

Components of error in predictions of species compositional change

McCune, Bruce

Department of Botany and Plant Pathology, Weniger Hall 355, Oregon State University, Corvallis,
OR 97331-2902 USA; Tel. + 1 503 737 1741; Fax: + 1 503 737 1735

Abstract. A method is given for measuring two components of error (rate and direction) in predictions of compositional change through time. Observed compositional change between two times can be represented as a vector between two points in multidimensional species space. The point at the tail of this vector is the species composition at one particular time. A vector of predicted compositional change will diverge from the vector of observed change to some degree. The error in the predicted rate of change is measured by the difference between the lengths of the two vectors. The error in the predicted direction of change is measured by the angle between the vectors. The cosine of this angle is a non-standardized correlation coefficient (r_n) between the predicted and observed species compositions. The quantity $1 - r_n^2$ measures the error in direction of the predicted dynamics without being influenced by the overall rate of change. These measures in Euclidean space have useful counterparts in city-block space. The method is illustrated by comparing actual long-term changes in Mid-western old-growth forests with the changes predicted by a growth and yield model, TWIGS.

Keywords: Distance measure; Eastern deciduous forest; Error; Forest dynamics; Long-term study; Mathematical modeling; Prediction error; Simulation model; Succession.

Introduction

Long-term studies of biotic communities provide us with a basis for testing models of slow community processes, such as changes in species composition (Austin 1981). This is most often done for models of single community parameters, such as diversity or total biomass, that are reasonably analyzed for error and reliability in isolation from other community parameters (e.g. Rauscher 1986; Reynolds & Chung 1986; Warwick & Cale 1988). Comparison of predicted with observed species composition is more difficult because it is a multivariate problem and because typical community data are non-normal, having many zeros and a

strong positive skew (Legendre & Legendre 1983; Beals 1984).

Compositional predictions can be compared to observed changes by representing compositional change as a vector in multidimensional species space. Predicted and observed vectors share the same origin (i.e. the predictions are based on the observations at some point in time) and are compared in their rate (length) and direction. These comparisons can be done in either Euclidean or city-block space. The comparisons in Euclidean space are in part a familiar parametric analysis, correlation (explained below). In contrast, the comparisons in city-block space are more difficult to visualize, but may be better suited to the peculiarities of species composition data.

In this paper, city-block space is first compared with Euclidean space. Next, the representation of compositional change as a vector is described. Then methods for measuring errors in the rate and direction of compositional vectors in the two spaces are given. Finally, the method is used to compare compositional predictions from a forest growth-and-yield model with observed long-term changes from case studies.

City-block vs. Euclidean space

The difference between city-block (= 'Manhattan') and Euclidean space will be discussed only briefly here, because it has been clearly presented elsewhere (Sneath & Sokal 1973; Greig-Smith 1983; Legendre & Legendre 1983; Beals 1984). The difference is most easily visualized in two dimensions. In Euclidean space, the shortest distance between two points is a straight line with length $c = (a^2 + b^2)^{1/2}$, where a and b are the distances between the two points in each of the two dimensions.

This is the Euclidean case of $p = 2$ for the Minkowski metric in two dimensions, $|c|^p = |a|^p + |b|^p$ (Sneath & Sokal 1973, p.125). The n -dimensional generalization is $c = (|d_1|^p + |d_2|^p + \dots + |d_n|^p)^{1/p}$ where d_i is the distance

between the two points in dimension i . In contrast, if $p = 1$ and $n = 2$ we have $c = |a| + |b|$, the metric for city-block space in two dimensions, i.e. the sum of the two sides of the right triangle defined by the two points. To go from 9th and Vine to 8th and Maple you must walk the block between 9th and 8th streets and the block between Vine and Maple, rather than walking diagonally through the block.

Although city-block geometry is foreign to most statistical procedures, relativized forms of city-block distance have been widely used in ecology as the Czekanowski, Bray-Curtis, or Sørensen group of similarity and distance coefficients (Greig-Smith 1983, p. 194; Legendre & Legendre 1983). These measures have empirically been shown superior to Euclidean distance for describing similarity relationships in compositional data (Beals 1973, 1984; Noy-Meir & Whittaker 1977; Gauch 1982a, p. 113; Faith, Minchin & Belbin 1987; Blondel & Farré 1988). One reason for this is that city-block measures, by not squaring differences in abundance, give less weight to outliers (large differences in species abundance) than does Euclidean distance (Legendre & Legendre 1983, p. 198).

In this study Euclidean distance and city-block distance are used in their most basic forms (coefficients D_1 and D_7 of Legendre & Legendre 1983). Because all of the distances in this study were fairly short, the well-known problems resulting from using these measures on very heterogeneous data sets were minimal.

Neither distance measure takes into account the correlation structure among species or differences in variance among species, as would a Mahalanobis distance or distance in a reduced-dimensional ecological space. Although this paper concerns vectors in a full n -dimensional species space ['floristic' distance of Økland (1986)], the method is also applicable to an 'ecological distance' (op. cit.).

