, the position of the stand in environment space will only move a limited distance from its original position, and this distance will depend on which environmental factors are being modified. Additionally, some environmental factors exhibit an elastic response and others exhibit a plastic response. Light, for example, is perfectly elastic. It is possible for vegetation to reduce available light at the ground to only a few percent of full sunlight; as soon as the vegetation is removed, however, the available light returns to its original value immediately. Alternatively, soil pH is a plastic site factor. Under a canopy of spruce trees (*Picea* spp.) soil pH may slowly but steadily decline until reaching a pH in equilibrium with the spruce trees and other fac-

tors. If the spruce trees are removed, the soil pH will not immediately return to its original value. Rather, the current value will be relatively unchanged, and future values of soil pH will depend in part on the vegetation established after the removal of the spruce trees.

Conceptually, both environmental elasticity and plasticity are modeled as vectors acting on the position of the site in environment space. Elastic vectors act to return a site to its original position in environment space, and thus act to resist movement away from the original position. The force of elastic vectors is roughly proportional to the distance of the site from the original position in environment space. Thus, they exert little resistance to small changes and increasingly greater resistance to large changes. Plastic environmental factors are essentially passive in that they retain their altered value after the cause of the change in value is removed. They may, however, exert variable resistance to change in a manner similar to a viscous medium, or friction. The orientation of plastic vectors is generally opposed to the vector sum of reaction. The actual trajectory of the site in environment space is thus the result of an *n*-dimensional vector summation of vegetation reaction, environmental elasticity, and environmental plasticity.

Additionally, in many vegetation types, the position of a site in vegetation space, and its trajectory through time, will be partly determined by herbivory from insects and larger animals, or pathogens. Accordingly, site favorability for some species may vary due to factors other than environmental modification. Considered on a local scale, herbivores may frequently migrate into or out of the vegetation being considered, and are best considered external to the system. Since their abundance will be only partly determined by the vegetation composition, it is best to model their effect as a vector acting directly on vegetation composition, without explicitly modeling the herbivores. On a larger scale, if the herbivore population can be considered to be regulated by the vegetation being considered, it is possible to add axes to the environment space, and consider the herbivores as a component of the environment, despite their biological nature.

Model behavior

Each species occupies a subspace of environment space which is delimited at its perimeter by the minimum requirements of the species, and which becomes increasingly favorable toward the species' optimum in the interior. Near the species' optimum the environment will support a greater biomass of the species, and the species will reproduce better than near the species' minimum requirements. The force of a species' reaction is roughly proportional to its biomass. If the species' reaction changes the environment of a site in such a manner as to move the site in environment space toward the species' optimum, the species will increase in biomass, driving the site more strongly toward the optimum, establishing a positive feedback loop.

For example, species with shade tolerant seedlings are capable of establishing under a mature canopy. Those seedlings which mature create or perpetuate shady conditions, which in turn encourages a high relative proportion of that species' reproduction, establishing a positive feedback loop.

Alternatively, a species' reaction may change the environment in such a way as to move the site away from the species' optimum, initiating a negative feedback loop. Such species tend to decline in biomass with time, or fail to reproduce. Rarely, however, will such a species drive itself to local extinction. As its biomass declines its reaction declines as well, and may become negligible before the species is locally extinct. Other species in the vegetation, however, may drive the environment beyond the amplitude of the species in question, enforcing local extinction.

For example, shade intolerant species metabolize best in full sunlight, but their growth unavoidably results in shade. The higher the biomass an environment initially supports, the less favorable the environment becomes for the subsequent generation, and the species declines. Other, more shade tolerant species, may reduce the available light below the requirements of the species in question, and cause local extinction.

The role of feedback loops in governing vegetation composition and environmental characteristics can become complex. Given generally modal plant responses to environmental factors, the increase or decrease of any factor may at first initiate a positive feedback loop which transforms into a negative feedback loop when the environmental factor passes the species optimum.

A more detailed mathematical treatment of these ideas could be presented, but my purpose is to pro-

mote an intuitive understanding and explore the implications of such a perspective for vegetation ecology. Roberts (1985) derived state transition equations for this system using fuzzy systems theory. Reed (1980) developed a similar model as the conceptual basis for a computer simulation model of forest succession. In a more *ad hoc* form, these ideas form the basis of forest simulation models referred to as 'gap models', which have been employed in a wide range of ecological studies (Shugart, 1984).

Discussion

Vegetation theory and vegetation dynamics

To some extent the environment is determined by factors inherent in the site and unchanging, such as the parent material of the soil. Many factors are elastic or plastic, however, and subject to modification by vegetation. The implication of environmental modification by vegetation is profound. Because the vegetation composition is determined in part by chance and modifies the environment in a manner determined by its composition, chance events can become embedded or canalized in the development of the vegetation and environment. The chance establishment of nitrogen-fixing species, or species which selectively accumulate nutrients, can affect the subsequent vegetation composition of a particular site for a long time. Thus, it cannot be assumed that two environments which were initially similar would subsequently support similar vegetation. As the vegetation developed on these sites, reaction may have modified these sites so that the current environments differ markedly. Two similar environments which become part of different feedback loops may diverge. Two dissimilar environments which become part of the same feedback loop may strongly converge with time.

From a dynamical systems perspective, stable species combinations are stable because they create or maintain a favorable environment for themselves, not because they are the species best adapted to an *a priori* environment. Unstable species combinations are unstable because they create an environment unfavorable to themselves and/or more favorable to a different combination of species.

Succession is not a development toward climax, but rather a change away from seral conditions. The successional pathways followed by vegetation, and the relation of vegetation composition to environment, are both determined by the properties of a fairly complex dynamical system which, however, can be characterized by a simple mechanist model.

McIntosh (1967) noted 'The community-type hypothesis and the continuum hypothesis may not be the only possible bases for a vegetation theory.' Essentially, both the community-type hypothesis and the continuum hypothesis can be viewed as dynamical systems which differ only in details of the particular system studied.

For community types to exist as other than subjective entities, vegetation must exhibit discontinuity over what was originally a continuous environment. The development of such a system would be encouraged by (1) few potential dominant species with markedly different physiologies, and (2) high environmental plasticity. For such a system, most points in environment space would fall within one positive feedback loop, governed by one or a few dominant species. With time, the environmental space would tend to become discontinuous, with the discontinuities imposed by vegetation reaction. Accordingly, with time, the vegetation would tend to assume certain modal compositions which result in a dynamic equilibrium with the modified environment.

Alternatively, for a continuum to exist, the vegetation must exhibit clinal variation that in some sense parallels a continuous environmental gradient. The development of such a system would be encouraged by (1) numerous potential dominant species, or species with similar physiologies, and (2) low environmental plasticity. Numerous potential dominants might allow situations where individual species' reactions largely cancel each other out so that the net community reaction is small. Alternatively, species with similar physiologies might create a situation where species' reactions were similar or substitutable. In either case, low environmental plasticity would ensure that the environment remains essentially constant, and therefore continuous.

The continuum hypothesis and the community type or association hypothesis have frequently been viewed as antithetical, yet both can be viewed as special cases of dynamical systems. Accordingly, it

is not necessary to hypothesize that the 'true' nature of vegetation is either continuous or discontinuous. Rather, either extreme or any point in between may obtain depending on the quantitative composition of the vegetation, the vegetation metabolism, and the environmental plasticity.

Vegetation analysis

One of the critical ideas of the dynamical systems perspective is that neither vegetation nor environment can be treated as an independent variable. The current state of both the environment and the vegetation composition depend on the previous states of both components. Accordingly, vegetation analysis should not employ the concepts of dependent and independent variables. Currently, the underlying model of many vegetation analyses is based on the algebra of multivariate functions. Such models assume that given a multivariate expression of environment there is one vegetation composition which is best adapted to the site. The task of vegetation analysis is then viewed as fitting the vegetation to environmental gradients, and the failure to achieve an exact fit to the functional model is attributed to 'noise' (Gauch, 1982). In light of a dynamical systems perspective of vegetation, a functional model can be characterized as environmental determinist.

The relationship between vegetation and environment is relational rather than functional. The environment determines which plant species can occur, but it does not determine which species will occur; the environment determines the maximum possible abundance of species which do occur, but it does not determine their actual abundance. As argued by Sukachev & Dylis (1964), the interpretation of vegetation as the direct product of a complex of physical site factors is erroneous. If two communities produce similar net or synthesized reaction vectors, either can exist in a given environment. Vegetation analyses based on a dynamical systems perspective should address the questions, 'Given an environment, what is the range of vegetation compositions possible, and what is the set of trajectories in environment space and vegetation space that correspond to these compositions?' Alternatively, 'Given a vegetation composition, what is the range of environments it can accommodate, and what is the set of associated trajectories?' I be-

lieve the natural mathematical framework for addressing such analyses is set theory, rather than multivariate algebra, because it defines explicit operators for relations rather than for functions.

Relation to ecosystem ecology

Developing a theoretical vegetation ecology on a dynamical systems perspective would allow vegetation ecologists to more easily integrate the results of ecosystem ecology, where the emphasis is generally on ecosystem function. Recognizing and quantifying the role of plants in modifying the environment, and the response of plants to the modified environment, leads to the possibility of simulation modeling as a significant tool in vegetation ecology. Additionally, increased awareness of the role of plant species in differentially modifying the environment may lead to contributions to ecosystem research by vegetation ecologists.

References

Austin, M. P., 1980. Searching for a model to use in vegetation analysis. Vegetatio 42: 11–21.

Austin, M. P., 1985. Continuum concept, ordination methods, and niche theory. Ann. Rev. Ecol. Syst. 16: 39–61.

Clements, F. E., 1916. Plant succession: an analysis of the development of vegetation. Publ. 242, Carnegie Inst., Washington D.C.

Cooper, W. S., 1926. The fundamentals of vegetation change. Ecology 7: 391–413.

Gauch, H. G. Jr., 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge.

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

Gleason, H. A., 1939. The individualistic concept of the plant association. Am. Midl. Nat. 21: 92–108.

Goodall, D. W., 1963. The continuum and the individualistic association. Vegetatio 11: 297–316.

Khil'mi, G. F., 1962. Theoretical forest biophysics. Transl. from Russian by R. Levin-Nadel. Israel Program for Scientific Translation, Jerusalem.

McIntosh, R. P., 1967. The continuum concept of vegetation. Bot. Rev. 33: 130–187.

Mueller-Dombois, D. & Ellenberg, H., 1974. Aims and methods of vegetation ecology. John Wiley and Sons, New York.

Reed, K. L., 1980. An ecological approach to modeling growth of forest trees. Forest Sci. 26: 35–50.

Roberts, D. W., 1985. Toward a dynamical systems vegetation theory. In: B. H. Banathy (ed.), Proceedings: Society for general systems research. Volume II. Systems inquiring: Applications, pp. 583–588.

Shugart, H. H., 1984. A theory of forest dynamics. Springer-Verlag, New York.

Sukachev, V. & Dylis, N., 1964. Fundamentals of forest bio-geocoenology. Transl. from Russian by J. M. McLennan. Oliver and Boyd, London.

Tansley, A. G., 1935. The use and abuse of vegetation terms and concepts. Ecology 16: 284–307.

Whittaker, R. H., 1953. A consideration of climax theory: The climax as a population and pattern. Ecol. Monogr. 23: 41–78.

Whittaker, R. H., 1967. Gradient analysis of vegetation. Biol. Rev. 42: 207–264.

Whittaker, R. H., 1974. Climax concepts and recognition. In: R. Knapp (ed.), Vegetation dynamics, pp. 137–154. Junk, The Hague.

Whittaker, R. H., 1978. Direct gradient analysis. In: R. H. Whittaker (ed.), Ordination of plant communities, pp. 7–50. Junk, The Hague.

Wuenscher, J. E., 1969. Niche specification and competition modeling. J. Theo. Biol. 25: 436–443.

Wuenscher, J. E., 1974. The ecological niche and vegetation dynamics. In: B. R. Strain & W. D. Billings (eds), Vegetation and environment, pp. 39–45. Junk, The Hague.

Models for the analysis of species' response to environmental gradients

M. P. Austin*
Division of Water and Land Resources, CSIRO, G.P.O. Box 1666, Canberra, A.C.T., 2601, Australia

Abstract

A procedural model for vegetation analysis is presented. Suggestions are made that analysis methods can test theory as well as examine vegetation-environment correlations. Gauch and Whittaker's propositions regarding species behavioural properties expected for an individualistic continuum are tested on a eucalypt forest data set. The data set is carefully stratified to control environmental heterogeneity. The shape and distribution of species response curves are then examined along a temperature gradient using 750 sites. The conclusions are:
(1) Bell shaped response curves to environmental gradients are not universal
(2) Positive-skewed curves are characteristic of major canopy species in eucalypt forest in southern New South Wales
(3) Species richness increases with temperature along the gradient
(4) Tests of other propositions regarding species modes and ranges are confounded by the change in species richness along the gradients
(5) More rigorous statistical analysis and analyses on other vegetation types are needed.

Introduction

Vegetation analysis, and in particular the use of numerical multivariate methods, is often dismissed as simply correlation (Levins & Lewontin, 1980), or as subjecting 'the data pile to ... multivariate analysis in the hope that ... something of value may be sifted out' (Salt, 1983). Goodall (1970) observed that 'the subject is likely to mark time until its biological aspects resume their rightful place i.e. as master rather than slave to mathematical and statistical considerations'. Vegetation analysis can be defined as the application of explicit numerical techniques to a data matrix composed of vegetation, plus other biotic and/or environmental variables recorded by sites. The thesis presented here is that vegetation analysis can make a contribution to current ecological issues.

Debate about the relative roles of competition, environment, predation and chance [Strong, 1983] in determining the composition of communities has been largely conducted by animal ecologists, though competition (Harper, 1977), environmental control of species distribution (Whittaker, 1975, 1978), and the role of grazing (Harper, 1977) have a long history of study in plant ecology (see Noy-Meir & Van der Maarel, 1987). However, vegetation scientists will not make a significant contribution to ecology unless they agree on their purpose, on

common procedures and on how vegetation theory should be tested. Often the plant ecologist is simply concerned to detect (understand?) the environmental factors which correlate with variation in vegetation composition, i.e. to generate hypotheses applicable only to that community and location. This form of analysis will not make a significant contribution to ecological theory, yet cannot be divorced from theory. Any vegetation study makes implicit assumptions about theory and what constitutes an appropriate methodology. Note theory is used here in the sense of 'a coherent group of general propositions used as principles of explanation for a class of phenomena'. What is currently needed is for vegetation scientists to recognize that their studies imply a certain procedural model or paradigm. Further, the component procedures of this model can be used to test specific aspects of ecological theory.

Procedural model for vegetation analysis

Figure 1 outlines the broad stages required in a procedural model for vegetation analysis; the specific purpose of the study and existing vegetation theory determine the decisions to be made at each stage. The purpose may be applied (e.g. mapping the plant communities of a nature reserve in relation to
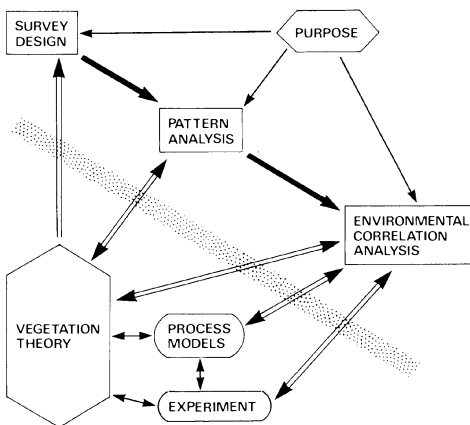
Fig. 1. Procedural model for vegetation analysis. Solid arrows indicate the usual steps taken. The shaded bar cuts those processes often neglected by vegetation analysts.

the conservation of a rare orchid) or it may be academic (e.g. testing Gleason's (1926) individualistic concept of species environmental distribution). In both cases, decisions have to be made on (1) the relevance of current mathematical methods to ecological theory and the purpose of the study, (2) appropriate sampling strategies to either test hypotheses or describe variation and (3) what environmental variables are to be measured and why. Such a sequence of steps is common to all vegetation analysis studies and explicit decisions are needed.

Each stage contains several components. The scope of decisions needed at each stage is briefly discussed in the following sections. Further discussion of survey design can be found in Austin (1984) and Noy-Meir (1971).

Survey design

This stage has three components: 1. Domain definition, 2. Sampling, 3. Attribute measurement. Domain definition refers to a statement of the region to be studied, the units of interest and the information required. Sampling presents a particular problem. While stratified random sampling is recognized as an ideal sampling procedure, practicality may demand that less optimal designs be used. How attributes are measured may restrict the questions which can be asked of the data. Hypothesis-testing derived from theory should determine survey design; for example a very special design would be required to test the concept of a vegetation continuum (cf. Austin, 1985).

Pattern analysis

Intrinsic pattern analysis is a form of exploratory data analysis (sensu Tukey, 1977) which seeks to expose trends, clusters or outliers in the floristic data matrix of species and sites obtained from a vegetation survey. Classification and ordination are the most popular methods for detecting such patterns. Contrary to some opinions (Gauch, 1982; Greig-Smith, 1983; Kershaw & Looney, 1985), the prob-

lems associated with choosing an ordination method consistent with current vegetation theory remain unresolved (Clymo, 1980; Minchin, 1983; Austin, 1985; Faith *et al.*, 1987). If we are to understand the composition, distribution and abundance of vegetation, 'pattern analysis' methods are needed to reduce the multivariate data to manageable units in a few significant dimensions. The patterns in these major dimensions may then be subjected to hypothesis-testing.

Environmental correlation analysis

Results from pattern analysis, either as classified groups or ordination trends, are examined for their correlation with environment. Various techniques are used in this step including simple graphic plotting (e.g. Direct Gradient Analysis (DGA), Whittaker, 1978), regression analysis and canonical correlation. This can be regarded as a search for some 'significant' relationship or as a careful testing of multiple hypotheses.

The development of suitable statistical models for use with DGA (Austin *et al.*, 1984; Austin, 1985; Austin *et al.*, in prep.) now allows the second alternative to be explored more fully than previously. Predictive models of the environmental determinants of vegetation can be formulated using modern statistical methods such as Generalized Linear Models (GLM, see below for further discussion). These statistical models, while allowing the rejection of hypotheses regarding the influence of certain factors, will only provide a description of the role of the significant factors.

Experimental tests

Most vegetation studies are only exploratory analyses of the observed floristic variation and the possible environmental correlates. Confirmatory studies are required, although the design of these has yet to be agreed on by plant ecologists. For vegetation scientists, it is necessary to cross the boundary shown in Fig. 1 between observational analysis and experimental or theoretical analysis as a matter of course when studying a problem (Austin, 1985).

Process models

Two types of process model need to be derived from exploratory analysis and experimental tests to ensure that an integrated form of ecological theory can be developed. The first type comprises environmental process models. Environmental indices (often termed scalars) that can be related to the physiological condition of plants are needed (Austin *et al.*, 1984). The second type of model is the species interaction or competition model; examples are Tilman (1982) based on exploitative competition, or Grime (1979) based on interference mechanisms. The forest-dynamics models (Shugart, 1984) combine features of physiological growth models and species interaction models. Validation of models of processes effecting long-lived species however poses problems.

Vegetation theory. Much of the current so-called vegetation theory is simply a set of descriptive propositions. For example, the continuum concept, with its associated ideas of bell-shaped response curves is devoid of any mechanistic processes which will generate the bell-shaped responses (Austin, 1985). Until process-oriented theories are available, vegetation analysis will remain dependent on descriptive concepts to guide survey design, pattern analysis and environmental correlation methods (Austin, 1980, 1984). The inconsistency between theory and methodology is clearly demonstrated by the continuing debate over ordination methods.

The role of vegetation analysis need not be simply exploratory; predictions based on a theory can be tested using the same procedural models. In the next section, the theoretical propositions put forward by Gauch & Whittaker (1972) as the observable properties of a vegetation continuum are tested using a specified survey design and simple graphical techniques for environmental correlation.

Evaluation of ordination methods and the role of species response curves

Evaluation of ordination methods has developed rather slowly from subjective evaluation of which method makes most ecological sense (i.e. fits the evaluator's pre-conceived ideas), through comparison with environmental ordination methods, to the use of artificial data sets where the generating model of vegetation response is known. Ordination methods have changed drastically as a result of the recognition that a species response curve to an environmental gradient is likely to be a non-monotonic curve and probably bell-shaped (Swan, 1970; Gauch, 1982; Austin, 1985). Since Swan's pioneering work, most ordination methods have been evaluated on the basis of artificial data sets with pseudo-Gaussian species responses. This descriptive model of vegetation response is based on the generalization of Swan's ideas developed by Gauch & Whittaker (1972).

Tests of current vegetation models

Given the popularity of the algorithm developed by Gauch & Whittaker (1972) for generating artificial data sets, the propositions about vegetation response on which it is based should be examined critically (see also Austin, 1976; Minchin, 1987). The major propositions are:
1. Species response curves approximate normal (Gaussian) curves i.e. the curves are symmetric and bell-shaped in relation to an environmental gradient
2. The modes of major species tend to have a uniform distribution along the environmental gradient, while the minor species modes are randomly distributed
3. The modal abundances (here, maximum observed or modal frequency), standard deviation (here, range) and the modal position (i.e. position of the maximum observed frequency on the environmental gradient), are uncorrelated.

Althouth alternative models are possible (Austin, 1976; Fresco, 1982; Minchin, 1983 unpubl. thesis),

none has been subjected to any rigorous scrutiny. The original descriptive model (Gauch & Whittaker, 1972) was based on subjective evaluation of direct gradient analysis studies about which reservations have been expressed (Austin *et al.,* 1984; Austin, 1985).

I test Gauch & Whittaker's (1972) major propositions by using a data set within which environmental variability is carefully controlled and where the major environmental determinants had previously been shown to be rainfall and temperature (Austin *et al.,* 1984). These two environmental variables (gradients) can be used to predict the probability of occurrence of eucalypt species in south-eastern New South Wales. Some initial analyses (Austin *et al.,* 1983; Austin *et al.,* 1984) with a data set of 1 286 observations demonstrated that species frequency-of-occurrence response curves tended to be bell-shaped though no test for skewness was made. In addition, it was shown that spurious bimodal response curves could be produced by certain sample distributions within the environmental space defined by mean annual temperature and mean annual rainfall.

Simple graphical analysis can address some of
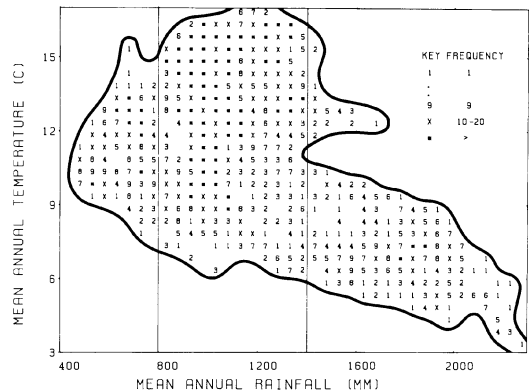
Fig. 2. Distribution of total vegetation sample of 6 080 plots in an environmental space defined by mean annual temperature and mean annual rainfall. The lines at 800 mm and 1 400 mm indicate the portion of the space from which the sample for analysis in this paper was taken. Note environmental conditions beyond boundary line do not exist in the study area. Numbers and symbols indicate number of plots in cells of size 0.5 °C and 100 mm rainfall: 10–19, * ≥ 20.

the issues, provided a large suitably stratified data set is available. Availability of suitable data limit the analysis to the use of presence/absence data. The shape of the species response curve to an environmental gradient when based on presence data is equivalent to defining the qualitative environmental realized niche (QERN) for the species with respect to that gradient. The original DGA studies (Whittaker, 1956, 1960; Whittaker & Niering, 1965) used elevation and an indirect ordination technique to provide a moisture index gradient with some stratification for lithology. With a large data set of

6080 plots (Austin *et al.,* in prep., see Fig. 2) it is possible to stratify on more of the environmental variables which might distort the response curves. Figure 3 shows the response curves of those species with maximum probability of occurrence greater than 0.3 i.e. major species, in relation to temperature where rainfall is between 800–1400 mm per annum. The sites used for Fig. 3 were selected on the following criteria:

a) Slopes ≥7°, excluding other topographic positions for example ridges and gullies

b) Radiation index value (measure of aspect, Aus-
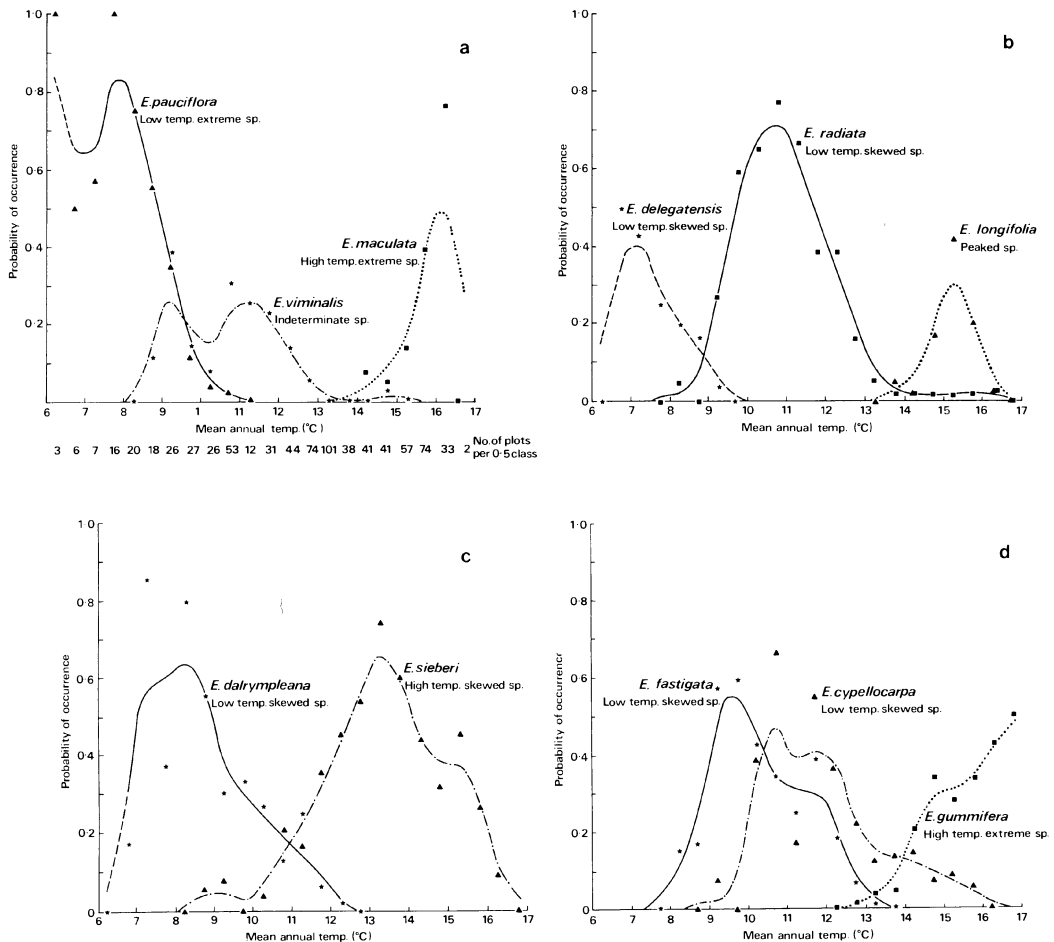
Fig. 3. Examples of the species response curves (QERN) to mean annual temperature of major canopy species from sclerophyll forest in south-eastern New South Wales, Australia. The plots used for these curves were selected from the 800–1400 mm rainfall region with slopes ≥7°, radiation indices between 0.9–1.05 and with no major disturbance.

tin *et al.,* (1984)) between 0.90–1.05, excluding hot (northern) and protected moist (southern) slopes

c) No known major disturbance

d) Within the 800–1400 mm rainfall by 6°–16.5 °C temperature region of the environmental space. This region was chosen to avoid problems of uneven sample distribution yet provide as large a sample as possible, in this case 750 sites (Fig. 2).

Further details of the study area can be found in Austin *et al.* (1984). A list of the species used in the analysis is available from the author.

Gauch & Whittaker's first proposition. There is a marked preponderance of skewed curves among the major species. A simple classification of these QERN curves on shape, noise and maximum modal frequency can be made (Table 1). In most cases, the skew is positive, i.e. the position of the mode (maximum) is less (at a lower temperature) than the mean. Examples include *Eucalyptus dalrympleana, E. cypellocarpa, E. fastigata, E. radiata* and *E. agglomerata* (Fig. 3). Negative skewness is uncommon (e.g. *E. sieberi*), though there are many species (Table 1) whose shape cannot be determined from this data set because of their positions at extreme high temperatures. Given the graphical evidence, statistical tests are unnecessary to prove the existence of departures from a bell-shaped curve. A variety of reasons may be responsible. There may be a complex interaction with rainfall over the range 800–1400 mm, with the mode shifting to progressively higher (or lower) temperatures with increasing rainfall. Examination of species reponse curves for the three rainfall classes 800–999, 1000–1199, 1200–1400 mm shows no such consistent patterns. A lurking variable may distort the response curve at some position along the gradient (cf. *E. pauciflora,* Austin *et al.,* 1984). However, the skewness occurs at all positions along the gradient so a separate factor would need to be postulated for every species.

I conclude from this graphical analysis that:

(1) when environmental variability is controlled, symmetric response curves are not universal for environmental factor-gradients, i.e. species response curves do not necessarily approximate a Gaussian curve,

(2) positive-skewed response curves are a characteristic of major canopy species of sclerophyll forest in south-eastern New South Wales, Australia, in relation to mean annual temperature with rainfall between 800–1400 mm.

No stronger generalizations are possible until further similar tests of the proposition have been made in other regions and vegetation types.

Gauch & Whittaker's second proposition. This states that the modes of major species are evenly

Table 1. Classification of species response (QERN) curves in relation to maximum modal frequency.

Maximal modal frequency	≥ 0.3	< 0.3
Indeterminate	1	9 (9)
Low temp. extreme	2 (1)	–
Low temp. skewed	7 (1)	–
Symmetric	2	5 (1)
High temp. extreme	2	1
High temp. skewed	10 (4)	3 (3)

( ) indicates number of species considered 'noisy'.

Table 2. Modal position of species in relation to temperature.

Temperature class	Maximal modal frequency		
	≥ 0.3	< 0.3	Total
7	1	1	2
7.5	–	1	1
8	–	–	0
8.5	–	–	0
9	2	–	2
9.5	1	1	2
10	–	1	1
10.5	3	–	3
11	–	2	2
11.5	–	–	0
12	–	4	4
12.5	–	–	0
13	3	1	4
13.5	–	1	1
14	–	1	1
14.5	2	1	3
15	2	2	4
15.5	–	–	0

distributed along environmental gradients while those of minor species tend to be randomly distributed. The modal positions for those species whose position is unequivocal are tabulated in Table 2. The results appear clustered. An approximate statistical test can be made using χ^2 and testing for departure from a random expectation of equal numbers in each temperature class. The value of $\chi^2 = 21.6$ with 17 degrees of freedom for the major species is not quite significant at 5% level. But the hypothesis put forward by Gauch & Whittaker is that species are regularly distributed and not random. This requires a one-tailed test for χ^2 *less* than the random expectation. The 5% level for accepting a regular departure from randomness is 8.67.

The proposition that species modes are regularly distributed is not acceptable for this data set, and the null hypothesis of random distribution of modes cannot be rejected. However, the evidence (high value of χ^2) suggests that the modes may have a contagious, rather than a random distribution.

Other propositions and the role of species richness. Given that the response curves are not pseudo-Gaussian, testing the values for the propositions regarding the standard deviation is inappropriate. The range of a species along the temperature gradient can be examined (Table 3), as can the positions of upper and lower limits of species (Table 4). No species occurs at all temperatures, the most frequent range for a species is between 3.5 and 5.0 °C. The lower limits of species are variously distributed. The upper limits are clustered at the highest temperatures.

There is a partial explanation for this that was not included in the original propositions. Species richness is not uniform along the temperature gradient. Though in a later version Gauch & Whittaker (1976) allow species richness to vary in their simulation algorithm, no consistent relationships could be established from Whittaker's original DGA studies (Whittaker, 1956, 1960; Whittaker & Niering, 1965; though see Peet, 1978). The tree line for eucalypts occurs at about 4 °C estimated mean annual temperature. Only one species occurs at the tree line, *E.*

Table 3. Distribution of size of species ranges in relation to temperature.

Temperature range (°C)	Major species ≥0.3	Minor species <0.3	Total
0.5	–	1	1
1	–	1	1
1.5	–	1	1
2	–	–	0
2.5	2	2	4
3	–	–	0
3.5	2	4	6
4	–	1	1
4.5	4	2	6
5	1	–	1
5.5	1	2	3
6	2	–	2
6.5	1	–	1
7	–	–	0
7.5	1	1	2
8	1	1	2
≥8.5	–	–	0

Table 4. Positions of species upper and lower limits in relation to temperature.

Temperature class (°C)	Lower limit			Upper limit		
	Major species ≥0.3	Minor species <0.3	Total	Major species ≥0.3	Minor species <0.3	Total
6.5	2	–	2	–	–	0
7	–	1	1	–	–	0
7.5	–	1	1	–	–	0
8	3	2	5	–	–	0
8.5	2	–	2	–	–	0
9	2	1	3	1	–	1
9.5	–	3	3	–	1	1
10	–	1	1	–	–	0
10.5	1	–	1	1	–	1
11	–	1	1	–	–	0
11.5	3	2	5	–	–	0
12	1	1	1	1	–	1
12.5	2	1	3	–	2	2
13	1	3	4	1	1	2
13.5	1	–	1	1	2	3
14	4	–	4	1	–	1
14.5	–	–	0	1	2	3
15	1	1	2	–	2	2
15.5	–	–	0	2	5	7
16	–	–	0	6	–	6

pauciflora subspecies *niphophila*, and species richness increases with temperature. It follows therefore that most species will have their upper limits at the higher temperatures. The mean species number per site (of all canopy species) per temperature class is plotted in Fig. 4 together with the regression line for all 750 plots. Thus, propositions regarding distribution of species modes, ranges and limits in relation to temperature are confounded with species richness. The number of species limits and modes must decline with decreasing temperature, while species ranges will also be constrained.

I conclude that:

1. Towards the limits of tree growth at lower temperatures in south-eastern Australia, tree species richness is a function of temperature

2. Further simulation algorithms for evaluation of ordination methods in relation to vegetation analysis should consider the implication of this, particularly as to how it effects other assumed properties of artificial vegetation data.

A fuller analysis of species richness of eucalypt forests in relation to temperature and rainfall is currently in preparation (Margules *et al.,* in press). Several studies have recently addressed the relationship of species richness to single environment factors (Grime, 1979; Tilman, 1982; Abramsky &
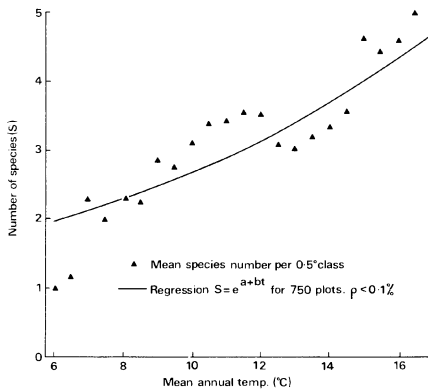
Rosenzweig, 1984). Margules *et al.* show that it is necessary to consider the major environmental gradients simultaneously if consistent results are to be obtained (see also Peet, 1978; Minchin, 1987). They show a complex response for eucalypt species number in relation to mean annual temperature (3 °–16.8 °C), mean annual rainfall (500–2300 mm) and relative radiation for a slightly smaller data set for the same region than that from which this subset was selected. The environmental stratification used here simplifies this relationship as it effects the pattern of species responses.

Two other propositions were put forward by Gauch & Whattaker (1972); that the modal values were lognormally or lograndomly distributed and that the values of the mode, modal position and 'standard deviation' for species were uncorrelated. The evidence for the frequency distribution of modal values is equivocal (Table 5). Tests for correlations between the modal values, modal positions and ranges of species showed some relationship. The maximum observed frequency (modal value) of a major species can be predicted from either the species range or its modal position using regression. When both measures were used, only range was significant (at a 0.1% level).

These results suggest that the modal value for a species is greater if the species temperature range is large or its modal position is at a low temperature

Fig. 4. Relationship between species richness and mean annual temperature of sclerophyll forest plots in south-eastern New South Wales. The data points ( ) are: Mean number of species per plot in each 0.5 °C class. Regression lin $S = e^{a + bt}$ is based on species number (S) for *all* 750 plots, as a consequence the points and regression line do not correspond.

Table 5. Frequency distribution for modal values of all species with well defined modes.

	Frequency
<0.1	4
0.1–0.19	5
0.2–0.29	9
0.3–0.39	6
0.4–0.49	1
0.5–0.59	4
0.6–0.69	1
0.7–0.79	5
0.8–0.89	1
0.9–1.0	(1)

(1) one species *E. pauciflora* reaches 1.0 at an extreme, i.e. not a well-defined mode.

or both. The maximum observed frequency in a temperature class usually corresponds to the mode for a particular species. The maximum observed frequency value in any temperature class is related to temperature. A regression equation relating maximum frequency and temperature is significant at a 1% level. The observed maximum frequency is greatest at lower temperatures. This is of course partially a consequence of the decline in species richness at low temperatures. However, there is no logical correlation. As species richness declines, so might the density of trees per unit area and hence the maximum observed frequency for a species might decline at low temperatures as density becomes low. This is not the case under the range of conditions studied.

The conclusion can be drawn from these data that the parameters of the response curves which define the QERN are not independent and are related to their position on the environmental gradient for *this vegetation type under the specified conditions.*

Discussion

The paper by Gauch & Whittaker (1972) is, in fact, a major statement of the individualistic continuum concept. Few other papers on vegetation analysis can be said to have addressed vegetation theory in recent years (Austin, 1980; Fresco, 1982). Theory has been left to other types of ecologists (e.g. Grime, 1973, 1979; Tilman, 1982; Brown, 1984). If we are to develop vegetation theory, then every concept needs to be questioned and tested.

By careful selective *stratification* of observations it has been possible to test a number of hypothetical propositions about the behaviour of vegetation using simple graphs, tabulations and statistical tests. However, the data, are subject to error and more robust and rigorous statistical methods such as GLM are needed to provide a clear unambiguous definition of the realized environmental niche of the species (Austin *et al.,* 1983; Austin *et al.,* 1984; see also Minchin, 1987). The next step should be to test the shape of the species QERN; can we distinguish asymmetric and symmetric response curves?

Table 6. Conclusions regarding species response curves for canopy species in eucalypt forest in south-eastern New South Wales.
N.B. When environmental variables are selectively stratified and mean annual temperature is the gradient.

1. Bell-shaped response curves are not universal
2. Positive-skewed curves are characteristic of major canopy species
3. Species richness is predictable from mean annual temperature
4. Test of propositions regarding species modes, and ranges are confounded by changes in species richness along the gradient
5. Maximum observed frequency is inversely related to temperature
6. Similar tests on other vegetation types are needed

Colleagues and I (Austin, Nicholls & Margules, in prep.) have used GLM with our eucalypt forest data to indicate how this might be done for several factors simultaneously.

Other propositions of Gauch & Whittaker (1972) have been found incorrect at least for sclerophyll forest in south-eastern N.S.W. using qualitative data (Table 6). The skewness of the response curves of eucalypts is more consistent with the hypothetical curves of Van der Maarel (1976) and Austin (1976). These could however be considered subcategories of the Mueller-Dombois & Ellenberg (1974) model. The issue must remain open, however, as there is evidence that some of the skewness may be due to compensating factors. For example the occurrence of *E. pauciflora*(snow gum) at high mean annual temperatures is considered to be associated with 'cold-air' drainage hollows. Vegetation analysis with DGA and GLM methods could be used to test such hypotheses. The other propositions are inadequate due to the influence of species richness and abundance effects. It has been argued before in relation to vegetation analysis (Austin & Greig-Smith, 1968; Austin, 1981; see also Peet, 1978) that these components of diversity are related to environment. Similar conclusions have been reached from experiments (Grime, 1979; Tilman, 1982; Austin & Austin, 1980). Margules *et al.* (in press) have extended the analysis of species richness to a three-dimensional environmental space and in-

dicate that (a) species richness is a joint function of environmental variables and (b) that the definition of appropriate guilds may be crucial to detecting species richness/environment relationships (cf. Minchin, 1987).

Methodological studies continue to be necessary but should be subordinate to the testing of existing theoretical concepts.

References

Abramsky, Z. & Rosenzweig, M. L., 1984. Tilman's predicted productivity – diversity relationship shown by desert rodents. Nature 309: 150–151.

Austin, M. P., 1976. On non-linear species response models in ordination. Vegetatio 33: 33–41.

Austin, M. P., 1980. Searching for a model for use in vegetation analysis. Vegetatio 42: 11–21.

Austin, M. P., 1981. The role of diversity problems in vegetation classification. In: A. N. Gillison & D. J. Anderson (eds), Vegetation classification in Australia, pp 125–139. ANU Press: Canberra.

Austin, M. P., 1984. Problems of vegetation analysis for nature conservation. In: K. Myers, C. R. Margules & I. Musto (eds), Survey methods for nature conservation, pp 101–130. CSIRO Division of Water and Land Resources, Canberra.

Austin, M. P., 1985. Continuum concept, ordination methods and niche theory. Ann. Rev. Ecol. Syst. 16: 39–61.

Austin, M. P. & Austin, B. O., 1980. Behaviour of experimental plant communities along a nutrient gradient. J. Ecol. 68: 891–918.

Austin, M. P. & Greig-Smith, P., 1968. The application of quantitative methods to vegetation survey. II. Some methodological problems of data from rainforest. J. Ecol. 56: 827–844.

Austin, M. P., Cunningham, R. B. & Good, R. B., 1983. Altitudinal distribution in relation to other environmental factors of several Eucalypt species in southern New South Wales. Aust. J. Ecol. 8: 169–180.

Austin, M. P., Cunningham, R. B. & Fleming, P. M., 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio 55: 11–27.

Brown, J. J., 1984. On the relationship between abundance and distribution of species. Amer. Nat. 124: 255–279.

Clymo, R. S., 1980. Preliminary survey of the peat-bog Hummell Knowe Moss using various numerical methods. Vegetatio 42: 129–148.

Faith, D. P., Minchin, P. R. & Belbin, L., 1987. Compositional dissimilarity as a robust measure of ecological distance. Vegetatio 69: 57–68.

Fresco, L. F. M., 1982. An analysis of species response curves and of competition from field data: some results from heath vegetation. Vegetatio 48: 175–185.

Gauch Jr, H. G., 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge.

Gauch Jr, H. G. & Whittaker, R. H., 1972. Coenocline simulation. Ecology 53: 446–451.

Gauch Jr, H. G. & Whittaker, R. H., 1976. Simulation of community patterns. Vegetatio 33: 13–16.

Gleason, H. A., 1926. The individualistic concept of the plant association. Bull. Torrey Bot. Club 53: 1–20.

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

Greig-Smith, P., 1983. Quantitative plant ecology. 3rd ed. Blackwell Scientific Publications, Oxford.

Grime, J. P., 1973. Competitive exclusion in herbaceous vegetation. Nature 242: 344–347.

Grime, J. P., 1979. Plant strategies and vegetation processes. John Wiley, Chichester.

Harper, J. L., 1977. Population biology of plants. Academic Press, London.

Kershaw, K. A. & Looney, J. H. H., 1985. Quantitative and dynamic plant ecology, 3rd ed. Edward Arnold, London.

Levins, R. & Lewontin, R., 1980. Dialectics and reductionism in ecology. In: E. Saarinen (ed.), Conceptual Issues in Ecology, pp 107–138. Reidel, Dordrecht.

Margules, C. R., Nicholls, A. O. & Austin, M. P. (in press). Diversity of *Eucalyptus* species predicted by a multi-variable environmental gradient. Oecologia.

Minchin, P. R., 1987. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 00: 000–000.

Mueller-Dombois, D. & Ellenberg, H., 1974. Aims and methods of vegetation ecology. Wiley, New York.

Noy-Meir, I., 1971. Multivariate analysis of semi-arid vegetation in south-eastern Australia: Nodal ordination by component analysis. Proc. Ecol. Soc. Aust. 6: 159–193.

Noy-Meir, I. & Van der Maarel, E., 1987. Relations between community theory and community analysis in vegetation science: some historical perspectives. Vegetatio 69: 5–15.

Peet, R. K., 1978. Forest vegetation of the Colorado Front Range: patterns of species diversity. Vegetatio 37: 65–78.

Salt, G. W., 1983. Roles: their limits and responsibilities in ecological and evolutionary research. Am. Nat. 122: 697–705.

Shugart, H. H., 1984. A theory of forest dynamics. Springer, New York.

Strong, D. R., 1983. Natural variability and the manifold mechanisms of ecological communities. Am. Nat. 122: 636–660.

Swan, J. M. A., 1970. An examination of some ordination problems by use of simulated vegetation data. Ecology 51: 89–102.

Tilman, D., 1982. Resource competition and community structure. Princeton University Press, Princeton N.J.

Tukey, J. W., 1977. Exploratory data analysis. Addison-Wesley, Reading, Mass.

Van der Maarel, E., 1976. On the establishment of plant community boundaries. Ber. Deutsch. Bot. Ges. 89: 415–443.

Whittaker, R. H., 1956. Vegetation of the Great Smoky Mountains. Ecol. Monogr. 26: 1–80.

Whittaker, R. H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecol. Monogr. 30: 279–338.

Whittaker, R. H., 1975. Communities and ecosystems, 2nd ed. MacMillan, New York.

Whittaker, R. H. (ed.), 1978. Ordination of plant communities. Handbook of Vegetation Science 5. 2nd ed. Junk, The Hague.

Whittaker, R. H. & Niering, W. A., 1965. Vegetation of the Santa Catalina Mountains, Arizona: A gradient analysis of the south slope. Ecology 46: 429–452.

The individualistic and community-unit concepts as falsifiable hypotheses*

Bill Shipley & Paul A. Keddy**
Department of Biology, University of Ottawa, Ottawa, Ont., Canada K1N 6N5

Abstract

This study reformulates the community-unit and individualistic concepts of plant communities as hypotheses concerning the distribution of species' boundaries along a gradient. These hypotheses are tested by an analysis of deviance on data derived from a direct-gradient analysis of a freshwater marsh plant community in Breckenridge, Quebec, Canada. Boundaries are clustered at certain intervals along the gradient ($p < 0.001$), contradicting the individualistic hypothesis. Upper boundaries are not consistently clustered at the same intervals as lower boundaries ($p < 0.001$), contradicting the community-unit hypothesis. Thus, neither of the two usual models of community structure explain the patterns found in Breckenridge Marsh, suggesting that the historical dichotomy is too limited. Hypotheses of pattern should be tested using inferential statistics. Hypotheses of mechanism should be tested by experimentation. The way out of the community-unit vs. individualistic community debate is to deny the dichotomy and to consider multiple working hypotheses of community structure.

Introduction

The debate concerning of nature of community organization has continued sporadically for more than 70 years (see Whittaker, 1962 for a review of the controversy). The Clementsian interpretation (Clements, 1916), also called the 'community-unit concept' (Whittaker, 1975) was accepted by the majority of ecologists during the first part of this

---

century. The works of Cain (1947), Egler (1947), Mason (1947), Whittaker (1951, 1956, 1967) and others quickly converted the majority of anglo-american community ecologists (McIntosh, 1967, 1975) to the 'individualistic concept' of community organization based primarily on the ideas of Gleason (1926, 1939). This transition to an individualistic viewpoint was due in part because, by using direct or indirect gradient analysis (Whittaker, 1967, the observed pattern of species replacements along a gradient was stated to be inconsistent with the patterns predicted by Clements.

In the context of testing between the two concepts mentioned above, direct gradient analysis still suffers from the problem of using subjective methods of analysing the observed patterns by failing to use inferential statistics to compare empirical data to the proposed models. In other uses of direct gradient analysis, Austin *et al.* (1984) have shown how Generalized Linear Modeling may be used to

test between competing hypotheses concerning species ecological response curves, but their method cannot be used directly to test between the community-unit and individualistic concepts (see also Austin, 1987).

In this paper we use a direct gradient analysis (Whittaker, 1967) of the emergent vegetation of a freshwater marsh to obtain data on the distribution of species' boundaries relative to a gradient-complex of water depth. After stating falsifiable hypotheses based on the predicted patterns of the individualistic and community-unit concepts, the observed pattern of boundaries along the gradient is compared to these hypotheses using an analysis of deviance (McCullagh & Nelder, 1983).

Hypotheses for the patterns implied by the two concepts

The community-unit concept claims that when species' distributions are plotted along some gradient or gradient-complex whose rate of change is constant, there exist groups of species. i.e. 'communities', which replace themselves along the chosen gradient (Whittaker, 1975). Within each grouping most species have similar distributions and the end of one group coincides with the beginning of another (Fig. 1).

The individualistic concept states that 'centers and boundaries of species' distributions are scattered along the environmental gradient' [Whittaker, 1975]. No distinct groups of species are predicted to exist therefore precluding the coincidence of one group with another.
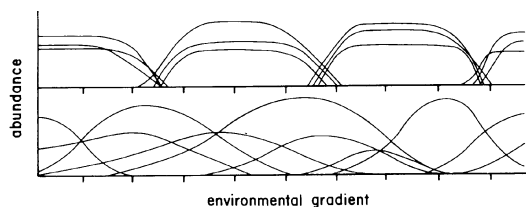
Fig. 1. Predicted distributions of species along an environmental gradient, according to the two concepts. Above: the community-unit concept; below: the individualistic concept (after Whittaker, 1975).

Following Pielou (1975, 1977), explicit hypotheses of these two concepts are formulated using the following terms: define 'upper boundary' as that point on the gradient where the distribution of a species (not a group) begins relative to an arbitrary direction on the gradient. Define 'lower boundary' of a species as that point on the gradient where its distribution ends.

The individualistic hypothesis states that:
(i) the average number of boundaries (both upper and lower) in each interval of the gradient should be equal except for random variation about the mean;
(ii) the number of upper boundaries per interval of the gradient should be independent of the number of lower boundaries.

The community-unit hypothesis states that:
(i) there should be significantly more boundaries (both upper and lower) in some intervals of the gradient than in others, i.e. boundaries are clustered;
(ii) the number of upper and lower boundaries per interval should increase and decrease together along the gradient.

The patterns predicted by these hypotheses are shown in Fig. 2. Note that the 2×2 table suggests at least two other possibilities for patterns along a gradient, as pointed out by Whittaker (1975).

Fig. 2. The number of community patterns which could be obtained from species whose boundaries are either (a) randomly distributed or clustered and whose upper and lower boundaries are either (b) coincident or independent. Note that the two historical community concepts do not exhaust the possibilities shown.

Methods

The study site is near Breckenridge, Quebec, which is approximately 40 km northeast of Ottawa, Ontario, Canada (45°48′00′′N, 45°57′30′′W). This marsh occurs on a very gently sloping shoreline. On the highest ground is a forest dominated by *Acer saccharinum* which is flooded during the high water period each spring. The study site itself consists only of herbaceous emergent species (see Fig. 4); this portion of the marsh extends for ca 200 m although not all of this section of marsh was actually sampled. Beyond 200 m from the edge of the tree canopy, floating leaved species such as *Nuphar variegatum* and *Potamogeton* spp. are found, giving way finally to submerged vegetation. There are no obvious vegetational discontinuities in this marsh; it is not at all like stepping off a bank and into the river bed. Furthermore, the section of the marsh actually studied was well within the average annual upper and lower water levels for this section of the river. Since many abiotic environmental factors are correlated with water depth (e.g., Hutchinson, 1975; Spence, 1982; Keddy, 1985) there are no obvious abiotic discontinuities either.

To obtain a random sample of the emergent community a 200 m line was established running along the edge of the tree canopy and divided into 5 m segments. Thirteen of these segments were chosen using a random number table to be the locations of belt transects running down into the marsh. Each transect was therefore at least 5 m distant from any other transect.

The water level on 24 July 1984 was marked and will be called the 'standard zero' water level; this corresponds to a Geodetic height (Water Survey of Canada) of 58.40 m based on water level records from the Brittania Bay, Ontario measuring station (number 02KF005) located in the same drainage basin. Using a surveying level, each transect was divided into 5 cm height intervals based on the standard zero mark. Each transect was surveyed down to a height of -40 cm and up to, but not including, the edge of the tree canopy. Transects were not extended into the forested region to avoid the confonding affect of an animal runway. Because the edge of the tree canopy varied along the marsh, the upper limit of each transect corresponded to a height of from 30 cm to 15 cm above the standard zero mark.

Transects were censused in the order in which they had been randomly chosen and in each 5 cm height interval of all transects the presence or absence of all species was recorded. All species found in the section of the marsh actually surveyed were herbaceous emergent plants; no transects penetrated into the region of the marsh containing floating-leaved or submerged vegetation.

Voucher specimens for all species were collected and sent to the Systematics Division of Agriculture Canada for verification and deposited in their herbarium (DAO). Note that two species, *Lysimachia terrestris* and *Lysimachia thyrsiflora,* could not be separated in their vegetative state and were therefore grouped together as *Lysimachia* spp.

In order to determine the location of boundaries it had to be assumed (Pielou, 1977; Underwood, 1978; Keddy, 1983) that a gap in the distribution of a species along a transect did not represent a true disjunction. In reality, only species that were uncommon in the marsh showed such gaps and so it seems reasonable to accept such an assumption. If the highest height interval censused contained a species then the upper boundary for that species could not be determined, nor could the lower boundary be found if the species was present in the lowest height interval censused. Finally, if a species had both its upper and lower boundary in the same height interval it was excluded from the analysis to avoid a bias in part (ii) of the two hypotheses.

The procedure described above yields a contingency table consisting of the number of upper and lower boundaries found in each relative height interval of each transect. These data were analysed by an analysis of deviance (Karson, 1982; McCullogh & Nelder, 1983).

An analysis of deviance is analogous to an analysis of variance but is not restricted to assumptions of normality in the error structure of the model. To test for significant systematic components of the data, one looks at the difference between the deviance of the data from a model not incorporating the factor of interest and the deviance of the data from a model in which the factor is included. The deviance is measured using the log-likelihood ratio, or G^2, which is asymptotically distributed as a chi-squared variate. Since the contingency table is three-dimensional, all but the third order interaction term is tested using marginal totals [Sokal & Rohlf, 1981]. In this study, all marginal totals were larger than five and so the (null) assumption of a chi-squared distribution for G^2 is reasonable. See Austin *et al.* (1984) for a brief description of analysis of deviance and Generalized Linear Models in general.

Part (i) of the two hypotheses can be evaluated by testing for a main effect of relative height along the gradient; if this effect exists then there are significantly more boundaries in some height intervals than in others. The interaction term between the type of boundary (upper or lower) and relative height is used to evaluate part (ii) of the two hypotheses. If this interaction term exists then the number of upper and lower boundaries per height interval does not increase and decrease together.

Since we collected data using relative height intervals, the actual distance of marsh included within an interval may vary depending on the slope of the shore; for instance, in Breckenridge Marsh the height intervals had distances from a minimum of 0.8 m to a maximum of 19.1 m. If boundaries are randomly distributed along the relative height gradient but some intervals are correlated with longer distances, then a spurious clustering of boundaries would be found. To solve this problem the distance occupied by each height interval was included as a covariate and its effects removed.

Finally, it should be noted that only one marsh was studied and so no inferences should be made concerning between-marsh patterns.

The Genstat V statistical package was used and the data were analysed on the Agriculture Canada computer network.

Results

Both upper and lower boundaries are clustered ($p < 0.001$; Fig. 3). The individualistic hypothesis is therefore rejected. As well, there was a significant interaction between the type of boundary (upper or lower) and the main effect of relative height on the gradient, indicating that the pattern of clustering with respect to upper boundaries differed from the

Table 1. Analysis of deviance of the boundary distributions in Breckenridge Marsh, Quebec, Canada. The 'change of deviance' column tests the significance of each term in improving the goodness-of-fit in each successive model. **: $p < 0.001$; n.s. $p > 0.05$.

Terms	d.f.	residual deviance	d.f.	change of deviance
Constant	297	433.4	–	–
Interval length (covariate)	296	420.5	1	12.8**
Height interval	283	365.0	13	55.6**
Boundary type	282	364.0	1	1.0[n.s.]
Height interval × Boundary type	269	293.7	13	70.3**

pattern of clustering amongst lower boundaries along the gradient ($p < 0.001$, Table 1). In fact, there was no correlation between the number of upper and lower boundaries per relative height interval ($r = 0.009$, $p > 0.38$). We must therefore reject the community-unit hypothesis as well.

Although the covariate, the actual distance of marsh included in each interval, was significant (Table 1), the effect was to increase the importance of clustering along the gradient since when the covariate was excluded from the analysis, the change of deviance due to relative height was decreased from 55.6 to 43.5. This is because some of the longest relative height intervals contained very few boundaries. In any case, the amount of deviance explained by the covariate was small.

In order to better visualize these results, Fig. 4 shows the distribution of each species plotted against relative height in the marsh, averaged over the 13 transects. It can be seen that there exists one group of species (from *Lycopus uniflorus* to *Lysimachia nummularia*) which all reach their lower distributional limits in the interval from 10 to 20 cm. Equally important is the absence of any species whose distribution begins at this part of the gradient. In the interval from -10 to -25 cm there occurs another group of species (from *Iris versicolor* to *Carex vesicaria*) who all reach their lower distributional limits. Here there also exists a group of species simultaneously beginning their distribution (from *Sparganium eurycarpum* to *Penthorum*
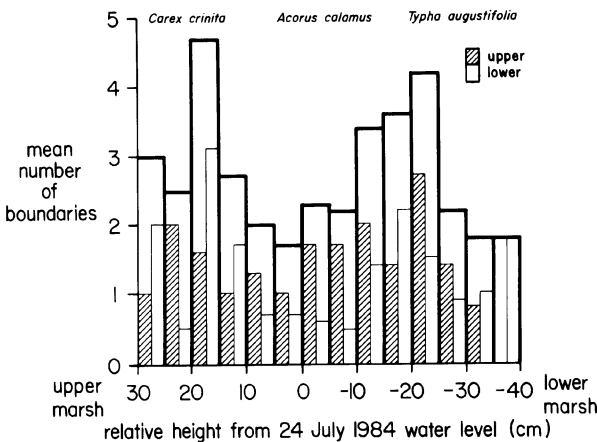
Fig. 3. The mean number of species boundaries in each 5 cm height interval plotted against relative height. Within each rectangle the mean number of upper boundaries (hatched) and lower boundaries (clear) are shown.

Lycopus uniflorus
Carex vulpinoidea
Potentilla anserina
Equisetum arvense
Aster simplex
Onoclea sensibilis
Plantago major
Carex projecta
Carex crinita
Lysimachia nummularia
Juncus filiformis
Iris versicolor
Scirpus americanus
Spartina pectinata
Eleocharis erythropoda Studel
Carex vesicaria
Eliocharis acicularis
Veronica scutellata
Acorus calamus
Alisma plantago-aquatica
Scirpus cyperinus
Sium sauve
Galium palustre
Scirpus acutus
Eleocharis smallii Britton
Sparganium eurycarpum
Carex retrorsa
Dulichium aurundinaceum
Sagittaria latifolia
Typha augustifolia
Rumex verticillatus
Polygonum hydropiperoides
Equisetum fluviatile
Scirpus fluviatilis
Penthorum sedoides
Glyceria borealis
Zinzania palustris
Pontedaria cordata
Lysimachia spp.
Agrostis stolonifera
Leersia orizoides
Polygonium coccineum
Lythrum salicaria

30   20   10   0   -10   -20   -30   -40

relative height from July 24 1984 water level (cm)

Fig. 4. The distribution of all plant species occurring in at least two transects, relative to the height gradient (upper marsh to the left, lower marsh to the right). These distributions are determined by calculating the average position of each species' upper and lower boundaries over the thirteen transects.

sedoides). It is this disconcordant pattern of upper and lower boundaries that causes the interaction term in Table 1.

In order to compare these results with those of other studies, we have computed Underwood's (1978) statistic. Using that test, lower boundaries are again found to be clustered ($p < 0.01$) but upper boundaries appear to be randomly distributed ($0.1 > p > 0.05$). Underwood's test is conservative in that it often fails to detect non-random patterns,

which explains the differences between the results of the two tests.

Discussion

In order to relate the results of this study to the debate concerning the individualistic and community-unit concepts it is necessary to consider the argument both from the perspective of pattern

and from the perspective of the mechanism (see Austin, 1985, 1987).

Hypotheses of pattern

The debate concerning pattern has usually been structured as a dichotomous choice between the predicted patterns of the community-unit and individualistic concepts (Curtis, 1955; Gleason, 1926; McIntosh, 1967; Whittaker, 1951; but see Whittaker 1975 for an exception). Such a dichotomy is misleading however, because it implies that evidence contradicting one expected pattern can be interpreted as evidence supporting the other. Such reasoning has been used to support the individualistic concept (Curtis, 1955; Whittaker, 1951, 1967). If we recognize that the two predicted patterns do not exhaust the number of possible patterns available (Whittaker, 1975) then one can no longer rely on negative evidence to support either of the two usual concepts of community organization along a gradient.

In this paper we have shown that the distribution of the species in Breckenridge Marsh are non-random, but do not correspond to the patterns predicted by the community-unit hypothesis. This demonstrates that the historical dichotomy is too limited, but might also be taken to suggest that Breckenridge Marsh is exceptional, since the results disagree with a large body of evidence apparently demonstrating individualistic distribution patterns (for example: McIntosh, 1975; Whittaker, 1962, 1975). Although this may be true, there are two reasons to believe that such a conclusion should be questioned.

First, the only other study in which both direct-gradient analysis and inferential statistics was used (Keddy, 1983) found boundary clustering in a lakeshore plant community. Three other studies have used inferential statistics to test boundary distributions, although the species ranges were not obtained using direct gradient analysis. Of these, Pielou & Routledge (1976) found clustering, Underwood (1978) found a random distribution, and Dale (1984) found a contiguity of upper and lower boundaries.

Second, the studies supporting the individualistic concept did not use inferential statistics to compare observed patterns with those predicted by the observed patterns with those predicted by the individualistic hypothesis and therefore may have failed to detect true non-random patterns. By subjectively scanning the distributions of many species along a gradient by eye, all but the most pronounced structure of such data would appear random, therefore biasing conclusions in favour of the individualistic concept.

Using pattern to deduce mechanism

In this paper we study the pattern of species' ranges relative to an obvious environmental gradient, but make no attempt to interpret our results as a test of any causal hypotheses. In the context of this debate however, pattern has generally been used to prove or disprove proposed mechanisms of community organization (Clements, 1916; Gleason, 1926; Whittaker, 1951, 1975). Certain assumptions are combined to produce a model of community organization, predicted patterns are derived as logical consequences of the model, and then the observed patterns of natural communities are compared with those predicted by the model. If the observed and predicted patterns are different then the underlaying mechanistic model is assumed to be wrong. If the observed and predicted patterns are similar then the underlaying mechanistic model is assumed to be correct (Whittaker, 1951). The problem with such an approach is that, given the large number of abiotic, biotic, and interactive factors to choose from, there are so many degrees of freedom that many different causal mechanisms could all 'explain' the same empirical pattern.

Consider our demonstration of boundary clustering in Breckenridge Marsh. This can be made consistent with the individualistic hypothesis by assuming that there are one or more abiotic factors which change rapidly at those points on the gradient where boundaries are clustered. There exists in Breckenridge Marsh, as in all natural communities, many abiotic factors which could possibly explain such a pattern: the degree of anoxia due

to flooding or to ice cover, the degree of moisture stress when the water level falls, changes in any one of many mineral nutrients or ratios of nutrients, disturbance due to ice damage or wave action in the spring, and so on. If it is possible to find a correlation of boundary clustering with a rapid change in any one of the large number of abiotic factors to choose from, should this correlation be interpreted as agreement with the individualistic concept? If some other abiotic factor is shown to change rapidly in a section of the marsh where there are relatively few boundaries, should we then conclude that the individualistic concept is still correct but that the factor in question plays no causal role? By arguing in such a way we render the individualistic concept unfalsifiable at the level of pattern analysis.

The problem is not solved by choosing strong environmental gradients such as the altitudinal gradient up mountainsides (Whittaker, 1951, 1956) or the gradient of relative water depth in a lake (Keddy, 1983) or in a marsh, in which many abiotic gradients are correlated to produce a gradient-complex. This is shown in Fig. 5 where the distribu-
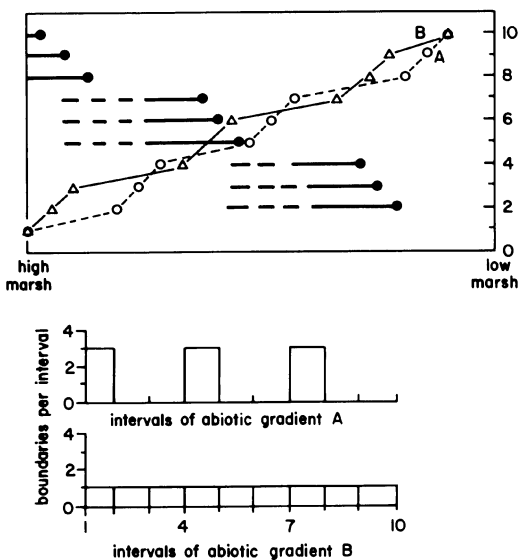
tions of various species are plotted along a hypothetical gradient, together with the values of two abiotic factors. If one were to do two different direct-gradient analyses of this 'gradient', then relative to factor A the boundaries would be clustered but relative to factor B the boundaries are regularly spaced. These two different results are obtained even though the two factors are positively correlated. Thus, the problem remains: the individualistic hypothesis is unfalsifiable when using pattern to deduce mechanism.

Hypotheses of mechanism

Whereas hypotheses regarding pattern can be tested using inferential statistics, hypotheses regarding mechanisms must be tested with experiments. Levins & Lewontin (1982) and Simberloff (1982b) restate the assumed mechanisms producing individualistic communities in an experimentally testable form. In Simberloff's words: 'species are individually distributed according to their interactions with their physical surroundings and ... only upon falsification of this hypothesis ought one to move on to more complex (physical environment × other-species interaction) ones'. Such a hypothesis has the advantage of not only being falsifiable but, contrary to Simberloff (1982a, b), has actually been falsified in many field experiments. These experiments test the assumed mechanism and not the predicted patterns which may arise as logical consequences from the mechanisms.

Sharitz & McCormick (1973) have demonstrated that the distribution of *Sedum smallii* on granite outcrops in the southeastern United States is controlled by a complex interaction of abiotic factors and competition with *Minuartia uniflora*. One boundary of *S. smallii* on these outcrops is determined by moisture stress and shallowness of the soil. The other end of its distribution is controlled through competitive exclusion from *M. uniflora*. Studies on the rocky intertidal zone (Connell, 1972; Lubchenco, 1980; Paine, 1984) have also demonstrated the importance of both biotic factors as well as interactions between biotic and abiotic factors in determining the structure of such communities.

Fig. 5. (a) Field distributions of nine species plotted in a hypothetical gradient along with the values of two correlated abiotic factors. (b) Distribution of species' boundaries obtained from direct-gradient analyses of this marsh relative to the two abiotic factors.

54

Grace & Whetzel (1981) give evidence that the distribution of *Typha angustifolia* is partly determined by competition with *Typha latifolia* in the ponds which they studied. Silander & Antonovics (1982), in a series of removal experiments, showed that the distributions of most of the plant species in a coastal plant community were profoundly affected by the types of neighbours that were removed. Snow & Vince (1984) and Vince & Snow (1984) used transplant experiments in an Alaskan salt marsh to demonstrate the interaction of biotic and abiotic factors in determining the structure of that community. Other papers reaching similar conclusions were reviewed by Connell (1983) and Schoener (1983). Finally, Turkington & Aarssen (1984 and references therein) present evidence that biotic interactions can produce coevolutionary differences between local genotypes, depending on the types of species which each interacts with, therefore biotic interactions may affect community structure on an evolutionary timescale as well.

There is therefore a large body of evidence which falsifies the causal assumptions of the individualistic concept and demonstrates integrated, that is causally dependent, responses of species within a community.

Hypotheses of pattern should be tested using inferential statistics. Hypotheses of mechanism should be tested by experimental manipulations, making sure not to equate correlation with causation (Levins & Lewontin, 1982). The way out of the community-unit vs. individualistic community debate is to deny the dichotomy and to consider multiple working hypotheses of community structure.

References

Austin, M. P., 1985. Continuum concept, ordination methods, and niche theory. Ann. Rev. Ecol. Syst. 16: 39–61.

Austin, M. P., 1987. Models for the analysis of species' response to environmental gradients. Vegetatio 69: 35–45.

Austin, M. P., Cunningham, R. B. & Flemming, P. M., 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio 55: 11–27.

Cain, S. A., 1947. Characteristics of natural areas and factors in their development. Ecol. Monogr. 17: 185–200.

Clements, E. E., 1916. Plant succession: an analysis of the development of vegetation. Publ. Carneg. Inst. 242.

Connell, H. H., 1972. Community interactions in marine rocky intertidal shores. Ann. Rev. Ecol. Syst. 3: 169–192.

Connell, H. H., 1983. On the prevalence and relative importance of interspecific competition: evidence from field experiments. Am. Nat. 122: 661–696.

Curtis, J. T., 1955. A prairie continuum in Wisconsin. Ecology 36: 558–566.

Dale, M. R. T., 1984. The contiguity of upslope and downslope boundaries in a zoned community. Oikos 43: 92–96.

Egler, F. E., 1947. Arid southwest Oahu vegetation, Hawaii. Ecol. Monogr. 17: 383–435.

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

Gleason, H. A., 1939. The individualistic concept of the plant association. Am. Midl. Na. 21: 92–110.

Grace, J. B. & Wetzel, R. G., 1981. Habitat partitioning and competitive displacement in cattails (*Typha*): experimental field studies. Am. Nat. 118: 463–474.

Hutchinson, G. E., 1975. A treatise on limnology, vol. 3. Limnological botany. Wiley, New York.

Karson, M. J., 1982. Multivariate statistical methods. Iowa State Univ. Press, Iowa.

Keddy, P. A., 1983. Shoreline vegetation in Axe Lake, Ontario: effects of exposure on zonation patterns. Ecology 64: 331–344.

Keddy, P. A., 1985. Wave disturbance on lakeshores and the within-lake distribution of Ontario's Atlantic coastal plain flora. Can. J. Bot. 63: 656–660.

Levins, R. & Lewontin, R., 1982. Dialectics and reductionism in ecology. In: E. Saarinen (ed.), Conceptual issues in ecology, pp. 107–138. Reidel, Dordrecht.

Lubchenco, J., 1980. Algal zonation in the New England rocky intertidal community: an experimental analysis. Ecology 61: 333–344.

Mason, H. L., 1947. Evolution of certain floristic associations in western North America. Ecol. Monogr. 17: 201–210.

McCullogh, P. & Nelder, J. A., 1983. Generalized linear models. Chapman and Hall, London.

McIntosh, R. P., 1967. The continuum concept of vegetation. Bot. Rev. 33: 130–187.

McIntosh, R. P., 1975. H. A. Gleason – 'individualistic ecologist' 1882–1975. His contributions to ecological theory. Bull. Torry Bot. Club 102: 253–273.

Paine, R. T., 1984. Ecological determinism in the competition for space. Ecology 65: 1339–1348.

Pielou, E. C. & Routledge, R. D., 1976. Salt marsh vegetation: latitudinal variation in the zonation patterns. Oecologia 24: 311–321.

Pielou, E. C., 1975. Ecological diversity. Wiley, New York.

Pielou, E. C., 1977. Mathematical ecology. Wiley, New York.

Schoener, T. W., 1983. Field experiments on interspecific competition. Am. Nat. 122: 240–285.

Sharitz, R. R. & McCormick, J. F., 1972. Population dynamics of two competing annual plant species. Ecology 54: 723–740.

Silander, J. A. & Antonovics, J., 1982. Analysis of interspecific interactions in a coastal plant community – a perturbation approach. Nature 298: 557–610.

Simberloff, D., 1982a. A succession of paradigms in ecology: essentialism to materialism and probabilism. In: E. Saarinen (ed.), Conceptual issues in ecology, pp. 139–153. Reidel, Dordrecht.

Simberloff, D., 1982b. Reply. In: E. Saarinen (ed.), Conceptual issues in ecology, pp. 139–153. Reidel, Dordrecht.

Snow, A. A. & Vince, S. W., 1984. Plant zonation in an Alaskan salt marsh I. Distribution, abundance and environmental factors. J. Ecol. 72: 651–667.

Sokal, R. R. & Rohlf, F. J., 1981. Biometry, 2nd ed. Freeman, San Francisco.

Spence, D. H. N., 1982. The zonation of plants in freshwater lakes. Adv. Ecol. Res. 12: 37–125.

Turkington, R. & Aarssen, L. W., 1984. Local-scale differentiation as a result of competitive interactions. In: R. Dirzo & J. Sarukhan (eds), Perspectives on plant population biology, pp. 107–127. Sinauer, Sunderland, MA.

Underwood, A. J., 1978. The detection of non-random patterns of distribution of species along an environmental gradient. Oecologia 36: 317–326.

Vince, S. W. & Snow, A. A., 1984. Plant zonation in an Alaskan salt marsh II. An experimental study of the role of edaphic conditions. J. Ecol. 72: 669–684.

Whittaker, R. H., 1951. A criticism of the plant association and climatic climax concepts. Northwest Sci. 25: 17–31.

Whittaker, R. H., 1956. Vegetation of the Great Smoky Mountains. Ecol. Monogr. 26: 1–80.

Whittaker, R. H., 1962. Classification of natural communities. Bot. Rev. 28: 1–160.

Whittaker, R. H., 1967. Gradient analysis of vegetation. Biol. Rev. 42: 207–264.

Whittaker, R. H., 1975. Communities and ecosystems, 2nd ed. Macmillan, New York.

Compositional dissimilarity as a robust measure of ecological distance

Daniel P. Faith, Peter R. Minchin & Lee Belbin*
CSIRO, Division of Water and Land Resources, G.P.O. Box 1666, Canberra, A.C.T. 2601, Australia

Abstract

The robustness of quantitative measures of compositional dissimilarity between sites was evaluated using extensive computer simulations of species' abundance patterns over one and two dimensional configurations of sample sites in ecological space. Robustness was equated with the strength, over a range of models, of the linear and monotonic (rank-order) relationship between the compositional dissimilarities and the corresponding Euclidean distances between sites measured in the ecological space. The range of models reflected different assumptions about species' response curve shape, sampling pattern of sites, noise level of the data, species' interactions, trends in total site abundance, and beta diversity of gradients.

The Kulczynski, Bray-Curtis and Relativized Manhattan measures were found to have not only a robust monotonic relationship with ecological distance, but also a robust linear (proportional) relationship until ecological distances became large. Less robust measures included Chord distance, Kendall's coefficient, Chi-squared distance, Manhattan distance, and Euclidean distance.

A new ordination method, hybrid multidimensional scaling (HMDS), is introduced that combines metric and nonmetric criteria, and so takes advantage of the particular properties of robust dissimilarity measures such as the Kulczynski measure.

Introduction

Many multivariate methods applied to community data operate, either explicitly or implicitly, on a matrix of compositional dissimilarities between samples. The degree of success in recovering ecological patterns in the data will depend on the nature and strength of the relationship between values of the chosen dissimilarity measure and the corresponding Euclidean distances between samples in ecological space ('ecological distances'). Equally important is the appropriateness of the assump-

tions about the form of this relationship which are implied by the mathematical model underlying the multivariate method.

In this study, we have used simulated vegetation data to examine directly the relationships between various coefficients of compositional dissimilarity and ecological distances. Of particular interest has been the relative *robustness* of coefficients to variations in the model of community variation in ecological space (cf. Minchin, 1987a). Attention was focused here on coefficients applicable to quantitative (abundance) data; the binary (presence-absence) case will be examined in a subsequent paper.

Our findings regarding the nature of the relationship between robust dissimilarity measures and ecological distance have consequences for the de-

sign of robust ordination methods. We propose a new method, hybrid multidimensional scaling (HMDS), that is based upon those aspects of compositional dissimilarity which are found to be robust predictors of ecological distance.

Background

While the importance of the choice of a dissimilarity measure in ecological ordination has been emphasized in previous work (Austin & Noy-Meir, 1971; Gauch, 1973; Austin, 1976; Fasham, 1977; Lamont & Grant, 1979; Prentice, 1977, 1980; Clymo, 1980; Beals, 1984), individual recommendations differ. Gauch (1973), Gauch & Whittaker (1972), and Beals (1973, 1984) argued that the Bray-Curtis measure (Bray & Curtis, 1957) is robust to non-linearities in species' response. The Relativized Manhattan measure (Renkonen, see Whittaker, 1952; Orlóci, 1974) is listed as an appropriate measure for non-linear data by Orlóci (1978). Austin & Greig-Smith (1968) found that a successive double standardization (species adjusted to equal maxima, then samples standardized to equal totals) is preferable to the use of the Bray-Curtis measure on unstandardized data. Orlóci (1974, 1978) has presented theoretical arguments in favour or Chord distance (sensu Orlóci, 1967), and Fasham (1977) found that a related measure, 'cosine theta' (Gower, 1967), out-performed several other measures in limited simulation trials with Gaussian models.

Swan (1970), Gauch (1973) and Ihm & Van Groenewoud (1975) used explicit models of community variation in ecological space and examined the relationship between compositional dissimilarity and ecological distance directly. However, their models assumed that species' responses are unimodal and symmetric (Gaussian). Evidence suggests that other types of responses (e.g. skewed, multimodal) are common (Austin, 1976, 1980, 1987; Minchin, 1987b), hence the generality of results from Gaussian models is questionable. Some other theoretical and simulation-based studies have used even more restrictive models. For example, the approach of Hajdu (1981) implies linear response curves.

Because species' response patterns in ecological space are unlikely to conform with any specific model (Greig-Smith, 1983; Austin 1976, 1985; Minchin, 1987b), the choice of a measure of compositional dissimilarity for ordination must consider the *robustness* of the measure's relationship with ecological distance over a range of species' response models. It follows that the development and evaluation of dissimilarity-based ordination methods must be guided by only those aspects of the relationship between dissimilarity and distance which are robust.

Ordination methods which have an in-built measure of compositional dissimilarity are prone to the problem of limited applicability arising from a restricted underlying model. The weaknesses of principal components analysis (PCA) can be interpreted in terms of the restrictive linear model implied by Euclidean distance (Noy-Meir & Austin, 1970). Correspondence Analysis (CA) can be formulated as a weighted principal co-ordinates analysis (PCoA: Gower, 1966), based upon Chi-squared distance (Chardy *et al.,* 1976). CA's behaviour will therefore depend on the relationship between Chi-squared distance and ecological distance. Surprisingly, Gauch *et al.* (1981) consider that the freedom to choose an appropriate dissimilarity measure for use in ordination is a *disadvantage*.

Ordination methods which permit a choice of dissimilarity measure may allow the combination of a scaling method that assumes a particular form of relationship between compositional dissimilarity and ecological distance, and a measure of dissimilarity that provides this relationship robustly. PCoA and so-called 'classical' multidimensional scaling (Torgerson, 1952) assume that the dissimilarity measure has a linear relationship with ecological distance. In contrast, nonmetric multidimensional scaling (NMDS, Shepard, 1962a, b; Kruskal, 1964a, b) assumes only *monotonicity*: a configuration is derived in which the distances between sample pairs are in *rank order* with their dissimilarities. Prentice (1977, 1980) argues that the weaker assumption of monotonicity should be preferred in view of our lack of knowledge about models of species response. Limitations of nonmetric methods have also been raised. The assumption of monotonicity may be so general that uninformative degeneracies in the solution are permitted (for a discussion, see Shepard, 1974). A balance must be reached between assuming too little, with a loss of useful information which can constrain the solution, and assuming too much, with a loss in the method's robustness. There do not appear to have been any tests of the robustness, over a wide range of response models, of the metric (linear) and nonmetric (monotonic) relationships of different dissimilarity measures to ecological distance. Such a study is a necessary first step in the solution to 'the

dissimilarity problem in ordination': what dissimilarity measure has a robust, informative, relationship with ecological distance; and what ordination method can take advantage of this relationship in its assumptions?

Methods

Dissimilarity coefficients evaluated

The dissimilarity measures evaluated in this study are listed in Table 1. Examination of two basic measures, Manhattan distance (MAN) and the complement of Kendall's coefficient (KEN), illustrate a problem which prompted consideration of many of the alternative measures in Table 1. MAN is defined by the sum of the absolute differences in abundance over all species. Such an index might be expected to reflect ecological distance, but a problem arises for larger ecological distances. When

two sites are sufficiently far apart in ecological space that they share no species, MAN yields a value that depends only on the total site abundances. Values of MAN can therefore suggest that one ecological distance is larger than another when the reverse is true. KEN initially appears to avoid spurious variation due simply to site totals, in that it reaches a constant maximum value when two sites share no species. However, for small ecological distances, when sites will tend to share many species, the actual value of KEN does reflect variation in site totals. Variation in site totals thus obscures predictive information about ecological distance for both simple measures.

A desirable measure would take a value of zero when ecological distance was zero, and some constant maximum value when ecological distance had increased to the point that shared abundance was zero (Beals, 1984). This theoretical argument has led us to focus on a number of measures that all have some form of standardization, such that the above properties are satisfied. These measures (Table 1) are the Kulczynski (QSK), Bray-Curtis (B-C), Chord distance (CHD), Canberra metric (CAN) and Relativized Manhattan (MAN:SAT) measures.

While the measures listed above are well constrained for large

Table 1. Basic measures with reference, abbreviation, standardizations used, and formula for dissimilarity between two objects, j and k, based upon attributes, $i = 1$ to N. Z is the number of attributes that are 0 for j and k. MAX_i is the maximum value of attribute i over all sites; MIN_i is the corresponding minimum. SPM is species adjusted to equal maximum abundance. SAT is sites standardized to equal totals. DBL is SPM followed by SAT. Equivalences of measure-standardization combinations reduced the total number of combinations to 29. For further explanation see text.

Name and reference	Abbreviation	Standardizations	Formula
Kendall (1970)	KEN	SPM	$\sum_i [MAX_i - \text{minimum } (X_{ij}, X_{ik})]$
Manhattan (Sokal & Michener, 1957)	MAN	SAT, SPM, DBL	$\sum_i |X_{ij} - X_{ik}|$
Gower metric (Gower, 1971)	GOW	SAT	$\sum_i [|X_{ij} - X_{ik}| / (MAX_i - MIN_i)]$
Euclidean (Sokal & Sneath, 1963)	EUC	SPS, SPM, SAT, DBL	$[\sum_i (X_{ij} - X_{ik})^2]^{1/2}$
Intermediate (Faith, 1984)	INT	SPM	$(\frac{1}{2}) \{ \sum_i |X_{ij} - X_{ik}| + MAX_i - \text{minimum } (X_{ij}, X_{ik})\}$
Quantitative symmetric (Kulczynski) — See for instance, Hajdu (1981)	QSK	SPM	$1 - (\frac{1}{2})\{[\sum_i \text{minimum } (X_{ij}, X_{ik})/\sum_i (X_{ij})] + [\sum_i \text{minimum } (X_{ij}, X_{ik})/\sum_i (X_{ik})]\}$
Bray-Curtis (Bray & Curtis, 1957)	B-C	SPM	$(\sum_i |X_{ij} - X_{ik}|)/[\sum_i (X_{ij} + X_{ik})]$
Chord (sensu Orlóci, 1967)	CHD	SPM	$[2 (1 - \{(\sum_i X_{ij} X_{ik})/[(\sum_i X_{ij}^2) (\sum_i X_{ik}^2)]^{1/2}\})]^{1/2}$
Canberra metric, Adkins form (Lance & Williams, 1967)	CAN	SPM, SAT, DBL	$[1/(N-Z)] \sum_i |X_{ij} - X_{ik}|/(X_{ij} + X_{ik})$
Chi-squared (Chardy *et al.,* 1976)	CSQ	SPM, SAT, DBL	$\{\sum_i (1/\sum_i X_{il})[X_{ij}/\sum_i X_{ij}) - X_{ik}/(\sum_i X_{ik})]^2\}^{1/2}$

and small ecological distances, they may differ in their predictivity of distance between these two extremes. The simulation study therefore included all these measures and, for comparison, KEN, MAN, EUC, the Gower metric (GOW) and two other measures INT and CSQ (Table 1). INT has been recommended as useful in ordination because of its intermediate properties relative to KEN and MAN (Faith, 1984). CSQ represents the implicit dissimilarity measure in the widely-used ordination method, correspondence analysis (CA).

Data standardizations

Because many dissimilarity coefficients have in-built standardizations, it is difficult to compare dissimilarity coefficients without also considering the effects of prior data standardizations. Consequently, we chose to study the ten coefficients listed above both on unstandardized data and in combination with each of three standardizations:
1. Species adjusted to equal maximum abundances (SPM).
2. Samples (sites) standardized to equal totals (SAT).
3. Bray-Curtis successive double standardizations: 1 followed by 2 (DBL).
    An additional standardization, of species to equal standard deviations (SPS), was only investigated in combination with Euclidean distance (EUC). This is the dissimilarity coefficient implied in an R-mode principal components analysis of a matrix of correlations among species. Subsequently, we refer to the combinations of standardizations and dissimilarity coefficients using the abbreviations given in Table 1.
    Some of the dissimilarity measures are mathematically equivalent after certain data standardizations. For example, after the SAT standardization the coefficients MAN, KEN, INT, B-C and QSK are indistinguishable. In addition, the in-built standardization of certain measures make them unaffected by some data standardizations. Thus GOW:SPM is equivalent to GOW and CHD:SAT is equivalent to CHD. Consideration of these equivalences resulted in a total of 29 unique combinations of standardization and dissimilarity coefficient (Table 1).

Models

We used a simulation approach similar to that of Minchin (1987a). The comprehensive simulation program, COENOS, is described in detail by Minchin (1987b). The model properties which were varied (Table 2) include length (beta diversity) of gradient, response curve shape, sampling pattern of the sites in ecological space, trends in sample totals, and degree of noise in the data. Limited computing resources prevented the examination of these properties in all factorial combinations and also restricted the amount of replication within each combination of properties. The resulting design was therefore not fully balanced. Table 3 shows the structure of the simulation experiment for two-dimensional models: the unidimensional experi-

Table 2. Model properties varied in the simulation study of quantitative dissimilarity measures. For a complete description of the simulation approach, see Minchin (1987b).

Properties varied in the generation of species sets
Beta diversity (R units of Minchin, 1987b)
1D models, 3 levels: 0.5, 1.0, 2.0
2D models, 6 levels: 0.5×0.25, 0.5×0.5, 1.0×0.5, 1.0×1.0, 2.0×1.0, 2.0×2.0

Ecological response shape. 3 levels:
1. Unimodal, symmetric. $\alpha = \gamma = 4.0$ on each gradient
2. Unimodal, skewed. Both α and γ uniformly distributed between 0.5 and 4.0
3. Interaction between species with unimodal, skewed responses, leading to some shouldered, bimodal and multimodal ecological responses
Interaction parameters (see Minchin, 1987b, for explanation):
    Modal abundance similarity interval = 0.75
    Mode proximity interval = 0.2
    1 to 3 interactors per species
    Interaction coefficients all equal to 0.15
    No negative interaction.

Properties varied in the generation of data matrices from species sets
Trend in sample totals. 4 levels:
1. No adjustments made to sample totals
2. Constant sample total over the ecological space
3. 3-fold linear trend across the ecological space
4. 3-fold unimodal trend, with maximum in the centre of the ecological space
Sampling pattern. 2 levels:
1. Regular grid in ecological space
2. Uniform random distribution over the ecological space
Quantitative noise. 3 levels:
1. No added noise
2. Normal errors with SD proportional to square root of original abundance. Constants of proportionality uniformly distributed between 0.25 and 0.75
3. Normal errors with SD proportional to square root of original abundance. Constants of proportionality uniformly distributed between 1.0 and 3.0

---

ment was similar. The model properties 'beta diversity' (6 levels) and 'response shape' (3 levels) were varied in all factorial combinations and three replicate species sets were produced in each of the 18 cells. The other model properties were varied during the generation of data matrices from these species sets, but not in all possible combinations. For example, all data matrices generated from the species sets with either 'symmetrical' or 'skewed' response shapes had 'regular' sampling patterns and 'no control' of sample totals. In total, the study was based on 255 data matrices with a single underlying gradient and 306 data matrices for which the ecological space had two dimensions.

Table 3. The design of the simulation experiment for two-dimensional models. Details of the model properties varied are given in Table 2.

Generation of species sets
Beta diversity (6 levels) × Response shape (3 levels) = 18 combinations
3 replicates per combination = 54 species sets

Generation of data matrices
1. Species sets with 'symmetric' or 'skewed' response shapes:
   Trend in sample total held at level 1 (no adjustment)
   Sampling pattern held at level 1 (regular grid)
   36 species sets
   × Quantitative noise (3 levels)
   = 108 data matrices

2. Species sets with 'interaction' response shapes:
   18 species sets
   × Trend in sample totals (3 levels: 2, 3 and 4)
   × Sampling pattern (2 levels, but level 2 only applied to the first replicate species set in each Beta diversity class and only with Quantitative noise at levels 1 and 3)
   × Quantitative noise (3 levels)
   = 198 data matrices

Analysis of robustness

For each data matrix, a matrix of ecological distances was computed as the Euclidean distances between samples in the simulated ecological space. In addition, compositional dissimilarity matrices were calculated from the simulated abundance data, using each of the 29 combinations of data standardization and dissimilarity coefficient listed in Table 1. Scatter plots of dissimilarity *versus* ecological distance were prepared and the strength of the linear and monotonic relationships between dissimilarity and ecological distance were assessed using Pearson product-moment correlation and Spearman rank correlation respectively. Most of the computations were performed using the Numerical Taxonomy Package (NTP; Belbin *et al.,* 1984).

The simulation results were analyzed in several ways:
1. Mean linear and rank correlations were computed for each measure, over all data matrices and within each level of each model property, then arranged in descending order.
2. Tables were produced showing the number of data matrices for which each measure achieved a difference in correlation compared with each other measure which was greater than a chosen threshold.
3. Multiway analyses of variance (ANOVA) were performed on either the correlation values for a chosen measure or the *difference* in correlation between a chosen pair of measures. Separate analyses were carried out within subsets of the simulation experiment for which the combinations of model properties were sufficiently balanced, in order to examine the effects of model properties and their interactions.

The first two approaches did not take account of the unbalanced design and the possible interactions between model properties, but they did serve to distinguish the consistently poor measures from the remainder. The relative performance of the better measures was then examined in more detail using ANOVA.

Results and discussion

Results are presented only for two-dimensional models: the results for the unidimensional models were similar. Fig. 1 shows an example of the relationships between dissimilarity and ecological distance for the coefficients MAN (Fig. 1a), KEN (Fig. 1b) and QSK (Fig. 1c). The data are from a two-dimensional model, with a beta diversity of $1 \times 1R$, skewed response surfaces, random sampling pattern and no noise. The general forms of the relationships are typical for these three measures. They clearly illustrate the contrast between QSK, with its in-built standardization by sample totals, and measures without any such standardization. QSK shows a roughly linear relationship with ecological distance from the minimum dissimilarity value of zero to the point where the dissimilarity reaches a maximum.

Table 4a shows the 29 dissimilarity measures listed in order of mean rank correlation over all 306 two-dimensional models. In Table 4b, they are ordered by mean linear correlation over only those 102 models for which the beta diversity of each gradient was no greater than 0.5R. At this level of beta diversity there are generally no sample pairs sufficiently far apart that they share no species.

In terms of *mean* rank correlation (Table 4a), the least successful measures were EUC, CSQ, INT and GOW, irrespective of prior data standardization, together with MAN and MAN:SPM. The measures with the highest mean rank correlations all included some form of standardization by species, either explicitly (SPM or DBL), or implicitly (variants of CAN). Excluding CAN, the best of the measures computed without prior data standardization were QSK and B-C. MAN:SAT performed similarly to these two measures and collectively QSK, B-C and MAN:SAT were more successful than CHD.

The measures with the lowest mean *linear* corre-
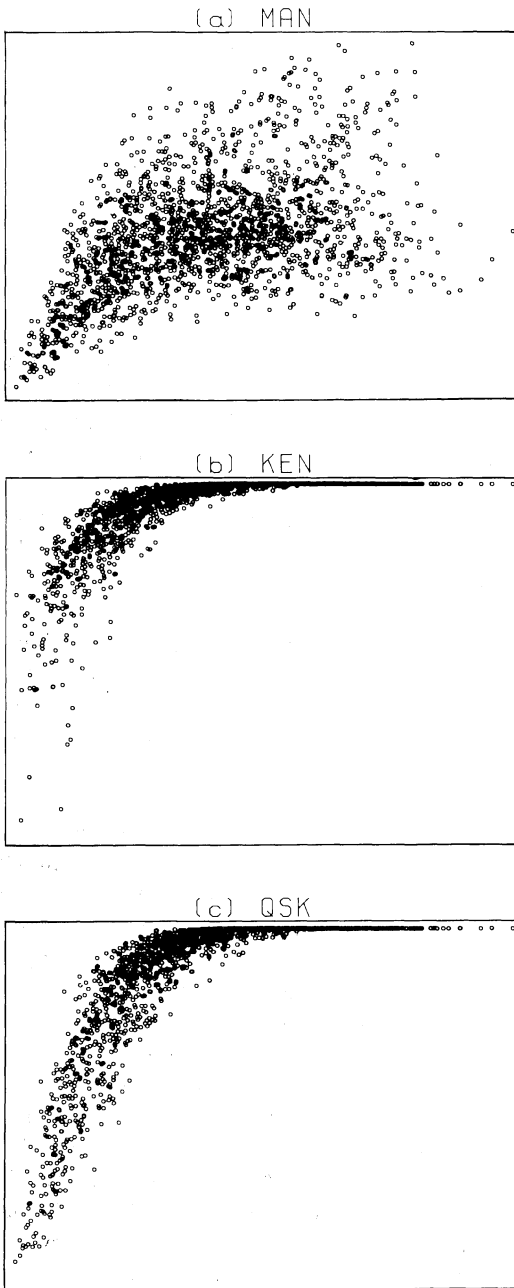
(a) MAN

(b) KEN

(c) QSK

Fig. 1. Relationship between compositional dissimilarity value (vertical axis) and 'target' ecological distance (horizontal axis) for three measures: (a) MAN; (b) KEN; and (c) QSK. Scales of axes are arbitrary. Each circle represents the dissimilarity value – distance value combination for a single pair of sites, for one simulated model. For further explanation see text.

Table 4. Dissimilarity measures ranked in order of (a) mean rank correlation with ecological distance over all 306 two-dimensional models and (b) mean linear correlation with ecological distance over the 102 two-dimensional models in which the beta diversity of the longest gradient did not exceed 0.5R. The abbreviations for the dissimilarity measures are explained in Table 1.

(a) Mean rank correlation		(b) Mean linear correlation	
1	QSK:SPM 0.8925	1	QSK:SPM 0.8722
2	B-C:SPM 0.8924	2	MAN:DBL 0.8684
3	MAN:DBL 0.8923	3	B-C:SPM 0.8682
4	CHD:SPM 0.8861	4	GOW:DBL 0.8491
5	CAN:SAT 0.8812	5	QSK 0.8325
6	CAN:DBL 0.8794	6	MAN:SAT 0.8321
7	CAN 0.8791	7	INT:SPM 0.8308
8	KEN:SPM 0.8637	8	CAN:SAT 0.8304
9	MAN:SAT 0.8551	9	CHD:SPM 0.8272
10	QSK 0.8550	10	CAN 0.8265
11	B-C 0.8548	11	CAN:DBL 0.8262
12	CHD 0.8158	12	B-C 0.8243
13	KEN 0.8157	13	GOW:SAT 0.8025
14	GOW:DBL 0.7487	14	KEN:SPM 0.7890
15	INT:SPM 0.7313	15	INT 0.7709
16	GOW:SAT 0.6917	16	EUC:DBL 0.7695
17	GOW 0.6523	17	CHD 0.7611
18	MAN:SPM 0.6522	18	GOW 0.7565
19	INT 0.6423	19	MAN:SPM 0.7564
20	CSQ:SAT 0.6325	20	CSQ:SAT 0.7539
21	EUC:SPM 0.6136	21	EUC:SAT 0.7442
22	CSQ 0.6087	22	KEN 0.7207
23	EUC:DBL 0.6046	23	EUC:SPM 0.7194
24	EUC:SAT 0.6022	24	CSQ 0.7099
25	CSQ:DBL 0.5926	25	MAN 0.6874
26	MAN 0.5617	26	CSQ:SPM 0.6636
27	CSQ:SPM 0.5605	27	CSQ:DBL 0.6625
28	EUC:SPS 0.5471	28	EUC:SPS 0.6424
29	EUC 0.4657	29	EUC 0.6079

lations (Table 4b) included all forms of EUC and CSQ, together with MAN, MAN:SPM, KEN, GOW, CHD and INT. Once again, the best measures included some type of standardization by species. Of the measures without species standardization, QSK and MAN:SAT had the highest mean linear correlations. Next best was B-C.

The relative performance of those measures which had the highest mean rank and linear correlations was examined in more detail using ANOVA. As an example, a summary of the analysis of the *difference* in rank correlation between QSK and CHD for models with symmetric and skewed response shapes is given in Table 5. In this case, the analysis indicates an interaction between beta

Table 5. Summary of the analysis of variance of the difference in rank correlation between QSK and CHD over the 108 two-dimensional models with 'symmetric' and 'skewed' response shapes.

Source of variation	DF	Mean square	Variance ratio	P
Between species sets				
Beta diversity (BD)	5	0.0108523	7.9	<0.001
Response shape (SH)	1	0.0013875	1.0	ns
BD.SH	5	0.0047089	3.4	0.01 – 0.05
Residual	24	0.0013760		
Total	35	0.0032062		
Within species sets				
Quantitative noise (NO)	2	0.0002449	1.9	ns
BD.NO	10	0.0001436	1.1	ns
SH.NO	2	0.0000371	0.3	ns
BD.SH.NO	10	0.0001039	0.8	ns
Residual	48	0.0001256		
Total	72	0.0001260		
Grand total	107	0.0011335		
Grand mean		0.0451		
Total number of observations		108		

diversity and response shape. Examination of the relevant mean values (Table 6) shows that:
1. QSK consistently achieved higher rank correlations than CHD (the difference is positive in all cells).
2. The difference in performance was most marked at the lowest levels of beta diversity, particularly for skewed response shapes.

The results and conclusions presented below are

Table 6. Means for beta diversity × response shape.

Beta diversity	Response shape	
	Symmetric	Skewed
0.5×0.25R	0.0592	0.1167
0.5×0.5R	0.0747	0.0425
1.0×0.5R	0.0288	0.0416
1.0×1.0R	0.0129	0.0393
2.0×1.0R	0.0241	0.0208
2.0×2.0R	0.0493	0.0311

Standard error of difference of means = 0.0175.
Least significant difference (0.05 level) = 0.0360.

based on similar analyses, although limited space does not permit the inclusion of the ANOVA summaries.

These analyses confirmed that QSK, MAN:SAT, and B-C were similar in both rank and linear correlations. However, for low beta diversity (0.5R or less) together with a unimodal trend in sample totals, we find greater linear and rank correlations for QSK and MAN:SAT compared with B-C. This result may indicate a greater robustness of these measures compared with B-C when site totals vary over short ecological distances.

The results clearly indicate the importance of standardizations, both implicit and explicit. Previous work on data standardization (e.g. Austin & Greig-Smith, 1968; Austin & Noy-Meir, 1971; Noy-Meir *et al.,* 1975; Orlóci, 1978) has been mainly concerned with the implications of standardization in the Euclidean 'species space' assumed by PCA and other techniques based on EUC. In contrast, we emphasize the effects of standardization on the relationship between compositional dissimilarity and ecological distance. For example, MAN:SAT

was one of the most robust and effective measures, whilst MAN and MAN:SPM were among the least successful. The most effective measures all included some form of standardization by sample total, which decreases their sensitivity to variations in sample total and causes such measures to take a constant maximal value for all sample pairs with no species in common.

In this study, standardization of species to equal maxima (SPM) also consistently improved both linear and rank correlations. This standardization tends to equalize the contributions of all species in the calculation of dissimilarities. Without the standardization, the dissimilarities reflect mainly those few species which have high modal abundances. However, we can not unequivocally recommend the use of SPM on field data. There is a danger that it may cause a deterioration in the predictive value of dissimilarities, by increasing the effective weight of sporadic or noisy species which never attain high abundances.

Conclusions

1. The three most effective and robust measures, in terms of both rank and linear correlation, are QSK, MAN:SAT and B-C. MAN:SAT and QSK are indistinguishable over our models, with slight improvement in performance for both over B-C. Because MAN:SAT is equivalent to QSK with the SAT standardization, we can recommend QSK as a basic robust measure, with the possible use of SAT and/or SPM standardizations. Further work with real and artificial data sets may distinguish among these options.
2. CHD is clearly less successful than QSK, MAN:SAT and B-C, in both rank and linear correlation. The use of the B-C measure in ordination has been critized (Orlóci, 1974, 1978) and the CHD distance recommended as an alternative (Orlóci, 1978). The greater robustness of B-C demonstrated here suggests that there has been undue emphasis upon CHD's metric properties and its absence of a changing scale of measure over different pairs of sites. The rela-

tively weak rank correlation for CHD also implies that the cosine theta (Gower, 1967) and the Geodesic measure (Orlóci, 1978) will also have a poor rank correlation, as they are monotonic with CHD.
3. CAN, with its in-built standardization by species, achieved higher rank correlations than QSK, MAN:SAT and B-C when the latter were applied without prior species' standardization. However, QSK:SPM, MAN:DBL and B-C:SPM consistently gave higher rank correlations than CAN. Furthermore, QSK and MAN:SAT had higher linear correlations than CAN at low beta diversities.
4. Of the measures compared, the least successful, in terms of both rank and linear correlation, were all forms of EUC, KEN, GOW and CSQ, together with MAN (when applied without the SAT standardization).

Implications for ordination

Ordination techniques such as principal components analysis (PCA), principal co-ordinates analysis (PCoA) and correspondence analysis (CA) produce configurations in which the distances between sample pairs are *proportional* to their compositional dissimilarities. Consequently, they would only be expected to recover the configuration of samples in ecological space if values of the dissimilarity coefficient were proportional to ecological distances. In PCA, the implied dissimilarity measure is some form of EUC (or possibly CHD), while CA is based on the CSQ measure. The relatively low linear correlations of EUC and CSQ with ecological distance, observed in this study, explain the poor performance of these techniques in simulation studies (e.g. Gauch & Whittaker, 1972; Gauch *et al.,* 1977; Fasham, 1977; Minchin, 1987a), especially at higher beta diversities.

None of the dissimilarity measures examined in this study has a linear relationship with ecological distance when beta diversity is sufficiently high that some samples have no species in common. However, QSK and MAN:SAT have reasonable linear correlations with ecological distance when beta

diversity is low (0.5R or less). These results imply that, when beta diversity is low, better recovery of ecological space can be achieved by applying a linear ordination technique (e.g. PCoA, metric multidimensional scaling) to QSK or MAN:SAT, rather than using PCA or CA.

The overall *rank* correlation between dissimilarity and ecological distance, as assessed in this study, should provide a basis upon which to predict the relative performance of dissimilarity measures in nonmetric multidimensional scaling (NMDS). The results suggest that QSK, MAN:SAT and B-C should give the most robust and effective ordinations, particularly for 'global' NMDS (Kruskal, 1964a, b; Fasham, 1977), which assumes a single monotone relationship over all sample pairs. Another variant, 'local' NMDS, fits a separate monotone regression for each row of the dissimilarity matrix (Sibson, 1972; Prentice, 1980; Minchin, 1987a). This study did not examine the local or row-wise rank correlations between dissimilarity and ecological distance, hence no firm evidence is available upon which to recommend the most appropriate measures for local NMDS.

We have commented separately on the implications of the linear and rank correlation results for ordination methods. On the one hand, QSK has a robust linear relationship with small ecological distance. On the other hand, QSK shows a robust overall monotonic relationship with distance. We now present a method that takes advantage of both of these aspects of QSK's relationship with ecological distance. This provides a solution to the 'dissimilarity problem' in ordination: the search for a combination of a method that assumes a particular relationship between dissimilarity and distance, and a measure that provides this relationship robustly.

We propose the following procedure, called hybrid multidimensional scaling (HMDS):

1. Compute a matrix of robust dissimilarities, using QSK (or perhaps MAN:SAT or B-C).
2. Make two copies of this matrix. In the first, set all dissimilarities as missing that are above a threshold value. Dissimilarities below this value are taken to have a linear relationship with distance. In the second matrix, keep all values.

3. Use these matrices as input to a multidimensional scaling program such as KYST (Kruskal *et al.,* unpubl. rep.). The first matrix, with missing values, is treated with the metric option (a *linear* regression of distance on dissimilarity) and the second with the non-metric option (a *monotonic* regression). These options can be used in either the 'global' or 'local' forms. The stress, or 'badness-of-fit' of any trial configuration is a combination of the stress contributions from the two matrices.

In practice, a starting configuration may be found using an initial run of NMDS. As another practical consideration, the choice of a fixed threshold may be replaced by a gradual downweighting of the linear constraint with increasing dissimilarity. We currently use a fixed threshold with HMDS, equal to a value of 0.8 with the measure QSK.

A combination of metric and nonmetric models in multidimensional scaling has previously been suggested by psychometricians as a strategy to avoid degeneracies found in NMDS solutions (see Shepard, 1974). The metric criterion is usually a polynomial fit and the nonmetric criterion of monotonicity is used as a constraint on the form of the polynomial. Similar uses of constraints with NMDS are described in Shepard (1974), and Borg & Lingoes (1980). HMDS differs from these approaches in using a prescribed dissimilarity measure that has a robust metric (linear) relationship with distance only over a certain range. The monotonic regression serves as the only direct constraint on the larger dissimilarities. Critical to the concept of HMDS is that a nonmetric analysis is combined with a metric analysis, but with the values above the threshold value in this latter matrix viewed as misleading and therefore set as missing.

An example of HMDS is illustrated in Fig. 2, for the same complex two dimensional model used for Fig. 1. A plot of the target ecological distances (horizontal axis) versus the ordination distances (vertical axis) shows a close linear relationship (Fig. 2a). Fig. 2b shows the plot of input QSK dissimilarities (vertical axis) against the distances in the ordination (horizontal axis). This may be compared with Fig. 1c, which plots the QSK dissimilari-
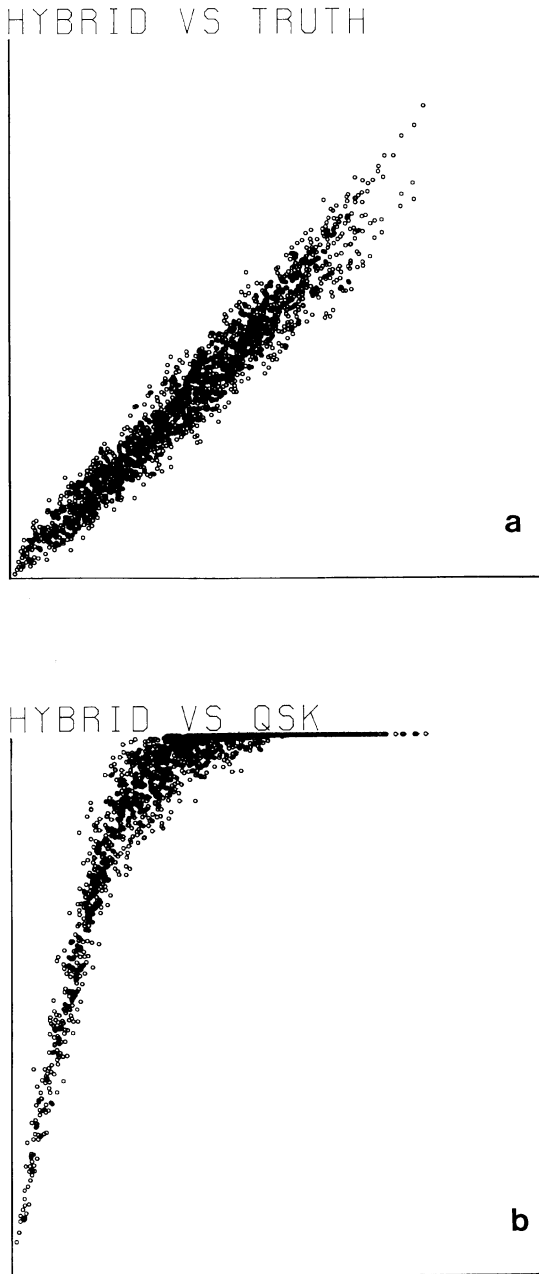
HYBRID VS TRUTH

a

HYBRID VS QSK

b

Fig. 2. (a) Plot of Euclidean distances between sites derived from HMDS configuration (vertical axis) versus ecological distances (horizontal axis); (b) Relationship between QSK dissimilarities used as input to HMDS (vertical axis) and the Euclidean distances from the HMDS configuration (horizontal axis). Simulated model used is that of Fig. 1.

ties against the target ecological distances. The metric contraint was enforced up to a threshold QSK dissimilarity value of 0.8, and the configuration distances have a strong linear fit to the dissimilarities up to this threshold. Above this threshold a monotonic relationship between dissimilarity and distance (Fig. 2b) is recovered similar to that for the target (Fig. 1c).

Careful comparisons based on computer simulations are needed to compare the robustness of HMDS with alternative methods. It is clear that meaningful comparisons can only be achieved if appropriate dissimilarity measures are used. In an extensive simulation study to be reported elsewhere (Minchin, Faith & Belbin, unpubl.), we demonstrate the relative robustness of HMDS compared with NMDS (both local and global), metric MDS, and network re-estimation (Williamson, 1978; Beals, 1984), even when all methods use the QSK measure. We hope in a later study to compare HMDS with another method, Parsimony Analysis (Faith et al., 1985) which is in theory highly robust to variations in the model of species' response.

In conclusion, we note that the robust dissimilarity measures identified by this study should be useful in connection with other multivariate methods, (e.g. cluster analysis and network methods) when applied to compositional data with the aim of recovering patterns in ecological space.

References

Anderberg, M. R., 1973. Cluster analysis for applications. Academic Press, New York.

Austin, M. P., 1976. Performance of four ordination techniques assuming three different non-linear species response models. Vegetatio 33: 43–49.

Austin, M. P., 1980. Searching for a model for use in vegetation analysis. Vegetatio 42: 11–21.

Austin, M. P., 1985. Continuum concept, ordination methods, and niche theory. Ann. Rev. Ecol. Syst. 16: 39–61.

Austin, M. P., 1987. Models for the analysis of species' response to environmental gradients. Vegetatio 69: 35–45.

Austin, M. P. & Greig-Smith, P., 1968. The application of quantitative methods to vegetation survey. II. Some methodological problems of data from rain forest. J. Ecol. 56: 827–844.

Austin, M. P. & Noy Meir, I., 1971. The problem of non-

linearity in ordination. Experiments with two-gradient models. J. Ecol. 59: 763–773.

Beals, E. W., 1973. Ordination: mathematical elegance and ecological naivete. J.Ecol. 61: 23–35.

Beals, E. W., 1984. Bray Curtis ordination: an effective strategy for analysis of multivariate ecological data. Adv. Ecol. Res. 14: 1–55.

Belbin, L., Faith, D. P. & Minchin, P. R., 1984. Some algorithms contained in the numerical taxonomy package NTP. CSIRO Division of Water and Land Resources, Canberra Technical Memorandum 84/23.

Borg, I. & Lingoes, J. C., 1980. A model and algorithm for multidimensional scaling with external constraints on the distances. Psychometrika 45: 25–38.

Bray, J. R. & Curtis, J. T., 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27: 325–349.

Chardy, P., Glemarc, M. & Laurec, A., 1976. Application of inertia methods to benthic marine ecology: practical implications of the basic options. Estuarine Coastal Mar. Sci. 4: 179–205.

Clymo, R. S., 1980. Preliminary survey of the peat-bog Hummell Knowe Moss using various numerical methods. Vegetatio 42: 129–148.

Faith, D. P., 1984. Patterns of sensitivity of association measures in numerical taxonomy. Math. Biosci. 69: 199–207.

Faith, D. P., Minchin, P. R. & Belbin, L., 1985. Parsimony and falsification in ecology: toward an assumption-free approach to the study of species' response to gradients. Stud. Plant Ecol. 16: 31–32.

Fasham, M. J. R., 1977. A comparison of non-metric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coenoclines and coenoplanes. Ecology 58: 551–561.

Gauch Jr, H. G., 1973. The relationship between sample similarity and ecological distance. Ecology 54: 618–622.

Gauch Jr, H. G. & Whittaker, R. H., 1972. Comparison or ordination techniques. Ecology 53: 868–875.

Gauch Jr, H. G., Whittaker, R. H. & Wentworth, T. R., 1977. A comparative study of reciprocal averaging and other ordination techniques. J. Ecol. 65: 157–174.

Gauch, H. G., Whittaker, R. H. & Singer, S. B., 1981. A comparative study of non-metric ordinations. J. Ecol. 69: 135–152.

Gower, J. C., 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325–338.

Gower, J. C., 1967. Multivariate analysis and multidimensional geometry. The Statistician 17: 13–28.

Gower, J. C., 1971. A general coefficient of similarity and some of its properties. Biometrics 23: 623–637.

Greig-Smith, P., 1983. Quantitative plant ecology. 3rd ed. Blackwell, Oxford.

Hajdu, L. J., 1981. Graphical comparison of resemblance measures in phytosociology. Vegetatio 48: 47–59.

Ihm, P. & Van Groenewoud, H., 1975. A multivariate ordering of vegetation data based on Gaussian type gradient response curves. J. Ecol. 63: 767–777.

Kendall, D. G., 1970. A mathematical approach to seriation. Philos. Trans. R. Soc. London A. 269: 125–135.

Kruskal, J. B., 1964a. Multidimensional scaling by optimizing goodness-of-fit to a non-metric hypothesis. Psychometrika 29: 1–27.

Kruskal, J. B., 1964b. Non-metric multidimensional scaling: A numerical method. Psychometrika 29: 115–129.

Lamont, B. B. & Grant, K. J., 1979. A comparison of twenty-one measures of site dissimilarity. In: L. Orlóci, C. R. Rao & W. M. Stiteler (eds), Multivariate methods in ecological work, pp. 101–126. International Co-operative Publishing House, Fairland, Maryland.

Lance, G. N. & Williams, W. T., 1967. Mixed data classificatory programs. I. Agglomerative systems. Aust. Comput. J. 1: 15–20.

Minchin, P. R., 1987a. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 69: 89–107.

Minchin, P. R., 1987b. Simulation of multidimensional community patterns: towards a comprehensive model. Vegetatio (in press).

Noy Meir, I. & Austin, M. P., 1970. Principal component ordination and simulated vegetational data. Ecology 51: 551–552.

Noy Meir, I., Walker, D. & Williams, W. T., 1975. Data transformations in ecological ordination. II. On the meaning of data standardization. J. Ecol. 63: 779–800.

Orlóci, L., 1967. An agglomerative method for classification of plant communities. J. Ecol. 55: 193–206.

Orlóci, L., 1974. Revisions for the Bray and Curtis ordination. Can. J. Bot. 52: 1773–1776.

Orlóci, L., 1978. Multivariate analysis in vegetation research. 2nd ed. Junk, The Hague.

Orlóci, L., 1980. An algorithm for predictive ordination. Vegetatio 42: 23–25.

Prentice, I. C., 1977. Non-metric ordination methods in ecology. J. Ecol. 65: 85–94.

Prentice, I. C., 1980. Vegetation analysis and order invariant gradient models. Vegetatio 42: 27–34.

Shepard, R. N., 1962a. Analysis of proximities: Multidimensional scaling with an unknown distance function. I. Psychometrika 27: 125–140.

Shepard, R. N., 1962b. The analysis of proximities: multidimensional scaling with an unknown distance function. II. Psychometrika 27: 219–246.

Shepard, R. N., 1974. Representation of structure in similarity data – problems and prospects. Psychometrika 39: 373–421.

Sibson, R., 1972. Order invariant methods for data analysis. J. R. Statist. Soc. B 34: 311–349.

Sokal, R. R. & Michener, C. D., 1957. The effects of different numerical techniques on the phenetic classification of bees of the *Hoplitis* complex (Megachilidae). Proc. Linn. Soc. London 178: 59–74.

Sokal, R. R. & Sneath, P. H. A., 1963. Principles of numerical taxonomy. Witt. Freeman and Co., San Francisco.

Swan, J. M. A., 1970. An examination of some ordination problems by use of simulated vegetational data. Ecology 51: 89–102.

Torgerson, W. S., 1952. Multidimensional scaling: I. Theory and method. Psychometrika 17: 401–419.

Whittaker, R. H., 1952. A study of summer foliage insect communities in the Great Smoky Mountains. Ecol. Monogr. 22: 1–44.

Williamson, M. H., 1978. The ordination of incidence data. Ecology 66: 911–920.

The analysis of vegetation-environment relationships by canonical correspondence analysis*

Cajo J. F. Ter Braak[1,2]**
[1]*TNO Institute of Applied Computer Science, Statistics Department Wageningen, P.O. Box 100, 6700 AC Wageningen, The Netherlands, and [2]Research Institute for Nature Management, P.O. Box 46, 3956 ZR Leersum, The Netherlands*

Abstract

Canonical correspondence analysis (CCA) is introduced as a multivariate extension of *weighted averaging ordination,* which is a simple method for arranging species along environmental variables. CCA constructs those linear combinations of environmental variables, along which the distributions of the species are maximally separated. The eigenvalues produced by CCA measure this separation.

As its name suggests, CCA is also a correspondence analysis technique, but one in which the ordination axes are constrained to be linear combinations of environmental variables. The ordination diagram generated by CCA visualizes not only a pattern of community variation (as in standard ordination) but also the main features of the distributions of species along the environmental variables. Applications demonstrate that CCA can be used both for detecting species-environment relations, and for investigating specific questions about the response of species to environmental variables. Questions in community ecology that have typically been studied by 'indirect' gradient analysis (i.e. ordination followed by external interpretation of the axes) can now be answered more directly by CCA.

Introduction

Direct gradient analysis relates species presence or abundance to environmental variables on the basis of species and environment data from the same set of sample plots (Gauch, 1982). The simplest methods of direct gradient analysis involve plotting each species' abundance values against values of an environmental variable, or drawing isopleths for each species in a space of two environmental variables (Whittaker, 1967). With these simple methods one can easily visualize the relation between many

species and one or two environmental variables.

Plant species experience the conditions provided by many environmental variables; therefore one might wish to analyse their joint effects. Multiple regression can be used for that purpose. However, despite some successful applications, e.g., Yarranton (1970), Austin (1971) and Forsythe & Loucks (1972), ordinary multiple regression has never become popular in vegetation science. Reasons for this include: (1) Each species requires separate analysis, so regression analysis may require an unreasonable amount of effort. (2) Vegetation data are often qualitative, or when they are quantitative the data contain many zero values for the plots at which a species is absent. In neither case do the data satisfy the assumption of a normal error distribution that is implicit in ordinary multiple regression. (3) Relationships between species and environmental variables are generally non-linear. Species abundance is often a single-peaked (bell-

shaped) function of the environmental variables. (4) Environmental variables are often highly correlated, and so it can be impossible to separate their independent effects. Generalized Linear Modelling (Austin *et al.,* 1984; Ter Braak & Looman, 1986) provides a solution for (2) and (3), but (1) and (4) remain. Whenever the number of influential environmental variables is greater than two or three, it becomes difficult to put results for several species together so as to obtain an overall graphical summary of species-environment relationships.

A simple method is therefore needed to analyze and visualize the relationships between many species and many environmental variables. Canonical correspondence analysis (CCA) is designed to fulfil this need. CCA is an eigenvector ordination technique that also produces a multivariate direct gradient analysis (Ter Braak, 1986). CCA aims to visualize (1) a pattern of community variation, as in standard ordination, and also (2) the main features of species' distributions along the environmental variables.

Ter Braak (1986) derived CCA as a heuristic approximation to the statistically more rigorous (but computationally fraught) technique of Gaussian canonical ordination, and also showed CCA's relation to correspondence analysis (CA), alias reciprocal averaging (Hill, 1973). In this paper a simple, alternative derivation of CCA is given starting from the method of weighted averaging (WA).

Theory

From weighted averaging to canonical correspondence analysis

Figure 1a shows an artificial example of single-peaked response curves for four species along an environmental variable (e.g. moisture). Species A occurs in drier conditions than species D. Fig. 1a shows presence-absence data for species D: the species is present at four of the sites.

How well does moisture explain the species' data? The fit could be formally measured by the deviance between the data and the curves, as in logistic regression (Ter Braak & Looman, 1986), but this idea will not be pursued here. Instead, a simple alternative based on the method of weighted averaging (WA) is used.
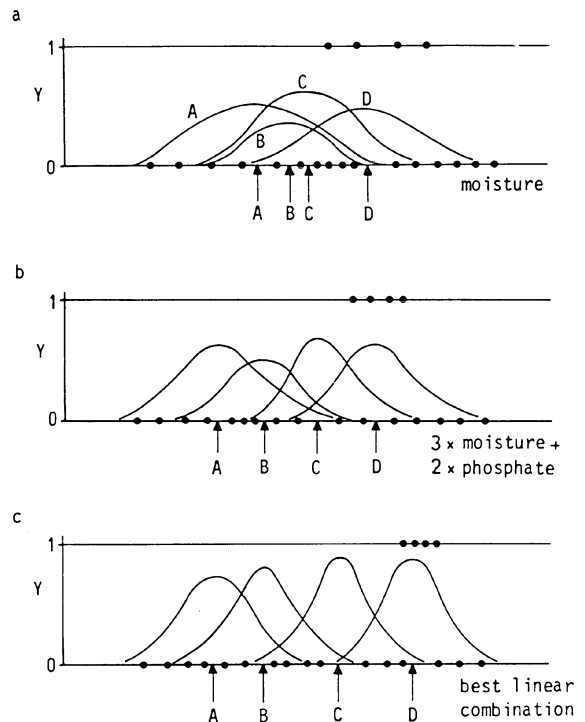
Fig. 1. Artificial example of single peaked response curves of four species (A–D) with respect to standardized environmental variables showing different degrees of separation of the species curves: (a) moisture; (b) a linear combination of moisture and phosphate, chosen apriori; (c) the best linear combination of environmental variables, chosen by CCA. Sites are shown by dots at $y = 1$ if species D is present and at $y = 0$ if species D is absent.

For each species a score can be calculated by taking the weighted average of the moisture values of the plots. For abundance data, this score is calculated as

$$u_k = \sum_{i=1}^{n} y_{ik}x_i/y_{+k} \qquad (1)$$

where u_k is the weighted average of the k-th (out of m) species, x_i is the (moisture) value of the i-th (out of n) site and y_{ik} is the abundance of species k at site i, and y_{+k} is the total abundance of species k. For presence-absence data the weighted average is simply the average of the moisture values of the plots in which the species is present. The weighted average

gives a first indication of where the species occurs along the moisture gradient (see the arrows in Fig. 1a). As a measure of how well moisture explains the species data, the *dispersion of the weighted averages* is used (see below). If the dispersion is large, moisture neatly separates the species curves, and moisture explains the species data well. If the dispersion is small, then moisture explains less.

To compare the explanatory power of different environmental variables, each environmental variable must first be standardized to mean 0 and variance 1. For technical reasons, weighted means and variances are used; each environmental variable is standardized such that

$$\sum_{i=1}^{n} y_{i+}x_i = 0 \quad \text{and} \quad \sum_{i=1}^{n} y_{i+}x_i^2/y_{++} = 1 \quad (2)$$

where y_{i+} is the total abundance at site i and y_{++} the overall total. The dispersion can now be written as

$$\delta = \sum_{k=1}^{m} y_{+k}u_k^2/y_{++} \quad (3)$$

By calculating the dispersion for each environmental variable one can select the 'best' variable.

Now suppose that moisture is the 'best' single variable in the artificial example. However, someone might suggest a better variable, that is a combination of two others (see, e.g., Loucks, 1962). In the artificial example a combination of moisture and phosphate, namely (3 × moisture + 2 × phosphate), is shown to give a larger dispersion than moisture alone (Fig. 1b); and consequently the curves in Fig. 1b are narrower, and the presences of species D are closer together, than in Fig. 1a. So it can be worthwhile to consider not only the environmental variables separately but also all possible linear combinations of them, i.e. all 'weighted sums' of the form

$$x_i = b_1z_{i1} + b_2z_{i2} + \ldots + b_pz_{ip} \quad (4)$$

where z_{ij} is the value of the j-th (out of p) environmental variable at site i, and b_j is the weight (not necessarily positive) belonging to that variable; x_i is the value of a compound environmental variable at site i. (It is assumed in equation (4) that each en-

vironmental variable is centered to a weighted mean of 0. Although not essential, it will also be convenient to standardize the environmental variables according to equation (2) so as to make the weights (b_j) comparable.)

CCA turns out to be *the technique that selects the linear combination of environmental variables that maximizes the dispersion of the species scores.* In other words, CCA chooses the optimal weights (b_j) for the environmental variables. In the Appendix it is shown that these optimal weights are the solution of the same eigenvalue equation as the one derived by another rationale in Ter Braak (1986), and that the first eigenvalue of CCA is actually equal to the (maximized) dispersion of species scores along the first CCA axis.

The second and further CCA axes also select linear combinations of environmental variables that maximize the dispersion of the species scores, but subject to the constraint of being uncorrelated with previous CCA axes. In principle, as many axes can be extracted as there are environmental variables.

From correspondence analysis to canonical correspondence analysis

CA also maximizes the dispersion δ in equation (3). But it does so irrespective of any environmental variable; that is, CA assigns scores (x_i) to sites such that the dispersion is absolutely maximum, the scores being standardized as in equation (2) (Nishisato, 1980). CCA is therefore 'restricted correspondence analysis' in the sense that the site scores are restricted to be linear combinations of supplied environmental variables.

A familiar algorithm to carry out CA is the reciprocal averaging algorithm (Hill, 1973). In Ter Braak (1986) this algorithm is extended with an additional multiple regression step so as to obtain the CCA solution. In each iteration cycle the trial site scores are regressed on the environmental variables (using y_{i+}/y_{++} as site weights) and the new trial scores are the fitted values of this regression. The FORTRAN program CANOCO (Ter Braak, 1985b) to carry out CCA is in fact just an extension of Hill's (1979) program DECORANA.*

CCA is restricted correspondence analysis, but the restrictions become less strict, the more environmental variables are included in the analysis. If $p \geq n-1$, then there are actually no restrictions any more; CCA is then simply CA. The arch effect may therefore crop up in CCA as it does in CA (Gauch, 1982). The method of detrending (Hill & Gauch, 1980) can be used to remove the arch and is available in the computer program

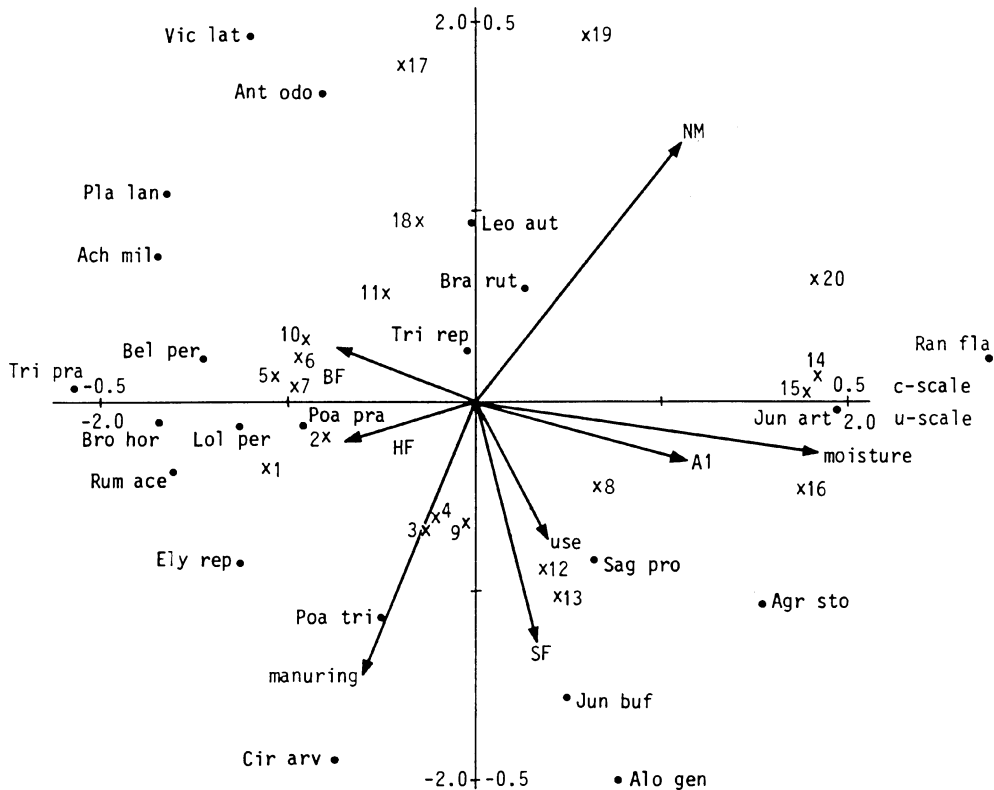*The program is available from the author at cost price.

72

Fig. 2. Dune meadow data: CCA ordination diagram with relevés (x), plant species (●) and environmental variables (arrow); first axis horizontally, second axis vertically. For relevé numbers see Table 1. Abbreviations are given as *underlining* in full names in Table 1. The c-scale applies to the environmental arrows, the u-scale to species and sites points. Eight infrequent species are not shown because they lie outside the range of this diagram.

CANOCO (Ter Braak, 1985b). But in CCA the arch can be removed more elegantly by dropping superfluous environmental variables. Variables that are highly correlated with the 'arched' axis (often the second axis) are most likely to be superfluous.

CA is very susceptible to species-poor sites containing rare species in that it places such aberrant sites (and the rare species occurring there) at extreme ends of the first ordination axes (Gauch, 1982), relegating the major vegetation trends in the data to later axes. CCA does not show this 'fault' of CA, provided the sites that are aberrant in species composition are not so aberrant in terms of the environmental variables.

Ordination diagram

The ordination diagram of CCA displays sites,

species and environmental variables (Fig. 2). The site and species points have the same interpretation as in CA. They display variation in species composition over the sites. The environmental variables are represented by arrows (Fig. 2). Loosely speaking, the arrow for an environmental variable points in the direction of maximum change of that environmental variable across the diagram, and its length is proportional to the rate of change in this direction. Environmental variables with long arrows are more strongly correlated with the ordination axes than those with short arrows, and so more closely related to the pattern of community variation shown in the ordination diagram.

Further insight into the ordination diagram of CCA can be obtained from yet another characterization of CCA. From equations (A.5) en (A.6) of the Appendix it follows that CCA is a
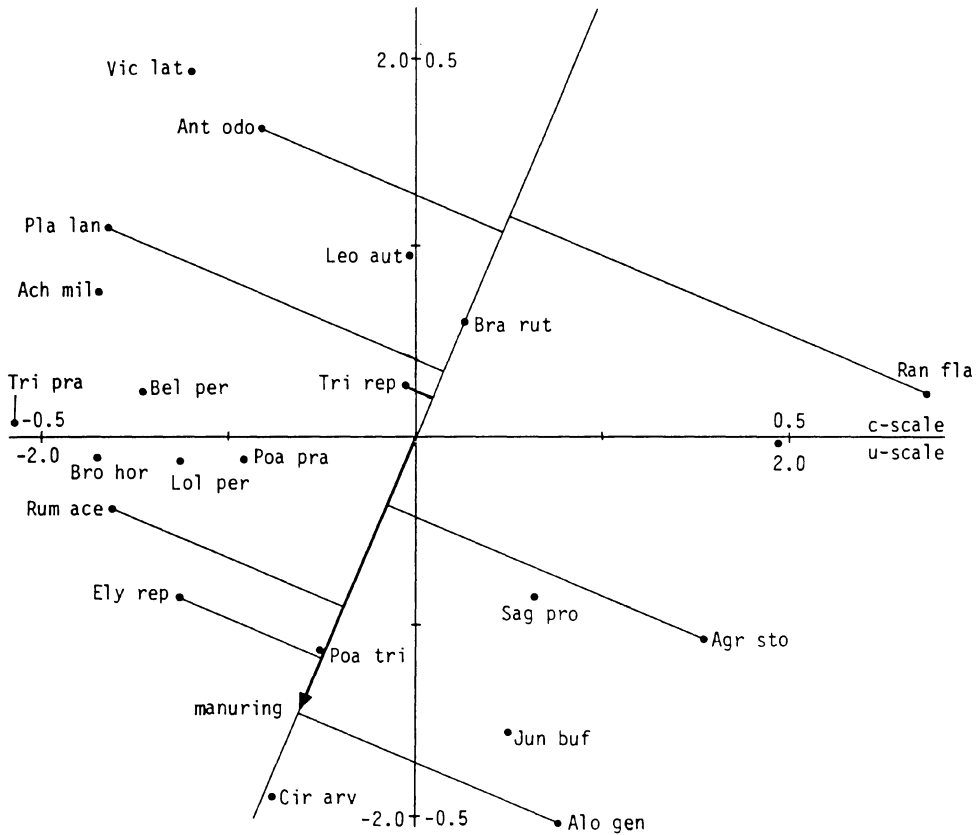
Fig. 3. Inferred ranking of the species along the variable quantity of manuring based on the biplot interpretation of Fig. 2. For explanation see the Ordination diagram section.

weighted principal components analysis applied to a matrix of species by environmental variables, the (k, j)-th element of which is the weighted average of species k with respect to environmental variable j (it is here assumed that each environmental variable is reduced to zero mean). CCA is a weighted analysis in the sense that species are given weights proportional to their total abundance (y_{+k}) and the environmental variables are weighted inversely with their covariance matrix. The intuitive advantage of the implicit species weights is that a weighted average for a species is imprecise when its total is low (Ter Braak & Looman, 1986) and is thus not worth much attention. Environmental variables are given equal weight irrespective of their variance or unit of measurement. (This type of weighting is also implicit in discriminant analysis (see Campbell & Atchley, 1981) and makes the analysis invariant to nonsingular linear transformations of the environmental variables). This characterization of CCA shows that the joint plot of species and environmental variables in the CCA ordination diagram can be interpreted similarly to a principal components biplot (Gabriel, 1971; Ter

Braak, 1983), allowing inference of the approximate values of the weighted averages of each of the species with respect to each of the environmental variables.

The most convenient rule for quantitative interpretation of the CCA biplot (Ter Braak, 1986) is therefore as follows: each arrow representing an environmental variable determines a direction or 'axis' in the diagram; the species points can be projected on to this axis (see Fig. 3). The order of the projection points corresponds approximately to the ranking of the weighted averages of the species with respect to that environmental variable. The weighted average indicates the position of a species' distribution along an environmental variable (Fig. 1), and thus the projection point of a species also indicates this position, although approximately.

Table 1. Dune meadow data: data table with species (rows) and relevés (columns of one digit width) arranged in order of their scores on the first axis of CCA. Relevé numbers are printed vertically. The abundance values, as used in the analysis, are on a 1 – 9 scale to replace the Braun-Blanquet symbols r, +, 1, 2m, 2a, 2b, 3, 4, 5. Thickness of the A1 horizon is divided into ten equal-sized classes (denoted 0 – 9). The values 1, 2 and 3 for agricultural use refer to hayfield, haypasture and pasture, respectively. For further explanation of the environmental variables see text.

	relevés
	1 111 11 11112
	51670217834923894560
*Tri*folium *pra*tense	2-52----------------
*Achi*llea *mi*llefolium	212243-2------------
*Bromus hor*deaceus	2--244----3---------
*Pl*antago *lan*ceolata	5-553-323-----------
*Rumex ace*tosa	5-63-------22-------
*Bel*lis *pere*nnis	2---23--222---------
*Ely*mus *repe*ns	44---4---446 -------
*Lo*lium *per*enne	2766657-2652--4-----
*Vicia lat*hyroides	----1-2-1-----------
*Poa pra*tensis	243444413544--24-----
*Ant*hoxanthum *odor*atum	4-324--4-------4----
*Cirsium arv*ense	----------2---------
*Poa tri*vialis	624547---655494---2-
*Tri*folium *repe*ns	2-52653-2213322261--
*Leon*todon *aut*umnalis	3--3335525222223622-2
*Brachy*thecium *ruta*bulum	2-622-4-62224-23-444
*Juncus buf*onius	---2-------443------
*Sagina pro*cumbens	------2---524223----
*Al*opecurus *gen*iculatus	-----2---723855----4-
*Hy*pochaeris *rad*icata	------22-------5----
*Aira pra*ecox	-------2------3----
*Sal*ix *repe*ns	--------2------3---5
*Agrostis sto*lonifera	---------483454-4475
*Juncus art*iculatus	----------4--4--334
*Cheno*podium *alb*um	------------1------
*Emp*etrum *nig*rum	--------------2----
*Ranunculus fla*mmula	-------------22-2224
*Eleocharis pal*ustris	-------------4-4584
*Cal*liergonella *cus*pidata	----------------4-33
*Pot*entilla *pal*ustris	----------------22--
thickness *A1*	40100001211133117930
moisture	11112112122445555555
quantity of *manuring*	24231210044123311131
agricultural *use*	12231231122122313231
*S*tandard *F*arming	01000000011011000010
*B*io-dynamic *F*arming	00001110000000000000
*H*obby *F*arming	10110000000100100000
*N*ature *M*anagement	00000001100000011101

The ordination diagrams of CCA and CA also share some of the shortcomings of WA (Ter Braak & Looman, 1986). The most important practical shortcoming is that species that are unrelated to the

ordination axes tend to be placed in the center of the ordination diagram and are not distinguished from species that have true optima there. This problem can easily be circumvented by looking at a species-by-site data table in which species and sites are arranged in order of their scores on one of the ordination axes (cf. Table 1).

The CCA ordination diagram is not in any way hampered by high correlations between species, or between environmental variables.

Applications

Exploratory use of the ordination diagram

Batterink and Wijffels (report) studied the possible relation between vegetation and management of dune meadows on the island Terschelling (The Netherlands).

A subset of their data is analysed here to illustrate the ordination diagram of CCA. This subset consists of 20 standard plots recorded in 1982, and 30 plant species (Table 1).

Five environmental variables were recorded: (1) thickness of the A1 horizon, measured in millimeters; (2) moisture content of the soil, scored on a five-point scale in a semi-objective manner; (3) quantity of manuring, scored on a five-point scale on the basis of a questionnaire sent to the owners of the meadows; (4) agricultural use, a nominal variable with three classes − hayfield, haypasture and pasture; and (5) type of management, a nominal variable with four classes − standard farming, biodynamic farming, hobby farming and nature management.

CCA cannot directly cope with ordinal variables, like moisture and manuring here. Ordinal variables must either be treated as if they were quantitative, or as nominal variables. Here they were treated as quantitative. Nominal variables, like type of management, must be transformed to dummy variables as shown in Table 1. For instance, the dummy variable 'nature management' indicates which meadows received that type of management. Agricultural use was however treated as a quantitative variable (Table 1), because haypasture was considered as an intermediate between hayfield and pasture.

Two values were missing in the environment data. CCA cannot cope with missing values, so relevés with missing values in the environment data must be deleted. To avoid deletion, missing values were replaced here by the mean of the corresponding variable over the remaining plots.

Despite the crude measurement of the environmental variables, they nicely explain the major variation in the vegetation. The first two eigenvalues of CCA (λ_1 = 0.46 and λ_2 = 0.29) were not much reduced in comparison with those of standard CA (0.54 and 0.40), and the two-dimensional configurations of species and sites in the ordination diagrams

looked similar. The most conspicuous difference was that relevés 17 and 19 were outliers in CA and not so much in CCA (Fig. 2).

The configurations of species and sites in CCA (Fig. 2) must be interpreted as in CA (Ter Braak, 1985a). For instance, from Fig. 2 *Sagina procumbens* can be expected to have its maximum abundance in the relevés close to its point in Fig. 2 (relevés 8, 12 and 13) and to be absent in relevés far from that point.

Figure 2 accounts for 65% of the variance in the weighted averages of the species with respect to each of the environmental variables. This percentage is calculated as in principal components analysis by taking $100 \times (\lambda_1 + \lambda_2)/(\lambda_1 + \ldots + \lambda_p)$. It can be deduced from Fig. 2, for example, that *Cirsium arvense, Alopecurus geniculatus* and *Elymus repens* mainly occur in the highly manured meadows, *Agrostis stolonifera* and *Trifolium repens* in intermediately manured meadows, and *Ranunculus flammula* and *Anthoxanthum odoratum* in little manured meadows (see Fig. 3). The other arrows can be interpreted similarly. From Fig. 2 it can thus be seen at once which species occur mainly under wetter conditions (those on the right hand side of the diagram) and which ones prefer drier conditions (those on the left hand side of the diagram).

Multi-species trend surface analysis

CCA can be used to detect spatial gradients in vegetation data. A spatial gradient can be specified by a linear combination of two orthogonal coordinates, say, the x-coordinate (z_1) and y-coordinate (z_2) of the relevés, i.e. by $b_1z_1 + b_2z_2$. The strongest spatial gradient in vegetation data might be defined as that combination of z_1 and z_2 that maximally separates the spatial distributions of the species, and can thus be estimated by taking the x- and y-coordinates as environmental variables in a CCA. Put another way, CCA searches for the direction of the strongest vegetation zonation (cf. Fig. 1).

Such an analysis was applied to counts of 13 arable weeds in summer barley in May 1983 in 96 plots (0.5 ×0.5 m) in the experimental field 'Doeksen' (50 m × 100 m) (B. Post, unpubl).

The first CCA axis was defined by b_1 = 0.0261 and b_2 = 0.0117, so that the gradient was estimated to make $\tan^{-1}(b_2/b_1)$ = 24° with the x-coordinate axis. Further, the first eigenvalue was six times the second eigenvalue, which indicated that the

gradient was a clear one. But, judged on the basis of the value of the first eigenvalue (λ_1 = 0.09), the amount of species turnover was quite small (cf. Gauch & Stone, 1979).

To verify the supposition that the gradient was related to moisture, percentage moisture was measured in the top soil (0–3 cm) in March 1985 (B. Post, unpubl). The strongest gradient in these moisture values had an angle of 34° with the x-coordinate axis and thus pointed approximately in the same direction as the gradient estimated by CCA from the 1983 weed data.

Vegetation succession

An example of application in a succession study on a rising sea-shore is found elsewhere in this volume (Cramer & Hytteborn, 1987). One of their questions was whether the vegetation succession tracks the land uplift (ca. 0.5 cm per year) or whether it lags behind.

This question was approached with detrended CCA with elevation and year as the 'environmental variables', through fitting the compound gradient x = b_1 × elevation + b_2 × year. The resulting weights were b_1 = 0.054 and b_2 = 0.041. Consequently, the equivalent change in vegetation per year is b_2/b_1 = 0.76 cm.

An approximate 95%-confidence interval for the change ranges from 0.4 cm to 1.1 cm and clearly includes the known land rise of ca 0.5 cm per year. The confidence interval was obtained from the standard errors of b_1 and b_2 in the final regression within the reciprocal averaging algorithm of CCA by using Fieller's theorem (see Finney, 1964, p. 27–29). The interval is presumably a little too short as it ignores that the CCA-axis is chosen optimally.

Discussion

CCA considerably extends the analytical power of ecological ordination. Questions like those tackled in the applications section above could formerly only be investigated by 'indirect gradient analysis', i.e. first extracting the ordination axes from the species data and subsequently interpreting the major axes in relation to environmental data – e.g. by regression analysis (Dargie, 1984), trend surface analysis (Gittins, 1968) or canonical correlation analysis (Carleton, 1984). Such two-step analyses ignore the minor axes of variation in community composition; yet 'minor' aspects of the variation

may still be substantial, especially in large data sets, and in some problems may be just the variation that one is actually interested in because of its relationship to particular external variables (see Jolliffe, 1982).

CCA works because species tend to have single-peaked response functions to environmental variables. When the response functions are simpler (e.g. approximately linear), the results can still be expected to be adequate in a qualitative sense, but it might then be advantageous to utilize instead the linear counterpart of CCA — redundancy analysis (Israëls, 1984). The weed data are a case in point. Because the number of species is quite small in that example, and the number of absences is small as well, these data could also be analysed from the beginning by canonical correlation analysis (Gittins, 1985). But canonical correlation analysis and redundancy analysis fail, when species do show single-peaked response functions (Gauch & Wentworth, 1976), i.e. in the case where CCA works best.

Appendix

Maximizing δ in Eq. (3) leads to CCA (Ter Braak, 1986) and CCA is a weighted principal components analysis applied to a matrix of weighted averages.

Let $\mathbf{Y} = \{y_{ik}\}$ and $\mathbf{Z} = \{z_{ij}\}$ be $n \times m$ and $n \times p$ matrices containing the species data and environmental data, respectively, and let $\mathbf{R} = \mathrm{diag}(y_{1+}, y_{2+}, \ldots, y_{n+})$. Each environmental variable is centered to a weighted mean of 0, i.e. $\mathbf{Z}'\mathbf{R}\mathbf{1}_n = \mathbf{0}$, where $\mathbf{1}_n$ is an n-vector containing 1's. Further, let $\mathbf{S}_{11} = \mathrm{diag}(y_{+1}, y_{+2}, \ldots, y_{+m})$, $\mathbf{S}_{12} = \mathbf{Y}'\mathbf{Z}$, $\mathbf{S}_{21} = \mathbf{Z}'\mathbf{Y}$, $\mathbf{S}_{22} = \mathbf{Z}'\mathbf{R}\mathbf{Z}$ and let $\mathbf{u}$ and $\mathbf{b}$ be vectors of order m and p, containing the species scores u_k and the weights b_j, respectively.

By inserting Eq. (4) in Eq. (1) we obtain

$$\mathbf{u} = \mathbf{S}_{11}^{-1}\mathbf{Y}'\mathbf{Z}\mathbf{b} = \mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{b} \qquad (A.1)$$

Hence,

$$\delta = y_{++}^{-1}\mathbf{u}'\mathbf{S}_{11}\mathbf{u} = y_{++}^{-1}\mathbf{b}'\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{b} \qquad (A.2)$$

which must be maximized with respect to $\mathbf{b}$, subject to Eq. (2). By inserting Eq. (4) in Eq. (2), we obtain $\mathbf{b}'\mathbf{Z}'\mathbf{R}\mathbf{1}_n = 0$, which is satisfied trivially because of the centering of Z, and

$$y_{++}^{-1}\mathbf{b}'\mathbf{S}_{22}\mathbf{b} = 1 \qquad (A.3)$$

The solution of this maximization problem is known to be the first eigenvector of the eigenvalue equation

$$(\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12} - \lambda\mathbf{S}_{22})\,\mathbf{b} = \mathbf{0} \qquad (A.4)$$

with $\delta = \lambda$ (see, for instance, Mardia *et al.*, 1979, theorem A.9.2). Eq. (A.4) is the centered version of Eq. (A5) in Ter Braak (1986). The latter equation has a trivial solution ($\lambda = 1$, $\mathbf{x} = \mathbf{1}_n$) and its nontrivial solutions satisfy Eq. (A.4) and Eq. (2). Therefore, maximizing δ leads to the first axis of CCA as defined in Ter Braak (1986). Further, maximizing δ subject to the constraint that the second axis is uncorrelated with the first axis (using weights y_{i+}, as in Eq. (2)) leads to the second eigenvector of (A.4), which is therefore identical to the second axis of CCA as defined in Ter Braak (1986), and so on for subsequent axes.

Let $\mathbf{W}$ be a $m \times p$ matrix containing the weighted averages of the species with respect to the environmental variables, i.e.

$$\mathbf{W} = \mathbf{S}_{11}^{-1}\mathbf{Y}'\mathbf{Z} \qquad (A.5)$$

The weighted principal components analysis of $\mathbf{W}$ described in the main text follows from the singular value decomposition

$$\mathbf{S}_{11}^{\frac{1}{2}}\,\mathbf{W}\,\mathbf{S}_{22}^{-\frac{1}{2}} = \mathbf{S}_{11}^{-\frac{1}{2}}\,\mathbf{S}_{12}\,\mathbf{S}_{22}^{-\frac{1}{2}} = \mathbf{P}\,\mathbf{\Lambda}^{\frac{1}{2}}\,\mathbf{Q}' \qquad (A.6)$$

where $\mathbf{P}$ and $\mathbf{Q}$ are orthonormal $m \times p$ and $p \times p$ matrices and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$. For convenience of notation it is assumed here that $p \leq m$. This singular value decomposition is just another way to solve (A.4) (see Mardia *et al.*, 1979, chapter 10). The coordinates of species k in the ordination diagram are given by the k-th row of the matrix

$$\mathbf{U} = y_{++}^{\frac{1}{2}}\,\mathbf{S}_{11}^{-\frac{1}{2}}\,\mathbf{P}(\mathbf{I} - \mathbf{\Lambda})^{-\frac{1}{2}}, \qquad (A.7)$$

and the coordinates of environmental variable j by the j-th row of the matrix

$$\mathbf{B}_e = y_{++}^{-\frac{1}{2}}\,\mathbf{S}_{22}^{\frac{1}{2}}\,\mathbf{Q}\,\mathbf{\Lambda}^{\frac{1}{2}}\,(\mathbf{I} - \mathbf{\Lambda})^{\frac{1}{2}} \qquad (A.8)$$

The pre- and post-multiplication factors involving y_{++} and $(\mathbf{I} - \mathbf{\Lambda})$ in Eqs. (A.7) and (A.8) are not essential for the biplot; they are included to obtain the scaling used in DECORANA (Hill, 1979, section 4,5). In Hill's scaling the coordinates of the sites are weighted averages of the species coordinates and the (weighted) variance of the coordinates of species present at a site is equal to 1 on average. Hill's scaling is used in Fig. 2.

References

Austin, M. P., 1971. Role of regression analysis in plant ecology. Proc. Ecol. Soc. Austr. 6: 63–75.

Austin, M. P., Cunningham, R. B. & Fleming, P. M., 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio 55: 11–27.

Campbell, N. A. & Atchley, W. R., 1981. The geometry of canonical variate analysis. Syst. Zool. 30: 268–280.

Carleton, T. J., 1984. Residual ordination analysis: a method for exploring vegetation-environment relationships. Ecology 65: 469–477.

Cramer, W. & Hytteborn, H., 1987. The separation of fluctuation and long-term change in the vegetation dynamics of a rising sea-shore. Vegetatio 69: 157–167.

Dargie, T. C. D., 1984. On the integrated interpretation of indirect site ordinations: a case study using semi-arid vegetation in south-eastern Spain. Vegetatio 55: 37–55.

Finney, D. J., 1964. Statistical methods in biological assay. Griffin, London, 668 pp.

Forsythe, W. L. & Loucks, O. L., 1972. A transformation for species response to habitat factors. Ecology 53: 1112–1119.

Gabriel, K. R., 1971. The biplot graphic display of matrices with application to principal component analysis. Biometrika 58: 453–467.

Gauch, H. G., 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge.

Gauch, H. G. & Stone, E. L., 1979. Vegetation and soil pattern in a mesophytic forest at Ithaca, New York. Am. Midl. Nat. 102: 332–345.

Gauch, H. G. & Wentworth, T. R., 1976. Canonical correlation analysis as an ordination technique. Vegetatio 33: 17–22.

Gittins, R., 1968. Trend-surface analysis of ecological data. J. Ecol. 56: 845–869.

Gittins, R., 1985. Canonical analysis. A review with applications in ecology. Springer Verlag, Berlin.

Hill, M. O., 1973. Reciprocal averaging: an eigenvector method of ordination. J. Ecol. 61: 237–249.

Hill, M. O., 1979. DECORANA – A FORTRAN program for detrended correspondence analysis and reciprocal averaging. Ecology and Systematics, Cornell University, Ithaca, New York.

Hill, M. O. & Gauch, H. G., 1980. Detrended correspondence analysis, an improved ordination technique. Vegetatio 42: 47–58.

Israëls, A. Z., 1984. Redundancy analysis for qualitative variables. Psychometrika 49: 331–346.

Jolliffe, I. T., 1982. A note on the use of principal components in regression. Appl. Statist. 31: 300–303.

Loucks, O. L., 1962. Ordinating forest communities by means of environmental scalars and phytosociological indices. Ecol. Monogr. 32: 137–166.

Mardia, K. V., Kent, J. T. & Bibby, J. M., 1979. Multivariate analysis. Academic Press, London.

Nishisato, S., 1980. Analysis of categorical data: dual scaling and its applications. University of Toronto Press, Toronto.

Ter Braak, C. J. F., 1983. Principal components biplots and alpha and beta diversity. Ecology 64: 454–462.

Ter Braak, C. J. F., 1985a. Correspondence analysis of incidence and abundance data: properties in terms of a unimodal response model. Biometrics 41: 859–873.

Ter Braak, C. J. F., 1985b. CANOCO – A FORTRAN program for canonical correspondence analysis and detrended correspondence analysis. IWIS-TNO, Wageningen, The Netherlands.

Ter Braak, C. J. F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167–1179.

Ter Braak, C. J. F. & Looman, C. W. N., 1986. Weighted averaging, logistic regression and the Gaussian response model. Vegetatio 65: 3–11.

Whittaker, R. H., 1967. Gradient analysis of vegetation. Biol. Rev. 42: 207–264.

Yarranton, G. A., 1970. Towards a mathematical model of limestone pavement vegetation. III. Estimation of the determinants of species frequencies. Can. J. Bot. 48: 1387–1404.

Ecological amplitudes of plant species and the internal consistency of Ellenberg's indicator values for moisture*

Cajo J. F. Ter Braak[1] & Niek J. M. Gremmen[2]**
[1]TNO Institute of Applied Computer Science, P.O. Box 100, 6700 AC Wageningen, The Netherlands
[2]Research Institute for Nature Management, P.O. Box 46, 3956 ZR Leersum, The Netherlands

Abstract

Two methods for estimating ecological amplitudes of species with respect to Ellenberg's moisture scale are discussed, one based on weighted averaging and the other on maximum likelihood. Both methods are applied to phytosociological data from the province of Noord-Brabant (The Netherlands), and estimate the range of occurrence of species to be about 4–6 units on the moisture scale. Due to the implicit nature of Ellenberg's definition of moisture, it is impossible to improve the indicator values in a statistically sound way on the basis of floristic data only. The internal consistency of the Ellenberg indicator values is checked by using Gaussian logit regression. For 45 out of the 240 species studied the indicator value is inconsistent with those of the other species. The same method is used to estimate the optima and amplitudes of species considered moisture-indifferent and of some species not mentioned by Ellenberg. Some of these 'indifferent' species show a remarkably narrow amplitude.

It is concluded that the Ellenberg indicator values for moisture form a reasonably consistent system.

Introduction

Ellenberg (1979) summarized the ecology of the Central-European vascular plants, by assigning to each species indicator values for light, temperature, moisture, nitrogen and acidity.

Ellenberg's indicator values are used to estimate the value of any of these environmental factors at a particular site by averaging the indicator values for this factor of all species present (e.g. Ellenberg, 1979, 1983; Persson, 1981; Smeets, Werger & Tevonderen, 1980; Böcker, Kowarik & Bornkamm, 1983). Plants often reflect temporally integrated environmental conditions and are therefore particularly useful indicators when values averaged over

time are needed. When the value of an environmental factor in the past is required, the only possible approach may be to base it on historical vegetation data.

During the development of a model simulating the effects of withdrawal of groundwater on the disappearance of plant species (Gremmen et al., 1985; Reijnen & Wiertz, 1984), we wished to know:

(1) do Ellenberg's indicator values for moisture and nitrogen correctly represent the optima of species for these factors in our study area,

(2) what is the ecological amplitude of each species for these factors, including species not mentioned by Ellenberg (1979)?

We will only discuss moisture values here. Clear-

ly, the same reasoning can be applied for other factors.

Ellenberg (1979) placed each species on a 12-point ordinal scale according to its distribution with respect to moisture (Table 1). It is not clear which characteristic(s) of the moisture regime (e.g. groundwater level, soil moisture content, and soil moisture deficit) were used in the definition of these classes. In practice the indicator values of Ellenberg's 'intuitive' scale seem to work well, however.

The implicit nature of Ellenberg's definition of moisture makes it impossible to check the correctness of the indicator values against actual measurements. Nevertheless, it is possible to check the internal consistency by comparing the indicator values of species that occur together: when a species mainly occurs together with species with higher (lower) indicator values, its indicator value is in comparison with those of the other species too low (too high). (When species have extreme indicator values this intuitive idea needs modification.) Alternatively, the consistency of the Ellenberg moisture values could be checked by studying the distribution of each species with respect to moisture. In this approach the moisture value of a site is calculated by averaging the indicator values of the species present. The indicator value of a particular species is clearly inconsistent with those of the other species when it deviates considerably from the center of the distribution of this species. This distribution also contains information on the ecological amplitude of the species for moisture.

In this paper this simple method is developed

further and compared with a more sophisticated maximum likelihood method, in which the species' distributions are modelled by Gaussian logit curves (Ter Braak & Looman, 1986). Both methods are applied to phytosociological (presence/absence) data from a diluvial part of The Netherlands to answer the questions stated above, the first of which being reformulated as: 'are Ellenberg's indicator values internally consistent in our study area?'.

Methods

Type of response curve

The relationship between the occurrence of a species and moisture may be shown in a *presence-absence response curve,* in which the probability $p(x)$ of occurrence of the species is plotted against moisture (x). Response curves may differ in shape and vary in complexity, but the response curves of species with respect to environmental variables are usually unimodal (Ellenberg, 1983; Whittaker, 1956). In this study we assume a unimodal response curve for each species with respect to moisture. In such curves, the width of the curve is proportional to the ecological amplitude and the position of its maximum is the indicator value. These two concepts lose their meaning in other response curves, such as bimodal or sigmoid curves.

Weighted averaging method

In the method of weighted averaging the *indicator value and ecological amplitude* of a species are defined as the *mean (M)* and *standard deviation (SD)* of the species' response curve. Thus, these characteristics are defined as if a response curve $p(x)$ were a statistical probability distribution (see Ter Braak & Barendregt, 1986). The moisture value of a relevé is estimated here as the average of Ellenberg's indicator values for moisture of all the species present in the relevé. Simplistic estimates of a species' indicator value and ecological amplitude would then be the sample mean and the sample standard deviation, respectively, of the moisture values of all relevés containing the species (Ter Braak & Looman, 1986). The newly calculated indicator values might then be compared with the indicator values given by Ellenberg (1979) to provide an informal test on the internal consistency of the latter. However, these estimates are too simple, because they neglect the distribution of the moisture values and their results may be misleading (Ter Braak & Looman, 1986). In an attempt to correct for the distribution of the moisture values, the moisture scale is divided into twelve classes, and the number of relevés, n_j, in each class j is counted. For any species a rough estimate of its response curve can then be obtained by calculating the fraction of relevés in each class that contain the species. These fractions can be displayed in a *response histogram* (Fig. 1). Improved estimates for the indicator value and ecological amplitude are then the mean and standard deviation of the response histogram. In this study the ecological amplitude is estimated in a slightly more subtle way, namely by

Table 1. Definition of Ellenberg's moisture values (Ellenberg, 1979).

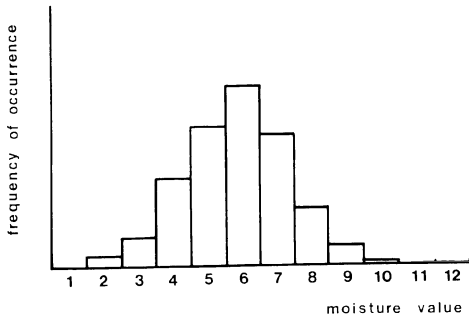|	
1	on extremely dry soils, e.g. bare rocks
2	in-between 1 and 3
3	on dry soils
4	in-between 3 and 5
5	on fresh soils, i.e. under intermediate conditions
6	in-between 5 and 7
7	on moist soils which do not dry out
8	in-between 7 and 9
9	on wet, often not well aerated soils
10	on frequently inundated soils
11	water plant with leaves mostly in contact with the open air
12	underwater plant, mostly totally immersed in water
×	indifferent

Fig. 1. Response histogram of a species with respect to moisture. The bars show the fraction of relevés in each moisture class which contain the species.