Compositional change as a vector

The species composition of a community at a moment in time (t_1) can be thought of as a point in 'species space', a space defined by as many axes of species abundance as there are species in the sample. A later sample (t_2) from the same geographic location as the first sample defines a second point in species space. An arrow (vector) connecting the two points summarizes the community's trajectory in species space from t_1 to t_2 .

The uses of 'vector' in this paper need clarification. In matrix terminology, each sample is a 'vector' in that it is a row or column of the species \times sample data matrix. No temporal trend is implied. The geometric basis for this use of 'vector' can be visualized as the line segment

connecting the origin of a coordinate system with a point defined by a row or column of numbers, each number describing the location of the point on a different axis (dimension). So, in matrix language, our samples at t_1 and t_2 are both vectors. By subtracting the vector at t_1 from the vector at t_2 ('centering' on t_1) the result is a vector in the sense of both matrix algebra and a change over time.

Vectors have frequently been used in the ordination literature of ecology to represent compositional change through time. They have been used to represent observed trajectories of communities through time, most often in an ordination-reduced two- or three-dimensional space (van der Maarel 1969; Allen, Bartell & Koonce 1977; Austin 1977; Lippe, de Smidt & Glenn Lewin 1985; McCune & Menges 1986; Lepš 1987; Malanson & Trabaud 1987; Webb, Steiger & Turner 1987; del Moral & Wood 1988; Smith & Iles 1988; Whittaker 1991) as well as simple structurally-based models of vegetation change (Goff & Zedler 1972; Carleton & Maycock 1978; Enright 1982; McCune & Allen 1985; Domon et al. 1986; Brisson, Bergeron & Bouchard 1988; Bergeron & Dubuc 1989) and complex dynamic models (Smith & Urban 1988). Points in species space have been connected along inferred successional trajectories represented by a chronosequence (Blondel & Farré 1988). Vectors in ordination spaces have also been used in other areas of ecology; for example, to represent changes in climatic responses of tree rings (Tessier 1989).

Changes in species composition typically have been represented by one or more straight-line vectors. In most cases the vectors are interpreted qualitatively, without a quantitative analysis of length and direction. Westman & O'Leary (1986), however, used quantitative characteristics of the vectors to measure components of resilience to disturbance.

The present paper extends these applications by showing how two vectors in species space, sharing a common origin, can be compared. In particular, I develop the example of the divergence of prediction from observation. This is a first step toward quantitative comparisons of community trajectories. Ultimately the methods need to be extended to trajectories composed of multiple segments.

Comparing predicted and observed compositional vectors

The method presented below has the following premises: (1) species composition of a community has been sampled at least twice (x_1 and x_2 ; Fig. 1). The two samples are the ends of the vector representing the

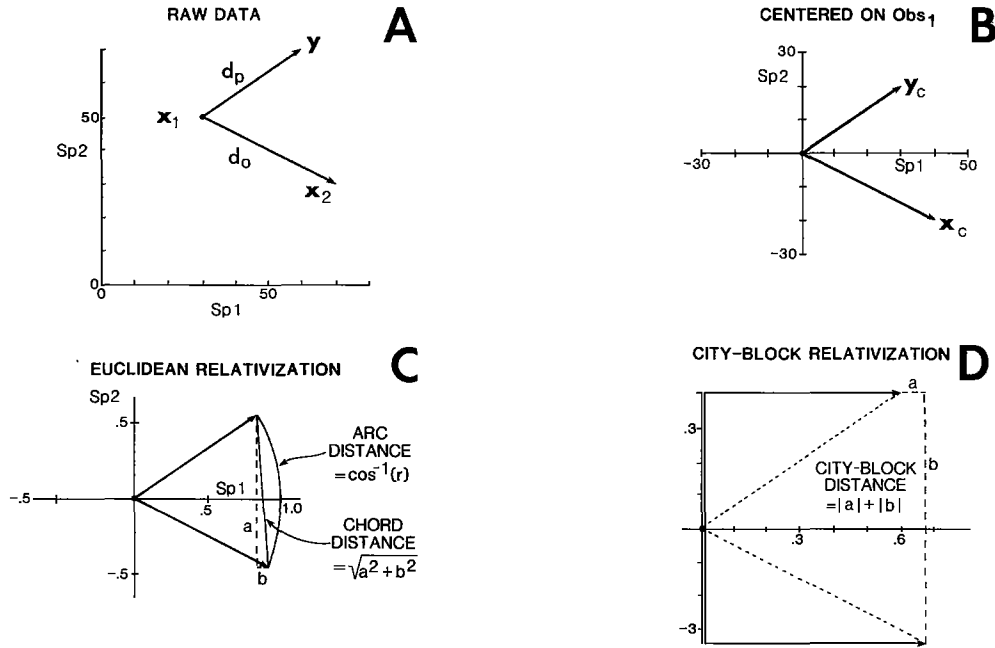


Fig. 1. Geometry of vector comparison in two dimensions, using a hypothetical example corresponding to Table 1. A. Vectors of predicted change and observed change diverging from observed species composition x_1 at t_1 . The axes labeled sp1 and sp2 represent raw abundance values for species 1 and 2 respectively. B. The same vectors as in 'A' but centered on x_1 . C. Euclidean measures of difference in direction after relativizing vectors to unit length. D. Difference in direction in city-block space after relativizing the vectors (solid right-angle lines) to equal totals.