Fig. 2. Gaussian-logit response curve (p(x) = probability of occurrence of the species at value x, p(max) = maximum probability of occurrence, x = environmental variable, t = tolerance, u = optimum).

using Ellenberg's indicator value of the species instead of the sample mean in the formula for the standard deviation:

$$\widetilde{SD}^2 = \sum_{i=1}^{n} \frac{y_i}{n_j}(\bar{x}_i - M_o)^2 / \sum_{i=1}^{n} \frac{y_i}{n_j} \quad (1)$$

where n is the number of relevés, $y_i=1$ or 0 depending on whether the species is present or absent in relevé i, $\bar{x}_i$ is the estimated moisture value and j the class of relevé i, n_j is the number of relevés in class j and M_o is Ellenberg's (1979) indicator value of the species. The latter is used in equation (1), instead of any newly computed indicator value, to avoid underestimation of the ecological amplitude. We also used some variants of equation (1), but the differences in the results did not seem to be of practical importance.

Maximum likelihood method

Ter Braak & Looman (1986) proposed to model the presence-absence response curve of a species by the Gaussian logit curve, in which the logit-transform of probability is a quadratic function. According to this model the probability p_{ik} that species k occurs in relevé i is (Fig. 2)

$$p_{ik} = 1/\{1 + c_k \exp [\tfrac{1}{2} (x_i - u_k)^2/t_k^2]\} \quad (2)$$

where u_k is the optimum (the value of x with highest probability of occurrence of species k) and t_k is the tolerance (a measure of ecological amplitude) of species k and x_i is the moisture value of relevé i. The maximum probability of occurrence of species k is $1/(1 + c_k)$. The Gaussian logit curve is symmetric. Its optimum is therefore identical to its mean. Also, its tolerance is almost identical to its standard deviation when the maximum of the curve is small (Ter Braak & Looman, 1986). The range of occurrence of a species is largely restricted to an interval of length $4t$ (Fig. 2).

The idea behind the maximum likelihood method is to fit Gaussian logit curves to the relevé data. This is done by varying 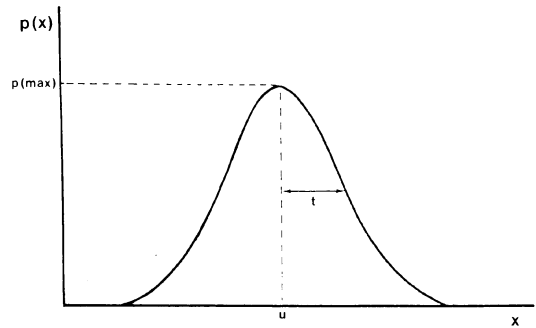the parameter values of the model in order to maximize the likelihood. The likelihood of a set of parameter values is defined as the probability of collecting the same data when this set of values were the true set of parameter values. In the present case the likelihood is taken to be the product of $p^y (1-p)^{1-y}$ over all relevés and species, with $p=p_{ik}$ and $y=1$ or 0 depending on whether species k is present or absent in relevé i. Logistic regression as utilized by Ter Braak & Looman (1986) is a special case of the maximum likelihood method, in which the species parameters (u_k, t_k and c_k) are estimated from data on species occurrence and known values of x_i. We could apply logistic regression here, using the moisture values from the weighted averaging method. However, in estimating the tolerances of the species it is more natural to assume, as in equation (1), that the optima are known, namely, that they are equal to Ellenberg's indicator values. From this assumption maximum likelihood estimates are derived for the moisture values of the relevés as well as for the tolerances and maxima of the species. The maximum likelihood estimates are obtained with an iterative algorithm:

(1) Start with the moisture values obtained by weighted averaging.

(2) Estimate the tolerance and maximum of each species from that species' data and the current moisture values.

(3) Estimate a new moisture value for each relevé from the floristic data, the species' optima and the current values for the tolerances and maxima of the species.

(4) Check whether the moisture values have changed, and if so, go back to step (2), otherwise stop.

In step (2) and step (3) the likelihood is maximized for each species and each relevé separately and, as a result, the total likelihood increases with each step. Step (2) resembles a Gaussian logit regression, but differs in that the optimum is given instead of being estimated. Step (3) of the maximum likelihood procedure has the attractive property that species with a small tolerance will have a greater effect on the estimation of the moisture value of a relevé than species with a large tolerance (cf. Ter Braak & Barendregt, 1986).

With the maximum likelihood method one can test statistically whether a species' optimum as specified by Ellenberg's indicator value is consistent with the indicator values of the other species. In this test the likelihood calculated above is compared with

a likelihood that is maximized also with respect to the value of the species' optimum (cf. Ter Braak & Looman, 1986). When the difference in residual deviance $(= -2$ log-likelihood) is larger than the critical value of a chi-square distribution with 1 degree of freedom, the species' optimum is shown to differ significantly from the value specified by Ellenberg (1979) and is therefore inconsistent with the indicator values of the other species. In principle this test can be carried out for each species in turn. However, in the present case, the test is very laborious because of the large number of parameters in the model. Because it is unlikely that the moisture values of the relevés will change much, when the second likelihood is maximized, they may just as well be kept fixed. Then, the statistical test amounts to comparing a species' indicator value with its optimum as estimated by a Gaussian logit regression of the data of this particular species on fixed moisture values. Instead of testing by deviance, we checked whether Ellenberg's indicator value lay within the 95%-confidence interval for the optimum. The construction of this interval is described by Ter Braak & Looman (1986). Such intervals were only constructed for species occurring in more than five relevés.

Data

In this study, 1 041 relevés (all from 1980 – 1982) were used representing the vegetation of the diluvial area in the western part of the province of Noord-Brabant, The Netherlands (Gremmen *et al.*, 1985) as follows: 323 relevés of woodland, 312 grassland, 250 marsh and ditch vegetation, 94 heathland and bog, and 62 other types. Quadrat size ranged from 4 m^2 in bog and grassland to 200 m^2 in woodlands.

Trees, large shrubs, and species that occurred less than 3 times were excluded. A total of 311 species remained, on average 13 per relevé; 280 of them had been assigned indicator values for moisture (Ellenberg, 1979). Most species have indicator values that are in the middle range (5 – 9). Of the species with more extreme moisture values 12% have an indicator value of 4 or less, and 16% have one above 9.

Results

The moisture values of the relevés estimated by the weighted averaging method showed a markedly uneven distribution, with many more 'wet' than 'dry' relevés (Table 2). These moisture values were strongly correlated $(r = 0.94)$ with those estimated by the maximum likelihood method, but as shown in Table 2, the estimated values for any single relevé may differ considerably (30% of the relevés differed by more than 0.5 unit, and 9% of the relevés by more than 1 unit).

Table 2. Comparison of the estimates of the moisture values of the relevés resulting from the weighted averaging method (x_{WA}) and the maximum likelihood method (x_{ML}). Entries refer to number of relevés.

x_{WA} x_{ML}	1	2	3	4	5	6	7	8	9	10	11	12	Total
1				1									1
2					3								3
3			1	11	5	2							19
4				1	13	1							15
5				1	67	36	1						105
6					28	220	61						309
7						15	122	13					150
8							43	99	6				148
9							16	108	76	4			204
10								3	24	10			37
11									7	19	11	1	38
12											10	2	12
Total	0	0	1	14	116	274	243	223	133	33	21	3	1041

Table 3. Comparison of the estimates of the species amplitudes from the weighted averaging method (SD, Equation (1)) and the maximum likelihood method (*t*, Equation (2)). Entries refer to number of species.

SD *t*	0.0 – 0.5	0.5 – 1.0	1.0 – 1.5	1.5 – 2.0	2.0 – 2.5	2.5 – 3.0	3.0 – 3.5	≥3.5	Total
0.0 – 0.5	6	30	1						37
0.5 – 1.0		30	20	1					51
1.0 – 1.5		18	51	7	4				80
1.5 – 2.0		1	32	27	2				62
2.0 – 2.5			7	13	4	1			25
2.5 – 3.0			2	3		2			7
3.0 – 3.5			1	2	2				5
≥3.5		1	1	5	1	1	1	3	13
Total	6	80	115	58	13	4	1	3	280

The simplistic estimate of a species' amplitude, that is the sample standard deviation (*SD*) of the moisture values of the relevés in which the species occurs, showed low correlation (0.2) with the more subtle estimate of *SD* by equation (1), which was on average 1.3 moisture scale unit. The maximum likelihood method tended to result in somewhat larger estimates of the amplitude than *SD* (Table 3). Species with indicator values of 11 and 12 had on average a markedly smaller tolerance than other species. This may be so because they are water plants.

In general the maximum probability of occurrence of a species estimated by the maximum likelihood method, was quite small; for only 23 (8%) of the species the maximum exceeded 0.50 and for 154 (55%) it was less than 0.10. Thus, the occurrence of most species cannot be predicted with confidence

from the moisture value of the site alone.

Figure 3 shows some typical examples of the response histograms and Gaussian logit curves fitted by the maximum likelihood method and by Gaussian logit regression.
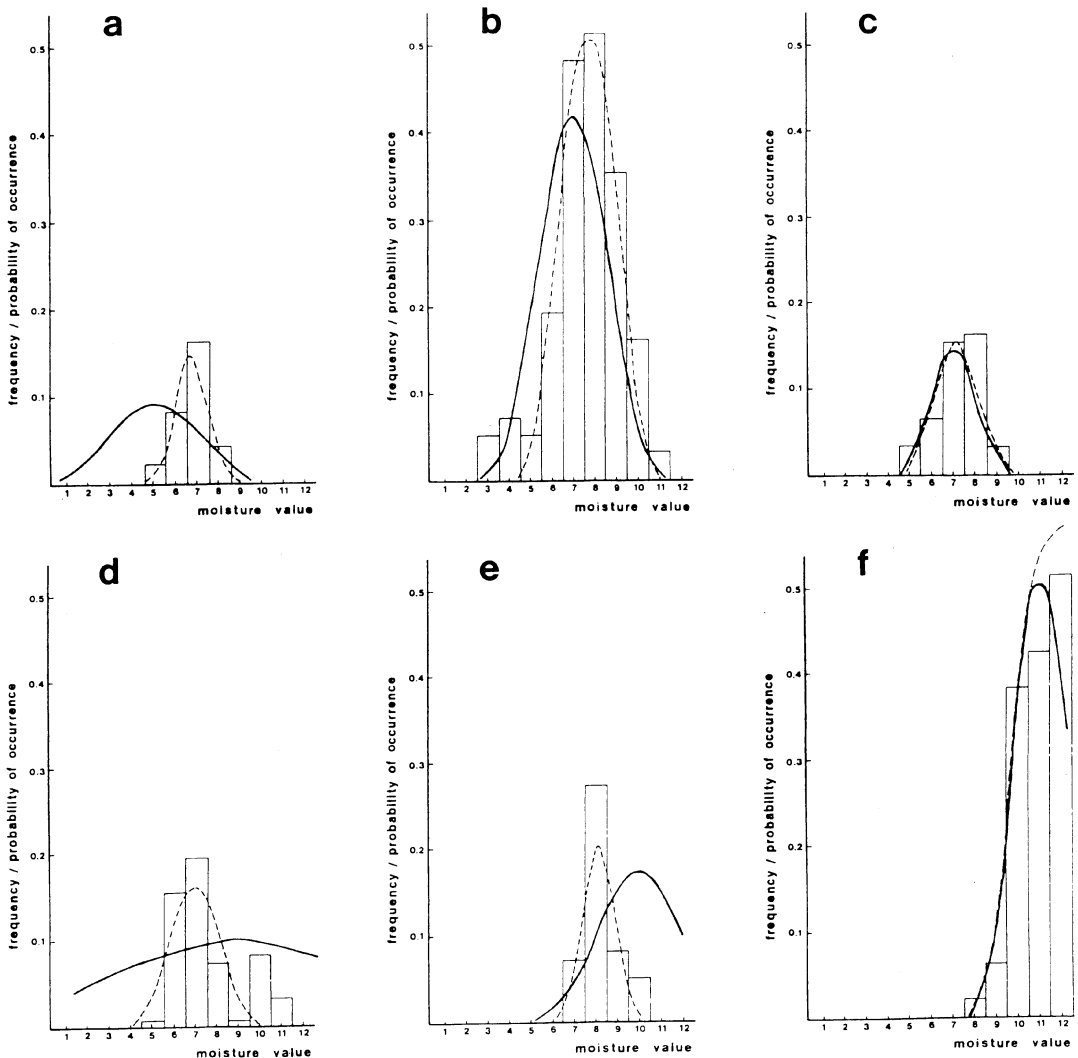
Fig. 3. Some examples of response histograms (bars) and estimated response curves. (— = response curve estimated by the maximum likelihood method; – – – = response curve estimated by Gaussian logit regression). F = Ellenberg moisture value, t = tolerance / estimated by the maximum likelihood method, SD = ecological amplitude estimated by the weighted averaging method. (a) *Heracleum sphondylium* $F = 5$; $t = 2.1$; $SD = 1.7$ (b) *Juncus effusus* $F = 7$; $t = 1.4$; $SD = 1.4$ (c) *Juncus subuliflorus* $F = 7$; $t = 1.0$; $SD = 0.9$ (d) *Alopecurus geniculatus* $F = 9$; $t = 5.5$; $SD = 2.2$ (e) *Iris pseudacorus* $F = 10$; $t = 1.8$; $SD = 1.8$ (f) *Lemna minor* $F = 11$; $t = 1.0$; $SD = 1.0$.

In some cases the mean of the response histogram deviates strongly from the indicator value of the species (Fig. 3a, e). In those cases the curve fitted by maximum likelihood with the species' indicator value taken as a fixed optimum, also deviates strongly from both the response histogram and the curve fitted by Gaussian logit regression. By using Gaussian logit regression 95%-confidence intervals for the optimum could be constructed for 175 (=73%) of the 240 species occurring in more than five relevés. For 45 (=26%) of these, Ellenberg's (1979) indicator value for moisture lay more than 0.5 unit outside this confidence interval. The extra 0.5 unit was used to allow for the fact that Ellenberg (1979) reports whole numbers. Thus for instance, an indicator value of 6.45 would be reported as 6. The indicator values of these species therefore are inconsistent with those of the other species. Table 4A gives information on species with an extreme deviation (≥ 1.7) between the Ellenberg moisture value and the estimated optimum. When no 95%-confidence interval could be calculated, the relationship between moisture and probability of occurrence was either non-significant (21 species) or sigmoid rather than unimodal (44 species), as judged by the deviance test at the 5%-level (cf. Ter Braak & Looman, 1986). No great inconsistencies in indicator value could be shown for species with a sigmoid relationship, because those with an Ellenberg indicator value of less than 7, showed a decreasing fitted response curve and those with an indicator value of 7 or more showed an increasing response curve (cf. Fig. 3f). It should be noted that a nonsignificant relationship or optimum may be due to a low frequency of a species in our data set and does not necessarily point to inconsistencies in Ellenberg's indicator values.

Gaussian logit regression was also used to check whether species Ellenberg (1979) considered indifferent, were also indifferent in our data set. For 28 of the 38 such species that occurred in 6 or more relevés, a 95%-confidence interval for the optimum could be calculated, and for 14 species the estimated tolerance was even less than 1.0 unit. Table 4B lists the species with the narrowest ecological amplitude ($t < 0.9$).

Our data set contained only three herbaceous species not mentioned by Ellenberg (1979) that occurred in more than 5 relevés; their indicator values were estimated by Gaussian logit regression (Table 4C).

Table 4. Ellenberg moisture value (*F*), estimated optimum, 95%-confidence interval for the optimum and estimated amplitude (tolerance) of a number of species. A. Species with a large discrepancy between Ellenberg moisture value and estimated optimum. B. Species with a narrow ecological amplitude, although regarded as indifferent by Ellenberg. C. Species not mentioned by Ellenberg.

Species name	F	optimum	interval	tolerance
A.				
Ornithopus perpusillus	2	4.1	3.5 – 4.3	0.5
Stellaria graminea	4	6.0	5.0 – 6.4	1.4
Alopecurus geniculatus	9	6.9	6.6 – 7.1	1.1
Iris pseudacorus	10	8.2	8.0 – 8.4	0.7
B.				
Anemone nemorosa	x	6.1	6.0 – 6.3	0.2
Melampyrum pratense	x	6.3	4.6 – 7.3	0.5
Bellis perennis	x	6.6	6.3 – 6.8	0.6
Prunella vulgaris	x	7.1	6.8 – 7.5	0.6
Ranunculus acris	x	6.9	6.7 – 7.0	0.7
Capsella bursa-pastoris	x	4.6	3.5 – 5.0	0.8
C.				
Eleocharis multicaulis	?	9.1	9.0 – 9.3	0.3
Epilobium obscurum	?	7.0	6.9 – 7.3	0.6
Myosotis laxa	?	7.8	7.5 – 8.1	0.7

Discussion

The ordinal scale of Ellenberg's indicator values

Ellenberg's indicator values are ordinal (strictly speaking values 11 and 12 are nominal); from the values in Table 1 we may infer which of two species prefers wetter conditions, but not the magnitude of the difference. But, in the methods applied here, the indicator values are treated as if they were quantitative, that is, as if they were measured on an interval scale. Durwen (1982) raised objections against such a quantitative treatment. In our opinion the ordinal nature of Ellenberg's moisture scale is far less important than the shape of the response curves, which should be symmetric (cf. Ter Braak & Barendregt, 1986). In the maximum likelihood method, a particular symmetric response curve was assumed — although response curves that are monotone by truncation, could also be dealt with. This condition of symmetry is equally important in the weighted averaging method, as mean and standard deviation are only useful characteristics for response curves that are more or less symmetric. After inspecting the response histograms of all species (cf. Fig. 3) we

concluded that the assumption of symmetry was not unreasonable, except, of course, for species with extreme optima. Therefore, we used the moisture indicator values of Ellenberg without transformation.

Comparison of the two methods

The weighted averaging method has three major problems. Firstly, as the number of relevés in each moisture class is not equal (Table 2), the estimates of the probability of occurrence in a class are not equally precise for all classes. The estimate of SD in equation (1) is closely related to the SD of the response histogram (Fig. 1), and it would seem reasonable to give less weight to classes with relatively few relevés. However, any such weighting policy, would make the estimator for SD again dependent on the distribution of the relevés, and thus cause bias.

A second problem is caused by relevés of extremely wet or extremely dry sites. The moisture values of these relevés will always be too low and too high, respectively, because only a few species indicate extreme conditions and many more species indicate conditions that are less extreme. Just by their numbers the probability of species of the latter group occurring at extreme sites is higher than of species indicating extreme conditions. This results in a general trend towards more moderate moisture values for extreme relevés, and this also results in a bias in the estimates for SD. Thirdly, the response histograms of species with an extreme indicator value will be truncated (cf. Fig. 3f) and it is not clear how the SD value of such species should be interpreted. The problem is partly one of definition, that is, when the response curve is truncated because more extreme conditions do not exist, it is not clear how SD should be defined, and partly one of estimation, namely when the response curve is truncated because more extreme conditions were not sampled, it is not clear how SD should be estimated. We do not know how to solve this problem in the weighted averaging method.

In the maximum likelihood method a specific model has to be adopted, in our case the Gaussian logit model. This is a disadvantage, since we do not really know the correct model. When the model is correct, the resulting estimates are better than in the weighted averaging method, but when it is incorrect, the meaning and quality of the estimates are unknown. We investigated the goodness-of-fit of the Gaussian logit curves obtained from the regressions with the usual chi-square test on the basis of observed and expected numbers of presence and absence in the 12 moisture classes. At the 5% level 72 species ($=27\%$) showed significant lack-of-fit. An example is *Alopecurus geniculatus* (Fig. 3d). The response histogram suggests gross deviations from the Gaussian logit curve in moisture classes 10 and 11, but these are due to only four occurrences. The important deviation is the low frequency of occurrence in moisture class 9. Despite the deviations, we believe that for our purpose and data the Gaussian logit model is a good compromise between model complexity and goodness-of-fit.

The problems in the weighted averaging method are largely solved automatically in the maximum likelihood method, where a truncated response curve is assumed to be part of a full Gaussian logit curve. However, an unexpected new problem arose, namely that the distribution of the moisture values of the relevés showed local minima near integer values.

This artifact (which is not apparent in Table 2) is because the Ellenberg (1979) indicator values are all integer values and in our method form the optima of the species' response curves. The maximum likelihood estimate of the moisture value of a relevé is based both on the species present and the species absent. When a species is present, it forces the estimate in the direction of the species' indicator value, whereas, when a species is absent, it forces the estimate away from the species' indicator value. Absence of a species usually has far less influence than presence, that is, when the maximum probability of occurrence of the species is low (Ter Braak & Barendregt, 1986). But the number of species absent in a relevé is large compared to the number of species present. If, for instance, the true moisture value of a relevé is 6.0, all species with an indicator value of 6 that are absent will force the estimate away from the value 6.0 and this force cannot be counteracted by the presence of a small number of species with this same indicator value. The maximum likelihood estimate thus tends to avoid the integer values. We believe that in the present study this artifact is not a very serious problem. Because the average width of the response curves is large as compared to the scale of these irregularities, the fitting of curves will still give a reasonable estimate of the species tolerance.

The maximum likelihood method has the additional advantage over the weighted averaging method by giving approximate standard errors of estimates, which makes it possible to test the internal consistency of the Ellenberg indicator values.

Improving the indicator values by ordination?

Clausman (1980) attempted to improve indicator values by an iterative procedure; he calculated moisture values for the relevés from the indicator values and then new indicator values from the moisture values, and then new moisture values from the new indicator values, and so on. This procedure is essentially an ordination method. For example, when weighted averaging is used in each calculation, the method amounts to reciprocal averaging. By consequence, the original meaning of the indicator values may get lost.

We applied detrended correspondence analysis (Hill & Gauch, 1980), to our data and found practically no correlation between the (initial) moisture values of the relevés and the (final) scores on the first axis ($r=0.01$). The first axis turned out to be highly correlated ($r=0.99$) with the nitrogen values of the relevés, estimated by averaging the Ellenberg indicator values for N, whereas the second axis was highly correlated ($r=0.99$) with the moisture values. Applied to our data, Clausman's (1980) method would have changed the Ellenberg's indicator values for moisture into indicator values for nitrogen, which is clearly unwanted! Consequently, ordination cannot be used to improve indicator values, except in the hypothetical case that it is certain that the main variation in the species data corresponds exactly to the factor one wants to improve the indicator values of. Therefore, we kept the indicator values fixed in both our methods and tested each species separately to see if its value was consistent with the indicator values of the other species.

Due to the implicit nature of Ellenberg's definition of moisture, it is impossible to improve the moisture values in a statistically sound way on the basis of floristic data only.

On generalizing the results

Our results show the ecological amplitude (*SD* or tolerance) of a species to be about 1.0 to 1.5 units on Ellenberg's moisture scale. Consequently, the range of a species' occurrence is estimated to be on average 4−6 units. It is difficult to say how these results are affected by conditions specific to our study area. The detrended correspondence analysis showed nitrogen to be the environmental variable that is most important for explaining the floristic

variation in our data. Consequently, the assumption in the maximum likelihood method of independence of the species is incorrect. Fortunately, nitrogen was practically uncorrelated with moisture, and therefore unlikely to have distorted the results to a large extent. The fact that moisture is shown to be the second most important environmental variable in our data set also gives some confidence in the results. In different geographical regions, the environmental variables that are most important for explaining the species distribution may differ. Especially when these factors are correlated with moisture, the estimates of the amplitude of a species with respect to moisture may differ because of distortion by these factors. In principle, the problem of other influential variables can be overcome in the maximum likelihood method by analysing more than one variable simultaneously. We may attempt this in the future.

Conclusion

The use of Ellenberg's moisture values on floristic data in estimating site moisture is an example of environmental calibration. Ellenberg's method of environmental calibration assumes a simple model of the responses of plant species to moisture: symmetric, unimodal response curves and equal amplitudes. This model does not include interaction effects of other environmental variables with moisture. A more precise calibration system necessarily has to include such interactions. Such a system could be derived from actual measurements of environmental variables and associated floristic data (Ter Braak & Barendregt, 1986), but would lose the simplicity and supposed general applicability of the Ellenberg system. May our results serve to increase the confidence with which Ellenberg's indicator values for moisture are used.

References

Böcker, R., Kowarik, I. & Bornkamm, R., 1983. Untersuchungen zur Anwendung der Zeigerwerte nach Ellenberg. Verh. Ges. Ökol. 9: 35−56.
Clausman, P. H. M. A., 1980. Ecologische interpretatie van vegetatieopnamen m.b.v. een computer. WLO-Meded. 7: 92−98.

Durwen, K.-J., 1982. Zur Nutzung von Zeigerwerten und art-spezifischen Merkmalen der Gefäszpflanzen Mitteleuropas für Zwecke der Landschaftsökologie und -planung mit Hilfe der EDV-Voraussetzungen, Instrumentarien, Methoden und Möglichkeiten. Arbeitsber. Lehrst. Landschaftsökol. Münster 5: 1–138.

Ellenberg, H., 1979. Zeigerwerte der Gefäszpflanzen Mitteleuropas. 2nd ed., Scripta Geobotanica 9, Göttingen.

Ellenberg, H., 1983. Vegetation Mitteleuropas mit den Alpen in ökologischer Sicht. 3rd ed., Ulmer Verlag, Stuttgart.

Gremmen, N. J. M., Reijnen, M. J. S. M., Wiertz, J. & Van Wirdum, G., 1985. Modelling for the effects of ground-water withdrawal on the species composition of the vegetation in the Pleistocene areas of The Netherlands. In: Ann. Rep. 1984. Research Institute for Nature Management, Arnhem, pp 89–111.

Hill, M. O. & Gauch, H. G., 1980. Detrended correspondence analysis: an improved ordination technique. Vegetatio 42: 47–58.

Persson, S., 1981. Ecological indicator values as an aid in the interpretation of ordination diagrams. J. Ecol. 69: 71–84.

Reijnen, M. J. S. M. & Wiertz, J., 1984. Grondwater en vegetatie: een nieuw systeem voor kartering en effectvoorspelling. (Engl. Summary) Landschap 1: 261–281.

Smeets, P. J. A. M., Werger, M. J. A. & Tevonderen, H. A. J., 1980. Vegetation changes in a moist grassland under altered water conditions. Biol. Conserv. 18: 123–142.

Ter Braak, C. J. F. & Barendregt, L. G., 1986. Weighted averaging of species indicator values: its efficiency in environmental calibration. Math. Biosci. 78: 57–72.

Ter Braak, C. J. F. & Looman, C. W. M., 1986. Weighted averaging, logistic regression and the Gaussian response model. Vegetatio 65: 3–11.

Whittaker, R. H., 1956. Vegetation of the Great Smoky Mountains. Ecol. Monogr. 26: 1–80.

An evaluation of the relative robustness of techniques for ecological ordination

Peter R. Minchin*
CSIRO Division of Water and Land Resources, G.P.O. Box 1666, Canberra, 2601, Australia

Abstract

Simulated vegetation data were used to assess the relative robustness of ordination techniques to variations in the model of community variation in relation to environment. The methods compared were local non-metric multidimensional scaling (LNMDS), detrended correspondence analysis (DCA), Gaussian ordination (GO), principal components analysis (PCA) and principal co-ordinates analysis (PCoA). Both LNMDS and PCoA were applied to a matrix of Bray-Curtis coefficients. The results clearly demonstrated the ineffectiveness of the linear techniques (PCA, PCoA), due to curvilinear distortion. Gaussian ordination proved very sensitive to noise and was not robust to marked departures from a symmetric, unimodal response model. The currently popular method of DCA displayed a lack of robustness to variations in the response model and the sampling pattern. Furthermore, DCA ordinations of two-dimensional models often exhibited marked distortions, even when response surfaces were unimodal and symmetric. LNMDS is recommended as a robust technique for indirect gradient analysis, which deserves more widespread use by community ecologists.

Introduction

Ordination techniques are commonly employed as research tools in the study of vegetation. A major objective of ordination in vegetation ecology is that which Whittaker (1967) termed indirect gradient analysis. When faced with the diversity of available

methodology, ecologists have sought to identify those techniques which are most appropriate for the purpose of indirect gradient analysis.

Three major approaches to the comparative evaluation of ordination techniques may be identified:

(1) The application of different ordination methods to sets of field data (e.g. Prentice, 1977; Clymo, 1980; Oksanen, 1983; Brown *et al.*, 1984).

(2) The comparison of vegetational ordinations with direct ordinations based on environmental indices (e.g. Loucks, 1962; Del Moral, 1980).

(3) The use of simulated data, derived from explicit models of community variation along environmental gradients (e.g. Swan, 1970; Austin, 1976; Fasham, 1977; review by Whittaker & Gauch, 1978; Prentice, 1980).

The first strategy suffers from the major limitation that there is no precise statement of the underlying gradient structure which a successful ordination is expected to recover. The ordination results are assessed on the basis of preconceptions about the major environmental relationships derived from previous work. It is seldom possible to make quantitative statements about sample positions on the underlying environmental gradients which are sufficiently precise to allow a sensitive comparison of the performance of different ordination methods and independent of the biases of the formal or informal methods of vegetation analysis used in previous work.

The second approach assumes that the axes chosen for direct ordination do indeed represent the major gradients to which vegetational composition is related. There is usually no way to assess the validity of this assumption. Furthermore, it is often extremely difficult to identify an environmental variable which is susceptible to measurement and which adequately expresses the dynamic complexity of inter-related factors which characterises many environmental gradients.

The third, the simulation approach, has a number of advantages:

(1) The expected result of ordination can be specified precisely, using the known co-ordinates of sites in the simulated environment space.

(2) Various model properties may be varied independently in order to study their effects on ordination success, both alone and in combination.

(3) The degree of stochastic variation (noise) can be controlled.

The major limitation is that the models employed may not be adequate simplifications of the natural situation (Austin, 1980; Greig-Smith, 1980). Since the initial work of Swan (1970), most simulation studies have used models which assume that species responses are Gaussian, that sites are spread uniformly throughout environment space and that community properties such as alpha diversity (number of species per sample) and sample total (total abundance of all species in a sample) do not vary systematically along gradients. All these assumptions may be questioned.

The available *evidence,* from direct observations

of species' response along recognised gradients and experimental studies with mixed communities, is insufficient to develop a general model of community change along gradients (Austin, 1980). The assumptions which Gauch & Whittaker (1972a, 1976) built into their computer programs for the generation of artificial data are still only hypotheses, the critical testing of which is just beginning (Minchin, 1983; Austin, 1985, 1987). It follows that the most one can expect to achieve in simulation studies is the identification of techniques which are *robust* to variation in those features of a vegetation-gradient model which fall within the bounds of current possibilities (cf. Austin, 1976, 1980, 1985). A robust method should be capable of achieving an adequate recovery of the underlying gradients over a range of model properties. This study assesses the comparative robustness of several ordination methods.

Methods

Models

The artificial data matrices used in this comparison were generated using an early version of COENOS, a flexible computer program for the simulation of community variation along environmental gradients (Minchin, 1983, 1987), which is available upon request. The program simulates species responses using a generalised beta-function, which can produce unimodal response surfaces of differing skewness. Interaction may be introduced between species, leading to ecological responses which are shouldered, bimodal or multimodal. COENOS can produce models with up to six gradients and there are flexible options for sampling patterns and the introduction of stochastic variation (noise). In addition, the variation along gradients of community properties such as alpha diversity and sample total can be controlled.

The following model properties were examined in this study:

(1) The number of underlying gradients (1 or 2) and their beta diversities or compositional lengths.

(2) The shape of species' ecological responses (symmetric/skewed, unimodal/bimodal/multimodal).

(3) The arrangement of sites in the simulated environment space (regular/random/clumped/restricted).

Table 1. The structure of the simulation experiments designed to examine the joint effects of model properties on the performance of ordination techniques. The beta diversity of gradients is expressed in *R* units, where a gradient of 1*R* is equal in length to the mean range of species occurrence. Quantitative noise levels are expressed in the *F* units defined by Gauch & Whittaker (1972a).

Models with a single simulated gradient (coenoclines)

Expt. no.	Model properties held constant	Model properties varied	Replicates per cell	Total no. of models
A1	Response curve shape (symmetric) Sampling pattern (regular) Trend in sample totals (not controlled)	Beta diversity (5 levels: 0.25, 0.5, 1, 2, 4*R*) Quantitative noise (3 levels: 0.0, 0.1, 0.2*F*) Qualitative noise (2 levels: absent, present)	1	30
A2	Beta diversity (1*R*) Sampling pattern (regular) Quantitative noise (0.0*F*) Qualitative noise (absent) Trend in sample totals (not controlled)	Response curve shape (6 lev.: symmetric, slight consistent skewness, ibid. extreme, mixed skewness, interaction between symmetric curves, ibid. skewed curves)	3	18
A3	Beta diversity (1*R*) Sampling pattern (regular) Trend in sample totals (not controlled)	Response curve shape (6 lev. as in A2) Quantitative noise (3 lev.: 0.0, 0.1, 0.2*F*) Qualitative noise (2 lev.: absent, present)	1	36
A4	Beta diversity (1*R*) Quantitative noise (0.0*F*) Qualitative noise (absent) Trend in sample totals (not controlled)	Response curve shape (4 lev.: symmetric, extreme consistent skewness, interaction between symmetric curves, ibid. skewed curves) Sampling pattern (4 lev.: regular, random, concentrated in centre, concentrated towards ends)	1	16
A5	Beta diversity (1*R*) Sampling pattern (regular) Quantitative noise (0.0*F*) Qualitative noise (absent)	Response curve shape (4 lev. as in A4) Trend in sample totals (3 lev.: not controlled, linear trend, parabolic trend)	1	12

Models with two simulated gradients (coenoplanes)

B1	Response surface shape (symmetric) Sampling pattern (regular) Trend in sample totals (not controlled)	Beta diversities (3 lev.: 0.33 × 0.33*R*, 1 × 0.33*R*, 1 × 1*R*) Quantitative noise (3 lev.: 0.0, 0.1, 0.2*F*) Qualitative noise (2 lev.: absent, present)	1	18
B2	Sampling pattern (regular) Quantitative noise (0.0*F*) Qualitative noise (absent) Trend in sample totals (not controlled)	Beta diversities (2 lev.: 1 × 0.33*R*, 1 × 1*R*) Response surface shape (5 lev.: symmetric, slight consistent skewness, ibid. extreme, mixed skewness, interaction between skewed curves)	3	30
B3	Sampling pattern (regular) Qualitative noise (absent) Trend in sample totals (not controlled)	Beta diversities (2 lev.: 1 × 0.33*R*, 1 × 1*R*) Response surface shape (5 lev. as in B2) Quantitative noise (3 lev.: 0.0, 0.1, 0.2*F*)	1	30
B4	Beta diversities (1 × 1*R*) Quantitative noise (0.0*F*) Qualitative noise (absent) Trend in sample totals (not controlled)	Response surface shape (3 lev.: symmetric, extreme consistent skewness, interaction between skewed curves) Sampling pattern (5 lev.: regular, random, concentrated in centre, concentrated around edges, T-shaped pattern, cross-shaped pattern)	1	18

(4) The type (quantitative/qualitative) and amount of noise.

(5) The trend of sample total along gradients (none/linear/parabolic).

A number of experiments were designed in which two or three of the properties were varied factorially while the others were held at a constant value. The structure of the experiments is summarized in Table 1. Computing limitations restricted the amount of replication. Since the robustness of ordination techniques to variation in response shape was of major interest, three replicates per treatment cell were employed in the two experiments on response shape (Table 1: A2, B2).

In all, 174 data matrices were produced, 87 with a single underlying gradient and 87 with two gradients. Some data matrices took part in more than one experiment (e.g. the noiseless, 1R data set in experiment A1 was also used as one of the symmetric response shape models in experiment A2). Each data set was subjected to ordination by each of the techniques listed below, with the exception that Gaussian ordination (GO) was applied to the unidimensional models only: the program used for GO produces only one ordination axis.

Complete details of model construction are given by Minchin (1983). In all models, the total number of species was adjusted to give an average of about 25 species per sample in the data sets created without noise. The modal abundances of species were allocated from a lograndom distribution, with limits of 1 to 100 for models with no interspecific interaction and 5 to 100 for interaction models. Species' modal positions were randomly distributed. In the models without interspecific interaction, those 15% of species with the largest response function integrals had their modes adjusted to a more even spacing (see Minchin, 1983, 1987 for more details). Ranges of occurrence on each gradient were allocated from a normal distribution, with the mean value determining the compositional length (beta diversity) of the gradient. Minchin (1987) introduced the 'R' unit to express the beta diversity of simulated gradients. A beta diversity value of 1R indicates a gradient whose length is equal to the mean range of species occurrence. There is no simple relationship between the R unit and other common measures of beta diversity, such as the half-change (Whittaker, 1960) or the sd unit (Hill & Gauch, 1980). However, 1R is approximately equal to 6 sd and 4.5 half-changes, when response functions are unimodal, fairly symmetrical and not grossly long-tailed or flat-topped. The standard deviation of species' ranges was set at 0.3 times the mean value for models without interspecific interaction and 0.5 times the mean for interaction models. The differences in the simulation parameters for interaction models, relative to those for non-interaction models, were determined after initial empirical trials. The aim was to produce interaction models with a reasonable number of complex ecological response functions but without too many 'extinctions' of species due to the interaction adjustments.

The choice of a lograndom distribution for modal abundances is based on analyses reported by Minchin (1983) and Gauch & Whittaker (1972a). No good evidence is available about the frequency distribution of ranges of occurrence. A normal distribution was accepted as a working hypothesis on the basis of preliminary, informal analyses of Gauch & Whittaker (1972a). The use of a random distribution for species modes accords with the results of Minchin (1983). Austin (1987) suggests that the distribution of modes tends to be clumped, but his analysis was restricted to a single functional guild (canopy trees).

Techniques compared

From the range of available ordination techniques, five were selected for comparative evaluation:

1. Principal components analysis (PCA) (Hotelling, 1933).

2. Principal co-ordinates analysis (PCoA) (Gower, 1966).

3. Detrended correspondence analysis (DCA) (Hill & Gauch, 1980), using the program DECORANA (Hill, 1979).

4. Gaussian ordination (GO), (Gauch et al., 1974), using Cornell Ecology Program 8B (Gauch, 1979).

5. Local non-metric multidimensional scaling (LNMDS) (Kruskal, 1964a, b; Sibson, 1972), using the program KYST (Kruskal et al., unpubl.).

PCA has been shown in previous simulation studies (e.g., Noy-Meir & Austin, 1970; Austin & Noy-Meir, 1971; Gauch & Whittaker, 1972b; Fasham, 1977) to produce distorted representations of underlying gradients unless beta-diversity is low. The reason for the distortion is that the mathematical model of PCA implies a linear relationship between compositional dissimilarity (as expressed by the Euclidean metric calculated from species data) and the separation of sites along environmental gradients. In fact the relationship is non-linear: as one moves further apart in environment space, the rate of increase in compositional dissimilarity tends to decline (Swan, 1970). In order to fit its linear model, PCA must represent gradients as curved, rather than linear trends.

Despite the recognition that the linear model of PCA is inappropriate unless beta diversity is very low, the method continues to be applied to community data by plant ecologists (e.g., Bradfield & Scagel, 1984; Van der Maarel et al., 1985) and is apparently popular with animal ecologists (e.g., Rotenberry & Wiens, 1980). This may be partly due to a lack of acceptance of the results of simulation studies, because the Gaussian models employed therein were regarded as unrealistic. Thus PCA was included in this study to test the expectation that its linear model would lead to similar distortions with a range of alternative non-linear response models.

The performance of PCA, as a method of indirect gradient analysis, varies according to the manner in which the data are

standardized (Austin & Noy-Meir, 1971). In this study, PCA was applied with each of three standardizations: (1) centred by species mean; (2) centred by species and standardized by species standard deviation (equivalent to an R-mode PCA of the correlation matrix between species) and (3) Bray-Curtis successive double standardization (i.e. species adjusted to equal maxima, then samples standardized to equal totals), followed by centring by species. The abbreviations PCA-C, PCA-CS and PCA-BC will be used below when referring to these three variants of PCA.

PCoA is effectively a generalization of PCA which allows the use of a much wider range of measures of compositional dissimilarity. In this study it was applied to the Bray-Curtis coefficient which is also known as percentage difference (Gauch, 1982) and the Czekanowski coefficient (see, e.g., Greig-Smith, 1983). This coefficient has been widely used in ecology. Values of the Bray-Curtis coefficient have a less curvilinear relationship with environmental separation (the distance between samples in environmental space) than do values of the Euclidean distance metric (Gauch, 1973; Faith et al., 1987). It was therefore expected that the use of this coefficient might increase the effectiveness of PCoA relative to PCA.

The remaining three techniques, GO, DCA and LNMDS, have been introduced to ecology as potential solutions to the problem of curvilinear distortion in linear ordinations. On the basis of simulation studies using Gaussian models (Hill & Gauch, 1980; Gauch et al., 1981) DCA is now commonly regarded as the 'state of the art' method (Gauch, 1982). It is gaining broad acceptance among ecologists (e.g. Walker & Peet, 1983; Beatty, 1984; Van der Maarel et al., 1985). However, the sensitivity of DCA to departures from the Gaussian model has not been assessed.

The underlying model of NMDS is relatively simple: given a matrix of resemblances (similarities or dissimilarities) between pairs of objects, NMDS constructs a configuration of points in a specified number of dimensions, such that the rank order agreement between the inter-point distances and the resemblance values is maximized. In ecological applications the objects are usually samples and the dissimilarities are calculated from the compositional data using some chosen coefficient. The epithet 'non-metric' refers to the fact that only the rank order of the input dissimilarities is utilized. This contrasts with methods of metric scaling (e.g. PCA, PCoA, correspondence analysis) where the distances between points in the derived configuration are *proportional* to the dissimilarities.

Gauch (1982; see also Gauch et al., 1981), who stated that NMDS assumes 'monotonicity, which is a weaker and better assumption than linearity but is still unrealistic for handling the Gaussian curve, which is ditonic', confused the model of species' responses to gradients with the model of the relationship between ordination distance and compositional dissimilarity. The monotonicity assumption of NMDS refers to the latter. NMDS does not make a direct assumption about the form of species response functions. In theory, NMDS can accommodate any type of response function, provided that the resulting relationship between compositional dissimilarity (as expressed by some dissimilarity coefficient) and sample separation in environment space remains approximately monotonic.

The 'global' variant of NMDS derives a configuration in which the distances between all pairs of sample-points are, as far as possible, in rank order agreement with their compositional dissimilarities. Any given pair of samples which are less similar in composition than some other pair should be placed further apart than that other pair in the ordination. The 'local' variant of NMDS (Sibson, 1972) has a more relaxed criterion: for each sample, the distances from its point to each other sample-point in the ordination should be in rank order with the compositional dissimilarities between that sample and each other sample. This variant allows for the possibility that the pattern of decline in compositional dissimilarity with increasing environmental separation may differ from point to point in environment space (Prentice, 1977, 1980).

In this study, NMDS was applied in the 'local' form and DCA ordinations were used to provide starting configurations. The Bray-Curtis coefficient was used as a measure of compositional dissimilarity, thus making the NMDS ordinations directly comparable to the ordinations by PCoA. Previous work with Gaussian models (Gauch, 1973) has shown that the Bray-Curtis coefficient has an approximately monotonic relationship with sample separation along gradients. Faith et al. (1987) used a range of models similar to those in this study to compare the rank correlations between compositional dissimilarity and sample separation for a variety of dissimilarity coefficients. The Bray-Curtis coefficient was among the most effective and robust of the measures compared. The coefficient attains a maximum value of 1.0 for all pairs of samples which have no species in common. Values of 1.0 are indeterminate, in the sense that they indicate only a lower bound on the gradient separation of a sample pair. For this reason, all values of 1.0 were coded as 'missing' in the present LNMDS ordinations. This results in the exclusion of such sample-pairs from the monotonic regressions of distance on dissimilarity. Stress formula 1 of Kruskal (1964a) was used and the 'primary' approach to tied dissimilarities was adopted (Kruskal, 1964a): if several pairs of samples have an identical compositional dissimilarity, they need not be given equal distances in the ordination. Model data sets with a single gradient were ordinated in both one and two dimensions, while models with two gradients were subjected to LNMDS in two and three dimensions.

Gauch et al. (1981) compared several available programs for NMDS and recommended ALSCAL (Young & Lewyckyj, 1979) as the fastest and most useful. Unfortunately, ALSCAL does not perform NMDS as originally proposed by Kruskal (1964a, b). ALSCAL uses an alternating least-squares algorithm, which maximizes the monotonic fit between *squared* ordination distances and *squared* compositional dissimilarities. Consequently, the larger dissimilarities receive relatively higher weight in the fitting process. These dissimilarities, between samples with few or no species in common, are the least informative. ALSCAL ordinations therefore tend to represent local structure poorly. Comparisons based on simulated data sets (Minchin, Faith & Belbin, unpublished) have shown that NMDS ordinations by KYST consistently recover the gradient structure much more successfully than ALSCAL ordinations. All LNMDS ordinations in this study were performed using the program KYST.

Assessment of fit

For the purpose of indirect gradient analysis, a successful ordination is defined as one in which the relative positions of samples matches their relative locations in environment space. Consequently, ordination performance may be assessed by comparing ordination configurations with the configuration of samples in the simulated environmental space.

A quantitative measure of the degree of fit between an ordination and the environmental configuration was obtained using Procrustean analysis (Schönemann & Carroll, 1970; Fasham, 1977). This technique fits one configuration to another using a combination of origin translation, rigid rotation and reflection of reference axes and uniform central dilation or contraction of scaling. The combination of transformations is found analytically, so as to minimize the sum of the squared distances between each point in the fitted configuration and its corresponding point in the target configuration. The RMS average of these displacements may be used as a measure of the discrepancy between the configurations (i.e. lower values indicate better fit). In presenting the results, the Procrustean discrepancy values are denoted by the symbol D, with a subscript indicating the number of dimensions in which the fit was performed.

For LNMDS ordinations performed with one more dimension than the number of gradients in the model, Procrustean analysis was performed in the higher dimensionality. This was achieved by adding an extra dimension to the simulated configuration, upon which each sample was given a score of zero. After fitting the ordination to the environmental configuration embedded in this space, the RMS displacement (D) was computed in the subspace defined by the original model gradient(s).

A limitation of D is that it can not distinguish situations where the lack of fit is due to a *systematic* distortion of the target configuration from those in which each point has been independently perturbed. In the first case, the fit is generally much better in some regions of the configuration than in others. Because of this limitation, the assessment of ordination performance was not restricted to comparison of the D values. All configurations were plotted and examined visually.

For unidimensional models, the degree of rank-order agreement between site locations along the simulated gradient and site scores on the first ordination axis was also assessed. Kendall's rank correlation coefficient (τ) was used for this purpose.

Results

PCA and PCoA

Curvilinear distortion of the underlying gradient(s) was evident in all PCA ordinations, irrespective of response function shape or other model properties. The distortion became more severe as beta diversity increased. Of the three variants of PCA examined, PCA-CS was the most resistant to distortion and PCA-C was the least successful. These results are consistent with those of earlier studies, based on Gaussian models (e.g. Noy-Meir & Austin, 1970; Austin & Noy-Meir, 1971; Fasham, 1977). It is evident that the curvilinear distortion of gradients by PCA is not due to peculiar properties of the Gaussian model.

PCoA with the Bray-Curtis coefficient performed similarly to PCA. In general, the degree of curvilinear distortion in PCoA ordinations was somewhat greater than in those produced by PCA-CS. For unidimensional models, the sequence of sites along the simulated gradient appeared as a curved path in the space defined by the first two axes of a PCA or PCoA ordination. Even when sites were regularly spaced along the model gradient they sometimes appeared bunched in PCA or PCoA ordinations, reflecting the curvature of the gradient into the third (or higher) dimensions.

With two underlying gradients, the consequences of curvilinear distortion by PCA and PCoA were more severe (Fig. 1). The planar pattern of sites in environment space was twisted and curved into three or more dimensions, making the recognition of the two-gradient structure in the resultant ordination impossible without prior knowledge of the underlying model. If the lines joining sets of sample points with equal co-ordinates on the second simulated gradient were deleted from Fig. 1, it would be very difficult to perceive the underlying gradients in the ordinations. The distortion is not difficult to recognize for unidimensional data sets, when an arch or horse-shoe shaped configuration is usually obtained in the first two dimensions. However, the planar structure of a two-dimensional data set can be twisted and curved into more than three dimensions, so that the shape of the configuration cannot be viewed in any two or three dimensional plot.

Gaussian ordination

The only combination of model properties in which GO consistently produced better results than
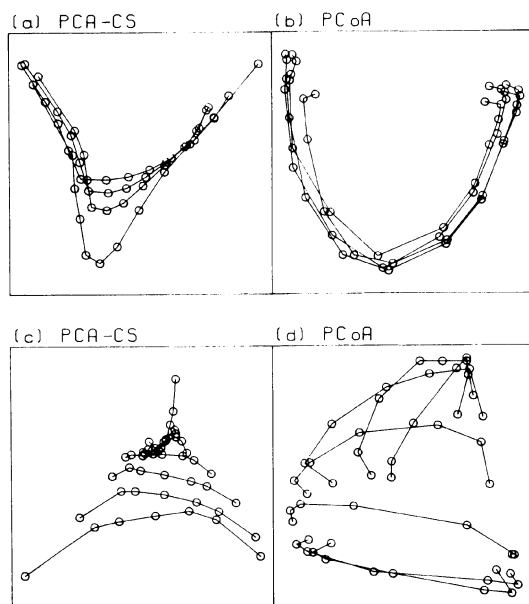
Fig. 1. PCA-CS and PCoA ordinations of a 1 × 0.33R (a, b) and a 1 × 1R (c, d) coenoplane. Both models had symmetric, unimodal response surfaces with no noise. Samples were arranged on regular 12 × 4 (a, b) and 7 × 7 (c, d) grids, respectively. Configurations on ordination axes 1 v 2 are shown, after being fitted to the simulated sampling pattern by Procrustean analysis. The lines join samples with equal co-ordinates on the second simulated gradient.

DCA and LNMDS was for data sets with a beta diversity of 1R or less, fairly symmetric response curves and either no noise or qualitative noise only. GO is relatively resistant to qualitative noise because only non-zero values are used in the fitting of Gaussian regressions, however the technique proved highly sensitive to quantitative noise.

Some performance statistics for GO in experiments A2 and A3 are given in Table 2. In the absence of noise, GO consistently recovered the correct rank order of samples only when response curves were symmetrical or of mixed skewness. The Procrustean discrepancy values in these situations were generally somewhat lower than those for DCA and LNMDS, indicating a better recovery of the inter-sample spacings.

The results in Table 2 clearly demonstrate the sensitivity of GO to the addition of quantitative

noise. For those noiseless data sets where GO achieved lower Procrustean discrepancies than DCA or LNMDS, the advantage of GO was reduced and usually reversed when noise was added. The failures of GO with noisy data sets were occasionally spectacular (e.g., Table 2, interaction model with noise level = 0.2 *F*).

DCA versus LNMDS

In view of the curvilinear distortion in PCA and PCoA ordinations and the restricted utility of GO, major interest centred on the relative performance of DCA and LNMDS. The balance between these two techniques depended on model properties, in particular the number and relative beta diversities of gradients, response shape and sampling pattern.

Response function shape

Some results for unidimensional models with a beta diversity of 1R and different response function shapes are given in Table 2. In the absence of noise, DCA recovered the rank order of samples perfectly *provided that all response functions were unimodal.* However, DCA failed in rank order recovery on two of the interaction coenoclines, which included some species with shouldered, bimodal or multimodal responses. LNMDS proved more robust to variation in response shape: when performed in one dimension, LNMDS achieved perfect rank order recovery over all shape categories. The Procrustean discrepancy values follow a similar pattern. LNMDS had slightly worse Procrustean fits than DCA for most of the symmetrical, mixed skewness and extreme skewness models, but consistently achieved better fits than DCA for the interaction models.

The effect of variation in response shape for two-dimensional models is exemplified by the results in Table 3. For 'rectangular' coenoplanes, with a beta diversity of 1 × 0.33R, neither DCA nor LNMDS had a consistent advantage when response surfaces were symmetrical or of mixed skewness. Under these conditions, relative performance in terms of Procrustean fit differed between replicates. Howev-

Table 2. Ordination performance statistics for some unidimensional models in experiments A2 and A3 (see Table 1). Both Kendall rank correlations (τ) and Procrustean discrepancy values (D_1) are tabulated. The effect of quantitative noise was only studied for the first replicate in each response shape category. All models have a beta diversity of 1R. Larger values of τ indicate better performance. For D_1, the opposite is true.

Ordination method		Ecological response function shape																			
		Symmetrical					Mixed skewness					Extreme skewness					Interaction (skewed)				
		Replicate					Replicate					Replicate					Replicate				
		1			2	3	1			2	3	1			2	3	1			2	3
		Noise level (F)					Noise level (F)					Noise level (F)					Noise level (F)				
		0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0
DCA	τ	1.00	0.99	0.98	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	0.99	0.96	1.00	1.00	0.98	0.93	0.90	1.00	0.99
	D_1	3.1	3.1	3.7	1.5	1.7	1.2	1.5	1.7	1.4	2.8	3.1	3.3	3.9	1.8	2.1	5.2	5.9	6.7	3.6	4.3
GO	τ	1.00	0.98	0.93	1.00	1.00	1.00	0.99	0.94	1.00	1.00	0.99	0.97	0.93	0.99	1.00	0.97	0.93	0.52	1.00	1.00
	D_1	2.4	3.0	4.4	1.0	1.3	1.2	3.0	3.7	1.5	1.9	4.5	5.3	6.2	4.0	2.7	3.7	5.2	19.9	2.9	3.4
LNMDS (1D)	τ	1.00	0.99	0.97	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	0.99	0.96	1.00	1.00	1.00	0.98	0.93	1.00	1.00
	D_1	3.3	4.0	4.7	1.9	2.6	2.6	2.8	2.7	2.7	2.7	2.8	3.1	4.1	2.2	2.8	4.2	3.6	4.6	2.7	1.8
LNMDS (2D)	τ	0.99	0.97	0.94	0.98	1.00	1.00	0.97	0.97	0.92	1.00	1.00	0.97	0.94	1.00	1.00	0.99	0.92	0.81	0.98	0.99
	D_1	4.3	4.8	5.5	2.6	3.9	1.8	2.6	2.4	6.1	5.1	3.4	4.3	5.4	3.2	3.5	3.0	5.4	8.2	4.0	2.9

er, LNMDS consistently achieved more accurate recovery of sample positions for the $1 \times 0.33R$ coenoplanes with extremely skewed response surfaces and for the $1 \times 0.33R$ coenoplanes in which interaction between species produced more complex response shapes.

Some ordination configurations for $1 \times 0.33R$ coenoplanes are shown in Figs. 2 and 3. In each case, the samples were located on a regular 12×4 grid in the simulated environment space. Figure 2 illustrates the lack of a consistent advantage for either technique with symmetrical and mixed skewness models. When DCA gave worse fits than LNMDS (Fig. 2b, f), it was generally due to a compression of variation along the second gradient towards one end of the first gradient. For those
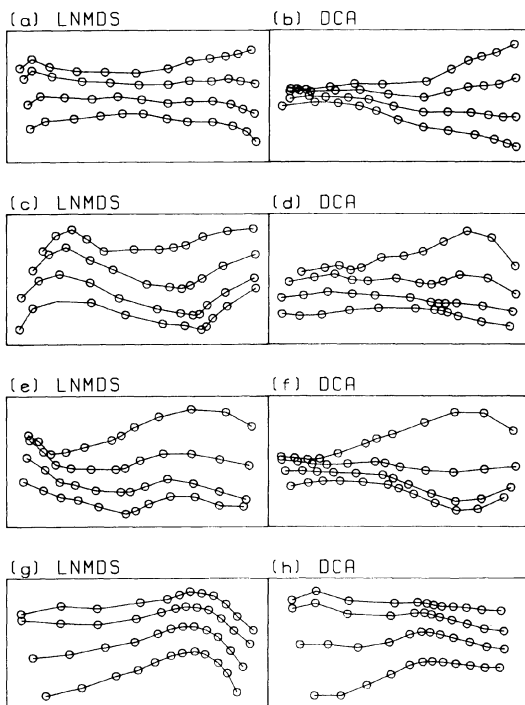
data sets where DCA achieved better results than LNMDS (Fig. 2d, h), the LNMDS configuration generally exhibited some systematic curvilinear distortion.

Figure 3 shows two examples of the consistently superior performance of LNMDS for $1 \times 0.33R$, extreme skewness and interaction models. The DCA solution for the extreme skewness model (Fig. 3b) once again shows compression at one end of the first gradient. For the interaction model illustrated (Fig. 3c, d), the DCA ordination is marred by curvilinear distortion at the left-hand side of the configuration.

The unequivocal superiority of LNMDS over DCA for $1 \times 1R$ coenoplanes is apparent from the results presented in Table 3. Irrespective of response surface shape, LNMDS always achieved lower Procrustean discrepancies than DCA, often markedly so. Some examples of ordination results for $1 \times 1R$ coenoplanes are shown in Figs. 4 and 5. For each of the models illustrated, the samples were located on a regular 7×7 grid in the simulated environment space.

Figure 4 shows the DCA and LNMDS ordinations of all three replicates with symmetrical, unimodal response surfaces. The regular grid was recovered reasonably well by DCA for the first replicate (Fig. 4b), but the DCA configurations for the other two replicates (Fig. 4d, f) are distorted. A peculiar feature of the distortion is the displacement of the samples in one corner of the grid to
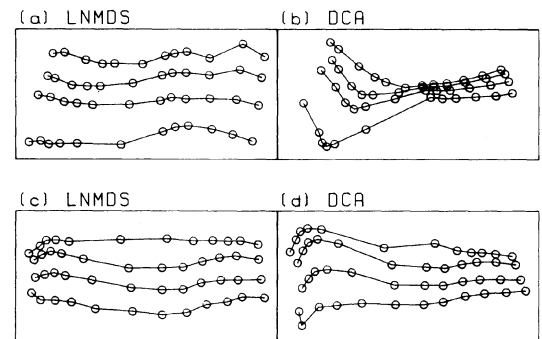

Fig. 2. LNMDS and DCA ordinations of some $1 \times 0.33R$ coenoplanes: (a) and (b) symmetric responses, replicate 1 (PCA-CS and PCoA ordinations of this model are shown in Fig. 1a, b); (c) and (d) symmetric responses, replicate 2; (e) and (f) mixed skewness, replicate 3; (g) and (h) mixed skewness, replicate 2. All models had no noise added and samples were arranged on a regular 12×4 grid.

Fig. 3. LNMDS and DCA ordinations of some $1 \times 0.33R$ coenoplanes: (a) and (b) extreme skewness, replicate 1; (c) and (d) interaction model, replicate 1. In both cases, samples were arranged on a regular 12×4 grid, and no noise was included.
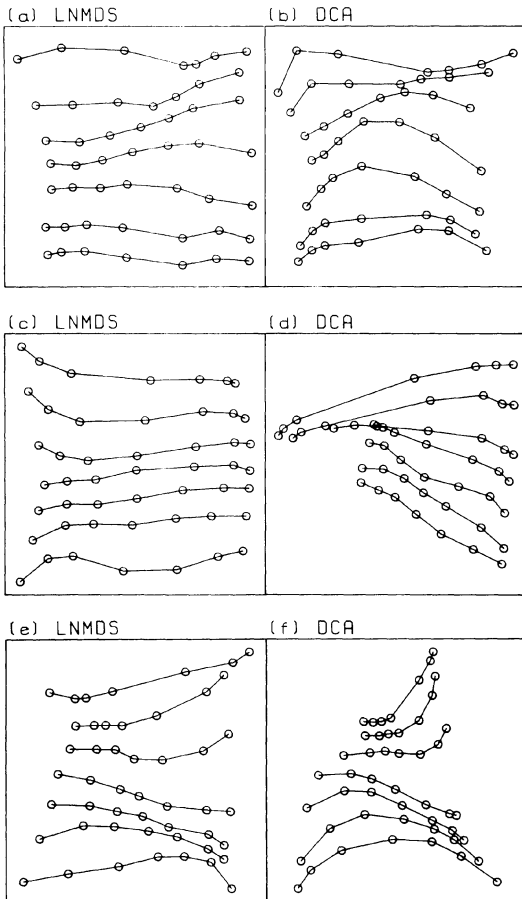
Table 3. Procrustean discrepancy values for ordinations of some two-dimensional models in experiments B2 and B3 (see Table 1). The effect of quantitative noise was only studied for the first replicate in each response shape category. Smaller values of D_2 indicate better performance.

Models with beta diversity = $1 \times 0.33R$

Ecological response function shape

Ordination method	Symmetrical					Mixed skewness					Extreme skewness					Interaction (skewed)				
	Replicate 1			2	3	Replicate 1			2	3	Replicate 1			2	3	Replicate 1			2	3
	Noise level (F)					Noise level (F)					Noise level (F)					Noise level (F)				
	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0
DCA D_2	6.3	6.4	6.4	5.3	7.3	7.1	7.6	9.3	8.2	6.5	9.4	9.2	8.7	8.3	6.5	5.8	6.1	7.5	6.6	5.4
LNMDS D_2 (2D)	4.3	4.2	12.2	7.2	5.7	8.0	9.0	11.1	9.7	6.0	5.8	6.4	6.8	7.3	5.1	5.1	5.7	7.1	5.0	4.5
LNMDS D_2 (3D)	5.3	4.7	13.9	8.0	6.8	11.1	11.4	13.0	10.5	6.4	5.9	6.2	7.6	13.3	5.9	6.2	7.3	6.9	5.4	8.2

Models with beta diversity = $1 \times 1R$

Ecological response function shape

Ordination method	Symmetrical					Mixed skewness					Extreme skewness					Interaction (skewed)				
	Replicate 1			2	3	Replicate 1			2	3	Replicate 1			2	3	Replicate 1			2	3
	Noise level (F)					Noise level (F)					Noise level (F)					Noise level (F)				
	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.2	0.0	0.0
DCA D_2	9.8	9.4	10.4	19.3	14.6	15.0	21.5	20.0	9.8	9.1	19.6	21.5	20.0	10.6	11.2	11.6	11.4	11.3	8.9	9.9
LNMDS D_2 (2D)	6.4	6.7	8.6	8.5	9.2	7.3	8.1	7.8	9.0	7.7	12.7	13.2	14.7	6.9	9.1	6.8	7.3	8.6	8.0	7.9
LNMDS D_2 (3D)	8.5	9.4	10.5	10.3	10.3	8.1	8.7	9.0	10.6	8.0	13.7	13.3	13.0	7.4	10.5	8.6	8.7	10.2	9.3	9.3

Fig. 4. LNMDS and DCA ordinations of replicate 1 × 1R coenoplanes with symmetric, unimodal response surfaces: (a) and (b) replicate 1 (PCA-CS and PCoA ordinations of this model are shown in Fig. 1c, d); (c) and (d) replicate 2; (e) and (f) replicate 3. In each case, samples were arranged on a regular 7 × 7 grid and no noise was added.
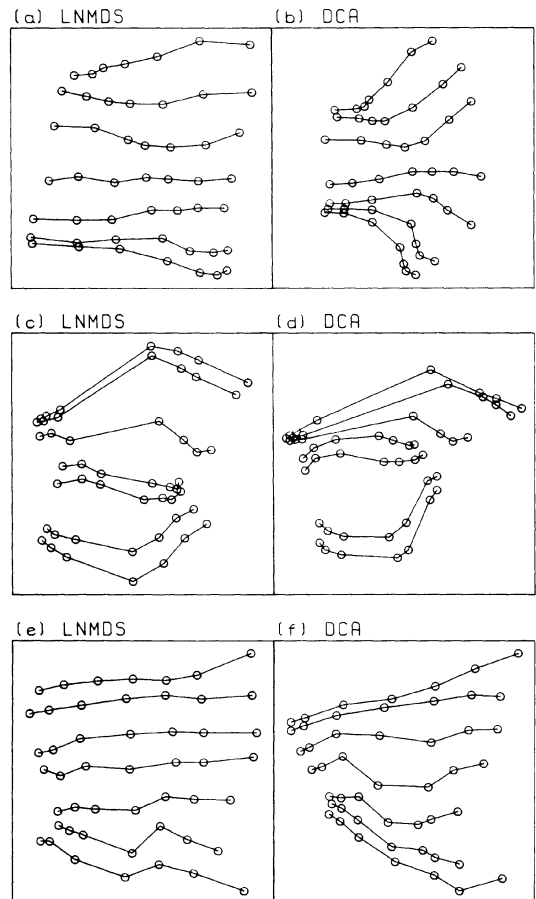
Fig. 5. LNMDS and DCA ordinations of some 1 × 1R coenoplanes: (a) and (b) mixed skewness, replicate 1; (c) and (d) extreme skewness, replicate 1; (e) and (f) interaction model, replicate 1. In each case, samples were arranged on a regular 7 × 7 grid and no noise was included.

form a narrow 'tongue' extending from one side of the configuration. The LNMDS ordinations (Fig. 4a, c, e) all display good recovery of the grid structure and there is no sign of the 'tongue' distortion produced by DCA.

Some examples of DCA and LNMDS ordinations of mixed skewness, extreme skewness and interaction coenoplanes with beta diversities of 1 × 1R are shown in Fig. 5. In each case the regular 7 × 7 grid is recovered better by LNMDS. The DCA ordination of the extreme skewness model

(Fig. 5d) shows a 'tongue' distortion similar to that observed for some of the symmetrical models. LNMDS corrects this distortion and achieves a tolerable reconstruction of the simulated grid (Fig. 5c). The performance of LNMDS on this data set was the worst for all of the regularly sampled, noiseless, 1 × 1R coenoplanes examined.

LNMDS usually achieved better recovery of gradients when performed in the correct dimensionality. When offered an extra dimension, LNMDS tended to curve the gradient structure in the higher-

dimensional space, although the distortion was never as severe as that observed in linear ordinations (PCA, PCoA). Consequently, there was generally a deterioration in the accuracy with which the environmental configuration was recovered on the first one or two axes (see Tables 2, 3, 4). The problem was more severe for coenocline models and coenoplanes in which one gradient was longer than the other. For most of the $1 \times 1R$ coenoplanes studied, the first two dimensions of a three-dimensional LNMDS still represented sample positions more accurately than the corresponding DCA solution (Table 3).

Quantitative noise

Both DCA and LNMDS displayed a considerable degree of resistance to the addition of quantitative noise. The effect of noise was to introduce random displacements of points in the ordination configurations, without altering the overall form of the configurations derived for the corresponding noiseless data sets (cf. Gauch *et al.*, 1981). The introduction of noise did not alter the relative performance of DCA and LNMDS, as described in the previous section (see Tables 2, 3). Despite the addition of noise, the overall structure of the simulated environment space is recovered remarkably well.

Qualitative noise

The effect of qualitative (presence-absence) noise on ordination performance was severe, especially for models in which beta diversity was low. Both DCA and LNMDS were unable to achieve an acceptable recovery of sample positions along simulated gradients with beta diversities of less than $1R$ in the presence of qualitative noise. (It is probable that the levels of qualitative noise applied in this study were unrealistically high: many data sets contained samples with fewer than five species, against a mean of ca. 25 in the noiseless data).

Sampling pattern

Experiments A4 and B4 (Table 1) revealed marked differences between DCA and LNMDS in their sensitivity to the pattern of sampling in the simulated environmental space. For unidimensional models,

the effects of variation in sampling pattern on both DCA and LNMDS were relatively minor. The relative performance of the techniques with regularly sampled coenoclines was maintained with the other sampling patterns. The results do not agree with those of Mohler (1981), who reported better gradient recovery by DCA when samples were concentrated towards the extremes of the gradient.

The results for $1 \times 1R$ coenoplanes were quite different, illustrating the danger of assuming that phenomena observed for unidimensional models should generalize to the multidimensional case. Some performance statistics for experiment B4 are given in Table 4. DCA displayed more variation in performance between sampling patterns than did LNMDS. LNMDS achieved better recovery of sample configurations than did DCA for all data sets, with the exception of two models with cross-shaped sampling patterns. DCA performed particularly poorly when samples were randomly distributed.

Of special interest are the results for $1 \times 1R$ coenoplanes in which samples were confined to a restricted region of the simulated environmental space. Two kinds of restricted sampling pattern were studied. In the first, sites were confined to a T-shaped region (Fig. 6a), so that the second gradient was only expressed towards the lower extreme of the first gradient. The second arrangement was cross-shaped, with the second gradient being expressed only near the centre of the first (Fig. 7a). Neither DCA nor LNMDS gave satisfactory results under this type of sampling, although LNMDS was generally more successful.

Figure 6 shows the DCA and LNMDS ordinations of the symmetric response coenoplane sampled by a T-shaped pattern. In the DCA configuration (Fig. 6c), the cross-bar of the T has been bent, giving the appearance of a long gradient with a shorter side branch. In addition, variation in the direction of the second gradient among samples forming the stem of the T has been suppressed. The LNMDS ordination (Fig. 6b) represents the T-shaped pattern reasonably well, although there is some curvilinear distortion of the stem of the T. Ordinations by DCA and LNMDS of the data set derived by cross-shaped sampling of the interaction coenoplane are shown in Fig. 7. In the DCA ordi-

Table 4. Procrustean discrepancy values for ordinations of some two-dimensional models in experiment B4 (see Table 1). All models had beta diversities of $1 \times 1R$ and no noise was added. Lower values of D_2 indicate better performance.

Ordination method	Ecological response function shape																	
	Symmetrical						Extreme skewness						Interaction (skewed)					
	Sampling Pattern						Sampling Pattern						Sampling Pattern					
	Regular	Random	Centre	Edges	T	Cross	Regular	Random	Centre	Edges	T	Cross	Regular	Random	Centre	Edges	T	Cross
DCA D_2	9.8	21.8	8.8	12.7	17.2	5.9	19.6	26.8	17.6	15.2	17.5	13.0	11.6	18.4	9.3	11.6	10.4	22.9
LNMDS D_2 (2D)	6.4	6.7	7.0	8.0	9.0	8.3	12.7	14.1	12.1	14.7	12.0	11.9	6.8	7.0	6.8	7.5	8.8	22.9
LNMDS D_2 (3D)	8.5	7.8	9.4	8.7	9.4	9.5	13.7	15.6	14.1	15.0	12.9	11.9	8.6	9.2	7.5	10.5	9.5	16.9

102

6 (a) SIMULATED   7 (a) SIMULATED

6 (b) LNMDS   7 (b) LNMDS

6 (c) DCA   7 (c) DCA
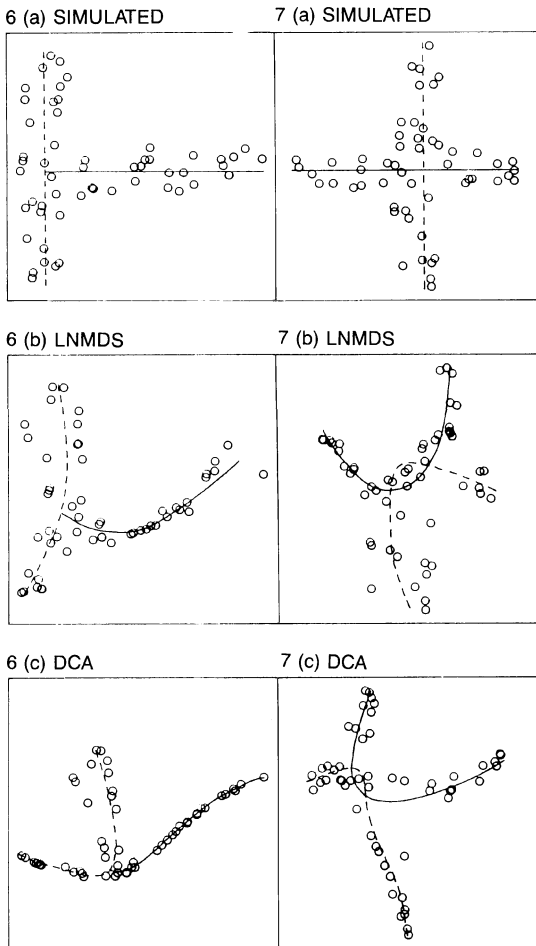
Fig. 6. Ordinations of a noiseless, 1 × 1*R* symmetric, unimodal model (replicate 1), with samples confined to a T-shaped region in environment-space: (a) sample arrangement in model; (b) LNMDS ordination; (c) DCA ordination. The solid and dashed lines indicate the direction of the first and second simulated gradients, respectively.

Fig. 7. Ordinations of a noiseless, 1 × 1*R* interaction model (replicate 1), with samples confined to a cross-shaped region in environment-space: (a) sample arrangement in model; (b) LNMDS ordination; (c) DCA ordination.

nation (Fig. 7c), both arms of the cross are bent through approximately 90° at their junction. One end of the arm marked by a dashed line extends into the third dimension of the ordination (not shown). The LNMDS ordination displays a similar

degree of distortion (Fig. 7b).

The poor recovery of T-shaped and cross-shaped coenoplane models by all techniques included in this study suggests the need for a greater emphasis in future studies on the implications for ordination of the shape of the underlying environmental space. In the past, it has been implicitly assumed that environmental space is completely (and often uniformly) covered by samples. However, it is often the case in nature that certain combinations of environmental conditions do not occur in a particular landscape (e.g., Austin *et al.,* 1984).

Trend in sample totals

The effect of systematic variation in sample totals was studied for unidimensional models only (Table 1, A5). Both LNMDS and DCA appeared to be relatively robust to two-fold linear and parabolic trends in sample totals. These results are preliminary and more work is required on this question, particularly with multidimensional models.

Conclusions

1. The linear ordination techniques PCA and PCoA are inappropriate for the purpose of indirect gradient analysis.

2. Gaussian ordination is very sensitive to quantitative noise and is not robust to departures from its assumed response model.

3. The currently popular DCA does not perform well with more complex response models and non-regular sampling schemes. The technique often produces distorted ordinations of 1 × 1*R* coenoplanes, even with symmetric, unimodal response surfaces.

4. Local non-metric multidimensional scaling (LNMDS), using the Bray-Curtis dissimilarity coefficient is the most robust and effective of the methods compared.

5. None of the techniques studied achieved satisfactory recovery of coenoplane models in which samples were restricted to T-shaped or cross-shaped regions of environment space.

Discussion

Linear ordination methods

This study confirms and extends the conclusions of previous work (see above) in showing the ineffectiveness of PCA due to curvilinear distortion (the 'type A' distortion of Orlóci, 1974). PCoA with the Bray-Curtis coefficient is similarly afflicted. The distortion is not confined to models with symmetric, unimodal responses, but also occurs with the more complex models examined. Gauch *et al.* (1981) refer to the published comments of J. C. Gower on a paper by Sibson (1972). According to Gower (see Gauch *et al.,* 1981) NMDS gives very similar results to metric scaling (i.e. PCoA), but at the expense of much more computation. However, in this study the two approaches only gave similar ordinations for models with very short gradients, where the relationship between Bray-Curtis dissimilarities and environmental distances was approximately linear. In all other cases, the LNMDS ordinations recovered the gradient structure much more effectively than PCoA.

It has been argued (Greig-Smith, 1980; Feoli & Feoli Chiapella, 1980; Van der Maarel, 1980) that the curvilinear distortion in PCA does not matter, provided that interpretation of PCA ordinations takes into account its probable occurrence. However, when there is more than one underlying gradient, it is very difficult to perceive that the points in the configurations such as shown in Fig. 1 fall on a surface irregularly curved into three or more dimensions, let alone to devise a procedure for mapping the points onto a plane.

Several authors have suggested that linear ordination methods are useful for indirect gradient analysis, provided that their application is restricted to data sets with low beta diversity (e.g., Austin & Noy-Meir, 1971). Unfortunately, it is often difficult to assess *a priori* whether the beta diversity of a data set is low enough for linear techniques to be applied with negligible distortion. In any case, when beta diversity *is* low, good recovery of the gradients can still be obtained using a robust non-linear technique, such as LNMDS: the monotonicity assumption of LNMDS embraces linearity as a

special case. There is little justification for the continued application of linear ordination methods to community data *for the purpose of indirect gradient analysis.* That is not to say that such methods are not useful for the analysis of other types of ecological data, for which the linear model is appropriate.

Gaussian ordination and other curve-fitting methods

In this study, Gaussian ordination (GO) was shown to lack robustness to departures from its assumed symmetric, unimodal response model, thus confirming the preliminary results of Austin (1976). Furthermore, GO is very sensitive to the addition of quantitative noise, much more so than DCA and LNMDS. This result is at variance with Gauch *et al.* (1974), who imply that GO is rather resistant to noise.

The program used for GO in this work (Gauch, 1979) produces only one ordination axis and is therefore inadequate for data sets with several underlying gradients. Extensions of GO to the multidimensional case are available, but there are theoretical and computational difficulties which remain largely unresolved. The approach of Ihm & Van Groenewoud (1975) makes unacceptably stringent assumptions about the response model. Orlóci (1978, 1980) has proposed an indirect algorithm, based on the application of metric multidimensional scaling to particular measures of compositional dissimilarity and ordination distance, derived under specific assumptions about the Gaussian species response model. Some initial simulation results (Fewster & Orlóci, 1983) suggest that the performance of the method is undesirably sensitive to the assumptions made in the derivation of dissimilarities and distances.

The maximum likelihood non-linear ordination method of Johnson & Goodall (1979; Goodall & Johnson, 1982) is similar in philosophy to GO, in that it attempts to derive a sample configuration which maximizes the fit of symmetric, unimodal response functions for each species. Simulation studies of the method have been restricted to models which conform with the assumed model

and no comparison has been made with other non-linear ordination techniques, such as LNMDS.

Further simulation studies are required to assess the *robustness* of variants of GO (and other 'curve-fitting' ordination methods) to variations in the species response model and their sensitivity to noise. The excellent performance of such methods with simulated data confirming with their assumed response models is no basis upon which to predict their effectiveness with real data.

Poor performance of DCA

There are at least two factors which may contribute to the lack of robustness and erratic performance of DCA observed in this study; (1) properties of the implied dissimilarity measure, and (2) the behaviour of the detrending and rescaling processes. Hill & Gauch (1980) claim that correspondence analysis, from which DCA is derived, 'makes no use of the concept of compositional distance'. This is untrue. Correspondence analysis can be formulated as a particular variety of principal coordinates analysis, in which compositional dissimilarity is measured using the Chi-squared distance metric and samples are weighted according to their totals (Chardy *et al.,* 1976). Correspondence analysis derives an ordination in which the distances between pairs of sample points are proportional to their Chi-squared distance values.

The appropriateness of Chi-squared distance as a measure of compositional dissimilarity in ecology may be questioned (Faith *et al.,* 1987). The measure accords high weight to species whose total abundance in the data matrix is low. It thus tends to exaggerate the distinctiveness of samples containing several rare species. Unlike the Bray-Curtis coefficient and related measures, Chi-squared distance does not reach a constant, maximal value for sample pairs with no species in common, but fluctuates according to variations in the representation of species with high or low total abundances. These properties of Chi-squared distance may account for some of the distortions observed in DCA ordinations.

DCA includes two empirical procedures which attempt *a posteriori* to rectify the curvilinear distortion of gradients: detrending and rescaling. Computational details are given by Hill (1979; see also Hill & Gauch, 1980). The procedures were apparently developed from a consideration of very simple models of species replacement along a single gradient. Both involve an arbitrary choice of the degree of segmentation of axes. From the results of this study, it appears that the effects of detrending and rescaling with non-trivial coenoplane models are not always desirable. The distortions of the underlying gradient structure in some DCA ordinations (e.g., Figs. 2b, 4d and 6c) may be attributable to the behaviour of either or both detrending and rescaling. For some of the models for which severe flattening of parts of the configuration occurred in DCA ordinations, several other DCA ordinations were performed by varying the number of detrending segments and the number of cycles of rescaling. None of these adjustments was successful in reducing the degree of distortion. In several published applications of DCA, (e.g., Robertson *et al.,* 1984; Gibson & Kirkpatrick, 1985) ordination configurations have flattened 'tongues', similar to that shown in Fig. 4d. These may be artifacts, due to the operation of detrending or rescaling.

Not all curvilinear structures which may appear in an ordination are distortions, arising from the non-linear relationship between dissimilarity and environmental distance. DCA has no way of distinguishing between 'horse-shoe' or 'arch' distortions and features of the environmental configuration which happen to be non-linear. There is a danger that DCA will introduce new distortions of its own. A good example is provided by the DCA ordination of the $1 \times 1R$ symmetric response coenoplane with a T-shaped sampling pattern, shown in Fig. 6c. The non-linear structure formed by the upright of the T and one side of the cross-bar has been flattened out, as if it were an arch distortion.

Reservations about the possible effects of the empirical adjustments in DCA have been expressed by some authors (e.g. Fewster & Orlóci, 1983). Nevertheless, DCA has become probably the most widely-used ordination technique for community data. A major contributing factor to the rapid acceptance of the method has been the distribution of

the program DECORANA (Hill, 1979), which is relatively easy to use and economical, in terms of both computing time and memory requirements. In the light of current results, the status of DCA as a satisfying solution to the problem of curvilinear distortion in ordination (cf. Gauch, 1982) must be questioned. Interpretation of DCA ordinations should take into account the possibility of artifactual distortions, due to the properties of the implied dissimilarity measure or the activities of detrending or rescaling.

The current results suggest a clear preference for LNMDS over DCA. DCA consistently outperformed LNMDS only for coenoclines with simple, unimodal response curves. On the basis of very limited simulations using Gaussian models, Orlóci et al. (1984) have also reported rather poor recovery of simulated gradients by DCA.

Gauch et al. (1981) compared DCA with several variants of NMDS, including the LNMDS technique examined here. They concluded that DCA was generally more successful than NMDS. However, their study considered only a small number of Gaussian response models, all with regular sampling patterns. No replication was apparently performed within each combination of model properties. The great variation in performance by DCA among replicate models, observed in this study (e.g. Tables 2 and 3; Fig. 4), highlights the danger of arriving at misleading conclusions if comparisons are restricted to a single replicate.

The use of DCA ordinations as initial configurations for LNMDS is another major factor distinguishing the present study from that of Gauch et al. (1981). They employed random starting configurations for LNMDS (using the program SIBSON) and apparently used metric scaling solutions (similar to PCoA) for the other variants of NMDS examined. The choice of a poor initial configuration can reduce the likelihood of NMDS achieving a solution with the best possible monotonic fit between ordination distance and dissimilarity. The iterative procedure may become trapped in a 'local optimum', where no small change in the configuration will decrease the stress, even though different con-

figurations exist with lower stress. It is possible that many of the NMDS ordinations obtained by Gauch et al. (1981) were local optima, resulting from the use of random starts or metric scaling solutions which exhibited severe curvilinear distortion.

Robustness of non-metric multidimensional scaling

Published applications of non-metric multidimensional scaling (NMDS) in ecology (in either the 'global' or 'local' form) are relatively rare (e.g., Prentice, 1977; Clymo, 1980; Field et al., 1982; Oksanen, 1983; Dargie, 1984), but in most cases the technique has been considered effective. Simulation studies based on Gaussian models have shown that NMDS can successfully recover gradients of high beta diversity, both in its 'global' (Fasham, 1977) and 'local' forms (Prentice, 1980). Austin's (1976) preliminary examination of some alternative response models provided an early indication of the relative robustness of 'global' NMDS.

This study has identified LNMDS, using the Bray-Curtis dissimilarity coefficient, as a robust technique for the analysis of community data when the aim is to recover the compositional dimensions associated with underlying environmental gradients. The relative merits of 'global' versus 'local' variants of NMDS and the possible effects of the choice of dissimilarity measure were not investigated. However, subsequent work has used a similar, but more extensive, simulation approach to compare the robustness of dissimilarity coefficients (Faith et al., 1987) and several forms of NMDS (Minchin, Faith & Belbin, unpublished). The results suggest a preference for 'local' over 'global' NMDS and the Bray-Curtis measure was among the most robust of the coefficients compared, in terms of its rank correlation with simulated environmental distance.

Non-metric multidimensional scaling is recommended as a robust technique for indirect gradient analysis which deserves more widespread use by community ecologists.

References

Austin, M. P., 1976. Performance of four ordination techniques assuming different non-linear species response models. Vegetatio 33: 43−49.

Austin, M. P., 1980. Searching for a model for use in vegetation analysis. Vegetatio 42: 11−21.

Austin, M. P., 1985. Continuum concept, ordination methods and niche theory. Ann. Rev. Ecol. Syst. 16: 39−61.

Austin, M. P., 1987. Models for the analysis of species' response to environmental gradients. Vegetatio 69: 35−45.

Austin, M. P., Cunningham, R. B. & Fleming, P. M., 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. Vegetatio 55: 11−27.

Austin, M. P. & Noy-Meir, I., 1971. The problem of non-linearity in ordination: experiments with two-gradient models. J. Ecol. 59: 763−773.

Beatty, S. W., 1984. Influence of microtopography and canopy species on spatial patterns of forest understorey plants. Ecology 65: 1406−1419.

Bradfield, G. E. & Scagel, A., 1984. Correlations among vegetation strata and environmental variables in subalpine spruce-fir forests − southeast British Columbia. Vegetatio 55: 105−114.

Brown, M. J., Ratkowsky, D. A. & Minchin, P. R., 1984. A comparison of detrended correspondence analysis and principal co-ordinates analysis using four sets of Tasmanian vegetation data. Aust. J. Ecol. 9: 273−279.

Chardy, P., Glemarec, M. & Laurec, A., 1976. Application of inertia methods to benthic marine ecology: practical implications of the basic options. Estuar. Coast. Mar. Sci. 4: 179−205.

Clymo, R. S., 1980. Preliminary survey of the peat-bog Hummell Knowe Moss using various numerical methods. Vegetatio 42: 129−148.

Dargie, T. C. D., 1984. On the integrated interpretation of indirect site ordinations: a case study using semi-arid vegetation in southeastern Spain. Vegetatio 55: 37−55.

Del Moral, R., 1980. On selecting indirect ordination methods. Vegetatio 42: 75−84.

Faith, D. P., Minchin, P. R. & Belbin, L., 1987. Compositional dissimilarity as a robust measure of ecological distance. Vegetatio 69: 57−68.

Fasham, M. J. R., 1977. A comparison of nonmetric multidimensional scaling, principal components analysis and reciprocal averaging for the ordination of simulated coenoclines and coenoplanes. Ecology 58: 551−561.

Feoli, E. & Feoli Chiapella, L., 1980. Evaluation of ordination methods through simulated coenoclines: some comments. Vegetatio 42: 35−41.

Fewster, P. H. & Orlóci, L., 1983. On choosing a resemblance measure for non-linear predictive ordination. Vegetatio 54: 27−35.

Field, J. G., Clarke, K. R. & Warwick, R. M., 1982. A practical strategy for analysing multispecies distribution patterns. Mar. Ecol. Prog. Ser. 8: 37−52.

Gauch, H. G., 1973. The relationship between sample similarity and ecological distance. Ecology 54: 618−622.

Gauch, H. G., 1979. Catalog of the Cornell ecology programs series. 11th ed. Ecology and Systematics, Cornell University, Ithaca, New York.

Gauch, H. G., 1982. Multivariate analysis in community ecology. Cambridge University Press, London and New York.

Gauch, H. G., Chase, G. B. & Whittaker, R. H., 1974. Ordination of vegetation samples by Gaussian species distributions. Ecology 55: 1382−1390.

Gauch, H. G. & Whittaker, R. H., 1972a. Coenocline simulation. Ecology 53: 446−451.

Gauch, H. G. & Whittaker, R. H., 1972b. Comparison of ordination techniques. Ecology 53: 868−875.

Gauch, H. G. & Whittaker, R. H., 1976. Simulation of community patterns. Vegetatio 33: 13−16.

Gauch, H. G., Whittaker, R. H. & Singer, S. B., 1981. A comparative study of nonmetric ordinations. J. Ecol. 69: 135−152.

Gibson, N. & Kirkpatrick, J. B., 1985. Vegetation and flora associated with localised snow accumulation at Mount Field West, Tasmania. Aust. J. Ecol. 10: 91−99.

Goodall, D. W. & Johnson, R. W., 1982. Non-linear ordination in several dimensions. A maximum likelihood approach. Vegetatio 48: 197−208.

Gower, J. C., 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325−338.

Greig-Smith, P., 1980. The development of numerical classification and ordination. Vegetatio 42: 1−9.

Greig-Smith, P., 1983. Quantitative plant ecology. 3rd ed. Blackwell, Oxford.

Hill, M. O., 1979. DECORANA − A FORTRAN Program for Detrended Correspondence Analysis and Reciprocal Averaging. Ecology and Systematics, Cornell University, Ithaca, New York.

Hill, M. O. & Gauch, H. G., 1980. Detrended correspondence analysis, an improved ordination technique. Vegetatio 42: 47−58.

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. J. Ed. Psych. 24: 417−441 & 498−520.

Ihm, P. & Van Groenewoud, H., 1975. A multivariate ordering of vegetation data based on Gaussian type gradient response curves. J. Ecol. 63: 767−777.

Johnson, R. W. & Goodall, D. W., 1979. Maximum likelihood approach to non-linear ordination. Vegetatio 41: 133−142.

Kruskal, J. B., 1964a. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29: 1−27.

Kruskal, J. B., 1964b. Nonmetric multidimensional scaling: a numerical method. Psychometrika 29: 115−129.

Loucks, O. L., 1962. Ordinating forest communities by means of environmental scalars and phytosociological indices. Ecol. Monogr. 32: 137−166.

Minchin, P. R., 1983. A comparative evaluation of techniques

for ecological ordination using simulated vegetation data and an integrated ordination-classification analysis of the alpine and subalpine plant communities of the Mt Field Plateau, Tasmania. Ph.D. thesis, University of Tasmania.

Minchin, P. R., 1987. Simulation of multidimensional community patterns: towards a comprehensive model. Vegetatio (in press).

Mohler, C. L., 1981. Effects of sample distribution along gradients on eigenvector ordination. Vegetatio 45: 141–145.

Noy-Meir, I. & Austin, M. P., 1970. Principal-component ordination and simulated vegetational data. Ecology 51: 551–552.

Oksanen, J., 1983. Ordination of boreal heath-like vegetation with principal component analysis, correspondence analysis and multidimensional scaling. Vegetatio 52: 181–189.

Orlóci, L., 1974. On information flow in ordination. Vegetatio 29: 11–16.

Orlóci, L., 1978. Multivariate analysis in vegetation research. 2nd ed. Junk, The Hague.

Orlóci, L., 1980. An algorithm for predictive ordination. Vegetatio 42: 23–25.

Orlóci, L., Kenkel, N. C. & Fewster, P. H., 1984. Probing simulated vegetation data for complex trends by linear and nonlinear ordination methods. Abstr. Bot. 8: 163–172.

Prentice, I. C., 1977. Non-metric ordination methods in ecology. J. Ecol. 65: 85–94.

Prentice, I. C., 1980. Vegetation analysis and order invariant gradient models. Vegetatio 42: 27–34.

Robertson, P. A., MacKenzie, M. D. & Elliot, L. F., 1984. Gradient analysis and classification of the woody vegetation for four sites in southern Illinois and adjacent Missouri. Vegetatio 58: 87–104.

Rotenberry, J. T. & Wiens, J. A., 1980. Habitat structure, patchiness and avian communities in North American steppe vegetation: a multivariate analysis. Ecology 61: 1228–1250.

Schönemann, P. H. & Carroll, R. M., 1970. Fitting one matrix to another under choice of a central dilation and a rigid motion. Psychometrika 35: 245–255.

Sibson, R., 1972. Order invariant methods for data analysis. J. Roy. Statist. Soc. B. 34: 311–349.

Swan, J. M. A., 1970. An examination of some ordination problems by use of simulated vegetational data. Ecology 51: 89–102.

Van der Maarel, E., 1980. On the interpretability of ordination diagrams. Vegetatio 42: 43–45.

Van der Maarel, E., Boot, R., Van Dorp D. & Rijntjes, J., 1985. Vegetation succession on the dunes near Oostvoorne, The Netherlands; a comparison of the vegetation in 1959 and 1980. Vegetatio 58: 137–187.

Walker, J. & Peet, R. K., 1983. Composition and species diversity of pine-wiregrass savannas of the Green Swamp, North Carolina. Vegetatio 55: 163–179.

Whittaker, R. H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecol. Monogr. 30: 279–338.

Whittaker, R. H., 1967. Gradient analysis of vegetation. Biol. Rev. 42: 207–264.

Whittaker, R. H. & Gauch, H. G., 1978. Evaluation of ordination techniques. In: R. H. Whittaker (ed.), Ordination of plant communities, pp. 227–336. Junk, The Hague.

Young, F. W. & Lewyckyj, R., 1979. ALSCAL-4 User's Guide. Data Analysis and Theory Associates, P.O. Box 446, Carrboro, North Carolina.

A hierarchical consideration of causes and mechanisms of succession

S. T. A. Pickett[1], S. L. Collins[2] & J. J. Armesto[3]
Department of Biological Sciences and Bureau of Biological Research, Nelson Laboratories, Rutgers University, New Brunswick, NJ 08903, USA
[1]*Present address: Institute of Ecosystem Studies, New York Botanical Garden, Cary Arboretum, Box AB, Millbrook, NY 12545, USA*
[2]*Department of Botany and Microbiology, University of Oklahoma, Norman, OK 73091, USA*
[3]*Laboratorio de Sistematica y Ecologia Vegetal, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile*
Keywords: Causal hierarchy, Mechanism, Model, Succession

Abstract

Questions of successional pattern and causality have been central concerns in vegetation ecology. In this paper we address the limits of the overextended models of Connell and Slatyer by discussing problems encountered in field tests. To help prevent such problems, we define the essential concepts needed to understand succession: pathway, cause, mechanism, and model. We then suggest a more complete enumeration of successional causes, and place them in a three-level hierarchy. The highest level in the hierarchy defines the general and universal conditions under which succession occurs: (1) availability of open sites, (2) differential availability of species, and (3) differential performance of species at the site. To provide a more detailed understanding of succession, each of these causes is decomposed into ecological processes. A further decomposition results in the third level of the hierarchy, which is required to elucidate the mechanisms of succession at particular sites and to make detailed predictions. The hierarchy allows the appropriate causes to be chosen to answer questions about succession at the desired level of generality or level of organization. Recognizing the appropriate level(s) in the hierarchy is critical for the successful explanation of succession, design of experiments, statement of predictions, construction of models and development of general theory.

Introduction

Succession has commanded much of the attention of plant ecologists since the inception of the discipline. The majority of that effort has focused on determining the patterns of vegetation change through time. The study of mechanisms has been slower to develop due to the long time periods involved, but also due to the dominance of theories that gave preeminence to the climax, to stage-wise turnover, and to facilitative interactions (e.g. Clements, 1916). In addition, the causes of succession

are complex, and this has hindered the development of a comprehensive body of theory.

Connell & Slatyer (1977) provided an antidote to several of the critical problems that had beset the study of successional causes. For example, they promoted an experimental approach to determining the mechanisms of succession and the view that succession can be the outcome of several mechanisms. However, application of the ideas of Connell & Slatyer (hereafter denoted C & S) has been problematical. Here we analyze these difficulties and clarify important concepts needed in the study of succession. (Occasional reexamination of the fundamental concepts in vegetation was long ago advised by Cooper, 1926). We then provide a causal hierarchy as a framework for the study of succession.

Basic definitions

Four concepts are needed to describe, explain or predict aspects of succession.

(1) *Pathway*: The temporal pattern of vegetation change. A pathway may show changes in community types, system states, populations, or other parameters of vegetation. A variety of diagrammatic pathways presented by Horn (1981) and the analysis by Boerner (1985) exemplify the idea. A pathway for a particular succession appears in Fig. 1.

(2) *Cause*: An agent, circumstance or action responsible for successional patterns. Mechanisms, defined below, are one sort of cause. Complexity in successional causality will be illustrated and organized in the hierarchy introduced later.

(3) *Mechanism*: An interaction that contributes to successional change. A mechanism of succession, as an agent of change, is thus an efficient cause in the classical sense.

An important caveat in using the term 'mechanism' is that the specific interaction that is termed a mechanism depends on the level of organization a study addresses (Allen & Starr, 1982; Allen, 1987). At the community level, a mechanism may be a general ecological process or interaction, such as competition, establishment, or mutualism. Importantly, Connell & Slatyer's (1977) terms facilitation, tolerance and inhibition fit this description. But when considered at lower levels or organization, such as the population or individual levels, general processes must be decomposed into more detailed mechanisms. For example, 'competition' can be divided into mechanisms of differential resource uptake, stress effects on resource use, alterations in growth rates and allocation, and subsequent changes in

resource availability. This example illustrates that what is called a mechanism on one level of organization can be considered an effect on another level. It is thus important to specify on what level of organization a mechanism is being sought. Such specification may avoid unproductive arguments, since ecological studies usually involve several levels of organization.

(4) *Model*: A conceptual construct to explain successional pathways or predict the course of particular successions by combining various mechanisms, specifying the relationship between the mechanisms and the pathway as well as one another. The general systems model of succession developed by MacMahon (1981), models for specific seres (Noble & Slatyer, 1980; Armesto & Pickett, 1985a), and simulation models (Shugart, 1986) all serve as examples (Fig. 1).Note that the phrase 'C & S model' refers specifically to the 'models of mechanisms' (their phrase).

Misapplication of Connell and Slatyer's models

Connell & Slatyer's (1977) paper (hereafter denoted C & S) has been the focus of some studies (e.g., Sousa, 1979; Hils & Vankat, 1982; Turner, 1983; Harris *et al.*, 1984), and has been a point of reference in many others (e.g. Keever, 1979).

The use of C & S in the literature is often broader than intended (e.g. Harris *et al.*, 1984; Kruger, 1983; see Breitburg, 1985). Fig. 1 in C & S, entitled 'models of the mechanisms of succession' summarizes their focus on mechanisms of species turno-
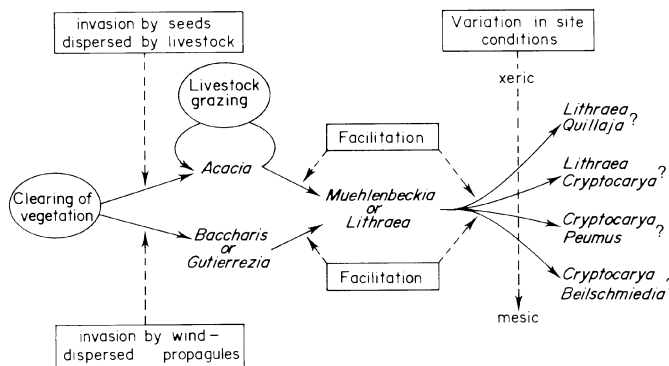

Fig. 1. A model of succession in Chilean matorral, composed of a pathway, indicated by solid arrows connecting stages labeled by their dominant genera. The mechanisms and causes of the succession are shown in boxes or circles, and their point of action in the pathway is shown by dashed arrows. (The difference between boxes and circles is unimportant for the current discussion). After Armesto & Pickett (1985a). Reproduced with permission of Revista Chilena de Historia Natural.

ver. The three models of C & S should be applied to specific plant-by-plant interactions (including the influence of mutualists and consumers) that result in successional turnover. They should not be applied to entire successions. Certainly, the C & S models are not alternative hypotheses about entire seres; attempting to 'test' them in that way is not likely to be productive (Quinn & Dunham, 1983). It is possible to analyze specific mechanisms of species replacement in succession within the broad categories of facilitation, tolerance and inhibition.

Another major limitation of the C & S models is the focus on species entry into the succession. Mechanisms by which species persist and mechanisms by which they yield space are not explicitly considered in the C & S models, although these events can be inferred from much of the rest of their discussion. The broad survey of successional causes by Clements (1916) provides the most inclusive framework for any discussion of successional mechanisms (Miles, 1979; MacMahon, 1981). Five non-teleological causes of succession appear in Clements' list: (1) Disturbance opening a site ('Nudation'), (2) Migration of propagules to the site, (3) Establishment of species at the site ('Ecesis'), (4) Interaction of organisms, (5) Alteration of the site by the organisms ('Reaction').

The nature of disturbance must be considered when succession is to be understood. C & S discuss disturbance in relation to community stability, but variations in features of disturbance do not discriminate among their three models of succession.

Migration is considered as a given or controlled factor by C & S, because they adopt an experimental approach. In field experiments the outcome of species removals or additions may depend in subtle ways on the presence or absence of species other than those being manipulated (Quinn & Dunham, 1983). In any event, understanding of the entire process of succession must include understanding the dynamics of invasion.

The C & S models are differentiable by the remaining three Clementsian causes. Establishment processes are key features of the three models. The obligate order of establishment is critical for separating the facilitation model from the other two. Finegan (1984) notes that facilitation may not require ordered establishment (i.e., relay floristics), but that refinement post dates the C & S models. Interaction among the species also differentiates the models. In the inhibition model, no invader can overcome the suppression of the initially established community, while in the tolerance and facilitation models, invasion is not depressed by the presence of the established plants. Finally, alteration of the environment by established plants has no impact in the tolerance model, although it does in the other two.

The examples to follow will consider both conceptual and practical problems in the application of the three mechanisms referred to in the C & S models. We will restrict our discussion to examples of inhibition. The complexity of all three mechanisms of species invasion are considered elsewhere (Finegan, 1984; Pickett *et al.*, in prep.).

A critical problem in interpreting information from field experiments designed to assess the three mechanisms arises because succession is fundamentally a plant-by-plant replacement process (Horn, 1976; Peet & Christensen, 1980) and interactions at the scale of neighborhoods are significant (Aarssen & Turkington, 1985a, b). Thus experiments focused on the entire assemblage of plants may fail to discriminate clearly among mechanisms (Breitburg, 1985). The clearest focus for discriminating among the mechanisms is at the level of individual replacement. Uncertainty resulting from confounding these scales appeared in the experiments of Hils & Vankat (1982) who were not able to confidently discriminate between the tolerance and inhibition mechanisms, although their results did contradict the predictions of the facilitation mechanism. Hils & Vankat (1982) suggested that different mechanisms may act simultaneously or consecutively. This is confirmed by Breitburg's (1985) analysis of marine cases and subsequent experiments on the fine scale (Armesto & Pickett, 1985b).

The role of species of *Rhus* in the invasion of trees into grasslands or oldfields (Petranka & McPherson, 1979; Werner & Harbeck, 1983), provides another example of difficulty in interpreting the inhibition mechanism. *Rhus typhina* increases the survivorship of trees invading a Michigan oldfield by thinning the dense herb cover that had formerly inhibited the trees (Werner & Harbeck, 1983). Thus, the interaction between *Rhus* and trees would be labeled facilitation. However, as Breitburg (1985) points out for marine cases, such interactions are asymmetrical. From the point of view of the grasses, the situation is inhibitory. Furthermore, if an experiment were done late in the interaction among the *Rhus*, grasses and trees, when the grasses had decreased substantially in cover, the trees would likely be released as a result of removing *Rhus*. The interaction would then be labelled inhibitory to the growth of tree seedlings. In order to fully understand the dynamics of tree invasion in an oldfield the mechanisms of invasion, persistence and mortality of the interacting species need to be known. Classifying the interaction as one of the three alternative types discerned by C & S may leave much unlearned about the interactions in a particular sere (Finegan, 1984).

Gap phase dynamics, common in mesic forests (Brokaw, 1985; Runkle, 1985) offer additional examples of the problems of interpretating the inhibition mechanism. Setting aside the problem of asymmetry in interactions, ascent of late

successional trees into the canopy depends not only on disturbance removing the inhibitory canopy individuals, but also on the tolerance of the late successional seedlings or sapling trees (e.g. Canham & Marks, 1985). Consequently, part of the entire interaction involves more than one mechanism (see Finegan, 1984).

For some questions concerning successional mechanisms and causes, it will be necessary to go beyond the particular C & S models.

A causal hierarchy

We suggest that the causes of succession be arrayed in a hierarchy for several reasons. A hierarchical arrangement spans all levels of inquiry. At the most inclusive level, the hierarchy presents a general explanation of succession, while at the most specific level, it comprises specific predictive factors. The hierarchy also encompasses all causes that operate at all levels of ecological organization. The specific research questions and objectives will determine where in the hierarchy causes are examined. Presentation of the hierarchy is not an admonition to incorporate all possible mechanisms or causes of succession in any particular study. Rather, it allows studies of part of the range of causes to be put in context, and guides the choice of interactions and constraints that must be considered to answer a given question.

The hierarchy has three levels. The highest, most general level is composed of the answers to the question, What causes succession? The answers that generally apply to any situation are that: (1) open sites become available; (2) species are differentially available to an open site; and (3) species behave differentially at the site. In the next level of the hierarchy, the three general causes of succession are divided into ecological processes or relationships. The first general cause, site availability, is determined by disturbance. The second differential species availability, is a function of the processes of dispersal and the dynamics of the propagule pool. The third, differential species performance, can be broken down into relations of the following: (1) resource availability; (2) ecophysiology; (3) life history strategy; (4) stochastic environmental stress through the sere; (5) competition; (6) allelopathy; and (7) herbivory and predation.

Each one of the processes or phenomena at the intermediate level in the hierarchy can in turn be understood in greater detail by examining the factors that determine its outcome and impact in a particular succession. It is these specific factors of the lowest, most detailed level of the hierarchy, that must be assessed or modeled to make specific predictions about the course of succession at a particular site. Similarly, certain of these factors need to be known to explain fine scale variation in succession. See Table 1.

We choose disturbance because for the purpose of illustration its importance in determining the course of succession has often been neglected (Vogl, 1980; Vitousek & White, 1981). The characteristics of disturbance requiring attention in understanding the course of successions include these (Sousa, 1984; White & Pickett, 1985): (1) severity; (2) size and shape; (3) timing relative to season, succession and past disturbance; and (4) spatial distribution of disturbed patches. The severity of disturbance is a measure of its impact on the vegetation (White & Pickett, 1985). To what extent the existing community is opened by disturbance determines for which potential colonists the new environment is suitable. It also determines whether vegetative or sexual propagules survive the disturbance and can contribute to the successional community (Noble & Slatyer, 1980). In some systems, e.g. mediterranean-type vegetation, the dynamics may be completely dependent on the predominant mode and impact of disturbance (e.g. fire in matorral, Armesto & Pickett, 1985a). Size of the area opened by disturbance will affect the environment of the site (Runkle, 1985; Denslow, 1980). Shape will additionally affect the physical environment, but will also influence the pattern of invasion of the site.

In addition to the simple physical characteristics of individual disturbances, the relationship of modes of disturbance to one another (e.g., Collins & Barber, 1986) and to other environmental and organism characteristics is important. The timing of the disturbance relative to season may influence its impact on the structure of the vegetation, the resources that are made available, the species that are particularly susceptible to its impact, and the suite of species that are potential immediate colonists (Keever, 1979). The timing of the disturbance relative to the successional status of the com-

Table 1. A hierarchy of successional causes. The highest level of the hierarchy represents the broadest, minimal defining phenomena. The intermediate level contains the mechanisms of change or causation of the highest level. The lowest level consists of the particular factors that determine the outcome of the intermediate-level processes, and are discernible or quantifiable at specific sites. Whether a particular process or factor advances or slows succession must be determined experimentally in specific instances or by generalization among comparable cases. Other processes or factors may be recognized in specific situations. For simplicity, interactions among factors at each level are not shown.

General causes of succession	Contributing processes or conditions	Defining factors
Site availability	Coarse-scale disturbance	Size, Severity, Time, Dispersion
Differential species availability	Dispersal	Landscape configuration
	Propagule pool	Dispersal agents, Time since disturbance, Land use
Differential species performance	Resource availability	Soil conditions, Topography, Microclimate, Site history
	Ecophysiology	Germination requirements, Assimilation rates, Growth rates, Population differentiation
	Life history strategy	Allocation pattern, Reproductive timing, Reproductive mode,
	Stochastic environmental stress	Climate cycles, Site history, Prior occupants
	Competition	Presence of competitors Identity of competitors Within-community disturbance Predators and herbivores Resource base
	Allelopathy	Soil characteristics, Microbes, Neighboring plants,
	Herbivory, disease and predation	Climate cycles, Consumer cycles, Plant vigor, Plant defense, Community composition, Patchiness

munity will also have important implications for subsequent succession. Timing of disturbance must also be considered in relation to the life histories of species in the community (Armesto & Pickett, 1985b). Disturbance may have different effects based on senescence, reproductive or architectural status of species in the community.

The frequency of disturbance in a landscape will likely influence the rate and course of succession through influences on the species pool, dispersal of species through the landscape, and configuration of patches of various successional ages (Pickett & Thompson, 1978; Forman & Godron, 1981). This aspect of the analysis of disturbance as a driving process of succession points out an important characteristic of disturbance that is helpful in application of the idea in general. It may apply to other components of the detailed level of the causal hierarchy as well. Each defining factor can be applied at a variety of spatial scales. For example, disturbance can be considered as a discrete event. In that case, a particular community is affected at a specific time. However, individual disturbances are a part of a disturbance regime that is discernible at coarser spatial and temporal scales (Levin & Paine, 1974). At the scale of entire landscapes, changing the disturbance regime may alter the dynamics of the successions that occur there. It is important to keep the distinction between disturbance regime and individual disturbances in mind to avoid confusion. Altering the magnitude of individual disturbances may have different effects on communities than altering the frequency and clumping of disturbances of various types and magnitudes in a landscape. In the first case, a disturbance is altered, and the response of one to several individuals is affected, whereas in the second case, the entire disturbance regime is altered and assemblages or landscapes may be affected.

Using the hierarchical array of successional causes proposed here, and recognizing the impact of scale on individual components of the hierarchy should enhance insight into the relationship of different causes to one another and lead to increased precision making predictions to be tested experimentally.

114

References

Aarssen, L. W. & Turkington, R., 1985a. Vegetation dynamics of neighbour associations in pasture-community evolution. J. Ecol. 73: 585–603.

Aarssen, L. W. & Turkington, R., 1985b. Biotic specialization between neighbouring genotypes in *Lolium perenne* and *Trifolium repens* from a permanent pasture. J. Ecol. 73: 605–614.

Allen, T. F. H., 1987. Hierarchical complexity: a non-euclidean conception of the data space. Vegetatio 69: 17–25.

Allen, T. F. H. & Starr, T. B., 1982. Hierarchy: Perspectives for ecological complexity. University of Chicago Press, Chicago.

Armesto, J. J. & Pickett, S. T. A., 1985a. A mechanistic approach to the study of succession in the Chilean matorral. Rev. Chil. Hist. Nat. 58: 9–17.

Armesto, J. J. & Pickett, S. T. A., 1985b. Experiments on disturbance in oldfield plant communities: Impact on species richness and abundance. Ecology 66: 230–240.

Boerner, R. E. J., 1985. Alternate pathways of succession on the lake Erie Islands. Vegetatio 63: 35–44.

Breitburg, D. L., 1985. Development of a subtidal epibenthic community: Factors affecting species composition and the mechanisms of succession. Oecologia 65: 173–184.

Brokaw, N. V. L., 1985. Treefalls, regrowth, and community structure in tropical forests. In: S. T. A. Pickett & P. S. White (eds), The ecology of natural disturbance and patch dynamics, pp. 53–69. Academic Press, New York.

Canham, C. D. & Marks, P. L., 1985. The response of woody plants to disturbance. In: S. T. A. Pickett & P. S. White (eds), The ecology of natural disturbance and patch dynamics., pp. 197–216, Academic Press, New York.

Clements, F. E., 1916. Plant succession: An analysis of the development of vegetation. Carnegie Inst. Pub. 242. Washington, D.C.

Collins, S. L. & Barber, S. C., 1986. The effects of disturbance on diversity in mixed-grass prairie. Vegetatio 64: 87–94.

Connell, J. H. & Slatyer, R. O., 1977. Mechanism of succession in natural communities and their role in community stability and organization. Am. Nat. 111: 1119–1144.

Cooper, W. S., 1926. The fundamentals of vegetation change. Ecology 7: 391–413.

Denslow, J. S., 1980. Patterns of plant species diversity during succession under different disturbance regimes. Oecologia 46: 18–21.

Finegan, B., 1984. Forest succession. Nature 312: 109–114.

Forman, R. T. T. & Godron, M., 1981. Patches and structural components for a landscape ecology. BioScience 31: 733–740.

Harris, L. G., Ebeling, A. W., Laur, D. R. & Rowley, R. J., 1984. Community recovery after storm damage: case of facilitation in primary succession. Science 224: 1336–1338.

Hils, M. H. & Vankat, J. L., 1982. Species removals from a first year oldfield plant community. Ecology 63: 705–711.

Horn, H. S., 1981. Succession. In: R. M. May (ed.), Theoretical ecology, 2nd ed., pp. 253–271. Sinauer Assoc., Sunderland, Mass.

Keever, C., 1979. Mechanics of succession on old fields of Lancaster County, Pennsylvania. Bull. Torrey Bot. Club 106· 229–250.

Kruger, F. J., 1983. Plant community dynamics in relation to fire. In: F. J. Kruger, D. T. Mitchell & J. V. M. Jarvis (eds), Mediterranean type ecosystems: the role of nutrients. Springer-Verlag, New York.

Levin, S. A. & Paine, R. T., 1974. Disturbance, patch formation, and community structure. Proc. Natn. Acad. Sci. USA. 71: 2744–2747.

MacMahon, J. A., 1981. Successional processes: Comparisons among biomes with special reference to probable role of animals. In: D. West, H. Shugart & D. Botkin (eds), Forest succession: concepts and application, pp. 277–321. Springer-Verlag, New York.

Miles, J., 1979. Vegetation dynamics. Chapman and Hall, London.

Noble, I. R. & Slatyer, R. O., 1980. The use of vital attributes to predict successional changes in plant communities subjected to recurrent disturbances. Vegetatio 43: 5–21.

Peet, R. K. & Christensen, N. L., 1980. Succession: a population process. Vegetatio 43: 131–140.

Petranka, J. W. & McPherson, J. K., 1979. The role of *Rhus copallina* in the dynamics in the forest-prairie ecotone in north central Oklahoma. Ecology 60: 956–965.

Pickett, S. T. A. & Thompson, J. N., 1978. Patch dynamics and the design of nature reserves. Biol. Conserv. 13: 27–37.

Quinn, J. F. & Dunham, A. E., 1983. On hypothesis testing in ecology and evolution. Am. Nat. 122: 602–617.

Runkle, J. R., 1985. Disturbance regimes in temperate forests. In: S. T. A. Pickett & P. S. White (eds), The ecology of natural disturbance and patch dynamics, pp. 17–33. Academic Press, New York.

Shugart, H. H., 1986. A theory of forest dynamics: The ecological implications of forest succession models. Springer-Verlag, New York.

Sousa, W. P., 1984. The role of disturbance in natural communities. Annu. Rev. Ecol. Syst. 15: 353–391.

Turner, T., 1983. Facilitation as a successional mechanism in a rocky intertidal community. Am. Nat. 121: 729–738.

Vitousek, P. M. & White, P. S., 1981. Process studies in succession. In: D. C. West, H. H. Shugart & D. B. Botkin (eds), Forest succession: concepts and application, pp. 266–276. Springer-Verlag, New York.

Vogl, R. J., 1980. The ecological factors that produce disturbance-dependent ecosystems. In: J. Cairns Jr. (ed.), The recovery process in damaged ecosystems, pp. 63–94. Ann. Arbor Sci. Publ., Ann. Arbor, Mich.

Werner, P. A. & Harbeck, A. L., 1982. The pattern of seedling establishment relative to staghorn sumac cover in Michigan old fields. Am. Midl. Nat. 108: 124–132.

White, P. S. & Pickett, S. T. A., 1985. Patch dynamics: An introduction. In: S. T. A. Pickett & P. S. White (eds), The ecology of natural disturbance and patch dynamics, pp. 3–13. Academic Press, New York.

The role of expert systems in vegetation science

I. R. Noble
Environmental Biology, Research School of Biological Sciences, PO Box 475, Canberra ACT 2601, Australia

Abstract

An area of artificial intelligence known as experts systems (or knowledge-based systems) is being applied in many areas of science, technology and commerce. It is likely that the techniques will have an impact on vegetation science and ecology in general. This paper discusses some of those impacts and concludes that the main effects will be in areas of applied ecology especially where ecological expertise is needed either quickly (e.g. disaster management) or across a wide range of ecological disciplines (e.g. land management decisions). Expert systems will provide ecologists with valuable tools for managing data and interacting with other fields of expertise. The impact of expert systems on ecological theory will depend on the degree to which 'deep knowledge' (i.e. knowledge based on first principles rather than on more empirical rules) is used in formulating knowledge bases.

Introduction

In the early 1980's the Japanese Ministry of Trade and Industry announced that it was supporting a wide ranging programme to develop the hardware and software resources for a 'fifth generation' of computers. This triggered a major increase in effort in an area of artificial intelligence known as expert systems as many other nations announced similar efforts. The impact that expert systems are likely to have on society − including all fields of science − over the next decade or so have been discussed by many authors (Weizenbaum, 1976; Feigenbaum & McCorduck, 1983; Duda & Short-cliffe, 1983; Lenat, 1984; Shannon *et al.,* 1985; Waterman, 1986). Here, I discuss some aspects of the application of expert systems (or knowledge-based systems as many prefer to call them) to ecology and especially that part of ecology that is involved in the prediction of the consequences of our actions in managing our environment.

What is an expert system?

There is a plethora of material describing expert systems in both the serious and popular scientific press and thus I will not attempt to review this material here. An expert system is a computer program capable of holding an apparently intelligent conversation with the user. It asks questions and the order of the questions changes with the responses given. Based on the knowledge held by the system and the answers to the questions, the system eventually states or validates a conclusion or decision and is able to explain how and why it reached this conclusion. Or more concisely it is a computer program designed to behave like professional experts.

An expert system can make use of a set of heuristic rules (i.e. 'rules of thumb') rather than a purely quantitative data base. It can be written in any of the common computer languages, despite some claims that that 'real' expert systems are written in LISP, PROLOG or a language similarly obscure to biologists. The program has two main components: a knowledge base, which is a series of often empiri-

cal rules or relationships, and an inference engine, which is code that is able to interact with the user and link the user's input to the knowledge base in order to answer some of the users' questions. There are advantages when writing expert systems in using a declarative (also called non-procedural) language such as PROLOG rather than an imperative (or procedural) language such as FORTRAN. Whereas in an imperative language the user must specify the steps to be taken in solving a problem (the algorithm), in a declarative language the user specifies only a description of the problem to be solved. The language itself provides the methodology to examine its data base and attempts to derive a solution. The main limitation in the development of declarative languages has been that they are slow to execute, but this is being overcome by advances in both computer software and hardware.

The major difference between an expert system and a process model, typical of the IBP programme and numerous other programmes, is best shown by example. Figures 1 & 2 show two versions of a section of a model of the damage to trees by fire in a forest community. The first describes the impact in strictly quantitative functions (i.e. a process model), while the second describes the same features in a mixture of quantitative and qualitative rules (production rules) more typical of a knowledge based system

In the process model knowledge about the system is encoded as mathematical formulae. The derivation of these formulae often require data that are difficult to obtain, or else 'guesstimates', which give the equations a false appearance of accuracy. In expert systems the knowledge is encoded as rules. There is usually some loss of accuracy, although more and more rules can be added to overcome this. However, the potential loss of accuracy

Fig. 1. A section of a process model of tree damage and mortality.

```
bark-thickness = FUNC1 (species)
bark-damage = FUNC2 (species, time-since-fire)
bark-remaining = bark-thickness − bark damage
effective-intensity = SEASONAL-EFFECT (season) * intensity
heating-effect = effective-intensity * FUNC3 (bark-remaining)
kill (species) = FUNC4.1 (heating-effect, species)
basal-sprout (species) = FUNC4.2 (heating-effect, species)
stem-sprout (species) = FUNC 4.3 (heating effect, species)
no-effect (species) = 1.0 − kill (species) − basal-sprout (species) − stem-sprout (species)
```

Fig. 2. Some production rules for tree damage and mortality

```
IF species is (Eucalyptus delegatensis OR E. fastigiata)
    THEN species-type is sensitive

IF intensity is no-scorch
    THEN no-effect

IF intensity is (crown-fire OR full-scorch) AND
  species-type is sensitive
    THEN all-killed

IF intensity is (crown-fire OR full-scorch) AND
  species-type is NOT (sensitive) AND
  EITHER (
  season is dry AND
    EITHER (
      dbh < limit
      THEN stem-sprout is uncommon
             basal-sprout is common
             killed is rare)
    OR
      dbh > = limit
      THEN stem-sprout is very − common
             basal-sprout is uncommon
             killed is practically-none )
    )
  OR (
  season is wet AND
    EITHER (
      dbh < limit
      THEN stem-sprout is common
             basal-sprout is common
             killed is practically-none )
    OR     (
      dbh > = limit
      THEN stem-sprout is very − common
             basal-sprout is rare
             killed is practically-none )
    )

IF previous-fire > 4 years-ago AND
  EITHER   (
    species is E. pauciflora       THEN limit is 35-cm
  OR
    species is E. dives            THEN limit is 20-cm
  OR
    species is E. dalrympleana THEN limit is 15-cm
  OR
      etc.
            )

IF previous-fire < = 4-years-ago AND
            etc.
```

is a problem only when we truly do know the system well enough the make precise predictions.

In the knowledge based model the biological system is described in terms of a series of production rules (i.e. IF situation true THEN this applies) and facts (e.g. the species *Eucalyptus delegatensis* is of

the sensitive species type). Proponents of expert systems argue that a knowledge based system more realistically mimics the human expert's use of knowledge. The example in Fig. 2 is not written in a particular language, but demonstrates one of the advantages of the declarative languages such as PROLOG in that the order of inclusion of rules is flexible — for example, the term 'limit' can be used in a rule which comes before the other rules which supply other essential information about limit. This makes it easy to update and modify such models.

When to use an expert system.

Table 1 shows a summary of the situations suitable for the application of expert systems (Forsyth, 1984). Do they apply to ecological work?

Diagnostic — Many ecological problems require that an item be classified or a choice be made between options. This is especially true in applied ecology where the question asked is often, 'which of a series of actions should be taken?' For example, should I conduct a prescribed burn in spring or autumn; should I burn today or not?

No established theory — I suspect that this is possibly an erroneous contrast, but nevertheless most ecologists would agree that much of ecology lacks a firmly established theory.

Data noisy and incomplete — No comment is needed.

Domain well bounded — This could be a problem in ecological applications because the domain under consideration when tackling ecological problems is usually very broad. Thus, expert systems will be able to provide advice on only small sections of wider problems. However, one of the long term goals of those working with expert systems is to link expert systems of different domains (e.g. Pereira *et al.,* 1984).

Human expertise scarce — This is true although unemployed postgraduates may disagree. However, many managers are making day-by-day decisions concerning ecological problems without the access

Table 1. A checklist of when to use knowledge based systems (based on Forsyth, 1984).

Suitable	Unsuitable
Diagnostic	Calculative
No established theory	Well established formulae
Data are noisy	Facts known precisely
Domain of knowledge well bounded	Domain not well bounded
Human expertise scarce and in demand	Expertise readily available

to ecological expertise which may be of assistance to them.

...and in demand — This is the real problem. Consultant ecologists are still relatively rare professionals and several factors are involved. First, many ecological problems do not require consultancy, but rather research. Thus the ecologists are called upon largely to provide data rather than to provide recommendations on decision making. That is, ecological expertise is too scarce in many situations for the consultancy role to have developed. Secondly, many managers consider themselves to be well acquainted with the numerous aspects of solving a land management problem and, thus, consider consulting a range of ecological specialists to be unnecessary. It is possible that expert systems may be developed to cover many of the specialist areas, thus making them more readily available to decision makers without the lengthy and expensive process of face to face consultation. These expert systems should be able to warn decision makers when more direct consultation is advisable.

Applications of knowledge based methodology in ecology have been limited largely to diagnostic problems. Starfield & Bleloch (1983) outlined an expert system to advise on prescribed burning. Noble (1985) has described an expert system that assists users to run a model which incorporates the vital attribute scheme (Noble & Slatyer, 1980) to predict vegetation change. Davis *et al.* (in press) have developed a knowledge based model which predicts aspects of fire intensity in tropical woodlands in northern Australia. This program forms part of a larger study to develop knowledge based systems to assist in the management of Kakadu National Park (Davis *et al.,* 1985; Walker *et al.,* 1985).

What will expert systems contribute to ecology?

I have already alluded to some of the impact that I think that the development of expert systems technology may have on ecology as a profession, but here I want to ask what might expert systems contribute to our understanding of ecological principles.

Expert systems may or may not contribute to ecological theory. If expert systems are used only to bring together a number of ecological rules-of-thumb and to package them in a way more readily available to a user then ecological understanding will advance very little. If, however, in our attempt to formulate the knowledge bases, we are forced to re-think the nature of ecological relationships then expert systems may have some impact. This is the basis of the debate about the role of 'deep' versus 'surface' knowledge in expert systems.

Deep versus surface knowledge

Most expert systems use rules with the form:

IF pattern THEN action

For example,

IF it is spring THEN don't burn

This sort of rule represents the surface knowledge of expertise in prescribed burning. The rule carries no insight into the processes that link the pattern 'it is spring' with the action 'don't burn'. It may be derived from simple empirical knowledge (i.e. experience) gathered over centuries.

The definition of deep knowledge is somewhat hazy but it is often described by example such as, deep knowledge includes the first principles to which a human expert will need to resort in order to solve difficult problems or to provide a creditable explanation of particular advice. More explicitly, deep knowledge often involves the use of rules of the form:

IF pattern-A & action THEN pattern-B will follow

For example,

IF spring foliage of species X is present & you burn THEN plant reserves will be depleted

IF it is spring & you deplete reserves of X

THEN summer growth will be poor

IF summer growth of X is poor THEN mortality increases

IF mortality of X increases & X is a desirable species THEN this is an unwanted result

IF result is unwanted THEN don't burn

The advantages of having deep knowledge built into the data base are several. If users are confronted by the rule

IF it is spring THEN don't burn

and they ask why, then the expert system can reply only

Don't burn in spring
BECAUSE it is spring

whereas with the deep knowledge rules the reply would be along the lines of

Don't burn in spring
BECAUSE it leads to an unwanted result
BECAUSE it leads to increased mortality of a desirable species
BECAUSE there has been poor summer growth
BECAUSE plant reserves were depleted in spring

Some users will then demand to know why poor summer growth leads to high mortality or why burning in spring depletes plant reserves, but there has to be a cut off point in any consultative system.

There is dispute among the expert systems' circles as to whether simple surface knowledge is sufficient to build useful expert systems. Most expert systems that have reached the production stage so far have been a collection of surface rules with a few additional rules to guide the inference engine of the expert system in efficiently consulting these rules. Chandrasekaran & Mittal (1983) have argued that, in medical diagnosis systems at least, it is not necessary to resort to deep knowledge to produce effective expert systems. Attarwala & Basden (1985) also discuss this topic in terms of causality and model detail based on their experiences in developing expert systems for corrosion control in industrial plants and tend to favour the use of deep knowledge.

The deep knowledge system will often allow

more generality in an expert system. For example, if we want to change a goal from that of protecting a species to eradicating it, then in the surface knowledge system we would have to change many of the rules relating to that species, whereas in the deep knowledge system we may need to change only the goal to be achieved, e.g. from achieving low summer mortality to achieving high mortality.

Deep knowledge will also provide more opportunities for interactions in ecological knowledge bases that combine several domains of expertise. For example, the above set of rules may interact with a set of rules in a domain dealing with the dynamics of a granivore. These rules may include:

If summer growth of X is poor THEN seed set is poor

IF seed set of X is poor THEN reproductive success of bird species Y is poor, etc.

Thus the two domains i.e. the impact of prescribed burning and the success of granivores are linked at this deeper level of knowledge.

If commercial pressure or simplistic expert systems engineering leads to ecological expert systems containing only surface knowledge then there is little possibility of a gain to ecological theory (as opposed to the practice of applied ecology). If we are forced to rethink and clearly state the inter-relationships between ecological processes in order to link them in a way that can provide advice (i.e. prediction) there is more to be gained.

Other impacts of expert systems

Starfield & Bleloch (1983), in the first paper on the application of expert system to ecology, suggested educational and communication advantages in building expert systems. These points are similar to the advantages listed for process modelling in the lead up to the IBP programme. Similarly, the claim that if expert systems theory forces ecologists to re-think ecological relationships then this will be of some benefit, is close to some of the early claims about process modelling — i.e. even if the models don't work we will still learn by building them.

It is sometimes argued that expert systems must be built by a new and special class of scientists known as knowledge engineers (Weiss & Kulikowski, 1984; Davis et al., in press). Thus we

have the equivalent to the 'synthesizers' of the IBP programme. At present the number of ecologists with skills appropriate to developing application packages based on expert systems are few and the tools crude. However, I doubt if this will remain the case as improved shells (software packages for developing expert systems) become available — a view supported by some of the expert systems workers themselves (e.g. Basden, 1983).

An aspect of expert systems technology that will have an impact on all professions that deal with large amounts of information, is their application to data base design. Commercial pressures are likely to lead to the development of relational data bases which use expert systems techniques to deduce additional connections between elements of the data base and to interact via a natural language interface. Like statistics, scuba tanks and word processors, these data bases will have an impact in the ecologist's ability to retrieve — and, hopefully, use — ecological information. Pereira et al. (1984) have begun a project in Portugal to develop a data base for environmental biophysical resource evaluation. In this they aim to bring the expertise of several disciplines, such as geology, hydrology, botany, zoology and microclimatology together in one expert system and to make this available to decision makers.

Another aspect of expert systems theory deals with systems that assist in the laborious tasks of interrogating experts and systematically organizing their knowledge. There are two broadly different approaches here. One is to aid the user in setting up the knowledge base. This involves assessing new rules againt those already in the knowledge base and warning of inconsistencies and incompleteness (e.g. omitting to tell the system facts that are so obvious to the expert that they are easily overlooked, such as that trees are usually much taller than grasses). TEIRESIAS is an example of such software (Davis & Lenat, 1982). The other approach is to provide the system with many case histories and algorithms for deducing, and even inducing, additional rules (e.g. Quinlan, 1983 for end games in chess). Most success in this area appears to be in diagnostic situations, e.g. an expert system to diagnose diseases of soy-bean (Michalski & Chilausky, 1980; Sammut, 1985). This learning approach is likely to have only limited application in ecology since we rarely have the large number of consistent

case histories to work with. However, relatively inexpensive software packages that implement some aspects of computer induction (e.g. 'EXPERT-EASE', & 'RULEMASTER; Waterman, 1986) are available for microcomputers and this may encourage ecologists to experiment with them (see McLaren, 1985 for an application of EXPERT-EASE).

Expert systems can also be used in training people. There have been some promising packages developed in this area but the subject falls outside this paper. However, the benefits to the ecological community of an expert system that guides the user through the complexities of experimental design, or of multivariate data analysis, should not be underestimated.

Discussion

Expert systems will have a major impact on applied ecology. Probably the most spectacular, and immediately challenging, problems will be in those aspects of environmental impact analysis dealing with disaster management. In these circumstances — e.g. wildfires or noxious spills — ecological information is needed quickly, it must be based on knowledge already held (i.e. there is no time for research), human experts may be unavailable and several domains of expertise may be involved. A high proportion of the first efforts to apply expert systems to ecology deal with aspects of fire management (Starfield & Bleloch, 1983; Davis *et al.* in press; Noble, 1985).

As applications increase, practitioners in the expert systems field will attempt to link the knowledge bases from disparate areas of ecology thus leading their more theoretically oriented colleagues to consider more carefully the unifying concepts of ecology. This claim was made for process modelling in its early days, but process modelling is a relatively restrictive tool. The necessity to quantify ecological knowledge was an insurmountable hurdle in many cases — or at least a useful and often valid excuse for not trying to achieve those unifying concepts. Expert systems don't carry the quantification restriction. They ask only that we can express our ideas in concise, logical rules.

References

Attarwala, F T. & Basden, A., 1985. A methodology for constructing expert systems. R&D management 15: 141–149.

Basden, A., 1983. On the application of expert systems. Int. J. Man-Machine Studies 19: 461–477.

Chandrasekaran, B. & Mittal, S., 1983. Deep versus compiled knowledge approaches to diagnostic problem solving. Int. J. Man-Machine Studies 19: 425–436.

Davis, J. R., Hoare, J. R. L. & Nanninga, P. M., 1985. The GEO-KAK fire behaviour and fire effects Expert System. In: J. Walker, J. R. Davis & A. M. Gill (eds), Towards an expert system for fire management at Kakadu National Park. CSIRO Div. Water and Land Resources Tech. Mem. 85/2. pp. 153–169. CSIRO, Canberra.

Davis, J. R., Hoare, J. R. L. & Nanninga, P. M., in press. Developing a fire management expert system for Kakadu National Park, Australia. J. Envir. Management.

Davis, R. & Lenat, D. B., 1982. Knowledge-based systems in artificial intelligence. McGraw-Hill, New York.

Duda, R. O. & Shortcliffe, E. H., 1983. Expert systems research. Science 220: 261–268.

Feigenbaum, E. A. & McCorduck, P., 1983. The fifth generation: artificial intelligence and Japan's computer challenge to the world. Michael Joseph, London.

Forsyth, R., 1984. The architecture of expert systems. In: R. Forsyth (ed.), Expert systems: principles and case studies. pp. 9–17. Chapman & Hall, London.

Lenat, D. B., 1984. Computer software for intelligent systems. Sci. Amer. 251: 152–160.

McLaren, R., 1985. Knowledge acquisition by computer induction. R&D Management 15: 159–166.

Michalski, R. S. & Chilausky, R. L., 1980. Knowledge acquisition by encoding expert rules versus inductive learning from examples: An experiment utilizing plant pathology. Int. J. Man-Machine Studies 12: 63–87.

Noble, I. R., 1985. Fire effects, vital attributes and expert systems. In: J. Walker, J. R. Davis & A. M. Gill (eds), Towards an expert system for fire management at Kakadu National Park. CSIRO Div. Water and Land Resources Tech. Mem. 85/2. pp. 96–103. CSIRO, Canberra.

Noble, I. R. & Slatyer, R. O., 1980. The use of vital attributes to predict successional changes in plant communities subject to recurrent disturbances. Vegetatio 43: 5–21.

Pereira, L. M., Oliveira, E. & Sabatier, P., 1984. Expert evaluation in logic of environmental resources through natural language. In: A. Elithorn & R. Banjeri (eds), Artificial and human intelligence. pp. 309–311. Elsevier Science Publishers.

Quinlan, J. R., 1983. Learning efficient classification procedures and their application to chess end games. In: R. S. Michalski, J. G. Carbonelli & T. M. Mitchell (eds), Machine learning. pp. 463–480. Tioga Press, Palo Alto.

Sammut, C. A., 1985. Concept development for expert system knowledge bases. Aust. Computer J. 17: 49–55.

Shannon, R. E., Mayer, R. & Adelsberger, H. H., 1985. Expert systems and simulation. Simulation 44: 275–284.

Starfield A. M. & Bleloch A. L., 1983. Expert systems: An approach to problems in ecological management that are difficult to quantify. J. Envir. Management 16: 261–268.

Walker, J., Davis, J. R. & Gill A. M. (eds), 1985. Towards an expert system for fire management at Kakadu National Park. CSIRO Div. Water and Land Resources Tech. Mem. 85/2. CSIRO, Canberra.

Waterman, D. A., 1986. A guide to expert systems. Addison-Wesley. Reading Massachusetts.

Weiss, S. M. & Kulikowski, C. A. 1984. A practical guide to designing expert systems. Chapman & Hall, London.

Weizenbaum, J., 1976. Computer power and human reason. W. H. Freeman & Co. New York.

123

Invasion models of vegetation dynamics

Robert Van Hulst*
Department of Biology, Bishop's University, Lennoxville, Qc, Canada J1M 1Z7

Abstract

Since communities change as a result of their successful invasion by new species it seems logical to attempt to predict future vegetation change by focussing on the invasion process. Several such invasion models are reviewed, and one particular model, based on dynamic game theory is developed further. This model can be used as an alternative to linear (e.g. Markov chain) models for the prediction of vegetation dynamics, and also to compare invasive abilities of species and resistance to invasion of communities. The main advantage of the model lies in the fact that it operates at a sufficiently high level of integration to allow for model calibration (in spite of the large number of underlying processes), and yet has an obvious population biological interpretation (in terms of the success of invading populations). The model can be calibrated using either time course data or experimental data, and it may be helpful in understanding what determines the fate of an invading population. It is used here to analyze two published vegetation ·dynamics data sets.

Introduction

Vegetation change can be (and has been) studied at different levels of integration. Community level analysis of the phenomenon has often led to biologically unrealistic and ineffective theories (cf. Drury & Nisbet, 1973; Miles, 1979). Many authors (e.g. Bazzaz, 1979; Noble & Slatyer, 1981; Sousa, 1979; Van der Valk, 1981) have, therefore, preferred a population level approach, thereby increasing our understanding of the processes involved, if not our ability to predict future vegetation change.

The models to be introduced here are based on an incremental approach: any new species that appears in the course of a succession has to fit into the framework of species already present and (or) prevailing abiotic conditions. Subsequently, the successful invader may then eliminate one or more

previously present species and change the abiotic conditions. To the extent that consistent combining rules between species exist, the species replacement process should in principle be predictable. While such predictability could, in theory, be derived from a low-level, mechanistic understanding of the crucial processes at work (competition, herbivory, plant induced environmental change, etc.), in practice such a detailed understanding is still remote, and may in fact never be attainable for many vegetation types.

Real plant communities are the result of processes of invasion and subsequent species assortment (Fowler & MacMahon, 1982). Such processes are open to experimental investigation: most plants can be transplanted relatively easily into various communities and the fate of these transplants can be monitored (Antonovics & Levin, 1980). What is required is a theoretical framework to study the fate of invaders and the invaded. Such a framework offers the attractive prospect of bridging the gap between population biology on one side (the in-

vader, its dispersal abilities, life cycle strategy, general and genetic demography), and community ecology on the other (competitive interactions, herbivory, community succession).

Invasion rates of species not yet (conspicuously) present in a community will generally be determined by four factors: i) the number of propagules already present and the rate of their arrival; ii) gap producing factors; iii) abiotic environmental factors; and iv) biotic factors. In what follows I shall concentrate on the effects of biotic factors, assuming that in many cases of interest the other factors will either not change dramatically over successional time or change in ways predictable from the changing species composition. This assumption may sometimes be unwarranted, (e.g. in the aquatic successions studied by Walker (1970), where the time scale approaches that of major climatic change) but we are left with enough cases where it undoubtedly can be justified.

Recent work in succession theory (e.g. Connell & Slatyer, 1977; Grime, 1979) has shown clearly the importance of inhibitory effects of an established community on potential invaders. Many communities appear to be remarkably resistant to invasion. Conversely, certain species are exceedingly good invaders in a wide range of community types (e.g. Van Hulst *et al.,* in press). What determines the resistance of a community to invasion, and the invasive ability of a species in a certain community? The two notions are difficult to quantify, because they extend over two levels of integration. By and large neglected by students of both population dynamics and community dynamics these properties nevertheless are basic to a proper understanding of the dynamics of the plant community. This again suggests the need for specific models to deal with invasion processes in ecology. Once such models have been developed invasive ability and resistance to invasion can be precisely defined and compared for different community types and species.

Invasion models

A number of questions have to be addressed before a mathematical model of a dynamic process can be formulated. These include the following: i) what is the time scale of interest? ii) what state variables are to be used? iii) is time taken to be continuous or discrete? and iv) is the model to be deterministic or stochastic? The models discussed below all operate at the successional time scale, glossing over short term ecological fluctuations, but treating species as fixed. The choice of appropriate state variables is more difficult to make. It may be argued that a dynamically sufficient description of community change has to include more than vegetation variables (Van Hulst, 1980). On the other hand, relevant environmental and animal population variables may be sufficiently predictable from vegetation variables to obviate the need for including them explicitly, especially at the longer, successional, time scale. The final answer to this question will be found in testing explicit models with different state variables. I will use continuous time deterministic models in what follows for reasons of mathematical expediency, but the sensitivity of the models to perturbations in the parameters will be briefly discussed.

Let us represent a plant community by a vector x containing the proportional cover, x_i, of each of its species, open ground being treated as merely another species. We are interested in deriving a dynamic for x. We do not know whether x is a dynamically sufficient descriptor, but let us assume it is. The dynamic behaviour of the community may then be represented as

$$\frac{dx}{dt} = f(x) \tag{1}$$

where x (x_i, $i = 1, \ldots n$), $x_i \geq 0$. Assume also that there is a stable equilibrium at $x = \hat{x}$ with all $x_i > 0$.

Following Allen (1976), Schaffer (1977), and Reed & Stenseth (1984) we now extend model (1) to include potential invaders into the community:

$$\frac{dx}{dt} = f(x, y)$$

$$\frac{dy}{dt} = g(x, y) \tag{2}$$

Here $y = (y_1, \ldots y_m)$ represents the cover of the invaders, assumed to be initially small. We also assume that $f(x,0) = f(x)$ and that $g(x,0) = 0$ (if there are no invaders model (1) applies).

Whether or not invasion is successful will depend on the stability of the equilibrium $x = \hat{x}$, $y = 0$ of model (2). If the equilibrium is stable y will return to 0 after the perturbation of initial colonization by y. If it is not stable at least some of the y_i may increase, possibly leading to subsequent changes in the x_i. To investigate the stability of the equilibrium of model (2) it is necessary to study the eigenvalues of the Jacobian matrix at the equilibrium $x = \hat{x}$, $y = 0$. This matrix has triangular form (Reed & Stenseth, 1984):

$$\mathbf{P} = \begin{bmatrix} \dfrac{\partial f}{\partial \mathbf{x}} & \dfrac{\partial f}{\partial \mathbf{y}} \\ \mathbf{O} & \dfrac{\partial g}{\partial \mathbf{y}} \end{bmatrix} = \begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{O} & \mathbf{R} \end{bmatrix} \qquad (3)$$

where the $\hat{}$ stands for evaluation at the equilibrium. Here:

$\mathbf{Q}$ is the community matrix (Levins, 1968) (interaction coefficients between species originally present)
$\mathbf{R}$ is the invasion matrix (Reed & Stenseth, 1984) (interaction coefficients between the invading species)
$\mathbf{S}$ measures the effects of the invaders on the species initially present
$\mathbf{O}$ is a matrix of zeroes, for the effects of the original species on the invaders are equal to zero, the invaders being very rare initially.

The eigenvalues of $\mathbf{P}$ are those of $\mathbf{Q}$ together with those of $\mathbf{R}$. We may assume that all eigenvalues of $\mathbf{Q}$ have negative real parts, the original community being at equilibrium. The success of the invaders is governed by the eigenvalues of $\mathbf{R}$: if at least one eigenvalue of $\mathbf{R}$ has positive real part, invasion will occur (if an eigenvalue has its real part equal to zero further requirements must be met to allow invasion; see Reed & Stenseth, (1984)).

Different functional forms for f and g in (2) have been used in the past, although a stability analysis of the extended model (2) has rarely been performed. Most popular have been models with f and g linear functions (e.g. Van Hulst, 1979). The dis-

crete and stochastic Markov chain model is also linear in the P_i, the probability of finding the system in state i. Where these models have been applied to vegetation dynamics data the fit of a constant coefficient model (or stationary chain) has usually proved to be poor (e.g., Usher, 1981; Van Dorp et al., 1985). If one reflects on the matter this is not surprising, because non-linearities are known to abound in ecological processes (e.g., Nisbet & Gurney, 1982). This suggests adopting a more complex functional form for f and g in model (2).

A game theory model

The fate of an invading population typically does depend on what other species are already growing at the site. There is no 'all-round' invader, although some species are able to invade a wide range of communities. A characteristic feature of many successions seems to be the dependence of invading species on just the right window in the successional sequence to enter the community and proliferate (Grime, 1979). The invading species may have been present for a long time in the form of seeds, rhizomes or suppressed individuals with the dominant species in the community preventing their increase. Only after a change in community composition (reflecting ultimately a change in environmental factors) can the new species increase. This dependence of the fate of an invader on the identity and abundance of the species already present parallels the dependence in game theory of the pay-off to a new player (or strategy) on the strategies already established. Particularly appropriate is the dynamic form of game theory, developed by Taylor & Jonker (1978), and Zeeman (1980, 1981).

Let us denote, as before, the proportional cover of a species i (including open ground as another 'species') by x_i, the community vector being x. Let us also introduce a **pay-off** matrix A $(a_{ij}, i,j = 1, \ldots n)$, where a_{ij} indicates the proportional pay-off to species i if surrounded by species j. Here 'pay-off' is to be interpreted as the 'per capita' (per unit cover) rate of increase of the cover of i. We can assume that species breed true (i.e. offspring play the same strategy as their parents), and that each species is

ecologically stable (each individual plays a fixed pure strategy).

We can now derive the pay-off (always in terms of the 'per capita' rate of increase in its cover) to species i within a community x as $\sum_j a_{ij}x_i = (Ax)_i$; the average pay-off to the species of community x as $\sum_i (Ax)_i = x'Ax$; and the average pay-off to the species of community x amidst those of community y as $x' Ay$ (see also Zeeman, 1981). If the rate of growth of species i is proportional to its pay-off, we can write the following equation (where the time scale has been chosen to set the proportionality constant equal to 1):

$$\frac{1}{x_i} \frac{dx_i}{dt} = (\mathbf{Ax})_i - \mathbf{x}'\mathbf{Ax}$$

or (4)

$$\frac{dx_i}{dt} = x_i [(\mathbf{Ax})_i - \mathbf{x}'\mathbf{Ax}]$$

a system of cubic differential equations. The behavioural repertoire of these equations is far richer than that of the linear models sometimes employed in succession modelling, and the model can mimic density dependence, alternative stable states and 'priority effects'. All the parameters of the equations are in the pay-off matrix A. If the pay-off to all species is increased equally the advantage of each species will not change. We may therefore reduce the principal diagonal of A to zero (by substracting a suitable constant from each column) without changing the dynamic imposed by A. Let us also define a **non-invadable community** as having a community vector x such that x cannot be invaded successfully by any one or more of the available species not in x. In game theory parlance a non-invadable community corresponds to an evolutionary stable strategy [Maynard Smith, 1982]. A community e will be non-invadable if for all $x \epsilon \Delta - e$ (in the model but not in e) either $x'Ae < e'Ae$ (e beats all other communities), or $x'Ae = e'Ae$ and $x'Ax < e'Ax$ (x and e are matched but e does better in x than x itself). Zeeman (1980) has shown that an ESS (= non-invadable community) is an attractor for the dynamics induced by (4), but that not every attractor is an ESS. Let us use the term **local climax** for an attractor. Biologically, a local climax is the end point of a successional sequence provided there are no major perturbations and changes in the species pool. The reason that a local climax is not necessarily an uninvadable community is that a perturbation in one or more of the species present may push the system away from the basin of attraction of the local climax, and possibly to another local climax. Conversely, a non-invadable community is always a local climax (in fact, it is a **global** climax).

One can also show the following (Zeeman, 1980). Given a species assemblage S; if S can form a non-invadable community with all s species present that community is unique and it forms the local climax of any mixed community with s species. An s species local climax is not necessarily a global climax since there may be other local climaxes with less than s species.

What are the advantages of the dynamic game theory formulation for succession modelling? First and foremost, it provides a mathematical model which attempts to explain community dynamics in terms of species properties, without, however, reducing community ecology to population ecology. It also provides a mathematical framework which is relatively well explored, and the model is structurally stable or can be arbitrarily closely approximated by a structurally stable model (Zeeman, 1980). Since small changes in the model parameters are unlikely to yield a model with radically different behaviour we can be much more confident about any conclusions drawn from the model than if this had not been the case. Finally, the model, although much richer than the linear models commonly used, remains easy to analyze. It is straightforward, for example, to find the basin of attraction of each local climax. The model parameters (the coefficients of the A matrix) are easily estimated from time course data, and, once this has been done the invasive ability of each species and the resistance to invasion of the community can be calculated explicitly.

The game theory model applied to succession data

Estimation of the coefficients of the invasion matrix A requires either data on temporal changes in

species abundances during succession (so that the $x(t)$ and $dx(t)/dt$ are known), or, alternatively, experimental data on changes in the $x(t)$ and $dx(t)/dt$ in plots of differing initial species composition. Reliable data on changes in cover of species inhabiting a site are understandably rare, especially in the case of long-lived species. Fortunately, however, changes in cover tend to be fairly gradual, so that we may safely interpolate for communities that have been observed only occasionally. The parameters of the game theory successional model are the coefficients of the pay-off matrix A. The model equations (4) are linear in A. If the dx_i/dt and the x_i for all n species are known at a sufficient number of points in time (minimally n points), then the coefficients can be found easily, either by using multiple linear regression analysis on more than n^2 data points, or, preferably, by using quasi-linearization techniques (Sage & Melsa, 1971). Estimation of the a_{ij} coefficients involves therefore the following steps: i) interpolation for each species i between the original data points; ii) numerical estimation of the time derivative at a sufficient number n of data points; iii) formation of a matrix Z of dimension mxn, n^2, containing the appropriate values for dx_i/dt and x_i at different times; and iv) estimation of the n^2 a_{ij} coefficients using quasi-linearization or multiple regression analysis.

The A matrix found in this way can then be studied in detail. If an independent data set is available for the same vegetation the model can now be tested simply by solving the system (4) with A as just derived numerically. The pay-off coefficients for the different species contain also biologically valuable information. A species that has all a_{ij} in its row of A greater than the corresponding row elements for every other species is able to replace any other species (Maynard Smith, 1982). Similarly, the average value of the elements of a_{ij} of a certain species i and for all the species j in community x can be taken as an index of invasive ability of species i for community x.

Conversely, resistance of a community to invasion by species presently absent may be investigated using stability analysis of the model (Rowe et al., 1985): if there exists a fixed point with all the com-

munity species present, and if this fixed point is stable with respect to the invading species, we will say that the community forms a local climax. Rowe et al. (1985) present an algorithm to find the stable points of equation (4) for any invasion matrix A.

I shall apply this method of analysis to two published successional data sets. As will be seen, each data set has its shortcomings, and the present analysis should therefore be considered as preliminary only.

Data set 1 is from Sousa (1979). Only 6 taxa are involved and open space has been treated as a 7th state. The data used are the unmanipulated controls: concrete blocks were set out in the intertidal zone near Santa Barbara, California, and the succession of species of macrophytic algae on the blocks and their percentage cover were scored over 2.5 years. The interpolated (cubic spline interpolator) data are reproduced in Fig. 1. The coefficients of the invasion matrix A were estimated using multiple linear regression. They are listed in Table 1. Note the *Ulva* spp. are efficient invaders of open space, but poor invaders of blocks already colonized by other species. Table 2 gives the invasive abilities ϕ (calculated as the average invasion index): the red alga *Rhodoglossum affine* is clearly the best invader. The single uninvadable community (found using the technique described above) is one with *Gigartina canaliculata* covering 35%, and *Rhodoglossum affine* covering the remainder of the rock. These are indeed the two species that Sousa (1979) found were often dominating the community. Note, however, that other species not in the present analysis and a changed perturbation regime may upset the above conclusions.

The second data set to be analyzed is due to W. G. Beeftink and his coworkers, and based on Hogeweg et al. [1985]. Yearly surveys were conducted of 33 permanent quadrats situated on sand flats that emerged from sea water after construction of a dam in the southwestern part of The Netherlands. The very detailed data offer a unique opportunity to study the invasion of new species in a changing community. Figure 2 shows how percentage cover of the 14 most abundant species and the percentage of open space changed from 1962 (just after the flats fell dry) to 1983. I used a cubic spline interpo-
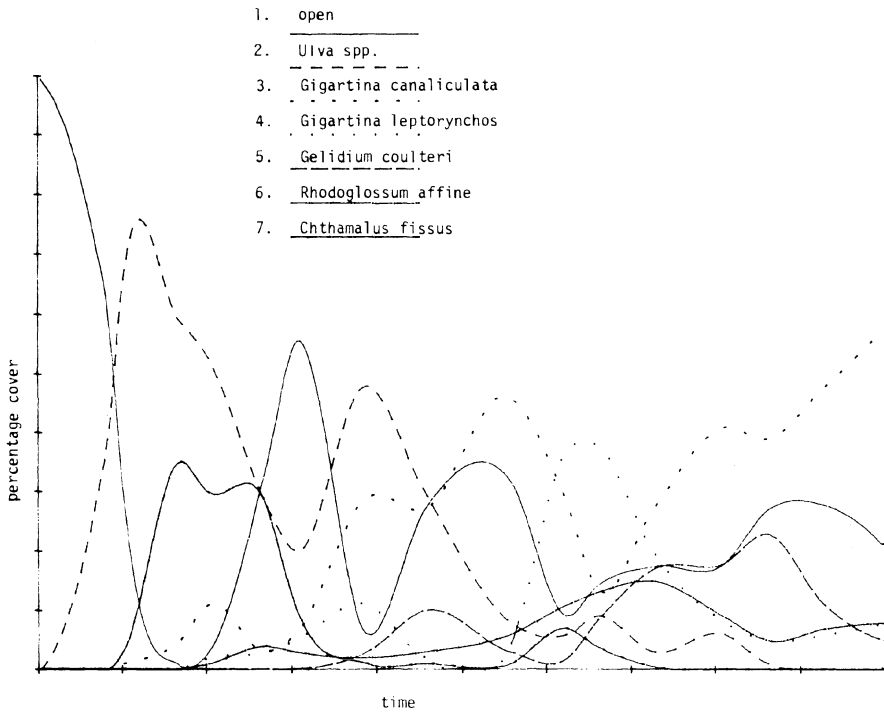
1. open
2. Ulva spp.
3. Gigartina canaliculata
4. Gigartina leptorynchos
5. Gelidium coulteri
6. Rhodoglossum affine
7. Chthamalus fissus

Fig. 1. Percentage cover of six algal taxa and percentage open space over the course of an algal succession (arbitrary time scale) on concrete blocks in the intertidal zone. Data are from Sousa (1979). Cubic spline interpolations were used to smooth the data curves.

Table 1. The invasion matrix *A* estimated for the data in Fig. 1. The *i, j* — th element of the matrix indicates the 'pay-off' (in terms of its growth rate) to species *i* due to its growing in the same community as species *j*. Note that species importance is measured as percentage cover, and that open space is treated as if it were another species.

	1	2	3	4	5	6	7
1. open	0.00	-0.22	0.04	0.15	0.44	-0.82	0.64
2. *Ulva* spp.	0.21	0.00	-0.18	-0.05	-0.69	0.71	-0.26
3. *Gigartina canaliculata*	0.30	0.54	0.00	-0.41	-0.10	1.23	-0.54
4. *Gigartina leptorynchos*	0.38	-0.22	0.90	0.00	-1.61	0.62	0.53
5. *Gelidium coulteri*	-0.29	-0.22	-0.37	0.40	0.00	-1.18	-0.10
6. *Rhodoglossum affine*	2.07	2.05	2.33	2.57	2.77	0.00	1.58
7. *Chthamalus fissus*	0.71	0.94	0.65	0.37	0.05	0.51	0.00

Table 2. Taxa from Table 1 with highest invasive ability ϕ. Invasive ability of a species *i* is measured as the average value of the invasion coefficients a_{ij} for all species *j* in the community.

Species	ϕ
Rhodoglossum affine	2.23
Chthamalus fissus	0.54
Gigartina canaliculata	0.17

lator to estimate the values between the sampling points. The invasion matrix for the 14 species and open space (estimated using multiple linear regression and with the diagonal set to 0) is shown in Table 3. I estimated the invasive ability ϕ for each species as described above by averaging its invasion coefficients. The four species with highest invasive ability are listed in Table 4. Note that these num-

1. open
2. Suaeda maritima
3. Spergularia marina
4. Juncus bufonius
5. Poa annua
6. Epilobium tetragonum
7. Epilobium adenocaulon
8. Epilobium hirsutum
9. Poa trivialis
10. Poa pratensis
11. Trifolium repens
12. Carex otrubae
13. Phragmites australis
14. Calamagrostis epigejos
15. Salix spp.
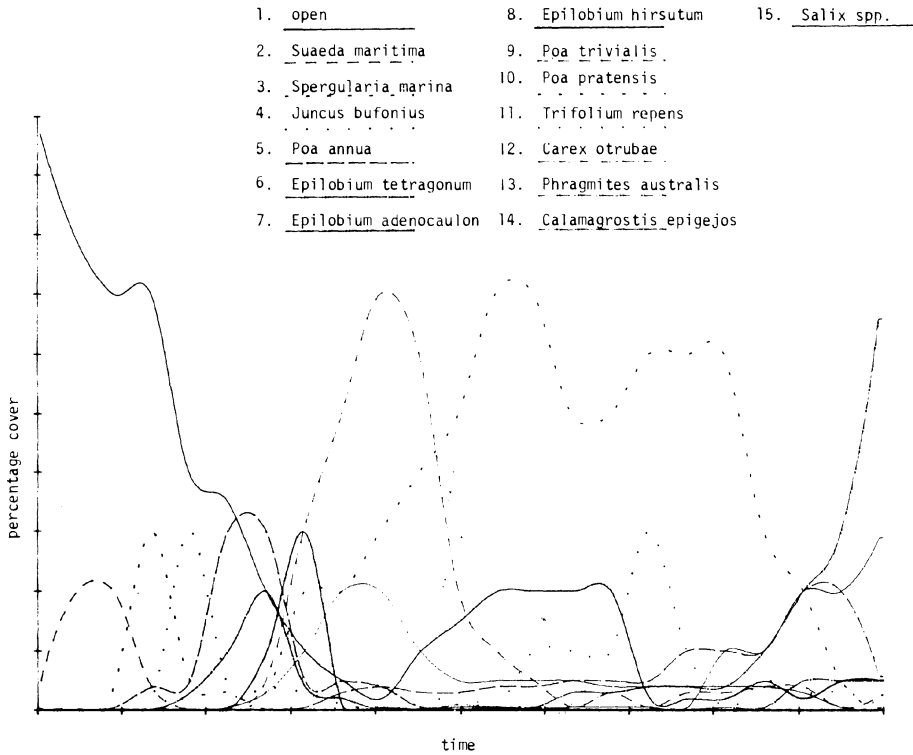
Fig. 2. Percentage of cover of 14 vascular plant taxa and open space over the course of a primary succession on an estuarian sand flat near the Dutch coast. Data are from Hogeweg *et al.* (1985). Cubic spline interpolations were used to smooth the data curves.

Table 3. The invasion matrix *A* estimated for the data of Hogeweg *et al.* (1985) Cf. Table 1 for further details.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. open	0.00	0.04	-0.04	-0.08	0.17	-0.43	-0.01	-1.24	0.26	0.04	-1.06	1.91	-11.56	1.01	-1.18
2. *Suaeda maritima*	1.15	0.00	-1.21	-0.92	0.36	-0.78	0.20	-2.00	0.03	-0.17	0.30	1.22	-0.21	-1.48	0.45
3. *Spergularia marina*	1.61	0.59	0.00	-0.94	0.62	-3.26	0.77	-0.37	0.07	-0.33	0.60	2.42	-0.39	-2.94	0.89
4. *Juncus bufonius*	1.77	-4.74	0.76	0.00	-3.07	4.69	0.03	-1.65	0.17	-0.31	0.69	2.48	-0.16	-3.22	1.26
5. *Poa annua*	-1.46	2.37	-0.59	0.07	0.00	-1.32	-0.38	0.22	-0.13	0.12	-0.09	-0.90	2.95	-0.06	0.42
6. *Epilobium tetragonum*	-1.04	-2.07	0.09	-1.21	-1.80	0.00	-0.09	-1.31	-0.17	0.23	-0.28	-1.74	0.30	1.74	0.06
7. *Epilobium adenocaulon*	-1.01	-1.20	-1.60	-1.71	5.23	-10.28	0.00	-3.02	0.16	0.21	-0.31	-0.34	-4.16	1.68	-2.04
8. *Epilobium hirsutum*	-0.38	-0.27	-0.59	-0.01	1.04	-2.19	0.00	0.00	-0.20	-0.02	0.02	0.08	0.01	-0.38	0.16
9. *Poa trivialis*	0.06	0.03	-0.33	0.72	3.31	-4.58	0.75	-0.4	0.00	-0.26	-0.16	3.08	-8.03	0.42	-1.48
10. *Poa pratensis*	0.03	-0.13	-0.30	-0.54	3.61	-4.29	0.35	-0.81	0.09	0.00	0.10	0.45	2.30	-0.42	-0.42
11. *Trifolium repens*	0.14	-0.01	-0.05	-0.06	1.55	-2.13	0.29	0.03	0.29	0.46	0.00	-6.07	-13.69	4.55	2.70
12. *Carex otrubae*	0.08	0.11	0.02	-0.04	1.01	-1.08	0.80	0.32	-0.00	0.08	0.12	0.00	2.90	0.42	-2.55
13. *Phragmites australis*	2.46	2.43	2.45	2.45	2.46	2.45	2.47	2.28	2.48	2.48	1.90	2.78	0.00	4.45	-2.55
14. *Calamagrostis epigejos*	-0.43	-0.42	-0.44	-0.44	-0.33	-0.59	-0.68	-0.12	-0.38	-0.27	-0.44	-2.02	1.28	0.00	0.89
15. *Salix spp.*	4.77	4.53	4.67	4.67	4.51	5.10	4.85	4.16	4.77	4.38	3.67	12.42	0.59	3.96	0.00

Table 4. Species from Table 3 with highest invasive ability ϕ (cf. Table 2).

Species	ϕ
Salix spp.	4.79
Phragmites australis	2.21
Carex otrubae	0.31
Poa annua	0.09

Table 5. Uninvadable communities (for the species analyzed) for the invasion matrix from Table 3. See text for further details.

Community 1:	27%	open
	21%	*Juncus bufonius*
	52%	*Poa trivialis*
Community 2:	16%	*Juncus bufonius*
	1%	*Poa annua*
	60%	*Poa trivialis*
	23%	*Trifolium repens*

bers measure ability to invade communities consisting of any of the other species in the analysis. If one wished to measure invasive ability with respect to a subset of these species one would have to average over the appropriate columns of the invasion matrix A.

The two communities that are stable with respect to invasion by any of the other species analyzed are listed in Table 5. Obviously, these communities need not be stable if the disturbance regime or the pool of available species change.

Discussion and conclusions

The game-theoretic invasion model introduced here relies heavily on high-level parameters estimated either from vegetation dynamics data or from perturbation experiments (species removal or transplantation experiments). The model has therefore a phenomenological flavour: the coefficients of the invasion matrix do have a clear biological interpretation, but they are not easily reduced to population level processes. Such a reduction, while instructive, would aid practical prediction of future vegetation change only in the simplest of situa-

tions: under realistic conditions the number and variety, and the inherent stochasticity of the underlying processes would militate strongly against predictive success.

Spatial phenomena are ignored in all models discussed here, as are disturbances. The study of patch dynamics and natural disturbances [Pickett & White, 1985] has greatly increased our understanding of the population biological basis of vegetation change. The focus of this study, however, is on the prediction of community composition, not on the detailed dynamics of local populations. If predictions of spatial distribution of species are required, then more detailed models are clearly in order. High level models, such as the one proposed here, must, of necessity, gloss over rarely occurring but potentially important disturbances such as fires, outbreaks of pests or pathogens, mudslides or volcanic eruptions. Regularly occurring disturbances such as the effects of herbivores, pathogens, and soil dynamics have more constant effects which can be incorporated in the model if they are specific to certain communities only.

Invasion based models in general, and the game theory based model proposed here in particular, may prove useful tools for the prediction and description of vegetation change. They provide alternatives to Markov chain models, which often show a poor fit to vegetation data and are purely phenomenological. More importantly, the focus on community change resulting from invasion processes may lead to a better understanding of both the community level process of succession and the population level features of invasive ability and invasion resistance. An experimental approach to the study of plant succession could rely on the determination of growth rates of a variety of invading species in different communities, and the subsequent derivation of the invasion matrix. Invasive abilities of different species, and resistance to invasion of different species groups can then be compared using the measures proposed here. Alternatively, an attempt can be made to predict future vegetation change, and the domain of applicability of the model can be explored. Work is presently in progress to do this for the case of temperate grassland communities.

References

Allen, P. M., 1976. Evolution, population dynamics, and stability. Proc. Nat. Acad. Sci. U.S.A. 73: 665–668.

Antonovics, J. & Levin, D. A., 1980. The ecological and genetic consequences of density dependent regulation in plants. Ann. Rev. Ecol. Syst. 11: 411–452.

Bazzaz, F. A., 1979. The physiological ecology of plant succession. Ann. Rev. Ecol. Syst. 10: 351–371.

Connell, J. H. & Slatyer, R. O., 1977. Mechanisms of succession in natural communities and their role in community stability and organization. Amer. Nat. 111: 1119–1144.

Drury, W. H. & Nisbet, I. C. T., 1973. Succession. J. Arnold Arbor. 54: 331–368.

Fowler, C. W. & MacMahon, J. A., 1982. Selective extinction and speciation: their influence on the structure and functioning of communities and ecosystems. Am. Nat. 119: 480–498.

Grime, P. J., 1979. Plant strategies and vegetation processes. Wiley, New York.

Hogeweg, P., Hesper, B., Van Schaik, C. & Beeftink, W. G., 1985. Patterns in vegetation succession: an ecomorphological study. In: J. White (ed.), The population structure of vegetation, pp. 637–666, Junk, Dordrecht.

Levins, R., 1968. Evolution in changing environments. Princeton University Press, Princeton.

Maynard Smith, J., 1982. Evolution and the theory of games. Cambridge University Press, Cambridge.

Miles, J., 1979. Vegetation dynamics. Chapman and Hall, London.

Nisbet, R. M. & Gurney, W. S. C., 1982. Modelling fluctuating populations. Wiley, New York.

Noble, I. R. & Slatyer, R. O., 1981. Concepts and models of succession in vascular plant communities subject to recurrent fire. In: A. M. Gill, R. H. Graves & I. R. Noble (eds), Fire and the Australian biota. Australian Academy of Science, Canberra.

Pickett, S. T. A. & White, P. S. (eds), 1985. The ecology of natural disturbance and patch dynamics. Academic Press, New York.

Reed, J. & Stenseth, N. C., 1984. On evolutionary stable strategies. J. Theor. Biol. 108: 491–508.

Rowe, G. W., Harvey, I. F. & Hubbard, S. F., 1985. The essential properties of evolutionary stability. J. Theor. Biol. 115: 269–285.

Sage, A. P. & Melsa, J. L., 1971. System identification. Academic Press, New York.

Schaffer, W. M., 1977. Evolution, population dynamics, and stability: a comment. Theor. Pop. Biol. 11: 326–329.

Sousa, W. P., 1979. Experimental investigations of disturbance and ecological succession in a rocky intertidal algal community. Ecol. Monogr. 49: 227–254.

Taylor, P. D. & Jonker, L. B., 1978. Evolutionary stable strategies and game dynamics. Math. Biosci. 40: 145–156.

Usher, M. B., 1981. Modelling ecological succession, with particular reference to Markovian models. Vegetatio 46: 11–18.

Van der Valk, A. G., 1981. Succession in wetlands: a Gleasonian approach. Ecology 62: 688–698.

Van Dorp, D., Boot, R. & Van der Maarel, E., 1985. Vegetation successions on the dunes near Oostvoorne, The Netherlands, since 1934, interpreted from air photographs and vegetation maps. Vegetatio 58: 123–136.

Van Hulst, R., 1979. On the dynamics of vegetation: succession in model communities. Vegetatio 39: 85–96.

Van Hulst, R., 1980. Vegetation dynamics or ecosystem dynamics: dynamic sufficiency in succession theory. Vegetatio 43: 147–151.

Van Hulst, R., Shipley, W. & Theriault, A. (in press). Why is *Rhinanthus minor* (Scrophulariaceae) such a good invader? Can. J. Bot.

Walker, D., 1970. Direction and rate in some British postglacial hydroseres. In: D. Walker & R. West (eds), The vegetational history of the British Isles, pp. 117–139. Cambridge University Press, Cambridge.

Zeeman, E. C., 1980. Population dynamics from game theory. In: Z. Nitecki & C. Robinson (eds), Global theory of dynamical systems, pp. 472–497, Springer, Berlin.

Zeeman, E. C., 1981. Dynamics of the evolution of animal conflicts. J. Theor. Biol. 89: 249–270.

Modeling of vegetation dynamics in the Mississippi River deltaic plain*

Marcel Rejmánek[1,2], Charles E. Sasser[2] & James G. Gosselink[2,**]
[1]*Botany Department, University of California, Davis, CA 95616, USA*
[2]*Coastal Ecology Institute, Louisiana State University, Center for Wetland Resources, LA 70803, USA*

Keywords: Atchafalaya, Delta development, Erosion, Louisiana, Markov model, Mississippi, Sedimentation, Succession, Wetland, Transition probability

Abstract

An analysis was made of vegetation phenomena associated with cyclic river delta building and abandonment in the Mississippi River deltaic plain, Louisiana. Markov models of vegetation succession were completed for the new Atchafalaya delta and for the abandoned Lafourche delta. Transition matrices representing different flood conditions in the Atchafalaya delta (1979–1984) were used separately and in combinations based on long term river discharge data. Effect of grazing was introduced into the model as a special transition matrix. Succession during active delta building is heavily dependent on a particular sequence of flood conditions and, apparently, on the intensity of grazing. Natural vegetation changes and land losses in the abandoned Lafourche delta were simulated using a transition matrix based on air photographs from 1945 and 1956. A general flow diagram of long cyclic vegetation changes in the Mississippi deltaic plain was completed.

Introduction

Deltaic plains of big rivers belong to the most variable and least predictable environments on earth (Bird, 1985; Wright, 1985). The complicated development of vegetation in river deltas reflects the unstable character of these landscapes. Realistic modeling of vegetation dynamics in river deltas is therefore necessarily limited to stochastic simulations. However, only long-term field studies can

provide data for the estimation of those parameters needed for stochastic models. Available data from the Atchafalaya delta, Louisiana, give an opportunity for the first estimation of transition probabilities between different types of vegetation under different river flow regimes.

The Atchafalaya delta represents one of the most dynamic geological events in historical times. Throughout this century, the flow of the Mississippi River has been shifting gradually to the Atchafalaya River. Since 1963, the flow of the Atchafalaya River has been stabilized at approximately 30% of the combined flow of the Mississippi and Red rivers by control structures located at Simmesport, Louisiana. Sediments carried by the Atchafalaya River have been filling Atchafalaya Bay. Between 1965 and 1972, the average annual flood discharge was 7500 m³/s, which carried an average annual sediment load of 42.6×10^6 metric tons (Roberts *et al.*, 1980). New islands which first emerged in 1973
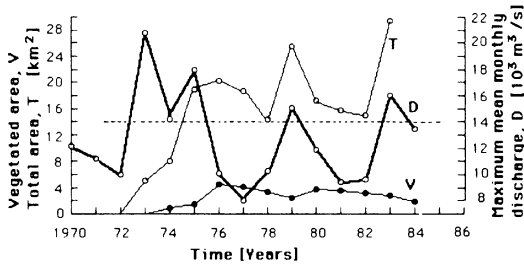
Fig. 1. Total exposed area of the Atchafalaya delta (T, area above mean sea level), vegetated area of natural islands (V) and maximum mean monthly discharge for the Atchafalaya River at Simmesport, Louisiana (D). The dotted line represents the approximate critical D value for positive growth of exposed area. The values of vegetated area of natural islands are based on LANDSAT and low level aerial imagery for the years 1973–1980 and estimated from relative changes of vegetated area covered by permanent plots during 1979 to 1984.

are the first stages in the formation of a new Mississippi delta in the center of Louisiana's coastal zone. This delta lobe has not been increasing monotonically (Fig. 1). In general, the growth of the delta only occurs during floods with mean monthly discharge greater than 14000 m³/s. Erosion or no growth takes place in intervening years with normal or sub-normal regimes. Islands situated adjacent to the navigation channel are influenced by dredged spoil. Their vegetated area is not shown in Fig. 1. The vegetated area of natural islands is still relatively small and follows, with a one yr delay, changes in the total area of land above mean water level. In addition, heavy grazing by nutria (coypu, *Myocastor coypus* (Molina), *Rodentia*), has affected vegetation dynamics recently. Vegetation development from 1973 to 1982 was described by Montz (1978) and Johnson *et al.*(1985).

The active delta building is in sharp contrast to prevailing subsidence and erosion of Louisiana coastal marshes which occupy the area of ancient Mississippi deltas successively formed and abandoned during the last 8000 years (Kolb & Van Lopik, 1966; Van Heerden, 1983; Baumann *et al.*, 1984). The Atchafalaya River is building a delta lobe in an area where Maringouin and Teche deltas were built and abandoned 7000 to 4000 years ago. While wetland loss is part of the natural geomorph-

ic scene, historically it has been balanced by land extension elsewere, so that since the last ice age the Mississippi River deltaic plain has grown enormously. In this century, this trend has been reversed and net loss of coastal wetlands is the rule. The present coastal land erosion rates are estimated to be 102 km² yr⁻¹, or 0.8% annually (Gagliano *et al.*, 1981).

Data and methods

A total of 110 permanent plots, each 1×1 m, placed at intervals 20–40 m along two perpendicular transects on each of four natural islands, have been maintained since 1979. Two of the islands, Hawk and Log islands, are on the west side of the delta; the other two, Rodney and Teal islands, are located on the east side of the delta. Cover estimates of vascular plant species and elevation changes in relation to the nearest benchmark were monitored in all plots during the period 1979–1984.

For the estimation of transition probabilities between particular vegetation stages, four main vegetation types were distinguished on the basis of multivariate analysis. They were characterized by dominance of *Cyperus difformis*, *Sagittaria* (*S. latifolia* and *S. platyphylla*), *Typha domingensis*, or *Salix nigra*. Remaining plant communities dominated by *Justicia ovata*, *Scirpus americanus*, *S. tabernaemontanii*, or *Leersia oryzoides* were considered as a fifth category ('Other'). Aquatic vegetation (*Najas guadalupensis*, *Vallisneria americana*, *Heteranthera dubia*, etc.) was very sparse and its appearance very inconsistent from year to year and from place to place. Aquatic vegetation was therefore classified in a sixth category, 'Non-vegetated', together with mud flats and open water. The proportion of quadrats remaining in the same category, or changing to another was determined for the five year-to-year changes during 1979–1984. The set of transformation probabilities for six vegetation categories constitutes a transition matrix. By multiplying a vector of abundances of these categories by a transition matrix, vegetation dynamics could be simulated. Because the matrices were very different, depending on flood (erosion or sedimentation) conditions, they were combined in a stochastic sequence derived from the analysis of long-term discharge data for the Atchafalaya River. Further details on the use of transition matrices can be found in Kemeny & Snell (1976), Austin (1980), Usher (1981), and Hobbs (1983).

A similar matrix was completed for transitions from solid marsh to open water in the abandoned Lafourche Mississippi delta, southwestern Barataria Basin. Air photographs from 1945 and 1956 were used (Dozier *et al.*, 1983). Marsh vegetation in photographs was classified according to digitized water area in 10000 quadrats (50×50 m).

Results

Simulation of vegetation succession based on four different transition matrices (Table 1) is shown in Figs. 2 and 3. Matrices 1980−1981 and 1981−1982 represent times of rather low river discharge, with erosion prevailing over sedimentation, on the aver-

Table 1. Transitions between the six states included in the model. C: *Cyperus*, S: *Sagittaria*, T: *Typha*, X: *Salix*, O: Other vegetation types, N: Non-vegetated.

Transition	Probability			
	'80 – '81 '81 – '82	82' – '83	'79 – '80	'83 – '84
C→C	0	0.500	0	0.170
C→S	0.570	0.500	0.375	0
C→T	0	0	0.500	0
C→X	0	0	0	0
C→O	0	0	0.125	0.510
C→N	0.430	0	0	0.320
S→S	0.794	0.840	0.857	0.428
S→C	0	0.072	0.024	0
S→T	0.029	0.028	0.024	0.042
S→X	0	0	0.024	0
S→O	0.029	0	0	0.042
S→N	0.148	0.060	0.071	0.488
T→T	0.850	0.625	0.900	0.499
T→C	0	0.125	0	0.167
T→S	0	0	0	0
T→X	0	0	0	0
T→O	0.150	0.250	0.100	0
T→N	0	0	0	0.334
X→X	0.660	0.920	1	0.860
X→C	0	0	0	0
X→S	0	0	0	0.140
X→T	0	0	0	0
X→O	0	0	0	0
X→N	0.340	0.080	0	0
O→O	0.500	0.500	0.500	0.383
O→C	0	0	0	0
O→S	0	0.160	0.500	0.083
O→T	0.100	0.180	0	0.167
O→X	0.400	0.160	0	0
O→N	0	0	0	0.367
N→N	0.600	0.900	0.229	0.700
N→C	0.100	0	0.183	0
N→S	0.300	0.100	0.500	0.250
N→T	0	0	0	0
N→X	0	0	0.088	0
N→O	0	0	0	0.050

age. These two matrices were very similar, and resulting simulations were practically indistinguishable. The average matrix was therefore used for all simulations. Matrix 1982−1983 represents high river discharge (maximum mean monthly discharge over 14×10^3 m³/s). Under such conditions, the whole subaerial area of delta is growing, but the vegetated area is slightly decreasing. During years of high floods, the prolonged submersion and reduced light penetration, combined with low river water temperatures, shorten the growing season and apparently inhibit germination and growth. The most remarkable difference between low and high river discharge is in contraction or expansion of stands dominated by *Cyperus difformis*. The first year after high river discharge (matrix 1979−1980) is characterized by a rapid increase of vegetated area.

Vegetation responses are different under different discharge-sedimentation conditions. It would be completely misleading if only one year-to-year transition were used for prediction of succession in deltaic environment. Moreover, the transition 1983−1984, which was predicted to be similar to that of 1979−1980 was completely different (Fig. 3). We believe that the heavy grazing by nutria is mainly responsible for this discrepancy. Exclosures established in 1980 by Fuller *et al.* (1985) in dense stands of *Sagittaria latifolia* were the only spots where *Sagittaria* was present in 1984. Surrounding, unprotected marsh, was converted to non-vegetated area. There were no differences in elevation (sedimentation/erosion) between exclosures and surrounding mud flats in 1984. Because a high number of nutria has been observed in the delta in the last years, we assume that an increase of their abundance in 1984 is the reason for decrease of vegetated area. This opportunistic feeder consumes a variety of marsh plants but tubers of *Sagittaria* species are preferred over other vegetation (Chabreck *et al.*, in press; Wentz, 1971). Nutria grazing may also be responsible for the lack of recovery of *Salix nigra* in 1984.

A more realistic simulation of vegetation dynamics should reflect patterns of river floods and subsequent sedimentation, erosion, and vegetation change. The final model (Figs. 4 and 5). combines
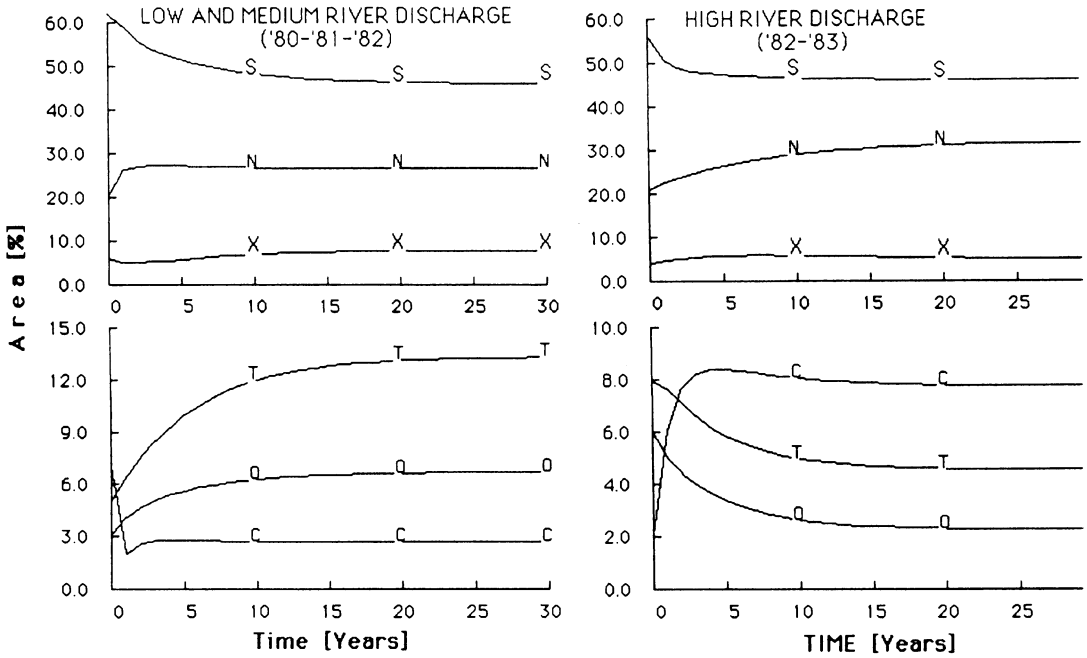
Fig. 2. Vegetation development on the Atchafalaya delta islands as predicted by the transition matrices for low/medium river discharge and for high river discharge. Vegetation categories are abbreviated as follows: S − *Sagittaria* spp., X − *Salix nigra*, T − *Typha domingensis*, C − *Cyperus difformis*, O − other types, N − nonvegetated area.
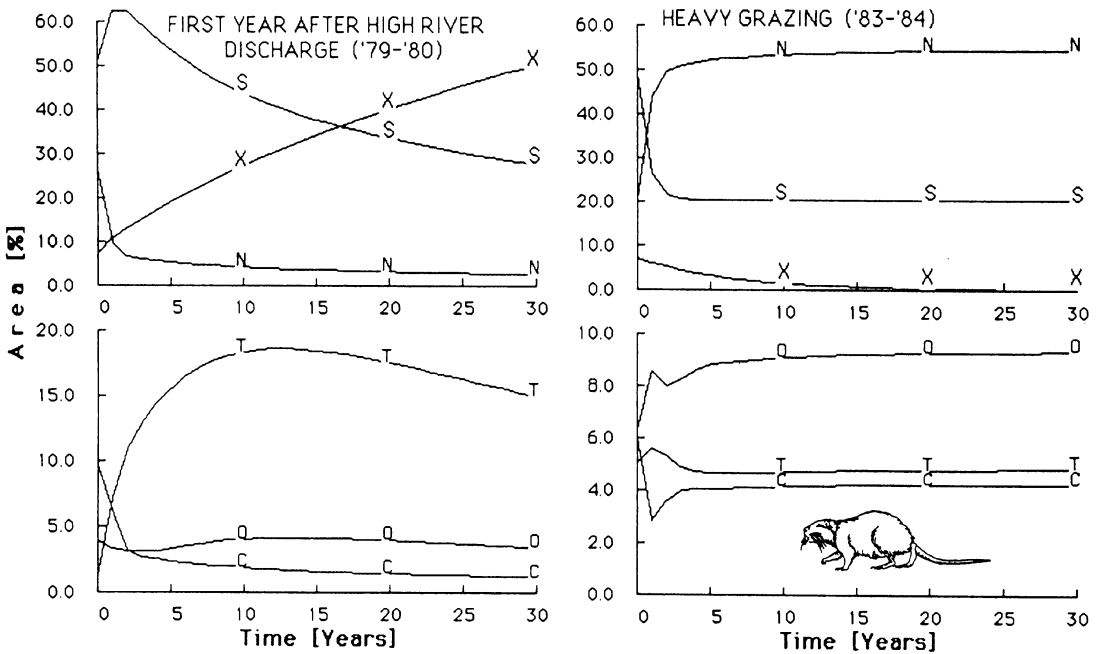

Fig. 3. Vegetation development on the Atchafalaya delta islands as predicted by the transition matrices for the first year after high river discharge and for heavy grazing by nutria. Abbreviations for vegetation categories are the same as in Fig. 2.
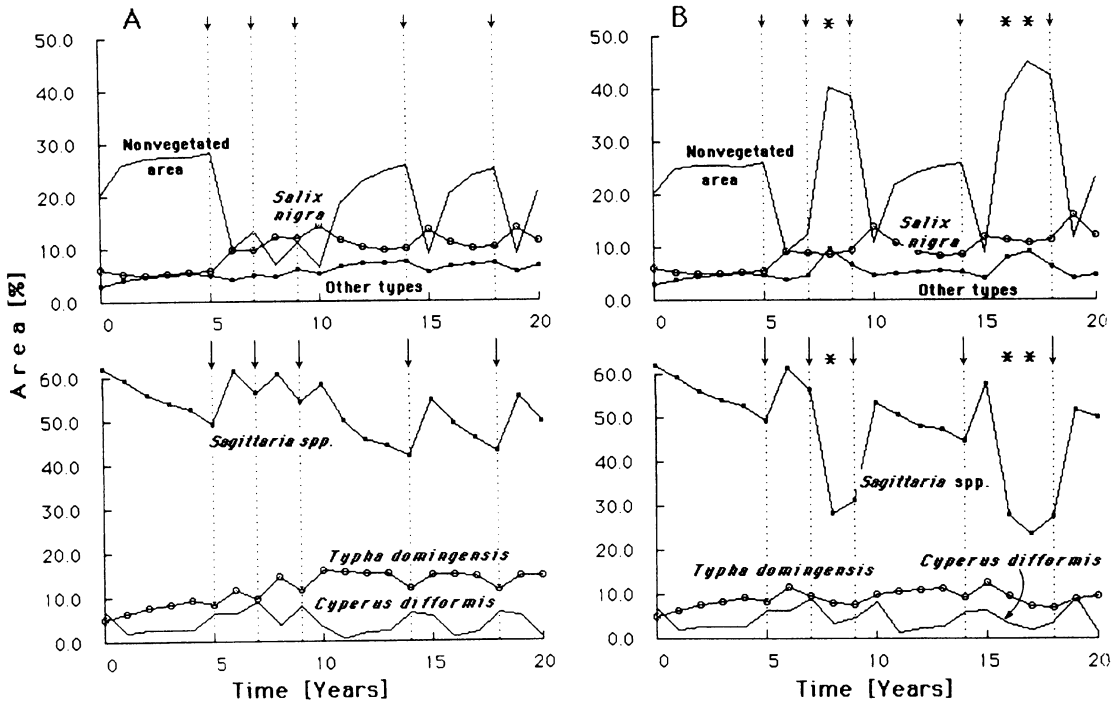
Fig. 4. An example of simulated vegetation development on the Atchafalaya delta islands by the model combining three transition matrices for different river discharge situations (A) and by a model combining the same three matrices with matrix for heavy grazing (B). Years with a high annual peak river flow are indicated by arrows; years with heavy grazing are indicated by asterisks.

three transition matrices corresponding to three discussed river discharge situations. The probability of switching between them is based on the analysis of long-term hydrological data. In particular, probability of high flood is 0.23. The simulation example in Fig. 4A shows the possible relative areal changes of six vegetation categories on islands existing in 1980. There was a more or less consistent increase of area occupied by *Salix nigra* and *Typha domingensis* vegetation types in all simulations during the next 20 years. The area dominated by *Sagittaria* spp. slightly decreased. The *Cyperus difformis* vegetation type (*C. difformis* and *Eleocharis pauciflora*) exhibited maximum relative fluctuations which reveals an L-strategy of component species (see Whittaker, 1975, p. 51). The *Typha domingensis* type exhibited the highest constancy. Heavy grazing (Fig. 4B) can change some proportions dramatically, namely non-vegetated and

Sagittaria dominated areas.

Changes of the actual area occupied by different vegetation categories are simulated assuming that a high discharge year ($>14\times10^3$ m³/s) causes the extension of the *Cyperus difformis* type area by 70%, and the first year after a high discharge year causes extension of existing area of all other vegetation types by 60%. Available data do not allow better estimates of vegetated area extensions. Resulting simulations of vegetated area changes (Fig. 5) are in agreement with analytical (Wang, 1984) and regressional (Wells *et al.,* 1982) model predictions of the Atchafalaya delta growth. Heavy herbivory (Fig. 5B) can change the vegetated area substantially: the difference in the potential area of *Sagittaria* spp. is most remarkable. Regarding data from the modern Mississippi subdeltas, it is very unlikely that some vegetation type not considered here could play an important role in the next 20 years of

138
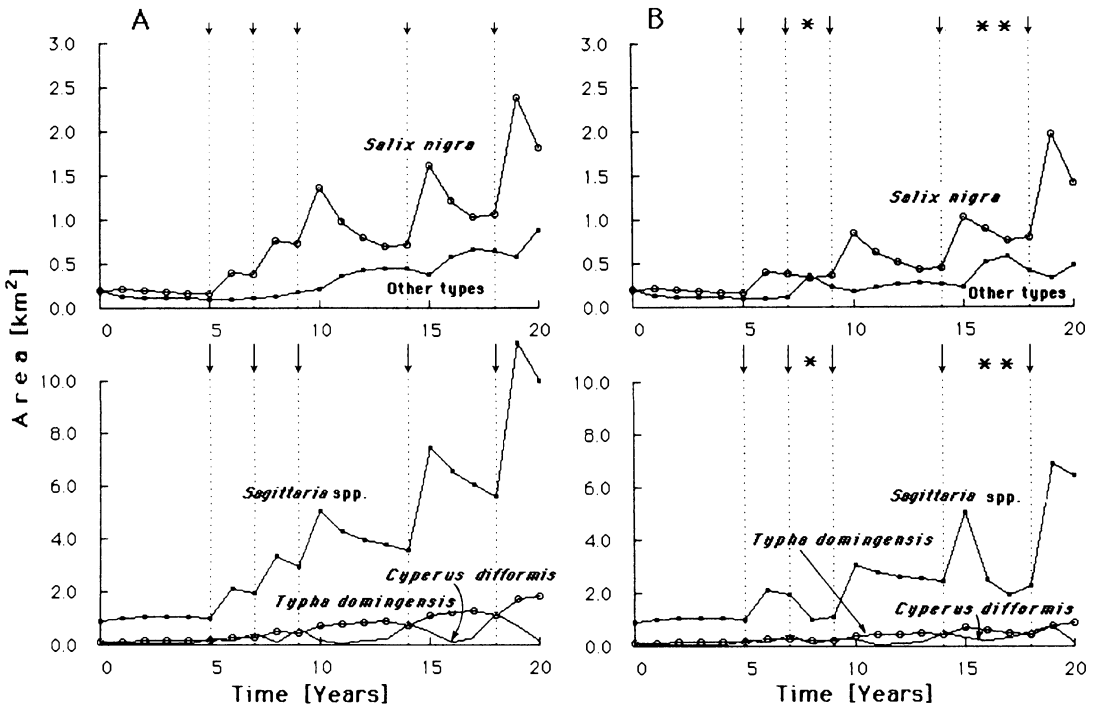
Fig. 5. An example of simulated vegetation development in the whole Atchafalaya delta by the model combining three transition matrices (A) and by the same model combined with the transition matrix for heavy grazing (B). Additional assumptions about discharge dependent vegetated area growth (See text) were applied in this model. Years with a high annual peak river flow are indicated by arrows; years with heavy grazing are indicated by asterisks.

Atchafalaya delta development. A possible exception might be a new type dominated by *Phragmites australis*.

A similar Markov model was completed for simulation of temporal changes in coastal marshes in southwestern Barataria Basin (Table 2). This area, covered by a mixture of saline, brackish, intermediate and fresh marshes, represents the Lafourche Mississippi delta formed about 2000 years ago and completely abandoned only 60 years ago. The simulation based on transition matrix 1945–1956 is shown in Fig. 6. Although this simulation predicts dramatic land losses, the prediction is optimistic compared with reality. Since the 1960s, intrusions of salt water, erosion and subsidence have been dramatically accelerated by the construction of many navigation and pipeline channels in Louisiana coastal marshes (Deegan *et al.,* 1984). Even the

1945–1956 period was not without human impact and because we have no undisturbed reference area, we will never know how much our simulation of the natural marsch disintegration is exaggerated.

Table 2. Matrix of the probabilities for transition between marshes differing in the proportion of water area. M1: solid marsh, M2: marsh with 5% – 10% water, M3: 10% – 25% water, M4: 25 – 80% water, OW: 80% – 100% water. Transitions between marshes, levees, beaches and developments are not shown.

		From				
		M1	M2	M3	M4	OW
To	M1	0.323	0.117	0	0	0
	M2	0.505	0.575	0.440	0.396	0.048
	M3	0.052	0.151	0.460	0.124	0.013
	M4	0.016	0.053	0.020	0.369	0.016
	OW	0.070	0.079	0.080	0.111	0.923
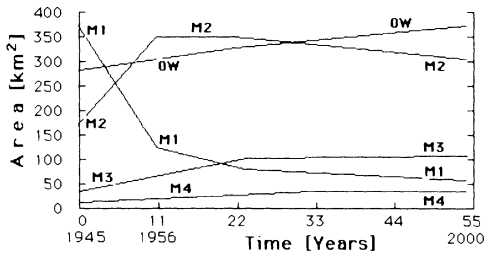
Fig. 6. Simulation of the Barataria Basin marsh development based on the transition matrix 1945–1956. Marsh categories are abbreviated as follows: M1 – solid marsh, M2 – marsh with 5% – 10% water, M3 – 10% – 25% water, M4 – 25% – 80% water, OW – 80% – 100% water.

Discussion and conclusions

Since the last glaciation, the Mississippi River has shifted its course several times and sediments have been transported to different portions of the Gulf of Mexico coast (Kolb & Van Lopik, 1966; Frazier, 1967). New deltas have been built and old deltas have been abandoned. We tried to model vegetation dynamics in two short periods of this gigantic geomorphic scene. Cyclic patterns of landform changes do not provide enough opportunities for autogenic and/or unidirectional succession. Evidence from the Mississippi deltaic plain and cyclic vegetation changes described from other deltas and floodplains (Drury, 1956; Waldemarson-Jensén, 1979) led us to the conception of cyclic vegetation succession in the Mississippi deltaic plain (Fig. 7) and to refusal of the earlier unidirectional and convergent succession scheme for the same area (Penfound & Hathaway, 1938).

Our studies and other available data (Gosselink, 1984; Neill & Deegan, in press) show that only about 65% of the area of an open bay is converted to fresh marsh (*Sagittaria* spp., *Typha* spp., *Phragmites australis, Panicum hemitomon,* etc.) when a new delta is formed and input of sediments is high. Apparently no more than 25% of the total area of a delta can be converted to swamp forests (*Salix nigra, Taxodium distichum, Nyssa aquatica, Acer rubrum* var. *drummondii,* etc.) and only about 1% reaches temporarily the 'climax' stage of hardwoods on natural levees (*Quercus virginiana, Q.*
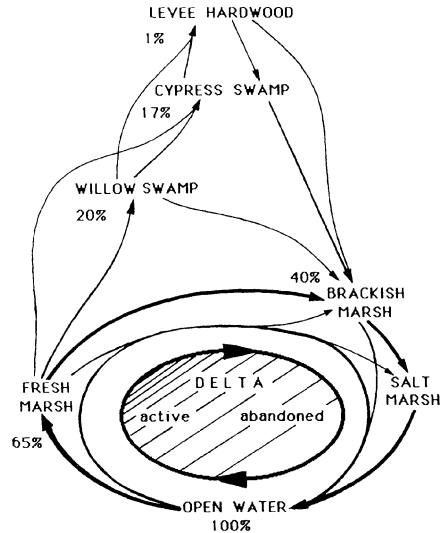
Fig. 7. Flow diagram of long term cyclic vegetation dynamics in the Mississippi deltaic plain. The values show the possible areas in percentage.

nigra, Celtis laevigata, Ulmus americana, etc.). Remaining woody vegetation (shrub communities dominated by *Myrica cerifera* or *Baccharis halimifolia* and *Iva frutescens*) can cover about 2% of the area in some periods of delta development. Later, when the river shifts its course and sedimentation is restricted, the whole area starts to sink because compression of sediments is not compensated by the external sediment supply. Salt water intrusions are more and more frequent. Fresh marshes and swamps are converted into brackish and salt marshes (*Spartina patens, Distichlis spicata, Juncus roemerianus, Spartina alterniflora,* etc.). Subsidence continues and erosion takes place until deltaic wetlands are converted into open water again. The complete cycle lasts several hundred to several thousand years.

References

Austin, M. P., 1980. An exploratory analysis of grassland dynamics: an example of lawn succession. Vegetatio 43: 87–94.

Baumann, R. H., Day, J. W. & Miller, C. A., 1984. Mississippi deltaic wetland survival: sedimentation versus coastal submergence. Science 224: 1093–1095.

Bird, E. C., 1985. Coastline changes: A global review. Wiley, New York.

Charbreck, R. H., Pilcher, B. K. & Ensminger, A. B. (in press). Growth, production, and wildlife use of delta duck potatoes in Louisiana. Proc. S. E. Assoc. Game Fish Comm.

Deegan, L. A., Kennedy, H. M. & Neil, C., 1984. Natural factors and human modifications contributing to marsh loss in Louisiana's Mississippi River deltaic plain. Environ. Manage. 8: 519–528.

Dozier, M. D., Gosselink, J. G., Sesser, C. E. & Hill, J. M., 1983. Wetland change in southeastern Barataria Basin, Louisiana, 1945–1980. Center for Wetland Resources, Louisiana State University, Baton Rouge. 102 pp.

Drury, W. H., 1956. Bog flats and physiographic processes in Upper Kuskowim River Region, Alaska. Contrib. Gray Herb. 178: 1–130.

Frazier, D. E., 1967. Recent deltaic deposits of the Mississippi River: their development and chronology. Gulf Coast Assoc. Geol. Trans. 17: 287–311.

Fuller, D. A., Sasser, C. E., Johnson, W. B. & Gosselink, J. G., 1985. The effects of herbivory on vegetation on islands in Atchafalaya Bay, Louisiana. Wetlands 4: 105–113.

Gagliano, S. M., Meyer-Arendt, K. J. & Wicker, K. M., 1981. Land loss in the Mississippi River deltaic plain. Trans. Gulf Coast Ass. Geol. Soc. 31: 295–300.

Gosselink, J. G., 1984. The ecology of delta marshes of coastal Louisiana: A community profile. Fish and Wildlife Service, U.S. Dept. of Interior, 134 pp.

Hobbs, R. J., 1983. Markov models in the study of post-fire succession in heathland communities. Vegetatio 56: 17–30.

Johnson, W. B., Sasser, C. E. & Gosselink, J. G., 1985. Succession of vegetation in an evolving river delta, Atchafalaya Bay, Louisiana. J. Ecol. 73: 973–986.

Kemeny, J. G. & Snell, J. L., 1976. Finite Markov chains. Springer, New York.

Kolb, C. R. & Van Lopik, J. R., 1966. Depositional environments of the Mississippi River deltaic plain – southeastern Louisiana. In: M. L. Shirley (ed.), Deltas, pp. 17–61. Houston Geological Society, Houston.

Montz, G. N., 1978. Vegetational characteristics of the Atchafalaya river delta. Proc. Louisiana Acad. Sci. 61: 71–84.

Neill, C. & Deegan, L. A. (in press). The role of Mississippi River delta lobe development in determining the habitat composition, diversity, and productivity of Louisiana coastal wetlands. Amer. Midl. Nat.

Penfound, W. T. & Hathaway, E. S., 1938. Plant communities in the marshes of Louisiana. Ecol. Monogr. 8: 1–56.

Roberts, H. H., Adams, R. D. & Cunningham, R. H. W., 1980. Evolution of the sand-dominated subaerial phase, Atchafalaya delta, Louisiana. Am. Assoc. Petroleum Geol. Bull. 64: 264–279.

Usher, M. B., 1981. Modelling ecological succession, with particular reference to Markovian models. Vegetatio 46: 11–18.

Van Heerden, I. L., 1983. Deltaic sedimentation in eastern Atchafalaya Bay, Louisiana. Center for Wetland Resources, Louisiana State University, Baton Rouge.

Waldemarson-Jensén, E., 1979. Succession in relationship to lagoon development in the Laitaure delta, North Sweden. Acta Phytogeogr. Suec. 66: 1–120.

Wang, F. C., 1984. The dynamics of a river-bay-delta system. J. Geophys. Res. 89: 8054–8060.

Wells, J. T., Chinburg, S. J. & Coleman, J. M., 1982. Development of the Atchafalaya River deltas: Generic analysis. Center for Wetland Resources, Louisiana State University, Baton Rouge.

Wentz, W. A., 1971. The impact of nutria (Myocastor coypus) on marsh vegetation in the Willamette Valley, Oregon. M.S. Thesis. Oregon State University, Corvallis.

Whittaker, R. H., 1975. Communities and ecosystems. 2nd ed. Macmillan, New York.

Wright, L. D., 1985. River deltas. In: R. A. Davis (ed.), Coastal sedimentary environments, pp. 1–76. Springer, New York.

Gophers and grassland: a model of vegetation response to patchy soil disturbance

R. J. Hobbs & V. J. Hobbs*
CSIRO, Division of Wildlife & Rangelands Research, LMB 4, P.O. Midland, W.A. 6056, Australia

Abstract

We present a computer model which simulates population processes and spatial patterning in response to gopher disturbance in an annual grassland community. The model includes the processes of seed production, dispersal, germination and seedling survival of four main grassland species. Runs show that soil disturbance by gophers affects both short-term spatial patterning and long-term species composition. The main determinant of species behaviour is their relative seed dispersal distances, and the model is relatively insensitive to initial species distribution or abundances. The model adequately simulated short-term community patterning, and provided predictions of long-term dynamics which can now be tested using longer-term field experiments.

Introduction

Various authors have recently emphasised the importance of disturbance at various spatial and temporal scales in shaping the dynamics of plant communities (e.g. Sousa, 1984; Pickett & White, 1985). Our studies of an annual grassland community on serpentine soil in Northern California have shown that continual disturbance of the soil by pocket gophers (*Thomomys bottae* Mewa) exerts a strong influence on the spatial patterning of the plant community through its effects on the population dynamics of the individual species (Hobbs & Mooney, 1985). To complement these studies, we have attempted to model the spatial and temporal dynamics of the grassland community using the field data available on plant population dynamics. In this paper we present a spatial model incorporating data on seed production, dispersal, germination and seedling survival for the major annual species

in the serpentine grassland. We use the model to simulate the observed effects of disturbance on community patterning and extend it to investigate the long-term effects of single and repeated disturbance on community dynamics.

System modelled

The model data are derived from studies on the annual grassland which occurs on serpentine outcrops on Jasper Ridge Biological Preserve in San Mateo County, Northern California. The climate is a Mediterranean-type with a mean annual rainfall of 480 mm and a virtually rain-free period from May to September. The grassland is dominated by annual forbs, which commence growth after the first significant rains at the end of the summer. Most species complete flowering and seed set by April or May, after which the plants die. A number of deeper rooted annual species flower later in the summer. Perennial bunch grasses are also present in the community and these also flower later in the summer.

Detailed descriptions of the grassland vegetation are given by McNaughton (1968) and Hobbs & Mooney (1985). The most abundant annual species at the site studied by Hobbs & Mooney (1985) were *Lasthenia californica* DC. ex Lindley, *Plantago erecta* Morris and *Calycadenia multiglandulosa* DC. *Bromus mollis* L. was also locally abundant. The area is subject to recurrent soil disturbance by gophers which dig tunnels 10–20 cm beneath the soil surface and form mounds of excavated soil. These mounds are 30–50 cm in diameter, and bury any existing vegetation to a depth of up to 10 cm.

Hobbs & Mooney (1985) found that fewer seeds of the dominant species (*Lasthenia* and *Plantago*) were dispersed onto gopher mounds due to the limited movement of seeds from within the closed vegetation. Species with taller flowering stalks (e.g. *Bromus* and *Calycadenia*) were more likely to colonise gopher mounds. Germination and survival of each species also differed between mounds and undisturbed habitats. These factors led to distinct spatial patterning of species abundances within the community, at a scale directly related to gopher activity (Fig. 1).

Model construction

Starting conditions

The model operates on a grid of squares, each representing 10 × 10 cm. This grid represents an area of serpentine grassland. In the following discussion the grid is 20 by 16 squares in size, although this can be expanded if required. The scale was cho-

sen to allow adequate representation of the processes of seed dispersal and to be large enough to allow inclusion of a number of gopher mounds in the grid.

The four species discussed above (*Lasthenia, Plantago, Calycadenia* and *Bromus*) are included in the model. These species are initially distributed randomly over the grid, each grid square being allocated a number of plants of each species which varies randomly between 0 and a maximum determined from field observations (Hobbs & Mooney, 1985). This provides a random pattern of abundance of each species independent of the other species. The maximum numbers of each species are also set to provide a total number of plants less than the maximum possible within a square. These maxima apply only to the starting numbers.

Gopher disturbance is then superimposed on the initial pattern of species abundances. Mounds are represented by areas of 4×4 grid squares, and are added at a rate representing an annual disturbance of one third of the total area, matching observed gopher activity. Mounds are distributed randomly over the area, and all plant abundances within mounds are set to zero, corresponding to the burying of existing vegetation. Where gopher disturbance is repeated through the model run, new random distributions are created each time. This may involve the repeated disturbance of any given area: again this matches the field situation.

Seed production and dispersal

Within each square, seed production by each species is determined by the number of plants present
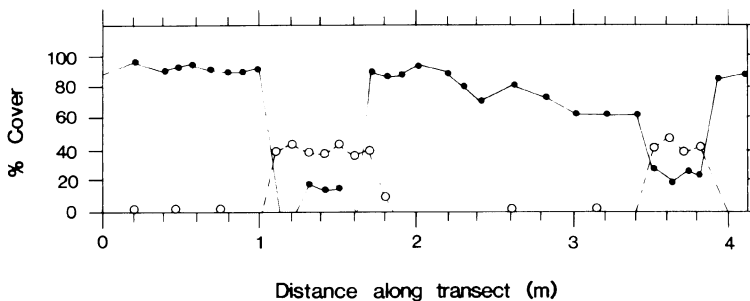
Fig. 1. Patterning observed in serpentine annual grassland at Jasper Ridge study site: percentage cover of *Lasthenia californica* (●) and *Bromus mollis* (○) in 10 × 10 cm quadrats demarcated along a 4 m transect (from Hobbs & Mooney, 1985).

and the number of seeds produced per plant. (Note that some of the propagules we call 'seeds' here are more properly described as fruits.) At low plant densities, maximum seed production is the product of the number of seeds per plant and the number of plants. Above a certain species-specific plant density the maximum seed production per square assumes a constant maximum value. The plant density at which this occurs is calculated from the field data on number of seeds per unit area. Actual seed production for each species within each square varies stochastically between zero and the maximum value.

The seeds produced within each square are then dispersed to surrounding squares on the basis of distance-related dispersal functions. For *Lasthenia* and *Bromus* these are derived from field data on dispersal distances. No empirical information was available for dispersal by *Plantago* and *Calycadenia*. The dispersal functions for these species are modified slightly from those of *Lasthenia* and *Bromus* respectively, since dispersal distances were found to be related more to height of seed release than to seed morphology (Hobbs & Mooney, 1985, see also Sheldon & Burrows, 1973). Mean maximum inflorescence heights were 12.2 cm for *Lasthenia*, 15.3 cm for *Plantago*, 19.9 cm for *Bromus* and 25.4 cm for *Calycadenia* (Hobbs & Mooney, 1985). Dispersal functions for each species are shown in Fig. 2. Numbers of seeds dispersed to each surrounding square are a random proportion of the number predicted by the dispersal function.

Germination and survival

Once seeds have been dispersed, a proportion of the seeds which land in each square germinates, and a further proportion of germinated seedlings survives to flowering. Germination and survival for each species are determined from field observations, and can either be constant or vary with density or with microhabitat. Here we discuss only constant germination and survival. Plant density thus affects only seed production, and this gives reasonable simulations of maximum plant densities; i.e. approximately similar to the starting maximum densities.

The numbers of plants which have survived in each square then constitute the new starting distribution for the following year.
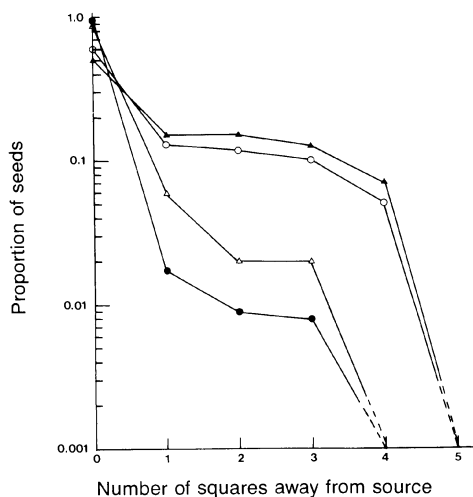
Fig. 2. Dispersal functions for transfer of seeds from source square (0) to surrounding squares: proportion of total seed production transferred to squares against number of squares away from source. ● *Lasthenia californica;* △ *Plantago erecta;* ○ *Bromus mollis;* ▲ *Calycadenia multiglandulosa.*

Model output

Output consists of a map of the location of new gopher mounds, a map of abundances for each species and a summary map of species dominances. Species dominance is defined simply in terms of numerical abundance; i.e. the most numerous species is classed as the dominant. Numbers of squares dominated by each species are given, and the average proportions of each species in different microhabitats (i.e. on and off gopher mounds) can be obtained.

Results and discussion

Because we are modelling an annual vegetation type, the time intervals represented by the model are easily defined, with one iteration of the model representing one year (i.e. April–April). The life cycle of the annual plants is easy to model, since all life history events occur at well-defined times. Similarly, since individual plants growing on the nutrient-poor serpentine soil spread little horizontally, spatial relationships present few problems compared with perennial or spreading species (e.g. Van Tongeren & Prentice, 1986). Outward spread

occurs only during seed dispersal. In practice, this is probably the case for all species except *Calycadenia,* which is longer lived and reaches a greater size than the other annual species in the model. In its present form, however, the model assumes no horizontal spread by individual plants.

Parameters used in the model runs discussed here are given in Table 1. Fig. 3 gives the starting distribution of dominance obtained using random allocation of species abundances to each square. *Lasthenia* and *Plantago* dominate the majority of squares because of their higher maximum numbers. With no disturbance, the distribution of individual species remains virtually unchanged in subsequent years and *Lasthenia* and *Plantago* continue to

dominate although local patterns of dominance shift slightly. If gopher disturbance is added, on the other hand, a clear difference in dominance patterns is found in the following year (Fig. 3b & c). In this case, *Bromus* becomes dominant on the gopher mounds and *Calycadenia* also increases in importance. The pattern of species dominance simulated in Fig. 3c closely approximates the patterns found in the field (Fig. 1 and Hobbs & Mooney, 1985). This suggests that the model is operating effectively, and that the parameters used give an adequate representation of the dynamics of the species modelled.

An important feature of the results is that the patterns of species dominance in response to gopher disturbance were generated entirely on the basis of interspecific differences in dispersal ability. Thus *Plantago* and *Lasthenia* were unable to disperse onto the gopher mounds to any great extent, whereas numerous *Bromus* and/or *Calycadenia* seeds were able to reach the mounds. No species interactions were included in these runs and plant density affected seed production only. This assumes that the annual community behaves as a collection of non-interacting species when considering the vegetation response to disturbance. While this may be an oversimplification, Platt (1975) also successfully modelled colonization of badger distur-

Table 1. Parameters used in model runs discussed in text. All plant densities are per model square; L = *Lasthenia;* P = *Plantago;* B = *Bromus;* C = *Calycadenia.*

	L	P	B	C
Starting maximum plant density	35	35	20	15
Number of seeds per plant	20	4	10	5
Maximum plant density with no density-dependent seed production	25	10	2	3
Percentage germination	15	60	30	30
Percentage survival	58	80	49	58
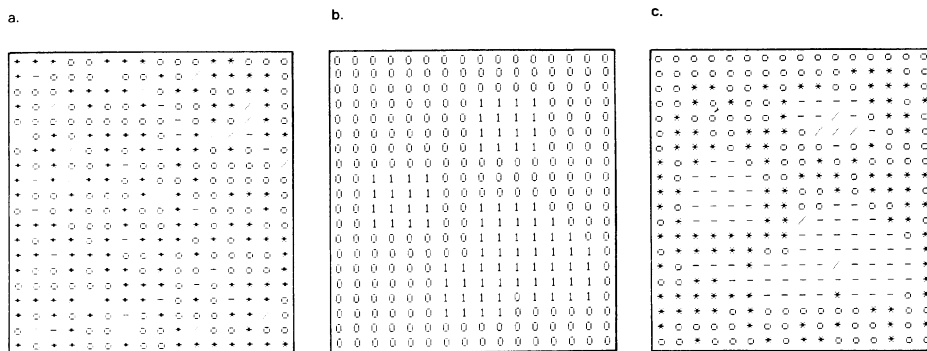
Fig. 3. Model output giving grid maps of species dominance and locations of gopher disturbance.
a. Initial pattern of dominance generated by random allocation of species abundances to each square: *Lasthenia californica* (*, dominant in 136 squares); *Plantago erecta* (o , dominant in 147 squares); *Bromus mollis* (−, dominant in 19 squares); *Calycadenia multiglandulosa* (/, dominant in 18 squares).
b. Location of randomly-generated gopher mounds; 6 mounds each 4×4 squares in area: 0, undisturbed; 1, gopher mound (i.e. no vegetation).
c. Pattern of dominance in the year following gopher disturbance shown in (b): *Lasthenia californica* (*, dominant in 111 squares); *Plantago erecta* (o , dominant in 128 squares); *Bromus mollis* (−, dominant in 75 squares); *Calycadenia multiglandulosa* (/, dominant in 6 squares).

bances using a non-interactive model, and recent studies on artificial soil disturbances (Rapp & Rabinowitz, 1985; Rabinowitz & Rapp, 1985a, b) have suggested that colonization may indeed be non-interactive. They suggested that species act independently of one another and very little competitive 'sorting out' takes place. The importance of species interactions in the serpentine grassland requires further investigation, and we plan to incorporate interactions in future model trials.

It should also be noted that the inclusion of a random element into the seed production and dispersal functions proved essential for the successful simulation of community dynamics. Ågren & Fagerström (1984) suggested that stochasticity in seed production may allow species with similar competitive abilities to coexist. This may then be a factor allowing the dominant species to coexist in the annual grassland.

Longer term runs were also performed to provide predictions for long-term experimental studies. Here we present results from runs representing three simple situations:

1. No gopher disturbance
2. One disturbance only
3. Repeated (annual) disturbance.

Fig. 4 illustrates the simulated dynamics of the four-species grassland community for each of these conditions (shown as numbers of squares dominated by each species at each time). Numerous simulations indicate that the results are not dependent on initial species distributions or abundances. With no gopher disturbance *Lasthenia* and *Plantago* dominate the entire area, with *Bromus* and *Calycadenia* present at very low levels of abundance (Fig. 4a). Even when the initial abundance of *Bromus* or *Calycadenia* is high, this declines very quickly and no spatial patterning is obvious. Where one disturbance event is added, there is a temporary increase in the abundance of *Bromus* as it becomes dominant on the newly formed gopher mounds (Fig. 4b). However, as *Lasthenia* and Plantago reinvade the gopher mounds, *Bromus* loses dominance and any patterning is lost after 3 – 4 years. Where there is annual gopher disturbance (Fig. 4c) dominance is no longer concentrated in *Plantago* and *Lasthenia,* and *Bromus* and *Calycadenia* assume much greater importance. The model predicts a dynamic relationship between *Bromus* and *Calycadenia* with *Calycadenia* becoming dominant on dis-
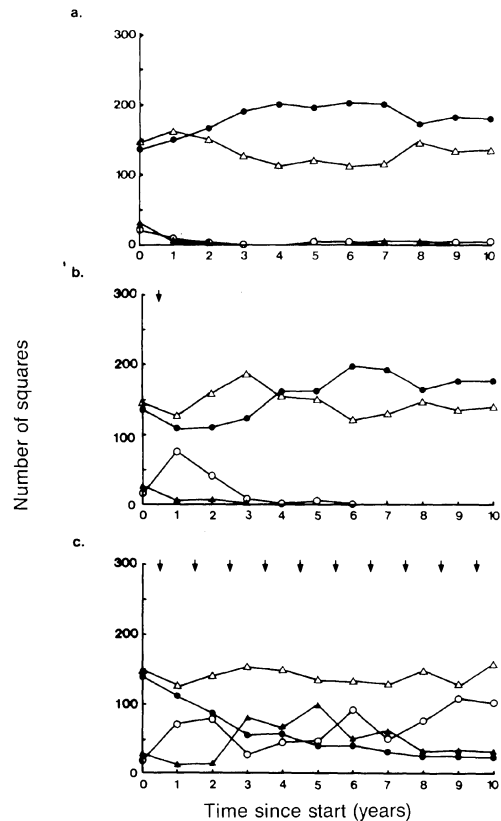
Fig. 4. Simulated long-term dynamics of a four-species grassland community with and without gopher disturbance: numbers of grid squares dominated by each species against time since start (years). a. No disturbance; b. Single disturbance (indicated by arrow); c. Annual disturbance (indicated by arrows). Symbols: ● *Lasthenia californica;* △ *Plantago erecta;* ○ *Bromus mollis;* ▲ *Calycadenia multiglandulosa.*

turbed areas only after 3 years of annual disturbance. Thereafer, patterns of dominance by *Bromus* and *Calycadenia* fluctuate. Overall, species patterning depends on the location of gopher mounds relative to those formed in previous years. Large mounds (i.e. where several mounds combine) or mounds formed in the same location for consecutive years remain dominated by *Bromus* or *Calycadenia* much longer.

Clearly, community composition and patterning is strongly affected by the addition of gopher disturbance. The effects of a single disturbance are relatively short-lived, however, and it is only when

146

disturbance is repeated that changes in the community persist. The results suggest that *Calycadenia* can form a significant part of the community only after several years of disturbance. We can therefore hypothesise that the present high level of *Calycadenia* abundance at the study site (Hobbs & Mooney, 1985) is due to the high level of gopher activity. We might therefore predict that where gophers are excluded, both *Calycadenia* and *Bromus* will decline in importance. This could however, take several years to become apparent due to lag effects from disturbances prior to exclosure.

The results of our studies fit with the more general conclusion that small-scale disturbance can play an important role in the dynamics of communities; e.g. Coppock *et al.* (1983), Collins & Barber (1985), Sousa (1985), Connell & Keough (1985).

The results discussed here represent preliminary trials of a model which can be expanded to explore many other aspects of grassland dynamics. The longer-term predictions of the model have not yet been tested, but data from long term exclosure trials at the study site will provide such a test. Further steps will be to integrate the community model into a broader ecosystem model (as attempted, for instance, by Grant & French, 1980); to check whether disturbance by gophers allows perennial bunch grasses to establish and persist in the annual community; and to explore the effects of frequency, timing and intensity of gopher disturbance on the plant community. Extending the model to a longer timescale could also allow the inclusion of fluctuations in the gopher populations and their effects on community dynamics.

Simulation of community patterning and dynamics using population processes provides a means of moving between different levels of the ecological hierarchy (Allen & Starr, 1982; Allen, 1986). By defining parameters at one level (population) we have simulated dynamics at the next (community). Although here we have modelled a relatively simple community of four species, the principle can be extended to more complex situations.

References

Ågren, G. I. & Fagerström, T., 1984. Limiting dissimilarity in plants: randomness prevents exclusion of species with similar competitive abilities. Oikos 43: 369–375.

Allen, T. F. H., 1986. Hierarchical complexity in ecology: a noneuclidean conception of the dataspace. Vegetatio 69: 17–25.

Allen, T. F. H. & Starr, T. B., 1982. Hierarchy: Perspectives for Ecological Complexity. University of Chicago Press, Chicago. 310 pp.

Collins, S. L. & Barber, S. C., 1985. Effects of disturbance on diversity in mixed-grass prairie. Vegetatio 64: 87–94.

Connell, J. H. & Keough, M. J., 1985. Disturbance and patch dynamics of subtidal marine mammals on hard substrata. In: S. T. A. Pickett & P. S. White (eds), The ecology of natural disturbance and patch dynamics, pp. 125–151. Academic Press, New York.

Coppock, D. L., Dettling, J. K., Ellis, J. E. & Dyer, M. I., 1983. Plant-herbivore interactions in a North American mixed-grass prairie. Oecologia (Berlin) 56: 1–9.

Grant, W. E. & French, N. R., 1980. Evaluation of the role of small mammals in grassland ecosystems: a modelling approach. Ecol. Modelling 8: 15–37.

Hobbs, R. J. & Mooney, H. A., 1985. Community and population dynamics of serpentine grassland annuals in relation to gopher disturbance. Oecologia (Berlin) 67: 342–351.

McNaughton, S. J., 1968. Structure and function in California grasslands. Ecology 49: 962–972.

Pickett, S. T. A. & White, P. S., 1985. The ecology of natural disturbance and patch dynamics. Academic Press, New York. 472 pp.

Platt, W. J., 1975. The colonization and formation of equilibrium plant associations on badger disturbances in a tall-grass prairie. Ecol. Monogr. 45: 285–305.

Rabinowitz, D. & Rapp, J. K., 1985a. Colonization and establishment of Missouri prairie plants on artificial soil disturbances II. Detecting small-scale plant-to-plant interactions and separating disturbance from resource provision. Amer. J. Bot. 72: 1629–1634.

Rabinowitz, D. & Rapp. J. K., 1985b. Colonization and establishment of Missouri prairie plants on artificial soil disturbances III. Species abundance distributions, survivorship, and rarity. Amer. J. Bot. 72: 1635–1640.

Rapp, J. K. & Rabinowitz, D., 1985. Colonization and establishment of Missouri prairie plants on artificial soil disturbances I. Dynamics of forb and graminoid seedlings and shoots. Amer. J. Bot. 72: 1618–1628.

Sheldon, J. C. & Burrows, F. M., 1973. The dispersal effectiveness of achene-pappus units of selected compositae in steady amids with convection. New Phytol. 72: 665–675.

Sousa, W. P., 1984. The role of disturbance in natural communities. Ann. Rev. Ecol. Syst. 15: 353–391.

Sousa, W. P., 1985. Disturbance and patch dynamics on rocky intertidal shores. In: S. T. A. Pickett & P. S. White (eds), The ecology of natural disturbance and patch dynamics, pp. 101–124. Academic Press, New York.

Van Tongeren, O. & Prentice I. C., 1985. A spatial simulation model for vegetation dynamics. Vegetatio 65: 163–173.

Description and simulation of tree-layer composition and size distributions in a primaeval *Picea-Pinus* forest

Rik Leemans & I. Colin Prentice[*]
Institute of Ecological Botany, University of Uppsala, Box 559, S-751 22 Uppsala, Sweden

Abstract

A structural description of a 0.5 ha stand in the primaeval coniferous forest of Fiby in central Sweden is reported, and a simulation model is used to test the hypothesis that the diameter and height distributions of the two most important tree species *Pinus sylvestris* L. (shade-intolerant) and *Picea abies* (L.) Karst. (shade-tolerant), are the outcome of simple successional processes following storm-felling 190 yr previously. The simulation model, FORSKA, is a 'gap model' treating light and other resource availability as homogeneous within patches. Simple formulations of dimensional relationships, vertical leaf area distributions within trees, light attenuation through the canopy and net assimilation as a function of light intensity allow FORSKA to simulate the species mixture, size distributions and vertical canopy structure of mixed-age, mixed-species forests. Parameters of the model were estimated from literature and from field data on height-diameter relationships, establishment rates and maximum ring-widths. The simulation generated a stand description similar to the real one in all essential characteristics. FORSKA allows vertical overlap between the crowns of nearby trees. The present simulation also differed from many gap model simulations in that the patch size was much greater than the maximum crown size, consistent with field observations that single treefall gaps have little effect; the persistence of shade-intolerant species in boreal forests generally depends on larger-scale disturbance events, such as large storm-gaps and forest fires.

Introduction

Fiby Urskog (59°54′N, 17°22′E) is one of the best preserved examples of a natural southern boreal/ boreo-nemoral coniferous forest in Sweden (Sjörs, 1965), located within a matrix of managed forests and farmland in the county of Uppsala. Although the forest is believed to be primaeval (in the sense that it has never been logged), its present state is thought to be the outcome of succession following major disturbance in 1795 by a wind storm that affected a large part of the county (Sernander, 1936; Hytteborn & Packham, 1985). We present data on the composition and structure of a 0.5 ha stand representative of one of the main forest types, and we use a simulation model to test the hypothesis that a few key processes – sapling establishment, tree growth, natural mortality and competition for light between species with different shade tolerances – are sufficient to predict the course of succession after the disturbance, and thus explain the present characteristics of the tree layer.

This test is also the first trial of the simulation model FORSKA, a general-purpose forest dynamics model under development at Uppsala University. FORSKA is a 'gap model' (Shugart, 1984) in the sense that the stand is simulated as an array of patches, each characterized by a distribution of individual trees. The processes by which the trees affect the environment of the patch (which in turn influence the growth of the trees) are formulated at a level of complexity appropriate to semi-quantitative prediction of mixed-age, mixed-species forest dynamics on a time scale of tens to hundreds of years (Prentice, 1986a).

The patch size in gap models is chosen such that light and other resources (averaged over the growing season) can reasonably be considered homogeneous horizontally within patches. The current state of the forest on a given patch determines the local resource environment, which in turn affects the state of the patch at the beginning of the next timestep. Establishment and mortality are treated as stochastic processes; growth is modelled using dimensional relationships in conjuncton with deterministic equations relating the net assimilation of individual trees to site characteristics, current biomass, and the current vertical profile of light intensity on the patch.

Most gap models are derived from JABOWA (Botkin et al., 1972), but the formulations of processes in JABOWA are by no means unique. FORSKA is simpler. It is strictly a successional model, i.e. it includes only the minimal set of processes required to simulate the internal dynamics of the tree layer. It is also purely deterministic, except for the stochastic simulation necessary to convert intensities and probabilities into events affecting individual patches and individual trees; unlike JABOWA and its descendants, FORSKA does not include arbitrary stochasticity that can conceal lack of fit between simulations and data. This simplicity makes FORSKA an appropriate tool for hypothesis testing. FORSKA also incorporates different formulations of growth, designed to improve comparability with physiological data (Prentice, 1986b). These differences are mostly not important for the present paper. One major difference, however, is in the approximation used for the verti-

cal structure of the forest. JABOWA treats each tree's foliage as being concentrated at the top of the tree. This approximation may be acceptable for many temperate or tropical broad-leaved trees, but it is inappropriate for coniferous forests at high latitudes, where low light angles, combined with pyramidal crown shapes, cause a much more even sharing of light among trees. FORSKA distributes each tree's foliage uniformly in the vertical down to a level determined by light intensity, and therefore allows an explicit simulation of the vertical structure of the forest.

No full description of FORSKA has yet been published. A complete mathematical specification would be out of place here; instead, we give a description of the basic principles and components of the model. A self-explanatory listing of the FORTRAN 77 code is available on request. A partial description also appears in Prentice (1986b).

Site description

Fiby Urskog is a nature reserve. Hytteborn & Packham (1985) describe the forest and its documented history. Its total area (64 ha) consists of a mosaic of two main forest types. One is a slow-growing, lichen- and other cryptogam-rich, *Pinus sylvestris* L. (Scots pine) forest, occurring on thin soils on granite ridges. The other is a more herb-species-rich *Picea abies* (L.) Karst. (Norway spruce) forest with scattered *Pinus, Betula pubescens* Ehrh. and *B. pendula* Roth. (birch), and *Populus tremula* L. (aspen), occurring on deeper soils over till and more recent sedimentary deposits. The tree population of the *Picea* forest is uneven-aged with standing dead trees and many rotten logs in various stages of decomposition on the forest floor. Few trees are older than 250 yr and many suffer from root or butt rot.

Methods

The model

FORSKA simulates stand dynamics through independent replicate simulations of discrete patches. The current state of each patch is defined by a list of simulated trees, each characterized by its species diameter at breast height (DBH) and height to the base of live crown (i.e. lowest foliage height, HBC). The entire state description of the stand is updated at each timestep, by a sequence of establishment, growth (including leaf and branch shedding), and mortality.

Establishment is stochastic. The species that establish themselves are those that can grow at the current average light intensity reaching the forest floor. Saplings of these species are established with an arbitrarily small DBH (here set to 1.0 cm). The number established follows a Poisson distribution with a fixed intensity. Establishment of new saplings on a patch is followed by growth of all trees on the patch. Various tree dimensions (DBH, height, basal area, stem biomass and leaf area) are related to one another by simple functions. Growth increases all of these dimensions as a result of a tree's net assimilation integrated in successive layers through the crown, scaled by a 'growth constant' (a function of species and site) and reduced by a factor related to total patch biomass (standing for resource depletion). (In the present version the growth constant is prescribed for each species for the particular site, but for some future applications it will be made an explicit function of species and environment.) The light response curve of net assimilation rate for each species is specified by three parameters: a maximum (light-saturated) rate, a compensation point and a 'half-saturation point'. The maximum rate is absorbed into the growth constant; the other two parameters define the species' shade-tolerance characteristics, including the requirement that the light intensity at the forest floor must be greater than the compensation point if the species is to establish. Competition for other resources is thus modelled in the same simplified way as in JABOWA: Growth rates of all individuals on a patch are slowed as total patch biomass approaches a ceiling. Competition for light is modelled more explicitly as a function of each tree's vertical distribution of leaf area in relation to the distribution of leaf area index for the entire patch and the species-specific light response of net assimilation rate. The growth efficiency of individual trees is also reduced as they become large, on the crude assumption that there is a 'cost factor' for total above-ground dry weight.

Mortality is a second stochastic process. The probability of death for an individual tree is inversely related to its growth increment during the current timestep. Above a pre-defined threshold growth rate the simulated probability of death is low and essentially growth-independent, standing for death due to various external catastrophes, such as windthrow. Below this threshold the simulated probability of death is much higher. (The relationship is a step function in most gap models; in FORSKA it is a reverse sigmoid function, the steepness of the transition is being controlled by a parameter.) The higher simulated probability of death of very slow-growing trees stands for death due to agents like pathogens whose effects are much more likely to be fatal in trees with an insufficient energy supply. Several natural mechanisms can contribute to slow growth, including the large size of old trees (through an increase in the proportion of non-photosynthesizing tissue), overshading by taller trees, sequestration of nutrient capital in the biomass, and external forces such as chronic pathogen damage and air pollution. Any of these mechanisms can indirectly cause enhanced death rate. The timestep length in FORSKA was set to 2 years since at least two successive low- or no-growth years are required to kill mature trees (Solomon et al., 1984). The model therefore assigns probabilities of death on the basis of two years' accumulated growth.

Prentice (1986b) discusses specific differences between FORSKA and the other gap models. Two differences relevant to this paper are:
1. The function most often used to describe the relation between DBH and height of individual trees is a second degree polynomial (Trorey, 1932; Ker & Smith, 1955; Botkin et al., 1972). FORSKA uses a Mitscherlich curve (Meyer, 1940). The polynomial has a maximum at a cer-

tain DBH and therefore cannot be applied to trees larger than this DBH; the growth equations in JABOWA are defined in such a way that such trees cannot exist. FORSKA, however, does not limit the maximum DBH trees can reach and therefore requires an asymptotic function. (The polynomial can be obtained from the Taylor series expansion of the Mitscherlich function and is a fairly good estimator of it over the range in which both curves are rising, cf. Fig. 2.)

2. FORSKA allows vertical overlap among the crowns of trees on the same patch. The leaf area of each tree is distributed uniformly between the total tree height (H) and HBC, which is initially zero. Tree growth is computed from the sum of net assimilation through successive vertical layers; the light available in each layer is determined from the cumulative leaf area index in all higher layers. The vertical integration step was set to 0.5 m. The growth of all trees on the patch can result in some trees (especially shade-intolerant trees) having their lowest leaves below compensation point, and therefore unable to support themselves. These leaves are dropped, and HBC rises to the level at which all leaves can support themselves. This process is irreversible, mimicking self-pruning. Trees with HBC > 0 have correspondingly less total leaf area for a given DBH.

The test data

We established a 0.5 ha plot (100 m × 50 m) within the *Picea* forest in the western part of Fiby Urskog. The dominant tree species in the canopy here is *Picea; Pinus* is less abundant (Table 1). There are only a few individuals of *Populus* and *Betula*. The latter two taxa are as shade intolerant as *Pinus;* to simplify the model test we lumped them with *Pinus*.

All individual trees were mapped, numbered and their coordinates within the plot determined. Total tree height, DBH, and HBC were measured on each individual. Height measurements were made with a Suunto clinometer. A height correction was made for abnormally leaning individuals. Small trees (1.3 m < H < 5.0 m) were measured directly with a measuring stick.

Ground level was measured with a theodolite at all corner points of a regular 2 m × 4 m grid covering the sample plot, and was found to vary through 5.6 m altogether. The relative base

Table 1. Some characteristics of the tree-layer of a 0.5 ha stand in the *Picea* forest at Fiby Urskog. 1: number of individuals > 1.3 m, 2: basal area (m² ha⁻¹), 3: maximum height (m), 4: mean height and range to the base of live crown (m).

Species	1	2	3	4
Picea abies	381	27.7	28.7	1.7 (0.0 – 8.3)
Pinus sylvestris	26	5.6	26.3	14.8 (10.1 – 19.0)
Betula spp.	8	1.2	23.5	12.3 (7.6 – 18.4)
Populus tremula	12	2.5	26.4	15.5 (11.7 – 20.8)
Total	427	37.0		

height (relative to the lowest point within the plot) of all trees was computed using a second degree polynomial regression on relative height between neighbouring grid-points and an interpolation from this regression towards the coordinates of each individual tree. We suspected that the upper canopy surface formed by all of the tree tops would have a smoother profile, implying less variation in relative height, than the ground surface. The measured heights of neighbouring trees might then differ because of variation in relative ground level, despite their crown tops reaching the same height in the canopy. The individual heights of canopy trees were therefore corrected for variation in ground level, and both uncorrected and corrected values were used as alternatives in estimating relationships between DBH and height (Table 2).

Ring counts were made from increment borings at the lowest possible height on several of the largest *Picea* and *Pinus* trees. These yielded ages between 175 and 190 years, consistent with the date of 1795 for a major disturbance to the forest.

Parametrizing the model

Values of 'site' parameters were taken from literature. Light extinction coefficients in mature forest stands are typically 0.3 – 0.5 in mid- to high-latitudes (Kira, 1975; Jarvis & Leverenz, 1983); we used a value of 0.4. Intensity of photosynthetically active radiation above the canopy was set to 400 μmol m⁻² s⁻¹, an approximate average value for the region during the growing season (Linder & Lohammer, 1981; Ågren, 1983). The conversion factor between above-ground biomass and the product (DBH)² × H was set at 0.025 kg (cm² m)⁻¹, roughly consistent with data of Hytteborn (1975) and Nihlgård & Lindgren (1977) from various types of forests in Sweden. Maximum biomass was set at 500 Mg ha⁻¹.

Table 3 gives a complete list of 'species' parameters as used in the routines for establishment and growth. The values in Table 3 were obtained partly from literature, partly from field observations.

Parameters of the height-DBH relationship were estimated by regression analysis of (uncorrected and corrected) tree height

Table 2. Regression of height on DBH in the sample plot: u = uncorrected, c = corrected values.

	model: $H = 1.3 + b\mathrm{DBH} + a\mathrm{DBH}^2$						model: $H = 1.3 + b(1 - e^{(-a\mathrm{DBH})})$				
	a (1) (m^{-1})	b (1)	R^2 (%)	max. height (m)	max. DBH (2) (cm)	initial slope (m cm^{-1})	a (1) (m^{-1})	b (1) m	R^2 (%)	max. height (m)	initial slope (m cm^{-1})
Picea u	-78.9 ± 5.6	91.5 ± 1.9	97.6	27.8	58	0.92	2.3 ± 0.4	40.2 ± 2.8	97.9	41.5	0.92
Picea c	-81.5 ± 5.2	93.3 ± 1.6	97.9	28.0	57	0.93	2.3 ± 0.2	41.6 ± 2.5	98.8	42.9	0.96
Pinus u	-70.6 ± 15.7	83.5 ± 6.3	98.2	26.0	59	0.84	2.9 ± 0.8	3.29 ± 5.5	97.8	34.2	0.95
Pinus c	-95.5 ± 16.1	100.6 ± 6.5	99.1	27.8	53	1.01	3.3 ± 0.7	33.7 ± 4.5	98.5	35.1	1.11

(1) ± standard error
(2) DBH at which maximum height is reached

and DBH data from the sample plot. Quadratic functions were fitted by standard polynomial regression and Mitscherlich functions by non-linear least-squares (Table 2, Fig. 1). There was little to choose between the alternative analyses in terms of goodness of fit to the data (all R^2 values were better than 97%). However, the R^2 values for the corrected height data were consistently slightly higher than those for the uncorrected height data. The polynomial regression yielded estimated maximum heights in the same range as the maximum observed heights on the plot (25 m–30 m), corresponding to trees of 50–60 cm DBH. In contrast, the non-linear regressions gave estimated maximum (asymptotic) heights of ca 35 m (*Pinus*) and ca 45 m (*Picea*), consistent with maximum heights observed in Sweden generally. We concluded that the Mitscherlich function is a more appropriate one for simulation purposes, and used estimated initial slopes and maximum heights from the non-linear regressions on corrected height data as values of the parameters describing the height-DBH relationship in the model.

The model also relates leaf area to (DBH)^2 by an asymptotic function. We set the initial slope of this function to 0.08 m^2 cm^{-2} for *Pinus* (based on data described statistically by Albrektsson, 1980) and 0.16 m^2 cm^{-2} for *Picea* (based on data in Reichle, 1981). Maximum leaf areas per tree were set to reasonable values of 125 m^2 (*Pinus*) and 250 m^2 (*Picea*) i.e. differing by the same factors as the slopes. Light compensation points were computed to give a maximum possible leaf area index of 2.5 for *Pinus* and 5.0 for *Picea*, consistent with data applying to these latitudes given by Cannell (1982). Half-saturation points were set at 330 μmol m^{-2} s^{-1} for *Pinus* (Troeng & Linder, 1982) and 100 μmol m^{-2} s^{-1} for *Picea* (Fuchs *et al.*, 1977).

Growth constants can be estimated from initial annual-ring widths for trees growing without significant overshading if the values of dimensional and light-response parameters are known. The oldest generation of trees in Fiby Urskog is presumed to have grown up in open conditions. Ring-widths for the first five years of growths of mature trees on the plot averaged ca 0.2 cm for both species. Given the values of the other parameters (Table 3), we computed growth constants of 78 cm^2 m^{-2} yr^{-1} for

Pinus and 11 cm^2 m^{-2} yr^{-1} for *Picea* from these ring widths. The difference may reflect the greater maximum photosynthetic rates typical of shade-intolerant plants. Finally, a guess at the cost factor for size was made by assuming that the carbon equivalent of ca 1/8 of the relative growth rate of mature trees forming a closed canopy is lost as respiration of non-photosynthesizing tissue (Daniel *et al.*, 1979).

The establishment rate of *Picea* was determined indirectly from the annual height increments of a number of saplings as determined by the height differences between the annually produced branch-rings and/or budscars, together with data on the heigth-class distribution of saplings, according to Nakashizuka (1983). Because of the absence of *Pinus* saplings in the sampling

Table 3. Species parameters in FORSKA and their values used in the test simulation.

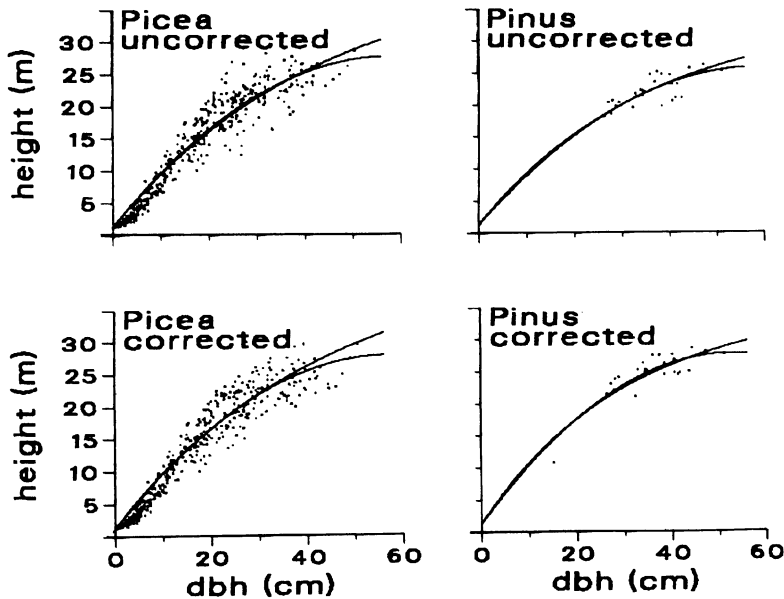|	*Pinus*	*Picea*	
Maximum height	35.1	42.9	m
Initial slope of height vs DBH	1.11	0.96	m cm^{-1}
Half saturation point	330	100	μmol m^{-2} s^{-1}
Compensation point	59	22	μmol m^{-2} s^{-1}
Growth constant	78	11	cm^2 m m^{-2} yr^{-1}
Maximum leaf area per tree	125	250	m^2
Initial slope of leaf area vs DBH^2	0.08	0.16	m^2 cm^{-1}
Cost factor for size	0.0033	0.0011	yr^{-1}
Establishment rate	20	20	ha^{-1} yr^{-1}

152

Fig. 1. Height-diameter relationships for *Picea abies* and *Pinus sylvestris* in the sample plot. The lower graphs are based on heights corrected for variation in ground level (see text). The curves are based on the regression analysis reported in Table 2.

plot, establishment rate for *Pinus* was set arbitrarily to the same value as for *Picea*.

Mortality parameters were set to values in the same order of magnitude as those used in other successful application of gap models. As with the establishment rates, no distinction was made between the two species. The minimum, growth-independent death rate was set to 0.0046 yr^{-1}, giving a hypothetical 1% chance of survival to 1000 years for trees growing with no external or internal constraints on growth. The maximum death rate (for non-growing trees) was set to 0.46 yr^{-1}, allowing a 1% chance of survival to 10 years. The critical DBH increment around which the death rate change most steeply was set to 0.04 cm^2 yr^{-1}; the parameter determining the steepness of the transition was set to 3.0, giving a fairly sharp change comparable with the step function used in other gap models.

The test

The simulated 0.5 ha stand was divided into five patches of 0.1 ha. We began the simulation with no trees, and ran the model for 190 simulated years, giving a stand description for comparison with field data from 1985. The simulation was carried out twice to give an indication of the model's 'natural variability' at a scale of 0.5 ha.

Results and discussion

Data-model comparison: Species abundances and general characteristics

Figure 2 shows the simulated successional trends (which because of the omission of *Betula* spp. and *Populus* are probably unrealistic for the first few decades). *Pinus* reached a maximum in basal area at around 100 simulated years and thereafter declined. *Picea* monotonically increased in basal area and leaf area index, but the curves were tending to level off by the end. The maximum height curves showed the kind of differences between species that would be expected from the differences in the parameters of the DBH-height relationships, given that the initial annual ring-widths of the two species were the same. *Pinus* grew faster in height during the initial model iterations, but by the end *Picea* had reached the same height as *Pinus;* if the run had been continued, *Picea* would soon have overtopped *Pinus.* The species' shade-tolerance characteristics are critical in determining their simulated
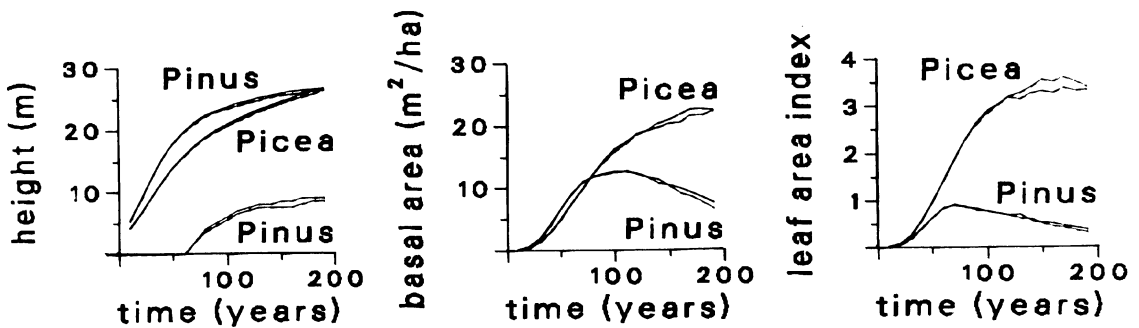
Fig. 2. Simulated trends in height, basal area and leaf area index of *Picea abies* and *Pinus sylvestris,* based on two simulations of a 0.5 ha stand composed of five replicate patches. The lower height curves for *Pinus* show the simulated increase in height of the base of live crown.

behaviour during succession. Light penetration through the canopy became insufficient for *Pinus* regeneration after about 60 simulated years; at the same time the base of the *Pinus* canopy started to rise. Mortality gradually took its toll of the remaining *Pinus* individuals, and by another 40 simulated years the total basal area of *Pinus* had reached its peak and began to decline. The simulated stand after 190 years is a realistic-looking mature *Picea* forest with some *Pinus* trees remaining from the earlier stages of succession. As in the real stand, no major gaps have yet appeared, that would allow *Pinus* and the other shade-intolerant species to establish once again.

The simulated quantities of *Picea* and *Pinus* and the overall structural characteristics of the stand (Table 4) are in reasonable agreement with the data (Table 1). The most noticeable discrepancy is in the canopy base heights of *Pinus,* which the model puts about 50% too low. This discrepancy probably

arose because the model assumes an unrealistic uniform vertical distribution of leaf area whithin each tree's crown. The qualitive difference between the canopy base heights of *Pinus* and *Picea* is correctly simulated, however. Light intensities in the lower canopy (both in the model and in the field) have reached sub-compensation point levels for *Pinus,* preventing the establishment of new *Pinus* saplings and raising the base of the *Pinus* canopy above that of *Picea.*

Data-model comparison: size distributions

The ability to simulate trends in the maximum heights and quantities of different tree species appears to be a robust feature of gap models (Shugart, 1984), but simulation of species specific size distributions provides a considerably more stringent test; there are many more ways for the

Table 4. Some characteristics of the simulated 0.5 stand.

species	number of individuals >1.3 m	basal area (m² ha⁻¹)	maximum height (m)	mean height to the base of live crown (m)	number of individuals >1.3 m	basal area (m² ha⁻¹)	maximum height (m)	mean height to the base of live crown (m)
	RUN 1				RUN 2			
Picea abies	239	22.4	26.5	0.0	250	22.2	26.2	0.0
Pinus sylvestris	40	7.7	26.3	7.8	35	6.7	26.6	8.0
Total	279	30.1			285	28.9		
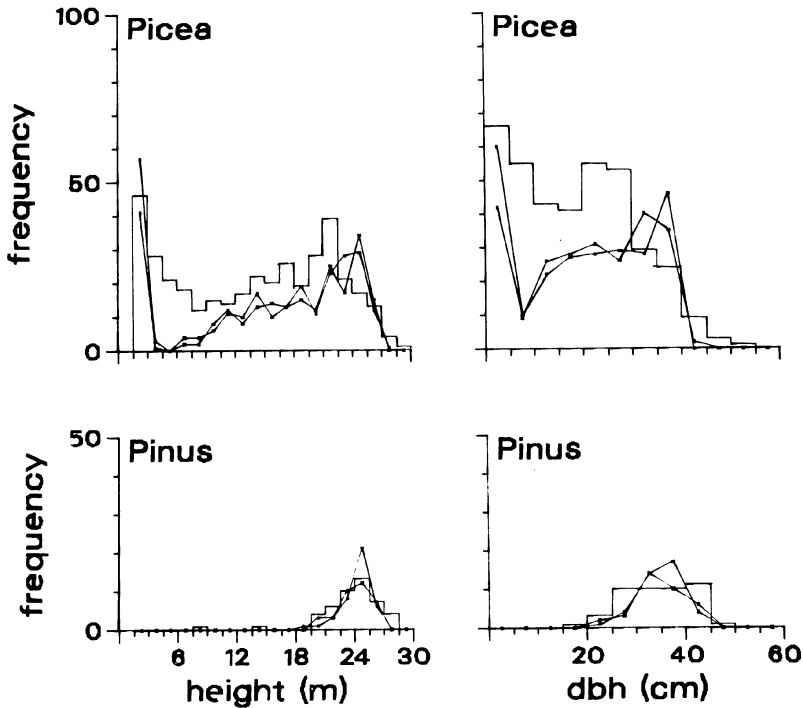
Fig. 3. Real and simulated height and diameter distributions of *Picea abies* and *Pinus sylvestris* in the *Picea* forest, Fiby Urskog. The histograms are based on data from the 0.5 ha sample plot. The lines are based on the two replicate simulations.

model to go wrong (H. H Shugart, pers. comm.) Figure 3 shows that the simulation captured the essential characteristic of the distributions of total tree height and DBH: it successfully reproduced the bimodal size distribution of *Picea* (characteristic of shade-tolerant canopy species in mature stands) and the unimodel distribution of *Pinus*, and got the size-classes frequencies for most part about right. However, the simulated forest has some of the canopy *Picea* trees too large, and also has too few *Picea* individuals in intermediate ('suppressed') size classes. This exaggerated bimodality may be a consequence of the simplification of the *horizontal* structure of the forest implied by the discrete-patch approximation. In the model, saplings that become established on a patch under an already dense canopy, closed by earlier established individuals, have a low growth increment and a near-zero probability of surviving and growing into the intermediate size classes. The horizontal structure

of the real canopy is more heterogeneous: there are many small and intermediate-sized gaps between individual tree crowns, in which newly established saplings tend to cluster (Leemans, 1986). Thus under the natural canopy, suppressed saplings have a higher probability of becoming larger trees, resulting in a somewhat less clearcut size-structure and a less deterministic temporal pattern of succession.

Our results show that the major observed differences between *Picea* and *Pinus* in importance (e.g. relative abundance), size distribution and canopy depth in the old-growth forest can all be explained as consequences of the small number of differences prescribed for the two species in the model (Table 3). To summarize these differences: (1) *Pinus* grows at first taller in relation to DBH, but reaches a lower maximum height than *Picea*. (2) The two species have light-response characteristics typical for the two primary tolerance classes. *Pinus* has a

higher compensation point and a higher half-saturation constant for light than *Picea*, and is thus less efficient at low light intensities, but is capable of much faster growth per unit leaf area than *Picea* at high light intensities. (3) Consistent with its lower compensation point (allowing a maximal leaf area index twice that of *Pinus*), *Picea* has twice as much foliage per unit basal area as *Pinus*, and a maximum leaf area per tree twice that of *Pinus*. These leaf-area differences become amplified when the two species compete in dense stands, since *Pinus* then progressively loses its lower branches. The agreement between the data and the simulation indicates that simple successional dynamics based on these few species-specific characteristics is sufficient to explain the major aspects of the present structure and composition of the tree-layer in the *Picea* forest.

Patch size and gap dynamics

Gap models are capable of simulating a wide variety of forest dynamics. They are named for their ability to simulate 'gap-phase replacement', whereby single trees eventually grow large enough to dominate their respective patches, so that the death of a single large tree can be sufficient to re-start succession with the establishment of new saplings including saplings of shade-intolerant species. This kind of dynamics is common in broad-leaved or mixed forests from temperate to equatorial latitudes. The appropriate patch size for gap models is implicitly set by the physical variables of vegetation height and average sun angle (Shugart & West, 1979). The FORET model, with a patch size of 0.083 ha, successfully simulates gap-phase dynamics in Appalachian forests at about 35°N (Shugart & West, 1977) where individual trees of several species can easily become large enough to become dominant on a patch. There is as yet no quantitative theory to fix the patch size exactly, but near 60°N the appropriate size is presumably at least 0.1 ha. Yet the largest possible *Picea* tree in our model has a leaf area of only 250 m^2 distributed through a leaf area index of 5, i.e. a crown area about 20 times smaller than the patch size. As a re-

sult, single deaths of *Picea* trees in the model do not have enough effect on light intensity to allow re-establishment of *Pinus*, and gap-phase replacement as it is usually understood cannot occur.

The absence of gap-phase replacement in the model is fully consistent with observations at Fiby and indeed with our general understanding of the dynamics of boreal forest systems, where narrow-crowned conifers dominate and sun angles are low. In Fiby Urskog, single deaths of large trees are observed from time to time; their effect is limited to a stimulation of the growth of suppressed *Picea* trees in the understorey (H. Hytteborn, pers. comm.). In contrast, gaps caused by storm-felling of >20 trees together bring sufficient light to the forest floor to allow the establishment of shade-intolerant *Pinus* and broad-leaved trees which compete with the remaining understorey trees. More generally, succession in boreal forests is typically a directional process towards dominance by the larger shade-tolerant species; disturbances on a fairly large scale (storm-felling and, especially, fire) are needed to re-start succession.

Gap models can thus simulate boreal forest dynamics as consequence of the interaction of patch size and tree size, provided that external disturbance is explicitly considered. A stronger theoretical basis for patch size would be welcome. However, model experiments in which only patch size was varied (Prentice, unpublished) proved that the structural simulation reported here was qualitatively robust with respect to patch variation between ca 0.04 and 0.2 ha. Size distributions quite different from the observed, and an excessive quantity of *Pinus*, were obtained in simulations with smaller patch sizes, confirming our original choice of 0.1 ha as a reasonable one.

References

Ågren, G. I., 1983. Nitrogen productivity of some conifers. Can. J. For. Res. 13: 494–500.

Albrektsson, A., 1980. Tallens biomass: storlek – utveckling – uppskattningsmetoder. Sveriges Lantbruksuniv. Inst. för Skogsskötsel, Rapp. 2, 189 pp.

Botkin, D. B., Janak, J. F. & Wallis, J. R., 1972. Some ecological consequences of a computer model of forest growth. J. Ecol. 60: 849–873.

156

Canell, M. G. R., 1982. World forest biomass and primary production data. Academic Press, New York.

Daniel, T. W., Helms, J. A. & Baker, F. S., 1979. Principles of silviculture. 2nd ed. McGraw-Hill, New York.

Fuchs, M., Schulze, E. D. & Fuchs, M. I., 1977. Spacial distribution of photosynthetic capacity and performance in a mountain spruce forest of northern Germany. II. Climatic control of carbon dioxide uptake. Oecologia 29: 329–340.

Hytteborn, H., 1975. Deciduous woodland at Andersby, eastern Sweden, above-ground tree and shrub production. Acta Phytogeogr. Suec. 61: 1–61.

Hytteborn, H. & Packham, J. R., 1985. Left to nature: forest structure and regeneration in Fiby urskog, Central Sweden. Arboricult. J. 9: 1–11.

Jarvis, P. G. & Leverenz, J. W., 1983. Productivity of temperate deciduous and evergreen forests. In: O. L. Lange, P. S. Novel, C. B. Osmond & H. Ziegler (eds), Physiological plant ecology IV, pp. 234–261. Springer, Berlin.

Ker, J. W. & Smith, J. H. G., 1955. Advantages of the parabolic expression of height-diameter relationships. For. Chron. 31: 235–246.

Kira, T., 1975. Primary production of forests. In: J. P. Cooper (ed.), Photosynthesis and productivity in different environments, pp. 5–40. Cambridge University Press, Cambridge.

Leemans, R., 1986. Structure of the primaeval coniferous forest of Fiby. In: J. Fanta (ed.), Forest dynamics research in western and central europe, pp. 221–230. Wageningen. PUDOC, Wageningen.

Linder, S. & Lohammer, T., 1981. Amount and quality of information on CO_2-exchange required for estimating annual carbon balance of coniferous trees. In: S. Linder (ed.), Understanding and predicting tree growth Stud. For. Suec. 160: 73–87.

Meyer, H. A., 1940. A mathematical expression for height curves. J. For. 38: 415–420.

Nakashizuka, T., 1983. Regeneration process of climax beech (*Fagus crenata* Blume) forests. III. Structure and development of sapling populations in different aged gaps. Jap. J. Ecol. 33: 409–418.

Nihlgård, B. & Lindgren, L., 1977. Plant biomass, primary production and bioelements of three mature beech forests in south Sweden. Oikos 28: 95–104.

Prentice, I. C., 1986a. Some concepts and objectives of forest dynamics research. In: J. Fanta (ed.), Proc. Forest dynamics research in western and central europe, pp. 32–41. Pudoc, Wageningen.

Prentice, I. C., 1986b. The design of a forest succession model. In: J. Fanta (ed.), Forest dynamics research in western and central europe, pp. 253–256. Wageningen, PUDOC, Wageningen.

Reichle, D. E., 1981. Dynamic properties of forest ecosystems. Cambridge University Press, Cambridge.

Sernander, R., 1936. Granskär och Fiby urskog, en studie över stormluckornas och marbuskarnas betydelse i den svenska granskogens regeneration. Acta Phytogeogr. Suec. 8: 1–232.

Shugart, H. H., 1984. A theory of forest dynamics. Springer, New York.

Shugart, H. H. & West, D. C., 1977. Development of an Appalachian deciduous forest succession model and its application to assessment of the impact of the chestnut blight. J. Envir. Managem. 5: 161–179.

Shugart, H. H. & West, D. C., 1979. Size and pattern of simulated forest stands. For. Sci. 25: 120–122.

Sjörs, H., 1965. Forest regions. Acta Phytogeogr. Suec. 50: 48–63.

Solomon, A. M., Tharp, M. L., West, D. C., Taylor, G. E., Webb, J. W. & Trimble, J. L., 1984. Response of unmanaged forests to CO_2-induced climate change: available information, initial tests, and data requirements. U.S. Dep. Energy, Techn. Rep. 009, 93 pp.

Troeng, E. & Linder, S., 1982. Gas exchange in a 20-year old stand of Scots pine. II. Variation in net photosynthesis and transpiration within and between trees. Physiol. Plant. 54: 15–23.

Trorey, L. G., 1932. Mathematical expression for the construction of diameter-height curves. For. Chron. 18: 3–14.

The separation of fluctuation and long-term change in vegetation dynamics of a rising seashore*

Wolfgang Cramer[1,2] & Håkan Hytteborn[1]** ***
[1]*Institute of Ecological·Botany, Uppsala University, Box 559, S-751 22 Uppsala, Sweden*
[2]*Present address: Department of Geography, University of Trondheim, N-7055 Dragvoll, Norway*

Abstract

Field layer vegetation in four transects across a rising seashore of the island Skabbholmen (59°47′N, 19°12′E) in the Stockholm archipelago, eastern central Sweden, was sampled at the beginning and end of a six-year period. The data were analyzed by canonical correspondence analysis (CCA) with two external predictor variables, year of sampling and elevation (a proxy for the longer-term trend of change). During the sampling period, the vegetation changed in the direction of the longer-term trend at a rate consistent with the known rate of land uplift. However, a major part of the observed change was in a different direction, reflecting a shorter-term response to disturbance. The analysis effectively separated different types of floristic dynamics related to processes with different time-scales. A comparison with (unconstrained) correspondence analysis (DCA) illustrated the interpretive advantages of multivariate direct gradient analysis over conventional floristic ordination.

Introduction

There is a large amount of data on the general trend of long-term vegetation change on rising Baltic and Bothnian seashores (Ericson & Wallentinus,

1979). These data mainly come from repeated sampling of the shore vegetation zonation during long periods (11–30 years, e.g., Brunberg-Schwanck & Bärlund, 1948; Schwanck, 1974; Ericson, 1980), and repeated floristic inventories on rising islands (e.g., Valovirta, 1937; Luther, 1961; Vartiainen, 1980). On a time-scale of tens to hundreds of years, field layer species populations have been shown to 'migrate' downwards along the sea-land gradient at about the same speed as new land emerges from the sea (Ericson, 1981). Also establishment of more long-lived woody species tracks the downward trend in relative sea level (Cramer, 1985; Verwijst & Cramer, 1986).

On a shorter time-scale (<10 years), the shoreline does not move smoothly, but is subject to fluctuations caused by different years' weather conditions. During short periods, sea level may be seen more as providing a disturbance regime with irregu-

larly recurring catastrophic destruction of the established vegetation (White, 1979) than as a continuous environmental change. Some data have been published on the short term responses of species populations to sea level changes (Ericson, 1980, 1981). We present here an application of a general technique to study such changes on a community level and to see whether the long-term trend of vegetation composition (as reflected in the spatial changes along the transect) can be separated from the short-term changes observable during six years.

Analysis of vegetational variation in space and time

Multivariate analyses of floristic data from more than one point in time have followed two principal lines:

1. The 'classification approach' clusters the samples and estimates transition frequencies between clusters, either based on the structure of the first analysis [e.g., Williams *et al.*, 1969; Londo, 1974; Persson, 1984], or on the structure of all samples at all sampling dates (e.g., Van Noordwijk-Puijk *et al.*, 1979; Persson, 1980; Austin & Belbin, 1981; Zhang, 1984). The two variants have been discussed by Williams *et al.* (1969). Transition diagrams often form comprehensible summaries of such studies, but it may be difficult to apply the approach to data with continuous floristic variation.

2. The 'ordination approach' locates all samples (from all sampling dates) in a multidimensional space on the basis of differences in species composition. The location of the samples from the same site but from different dates are compared graphically, e.g. by drawing time trajectories (Van der Maarel, 1969, 1975, 1980; Austin, 1977; Persson, 1980, 1984; Van Hecke *et al.*, 1981; Zhang, 1984; Gunnlaugsdottir, 1985). Such diagrams give an impression of the relative rate of change of species composition of the sites and of the increase or decrease of individual species' abundances. Information about the underlying

processes can be gained by correlating the ordination axes with known environmental gradients ('indirect gradient analysis').

This approach is in fact a simplification of a three-dimensional (species × sites × times) matrix (c.f., Williams & Stephenson, 1973). Swaine & Greig-Smith (1980) noted the difficulty in separating temporal change from environmentally determined gradients in such ordinations and provided a mathematical method to overcome this.

An extension of the ordination approach is provided by a new technique for multivariate direct gradient analysis, namely canonical correspondence analysis (CCA – Ter Braak, 1985, 1986, 1987). CCA combines the analysis of floristic gradients with external (often environmental) data for the sites. The axes extracted by CCA represent those directions of variation in species composition that are related to supplied external variables. Like its parent method, correspondence analysis (Hill, 1979a; Hill & Gauch, 1980), CCA can be combined with a detrending to remove arch effects.

In our study, the continuous appearance of both the spatial floristic gradients and the long-term trend, suggested the application of a gradient analysis technique. We therefore applied (detrended) CCA (DCCA) with two external variables, standing for two scales of change: (1) elevation (in the first sampling season – 1978) as a proxy for each sample's location on the long-term trend, and (2) year with the states '0' for the first sampling 1978 and '6' for the second in 1984, as an index of the actual change during our six-year period of observation. Thus, we aimed for an ordination of the entire data which would explicitly indicate the extent of congruence between floristic variation related to elevation (and therefore to the long-term trend caused by land uplift) and to shorter-term changes in the environment.

Materials and methods

Study area

The island Skabbholmen is situated in the northern Stockholm Archipelago (59°47′N, 19°12′E). The region falls into the

boreo-nemoral zone (Sjörs, 1963) and belongs to the inner archipelago ('the coniferous forest zone', cf. Du Rietz, 1925).

Vegetation sampling

In 1978 and 1984 we sampled the field layer vegetation of (in total) 135 1-m^2 plots (in the following called 'sites') along 4 transects, all running from slightly below mean water level into old-growth forest. The transects were chosen to represent different degrees of exposure and different vegetation gradients. Under the tree canopy, the vegetation was sampled in June, when the field layer of these spring − geophyte woodlands is maximally developed. The shore vegetation was sampled in early August. The sites were placed adjacent to each other in the open vegetation of the shore and the nearby woodland, and with 5 m-spaces in the forest. In each site, we noted vascular plants with cover degrees of the ordinal five-degree Hult − Sernander − Du Rietz scale. The total moss cover was estimated in %.

Geomorphic dynamics of the shore and fluctuations in water level

The eastern shore of Skabbholmen is exposed to 'Havssvalget', a several km wide strait where relatively high storm waves can develop. Most of this shore belongs to the type 'moderately exposed shore' (Cramer, 1980) and is characterized by a coarse, stony substrate indicating that erosion is more common than accumulation in the wave-wash zone. There are no conspicuous cliffs. In shallow bays, some sandy and organic material accumulates.

No tides with relevance to shore vegetation occur in the Baltic. Seasonal water level fluctuations usually result in a continuous lowering of water level during winter, with a minimum around May, then a rather rapid increase during summer, followed by an unstable period during autumn and the yearly maximum around the end of the year. Differences from year to year are great (Fig. 1). On a longer time-scale, post-glacial isostatic land uplift causes an apparent lowering of sea-level, at a rate of ca. 5.0 mm/yr in the study area (calculated from Åse (1964). Due to eustatic changes, this relative land uplift rate has changed during historical times, but the differences have been small during the last 200−300 years (Åse, 1980).

Sampling of environmental data

At both sampling dates (1978 and 1984), we estimated percentage of cover of stones (>2 cm) at the soil surface, of (allochthonic) drift material and of (autochthonic) plant litter. We measured elevation above present water level for each site in August 1984. For comparison of the habitat characteristics of seashore plants during the period of 1975−1984, we compared water levels of the period August 1975−July 1978 with those of the period August 1981−July 1984, i.e. for the three-year-periods prior to each of our samplings (Fig. 1). Data were obtained from the tide-gauge 'Forsmark', 85 km NNW of Skabbholmen. The data were transformed for the difference in relative land uplift between the tide gauge and the study area, by subtraction of (1.1 × 'No. of years since 1970') mm.

Data analysis

To remove the largest and most obvious part of floristic heterogeneity in the data set, we separated the main matrix into two subsets, 'shore' and 'forest'. All subsequent calculations were based on these two subsets, comprising (after removal of 3 sites that had no vegetation at either sampling date) 63 and 69 sites respectively. The borderline between these units was drawn on the basis of a TWINSPAN − analysis (Hill, 1979b) of the 1978 material.

Our application of CCA involved the same straightforward matrix simplification as was discussed above (each observation = one sample), but now combined with the treatment of elevation and year as external variables. We treated change along the elevation gradient as an estimate of the long-term average rate of vegetation change (evidence supporting this assumption for woody species populations in the same environment is given by Cramer (1985, 1986); Verwijst & Cramer, (1986). The introduction of elevation as an external variable in CCA required the ordination scores to be a linear combination of both elevation and year. Note that elevation had been measured only once − this implies that elevation and year were essentially uncorrelated in this treatment. Elevation and year can thus be seen as expressions of floristic change on different time-scales: elevation corresponding to the longterm trend caused by land uplift, and year referring to the actually observed change in species composition during a six year period. Slopes and approximate significances of the relations between these variables and the ordination scores are given in the standard output.

Investigation of the regression between species scores and the two external variables should then allow one to see to what extent the overall change in floristic composition during six years is due to the long-term trend or to other processes causing fluctuations on shorter time-scales. A more spatially detailed analysis can be made by comparing locations for sites in the two different years in the ordination diagram and relating these time trajectories to more local changes in environment.

160

Table 1. Transect 3, comparison table. In all columns the left entry is the 1978 value, the right one is from 1984. Species abundances according to the ordinal five-degree Hult − Sernander − Du Rietz scale. Plot numbers indicate distance from lowest plot in transect. Submergence calculated as percentage of all daily means during three years prior to each sampling.

```
Plot number                     1  2  3  4  5  6  7  8  9  10 11 12 13 14 15 16 17 18 19 20    21 22 23 24 25 26 27 28 29 30 35 40 45 50

Aster tripolium                11 11 11 11 11 11 11 1. 11 1. 1.
Triglochin maritima            11 11 11 11 2. 11 31 11 11 11 11 ..
Agrostis stolonifera           11 11 11 11 11 11 21 21 11 11 .1 .1 11 .. .1 11 1.
Deschampsia bottnica           .. 11 .1 .1 11 11 31 31 1. 11 ..
Eleocharis uniglumis           .. 11 11 21 31 21 21 41 11 1. 11 1.
Glaux maritima                 .. .. .1 11 11 11 11 12 12 12 11 11 11 11 1. 1.
Juncus gerardi                 .. .. .. .1 .1 .1 .1 11 11 31 31 31 11 11 11 11 1. 1.
Plantago maritima              .. .. .. .. 1. 1. 1. 11 1. 11 1. 11 1. 1.
Scirpus lac.ssp.tabernaem.     .. .. .. .. .1
Juncus articulatus             .. .. .. .. .1 .1 .1 11 11 11 11 1. 11 11 ..
Agrostis gigantea              .. .. .. .. .1 .. .1 .1 11 11 11 11 31 31 21 11 12 11    12 12 .1 .1
Carex extensa                  .. .. .. .. 21 11 1. 11 ..
Festuca rubra                  .. .. .. .1 .1 .1 11 23 43 4. 34 13 11 11 11 1.    .1
Plantago major                 .. .. .. 1. 1. 11 .. 1. .1 ..
Filipendula ulmaria            .. .. .. 11 1. 11 .1 .2 .1 11 23 24 13 13 14    25 13 15 15 24 11 12 14 12 1. .. 1. .. ..
Centaurium littorale           .. .. 11 11 1. 1.
C. pulchellum                  .. .. 11 1. .1 .1 .. .1 ..
Parnassia palustris            .. .. .1 .1 11 11 11 11 11 11 11 11 11 .1
Odontites verna ssp.litor.     .. 1. 1. 1. 1.
Leontodon autumnalis           .. .1 11 11 11 11 1.
Samolus valerandi              .. .1
Calamagrostis stricta          .. .1 .1 .. .1 .1 .1 .1
Lythrum salicaria              .. 1. 11 11 11 .1 1.
Alnus glutinosa seedlings      .. .1 .2 .. 11 .. .1 .1 ..
Sagina nodosa                  .. 1. 1. 1. 1.
Carex serotina ssp.pulch.      .. 1. 1. 1.
Sagina procumbens              .. 1. .. 1. .. .1 11
Valeriana salina               .. .1 .. 11 11 1.    1.
Fraxinus excelsior seedl.      .. 1.    1. 11
Potentilla anserina            .. 1.
Galium palustre                .1 .. 11 11 1. 1.    .1
Taraxacum vulgatum coll.       1.    .1
Angelica sylvestris            11 .1 11 11 .. 31 31 41    2. 11 11 21 .1 .. .1
Calamagrostis epigejos         .1 11 11 41 51 21 21 31    31 11 1. 31 1. 1.
Hierochloe odorata             .1 .1 .1 .1
Ophioglossum vulgatum          .. .1 .1 .1
Rubus caesius                  .1 11 12 13    12 11 .1 .1
Moehringia trinervia           1.    .1 .1 .1
Rubus idaeus                   .1 .2 .3    22 23 31 33 12 13 .2 .1 .3 13 .. .. 1.
Sonchus arvensis               .1    1.
Cerastium fontanum             .1 .1    1. .. .1 .. .1
Stellaria media                1.    1.
Elymus caninus                 1.    1. 11 .1 .1 1. .. .1
E. repens                      .1    .1 .1
Silene dioica                  .1    1. .. .1 .. 1. 1. .. .. 1. ..
Mercurialis perennis           .. 1. 21 11 51 55 51 53 53 54 55 55 55
Allium ursinum                 .. 1. 1. 4. 44 53 55 55 55 54 55 55 55
Cirsium arvense                .1 .1 .2
Milium effusum                 .. 1. .. .1
Stachys sylvatica              .1 .1 .1 .. .1 1. 11 .. 11 .1 ..
Dactylis glomerata             .1
Anemone nemorosa               11 1. 11 22 24 55 54 44
Corydalis spp.                 .1 .1 11 11 11 .2 12 11
Gagea lutea                    .1 .. .1 .1 11 .1 .1 .1
Acer platanoides seedl.        .1 .1 1. ..
Paris quadrifolia              1. .1 1. .. .1 ..
Ranunculus auricomus           11 1. ..
Urtica dioica                  2. 1.
Ranunculus cassubicus          11 ..


Mosses %

Stones %

Drift material %

Litter %

submergence (% days)

Elevation (m)
```

Results

Vegetation of the lower shore

Instead of showing distinct vegetation belts, shore plant populations (in both years) generally had overlapping distributions (cf. Table 1). Lowest on the shore, *Eleocharis uniglumis* dominated, gradually passing over to a *Juncus gerardi* belt. Higher up, *Festuca rubra* dominated. In the shade of the outermost *Alnus glutinosa* bushes, *Calamagrostis epigejos, Filipendula ulmaria* and *Angelica sylvestris* were prominent. In the more sheltered transects, a *Phragmites australis* community occurred, partly in dense stands with hardly any other vascular plant species, partly with more scattered individuals mixed with *Scirpus lacustris* spp. *tabernaemontani, S. uniglumis* and *Galium palustre.*

In the most exposed transect, the lowest parts (at and slightly above mean water level) were poor in species with *Deschampsia bottnica* and also *Phragmites.* In the upper parts, a community with *Phalaris arundinacea* and *Sonchus arvensis* had developed.

Vegetation of the woodland zone

The field and shrub layer vegetation in the woodlands close to the open shore was characterized by *Rubus idaeus* and *Elymus caninus. Rubus caesius* was common both on the shore and in the forest. Higher up, *Mercurialis perennis* and *Allium ursinum* dominated. In the sites with dense spruce forest, the field layer consisted mainly of scattered individuals of *Dentaria bulbifera* and *Trientalis europaea.* The more open woodlands of the exposed transects had a field layer more species-rich and with higher species abundances of, among others, *Geranium sylvaticum, Veronica chamaedrys* and *Laserpitium latifolium.*

Environmental change

Time-scales of change. Due to the strong effect of
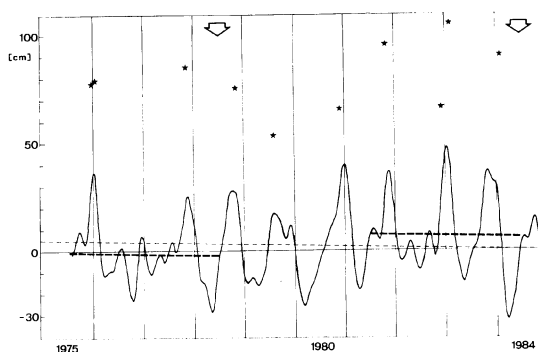
Fig. 1. (Smoothed) curve of water levels at Forsmark, 85 km NNW of the study area from 1975 to 1984. Arrows mark the two sampling dates, stars each year's absolute maximum, the thin dashed line theoretical zero level, the thick dashed lines mean over three year periods prior to each sampling, related to theoretical zero level. Data from the Swedish Meteorological Survey (SMHI).

sea levels on the shore environment, changes in habitat conditions can be observed on the various time-scales corresponding to the various time-scales of sea level change (Fig. 1):

1. The seasonal rhythm of low water during spring and summer, variable water levels in late summer and autumn, and high water levels, storm waves and ice-push in winter limit establishment of plants on new areas to early summer and require tolerance against drowning and mechanical destruction for biennial or perennial species.
2. The year-to-year variations of the physical environment, such as water levels, wind velocities, precipitation and radiation, affect the chances for establishment and persistence of coastal plants differentially in different years.
3. The effects of seashore displacement, caused by land uplift, can be observed over periods longer than 10 years and lead to the general apparent 'downward migration of plants' (Ericson, 1980).

Water level changes in the course of our study. During recent centuries, the general trend of land uplift at Skabbholmen resulted in an apparent long-term lowering of sea level of ca. 5.0 mm/yr

162
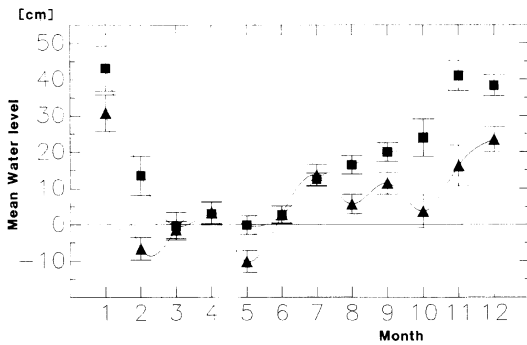
Fig. 2. Mean water levels for months in the three-year periods prior to each vegetation sampling with 95%-confidence intervals. Triangles: July 1975 to June 1978, squares: July 1981 to June 1984. Data from the Swedish Meteorological Survey (SMHI).

(calculated after Åse, (1964)). In the Baltic Sea, there are great differences in mean water level from year to year, however, and these create difficulties in observing the trend of seashore displacement during periods shorter than 10 years.

From the long-term trend, we predicted a mean water level ca. 3 cm lower for the second sampling, compared with the first. Figure 1 shows the difference in the course of water levels during the period 1975 to 1984. The graph illustrates that water levels during the second period have been higher than the predicted value; and in fact higher than during the first period. Comparison of monthly means above theoretical zero level (= mean water level corrected by land uplift, Fig. 2) showed that this difference was levelled out in March, April, June and July; the greatest differences occurred in winter.

Changes in shore geomorphology of the island. We noted the following indications of geomorphological change in the four transects between 1978 and 1984: in transect 3, the stones that 1978 had been bare in the lowest part of the transect (sites 1 to 10) were, in 1984, covered by coarse sand and some drift material. The moss belt in sites 12 to 14 disappeared, another one was now found in sites 16 to 19. The accumulation of drift material around site 22 had been moved to a new belt higher up (sites 24 to 26), leaving more bare stones than be-

fore. Much of the litter in this zone was covered by the drift material (Table 1). In transect 4, erosion (most probably during a heavy storm with high water level in January 1983) had affected the middle part, near to the edge of – and inside the woodland: the topsoil had been removed almost entirely, many tree roots now lay bare, and the percentage of stones at the soil surface had increased. In the forest, a new drift accumulation (ca. 20 cm high) was created. In total, there were indications of heavy wave action higher on the shore in 1984 than in 1978. This may be due to both the generally higher water levels in winter and the extreme storm in January 1983.

Field layer response to environmental change

If there is a significant trend of short-term temporal change in species composition of all sites, then the regression coefficient (called the 'canonical coefficient') of time in years on at least one of the ordination axes should be significant. Standardized canonical coefficients and approximate significance levels (as given by CANOCO) are listed in Table 2 for both shore and forest data. The significances show that the first axis is related to both variables in both subsets. The coefficients indicate a strong dominance of elevation on this axis. The second axis is in both cases dominated by year.

Table 2. Canonical coefficients (= standardized regression coefficients of the ordination scores on the external variables) in DCCA of the transect data with year and elevation as external variables.

	Canonical coefficients	
	Shore	Forest
First axis		
Elevation	1.00*	0.68*
Year	0.12*	– 0.11*
Second axis		
Elevation	0.04	– 0.07*
Year	– 0.34*	– 0.22*

* significant ($P<0.05$).

In the ordination diagram (Fig. 3), the main direction of change for each of the external variables in the shore subset is illustrated by the thick arrows in the center of the diagram (called a biplot). The length of these arrows corresponds to the relative importance of these two variables in determining the floristic variation. For the sake of simplicity, only the species and samples from transect 3 are shown in the diagram.

The thin arrows show the actual change of each sample. Relative to the first axis, the direction of these arrows gives an estimate of the relation be-

Fig. 4. The site trajectories of transect 3 (shore) with sample numbers, based on the same DCCA ordination of the total shore data as in Fig. 3.

tween six-year change in this site and the main pathway of long-term vegetation change, as far as it can be predicted from elevation. An arrow perpendicular to the first axis indicates that no change corresponding to the long-term trend has occurred at this site. Figure 4 shows the same samples as in Fig. 3, but with numbers to allow their identification with Table 1.

The (standardized) coefficients of the multiple regressions of external variables on ordination scores given in Table 2 can be used to study the relative contribution of the external variables to the dispersion of the sample scores along each of the ordination axes. For this, we calculated the unstandardized coefficients by dividing the standardized coefficients by the standard deviations of the corresponding variables (Table 3).

The resulting equations show the linear combinations of the two variables in their relation to the sample scores:

Fig. 3. Diagram of the DCCA ordination for the shore data set (Axes 1 and 2). Triangles mark species occurring in transect 3, squares and thin arrows mark site trajectories of six-year change in transect 3. Thick arrows denote main direction of change and relative importance of the two external variables.

Table 3. Means and standard deviations for the external variables

	Shore		Forest	
	mean	SD	mean	SD
Elevation (cm)	22.3	18.5	173.0	91.0
Year	3.0	3.0	3.0	3.0

Shore: $\quad x = 0.054\,z_1 + 0.040\,z_2 \qquad (1)$

Forest: $\quad x = 0.008\,z_1 - 0.035\,z_2 \qquad (2)$

where x is the first axis sample score, z_1 is elevation in cm and z_2 is year in years. x expresses species turnover along the first ordination axis in terms of standard deviations (SD) of the mean species response. The average rate of change along the first axis each site went through from 1978 to 1984 can be estimated by substituting z_1 and z_2 by the elevation and year value for any site in these two years and calculating the difference between the two resulting x-values. As elevation in our model was given the same values for both sampling years, the result will only be dependent on the difference in z_2 values. Substituting 0 and 6 for z_2 will therefore give an average rate of change along the first axis of 0.24 SD (shore) and 0.21 SD (forest).

From the same equations (1) and (2), the ratio of the regression coefficients for year and elevation gives an expression for the average rate of vertical change of species distributions (i.e., the elevation difference that gives an equivalent floristic change as a time lag of one year). The coefficient of variation for this ratio (approximated as the square root of the sum of squares of the two single coefficients of variation) yields 95%-confidence intervals for the results (Ter Braak, pers. comm.):

Shore: $\quad$ 0.74 ± 0.37 cm/yr

Forest: $\quad$ −4.71 ± 3.44 cm/yr

The short-term rate of vertical change of species distributions along the first axis for the shore vegetation was thus (a) different from zero and (b) included the present rate of seashore displacement (0.5 cm/yr). Within the margins of error, we conclude that the shore vegetation had changed significantly during six years in the direction of the long-term floristic trend and that the rate of change in this direction was consistent with the known rate of land uplift.* For the forest vegetation, we attempted no further interpretation of this result, because the negative value and its large confidence interval (resulting from the small regression coefficient with

elevation) indicate a low reliability.

The first-axis component of the 'year vector' in Figs. 3 and 4 relates to the change discussed in the previous section. The second, larger component, however, illustrates the relatively large importance of temporal change in floristic composition that is independent from the first component, i.e., changes deviating from the long-term trend. Species in the upper half of the diagram were generally reduced in abundance; those in the lower half gained.

Comparison of direct and indirect gradient analysis

We compared the results of DCCA with those from a standard DCA ordination, unconstrained by external variables. DCCA yielded eigenvalues for the first axes that were slightly lower than DCA for the first axis (Table 4), but considerably lower for the second. A comparison of gradient lengths shows more directly that the DCCA axes (especially the second axis) are related to a smaller amount of (estimated) species turnover than those of DCA. Note that the third and fourth DCCA axes were predefined as residual axes, i.e., they account for components of the floristic variation not accounted for by the two external variables. Table 4 also shows the correlation coefficients between year and elevation and the first two axes of both analyses.

The main conclusion from this comparison is as follows: the first axes in both DCA and DCCA (with similar eigenvalues) were highly correlated with elevation and represent the dominating floristic trend in the data sets. The second axes in

Table 4. Comparison between DCA and DCCA: eigenvalues, gradient lengths and species-environment correlations.

	Shore		Forest	
	DCA	DCCA	DCA	DCCA
Eigenvalues				
1st Axis	0.654	0.564	0.605	0.410
2nd Axis	0.466	0.104	0.313	0.057
3rd Axis	0.281	0.524*	0.215	0.498*
4th Axis	0.181	0.287*	0.154	0.252*
Gradient lengths (SD)				
1st Axis	5.5	4.4	5.9	3.3
2nd Axis	5.7	0.9	2.8	0.7
3rd Axis	3.8	5.0*	3.1	4.4*
4th Axis	3.2	3.5*	2.4	3.4*
Correlation coefficients with 1st Axis				
Elevation	0.837	0.939	0.766	0.852
Year	0.103	0.068	−0.060	−0.112
With 2nd Axis				
Elevation	0.316	0.142	0.134	−0.183
Year	0.042	−0.737	−0.144	−0.563

* residual axes.

DCCA, however, have smaller eigenvalues, very short gradient lengths and high correlations with year, while the second DCA-axes have higher eigenvalues, longer gradients and no such clear correlations. This, with respect to the relatively high eigenvalues of the (residual) third and fourth axes in DCCA indicates that other environmental variables (e.g., substrate or exposure) result in greater variation in species composition than does the overall change during six years.

Discussion

Rate and direction of vegetation change

In an environment characterized by a high degree of disturbance, we can not expect a clear shift of all species response intervals along the gradient of long term change (elevation) during six years. In the shore subset, the main trend of floristic change could be separated into two parts: (1) a small, but significant trend based on a migration of species populations downward with a vertical rate of ca. 0.7 cm/yr and (2) a greater change that was not related to elevation, and so must reflect a short-term fluctuation. The relative contribution of each of these two aspects of species composition change can also be studied for individual samples. In Figs. 3 and 4, samples changing in agreement with the long-term trend of succession have arrows from left to right, while samples with deviating change move from top to bottom.

The direction of change in sites 3 to 9 was more clearly related to elevation than in the sites higher on the shore. Table 1 and Fig. 3 indicate that this is mostly due to the colonization of a number of graminoids, such as *Deschampsia bottnica, Juncus gerardi, J. articulatus, Agrostis gigantea* and *Festuca rubra.* Some of these species also show corresponding decrease in the higher parts of their distribution, e.g. *Juncus gerardi* and also *Eleocharis uniglumis.* We observe a distinct shift in elevational distribution of most species. Although inundation was more frequent during three years prior to the second sampling, the increase in cover of sediments (reflected by the apparent decrease of stone cover) allowed a higher species richness. Note the deviating behaviour of site 6 that is explained by decrease in cover only (*Plantago maritima* is the only disappearing species).

Sites 10 to 20, by contrast, show very little change along axis 1, their time trajectories gradually turn to positions almost perpendicular to the arrow of change in elevation. On the other hand, the trajectory length is greater here than lower down the shore. A few species act as colonizers again, but they do not come from belts higher up (e.g., *Hierochloe odorata, Ophioglossum vulgatum,* the apparent colonization of *Alnus glutinosa* somewhat misleads here, because the shrub and tree layer had not been sampled). Especially site 12 and also its neighbours showed a loss in species number (also moss cover decreased strongly), but these species were not colonizing elsewhere. In general, the change of species composition in this belt had little to do with the long-term trend of migrating species populations. The main cause for this appeared to be the erosion due to the storm in January 1983. This picture may also be somewhat distorted for

those species that had a distribution reaching into the woodland that had been excluded here.

The results from the forest suggest that short term vegetation dynamics (though in the same order of magnitude as on the shore) in the forest were determined by other factors than seashore displacement. An example is given by the change of some species abundances (e.g., the increase of the open shore species *Filipendula ulmaria* and the corresponding decrease of the woodland species *Allium ursinum*) on the organic drift material accumulation (sites 24–26) that was observed in the forest part of transect 3 (Table 1). Due to the great variability in forest types, the probable existence of gap dynamics and the shortage of other external data, we do not discuss the possible pathways of change in the forest vegetation any further here.

Comparison between direct and indirect gradient analysis

The usual interpretation of DCA ordination results (indirect gradient analysis) is built on the idea that the spread of sites along axes reflects successive species replacement along a hypothetical environmental gradient. CCA also fits a unimodal response model, but with respect to combinations of supplied external variables. Due to the extra constraint in CCA, any difference between the floristic gradients related to the external variables and those calculated by DCA will cause lower eigenvalues in CCA than in DCA (Table 4). The ordination axes, although corresponding to a smaller fraction of the total floristic variance, are useful to explore the effect of external variables that are not necessarily dominating the vegetational gradients. A comparison to unconstrained DCA is needed, though, to estimate the importance of floristic trends others than the ones expressed by the constrained axes.

The environmental variable elevation was the basis of our sampling plan and turned out to have the most profound influence on the distribution of species in both DCA and CCA, as it determined the first axes almost completely (Table 4). This is not surprising on a seashore where many species response intervals can be included in their entirety by

transects across the shore. Field observations and inspection of our tables gave us little reason to doubt the assumption that most species, in our case, actually do respond with a roughly bell-shaped, unimodal curve to this gradient, as is required for success in the use of DCA as an ordination method (Hill & Gauch, 1980).

Correlations of the short-term time gradient with ordination axes were much more clear in CCA than in DCA: they were both more highly correlated with one of the axes, and the same axis was less related to the other supplied variable. This possible separation of effects gives an advantage of CCA in interpretation for two reasons. First, while DCA aims for 'maximum floristic variance expressed in the ordination', CCA aims for 'maximum correlation between external variables and that part of the floristic variance they can explain'. Hence in CCA a smaller amount of information from the input matrix is expressed in the ordination, but the result is, in turn, better related to the supplied data about the environment. For an ecological study, we thus end up with a less noisy image of the species' responses. Second, the introduction of more than one gradient can create difficulties in the interpretation of pure species/sample ordinations. These difficulties may be especially troublesome in studies of three-dimensional (species × sites × times) matrices (Williams & Stephenson, 1973). Temporal gradients can then be confused with spatial gradients (as was pointed out by Swaine & Greig-Smith, (1980)). In DCA, the impact of external gradients on the ordination can be studied either graphically or with help of (multiple) regression analyses between axes and external variables. CCA allows a more direct approach to the problem, since it provides (1) clear information on the relations between the external gradients and the ordination scores, (2) ordination diagrams (biplots) that clearly illustrate the relative importance of the supplied external variables and their main direction of change (Fig. 3), and (3) a better separation between the influences of the external variables (cf. Purata, 1986).

References

Åse, L.-E., 1964. Nivåförändringen i östra Svealand och Åland beräknad med utgångspunkt från äldre kartmaterial. Ymer 84: 121–182.

Åse, L.-E., 1980. Shore displacement at Stockholm during the last 1000 years. Geogr. Ann. 62A: 83–91.

Austin, M. P., 1977. Use of ordination and other multivariate methods to study succession. Vegetatio 35: 165–175.

Austin, M. P. & Belbin, L., 1981. An analysis of succession along an environmental gradient using data from a lawn. Vegetatio 46: 19–30.

Brunberg-Schwanck, B. & Bärlund, U., 1948. Vegetation och landhöjning. In: Skärgårdsboken, pp 257–291. Nordenskiöld-Samf. i Finland, Helsinki.

Cramer, W., 1980. Die Wechselwirkungen zwischen Strandverschiebung und Strandvegetation. Acta Phytogeogr. Suec. 68: 43–50.

Cramer, W., 1985. The effect of sea shore displacement on population age structure of coastal *Alnus glutinosa* (L.) Gaertn. Holarct. Ecol. 8: 265–272.

Cramer, W., 1986. Altersstruktur von *Alnus glutinosa* (L.) Gaertn. unter dem Einfluss der Strandverschiebung an der schwedischen Ostseeküste. Verh. Ges. f. Ökol. 14: 519–522.

Du Rietz, G. E., 1925. Die Hauptzüge der Vegetation des äusseren Schärenhofes von Stockholm. Sv. Bot. Tidskr. 19: 317–369.

Ericson, L., 1980. The downward migration of plants on a rising Bothnian sea-shore. Acta Phytogeogr. Suec. 68: 61–72.

Ericson, L., 1981. Aspects of the shore vegetation of the Gulf of Bothnia. Wahlenbergia 7: 45–60.

Ericson, L. & Wallentinus, H.-G., 1979. Sea-shore vegetation around the Gulf of Bothnia. Wahlenbergia 5.

Gunnlaugsdottir, E., 1985. Composition and dynamical status of heathland communities in Iceland in relation to recovery measures. Acta Phytogeogr. Suec. 75.

Hill, M. O., 1979a. DECORANA, a FORTRAN program for detrended correspondence analysis and reciprocal averaging. Ecology and Systematics, Cornell University. Ithaca, New York.

Hill, M. O., 1979b. TWINSPAN, a FORTRAN program for analysing multivariate data in an ordered two-way table by classification of individuals and attributes. Ecology and Systematics, Cornell University. Ithaca, New York.

Hill, M. O. & Gauch, H., 1980. Detrended correspondence analysis, an improved ordination technique. Vegetatio 42: 47–58.

Londo, G., 1974. Successive mapping of dune slack vegetation. Vegetatio 29: 51–61.

Luther, H., 1961. Veränderungen in der Gefässpflanzenflora der Meeresfelsen von Tvärminne. Acta Bot. Fenn. 62.

Persson, S., 1980. Succession in a south Swedish deciduous wood: a numerical approach. Vegetatio 43: 103–122.

Persson, S., 1984. Vegetation development after the exclusion of grazing cattle in a meadow area in the south of Sweden. Vegetatio 55: 65–92.

Purata, S. E., 1986. Studies on secondary succession in Mexican tropical rain forest. Thesis Uppsala.

Schwanck, B., 1974. Vegetationens förändring på de framväxande skären. In: Skärgård i omvandling, pp 79–85. Nordenskiöld-Samf. i Finland, Helsinki.

Sjörs, H., 1963. Amphi-atlantic zonation, nemoral to arctic. In: A. Löve and D. Löve (eds), North Atlantic biota and their history, pp 109–125. Pergamon Press, Oxford.

Swaine, M. D. & Greig-Smith, P., 1980. An application of principal components analysis to vegetation change in permanent plots. J. Ecol. 68: 33–41.

Ter Braak, C. J. F., 1985. CANOCO – A FORTRAN program for canonical correspondence analysis and detrended correspondence analysis. IWIS-TNO, Wageningen, The Netherlands.

Ter Braak, C. J. F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167–1179.

Ter Braak, C. J. F., 1987. The analysis of vegetation-environment relationships by canonical correspondence analysis. Vegetatio 69: 69–77.

Valovirta, E. J., 1937. Untersuchungen über die säkuläre Landhebung als pflanzengeographischer Faktor. Acta Bot. Fenn. 20.

Van der Maarel, E., 1969. On the use of ordination models in phytosociology. Vegetatio 19: 21–46.

Van der Maarel, E., 1975. Small-scale changes in a dune grassland complex. In: W. Schmidt (ed.), Sukzessionsforschung. pp 123–132. Cramer, Vaduz.

Van der Maarel, E., 1980. Vegetation development in a former orchard under different treatments: a preliminary report. Vegetatio 43: 95–102.

Van Hecke, P., Impens, I. & Behaeghe, T. J., 1981. Temporal variation of species composition and species diversity in permanent grassland plots with different fertilizer treatments. Vegetatio 47: 221–232.

Van Noordwijk-Puijk, K., Beeftink, W. G. & Hogeweg, P., 1979. Vegetation development on salt-marsh flats after disappearance of the tidal factor. Vegetatio 39: 1–13.

Vartiainen, T., 1980. Succession of island vegetation in the land uplift area of the northernmost Gulf of Bothnia. Acta Bot. Fenn. 115.

Verwijst, T. & Cramer, W., 1986. Age structure of woody species in primary succession on a rising Bothnian sea-shore. In: J. Fanta (ed.), Forest dynamics research in western and central Europe, pp 145–163. Pudoc, Wageningen.

White, P., 1979. Pattern, process and natural disturbance in vegetation. Bot. Rev. 45: 229–299.

Williams, W. T., Lance, G. N., Webb, L. J., Tracey, J. G. & Dale, M. B., 1969. Studies in the numerical analysis of complex rain-forest communities. III. The analysis of successional data. J. Ecol. 57: 515–535.

Williams, W. T. & Stephenson, W., 1973. The analysis of three-dimensional data (sites × species × times) in marine ecology. J. Exp. Mar. Biol. Ecol. 11: 207–227.

Zhang, L., 1984. Vegetation ecology and population biology of *Fritillaria meleagris* L. at the Kungsängen Nature Reserve, eastern Sweden. Acta Phytogeogr. Suec. 73.

Regeneration dynamics of beech forests in Japan*

Tohru Nakashizuka**
Vegetation Laboratory, Forestry and Forest Products Research Institute, P.O. Box 16, Tsukuba Norin Kenkyu Danchi-nai, Ibaraki, 305 Japan

Abstract

A conceptual model of the regeneration of beech (*Fagus crenata*) forests in Japan is presented; stand structure, tree population dynamics and characteristics of gap formation are summarized. Beech forest can be recognized as a regeneration complex, and three phases: gap, building and mature, are distinguished through the structure of regenerating populations. The mean return interval is 100–200 yr. Disturbances by typhoons and dwarf bamboos (*Sasa*) on the forest floor may greatly affect the synchrony of regeneration and, hence, the stability of beech forests in Japan.

Introduction

One of the recent trends in the study of forest vegetation is the emphasis on gap formation and disturbance. Canopy gap formation, which is associated with the death of canopy trees, until recently was one of the least investigated aspects of forest regeneration. Climax forests were regarded as in steady state according to the classical view (Clements, 1916; Tansley, 1920). Recent papers, however, have documented the structural heterogeneity and time-variability of natural vegetation in the world, and emphasized the role of natural disturbance (Pickett & Thompson, 1978; White, 1979).

Analysis of every phase of the life history of constituent species is needed in order to understand the regeneration dynamics of forests. Beech (*Fagus crenata*) forest, which is common natural vegetation in the cool-temperate zone of Japan, is one of the most intensively investigated forest types in

terms of its regeneration dynamics. Stand structure and tree age distribution (Nakashizuka & Numata, 1982a, b; Honma & Kimura, 1982; Hara, 1983), seedling survivorship (Maeda & Miyakawa, 1971), sapling population dynamics (Nakashizuka, 1983, 1984b) and gap formation (Nakashizuka, 1984a) have been reported. This paper provides a conceptual model of beech forest dynamics, drawing on these various sources of data and adding some original ones.

A unique characteristic of beech forests in Japan is the overwhelming dominance of dwarf bamboo (*Sasa* spp.) on the forest floor (Shidei, 1974). The inhibiting effect of *Sasa* on tree regeneration has been well reported from a silvicultural viewpoint (e.g. Maeda & Miyakawa, 1971). The monocarpic life history of the bamboo is also significant in forest dynamics. The effects of *Sasa*, and its interaction with disturbance, are discussed here.

Stand structure

Mosaic forest structure as described by Watt (1947) can also be recognized in Japanese beech forests (Nakashizuka & Numata, 1982a, b).

Nakashizuka & Numata showed that large trees in a forest are distributed randomly or even uniformly, while the trees of small and intermediate size are distributed contagiously (Fig. 1). Annual ring analysis showed that trees in each contagious patch emerged within a period of 20–60 yr. Dense populations of young trees are established in gaps, especially when the undergrowth is sparse (Nakashizuka, 1984b). Such a structure was also reported in several other beech forests in Japan (Honma & Kimura, 1982; Hara, 1983). Dense populations in gaps apparently develop into the contagious structure of intermediate sized trees, and eventually to randomly distributed large trees. Thus we could see the mosaic structure, the so-called 'regeneration complex' (Watt, 1947), which is a common characteristic of many kinds of forests in the world (Whitmore, 1982).

The age distribution of trees in several beech forests in Japan suggested some synchronization of forest regeneration (Nakashizuka & Numata, 1982a, b; Honma & Kimura, 1982). Two examples were shown in Figs 2 and 3, in which the data by Nakashizuka & Numata (1982a, b) and original data (obtained by the same methods) are synthesized. Trees smaller than 10 cm in DBH are neglected in the figures. (As these small trees occupied age classes younger than 100 years at most (mostly younger than 60 years), the distributions of trees older than 100 years indicate effectively the complete ones (Nakashizuka & Numata, 1982a).) Age distributions of trees in three stands (0.64–1.0 ha area) within 1–2 km of each other at Kayanodaira, central Japan, were all similar, having a mode in the age classes about 140–200 years (Fig. 2). Also in four stands (0.24–0.36 ha) within
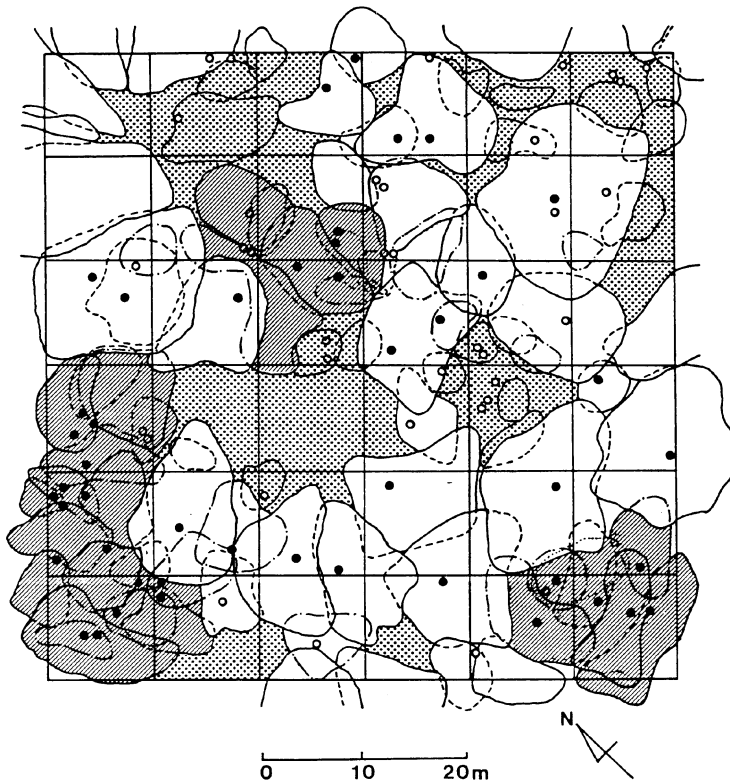
Fig. 1. Crown projection diagram of a forest at Mt. Moriyoshi. Solid and open circles indicate the location of stem bases of beech and other species, respectively, over 10 cm in DBH. Cross-hatching shows the contagious structure of middle sized trees and stippling indicates gaps. Revised from Nakashizuka (1982b).
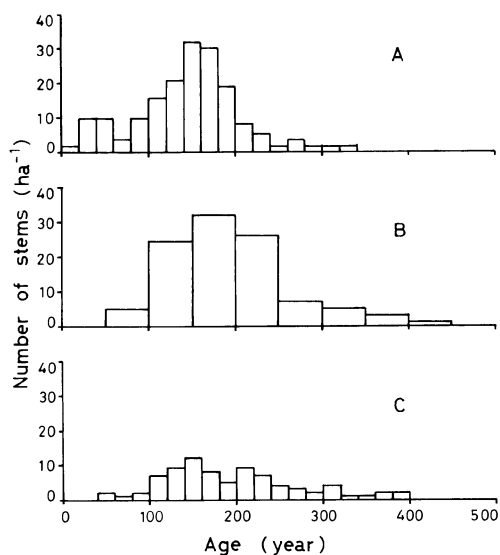
Fig. 2. Age distributions of trees over 10 cm in DBH in the quadrats at Kayano-daira. A: Data from Nakashizuka & Numata (1982a) in 0.64 ha area. B: Data from Nagano District Forestry Office (1975) in 1.0 ha area. C: Data by Laboratory of Ecology, Chiba University in 1.0 ha area.

Fig. 3. Age distributions of trees over 10 cm in DBH in the quadrata at Mt. Moriyoshi. Shaded and open parts show the stems of beech and other species, respectively. Quadrats A (0.28 ha) and B (0.36 ha) are based on the data of Nakashizuka & Numata (1982b). C (0.24 ha) and D (0.30 ha) are based on original data of the author.

an area of several sq. km at Mt. Moriyoshi, northeastern Japan, tree age distributions had two modes, i.e. 120–160 yr, and 200–220 yr (Fig. 3).

To summarize, a beech forest may have a homogeneous age structure in an area of several sq. km, suggesting that the regeneration in gaps tends to synchronize. The causes of this phenomenon will be discussed later.

Seedlings under a closed canopy

Fagus crenata, like other climax hardwood species, is classified as a species requiring banks of persistent seedlings (see Grime, 1979). In the forest at Mt. Moriyoshi, where dwarf bamboo and other understory species were removed by grazing, beech seedlings can live as long as 15 years under closed canopies (Table 1, see also Nakashizuka & Numata, 1982b). As long as recruitment continues, a considerable number (10 to 100 m^{-2}) of seedlings appear on the forest floor when the undergrowth is sparse. But in the forest at Ohdaigahara, where *Sasa borealis* dominates the forest floor, seedling

density was very low (0.19 m^{-2}), about a half of which was the current year's seedlings (data obtained by the author from an average of 80 1 m^2 quadrats, following the enumeration method of Nakashizuka & Numata, 1982b). Maeda & Miyakawa (1971) reported that more than 90% of germinated seedlings died within a year under a cover of *Sasa kurilensis*. The cover of *Sasa*, therefore, prevents the establishment of seedling banks of beech. A similar situation was reported in Chilean *Nothofagus* forest where regeneration was prevented by another dwarf bamboo, *Chusquea* (Veblen et al., 1980).

Under a closed canopy, however, the height of beech seedlings is very small even when the undergrowth is sparse; only 20 cm height was reached by ten-year-old seedlings (Table 1). Seedlings cannot grow rapidly unless they are released in gaps (Nakashizuka, 1983).

Table 1. Age distribution and mean height of each age class of beech seedlings under closed canopies at Mt. Moriyoshi.

Age (yr)	Quadrat-1		Quadrat-2	
	Density (m^{-2})	Mean height (cm)	Density (m^{-2})	Mean height (cm)
0				
1	10.55	8.5	2	7.8
2				
3	0.04	13.0	82	13.2
4	0.80	10.5		
5			10	16.5
6	0.32	14.3	2	13.1
7	0.04	15.0		
8	0.04	18.0	1	21.5
9			1	17.2
10			2	21.6
11				
12				
13			2	33.2
14				
15			1	48.9
Total	11.8		104	

Population dynamics in gaps

Population dynamics through the regeneration process discussed here is based on the data obtained at Mt. Moriyoshi (Nakashizuka, 1984b). Cattle and horses had been put into the beech forest from about 1920 to about 1980. As they grazed dwarf bamboo and other understory plants, which prevent the regeneration of trees, we could observe the successful regeneration of beech. The forest at Mt. Moriyoshi was thus appropriate to analyse the population ecology of beech when released from suppression by the understory.

Detailed studies on the age structure of sapling populations in gaps showed that the regenerating populations in gaps were composed of saplings germinated from 10−15 yr before until 10−15 yr after the gap formation (Nakashizuka, 1983, 1984a, b). This fact matches well the observation that seedlings under closed canopies can survive up to about 15 yr. In 10−15 yr after the gap formation, the dense foliage of the regenerated population apparently prevents subsequent recruitments under it. Thus, the population established in such a way could be treated as a single cohort.

Nakashizuka (1984b) recognized three phases in the course of regeneration based on age, density, biomass and thinning relations of the populations. The first phase is characterized by the recruitment of seedlings germinated after the gap formation in the initial 10−15 yr. Adding to the advance-regenerated seedlings, recruitment continues with low mortality until the population attains its full density state (Fig. 4). The second phase is characterized by the fast growth of saplings and a rapid decrease in density. After the population attains its full density, self-thinning begins according to the −3/2 power law of Yoda *et al.* (1963). In 70−80 yr after the gap formation, the biomass approaches an asymptote, and the density decrease becomes slow: this is the third phase.

The first, second and third phases correspond to gap, building and mature phases, respectively, as described by Watt (1947) and Whitmore (1982). These phases can be defined not only physiognomically, but also quantitatively.

Gap formation

It is first necessary to define a 'gap'. Nakashizuka (1984a) defined a gap as an area larger than 10 m^2 without crowns over 15 m in height. (The recognition and finding of smaller or older gaps may be subjective and vague, and when a sapling population reaches about 15 m in height, the canopy of the regenerating population is almost continuous to the forest canopy.) I follow my earlier definition here.

In general, beech forests in Japan regenerate in gaps, made by one or several canopy trees, up to 500 m^2 in area. In the forest at Mt. Moriyoshi the maximum size of a gap was about 470 m^2, and 20.5% of the total forested area was gaps (Nakashizuka, 1984a). The same statistics at Ohdaigahara (obtained by the author following a similar method to Nakashizuka, 1984a) were about 200 m^2 and 18.6%, respectively (Fig. 5). Many mature forests in Japan seem to have a gap area 10−20% of the forested area. By contrast, in tropical forests, gap area usually occupies less than 10% of the forested area (Hartshorn, 1978; Hubbell, 1979; Poore, 1968; Whitmore, 1975). In temperate hardwood forest, however, the proportion of gap area has been reported to be 3.2−24.2% (Runkle,
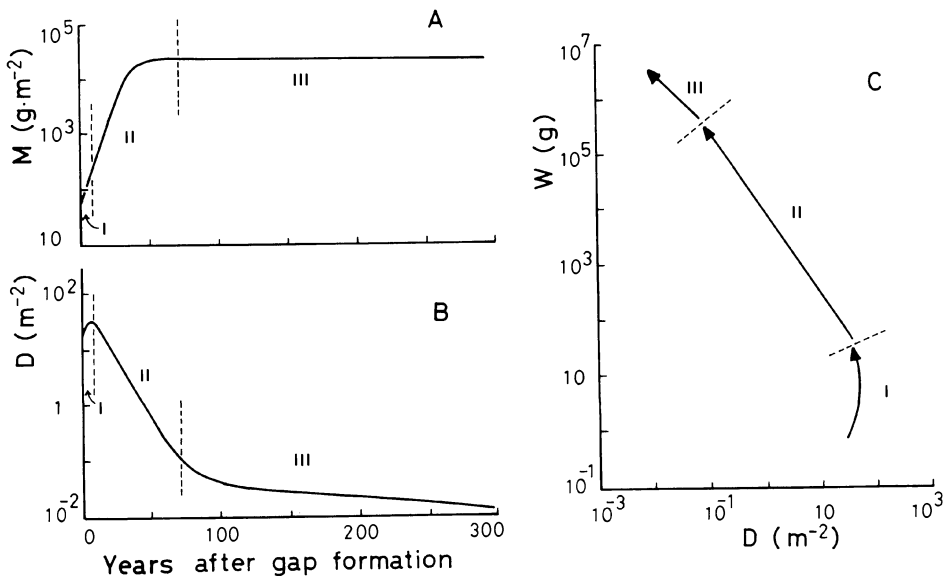
Fig. 4. Structural change of regenerated populations in gaps. A: Growth of aboveground biomass (M) after the gap formation. B: Change of stem density (D) after the gap formation. C: Relation between mean aboveground weight (W) and stem density (D). I, II and III are the regenerational phases distinguished (after Nakashizuka, 1984b).

1982), 17.6% (Vankat *et al.*, 1975) or 9% (Williamson, 1975), values that are comparable with or slightly smaller than those in beech forests in Japan.

Typhoons seem to be a major factor making gaps, because about 60–70% of gap-making trees had snapped stems or were uprooted in the forests at Mt. Moriyoshi (Nakashizuka, 1984b) and at Ohdaigahara. However, 10% or more of the gap makers were found as standing dead, suggesting death from senescence. Falinski (1978) reported a considerable proportion of stem-broken trees had died before the breaking of their stems. Therefore, the proportion of trees killed by typhoons may in fact be smaller than 60–70%. This is a question to be answered in an ongoing census of gaps as they form.

The rate of gap formation and the turnover time (mean return interval) of canopy in the forest at Mt. Moriyoshi were estimated to be 41–82 m² ha⁻¹ yr⁻¹ and 97–194 yr, respectively, by making simple models of the gap filling process; the model that gaps are filled only by regenerating saplings gives the low estimate of gap formation rate and the high estimate of mean return time, while the model that gaps are filled by encroachment of the surrounding canopy plus regenerating saplings gives the high estimate of gap formation rate and the low estimate of mean return time (Nakashizuka, 1984a). In the forest at Ohdaigahara, those values were estimated to be 31–62 m² ha⁻¹ yr⁻¹ and 132–263 yr, respectively, using similar assumptions. These values are rather larger than those obtained in tropical forests (Hartshorn, 1978; Brokaw, 1982), but comparable to those in temperate forests (Naka, 1982; Runkle, 1982; Romme & Martin, 1982).

Nakashizuka (1984a) showed that the turnover time of the forest at Mt. Moriyoshi was also comparable to the values estimated from annual increment and aboveground biomass (122 yr), annual death and aboveground biomass (192 yr) and mean age of canopy trees (207 yr). The dynamic equilibrium of a climax forest, if it exists, must be maintained by the partial destruction (gap formation) and repair (regeneration in gaps) of the forest. The biomass of forest in such a state must be smaller than the asymptote of biomass growth in
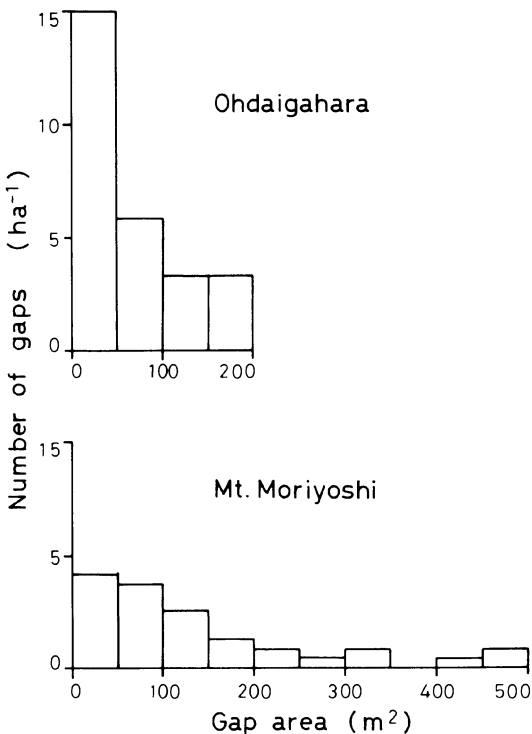
Fig. 5. Size distribution of gaps in the forests at Ohdaigahara (studied by the author in a 1.2 ha area) and Mt. Moriyoshi (2.4 ha, after Nakashizuka, 1984a).

individual patches within the mosaic (Fig. 4), as predicted by Peet (1981).

The interaction of disturbance and dwarf bamboo

The age distribution of trees in the forests at Kayano-daira and Mt. Moriyoshi had modes suggesting some synchronized regenerations in gaps (Figs 2, 3). There are several possible reasons for this phenomenon. A single large catastrophic disturbance in the past might conceivably be the cause of this. Certainly, a single very large typhoon has destroyed subalpine coniferous forests in as much as a several hundred ha area in Japan (Tamate *et al.*, 1977). However, no record like this exists for Japanese temperate hardwood forests, where typhoons create gaps by felling single or small groups of canopy trees. Watanabe *et al.* (1985) showed that

even a strong typhoon, which caused major damage in central Japan, broke only 39 trees in a 6.25 ha area of beech forest. Therefore, the synchronous regeneration seems not to be caused by a single catastrophic typhoon.

Another factor must be considered, however, when we discuss the regeneration of beech forests in Japan: the dominance and synchronous death of dwarf bamboos (*Sasa* spp.) on the forest floor. Their dominance prevents beech from establishing a seedling bank. In a forest with a dense cover of *Sasa*, few seedlings will regenerate even in new gaps, because the cover of *Sasa* will become still denser. It would take a longer time to fill the gap with *Sasa* than without *Sasa*. Therefore, forest with *Sasa* would have larger proportion of gap area forest without *Sasa*, given the same rate of gap formation. In other words, the canopy of the forest with *Sasa* would become an open forest. This mechanism may explain why the proportion of gap area in beech forests in Japan is higher than the norm for hardwood forests.

Sasa spp. flower once in several decades and die over a large area, as much as several sq. km (Numata, 1970). If *Sasa* in a beech forest dies, regeneration will begin almost at the same time both in gaps that have not yet been filled by beech, and in gaps that are made before the recovery of *Sasa*. This seems to be the cause of the wide mode in age distribution of trees in beech forests.

The cover of *Sasa* thus amplifies the disturbance-mediated fluctuation of the community parameters of the forest (biomass or age structure of trees). As yet little is known in detail about the life history of *Sasa* in relation to the disturbance regime in the beech forests of Japan; long-term study over large areas is needed.

References

Brokaw, N. V. L., 1982. Treefalls: Frequency, timing, and consequences. In: E. G. Leigh Jr, A. S. Rand & D. M. Windsor (eds), The ecology of a tropical forest: Seasonal rhythms and long-term changes. pp. 101–108. Smithsonian Inst. Press, Washington, D.C.

Falinski, J. B., 1978. Uprooted trees, their distribution and influence in the primeval forest biotope. Vegetatio 38: 175–183.

Hara, M., 1983. A study of the regeneration process of a Japanese beech forest. Ecol. Rev. Sendai 20: 115–129.

Hartshorn, G. S., 1978. Tree falls and tropical forest dynamics. In: P. B. Tomlinson & M. H. Zimmerman (eds), Tropical trees

as living systems. pp. 617–638. Cambridge Univ. Press.

Honma, S. & Kimura, M., 1982. Analysis of structure and regeneration process of beech (Fagus crenata Bl.) forest. In: Effects of forest on regulating environment, No. 2: 7–14 (in Japanese with an English summary).

Hubbell, S. P., 1979. Tree dispersion, abundance and diversity in a tropical dry forest. Science 203: 1299–1309.

Maeda, T. & Miyakawa, K., 1971. New sylvicultural techniques for natural regeneration of beech forests. In: New sylvicultural techniques for natural regeneration. Sobun, Tokyo, pp. 179–252 (in Japanese).

Naka, K., 1982. Community dynamics of evergreen broadleaf forests in southwestern Japan. I. Wind damaged trees and canopy gaps in an evergreen oak forest. Bot. Mag. Tokyo 95: 385–399.

Nakashizuka, T., 1983. Regeneration process of climax beech (Fagus crenata Blume) forests III. Structure and development process of sapling populations in different aged gaps. Jap. J. Ecol. 33: 409–418.

Nakashizuka, T., 1984a. Regeneration process of climax beech (Fagus crenata Blume) forests IV. Gap formation. Jap. J. Ecol. 34: 75–85.

Nakashizuka, T., 1984b. Regeneration process of climax beech (Fagus crenata Blume) forests V. Population dynamics of beech in a regeneration process. Jap. J. Ecol. 34: 411–419.

Nakashizuka, T. & Numata, M., 1982a. Regeneration process of climax beech forests I. Structure of a beech forest with the undergrowth of Sasa. Jap. J. Ecol. 32: 57–67.

Nakashizuka, T. & Numata, M., 1982b. Regeneration process of climax beech forests II. Structure of a forest under the influences of grazing. Jap. J. Ecol. 32: 473–482.

Numata, M., 1970. Conservation implications of bamboo flowering and death in Japan. Biol. Conserv. 2: 227–229.

Peet, R. K., 1981. Change in biomass and production during secondary forest succession. In: D. C. West, H. H. Shugart & B. D. Botkin (eds), Forest succession: Concepts and application. pp. 324–338. Springer-Verlag, Berlin.

Pickett, S. T. A. & Thompson, J. N., 1978. Patch dynamics and the design of nature reserves. Biol. Conserv. 13: 27–37.

Poore, M. E. D., 1968. Studies in Malaysian rainforest. J. Ecol. 56: 143–189.

Romme, W. H. & Martin, W. H., 1982. Natural disturbance by treefalls in old-growth mixed mesophytic forest: Lilley Cornett Woods, Kentuckey. In: R. N. Muller (ed.), Central Hardwood Forest Conference IV Proceedings, pp. 367–383.

Runkle, J. R., 1982. Patterns of disturbance in some old-growth mesic forests of eastern North America. Ecology 63: 1533–1546.

Shidei, T., 1974. Forest vegetation zones. In: M. Numata (ed.), The flora and vegetation of Japan. Copub. Kodansha, Tokyo & Elsevier Scientific, Amsterdam, pp. 87–124.

Tamate, S., Kashiyama, T., Sasanuma, T., Takahashi, K. & Matsuoka, H., 1977. On the distribution map of forest wind damage by typhoon No. 15, 1954 in Hokkaido (research note). Bull. Gov. For. Exp. Sta., Tokyo 289: 43–67.

Vankat, J. L., Blackwell Jr, W. H. & Hopkins, W. E., 1975. The dynamics of Hueston Woods and a review of the question of the successional status of the southern beech-maple forest. Castanea 40: 290–308.

Veblen, T. T., Schlegel, F. M. & Escobar, R. B., 1980. Structure and dynamics of old-growth Nothofagus forests in the Valdivian Andes, Chil. J. Ecol. 68: 1–31.

Watanabe, R., Nakashizuka, T., Honma, S., Hara, M. & Yoda, S., 1985. Studies on the Fagus crenata forest in Kayanodaira II. Fallen trees by wind of the typhoon No. 10 in 1982. Bull. Inst. Nature Educ. Shiga Heights, Shinshu Univ. 22: 15–18 (in Japanese with an English summary).

Watt, A. S., 1947. Pattern and process in plant community. J. Ecol. 35: 1–22.

White, P. S., 1979. Pattern, process, and natural disturbance in vegetation. Bot. Rev. 45: 229–299.

Whitmore, T. C., 1975. Tropical rainforests of the Far East. Clarendon Press, Oxford, 281 pp.

Whitmore, T. C., 1982. On pattern and process in forests. In: E. I. Newman (ed.), The plant community as a working mechanism. pp. 45–59. Blackwell, Oxford.

Williamson, G. B., 1975. Pattern and seral composition in an old-growth forest. Ecology 56: 727–731.

Yoda, L., Kira, T., Ogawa, H. & Hozumi, K., 1963. Self thinning in overcrowded stands under cultivated and natural conditions (Intraspecific competition among higher plants 11). J. Biol. Osaka City Univ. D8: 161–178.

The appearance and disappearance of major vegetational assemblages: Long-term vegetational dynamics in eastern North America

Thompson Webb III*
Department of Geological Sciences, Brown University, Providence, RI 02912–1846, USA

Abstract

Contoured maps of pollen percentages from eastern North America illustrate that the vegetation changed continuously during the past 18 000 years. The maps show that the geographic distribution for selected pollen types from 500 years ago parallels the major patterns in the vegetation, and that the correspondence is good enough to justify using maps of fossil pollen to interpret the patterns and composition of the broad-scale vegetation from earlier times. Data for *Artemisia,* Cyperaceae, *Picea, Betula, Alnus, Abies,* and *Pinus* pollen illustrate how the changing location and abundance of their plant taxa and populations altered the spatial pattern, local composition, and the overall structure of the vegetation and thus led to the disappearance as well as appearance of major biomes and ecotones. For example, the *Picea* parkland biome that existed from 18 000 to 12 000 yr B.P. within a broad region south of the ice sheet disappeared within 2000 years after 12 000 yr B.P., and the modern boreal forest biome began to develop across a large area of central Canada only after 6000 yr B.P. On a time scale of 10^4 to 10^6 yr, these types of vegetation changes involve a continual overlapping and separation of the abundance distributions for different taxa. Such changes are likely to have been a feature of vegetational dynamics for millions of years.

Introduction

Contoured maps of pollen percentages from networks of radiocarbon-dated sediment cores illustrate how the vegetation has changed during the past 18 000 years within eastern North America and Eurasia (Bernabo & Webb, 1977; Gaudreau & Webb, 1985; Huntley & Birks, 1983; Peterson, 1983; Ralska-Jasiewiczowa, 1983; Webb, 1985). Temporal sequences of the maps record the numerical as well as spatial expansion and contraction of individual plant taxa and depict the changing pattern of abundance gradients as the range boundaries and density centers for various taxa change in location and areal coverage leading to new associations among taxa. During the past 18 000 years, the changes were large enough to result in the appearance and disappearance of major biomes and ecotones (Ritchie, 1976; Davis, 1983b).

Recent paleoclimatic studies have indicated the extent of climatic change during the past 18 000 years (CLIMAP, 1981; Hecht, 1985; Street & Grove,

1979), and various studies are showing what the past climatic patterns were (Kutzbach & Guetter, 1986; Kutzbach & Wright, 1985). In eastern North America, much of the climatic change from 18000 to 6000 yr B.P. was strongly influenced by the retreat and disappearance of the Laurentide ice sheet, which was near its maximum extent at 18000 yr B.P. The large climatic changes during this time period make it an ideal one to use in developing theory about how long-term climatic change affects the vegetation (Bartlein *et al.,* 1984, 1986; Davis *et al.,* 1986; Prentice, 1983, 1986; Ritchie, 1986; Overpeck *et al.,* 1985; Solomon & Webb, 1985; Webb, 1986). Radiocarbon dating within this time span has a precision fine enough to permit mapping in intervals upto 500 years (Webb, 1982). Because mapping at this time scale is impossible for any previous 18000 year interval, the past 18000 years gains added importance in providing a window for viewing comparable intervals during the more distant geological past.

The focus of my paper is on describing some of the changes in vegetational composition and structure in the areas of eastern North America where *Picea* trees have been numerous within the past 18000 years.

Data and methods

A sequence of isopoll maps with contours of equal pollen percentage illustrates the distribution of the pollen types at each 2000 year interval from 18000 to 2000 yr B.P. The final map in the series is for 500 yr B.P., which is just before European settlement in North America, and thus illustrates the 'pre-settlement' vegetation patterns before widespread logging and agriculture modified the pollen record in much of eastern North America south of 50°N (Bernabo & Webb, 1977; Van Zant *et al.,* 1979; Gaudreau & Webb, 1985).

The pollen data came from a network of lakes and bogs where sediments had been collected and radiocarbon-dated, and the data set represents the work of over 80 palynologists since 1960. The maps were produced from computer files listing the pollen counts and radiocarbon dates for each site (Webb *et al.,* 1983a, b; Gaudreau & Webb, 1985). First, the dates were used to estimate the age for each pollen sample, and then the pollen percentages from adjacent pollen samples were linearly interpolated to obtain estimates of the pollen percentages for each 500 year interval (e.g. 500, 1000, 1500, ..., 18000 yr B.P.) for which data existed at each site. In general, no data were mapped

at a site for time intervals more than 500 years older than the oldest radiocarbon data at that site. The pollen percentages were calculated from a pollen sum of all tree, shrub, and non-aquatic herb pollen. At a few sites, locally over-represented wetland taxa (e.g. *Nyssa, Larix*) were deleted from the pollen sum. Initial versions of the maps showed anomalously high values for these taxa at single sites.

Site density varies depending on the dates being mapped. The total number of available sites increases from 15 at 18000 yr B.P. to 70 at 12000 yr B.P. and to over 250 from 10000 yr B.P. to 500 yr B.P. (Fig. 1). See Webb (1985) for a table and bibliography for the sites with data available at 6000 yr B.P.; Webb et al. (1983a, b) and Gaudreau & Webb (1985) provide lists of the sites and their dating control in southern Quebec and in central and northeastern United States. Three contours for each pollen type were chosen to illustrate the key features of the distribution for that taxon. The differential scaling of pollen types depended upon the overall abundance of the type in the pollen record. *Pinus* pollen is the most abundant type. Its maps had contours of 20% and 40%, but it is too ubiquitous for a low-valued contour to be informative. For most taxa, the contour with the lowest value lies near or outside the range boundary for the plants, the intermediate contour marks areas where the plants are in low to moderate numbers, and the contour with the highest value identifies the region of high abundances (i.e. the density center) for the taxon.
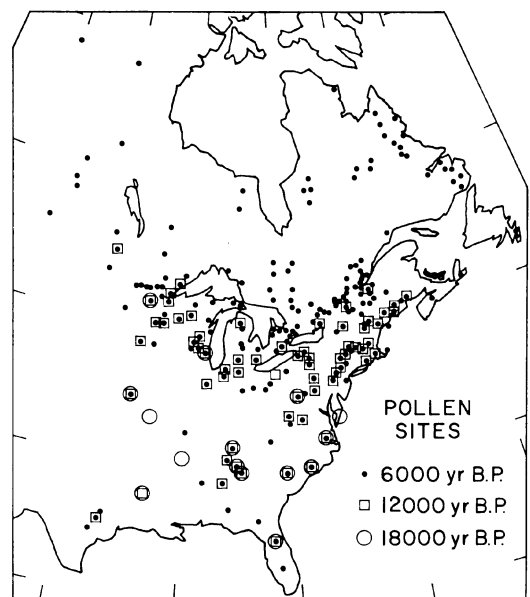
Fig. 1. Location of sites with pollen data for 6000 yr B.P. (black dots). Open squares mark sites with data for 12000 yr B.P. and open circles mark sites with data for 18000 yr B.P.
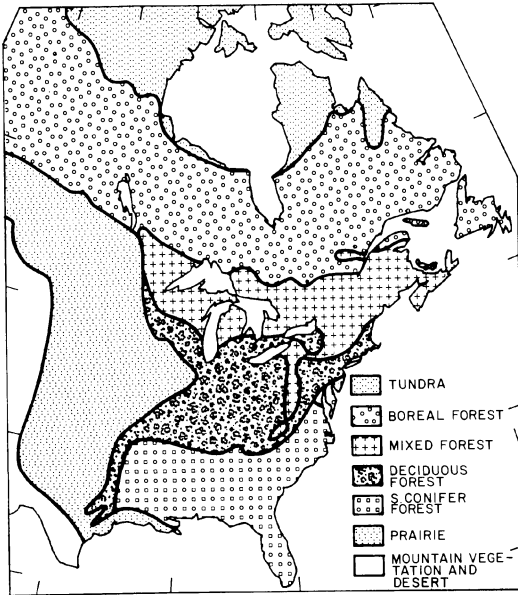
Fig. 2. Generalized vegetation map for eastern North America (modified from Webb & Bernabo, 1977).

Results

500 yr B.P.

Maps of pollen data for 500 yr B.P. show that patterns in the abundances for major pollen types parallel the patterns in the major vegetational regions (Figs. 2, 3). The pollen samples in the region of prairie vegetation contain high values of *Artemisia* and Cyperaceae pollen; the sites in the tundra and forest-tundra have high values of Cyperaceae, *Betula,* and *Alnus* pollen; the sites in the boreal forest exhibit high values of *Picea, Abies,* and *Betula* pollen; the sites in the mixed conifer-hardwood forest contain high values of *Betula* and *Tsuga* pollen in the east and of *Pinus* pollen in the west; the sites in the deciduous forest are dominated by high values of *Quercus* pollen; and the sites in the southern conifer forests have high values of *Pinus* pollen. Steep gradients in the pollen abundances mark the ecotones between the vegetation regions. Six biomes and five ecotones are evident, and the mixed
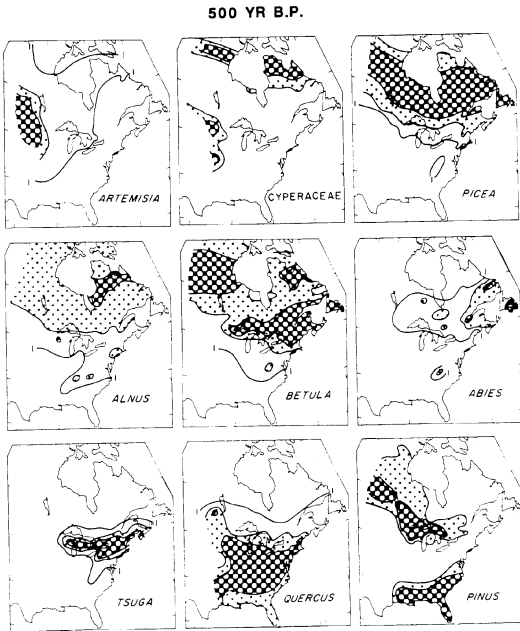
500 YR B.P.

Fig. 3. Maps with isopolls (contours of equal pollen percentage) for 500 yr B.P. for *Artemisia* (1, 5, 10%), Cyperaceae (1, 5, 10%), *Picea* (1, 5, 20%), *Alnus* (1, 5, 20%), *Betula* (1, 10, 20%), *Abies* (1, 2, 6%), *Tsuga* (1, 5, 20%), *Quercus* (1, 5, 20%), and *Pinus* (20, 40%) pollen. Numbers in parentheses give percentages for isopolls. Stippling highlights regions with intermediate (white with black dots) and high (black with white circles) percentages.
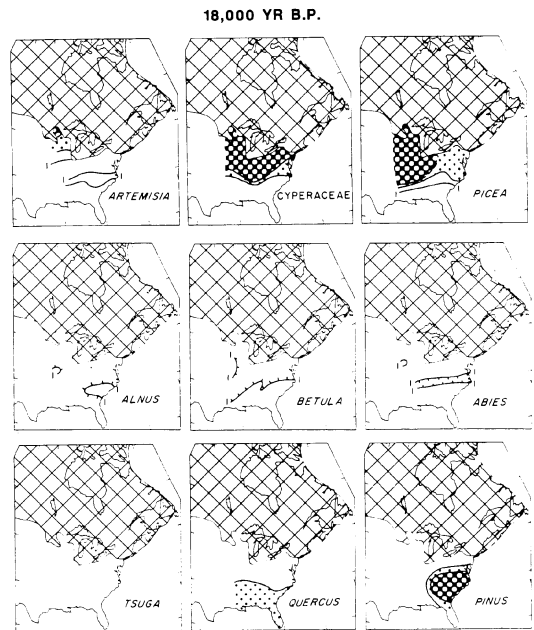
18,000 YR B.P.

Fig. 4. Maps with isopolls for 18 000 yr B.P. for *Artemisia, Cyperaceae, Picea, Alnus, Betula, Abies, Tsuga, Quercus* and *Pinus* pollen. See Fig. 3 for labeling of isopolls and stippled regions for each pollen type.

180

conifer-hardwood forest is further divided into eastern and western sectors with *Pinus* pollen highest in the west. The ability of pollen data to record the modern patterns in the vegetation demonstrates the general sensitivity of the data to vegetational patterns and makes it likely that isopoll maps from past times record the prevailing vegetational patterns. In describing the isopoll maps, I therefore followed the practice of Webb *et al.* (1983a, p. 142) and provided a narrative description of the vegetation changes illustrated by the maps of pollen percentages.

18 000 yr B.P.

The patterns in the 9 pollen types at 18 000 yr B.P. (Fig. 4) differ strongly from those at 500 yr B.P. (Fig. 3). These differences are evident despite a de-crease in site number from 250 (Fig. 2) to 15 (Fig. 3). The ice sheet extended south to about 41 °N and was probably over 3 km in maximum height in central Canada (Denton & Hughes, 1981). South of it *Artemisia,* Cyperaceae, and *Picea* pollen were abundant over a wide region and indicate growth of a *Picea* parkland quite different from the boreal forest of today. The *Picea* trees must have grown far enough apart to allow for the relatively abundant growth of herbs and the ready dispersal of herb pollen. Neither *Betula* nor *Alnus* pollen were abundant in this region or elsewhere. Northern pines grew abundantly in forests from Georgia to the Carolinas (Watts, 1983), and Smith (1985) has recently provided evidence of abundant *Pinus* pollen from a site in the Ozarks that dates back to 17 000 yr B.P. Along the ecotone between the *Picea* parkland and the *Pinus* forests, three sites possessed greater than 1% *Abies* pollen. This evidence
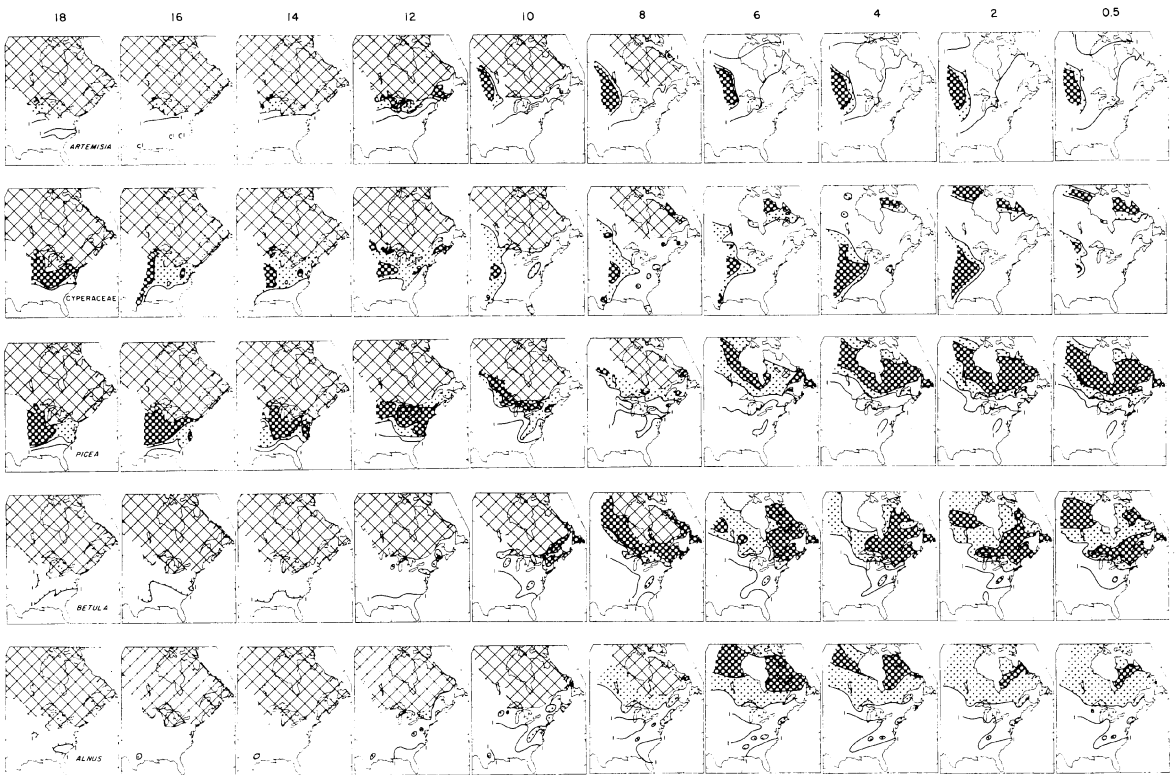
Fig. 5. Isopoll maps from 16 000 to 500 yr B.P. for *Artemisia,* Cyperaceae, *Picea, Betula* and *Alnus.* See Fig. 3 for labeling of isopolls and stippled regions for each pollen type.

may indicate a center for *Abies* populations at a low level of overall abundance (Delcourt & Delcourt, 1984), but the strong under-representation of *Abies* tree abundances by *Abies* pollen abundances (Webb *et al.*, 1981) makes this interpretation tentative at best.

Quercus and *Carya* pollen were somewhat abundant east of Louisiana, but no published terrestrial site south of 32 °N has provided information on these taxa for 18 000 yr B.P. (Watts, 1983). The best fossil pollen data published so far are from Sheelar Lake (Watts & Stuiver, 1980) in north Florida, which recorded 9% *Quercus*, 3% *Carya*, and 75% *Pinus* pollen at 18 500 yr B.P. but had an hiatus in sedimentation from then to 14 500 yr B.P. I used these data but not those from Goshen Springs. Evidence for a full glacial pollen assemblage similar to the Holocene deciduous forest is therefore lacking (Davis, 1983b).

Several pollen types that are fairly abundant over broad areas today have either low values at isolated sites at 18 000 yr B.P., e.g. *Tsuga* (Fig. 4) and *Fagus*, low values at several adjacent sites, e.g. *Betula*, *Alnus*, and *Abies* (Fig. 4) and also *Ostrya/Carpinus*, *Fraxinus*, *Ulmus*, and *Castanea*. The climatic conditions 18 000 yr B.P. were colder and drier than today (Kutzbach & Wright, 1985) and were probably unfavorable to wide-spread growth of large numbers of individuals of these taxa. The main vegetation pattern at 18 000 yr B.P. was the contrast between the *Picea* parkland in the north and *Pinus* forest or woodland in the south.

Map sequence from 18 000 to 500 yr B.P.

Maps between 18 000 and 500 yr B.P. show the timing and patterns of change as the vegetation shifted from its full glacial distribution to its distribution just before European settlement. These maps illustrate temporal changes in location and abundance for various taxa and the resultant changes in the overlapping of distributions among taxa. These individualistic changes among the different taxa led to profound changes in the pattern, composition, and structure of the vegetation that produced maximum amounts of *Picea* pollen. Similar changes have occurred in other vegetational regions (Davis, 1983b).

Picea, Artemisia, and Cyperaceae. At 18 000 yr B.P. the highest values of *Picea* pollen were centered in Missouri (King, 1973), but a second population center also existed on the east coast as is evident in the map for 16 000 yr B.P. (Fig. 5). By 12 000 yr B.P., *Picea* values had increased over a broad region south of the ice sheet, and the orientation of the main abundance gradients for *Picea* pollen was primarily north-south. This orientation changed to northeast-southwest by 10 000 yr B.P. as the band of high values became narrower and was located just in front of the ice sheet in the west. After 10 000 yr B.P. the values of *Picea* pollen generally decreased until 7000 yr B.P. when only 5 isolated sites had values above 20%. By 6000 yr B.P., with central Canada free of ice, new regions with high
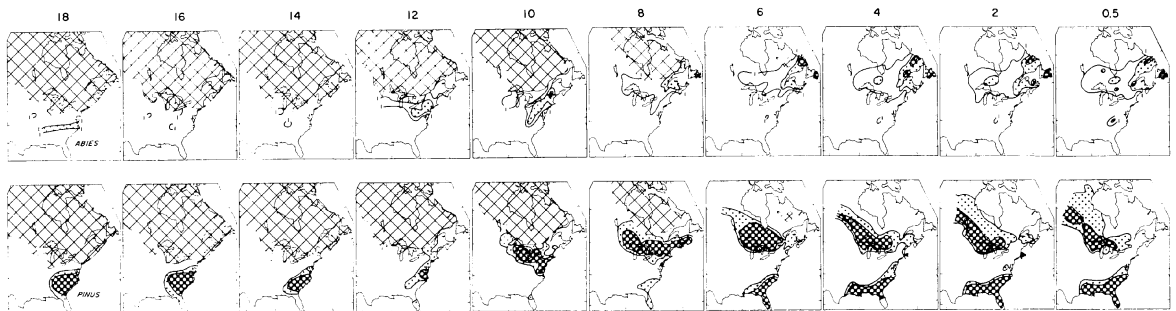
Fig. 6. Isopoll maps from 18 000 to 500 yr B.P. for *Abies* and *Pinus* pollen. See Fig. 3 for labeling of isopolls and stippled regions for each pollen type. Cross-hatching marks area with Laurentide ice sheet.

values of *Picea* pollen appeared both along the Labrador coast and south and west of Hudson Bay. *Picea* populations then expanded both to the north and south as the modern boreal forest developed. The changes for *Picea* populations were gradual from 18 000 to 12 000 yr B.P., abrupt from 12 000 to 10 000 yr B.P., gradual from 10 000 to 7000 yr B.P., abrupt between 7000 and 6000 yr B.P. and again gradual up to 500 yr B.P.

Evidence from other pollen types indicates the nature of the vegetational changes associated with these changes in *Picea* populations. *Artemisia* and Cyperaceae pollen had high to intermediate values in the north and low values in the south from 18 000 to 12 000 yr B.P. (Fig. 5). The distributions of both of these herbaceous types coincide with part or all of the region with high *Picea* pollen. At 18 000 yr B.P., high values of *Artemisia* pollen occurred only in Minnesota, but by 12 000 yr B.P., its region of high values extended as a narrow band all along the southern edge of the ice sheet and had a marked north-south gradient. Cold, dry environmental conditions along the ice front probably favored the prolific growth of arctic *Artemisia* plants.

In this same region and much further to the south, Cyperaceae plants grew in moderate to high numbers, and the coincidence between the patterns of *Picea* pollen and these herbaceous types suggests that a broad region of *Picea* woodland or parkland existed south of the ice sheet from 18 000 to 12 000 yr B.P. (Fig. 5). The spacing of the *Picea* trees was wide enough to permit the growth of the herbs and the dispersal of their pollen. Few *Betula* or *Alnus* shrubs or *Abies* trees grew in this region until 12 000 yr B.P. when moderate values for the pollen of each of these types first appeared in the northeastern United States (Figs. 5, 6). This vegetation differed markedly in both composition and structure from the modern boreal forest or forest-tundra. In these, *Picea, Pinus, Abies, Betula,* and *Alnus* are prominant today, but *Artemisia* is infrequent (Fig. 3). The *Picea* parkland was a favorable region for browsing by mastodons, and loess accumulated in it near major river systems in the midwestern United States (Porter, 1983).

Between 12 000 and 10 000 yr B.P., a major change occurred in the vegetation. The band of high values of *Picea* pollen at 10 000 yr B.P. marks a region of closed forest for the first time because the distributions of *Picea* pollen and herb pollen cease to overlap. The main orientation of the abundance gradients for *Artemisia* and Cyperaceae pollen shifted from north-south to east-west as the region of modern prairie developed. The arctic and boreal species producing the *Artemisia* and Cyperaceae pollen in the west were probably replaced by prairie species. This 90° change in the abundance gradient for these types suggests a major environmental change as hot dry summer conditions became important on the central plains and the cool dry conditions in front of the ice sheet became less pronounced.

Betula and Alnus. Between 12 000 and 10 000 yr B.P., the values for *Betula* and *Alnus* pollen increased in the east north of the highest *Picea* values. High values of Cyperaceae pollen also occurred in southern Labrador as the modern tundra and forest tundra began to develop. This development continued when the region of high values of Cyperaceae, *Betula,* and *Alnus* moved northward along the Labrador coast at 8000 yr B.P. and then expanded across the recently deglaciated parts of central and northern Canada by 6000 yr B.P. South of the ice sheet, the vegetational composition changed relatively quickly from 10 000 yr B.P. when *Picea* populations were prominent to 8000 yr B.P. when *Betula* trees and shrubs were prominent to 6000 yr B.P. when *Alnus* populations expanded rapidly across the deglaciated central region (Fig. 5). After 6000 yr B.P., the modern boreal forest developed (Ritchie, 1976) with the expansion of *Picea* populations both north and south, the confinement of high *Alnus* values to the modern forest-tundra zone, and the separation of the *Betula* pollen densities into three centers. Shrub-*Betula* populations became dominant in the northeastern forest-tundra zone, *Betula papyrifera* populations grew in the western region of boreal forest, and *B. papyrifera* populations graded southward into dominant *B. leutea* populations in the east-central region. The *B. papyrifera* populations grow today with the *Picea* and *Abies* in the southern boreal forest, and the *B. leutea* populations grow most

prolifically in the mixed conifer-hardwood forests of the east. During this same time period, the modern tundra also developed in the north. Note the westward expansion of the region with high values of Cyperaceae pollen after 8000 yr B.P.

Abies and Pinus. Today in the southeastern boreal forest, *Abies* populations are prominent along with *Picea* populations. The *Picea-Abies* association in the east first appeared at 12000 yr B.P. in the mid-Atlantic region. It expanded north and south to form a band along the Appalachians at 10000 yr B.P. (Fig. 6). By 8000 yr B.P., the geographic distribution of *Abies* populations was similar to their distribution today. In contrast, the regional distribution of *Picea* trees became similar to their modern one sometime after 6000 yr B.P. After 6000 yr B.P., *Abies* populations expanded north of the Great Lakes and also in New England (Davis *et al.,* 1980), southern Quebec (Webb *et al.,* 1983b), and the central Appalachians (Delcourt & Delcourt, 1985). Although significant numbers of *Picea* and *Abies* trees have grown together in some part of their range since 12000 yr B.P., each taxon followed a separate history in gaining its modern distribution (Davis, 1983b).

Just as the association between *Picea* and *Abies* populations varied through time, so too has the association between *Picea* and *Pinus* populations. Today, *Pinus* pollen percentages are highest in south-central Canada and overlap the region of moderate to high values of *Picea* pollen west of Hudson Bay (Figs. 5, 6). At 18000 yr B.P., the arrangement was quite different; the region of highest *Pinus* values was along the east coast, *Picea* values were highest in the west, and *Abies* values were low everywhere.

Between 18000 and 12000 yr B.P., *Pinus* populations decreased in abundance and area and became confined to the east coast. In the next 2000 years, the population center for *Pinus* broadened and shifted westward. A region of *Pinus* forests developed south of the narrow band of *Picea* forests. By 8000 yr B.P., the population center for *Pinus* had expanded northward, and *Pinus* forests grew over a wide area from Nova Scotia to Manitoba while just to the north *Picea* populations had declined. By 6000 yr B.P., *Pinus* populations had decreased in the east, and the population center had shifted westward to its present location in the northern Great Lakes region and central Canada. At the same time, the number of *Pinus* trees increased in the southeastern United States as a well-pronounced ecotone with *Quercus* forests and a new biome became established there (Figs. 4, 6).

Discussion

Recent research concerning climatic changes within the past 10^4 to 10^7 years indicates how orbitally induced seasonal changes in solar insolation have produced major climatic changes with periodicities of 20000 to 400000 years (Hays *et al.,* 1976). These changes arise from small, but regular, variations in the tilt of the earth's rotational axis (period of 41000 yr) and in the elliptical nature of the earth's orbit about the sun (period of 100000 yr) as well as from the precession of the axis of the earth's orbit (period of 20000 yr). Within the precession cycle, the earth moves from being closest to the sun in January, as it is now, to being closest in July, as it was 10000 years ago. At that time the northern hemisphere on average received 8% more radiation in summer than it does today (Kutzbach & Street-Perrott, 1985). The land-sea contrast in specific heat results in these changes in seasonal radiation alternately weakening and strengthening the monsoons in the tropics, where immediate climatic changes are induced (Kutzbach, 1981; Kutzbach & Street-Perrott, 1985).

The orbitally induced changes in monsoonal climates have always influenced global climates, with a major impact in the tropics (Berger *et al.,* 1984). Since at least the Miocene, the radiation changes have also induced the build up and melting of ice sheets in high latitudes. The long response times (10^3 to 10^4 yr) of the ice sheets introduced major lags into the climate system and its response to external forcing, but the imprint of orbital forcing is still clear in pacing of the variations in global ice volume (Imbrie & Imbrie, 1980; Imbrie, 1985).

By use of climate models and physical reasoning, paleoclimatologists are beginning to be able to

sketch out how the relatively simple radiational forcing (the sum of three periodic processes), which is uniform within latitudinal belts, produces a spatially regular but more heterogenous response in the atmosphere, oceans, and ice sheets. This response includes (1) continuous and sometimes abrupt changes in the atmospheric circulation; and (2) spatial patterns of different size for temperature and moisture differences between today and past times. The vegetation then imposes a whole new set of patterns by its response to these temporal and spatial variations in climatic variables (Bartlein et al., 1986). Not only does each plant species respond differently to climate, but it is also differentially sensitive to seasonal variations in moisture and temperature within its range as well as at its range boundary (north vs. south or east vs. west). No single climatic variable controls the spatial and temporal distribution of a pollen type or plant species.

The pollen maps (Figs. 5, 6) show how the associations among taxa varied as their constituent plant populations changed in location and abundance in response to climatic change. On time scales of 10^4 to 10^5 years, formations and their associated ecotones are ephemeral features, and plant taxa are capable of growing in a variety of locations and associations (Watts, 1973; Davis, 1983b). The individualistic behavior of taxa was prominent over the past 18000 years, and the changing mixture of taxa caused the composition and patterns to vary within the vegetation across eastern North America (Davis, 1976). During times of rapid climatic change, e.g. 12000 to 10000 yr B.P., the vegetation responded rapidly, and distribution patterns of individual taxa changed in orientation to reflect the new climatic gradients (Figs. 5, 6). For example, the movement of the pine populations between 12000 and 10000 yr B.P. indicates strong climatic control as seasonal temperature and atmospheric circulation patterns responded both to the retreating and lowering ice sheet and to the changing seasonal radiation (Webb, 1984; Webb, Bartlein Kutzbach, in press). After 8000 yr B.P., increased winter temperatures probably favored the increased numbers of pine trees in the previously oak-dominated forests of the southern United States (Webb et al., in press).

The magnitude and kinds of vegetational changes between 18000 yr B.P. and 500 yr B.P. (Figs. 3–6) presumably have occurred many times during the Quaternary (1.8×10^6 yr B.P. to present day) as the populations of plant taxa tracked the environments favorable to their growth. The specific sequence and timing of vegetation changes during the past 18000 years are unique to this interglacial, and mainly reflect the unique sequence of regional climate changes induced by the particular combination of orbital forcing (i.e. changes in precession, tilt, and eccentricity of the earth's orbit) that differentiates the past 18000 years from previous interglacials (Berger, 1978; Imbrie, 1985). Because the unique aspects of the past 18000 years arise from a unique combination of established processes and not from a set of unique processes, observations about the general behavior of climate and of taxa during this time period provide insights applicable to any 18000 year period within the Quaternary and even the Tertiary. This conclusion holds even though the presence and retreat of the Laurentide ice sheet had a major impact on North American climates during the past 18000 years. Certain areas and time intervals within this period experienced climate changes little influenced by the presence of the ice. For instance, the climate changed significantly in the tropics where the climatic effects of the large ice sheets are minimal (Kutzbach & Street-Perrot, 1985), and the climate and vegetation also changed significantly during the past 6000 years (Fig. 5) when global ice volume was relatively constant. These climate changes are not unique to the Quaternary, and their associated vegetational responses may be representative of vegetation changes from the Tertiary and earlier.

The vegetational dynamics recorded on the pollen maps (Figs. 5, 6) pose an interesting problem for modeling. Webb et al. (in press) have recently used response functions that can estimate pollen percentages from linear combinations of temperature and precipitation values. The challenge is whether a synoptically consistent set of climatic conditions can be found that, via the response functions, can generate the spatial and temporal patterns illustrated on the maps. Initial research to explore this possibility has used the response functions to esti-

mate pollen diagrams from the patterns of past climatic change that have been simulated by general circulation models (Webb *et al.*, in press). Model simulations exist for each 3000 year interval from 18 000 yr B.P. to present (Kutzbach & Guetter, 1986). Preliminary results for data from eastern North America indicate that certain of the model-generated results match many of the observed changes in the pollen record. The model results show the abrupt change in patterns between 12 000 and 10 000 yr B.P. and reproduce many of the changes in the abundances of *Picea* and *Quercus* pollen. Other features on the pollen maps, especially those sensitive to precipitation changes, are not as well reproduced by the model results. The implications of these initial results are under investigation.

The role of succession and disturbance

Ecological texts generally restrict their discussion of vegetational dynamics to primary and secondary succession (e.g., Mueller-Dombois & Ellenberg, 1974; Kershaw & Looney, 1985). These processes occur, but the eastern North American pollen maps as well as comparable maps for Europe (Huntley & Birks, 1983) show that neither process explains the major vegetational changes of the past 18 000 years. The maps would have to be drawn for each 50 year interval in order for variations due to gap-phase processes and fire-induced succession to appear in a sequence of maps. Such variations would little affect the broad-scale patterns and would appear as a rapid flickering of the mapped values, were a film made from such a sequence of maps from the past 20 000 years. Although important within stands over decades and centuries, successional changes comprise a continuously active but relatively minor component of the continent-wide variations in ancient vegetation.

Fires, disease, and wind-throws continuously create either gaps within forests or bare patches in open vegetation and thus produce openings that can be colonized by new taxa as well as exploited by established taxa. Invading species depend upon the continual occurrence of openings. At the scale of a woodland stand, fire, disease and wind-throws often appear as singular events, but at the continental scale, these disturbance factors are continuously active agents with characteristic frequencies that change in space and time. Climatic change can alter these frequencies, and individual disturbances can be the agents that accelerate vegetation change once the climate has changed. Disturbance events tend to affect small areas ($<10^3$ km^2). In some cases, however, pollen data record synchronous (± 300 yr) wide-ranging disturbance events, such as a large decrease in abundance over an entire species range, that may have resulted from the rapid spread of a disease (Davis, 1981; 1983a). The pattern of these abundance changes in time and space contrasts with those that are climatically induced, because the latter tend to be time-transgressive and to result in shifts in the location of both the population center and the range boundaries of the taxon (Webb, 1982).

Vegetation dynamics and evolution

Evolutionary change has occurred within the context of these millenial-scale climatic and vegetational dynamics of the past million or more years. The development of new genotypes is probably a continuous process as taxa adapt to their new locations and associations, but speciation within a genus is a relatively rare event, because the paleontological record reveals an average life time for plant species of 1 to 10 million years (Tiffney, pers. comm.). Each species must therefore be capable a) of tracking these millenial-scale changes in climate (Davis, 1983b; Webb, 1986) and b) of coping with major changes in vegetational composition and structure. For certain species, this has involved alternation between growing in dense populations across wide areas and growing in sparse populations in isolated microhabitats. The latter situation is a precondition for allopatric speciation, but its time scale of occurrence may be too short for the separate populations (a) to develop major morphological differences; (b) to become genetically isolated from the other populations. Millenial-scale climatic change can therefore create some of the necessary condi-

tions for plant speciation, but the environmental changes may be too frequent to allow the process to go to completion. Longer term geological and environmental processes (e.g. plate tectonics) probably play an important role in speciation.

As research into the effects of climatic variations on evolution continues, the different vegetative and vegetational responses to climatic change on various time scales must be recognized. On an annual time scale, trees are sessile organisms that use phenological and physiological mechanisms to survive annual changes in climate. On time scales of 10^4 to 10^5 years, tree species are comprised of many populations with dispersal and establishment capabilities that allow the species to be mobile enough to track the millenial-scale changes in climate. On time scales 10^3 to 10^4 times longer, most tree species can no longer survive the full range of climatic change, and speciation (whether rapid or gradual) becomes an important mechanism in allowing the genus or higher taxon to continue. Evolutionary processes are active on all time scales; but, on relatively short time scales, phenological, physiological, and ecological processes dominate and modify the rate of evolution by allowing species to survive large environmental changes. In light of the ephemeral nature of vegetational assemblages, research is needed to demonstrate the specific ways in which vegetational dynamics, such as succession or the changing composition and structure of communities or formations, affect evolutionary processes.

References

Bartlein, P. J., Prentice, I. C. & Webb III, T., 1986. Climatic response surfaces based on pollen from some eastern North American taxa. J. Biogeogr. 13: 35–57.

Bartlein, P. J., Webb III, T. & Fleri, E. C., 1984. Holocene climatic change in the northern Midwest: pollen-derived estimates. Quat. Res. 22: 361–374.

Berger, A., 1978. Long term variations of daily insolation and Quaternary climatic changes. J. Atms. Sci. 35: 2362–2367.

Berger, A., Imbrie, J., Hays, J., Kukla, G. & Saltzman, B. (eds), 1984. Milankovitch and climate. Reidel, Dordrecht.

Bernabo, J. C. & Webb III, T., 1977. Changing patterns in the Holocene pollen record from northeastern North America: a mapped summary. Quat. Res. 8: 64–96.

CLIMAP Project Members, 1981. Seasonal reconstructions of the Earth's surface at the Last Glacial Maximum. Geol. Soc. Amer. Map and Chart Series 36.

Davis, M. B., 1976. Pleistocene biogeography of temperate deciduous forests. Geosci. Man 13: 13–26.

Davis, M. B., 1981. Outbreaks of forest pathogens in Quaternary history. IV International Palynological Conference, Lucknow (1976–77) 3: 216–227.

Davis, M. B., 1983a. Holocene vegetational history of the eastern United States. In: H. E. Wright Jr (ed.), Late Quaternary environments of the United States. Vol. 2. The Holocene, pp. 166–181. University of Minnesota Press, Minneapolis.

Davis, M. B., 1983b. Quarternary history of deciduous forests of eastern North America and Europe. Ann. Missouri Bot. Gard. 70: 550–563.

Davis, M. B., Spear, R. W. & Shane, L. C. K., 1980. Holocene climate of New England. Quat. Res. 14: 240–250.

Davis, M. B., Woods, K. D., Webb, S. L. & Futyma, R. B., 1986. Dispersal versus climate: expansion of Fagus and Tsuga into the upper Great Lakes region. Vegetatio 69: 93–103.

Delcourt, H. R. & Delcourt, P. A., 1984. Late-Quaternary history of the spruce-fir ecosystem in the Southern Appalachian Mountain region. In: P. S. White & J. Wood (ed.). The southern Appalachian Spruce-Fir ecosystem: its biology and threats. U.S. Research Resource Management Report Ser-71: 22–35.

Delcourt, H. R. & Delcourt, P. A., 1985. Quaternary palynology and vegetational history of the southeastern United States. In: V. M. Bryant & R. G. Holloway (eds), Pollen records of Late-Quaternary North American sediments, pp. 1–37. American Association of Stratigraphic Palynologists Foundation, Dallas.

Denton, G. H. & Hughes, T. J. (eds), 1981. The last great ice sheets. J. Wiley & Sons, New York.

Gaudreau, D. C. & Webb III, T., 1985. Late-Quaternary pollen stratigraphy and isochrone maps for the northeastern United States. In: V. M. Bryant & R. G. Holloway (eds), Pollen records of Late-Quarternary North American sediments, pp. 247–280, American Association of Stratigraphic Palynologists Foundation, Dallas.

Hays, J. D., Imbrie, J. & Shackleton, N., 1976. Variations in the earth's orbit: pacemaker of the ice volume cycle. Science 194: 1121–1132.

Hecht, A. D. (ed.), 1985. Paleoclimate analyses and modeling. J. Wiley & Sons, New York.

Huntley, B. & Birks, H. J. B., 1983. An atlas of past and present pollen maps for Europe: 0–13 000 years ago. Cambridge University Press, Cambridge.

Imbrie, J. & Imbrie, J. Z., 1980. Modeling the climate response to orbital variations. Science 207: 943–953.

Imbrie, J., 1985. A theoretical framework for the Pleistocene ice ages. J. Geol. Soc. London 142: 417–432.

Kershaw, K. A. & Looney, J. H. H., 1985. Quantitative and dynamic plant ecology. E. Arnold Ltd., London.

King, J. E., 1973. Late Pleistocene palynology and biogeography of the western Missouri Ozarks. Ecol. Monogr. 43: 539–565.

Kutzbach, J. E., 1981. Monsoon climate of the early Holocene:

climate experiment with the earth's orbital parameters for 9000 years ago. Science 214: 59–61.

Kutzbach, J. E. & Street-Perrott, F. A., 1985. Milankovitch forcing of fluctuations in the level of tropical lakes from 18 to 0 kyr B.P. Nature 317: 130–134.

Kutzbach, J. E. & Wright Jr, H. E., 1985. Simulation of the climate of 18000 yr B.P.: results for North American/North Atlantic/European sector and comparison with the geologic record. Quat. Sci. Rev. 4: 147–187.

Kutzbach, J. E. & Guetter, P. J., 1986. The influence of changing orbital parameters and surface boundary conditions on the climate simulations for the past 18000 years. J. Atms. Sci. 43: 1726–1759.

Mueller-Dombois, D. & Ellenberg, H., 1974. Aims and methods of vegetation ecology. J. Wiley & Sons, New york.

Overpeck, J. T., Webb III, T. & Prentice, I. C., 1985. Quantitative interpretaton of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. Quat. Res. 23: 87–108.

Peterson, G. M., 1983. Holocene vegetation and climate in the western USSR. Ph.D. Thesis, University of Wisconsin, Madison.

Porter, S. C. (ed.), 1983. Late-Quaternary environments of the United States. Vol. 1. The Late Pleistocene. Univ. of Minnesota Press, Minneapolis.

Prentice, I. C., 1983. Postglacial climatic change: vegetation and the pollen record. Progr. Phys. Geogr. 7: 273–286.

Prentice, I. C., 1986. Vegetation responses to past climatic variation mechanisms and rates. Vegetatio 67: 131–141.

Ralska-Jasiewiczowa, M., 1983. Isopollen maps for Poland: 0–11000 yr B.P. New Phytol. 94: 133–175.

Ritchie, J. C., 1976. The late-Quaternary vegetational history of the western interior of Canada. Can. J. Bot. 54: 1793–1818.

Ritchie, J. C., 1986. Vegetation-climate equilibrium in relation to palaoecology. Vegetatio 67: 65–74.

Smith Jr, E. N., 1985. Paleovegetation zonation and tree-species migrations in southeastern Missouri: 17110 yr B.P. to the present. Palynology 9: 253.

Solomon, A. M. & Webb III, T., 1985. Computer-aided reconstruction of late-Quaternary landscape dynamics. Ann. Rev. Ecol. Syst. 16: 63–84.

Street, F. A. & Grove, A. T., 1979. Global maps of lake-level fluctuations since 30000 B.P. Quat. Res. 10: 83–118.

Van Zant, K. L., Webb III, T., Peterson, G. M. & Baker, R. G., 1979. Increased *Cannabis/Humulus* pollen, an indicator of European agriculture in Iowa. Palynology 3: 229–233.

Watts, W. A., 1973. Rates of change and stability in vegetation in the perspective of long periods of time. In: H. J. B. Birks

& R. G. West (eds), Quaternary plant ecology, pp. 195–206. Blackwell, Oxford.

Watts, W. A., 1983. Vegetational history of the eastern United States 25000 to 10000 years ago. In: S. C. Porter (ed.), Late-Quaternary environments of the United States. Vol. 1. The late Pleistocene, pp. 294–310. University of Minnesota Press, Minneapolis.

Watts, W. A. & Stuiver, M., 1980. Late Wisconsin climate of northern Florida and the origin of species-rich deciduous forest. Science 210: 325–327.

Webb III, T., 1982. Temporal resolution in Holocene pollen data. Third North American Paleontological Convention, Proc. 2: 569–572.

Webb, III T., 1984. Late-Quaternary changes in climate seasonality in eastern North America. American Quaternary Association. Program and Abstracts, Boulder, CO, pp. 138–139.

Webb III, T., 1985. A Global paleoclimate data base for 6000 yr B.P. U. S. Department of Energy Report TR-018. Department of Energy, Washington, DC.

Webb III, T., 1986. Is vegetation in equilibrium with climate? How to interpret late-Quarternary pollen data. Vegetatio 67: 75–91.

Webb III, T. & Bernabo, J. C., 1977. The contemporary distribution and Holocene stratigraphy of pollen in eastern North America. In: W. C. Elsik (ed.), Contributions of stratigraphic palynology. Vol. 1. Cenozoic Palynology, pp. 130–146. American Association of Stratigraphic Palynologists, Contr. Ser. No. 5A, AASP Foundation, Dallas.

Webb III, T., Howe, S. E., Bradshaw, R. W. H. & Heide, K. M., 1981. Estimating plant abundances from pollen percentages: the use of regression analysis. Rev. Palaeobot. Palyn. 32: 269–300.

Webb III, T., Cushing, E. J. & Wright Jr, H. E., 1983a. Holocene changes in the vegetation of the Midwest. In: H. E. Wright Jr (ed.), Late-Quaternary environments of the United States, Vol. 2, The Holecene, pp. 142–165. Univ. of Minnesota Press, Minneapolis.

Webb III, T., Richard, P. J. H. & Mott, R. J., 1983b. A mapped history of Holocene vegetation in southern Quebec. Syllogeus 49: 273–336.

Webb III, T., Bartlein, P. J. & Kutzbach, J. E., in press. Climatic change in eastern North America during the past 18000 years: comparisons of pollen data with model results. In: H. E. Wright, Jr. & W. F. Ruddiman (eds.), North America and adjacent oceans during the last deglaciation. Geol. Soc., Amer., Decade of North American Geology.

Climate and plant distribution at global and local scales

F. I. Woodward[1] & B. G. Williams[2], *
[1]*Department of Botany, University of Cambridge, Downing Street, Cambridge, CB2 3EA, U.K.;*
[2]*Department of Physical Chemistry, University of Cambridge, Lensfield Road, CB2 1EP, U.K.*

Abstract

This paper investigates, with predictive models, the utility of ecophysiological responses to climate as predictors of plant distribution. At the global scale responses to extreme minimum temperatures and to the hydrological budget effectively predict the distribution limits of the major vegetation types of the World. A minimum temperature of $-15\,°C$, for example, appears critical in controlling the poleward spread of vegetation that is dominated by evergreen broadleaved species; however, the presence or absence of more frost resistant species, such as those that are deciduous broadleaved, is not obviously explained in terms of extremes of climate. In such cases, predicting the competitive relationships between species is necessary and dependent on the climatic sensitivity of population dynamics.

Introduction

The distribution of the major vegetation types of the World has been well documented for nearly two centuries since the account of Von Humboldt & Bonpland (1805). This early account attempted to explain the varying patterns of plant distribution in terms of global patterns in climate. The notion that the dominant control of the distribution of vegetation is by climate is therefore as old as the maps which describe the distributions. However an understanding, or even a description, of the mechanisms by which climate may exert this control is still lacking. Even the more modern approaches of Holdridge (1967) and Box (1981) only provide very broad correlations between plant distribution and the many features of climate recorded at meteorological stations, and do not establish a fundamental physiological basis for the correlations. Neverthe-

less strong and universal correlations do exist between the distribution of vegetation and two major features of climate: temperature and precipitation. These two features of climate might therefore be useful starting points for investigating the mechanisms by which climate may control plant distribution.

Physiological bases

Precipitation

Spurr & Barnes (1980) describe correlations between the geographical limits of forests in northern America and annual precipitation. In central areas of the United States of America the lower limit of rainfall for full forest is 640 mm, declining with increasing latitude. This decline is related to the fall in temperature with latitude (Müller, 1982), which leads to lower rates of evaporation and transpiration. Moreover the throughfall of water to the for-

est floor increases at smaller totals of precipitation. Of this total throughfall, some is lost from the forest as run-off, whilst the remainder percolates into the soil where it is available for uptake by roots. The overall sum of precipitation minus evaporation (from leaf and soil surfaces) and run-off will be defined as the hydrological budget, and is a measure of the water that is available to plants.

In more detail Grier & Running (1977) and Waring *et al.* (1978) have shown that the leaf mass in a plant canopy is closely linked with the hydrological budget. Woodward (1986) has demonstrated the converse, that the hydrological budget may be used to predict the leaf mass in a plant canopy. These predictions were based on the Monteith modified Penman equation (Monteith, 1965), using observed measurements on stomatal resistances to water loss, the response of stomatal resistance to irradiance, leaf area and irradiance profiles through the plant canopy, the response of the leaf boundary layer resistance to wind speed and using local measurements of climate (from Müller, 1982).

The link between the hydrological budget and leaf mass may be understood in physiological terms. Observations on a range of plant species (Merrill, 1945; Addicott & Lyons, 1973; McMichael *et al.*, 1973) show feedback between leaf mass and the hydrological budget; increasing the supply of water to plants leads to leaf growth and an increase in leaf mass, whilst a decreased supply leads to leaf abscission and a reduction in leaf mass. Leaf mass may be used to predict vegetation type, with high leaf mass corresponding to forests and gradual reductions in leaf mass first corresponding with shrub vegetation and subsequently herbaceous and finally desert vegetation (Schulze, 1982).

Temperature

Many aspects of the thermal climate may influence plant distribution and so it appears improbable that the simple predictive approach which is effective for the hydrological budget can also be effective for temperature. However many measurements have shown that the annual minimum temperature may effectively limit plant distribution and perhaps

vegetation type, by exceeding the lethal threshold for survival. This has been shown for many species, in particular by Sakai (e.g. Sakai & Weiser, 1973; Sakai, 1978; Sakai & Wardle, 1978; and reviewed by Woodward, 1986).

This approach is illustrated by the low temperature thresholds for mortality of 220 arboreal species (Fig. 1). Four critical temperatures which may be critical in controlling the distribution of a particular life-form of vegetation are shown. When the minimum temperature falls in the region of 0 to 10 °C, mortality occurs for plants that are chilling-sensitive (Levitt, 1980). These species are typically evergreen when there are no significant periods of drought, but deciduous during periods of drought (Axelrod, 1966; Doley, 1981).

The next significant threshold is at −15 °C, the lowest temperature at which the majority of broad-leaved, evergreen species can survive. No specific process appears to coincide with this limit, although changes in cell water potential and the increasing probability of intracellular ice formation (Levitt, 1980) may be critical. The dominant physiognomy in sites which are subjected to frosts as low as −15 °C is evergreen, frequently with species which are also evergreen through drought (Larcher, 1981).

Water may be supercooled in vitro to a limit of about −40 °C at which point spontaneous nucleation occurs. The observations by Sakai support the view that supercooling occurs in overwintering plant tissues such as buds and that mortality occurs once ice nucleation occurs in the primordial tissue of the bud. In areas where the minimum temperature falls in the range of −15 °C and −40 °C the dominant life-form is broadleaved and winter-deciduous (Sakai, 1978) and is dependent on the ability to avoid intracellular ice formation by supercooling.

A number of species are capable of surviving temperatures below −40 °C, often with no observable limit for survival. The majority of the species in this category are conifers from the boreal regions but shade intolerant, pioneer species in the genera of *Populus* and *Betula,* which occur in the boreal region, also occur in this category (Sakai, 1978, 1979).
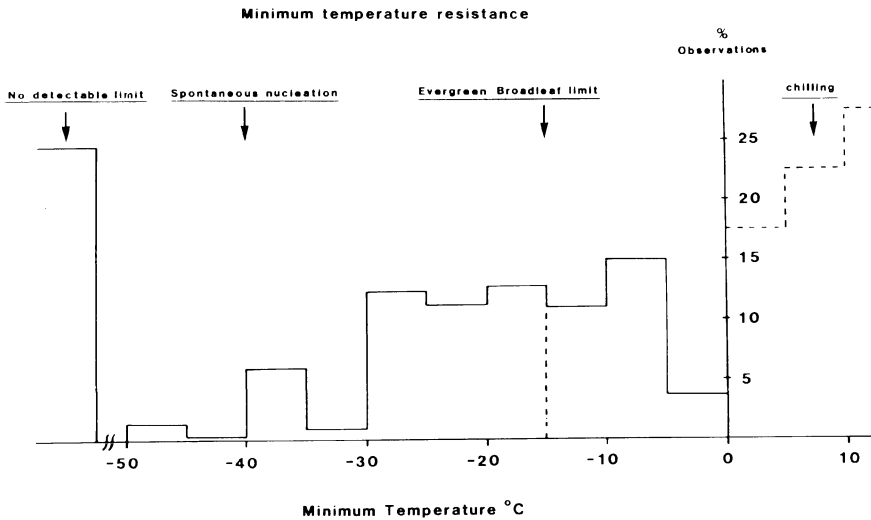
Fig. 1. Observations on the minimum temperature resistance of 220 arboreal species (———); projected response to chilling temperatures (— — —). Observable limits (↓).

At latitudes closest to the poles the ground is covered by snow to an extent which increases with latitude. Areas with a constant snow cover do not support the growth of flowering plants; however, in areas with sufficient summer warmth to melt the snow, flowering plants do occur. The ability of these species to survive at the extreme range of habitability is controlled by the period of snow free cover and by their capacity to complete leaf growth and perhaps reproduction in this short period. Observations of plant phenology and climate (e.g. Bliss, 1956; Callaghan & Collins, 1981; Tieszen *et al.*, 1981) suggest that a simple measure of the heat sum (the product of mean temperature and time) is an adequate descriptor of the poleward limit of forest and the onset of tundra. Tundra has not been sub-divided here, although Woodward (1986) discusses further sub-divisions. A heat sum equal to 50 month degrees has been used to estimate the transition from forest to tundra.

There is therefore a clear case for the impact of low temperatures in controlling the physiognomy of vegetation when moving in a poleward direction. However it is not clear why, for example, the most frost resistant species of conifers are not dominant species in warmer climates. Observations on the growth of these species in Botanic Gardens (e.g. Schnelle *et al.*, 1984) and on the growth of such species in vegetation (Satoo, 1983) suggest that the cost of greater frost resistance is at the expense of competitive ability. Specific biochemical processes appear to be involved in frost resistance (Senser & Beck, 1977) and these require a source of energy which might otherwise have been directed to growth. It is proposed therefore that whereas minimum temperatures are critical in controlling the poleward spread of a particular physiognomy, the equatorial spread may be controlled by competition between species which differ in cold tolerance (MacArthur, 1972).

Predicting global vegetation

Predictions from records of temperature

The observed limits to the survival of species from the major vegetation types of the World (Fig. 1) have been summarised in Table 1. These limits will be used to predict the distribution of the major vegetation types of the World. The aim is to test the hypothesis that climate acts through these ecophys-

Table 1. Defining vegetation type from observations of annual minimum temperature.

Minimum temperature range	Probable vegetation type
1. $\geq 10\,°C$	Evergreen, broadleaf
2. $\geq 0\,°C$ and $<10\,°C$	Evergreen, chilling resistant
3. $\geq -15\,°C$ and $<0\,°C$	Evergreen, frost resistant, broadleaf
4. $\geq -40\,°C$ and $< -15\,°C$	Deciduous, broadleaf
5. $< -40\,°C$	Evergreen or deciduous, needle leaf
6. Month degrees ≤ 50	Tundra

iological processes to control the distribution of vegetation. Lack of agreement between predictions of vegetation, based on climatic observations and ecophysiological responses, and on observations of vegetation will be cause to reject the hypothesis. However the final proof of the hypothesis must await experimental analysis.

A set of climatic records from the World's meteorological stations (Müller, 1982) have been used as the climatic data base for these predictions. Cli-matic records were stored in the University computer (IBM 3081) and used to predict the vegetation at the sites of the meteorological stations on the basis of the criteria presented in Table 1. The global pattern of vegetation on this basis was then mapped using a software package developed by Dr A. G. Smith (Department of Earth Sciences, University of Cambridge) and is presented in Fig. 2. The predicted distribution of vegetation may be compared with an equivalent map of observations (Fig. 3), based on that published in Polunin (1960).

A comparison between the two maps shows a number of areas of similarity, such as the zones of tundra, conifers and in some cases for deciduous and evergreen forests. However a number of areas do not coincide, particularly in dry regions. It is clear that the predictions should be modified to include the hydrological balance.

Predictions from records of temperature and the hydrological balance

Woodward (1986) has shown that the leaf mass

PREDICTIONS FROM OBSERVATIONS OF TEMPERATURE

Symbols
▲ Tundra
■ Conifer
▼ Deciduous Broadleaf
◆ Evergreen Broadleaf
● Evergreen (Broadleaf) Frost Tolerant
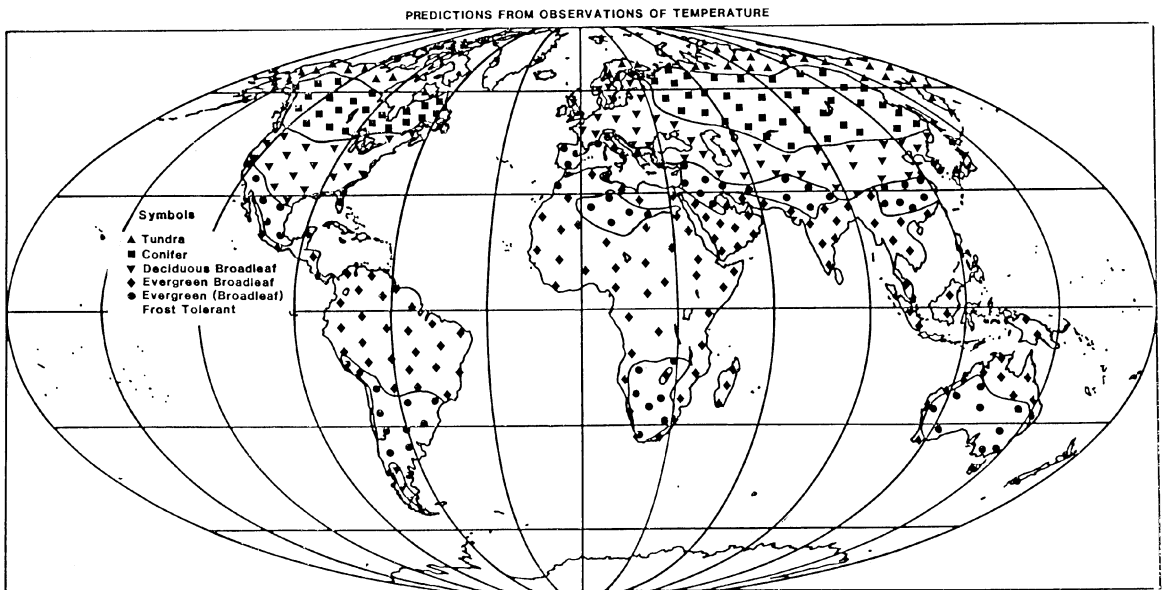
Fig. 2. Predictions of global vegetation on the basis of temperature. (▲), tundra; (■), conifer; (▼), deciduous broadleaf; (◆), ever-green broadleaf; (●), evergreen broadleaf, frost tolerant.
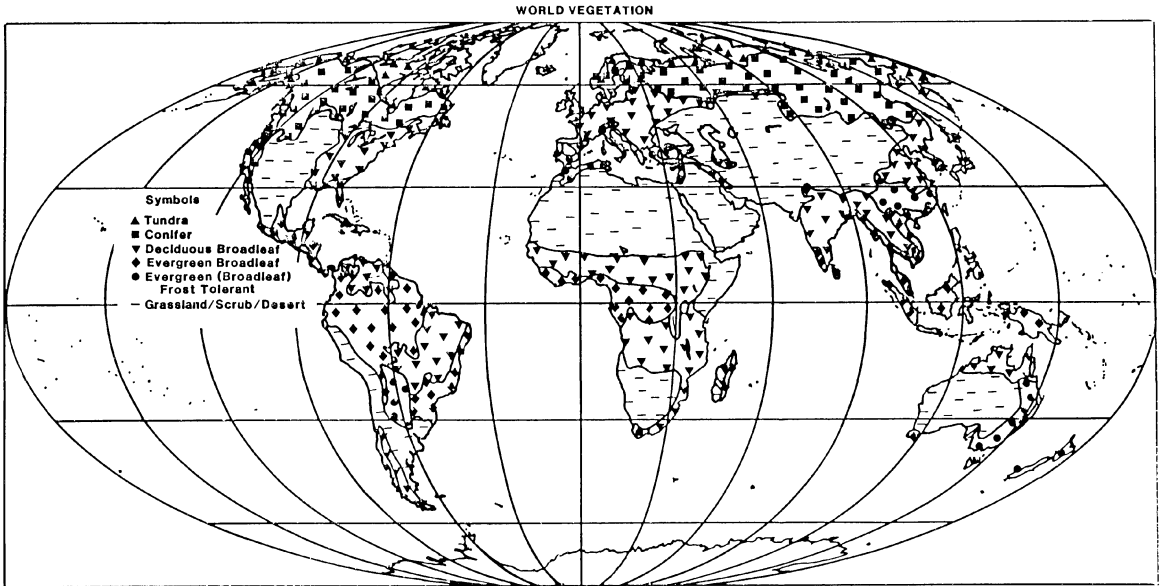
WORLD VEGETATION

Fig. 3. Map of the Vegetation of the World from Polunin (1960). Symbols as for Fig. 2 and (−), grassland/scrub/desert.

PREDICTIONS FROM OBSERVATIONS ON TEMPERATURE, PRECIPITATION AND ANNUAL WATER BALANCE
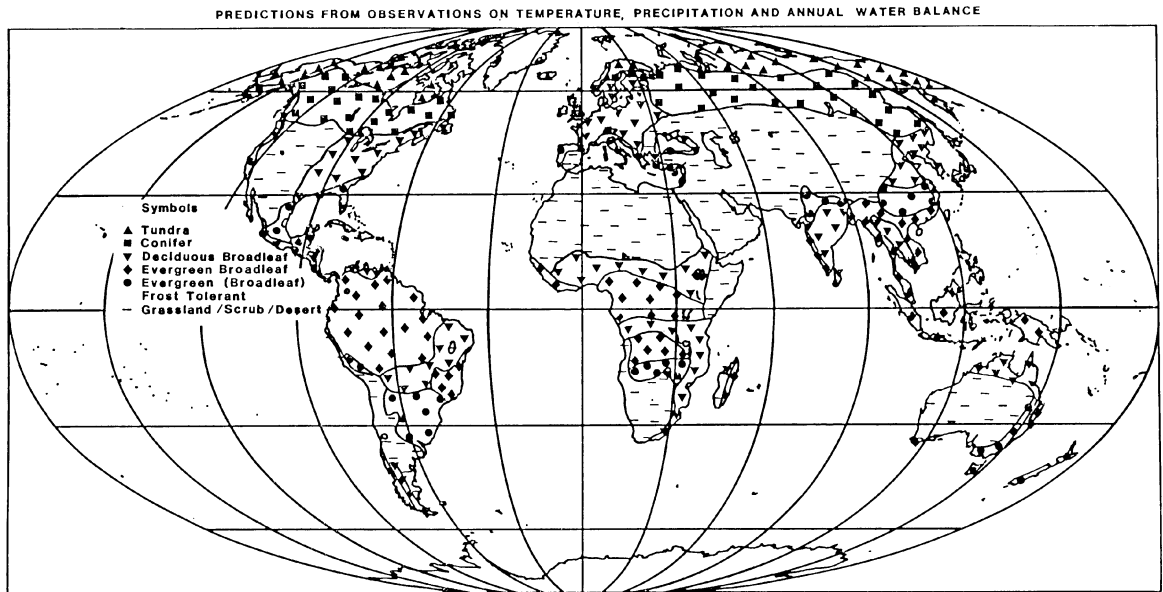
Fig. 4. Predictions of global vegetation on the basis of temperature, precipitation and the annual water balance (precipitation − evaporation). Symbols as for Figs. 2 and 3.

(leaf area index) of native vegetation can be predicted quite closely using the Monteith modified Penman equation, observations on stomatal resistances and canopy leaf area densities and measurements of local climate. These predictions and their data bases are too extensive for consideration here. However they can be reduced considerably, without a large loss in accuracy, to a short table of the minimum annual precipitation for a particular vegetation type, and a simple measure of the annual hydrological budget, estimated as precipitation minus potential evapotranspiration (Müller, 1982) and termed the water balance (Table 2). These limits are similar to those which may be calculated from observations of the hydrological balance of forests (Rutter, 1968; Doley, 1981) and have been used to predict vegetation types on the basis of site records of annual precipitation and calculations of evaporation (Müller, 1982).

Annual precipitation should equal or exceed 600 mm for forests in all areas, except where the minimum temperature falls below $-40\,°C$, where the limit is 400 mm (Woodward, 1986). If potential evaporation exceeds annual precipitation, with a negative water balance, even though the threshold for precipitation is exceeded, then the prediction changes to a drought-deciduous vegetation in regions that are not subjected to frost. In the colder

regions the life form of the vegetation is predicted to change to grassland or scrub when the precipitation threshold is not reached, irrespective of the water balance, although it generally emerges that potential evaporation exceeds precipitation in these areas. The global predictions on the basis of these limits are shown on Fig. 4.

A comparison between the vegetation map based on predictions incorporating temperature, precipitation and the water balance and the map of observed vegetation types (Fig. 3) shows a considerable degree of correspondence. The large expanses of shrub or grassland are now well defined, as are the forest transitions.

The importance of population dynamics

The distribution of life forms has been defined in ecophysiological terms, implying a mechanistic base to the predictions. However it is also possible that the responses which have been considered are closely correlated with different but more critical limiting factors. The validity of the present predictions can only be established by experimental analyses on the effects of climate on vegetation.

Unfortunately, vegetation is an unwieldy tool for such analyses and consequently efforts in this direction have been concentrated on species, or subdivisions thereof. Such studies are valuable but are distant from a study of the integrated effects of climate on all of the species in a particular tract of vegetation. If, for example, there is a general global cooling of climate, then species at the limits of their distribution in their associated vegetation may be subjected to extreme and lethal changes in climate. If in addition the species differ slightly in their lethal thresholds then gaps occur in the vegetation where the most sensitive species have died. Species which are more resistant to climatic extremes will then fill this gap. The end result of this infill may either be the retention of the existing vegetation, but with a change in the composition of species with the same life form, or a change to a different type of vegetation and range of species.

In these terms the climatic control of vegetation type is seen as a process which is mediated through

Table 2. Further definition of vegetation types on the basis of minimum temperature, annual precipitation and water balance (type 1 positive, all other types negative).

Minimum temperature range	Annual precipitation mm	Probable vegetation type
1. $\geq 0\,°C$	≥ 600	Evergreen, broadleaf
2. $\geq 0\,°C$	≥ 600	Deciduous, broadleaf
3. $\geq 0\,°C$	< 600	Grassland/shrub
4. $\geq -15\,°C, < 0\,°C$	≥ 600	Evergreen, frost resistant, broadleaf
5. $\geq -15\,°C, < 0\,°C$	< 600	Grassland/shrub
6. $\geq -40\,°C, < -15\,°C$	≥ 600	Deciduous, broadleaf
7. $\geq -40\,°C, < -15\,°C$	< 600	Grassland/shrub
8. $< -40\,°C$	≥ 400	Evergreen or deciduous, needle leaf
9. $< -40\,°C$	< 400	Grassland/shrub

such population processes as gap creation and fill, although resulting from a mechanism which may be understood in ecophysiological terms and which has predictive efficiency in defining global types of vegetation. An ecological study based solely on the sensitive species would therefore explain the occurrence of a gap in the vegetation but could not predict the effect of climate on the vegetation as a whole. In addition the climatic extreme may only be effective at one particular stage of the life cycle, so if this occurred at the seedling stage, effectively preventing regeneration but with a mature phase which was resistant, then the change in the vegetation would be slow and dependent on plant longevity. It is clearly important, therefore, to study the impact of climate on all stages of the life cycle, in order to provide a more complete understanding of the control of plant distribution. Narrowing down the spatial scale of study in this way may also serve to explain why the dominant vegetation may be adequately predicted, as shown in Fig. 4, but without explanation of the lack of dominance of vegetation from a more extreme, e.g. colder, climate.

Life cycle dynamics

Ecophysiological studies, such as those which investigate the effects of temperature or irradiance on plants, are usually confined to experiments on individual plants. This restriction is because of the logistics of experimental design. Though of great value, such experiments may have little relevance at the population level. Modelling life cycle dynamics can help in dissecting this problem.

Figure 5 presents a diagrammatic version of the life cycle of a perennial plant as presented in Woodward & Jones (1984). The probability that an individual will pass from one stage of the life cycle to the next is indicated as p_i, where i is the number of the stage in the life cycle, starting from the seed bank and ending at seed production. All stages, except seed production, are transition probabilities. Seed production is the product of the probability of setting seed and the seed number produced by the plant.
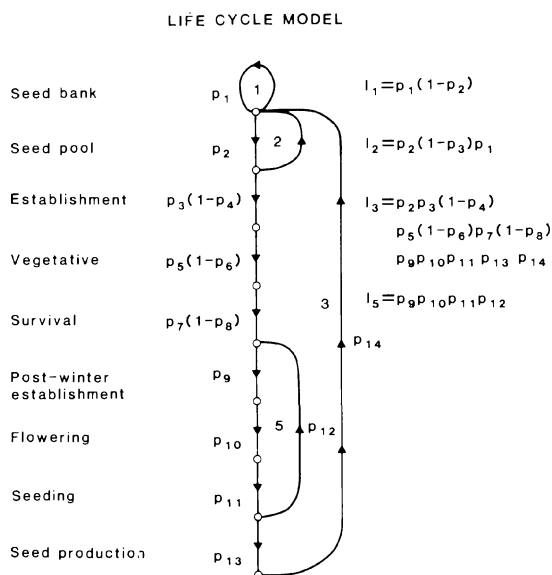
LIFE CYCLE MODEL

Fig. 5. Life cycle model. p_i, transition probability of state i; l_j, subloop j of the life cycle, defined in terms of the product of the appropriate transition probabilities.

The life cycle is also defined in terms of a number of loops l_j, where j is the loop defined in terms of the probabilities which it encompasses. So, for example, loop 1 predicts the population of seeds which remain in the seed bank, a population which is controlled by seed longevity (p_1) and the emergence of the seed from the seed bank (p_2). Individuals in a population of a perennial and polycarpic species will survive through all of the loops. Loop 5 describes the polycarpic behaviour of sexually mature individuals and may be included with loop 3, which in isolation describes a monocarpic life cycle, to define a fourth loop, loop 4, for perennial and polycarpic species.

Woodward & Jones (1984) attempted to explain the altitudinal distribution of a number of species by observing, and making future predictions of, the population dynamics of the species at different altitudes. Field experiments started with the creation of gaps in the native vegetation, followed by seeding at one species per gap, at a range of densities. This work also attempted to account for the same feature posed by mixtures of different physiogno-

mies, which is the natural absence of species from cool climates (mountain species) in warmer climates (lowland sites), even though no aspect of the lowland climate is lethal (Woodward, 1986).

The probabilities of surviving each stage of the life cycle (Fig. 5) were recorded, including sensitivities to population density and to recorded variations in climate and were subsequently used to predict population growth using Leslie matrices. The data presented in Woodward & Jones (1984) have been re-evaluated to predict the annual rates of population increase in the 5 loops shown in Fig. 5.

Five examples of species at different altitudes and different initial densities are presented (Table 3). For *Eupatorium cannabinum*, at an altitude of 610 m, plant survival does not extend beyond loop 2 because of 100% mortality during winter frosts. This fundamental property of the species is not influenced by population density and defines the absolute distributional limit of the species, in the same manner as described for various vegetation types. At a lower altitude of 360 m the species is able to complete its life cycle, although only when the initial density of the species is low. At high density, winter mortality is again critical at the end of the first growing season, implying the plants must be a critical size in order to survive low temperatures.

Table 3. Interspecific variations in life cycle probabilities for 1) *Eupatorium cannabinum* at 610 m, 2) *E. cannabinum* at 360 m, 3) *Oxyria digyna* at 610 m, 4) *Potentilla reptans* at 610 m, and 5) *Saxifraga hypnoides* at 610 m.

Plant density m^{-2}	Loop Number					Treatment
	1	2	3	4	5	
157	0.10	0.26	0	0	0	1
47 100	0.10	0.26	0	0	0	
157	0.10	0.28	4.06	2.26	0.63	2
47 100	0.10	0.19	0	0	0	
157	0.10	0.01	1.95	1.25	0.49	3
47 100	0.10	0.17	0.56	0.62	0.36	
157	0.35	0.07	0	0	0	4
47 100	0.35	0.12	0.16	0.22	0.26	
157	0.25	0.14	3.44	0	0.23	5
47 100	0.25	0.14	1.61	0.94	0.18	

The third case, for the Arctic-Alpine species of *Oxyria digyna*, is frequently observed (e.g., Harper, 1977) and shows that population growth decreases with density. In this case it is also predicted that the seed bank (loop 2) will increase with density because of a reduction in seed germination caused by interference from mature plants.

The reverse effect of the first case is shown for *Potentilla reptans* at an altitude of 610 m. In this case the species is predicted to become extinct when growing at low densities. However, unlike *E. cannabinum*, an increase in population density will lead to longer term survival, with all loops showing positive coefficients. For *P. reptans* winter survival is not the sole cause of the demise of the species. A number of features of the life cycle, such as a low probability of seed emergence from the seed bank, low winter survival and low seed production, are all effective in reducing population growth, features which may evidently be reversed by increasing population density.

A similar response is also shown for the fifth case of *Saxifraga hypnoides* at 610 m. The predictions indicate that when a population establishes from a low, initial density the life cycle is short lived with the probability of individuals passing through loop 4 falling to zero. Increasing the initial population density reduces the annual increase through loop 3 but also leads to an increase in the longer lived component of the population, which survives through loop 4.

These five cases provide clear predictions of the density-dependent, ecophysiological responses of a number of species. It may be concluded that studies on species within their natural ranges must seek to explain the modifying effects of population density and dynamics on ecophysiological responses to climate. This may be unnecessary in situations at, or beyond, the climatic limit of the species.

References

Addicott, F. T. & Lyons, J. L., 1973. Physiological ecology of abscission. In: T. T. Kozlowski (ed.), Shedding of plant parts, pp. 85–124. Academic Press, New York.

Axelrod, D. I., 1966. Origin of deciduous and evergreen habits in temperate forests. Evolution 20: 1–15.

Bliss, L. C., 1956. A comparison of plant development in

microenvironments of Arctic and Alpine tundra. Ecol. Monogr. 26: 303–337.

Box, E. O., 1981. Macroclimate and plant forms: An introduction to predictive modeling in phytogeography. Junk, The Hague.

Callaghan, T. V. & Collins, N. J., 1981. Life cycles, population dynamics and the growth of tundra plants. In: L. C. Bliss, O. W. Heal & J. J. Moore (eds), Tundra ecosystems: a comparative analysis, pp. 257–284. Cambridge University Press, Cambridge.

Doley, D., 1981. Tropical and subtropical forests and woodlands. In: T. T. Kozlowski (ed.), Water deficits and plant growth, Vol. VI pp. 209–323. Academic Press, New York.

Grier, C. C. & Running, S. W., 1977. Leaf area of mature northwestern coniferous forests: relation to site water balance. Ecology 58: 893–899.

Harper, J. L., 1977. Population biology of plants. Academic Press, London.

Holdridge, L. R., 1967. Life zone ecology (rev. ed.). Tropical Science Center, San Jose, Costa Rica.

Larcher, W., 1981. Low temperature effects on Mediterranean sclerophylls: an unconventional viewpoint. In: N. S. Margaris & H. A. Mooney (eds), Components of productivity of mediterranean climate regions – basic and applied aspects, pp. 259–266.

Levitt, J., 1980. Responses of plants to environmental stresses. Vol. 1. Chilling, freezing and high temperature stresses. 2nd ed. Springer, Berlin.

MacArthur, R. H., 1972. Geographical ecology: Patterns in the distribution of species. Harper & Row, New York.

McMichael, B. L., Jordan, W. R. & Powell, R. D., 1973. Abscission processes in cotton: induction by plant water deficit. Agron. J. 65: 202–204.

Merrill, E. D., 1945. Plant life of the Pacific world. Macmillan, New York.

Monteith, J. L., 1965. Evaporation and environment. In: G. E. Fogg (ed.), The state and movement of water in living organisms. Symposium of the Society for Experimental Biology, 19, pp. 205–234. Cambridge University Press, Cambridge.

Müller, M. J., 1982. Selected climatic data for a global set of standard stations for vegetation science. Junk, The Hague.

Polunin, N., 1960. Introduction to plant geography and some related sciences. Longmans, London.

Rutter, A. J., 1968. Water consumption by forests. In: T. T. Kozlowski (ed.), Water deficits and plant growth Vol. II. pp. 23–84. Academic Press, New York.

Sakai, A., 1978. Freezing tolerance of evergreen and deciduous broadleaved trees in Japan with reference to tree regions, Low Temp. Sci., Ser. B. 36: 1–19.

Sakai, A., 1979. Freezing avoidance mechanism of primordial shoots of conifer buds. Pl. Cell Physiol. 20: 1381–1390.

Sakai, A. & Wardle, P., 1978. Freezing resistance of New Zealand trees and shrubs. N. Z. J. Ecol. 1: 51–61.

Sakai, A. & Weiser, C. J., 1973. Freezing resistance of trees in North America with reference to tree regions. Ecology 54: 118–126.

Satoo, T., 1983. Temperate broad-leaved evergreen forests of Japan. In: J. D. Ovington (ed.), Ecosystems of the world. 10. Temperate broad-leaved evergreen forests pp. 169–189. Elsevier, Amsterdam.

Schnelle, F., Baumgartner, A. & Freitag, E., 1984. Arboreta Phaenologica 28.

Schulze, E.-D., 1982. Plant life forms and their carbon, water and nutrient relations. In: O. L. Lange, P. S. Nobel, C. B. Osmond & H. Ziegler (eds), Encyclopedia of Plant Physiology, Vol. 12B, pp. 616–676.

Senser, M. & Beck, E., 1977. On the mechanisms of frost injury and frost hardening of spruce chloroplasts. Planta 137: 195–201.

Spurr, S. H. & Barnes, B. V., 1980. Forest ecology. 3rd ed. Wiley, New York.

Tieszen, L. I., Lewis, M. C., Miller, P. C., Mayo, J., Chapin III, F. S. & Oechel, W., 1981. An analysis of processes of primary production in tundra growth forms. In: L. C. Bliss, O. W. Heal & J. J. Moore (eds), Tundra ecosystems: a comparative analysis, pp. 285–356. Cambridge University Press, Cambridge.

Von Humboldt, A. & Bonpland. A., 1805. Essai sur la géographie des plantes; accompagné d'un tableau physique des régions equinoxales. Paris.

Waring, R. H., Emmingham, W. H., Gholz, H. L. & Grier, C. C., 1978. Variation in maximum leaf area of coniferous forests in Oregon and its ecological significance. For. Sci. 24: 131–140.

Woodward, F. I., 1986. Climate and plant distribution. Cambridge University Press, Cambridge.

Woodward, F. I. & Jones, N., 1984. Growth studies of selected plant species with well-defined European distributions. I. Field observations and computer simulations on plant life cycles at two altitudes. J. Ecol. 72: 1019–1030.

Alternate plant life history strategies and coexistence in randomly varying environments

Stephen Ellner* **
Biomathematics Graduate Program, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, USA

Abstract

Environmental fluctuations can in theory allow the coexistence of ecologically similar species by 'time-sharing' a niche, as envisioned by Hutchinson. The evolution of this situation is studied in a competition model, using as an example the evolution of seed germination strategies. Coexistence occurs via the evolution of 'low-risk' and 'high-risk' strategies for dealing with the variability by different species. Coexistence is promoted by intermediate levels of variability or disturbance, and by a trade-off between seed yield and seed survivorship. These results may be applicable also to other low vs. high risk life history options in unpredictably varying environments, such as: stress resistance vs. potentially rapid growth, high adult survivorship vs. high reproductive output. The model's predictions differ from those obtained without consideration of life history evolution in response to environmental variability, and are consistent with some recent studies of plant strategies in intermittently stressed communities.

Introduction

How might environmental variability act to promote the long-term coexistence of ecologically similar species? Essentially two different theoretical answers to this question have been offered:

1. It does not. In fact, even very small amounts of variability set drastic limits to the similarity of coexisting species. This was first asserted by Klopfer & MacArthur (1961), to explain the high diversity of tropical bird communities. In their argument, environmental variability is viewed as 'noise', accidental variations in resource abundances or other environmental factors which may lead to 'chance' extinctions of species. Later May & MacArthur (1972) and May (1973, 1974) claimed to have proved mathematically that small amounts of noise have drastic effects on coexistence in models of community dynamics. However, their analysis was incorrect and a correct analysis shows much smaller effects of noise (Turelli, 1978, 1981) in Lotka-Volterra type models.

2. Environmental variability may allow coexistence by 'time-sharing' of a niche. Hutchinson (1961) proposed this idea as a possible explanation for high species diversity in plankton communities, less than 100 pages away from Klopfer & MacArthur (1961). Hutchinson's hypothesis was that competitive exclusion could be postponed indefinitely,

TIME-SHARING A NICHE
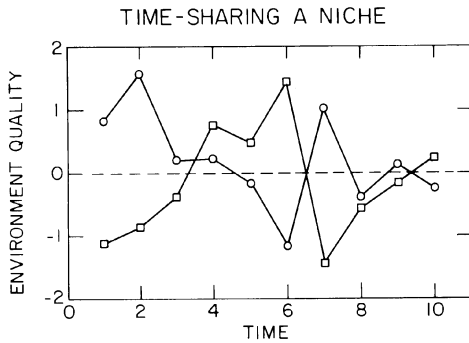
Fig. 1. Time-sharing an ecological niche in a randomly varying environment. The species have different requirements for optimal performance, with the result that above average years for one species (boxes) tend to be below average for the other (circles) and vice versa.

if conditions regularly changed so that different species are favored at different times. Figure 1 shows a theoretical example of this sort of environmental variability. The variation is highly stochastic, but a good time for one species tends to be a bad time for the other, so each population may increase during favorable times. Hutchinson argued verbally that the species might then coexist indefinitely. His argument is now supported by rigorous mathematical analyses of a variety of models for populations in fluctuating environments, both deterministic (e.g., De Mottoni & Schiaffino, 1981; Cushing, 1980, 1984) and stochastic (e.g., Slatkin, 1978; Chesson & Warner, 1981; Chesson, 1982, 1986; Ellner, 1984; Shmida & Ellner, 1984; Ågren & Fagerström, 1984). Coexistence occurs in these models by what Chesson has called the 'storage effect': some stage in the life history 'stores up' the gains made during favorable periods and has relatively constant survivorship through unfavorable periods. The storage stage in plants may be, e.g., a seed bank, or long-lived adults which survive stressful periods by inactivity (Evenari *et al.,* 1971). Recently Tilman (1982) and Abrams (1984) have considered models for resource competition with fluctuating resource supply rates, which are very similar in spirit to Hutchinson's verbal model.

Unequivocal empirical support for the 'time-sharing' hypothesis is still lacking. The reasons for

this, it seems, are that until recently (1) simpler theories of community structure in constant environments dominated theoretical and empirical work, and (2) clear and general principles, that could be tested in practical field-studies, had not yet emerged from fluctuating environment theories. However, environmental fluctuations of the sort needed for the 'time-sharing' hypothesis have now been amply documented (e.g., Grubb, 1977; Shmida & Ellner, 1984 in plant communities).

This paper is another theoretical study, but its aim is practical. I want to suggest to empirical ecologists a class of documented life-history phenomena which seem (in theory) to be likely candidates for coexistence by time-sharing. The phenomena were found by posing the question: how might evolution in a fluctuating environment lead to 'time-sharing' of a niche? In the situation of Fig. 1, the answer is clear. Within the normal range of environmental conditions (temperature, nutrient supply rates, etc.), different species are adapted to perform optimally under different sets of conditions. This is analogous to the classical idea (Whittaker, 1975) of niche differentiation along a habitat gradient such as elevation or moisture.

Suppose, however, that the variability takes the form of an unpredictable, intermittent stress or disturbance — such as drought, flooding, herbivore attack, or wave action — which reduces the performance of all species. In such cases, natural selection may lead to the evolution of alternative 'strategies' for coping with the variability. The 'high-risk' strategist is superior in the absence of stress. It gambles on (unpredictable) future conditions being favorable, and reaps large benefits if they are. The 'low-risk' strategist is more cautious and better equipped to deal with stress, but is inferior to the 'high-risk' strategist when conditions are favorable. The result, diagrammed in Fig. 2, is an alternation of relative advantage between low-risk and high-risk strategists, which may lead to coexistence of the sort envisioned by Hutchinson. Some possible examples of low-risk vs. high-risk alternatives are listed in Table 1.

The theoretical development here is phrased in terms of seed germination, for specificity and to make use of previous theory. Species differ in their
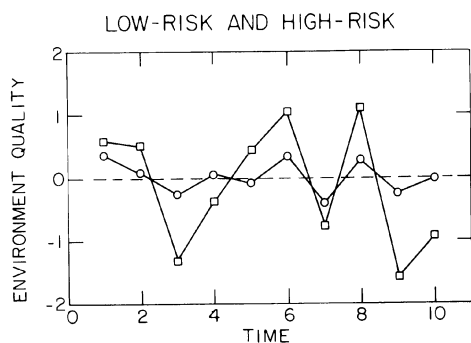
LOW-RISK AND HIGH-RISK

Fig. 2. Low risk and high risk strategists in a randomly varying environment. The high risk species (boxes) is superior under favorable conditions, while the low risk species (circles) is less affected by unfavorable conditions.

Table 1. Possible examples of low-risk versus high-risk strategies in intermittently stressful habitats.

Effect of stress	Low risk	High risk
Adult mortality	Stress resistance ('S' strategy of Grime, 1979)	Potentially rapid growth relative to competitors ('C' or 'R' strategies).
Juvenile or seedling mortality, especially of smaller individuals	Few, large seeds	Many small seeds
	Vegetative reproduction	Sexual reproduction
	Low reproductive effort, high adult survivorship	High reproductive effort, low adult survivorship
	Cautious, delayed germination	Easy, rapid germination

annual germination fraction (*G*) and annual survivorship of dormant seeds (*s*). Germinated seeds face an unpredictable, randomly varying environment, so the low-risk and high-risk strategies are low vs. high seed germination fractions.

Differences in germination can lead to coexistence in stochastic models of competing species (Ellner, 1984); the question here is whether such differences might arise and persist, under natural selection. These evolutionary considerations distinguish the present paper from all previous analyses

of coexistence in randomly varying environments. It introduces a new constraint, and leads to new and different predictions, some of which appear to be consistent with recent studies of plant strategies in intermittently stressed communities.

Because of space limitations, all formal proofs are omitted (the methods of analysis are outlined in the Appendix), and I have not attempted a complete review of the relevant empirical literature.

The model

The model describes the population dynamics of an annual plant species, in a spatially uniform habitat with random environmental fluctuations. The biological rationale for the model is more fully developed in Ellner (1985).

The density (number/area) of live seeds of species *i* in year *t*, denoted $X_i(t)$, is censused immediately before the start of the growing season. A fraction G_i of those seeds germinate. The seedlings compete among themselves and with other plants in the habitat, set seed, and then die before the start of the next growing season. The ungerminated seeds are assumed to have annual survivorship $s_i (0 < s_i < 1)$ which is constant (i.e., independent of seed age, density, etc.). The population dynamics are then

$$X_i(t + 1) = X_i(t) [G_i Y_i(t) + (1 - G_i)s_i] \qquad (1)$$

where $Y_i(t)$ is the per capita seed yield (seeds produced per germinated seed) of species *i* in year *t*.

Cohen (1966, 1968) studied model (1) under the assumptions that the $Y_i(t)$, $t = 0, 1, \ldots$, are independent, identically distributed random variables. Here, seed yield is assumed to follow the 'reciprocal yield law' $Y = K/(C + x)$, where *K* and *C* are nonnegative constants and *x* is the seedling density. Environmental variability is modelled by allowing *K* to vary randomly between years. The per capita seed yield is then given by

$$Y_i(t) = K_i(t)/(C_i + \sum_{j=1}^{n} \alpha_{ji} G_j X_j(t)). \qquad (2)$$

202

The competition coefficients α_{ji} measure the competitive impact of a species-j seedling relative to the impact of a species-i seedling. The use of a single coefficient to describe interspecific competition obscures effects of spatial patterning which may strongly influence the dynamics (Shmida & Ellner, 1984), but for our present purposes it is a useful metaphor for the average intensity of interspecific competition.

It is convenient to re-write the model in terms of N, the density of seedlings, rather than X, the density of seeds. These are related by $N_i(t) = G_i X_i(t)$. Combining (1) and (2), the model in terms of N_i is:

$$N_i(t+1)=N_i(t)\left[\frac{G_i K_i(t)}{C+\sum\limits_{j=1}^{n}\alpha_{ji}N_j(t)}+(1-G_i)s_i\right]. \quad (3)$$

We will call (3) the RYL model, for 'reciprocal yield law.' With no loss of generality, units of population density are chosen for each species so that $\alpha_{ii} = 1$ for all i.

Germination theory

A theory of seed germination for annual plants has been developed by Bulmer (1984) and Ellner (1985), based on the RYL model (3) and the concept of an evolutionary stable strategy (ESS). The ESS concept refers to a single population, composed of various subpopulations with different 'strategies'. An ESS is defined by the property that any alternative strategy, if it is introduced into the population as a rare mutant, will decrease in frequency to eventual extinction.

In our case, the alternative strategies are different values of the germination fraction $G_i(G_1 \neq G_2)$, but $\alpha_{12} = \alpha_{21} = 1$ and $s_1 = s_2$ because all are the same species. G_1 is an ESS if, for any $G_2 \neq G_1$, $X_2(t) \to 0$ whenever a few X_2's are added to a population consisting mostly of X_1's. Clearly, the ESS is a function of s, C, and the distribution of $K(t)$. The ESS is presumed to represent the endproduct of natural selection: mutations occur and advantageous alleles spread through the population, until
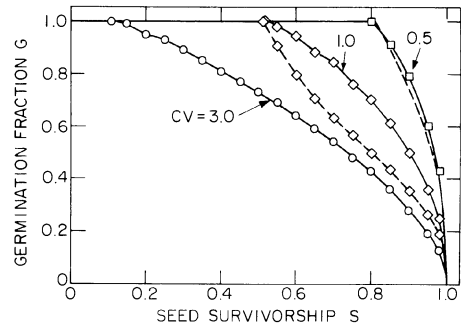
Fig. 3. Evolutionarily stable (ESS) annual germination fraction in the RYL model, plotted as a function of the seed survivorship s. Parameter values $C = .05$, $K(t)$ lognormally distributed with coefficient of variation $CV = .5$, 1.0, or 3.0. The dashed lines are the 'small-fluctuations' approximate solution.

finally the ESS is reached and new mutants are selected against. See Levin (1978), Maynard Smith (1982) and Ellner (1985) for the assumptions and limitations of ESS theory in general and as applied to seed germination.

Typical results of ESS germination theory for the RYL model are illustrated in Figs. 3 and 4. In Fig. 3, $K(t)$ is lognormally distributed and the ESS germination fraction is plotted as a function of the seed survivorship s. The ESS germination fraction decreases with increasing seed survivorship, and with increased environmental variability as measured by the coefficient of variation of $K(t)$ (the coefficient of variation is CV = standard devia-
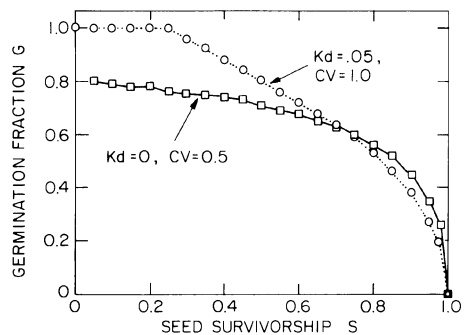

Fig. 4. As in Fig. 3, but with a 20% probability of disturbance. $K(t) = K_d$ in disturbance years, lognormally distributed in non-disturbance years; $C = .2$.

tion/mean). The ESS germination fraction ranges from 1 at low seed survivorship to zero in the limit of 100% seed survivorship, and near $s = 1$ the ESS germination fraction changes rapidly with small changes in seed survivorship. The unrealistic point $s = 1$ (100% eternal survival of dormant seeds) is not included in the model; the figures are showing that the ESS G becomes arbitrarily small at values of s near 1.

These numerical results are bolstered by an analytical approximation (dashed lines) which is asymptotically accurate for 'small' variability and/or s near 1 (Ellner, 1985), and predicts similar behavior for any distribution of $K(t)$. The same approximation is used below to analyze coexistence; it is quite accurate for $CV \leq .5$ and does not begin to go awry until $CV > .8$ (which is a fairly large variability) for lognormally distributed $K(t)$. Another useful approximation is setting $C = 0$. If units are chosen so that the mode of $K(t)$ is at 1, as we do in all our calculations, $1/C$ is the modal number of seeds per adult in the absence of competition. Thus C near 0 is the rule for most annuals (Harper, 1977) and setting $C = 0$ gives a good approximation.

In Fig. 4, catastrophic 'disturbance' years are added to the environmental variability: one year in 5 on average, $K(t)$ takes a low value K_d. If disturbances are absolute ($K_d = 0$), some dormancy is favored regardless of the mortality of dormant seeds. Less severe disturbances ($K_d > 0$) have effects similar to an increase in the overall variability (compare Fig. 4, CV = 1 with Fig. 3, CV = 3): dormancy is favored over a broader range of seed survivorships, and the ESS germination fraction is smaller.

The predictions of the RYL model differ markedly from Cohen's (1966, 1968) results on 'optimal' germination in a density-independent model (essentially, model (3) with all $\alpha_{ij} = 0$: yield is not reduced by competition with other plants). Cohen's model generally predicts much higher germination fractions and much smaller changes in the germination fraction as seed survivorship varies, and there are also many qualitative differences with the RYL model's predictions (Bullmer, 1984; Ellner, 1985).

Coexistence of low-risk and high-risk germination strategies

The ESS germination theory leads also to predictions about coexistence of species, and about the role of environmental variability in their coexistence. Consider two species in model (3), with equal competition coefficients $\alpha_{12} = \alpha_{21} = \alpha$ (symmetric competition) but $s_1 \neq s_2$, $G_1 \neq G_2$ (different species). If we ignore environmental variability, and analyze the model with $K(t)$ set equal to the mean or mode of its distribution, there is a critical value of α, α_c^* (for 'constant environment'), such that coexistence occurs for $\alpha < \alpha_c^*$ but not for $\alpha > \alpha_c^*$. However, the true critical value of α is a different number α_v^* which takes account of the effects of environmental variability. α_v^* is determined by the requirement that each species increase in numbers when rare, so neither goes extinct (see the Appendix).

The analysis of coexistence is as follows. For given seed survivorships s_1 and s_2, and a given pattern of environmental variability, the species' germination fractions G_1, G_2 are determined by the ESS theory. The critical values α_c^* and α_v^* are then calculated. If $\alpha_v^* > \alpha_c^*$ coexistence of the species is promoted by environmental variability. The results are displayed by graphing the coexistence set CS, which consists of all (s_1, s_2) values such that $\alpha_v^* > \alpha_c^*$.

This definition of the coexistence set seems to be appropriate for interpreting observations on real communities. The fieldworker studies species which have evolved in response to their environment, including the variability. Environmental variability plays a role in their coexistence, if the elimination of variability would result in loss of species. The coexistence set represents the set of species pairs whose coexistence may be dependent on variability. Conditions which produce a large coexistence set will be said to 'promote coexistence'.

So far, the species in the model have evolved 'in isolation': the ESS germination fractions are based only on intraspecific competition without any effect of the other species. As an opposite extreme (with reality hopefully somewhere in between), I

also considered species which were perfectly coevolved to each other. In simulations, the species were started at their ESS germination fractions. The germination fractions were adjusted every 500 generations until the species reached 'coevolutionarily stable strategies' at which neither species could be invaded, and the coexistence set was determined for the resulting germination fractions.

The model was analyzed using the approximations described ('small' variability and $C = 0$), and the approximations were checked by numerical calculations at higher variability (up to $CV = 3$) and $C \leq .2$. The main results are as follows:

1. Coexistence is most promoted at intermediate levels of variability. If the variability consists of occasional disturbances, this means intermediate frequency and/or intermediate severity of disturbances.

2. Coexistence is promoted by a trade-off between seed survivorship and seed yield. Such a trade-off is intuitively plausible. If high-survivorship seeds are more 'costly' to produce than low-survivorship seeds, the total number of seeds produced by a parent will decrease with increasing seed survivorship.

3. Coevolution does not demote coexistence, but may not promote it much. Coevolution was studied only by simulation, so the results are not unequivocal, but it seems that coevolution has very little effect on germination fractions near the borders of the coexistence set. Coevolution always expanded the coexistence set, usually by only a little.

These results are illustrated in Figs. 5–7, for lognormally distributed $K(t)$ possibly with catastrophes. The coexistence set is symmetric about the line $y = x$ (i.e., $(s_1,s_2) \in$ CS if and only if $(s_2,s_1) \in$ CS) because of the 'symmetric competition' assumption, so only the portion of the coexistence set where $s_1 \leq s_2$ is graphed (below the dashed line $y = x$ in all figures).

Figure 5 exemplifies the effects of increasing variability, in this case an increasing severity of disturbances. One year in four $K(t)$ takes the value K_d, and otherwise $K(t)$ is lognormally distributed with $CV = .5$. As K_d decreases, representing worse and worse disturbances, the coexistence set moves from high to low values of (s_1,s_2) and eventually
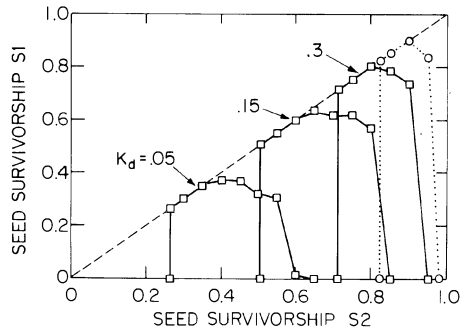
Fig. 5. Coexistence sets for the RYL model with a 25% probability of disturbance, $K(t) = K_d$ in disturbance years, lognormally distributed with $CV = .5$ in non-disturbance years; $C = .15$. The dotted curve outlines the coexistence set when there are no disturbances.
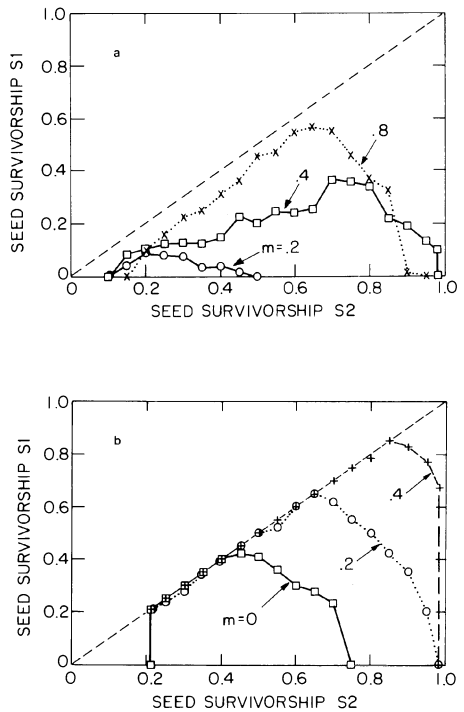
Fig. 6. Coexistence sets for the RYL model with a trade-off between seed yield and seed survivorship. a: $K(t)$ as in Fig. 5, but with $K_d = .01$; $b = 1$ and $m = .2, .4, .8$. b: $K(t)$ lognormally distributed with $CV = 2.0$. $m = 0$ (solid), $m = .2$, $b = 1$ (dots), $m = .4$, $b = 2$ (dashed).
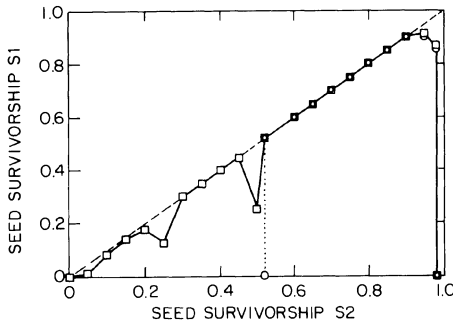
Fig. 7. Coexistence sets with coevolution (solid) and without coevolution (dots). Lognormally distributed $K(t)$ with $CV = 1.0$; $C = 0.05$, $m = .25$, $b = 3$. The coexistence set with coevolution typically is irregularly shaped as in this example.

gets squeezed into the lower left-hand corner. An increased frequency of disturbance (with K_d unchanged) and/or increased variability in non-catastrophe years have very similar effects.

The 'migration' of the coexistence set is due to the increased value of dormancy and of seed survivorship in more variable or more frequently stressed habitats. At low variability, the optimal strategy is 100% germination for all species, so coexistence of alternate strategies is impossible, unless s_2 is very near 1 (see Fig. 3). The coexistence set is then limited to small bands near $s_1 = 1$ or $s_2 = 1$. As the variability is increased, dormancy is favored over a large range of seed survivorships, so the coexistence set can expand. But with higher variability or more frequent stress, much dormancy is favored and so species with high seed survivorship have a tremendous advantage. Eventually, the coexistence set is limited to a small region near $s_1 = s_2 = 0$. If one species has a higher seed survivorship, that is such an overwhelming advantage that any species with lower seed survivorship is driven to extinction.

This observation also explains Result 2, that coexistence is promoted by a trade-off between seed survivorship and seed yield. The trade-off decreases the advantage of high-s species in high-variability conditions, so the coexistence set is expanded to include higher seed survivorships. The trade-off was modelled in the simulations by setting:

$$K_i(t) = K(t)(1 - m \cdot s_i^b). \qquad (4)$$

In (4), m is the fraction of yield lost at $s = 1$ compared to $s = 0$, and b controls the shape of the trade-off. Figure 6a shows the effects of the trade-off in an environment with frequent severe disturbances ($K_d = .01$ one year in four). With a weak trade-off ($m = 0.2$) the coexistence set is small, but with stronger trade-offs it expands. Figure 6b is a similar example, with no disturbances but more variability; note that the coexistence set may include the vast majority of possible competitors.

Coevolution of the species also could produce large coexistence sets (e.g., Fig. 7). The coevolution simulations greatly overestimate the potential for coexistence in such cases, since the species are not allowed to go extinct before the coevolution occurs. Outside the non-coevolved coexistence set, however, extinction would be expected for $\alpha = \alpha_c^*$ unless the germination fraction evolves rapidly. Also, Fig. 7 shows an exceptionally large effect of coevolution. The irregular shape of the coexistence set is typical, but inexplicable; I do not even have any guesses. The important conclusion from the coevolution simulations is that coexistence predicted via ESS theory is not destabilized by coevolution.

Discussion

The theoretical results for coexistence of alternative germination strategies may be expected to apply more generally to low-risk and high-risk strategies such as those listed in Table 1. For example, model (3) is formally equivalent to a 'reproductive effort' model for perennials with constant adult survivorship and variable, density-dependent juvenile recruitment (the seed bank becomes the adult, and germination is reproduction). These assumptions are realistic for many plant and animal species (Chesson, 1985).

The model predicts that natural selection may lead to coexistence by 'time-sharing' a niche in a randomly fluctuating environment, via the evolution of low-risk and high-risk strategies for coping with the variability. Low vs. high is relative — in more variable environments, all species must adopt a more cautious, lower-risk strategy. The divergent

strategies arise due to intrinsic differences between the species (e.g., the different seed survivorship in this paper's model), which predispose them to evolving different strategies.

Coexistence is promoted by a trade-off between (a) the capacity to maximize gains (e.g., growth, reproduction, nutrient uptake) during favorable periods and (b) the capacity to minimize losses (e.g., mortality, nutrient leakage) during unfavorable periods. The trade-off between seed yield and seed survivorship (Fig. 6) is an example; in the 'reproductive effort' interpretation of the model, the trade-off is between adult survivorship and adult fecundity.

Coexistence is most promoted usually at intermediate levels of variability or intermittent stress. This prediction is particularly interesting, because it differs from the predictions of models which do not include the evolutionary dynamics. In models where the species' life-history parameters are static, coexistence is most promoted either at maximum or at minimum levels of variability [Chesson, 1985]. In this paper's model, the species are allowed to evolve to an ESS determined by the level of variability. Intermediate levels of variability are then most conducive to the evolution of coexisting alternative strategies.

Several recent studies support the qualitative prediction, that low-risk and high-risk strategists can co-occur in intermittently stressed habitats. Boorman (1982) found that the biennial and perennial species of a dune-grassland had a broad range of growth strategies. Some had potentially rapid growth, while others had low growth rates but tolerated extreme conditions. These were classed as 'C' and 'S' strategists, respectively, according to Grime's (1979) classification, but could also be described as high-risk and low-risk strategists as defined here. Ernst (1983) similarly found wide variability among dune annuals in their responses to nutrient deficiency and in their tolerance for burial under sand by wave action. Other studies of coastal vegetation have found large interspecific differences in responses to stresses from salinity, immersion, and sulfide (Schat & Scholten, 1985; Havill *et al.*, 1985; Groenendijk, 1985; Rozema *et al.*, 1985), and in the extent to which a permanent reserve of

viable seeds is maintained (Schat, 1983; Ernst, 1985). In many of those cases, however, the different strategies may simply be associated with different habitats or different vegetation zones along the salinity gradient inland from the coast (e.g., Rozema *et al.*, 1985).

Menges & Walker's (1983) study of floodplain herbs found coexistence of alternative strategies at middle elevations, but not at higher or lower elevations. Since elevation is inversely related to flooding frequency in their study areas, their results support the prediction that coexistence of alternative strategies is most likely at intermediate frequencies of stress. High elevations were dominated by 'competitive' strategists, according to Grime's (1979) classification, while 'ruderal' and 'stress-tolerant' strategists shared the ground at middle elevations. In the lowest elevation zone (20–39 cm; below that, aquatic plants dominate), stress-tolerant species predominate and only one ruderal (*Polygonum punctatum*) has its distributional optimum in that zone (Table 2 of Menges & Waller, [1983]; they did not discuss the loss of ruderal species below 39 cm).

Many of the best-studied examples of alternative reproductive strategies within a community concern short-lived grassland forbs. Studies of chalk grassland forbs in the Netherlands (reviewed by During *et al.*, 1985) have been summarized by arranging the species along a gradient from 'savers' to 'spenders,' which is very similar to the low-risk vs. high-risk comparison in Table 1 (although they were done independently). 'Savers' are characterized (relative to 'spenders') by a longer lifespan; fewer, larger seeds; reproduction spread over several years; a persistent seed bank; and less reaction to spatial and temporal variations (During *et al.*, 1985).

The species differ in responses to climatic fluctuations, which fits this paper's theory, but there is no evidence yet on the role of competition between spenders and savers, if indeed there is any [H. During, pers. comm.]. Competition with the matrix-forming perennials, and the annual mowing of the community, may keep the forbs at densities too low for competition within the guild to have any importance (During *et al.*, 1985; H. During, pers. comm.).

Observations of chalk grassland in England also suggest that the short-lived species suffer more interference from the matrix-forming perennials than from other short-lived species (Grubb et al., 1982).

The low-risk strategist as defined here is conceptually very similar to Grime's 'stress-tolerant' strategist. The results therefore can be viewed as an extension of Grime's (1979) ideas about plant strategies in relation to habitat. Grime (1979) suggested that coexistence of alternative strategies might be explained by phenological divergence – species active in different seasons face different conditions, and therefore adopt different strategies. Our model applies to species active during the same season, under conditions which are stressful in some years. One possible result, according to the model, is coexistence of stress-tolerant 'low-risk' species with 'high-risk' species that adopt the strategy appropriate to non-stressed conditions.

Finally, it must be emphasized that this paper addresses only one of many possible mechanisms for the coexistence of ecologically similar plant species (Shmida & Ellner, 1984), and no claim is made that high-diversity plant communities may be explained solely on the basis of the theory presented here.

Appendix Methods of analysis

The analyses are all based on the 'mutual invisibility' criterion for coexistence in randomly varying environments, conjectured by Turelli (1981) and recently proved by Chesson & Ellner (manuscr.) under conditions satisfied by this paper's models. Two boundary growth rates

$$v_i = \lim_{x \to 0} E[\ell n(X_i(t+1)/X_i(t)) \,|\, X_i(t) = x] \quad (5)$$

determine the outcome. In (5), 'E' denotes mathematical expectation with respect to the random environmental variability and the random fluctuations in the numbers of competitors, in the limit $X_i \to 0$. If $v_1 > 0$ and $v_2 > 0$, the species coexist in the sense of 'stochastic boundedness' as defined by Chesson [1982]. If $v_1 < 0$ and $v_2 > 0$, species 1 fails to invade a population dominated by species 2.

In the ESS germination analyses, the species differ only in their value of G, and the boundary growth rates are functions $v_i(G_1, G_2)$. G_1 is an ESS if $v_2(G_1, G_2) < 0$ for any $G_2 \neq G_1$; this defines a unique G_1, unless the model is such that no value of G allows persistence [Ellner, 1985]. The v_i can be calculated approximately, in the case of 'small fluctuations' in $K(t)$, by Taylor

expansion about the mean of X_1 and K (Ellner, 1985, following Turelli, 1981).

The coexistence set was approximated as follows. For any G_i, s_i, α and distribution of $K(t)$, $V = \min(v_1, v_2)$ was approximated by the small fluctuations procedure. α_c^* is defined by $V = 0$ for $K(t)$ held constant at its mean. If $V > 0$ at $\alpha = \alpha_c^*$ then $\alpha_v^* > \alpha_c^*$. The small fluctuations formula for the ESS germination fraction is then substituted into the condition $V > 0$ at $\alpha = \alpha_c^*$. This approximate description of the coexistence set was simplified by Taylor expansion about $C = 0$, retaining only the first two terms. The first term is messy but tractable and the second term contributes a uniformly $O(C)$ perturbation which does not effect the qualitative results for C small.

References

Abrams, P., 1984. Variability in resource consumption rates and the coexistence of competing species. Theor. Pop. Biol. 25: 106–124.

Ågren, G. I. & Fagerström, T., 1984. Limiting dissimilarity in plants: randomness prevents exclusion of species with similar competitive abilities. Oikos 43: 369–375.

Boorman, L., 1982. Some plant growth patterns in relation to the sand dune habit. J. Ecol. 70: 607–614.

Bulmer, M. G., 1984. Delayed germination of seeds: Cohen's model revisited. Theor. Pop. Biol. 26: 367–377.

Chesson, P. L., 1982. The stabilizing effect of a random environment. J. Math. Biol. 15: 1–36.

Chesson, P. L., 1986. Environmental variation and the coexistence of species. In: T. Case & J. Diamond (eds), Community ecology, pp. 240–256. Harper & Row, New York.

Chesson, P. L. & Warner, R. R., 1981. Environmental variability promotes coexistence in lottery competitive systems. Am. Nat. 117: 923–943.

Cohen, D., 1966. Optimizing reproduction in a randomly varying environment. J. Theor. Biol. 12: 110–129.

Cohen, D., 1968. A general model of optimal reproduction in a randomly varying environment. J. Ecol. 56: 219–228.

Cushing, J., 1980. Two species competition in a periodic environment. J. Math. Biol. 10: 385–400.

Cushing, J., 1984. Periodic two predator, one prey interactions and the time sharing of a resource niche. SIAM J. Appl. Math. 44: 392–410.

De Mottoni, P. & Schiaffino, A., 1981. Competition systems with periodic coefficients: a geometric approach. J. Math. Biol. 11: 319–335.

During, H. J., Schenkeveld, A. J., Verkaar, H. J. & Willems, J. H., 1985. Demography of short-lived forbs in chalk grassland in relation to vegetation structure. In: J. White (ed.), The population structure of vegetation, pp. 341–370. Junk, Dordrecht.

Ellner, S., 1984. Asymptotic behavior of some stochastic difference equation population models. J. Math. Biol. 19: 169–200.

Ellner, S., 1985. ESS germination strategies in randomly varying environments. Theor. Pop. Biol. 28: 50–79, 80–116.

Ernst, W. H. O., 1983. Anpassungsstrategien einjähriger Dünenpflanzen. Verh. Ges. Ökologie (Mainz 1981) 10: 485–495.

Ernst, W. H. O., 1985. Some considerations of – and perspectives in coastal ecology. Vegetatio 62: 533–545.

Evenari, M., Shanan, L. & Tadmor, N., 1971. The Negev: The challenge of a desert. Harvard U. Press, Cambridge, MA.

Grime, J. P., 1979. Plant strategies and vegetation processes. Wiley, New York.

Groenendijk, A. M., 1985. Ecological consequences of tidal management for the salt-marsh vegetation. Vegetatio 62: 415–424.

Grubb, P. J., 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. Biol. Rev. 52: 107–145.

Grubb, P. J., Kelly, D. & Mitchley, J., 1982. The control of relative abundance in communities of herbaceous plants. In: E. I. Newman (ed.), The plant community as a working mechanism, pp. 79–97, Blackwell, Oxford.

Harper, J. L., 1977. Population biology of plants. Academic Press, New York.

Havill, D. C., Ingold, A. & Pearson, J., 1985. Sulfide tolerance in coastal halophytes. Vegetatio 62: 279–285.

Hutchinson, G. E., 1961. The paradox of the plankton. Am. Nat. 95: 137–145.

Klopfer, P. H. & MacArthur, R. H., 1961. On the causes of tropical species diversity: niche overlap. Am. Nat. 95: 223–226.

Levin, S. A., 1980. Some models for the evolution of adaptive traits. In: C. Barigozzi (ed.), Vito Volterra Symposium on mathematical models in biology. Proc., Rome 1979, pp. 56–72. Springer, New York.

May, R. M., 1973. Stability in randomly fluctuating versus deterministic environments. Am. Nat. 107: 621–650.

May, R. M., 1974. On the theory of niche overlap. Theor. Pop. Biol. 5: 297–332.

May, R. M. & MacArthur, R. H., 1972. Niche overlap as a function of environmental variability. Proc. Natl. Acad. Sci. USA 69: 1109–1113.

Maynard Smith, J., 1982. Evolution and the theory of games. Cambridge University Press, Cambridge.

Menges, E. S. & Waller, D. M., 1983. Plant strategies in relation to elevation and light in floodplain herbs. Am. Nat. 122: 454–473.

Rozema, J., Bijwaard, P., Prast, G. & Broekman, R., 1985. Ecophysiological adaptations of coastal halophytes from foredunes and salt marshes. Vegetatio 62: 499–521.

Schat, H., 1983. Germination ecology of some dune-slack pioneers. Acta Bot. Neerl. 32: 203–212.

Schat, H. & Scholten, J., 1985. Comparative population ecology of dune slack species: the relation between population stability and germination behaviour in brackish environments. Vegetatio 61: 189–195.

Shmida, A. & Ellner, S., 1984. Coexistence of plant species with similar niches. Vegetatio 58: 29–55.

Slatkin, M., 1978. The dynamics of a population in a Markovian environment. Ecology 59: 249–256.

Tilman, D., 1982. Resource competition and community structure. Princeton University Press, Princeton.

Turelli, M., 1978. A reexamination of stability in randomly varying versus deterministic environments with comments on the stochastic theory of limiting similarity. Theor. Pop. Biol. 13: 244–267.

Turelli, M., 1981. Niche overlap and invasion of competitors in random environments. I. Models without demographic stochasticity. Theor. Pop. Biol. 20: 1–56.

Whittaker, R. H., 1975. Communities and ecosystems. MacMillan, New York.

Beyond reductionism and scholasticism in plant community ecology

Paul A. Keddy*
Department of Biology, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5

Abstract

Recent progress in plant population biology is often used to argue that most questions in community ecology should be answered using a reductionistic approach. This progress may, however, be attributed to at least two other factors: (1) greater emphasis on clear questions with testable alternatives, and (2) greater agreement upon the important variables which describe the systems of interest. In plant community ecology there has been a tendency to collect data rather than pose clear questions, and there is lack of agreement on which community properties are most useful to measure. As a consequence, there is sometimes a tendency to debate concepts instead of posing clear questions. Other more productive avenues for community level ecology exist.

Introduction

Since the publication of Harper's (1967) paper entitled 'A Darwinian approach to plant ecology', there has been rapid growth in plant population biology (e.g., Harper, 1977; Stearns, 1977; Solbrig *et al.,* 1979; Dirzo & Sarukhan, 1984). This growth is frequently contrasted with that in plant community ecology and/or vegetation science, followed by the conclusion that reductionism is therefore the avenue to advancement in community ecology.

My own interpretation of Harper's paper is that the reductionism – holism interpretation overlooks two other equally important messages. Both of these messages apply equally well to any level of organization being investigated. Here I will discuss them in the context of the community level of organization. Studies at this level of organization are variously called 'plant community ecology' or 'vegetation science', with vegetation science of-

ten associated more strictly with phytosociology. Since these comments are directed simply at the community level of organization, I will not emphasize such distinctions.

Let us reconsider two aspects of Harper's paper in turn. He first reiterates the conceptual framework of population biology (evolution) and using it poses a series of questions. The same conceptual framework is then used to propose dependent variables which must be measured to answer these questions.

Questions

There is frequently less emphasis on precise questions at the community level of organization. This lack of emphasis on questions, or testable hypotheses, may result from what might initially seem to be a strength of such research: the relative ease with which data can be collected. It is so easy to collect descriptive data in plant community ecology, however, that there is always the temptation to start sampling quadrats just to see what happens or in

hope that the data will be useful someday. There is also, as a consequence, a strong emphasis on statistical methods for manipulating large data sets. Neither vegetation description, nor new statistical techniques, can themselves produce scientific advancement. They should be treated as just two of many possible tools available for answering questions. Perhaps community ecology would develop more rapidly if we all hung up our quadrats for a few years and instead tried to decide what questions we need to answer.

This is not to completely deny the value of purely descriptive studies. Phytosociology has at least two useful roles. The first is to provide a vegetation classification for resource management, including the systematic selection and management of nature reserves. The second is to generate hypotheses regarding vegetation processes which can then be experimentally tested. While there are many examples of the former, there are relatively few of the latter. Goldsmith's studies on sea cliff vegetation (1973a, b, 1978) illistrate the way in which descriptive multivariate studies can be used to generate questions, and the role of experiments in answering such questions. Many purely descriptive studies, however, do not pose meaningful questions about ecological processes, much less propose how such questions could be answered experimentally.

Measurable dependent variables

The second strength of Harper's paper is the emphasis upon measurable dependent variables. Evolution provides strong clues as to what variables are important to measure: they are traits which determine an organisms's contribution to future generations. In the absence of a similar conceptual framework at the community level of organization, there has been a tendency to concentrate on descriptions of species composition as if only taxonomic names could describe communities. In contrast, when more holistic variables such as diversity and biomass are considered, patterns at the community level emerge clearly (e.g., Al-Mufti *et al.,* 1977; Grime, 1973, 1979; Silvertown, 1980; Tilman, 1982, pp. 123–132, Wheeler & Giller, 1982; Del

Moral, 1983, 1985; Menges & Waller, 1983; Wilson & Keddy, 1986). The resulting relationships are not only consistent across many vegetation types, but also allow predictions of vegetation responses to changes in factors such as grazing or fertilization.

Variables such as biomass and diversity are not the only ones which can describe vegetation; Grime (1979) proposes that traits such as physiology, life form, phenology, demography and morphology vary predictably within plant communities. Systematically screening large numbers of species for specific traits then provides new variables for describing communities; two recent examples are maximum relative growth rate (Grime & Hunt, 1975) and germination patterns (Grime, 1979; Grime *et al.,* 1981). These methods of 'comparative plant ecology' have the potential to produce important new dependent variables for describing plant communities.

Concepts and theories

The interaction between clearly defined questions and measurable traits of vegetation cannot be over-emphasized. Clearly defined questions generate alternative hypotheses which are testable, and measurable variables provide the means to test the hypotheses. If either of these is missing, the process of discovery is hindered, if not entirely halted. Thus, some of the questions asked in vegetation science (e.g. continuum vs. community unit concept) have not been productive because they relate to concepts rather than hypotheses (sensu Peters, 1980). That is, they are stated in a non-testable form rather than as explicitly defined relationships among measurable quantities. In examining the decades of publications on concepts such as the continuum vs. community unit organization of vegetation, it seems reasonable to conclude that 'we have become modern scholastics interminably discussing questions which cannot be solved or tested scientifically' (Peters, 1980; Stearns, 1976). In re-examining the continuum as opposed to the community unit concept, for example, Shipley & Keddy (1987) have argued that existing concepts

generate four mutually exclusive hypotheses rather than two. Moreover, when inferential statistics are used to try and test these hypotheses, both the 'continuum' pattern and 'community' pattern are falsified and a third alternative accepted.

In conclusion, I have suggested that the two most important points in Harper's paper have little to do with reductionism, but rather with the application of certain scientific methods. These scientific methods are not in any way new, but are reiterated here for two reasons. First, there is no reason to assume that plant population ecology is the only, or even the best route to explore plant communities – indeed, it has conceptual and methodological problems of its own (Stearns, 1976, 1977). Second, community ecologists have, on occasion, consumed large amounts of time and energy discussing poorly defined concepts rather than posing clear questions and specifying the variables which must be measured to answer these questions.

The search for theories and models in vegetation science (the theme of this symposium) seems most likely to succeed if clear questions about community level phenomena are posed using measurable variables. What is less clear is which questions, and which variables, will produce the rapid progress, rigorous predictions and broad generalizations we require.

References

Al-Mufti, M. M., Sydes, C. L., Furness, S. B., Grime, J. P. & Band, S. R., 1977. A quantitative analysis of shoot phenology and dominance in herbaceous vegetation. J. Ecol. 65: 759–791.
Del Moral, R., 1983. Competition as a control mechanism in subalpine meadows. Am. J. Bot. 70: 232–245.
Del Moral, R., 1985. Competitive effects on the structure of subalpine meadow communities. Can. J. Bot. 63: 1444–1452.
Dirzo, R. & Sarukhan, J., 1984. Perspectives on plant population ecology. Sinauer Associates, Sunderland, Mass., U.S.A.

Goldsmith, F. B., 1973a. The vegetation of exposed sea cliffs at South Stack, Anglesey. I. The multivariate approach. J. Ecol. 61: 787–818.
Goldsmith, F. B., 1973b. The vegetation of exposed sea cliffs at South Stack, Anglesey. II. Experimental studies. J. Ecol. 61: 819–829.
Goldsmith, F. B., 1978. Interaction (competition) studies as a step towards the synthesis of sea–cliff vegetation. J. Ecol. 66: 921–931.
Grime, J. P., 1973. Control of species density in herbaceous vegetation. J. Envir. Manage. 1: 151–167.
Grime, J. P., 1979. Plant strategies and vegetation processes. Wiley, New York.
Grime, J. P. & Hunt, R., 1975. Relative growth rate: its range and adaptive significance in a local flora. J. Ecol. 63: 393–422.
Grime, J. P. et al., 1981. A comparative study of germination characteristics in a local flora. J. Ecol. 69: 1017–1059.
Harper, J. L., 1967. A Darwinian approach to plant ecology. J. Ecol. 55: 247–270.
Harper, J. L., 1977. Population biology of plants. Academic Press, London.
Menges, E. S. & Waller, D. M., 1983. Plant strategies in relation to elevation and light in floodplain herbs. Am. Nat. 122: 454–473.
Peters, R. H., 1980. Useful concepts for predictive ecology. Synthese 43: 257–269.
Shipley, B. & Keddy, P. A., 1987. The individualistic and community-unit concepts as falsifiable hypotheses. Vegetatio 69: 47–55.
Silvertown, J., 1980. The dynamics of a grassland ecosystem: Botanical equilibrium in the park grass experiment. J. Appl. Ecol. 7: 491–504.
Solbrig, O. T., Jain, S., Johnson, G. G. & Raven, P. H., 1979. Topics in plant population biology. Columbia University Press, New York.
Stearns, S. C., 1976. Life history tactics: a review of the ideas. Quart. Rev. Biol. 51: 3–47.
Stearns, S. C., 1977. The evolution of life history traits: a critique of the theory and a review of the data. Ann. Rev. Ecol. Syst. 8: 145–171.
Tilman, D., 1982. Resource competition and community structure. Princeton University Press, Princeton.
Wheeler, B. D. & Giller, K. E., 1982. Species richness of herbaceous fen vegetation in Broadland Norfolk in relation to the quantity of above ground material. J. Ecol. 70: 179–200.
Wilson, S. D. & Keddy, P. A., 1986. Measuring diffuse competition along an environmental gradient: results from a shoreline plant community. Am. Nat. 127: 862–869.

Some models of catastrophic behavior in exploited forests

Marino Gatto & Sergio Rinaldi*
*Centro Teoria dei Sistemi, CNR, Politecnico di Milano, 20133 Milano, Italy***

Abstract

This paper presents some models which show that small variations of the human exploitation of a natural forest can give rise to dramatic changes in the forest biomass. Two simple mechanisms for catastrophes, already pointed out for other ecosystems, are briefly discussed at the beginning of the paper. A more interesting mechanism for catastrophes is then presented. It is based on a model of the dynamics of forest biomass and soil nutrients and on the fact that mortality may become very high when soil acidity exceeds a threshold. In particular, it is shown that an increase of the exploitation may give rise to a catastrophic collapse of the forest if the exogenous nutrient inflow, which brings about also acidic deposition, is sufficiently high. Moreover, such a catastrophe is irreversible, i.e., regeneration is not possible, if the nutrient inflow is too high.

Introduction

Both theory and observation indicate that some ecosystems are likely to possess several different equilibrium points (see May, 1977). This implies that substantial natural disturbances (such as, for example, wind, fire, pest, or release of toxic substances in the case of a forest) may perturb an ecosystem from a stable equilibrium into the region of attraction of another stable equilibrium. Thus, in these cases, the state of the system at a specified time strongly depends upon the time and amplitude of the past historical accidents. But the existence of multiple stable equilibria may also be the cause of more subtle and intriguing changes, called catastrophes, which are caused by a small variation of a parameter (see below).

In this paper we outline three mechanisms which cause catastrophes in the dynamics of models of natural forests subject to human exploitation. Two such mechanisms have already been discussed by Clark (1976) and Noy-Meir (1975) with reference to other ecosystems and involve just forest biomass dynamics and the specific mode of forest exploitation: the first is based on the inverse density-dependence in the relationship between the growth rate per unit biomass and the total biomass, while the second is related with the concavity of the harvesting function of the exploiters. The third mechanism is more complex and takes into account also the dynamics of soil nutrients. In this case the forest collapse is based on the fact that tree mortality may become very high when soil acidity exceeds a given threshold. Since soil acidity may depend upon the exogenous inflow of nutrient due to natural sources, fertilization and acidic deposition, also the possibility of collapse and regeneration may depend upon this inflow.

The mathematical models we use are very crude. We take in fact a short-cut approach to the problem by lumping many different phenomena into a single

equation of biomass growth. This simplicity, however, allows us to detect the stability properties of all the equilibria of a forest ecosystem, a condition which is necessary to point out the potential catastrophes. It should be clear, however, that the aim of this work is only conceptual and qualitative (as any formal theory in ecology): our model is not meant to rival, on a quantitative ground, the complex and detailed simulation models that are currently built and used (see Shugart, 1984, for a review).

On catastrophes

Catastrophe theory (see Thom, 1972; Zeeman, 1977) has started to be used in ecology (e.g., Jones, 1977; Bazin & Saunders, 1978; Casti, 1982: Kempf *et al.,* 1984). It can be introduced as follows.

A basic ingredient is a measure E of forest exploitation, which can be quantified in various ways, for instance, the percentage of mature trees that can be cut per year, or the amount of labor and capital employed in felling trees.

If exploitation is kept constant over time, the state of the system, i.e., the variables that fully characterize the ecosystem (tree standing biomass, harvested biomass, nutrient concentration, etc.) will converge, in the absence of natural disturbance, to one of its possibly many stable equilibria. When exploitation is smoothly and slowly increased also these equilibria will vary smoothly. But sometimes a small (infinitesimal) perturbation of the exploitation E from a particular given value $\bar{E}$ entails disappearance of the equilibrium. In other cases, the equilibrium which is stable for $E < \bar{E}$ becomes critical for $E = \bar{E}$ and unstable for $E > \bar{E}$. Now, if we imagine that an exploited forest is in a stable equilibrium for $E < \bar{E}$ and that this equilibrium disappears or becomes unstable for $E > \bar{E}$, we can say that we have a catastrophe. In fact, as soon as the exploitation E is higher than $\bar{E}$, the state of the system moves toward a different stable equilibrium which can be quite far from the one in which the system was before the perturbation. This dynamic transition from one equilibrium to another is a catastrophe. Often there are two critical values of the exploitation: a typical situation is as in Fig. 4 which refers to a model detailed later on. The solid curves describe the values of forest biomass at equilibrium for different exploitations: notice that in between

the critical values E' and E'' there are two stable equilibria, one of which is extinction. If exploitation is slowly increased from zero, biomass T smoothly declines (see open circles 1, 2, 3) until the critical value E'' is reached (closed circle 4) where a catastrophic transition from T'' to extinction occurs because for bigger exploitation pressure only the null biomass is a stable equilibrium. After this collapse, decreasing the exploitation will have no effect (the state of the ecosystem is 'trapped' in the extinction equilibrium) until the critical value E' is reached where biomass can regrow to the level T'. If E' is zero the catastrophe is called *irreversible* for obvious reasons, *reversible* otherwise.

Growth and harvesting functions

We consider a forest, not at the level of a single stand, but as a mosaic tessellated with many small patches. In this case it is not unrealistic to assume that the biomass response of the forest landscape can be described by a single differential equation (see Shugart, 1984, pp. 134–138).

Let $T(t)$ be the total standing biomass in the forest at time t. Then, if the forest is not exploited, the rate of change dT/dt of biomass is simply the natural growth rate G which is here assumed to be a function of forest biomass only. In the normal compensatory case the growth rate per unit biomass $G(T)/T$ is a decreasing function of total biomass (density-dependence operates via increased mortality and/or decreased fertility) and $G(T)$ is concave as shown in Fig. 1. Sometimes, however,
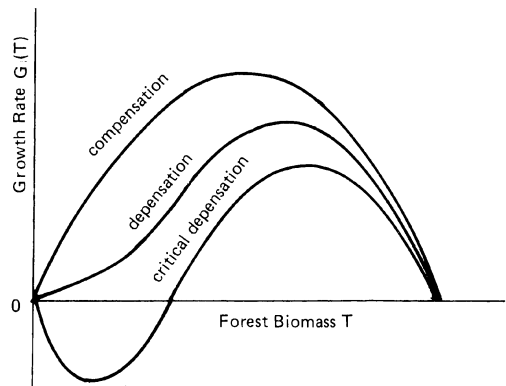

Fig. 1. Different types of natural growth rate functions.

crowding can be, within certain limits, not deleterious but favorable to the well-being of each tree. For instance, Silvertown (1982) reports two studies on *Pinus ponderosa* which show that the probability of a seed escaping predation by chalcid wasps (Fowells & Schubert, 1956) and by Abert squirrels (Larson & Schubert, 1970) is an increasing function of crop size. Also, other mechanisms come to mind like resistance to wind throw and moisture retaining capacity in the soil. Therefore, the effect of density is contrary to usual (inverse density-dependence). In this case it is possible that the growth rate per unit biomass $G(T)/T$ is first increasing with T and then decreasing: this phenomenon is called *depensation*. When, in the limit, $G(T)$ is initially negative, the so-called *critical depensation* takes place. Thus, if there is depensation the natural growth rate $G(T)$ is first convex and then concave as shown in Fig. 1.

When the forest is exploited, the rate of change dT/dt of standing crop is the difference between natural growth rate G and harvesting rate H. In real world situations H may be a rather complex function of age distribution, density, and current economic value of the trees present in the forest. Nevertheless, in order to describe the forest growth with a single equation we must assume that the harvesting rate H can be specified as a function of biomass only, i.e., $H=H(T)$. This function will be called *harvesting policy* in the following. Moreover, let us write

$$H(T)=E\,h(T) \tag{1}$$

where the *harvesting function h(T)* describes the dependence of harvest upon biomass and the multiplicative parameter E, called *exploitation*, somehow represents the aggressiveness of the exploiters and hence the stress exerted on the forest. Variations over time of the harvesting policy $H(T)$ may be due to variations of the harvesting function $h(T)$ or to variations of the exploitation E, the object of this paper.

There are several possible shapes for the harvesting function $h(T)$. The simplest is related to the use of *percent quotas*, which means that a fixed percentage of the standing trees can be felled each year. In this case the harvesting policy is simply given by

$$H=ET \tag{2}$$

where the exploitation E is proportional to the percentage of standing biomass that can be harvested each year (Fig. 2, curve a).

A second shape for $h(T)$ follows from the assumption that E represents, in suitable units, the amount of labor and capital invested. Then the harvesting rate cannot indefinitely increase with biomass, because, given E, a certain finite harvesting capacity cannot be exceeded. Thus the harvesting function must be a saturating one. A possible analytical choice (see Holling, 1959) is

$$H=Eh=ET/(\alpha+\beta T) \tag{3}$$

with α and β positive parameters (Fig. 2, curve b).

Finally, if one considers that lumbering may not be rewarding when the forest biomass is low (e.g., because this biomass mostly consists of young, nonmarketable trees), a third possible shape for the harvesting function is a sigmoid one (Fig. 2, curve c).

Two simple catastrophes

Let us now consider the case of a forest with depensation and let us suppose that the harvesting policy is linear (see Eq. (2)), i.e.,

$$dT/dt = G(T)\text{-}ET$$

Thus, the steady states corresponding to different degrees of exploitation can easily be found by intersecting the growth function $G(T)$ with the straight lines through the origin (see Fig. 3).
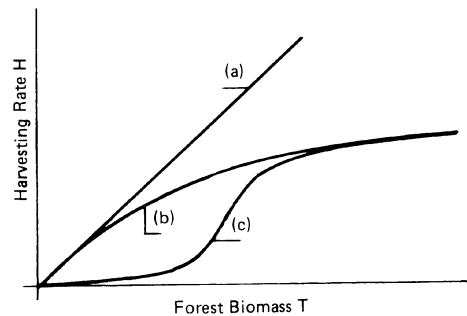

Fig. 2. Different types of harvesting policies: (a) *linear* (percent quotas), (b) *concave*, (c) *sigmoid*.
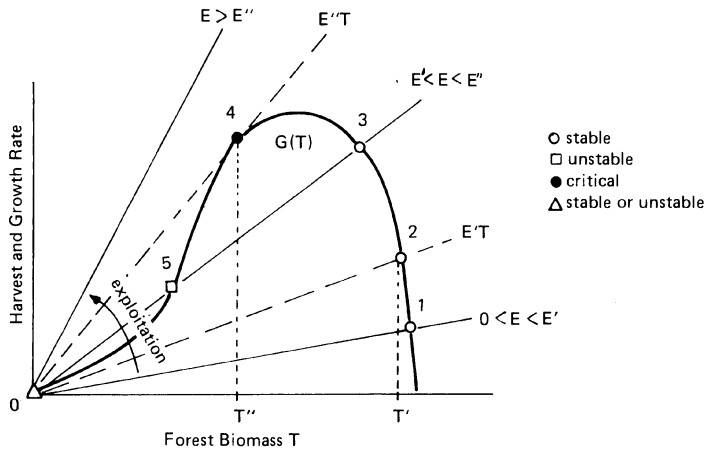
Fig. 3. Depensation in the growth rate *G(T)* and multiple equilibria.

If the exploitation is low, i.e., $0 < E < E'$, where at the origin (see Fig. 3), there are two equilibria: the extinction (point 0) which is *unstable* and a nontrivial steady state (point 1) which is *stable*. For higher exploitation, i.e., when $E' < E < E''$ where E'' is defined in Fig. 3, a third intermediate unstable equilibrium (point 5) moves in, and extinction (point 0) becomes stable. Finally, if the exploitation is equal to E'' the two non-trivial equilibria (points 3 and 5) coincide, thus giving rise to a *critical* equilibrium (point 4), while if the exploitation is very high, i.e., $E > E''$ there is only one stable equilibrium, the extinction.

Correspondingly the mappings of the stable biomass *T* and of the harvest rate *H* versus exploitation are as shown in Fig. 4. It is apparent that the values E' and E'' of the exploitation mark the appearance of a catastrophe. When the exploitation is slowly increased the standing crop smoothly decreases (see points 1, 2, 3, 4 along the growth curve of Fig. 3) until the critical value T'' is reached (see Fig. 4). At that point if the exploitation is further increased, trees die and the forest collapses (recall that point 0 of Fig. 3 is stable for $E > E''$). Now, if the exploitation E is reduced after the new equilibrium $T = 0$ has been reached or approached, trees do not grow any more unless ex-
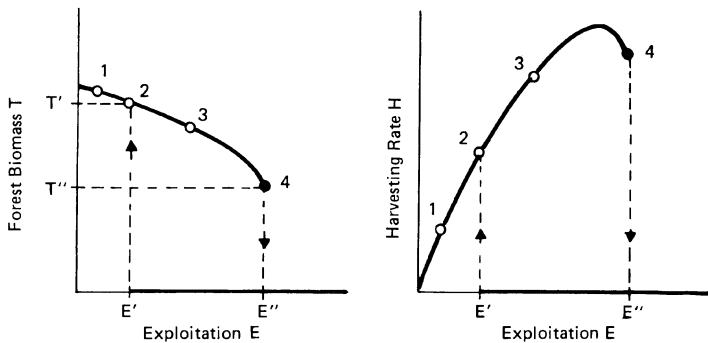
Fig. 4. Standing crop *T* and harvesting rate *H* vs. exploitation *E* when depensation is present. The values E' and E'' of the exploitation mark the critical values of the fold catastrophes. The points 1, 2, 3 and 4 are the same as in Fig. 3.

217

ploitation is substantially decreased. In fact, only when E becomes smaller than the critical value E'', point 0 in Fig. 3 becomes unstable again and the tree biomass sharply increases toward the value T'' (see Fig. 4). This type of catastrophe, which essentially points out the existence of a *hysteresis*, is called *fold catastrophe*.

When depensation is critical, extinction is always a stable steady state and this implies that the catastrophe is *irreversible*. In words, this means that in this case regeneration is not possible once the critical exploitation E'' has been exceeded.

As a second example we now consider the case in which the harvesting funtion $h(T)$ is concave. This is in practice the problem considered by Noy-Meir (1975) with grazed vegetation replaced by trees and herbivores by man. Since we have already shown that depensation can give rise to catastrophes, we assume that the natural growth rate is of standard compensatory type. Although it is not necessary we further consider, for the sake of simplicity, the case of *logistic growth* for tree biomass. Thus, the forest dynamics is given by

$$\frac{dT}{dt} = rT(1 - \frac{T}{K}) - \frac{ET}{\alpha + \beta T} \qquad (4)$$

where K is the *carrying capacity* and r the *intrinsic rate of increase*. The equilibria can be found by intersecting the curves $H = E\ h(T)$, corresponding to the particular harvesting function (5), with the logistic parabola $G(T)$ (see Fig. 5). For a particular value E' of the exploitation ($E' = r\alpha$) the two curves $G(T)$ and $E\ h(T)$ have the same tangent at the origin, as shown in Fig. 5. By comparing the curvatures of these two curves at the origin (i.e., the second derivatives of the functions $G(T)$ and $Eh(T)$) it is easy to prove that if $\alpha/K > \beta$ there are no multiple stable equilibria and hence no catastrophes; buf if

$$\alpha/K < \beta \qquad (5)$$

the outcome is as shown in Fig. 5. Since this figure is structurally equivalent to Fig. 3, we can conclude that when Eq. (5) is satisfied a reversible catastrophe takes place and forest biomass and harvest rate vary with exploitation as shown in Fig. 4.
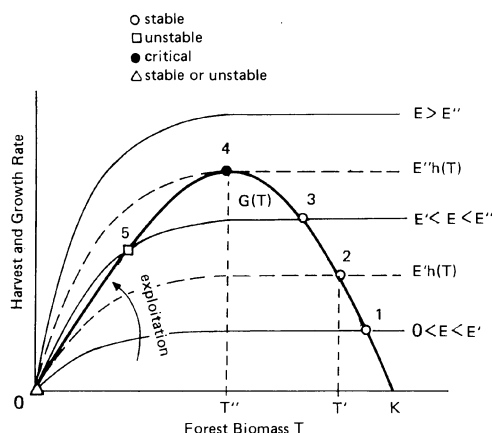

Fig. 5. Concave harvesting policies $E\ h(T)$ and multiple equilibria.

Soil nutrients and catastrophes

In this section we account for the dynamics of nutrients in the forest and we show that, under suitable circumstances, the interactions between standing crop, nutrients, and acid precipitation can give rise to catastrophes.

The model

Since we have already shown that the concavity of the harvesting function can be *per se* the cause of catastrophes, in order to stress the role of nutrients we now assume that the harvest H is linearly related to forest biomass (see Eq. (2)). The second basic assumption is that the nutrient pool for the standing biomass can be described by the concentration N of one nutrient which is homogeneously distributed in the soil. This hypothesis is crude although it is well-known that usually only one element (most often nitrogen) limits productivity in forests (see, for instance, Cole & Rapp, 1981; Ågren & Kauppi, 1983; Vitousek & Matson, 1984).

The model is as follows

$$dN/dt = W - aN - bNT + cm(N)T \qquad (6a)$$

$$dT/dt = [ebN - dT - m(N) - E]T \qquad (6b)$$

where a, b, c, d, e are positive parameters, E is the

exploitation, W is the input to the nutrient pool from sources other than decomposition, and $m(N)$ is the mortality. In Eq. (6a) the negative term $-aN$ represents the leaching losses, which at least for nitrogen are proportional to the nutrient concentration in the soil.

The second negative term $-bNT$ in Eq. (6a) is the nutrient uptake from the soil. The underlying assumption is that the root system is proportional to the above-ground biomass and that each root pumps up a constant amount of water per unit biomass so that the total uptake is proportional to T and N (*passive uptake*). This is, at least partially, unrealistic since a tree can develop, if needed, roots in a soil layer which is temporarily free from a too high nutrient concentration. However, when W and hence N are sufficiently high the nutrient reaches also those layers and our assumption becomes less crude.

If the amount of nutrient contained in each unit of standing biomass is roughly constant (as shown by Sprugel, 1984 with reference to balsam fir), then the nutrient transferred back to the soil by decomposition is proportional, through a multiplicative coefficient, c, to the dying forest biomass $m(N)T$. Here we make the strong assumption that decomposition is fast, since the flow of nutrient from the decomposable dead biomass enters the balance of the soil pool (Eq. (6a)) without any time delay. It should be noticed, however, that the components of the dead biomass which are most readily decomposed (foliage and branches) are usually richest in nutrient content (Sprugel, 1984). For these components mineralization is achieved with a time delay which is small when compared with tree lifetime. If, on the contrary, time scales were comparable, one should introduce a more sophisticated model with a distinction between organic and inorganic nutrient pools. The dying tree biomass would be an input to the organic pool, which would slowly release a flow of inorganic nutrient to the other pool.

In Eq. (6b) the term $ebN - dT$ represents the primary productivity. Therefore, the production of new biomass per unit biomass is enhanced by the nutrient uptake bN (the parameter e being an efficiency conversion factor), but inhibited when the total standing biomass is too high. This density dependence stands for root competition and increased shading, which limits the amount of energy captured from the surrounding environment by each tree unit (see Silvertown, 1982).

The final and perhaps most important, comment on the model is deserved by tree mortality m, which allows us to take acidic deposition into account. As is well known, one of the most important mechanisms for increased tree mortality is believed to be abnormal soil acidity (Matzner & Ulrich, 1981), because when pH falls below 5 aluminium solubility sharply increases. Owing to air pollution, precipitation is both carrying nutrient (e.g., nitrogen compounds) and acidity. Therefore increased emissions imply a bigger nutrient inflow W to the soil but also a greater acidic input. It follows that increasing N concentration is very often paralleled by increasing soil acidity. All this supports the assumption that mortality is a function of N, which is roughly constant up to a certain value and thereafter very sharply increasing.

It should be clear that the results which follow from the analysis of this model are direct consequences of the specific combination of phenomena we have taken into consideration. There is a whole array of other mechanisms, in particular due to acidic precipitation, which influence the growth and mortality of trees. Certainly, all of them play a role in a forest dynamics. Therefore, the link between the formal conclusions that will follow and the reality of a forest may be rather weak, unless the phenomena represented in our model are dominant, so that the real forest behaves at least qualitatively like our hypothetical (6).

Analysis of equilibria

The equilibria $(\bar{N}, \bar{T})$ can easily be found by annihilating the two derivatives dN/dt and dT/dt in Eq. (6). Geometrically, the equilibria can be visualized in the state space (N, T) as the intersections of the two *isoclines*, which are Eq. (6a) with $dN/dt = 0$ and Eq. (6b) with $dT/dt = 0$, respectively. Thus, the first isocline is given by

$$T = \frac{1}{c} \frac{W - aN}{\frac{b}{c}N - m(N)} \tag{7a}$$

while the second one is given by

$$T = 0 \qquad T = \frac{1}{d}[ebn - m(N) - E] \tag{7b}$$

The isoclines (7a) can be parametrized with respect to the nutrient inflow W, while the isoclines (7b) depend upon the exploitation E. In Fig. 6 we show how these isoclines can be found starting from the mortality function $m(N)$ and how they vary with W and E, respectively.

Once W and E are fixed at two prescribed values, we can have one or more intersections of the isoclines. In Fig. 7, for example, we have a case with two equilibria: point 1 corresponds to tree extinction in the forest, while point 2 represents a viable forest. The stability analysis of these equilibria

is very simple and can be carried out via *linearization,* but is not detailed here.

Catastrophes in forest ecosystems

We can now focus on the problem we like to discuss in this paper, namely the possibility of catastrophes in forest ecosystems. If we consider the exploitation E as a slowly varying parameter we can have five different cases depending upon the value of the nutrient inflow W. This is summarized in

(a)

(b)
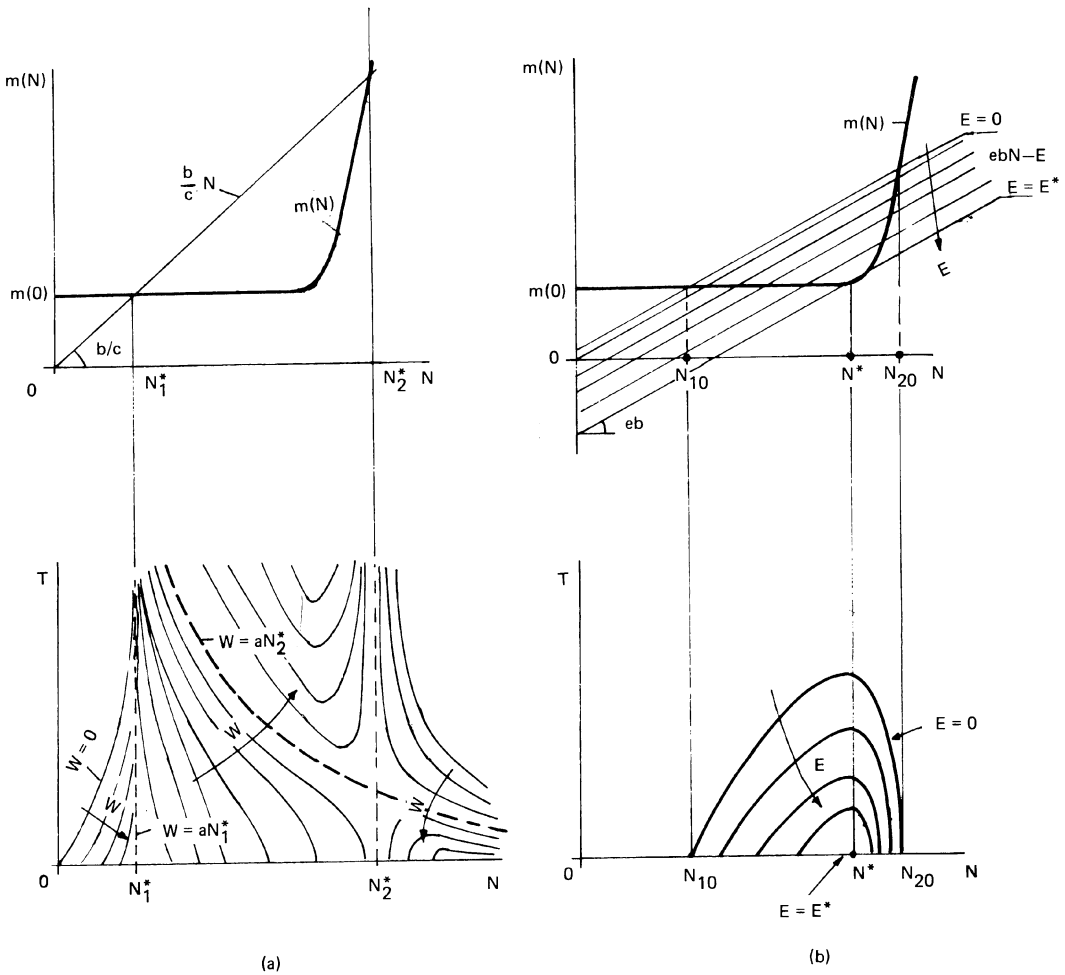
Fig. 6. (a) The isoclines $dN/dt = 0$ when W is varied; (b) the isoclines $dT/dt = 0$ when E is varied.
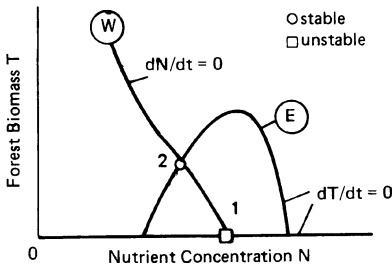
Fig. 7. The equilibrium points 1 and 2 as found by intersecting the isoclines.

Fig. 8 which must be read column by column starting from the left side. In case (i) when $W < aN_{10}$ there is only one stable equilibrium point, namely the forest extinction, and this is obviously true for all values of E. A similar situation occurs in case (v) which is characterized by very high values of W, namely $W > aN^{**}$. This means that our hypothetical forest cannot survive if the nutrient inflow to the system is either too low or too high. On the contrary, the forest is viable in the three other cases, i.e., when $aN_{10} < W < aN^{**}$ provided the exploitation is not too high. In case (ii) the stable non-

trivial equilibrium is characterized by a smooth decline to zero of forest biomass. Therefore, in this case, there is no catastrophe when the exploitation is slowly increased. On the contrary, in case (iii) we have a catastrophe since the forest collapses for $E > E''$. Nevertheless, this catastrophe is reversible as regeneration takes place if the exploitation is lowered below the critical value E'. Finally, in case (iv) the catastrophe is irreversible since the nutrient inflow W is so high that extinction is always a stable equilibrium point. All this can be simply summarized by saying that the chances of a forest collapse due to overexploitation are enhanced by high nutrient (hence acidity) inflow into the soil.

A similar analysis can be done by considering W as a slowly varying parameter and fixing the exploitation E at given value. Fig. 9 shows the result, namely a reversible fold catastrophe. This may be taken as a demonstration that exploited natural forests are potentially exposed to catastrophic collapses if acidic deposition becomes heavier and heavier.

The results of the overall analysis can be summed up by defining all the regions corresponding to different modes of behavior in the space of the two parameters E and W. These regions are shown in
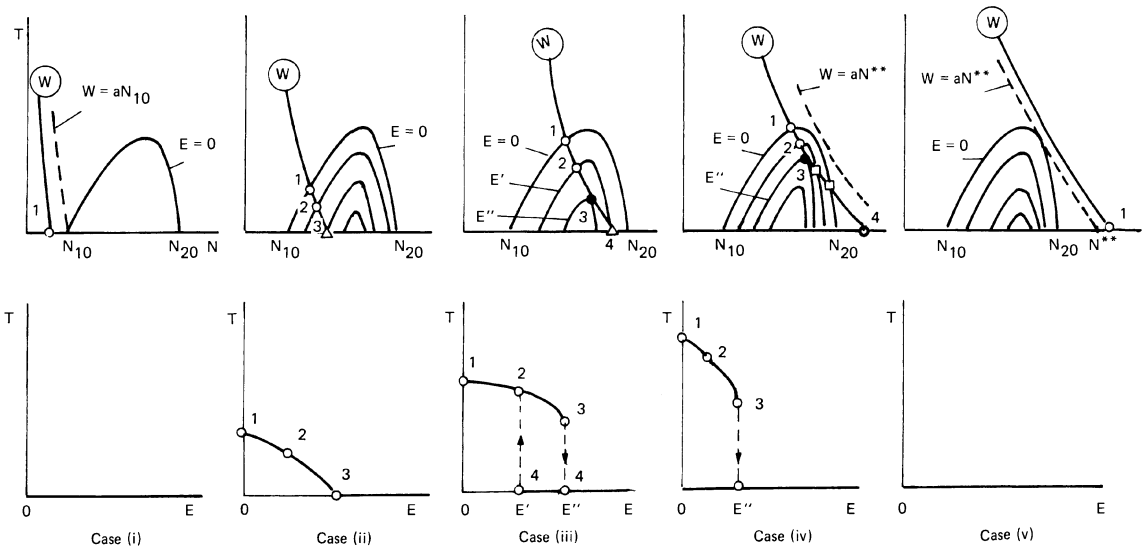

Fig. 8. The five different modes of behavior. (i): forest extinction is the only stable equilibrium; (ii): forest can survive but biomass smoothly declines to zero when exploitation is increased; (iii): reversible fold catastrophe: (iv): irreversible fold catastrophe; (v): forest extinction is the only stable equilibrium.
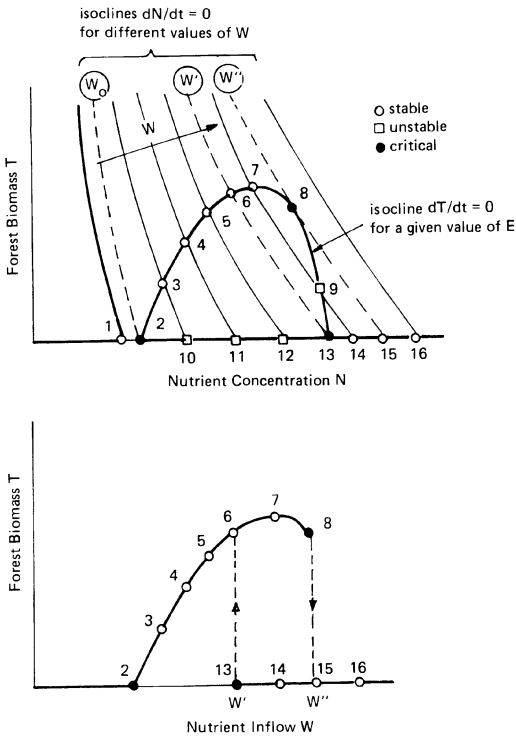
Fig. 9. (a) The equilibria in the state space for different values of *W;* (b) biomass *T* vs. nutrient inflow *W:* the points *W'* and *W''* mark the catastrophes.

Fig. 10 where the two solid lines represent the potential catastrophes: the highest one corresponds to a collapse when it is crossed from below (increase of acidic deposition) or from the left (increase of exploitation), while the lowest one corresponds to forest regeneration when it is crossed from above or from the right. The two lines form a cusp at point (E^*, aN^*) and for this reason this type of catastrophe has been named *cusp catastrophe.* The figure shows that for any given nutrient inflow W there is an upper limit to the exploitation E, above which our hypothetical forest is doomed to extinction. In particular if $aN_{10} \le W \le aN^*$ the biomass tends to zero when the exploitation approaches its upper limit, while if $aN^* < W < aN^{**}$ a catastrophic collapse occurs if this upper limit is exceeded. Conversely, if the exploitation E is smaller than the limit value E^* there exists an interval $[W_1, W_2]$ of nutrient inflows for which the forest is viable (see Fig. 10). This interval nevertheless becomes smaller and smaller when the exploitation is increased. In conclusion, it looks like human exploitation and acidic deposition have a kind of negative synergistic effect.
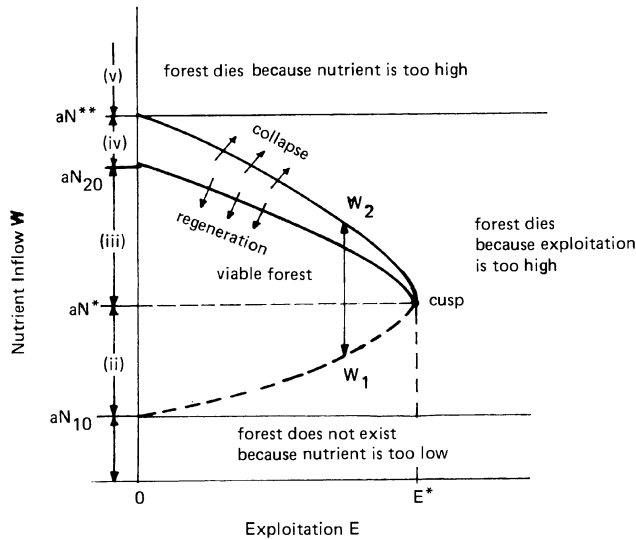

Fig. 10. The cusp catastrophe and the regions with different modes of behavior. The partition (i), (ii), ..., (v) of the vertical axis makes reference to the classification of Fig. 8.

Concluding remarks

Our modeling exercise shows that overexploitation of a hypothetical natural forest can lead to catastrophic and possibly unexpected collapses. The basic mechanisms for this are many and quite diversified. Among them we have pointed to depensation in the growth curve and concavity of the harvesting function. But more interesting is the catastrophe which is predicted by a nutrient-biomass model when mortality strongly increases with soil acidity. This case has been analyzed in detail in the paper and the main conclusion is that the catastrophe is a 'cusp' with respect to exploitation and acidic deposition. This means that in such a forest a collapse can occur when exploitation and/or acidic deposition are increased.

Investigations of other mechanisms which might generate catastrophes in exploited or unexploited forests should be possible developing similar modeling exercises. Among them we certainly have a number of important and interesting mechanisms for increased tree mortality (like damage to foliage or uptake of toxic substances) which are directly or indirectly related to the increase of acidic precipitation.

References

Ågren, G. I. & Kauppi, P., 1983. Nitrogen dynamics in European forest ecosystems: considerations regarding anthropogenic nitrogen depositions. CP–83–28, Int. Inst. Appl. Systems Analysis, Laxenburg, Austria.

Bazin, M. J. & Saunders, P. T., 1978. Determination of critical variables in a microbial predator-prey system by catastrophe theory, Nature, 275: 52–54.

Casti, J., 1982. Catastrophes, control and the inevitability of spruce budworm outbreaks. Ecol. Model. 14: 293–300.

Clark, C. W., 1976. Mathematical bioeconomics, Wiley, New York.

Cole, D. W. & Rapp, M., 1981. Elemental cycling in forest ecosystems. In: D. E. Reichle (ed.), Dynamic properties of forest ecosystems. pp. 341–409, Cambridge University Press.

Fowells, H. A. & Schubert, G. H., 1956. Seeds crops of forest trees in the pine region of California. Techn. Bull. US Dept. of Agriculture, 1150.

Holling, C. S., 1959. Some characteristics of simple types of predation and parasitism. Can. Entomol. 91: 385–398.

Jones, D. D., 1977. Catastrophe theory applied to ecological systems. Simulation, 29: 1–15.

Kempf, J., Duckstein, L. & Casti, J., 1984. Relaxation oscillations and other non-Michaelian behavior in a slow-fast phytoplankton growth model. Ecol. Model. 23: 67–90.

Larson, M. M. & Schubert, G. H., 1970. Cone crops of ponderosa pine in Central Arizona, including the influence of Abert squirrels. US Dept. of Agriculture, For. Serv. Res. Paper, 58.

Matzner, E. & Ulrich, B., 1981. Effect of acid precipitation on soil. In: R. A. Fazzolare & C. B. Smith (eds), Beyond the energy crisis, vol. 2, pp. 555–564 Pergamon Press, Oxford.

May, R. M., 1977. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. Nature, 269: 471–477.

Noy-Meir, I., 1975. Stability of grazing systems: an application of predator-prey graphs. J. Ecol. 63: 459–483.

Shugart, H. H., 1984. A theory of forest dynamics. Springer-Verlag, New York.

Silvertown, J. W., 1982. Introduction to plant population ecology, Longman, London.

Sprugel, D. G., 1984. Density, biomass, productivity, and nutrient-cycling changes during stand development in wave-regenerated balsam fir forests. Ecol. Monogr. 54: 165–186.

Thom, R., 1972. Stabilité structurelle et morphogénèse, Benjamin, New York. (English translation by D. H. Fowler: Structural Stability and Morphogenesis, Benjamin, New York, 1975).

Vitousek, P. M. & Matson, P. A., 1984. Mechanisms of nitrogen retention in forest ecosystems: a field experiment. Science, 225: 51–52.

Zeeman, E. C., 1977. Catastrophe theory: selected papers. Addison-Wesley, Reading, Massachusetts.

Author index