observed compositional change. (2) a prediction has been made of species composition (y) based on x_1 . The vector from x_1 to y represents the predicted compositional change from t_1 to t_2 . (3) It is desired to compare the two vectors in both the rate and direction of change (Fig. 1). A synopsis of the method is given below.

Using the simple distance measures described above in an unreduced n -dimensional space gives the following characteristics: 1. It does not account for the correlation structure among species. 2. Species are not standardized with respect to variance in abundance. 3. Underestimates are equally important as overestimates; the method does ^{not} account for possible asymmetries in the underlying distribution of errors in abundance.

Steps in the analysis

1. Begin with three points (x_1, x_2, y ; see above) in species space describing two compositional vectors (predicted and observed) diverging from a common point x_1 , the species composition at t_1 .
2. Calculate the length of each vector (d_p and d_o) as the distance between head and tail. (Vector lengths can be calculated in either city-block or Euclidean space.) Vector length divided by the time interval is the average rate of compositional change (c_p and c_o) over that interval:

$$c_p = d_p / (t_2 - t_1) \text{ and } c_o = d_o / (t_2 - t_1).$$

3. Subtract the observed rate from the predicted rate of change to obtain the error in rate. To express the prediction error (e_r) as a percentage of the observed rate, $e_r(\%) = 100(d_p - d_o) / d_o$.

4. Center data on the observed composition at t_1 by subtracting x_1 from both x_2 and y : $x_c = x_2 - x_1$ and $y_c = y - x_1$ (Fig. 1B). This centering alters possible ranges of the usual distance measures, since species can now have negative abundance values (when a species is less abundant at t_2 than at t_1).

5a. *Euclidean method.* Calculate error in direction (e_d) as the angle between the vectors (Fig. 1C), $e_d = \Phi = \arccos(r_n)$, where r_n is the non-standardized correlation coefficient between x_c and y_c

$$r_n = \frac{\sum_{i=1}^N x_i y_i}{\sqrt{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i^2}} \quad (1)$$

where x_i and y_i are abundances of species i at x_c and y_c , respectively. This formula can be derived from the law

Table 1. Transformations of simple abundance data in two-dimensional species space and comparison of predicted with observed compositional changes. This hypothetical data set corresponds to Fig. 1. Sp1 and sp2 are species 1 and 2. The raw data are first centered on the observed species composition at t_1 , then standardized by setting vectors to unit length in either Euclidean or city-block space.

Community	Raw Data		Centered on t_1		Standardized			
	sp1	sp2	sp1	sp2	Euclidean		City-block	
Observed, t_1x_1	30	50	0	0	0.000	0.000	0.000	0.000
Predicted, t_2y	60	70	30	20	0.023	0.015	0.600	0.400
Observed, t_2x_2	70	30	40	-20	0.020	-0.010	0.667	-0.333
					Euclidean		City-block	
Error in rate, e_r (%)					-19.4		-16.7	
Error in direction, e_d (%)					37.7		40.0	

of cosines and the Pythagorean theorem. The geometric interpretation of a correlation coefficient has been presented elsewhere (e.g. see Beals 1984). The formula for r_n is similar to the standard Pearson product-moment correlation coefficient, except that means are not subtracted from the individual terms. The standardization that is built into Pearson's r would, in effect, equalize the means of species abundance, $\bar{x}_c = \bar{y}_c$, altering the geometry of the diverging vectors.

Some authors may prefer to express the directional error as a percent of the maximum possible error ($\Phi = \pi$ radians), based on the coefficient of determination (r_n^2). One can scale r_n^2 such that if $\Phi = \pi$ radians ($r_n = -1$), then $e_d(\%) = 100\%$; if $\Phi = \pi/2$, then $e_d(\%) = 50\%$; if $\Phi = \pi/4$, then $e_d(\%) = 25\%$; and if $r_n = 1$, then $e_d(\%) = 0\%$. This scaling is accomplished by: $e_d(\%) = 50 * (1 - (r_n * |r_n|))$. Note that scaling this percentage by r_n [using $e_d(\%) = (1 - r_n^2)/2$] rather than r_n^2 does not maintain a proportionate relationship between percent error and the angle of divergence.

There are two alternative routes to e_d . These routes may be computationally more direct depending on previous data standardizations or one's needs for intermediate results.

a. Set the vectors to unit length in Euclidean space by dividing each element (x_i) of the vectors by $(\sum x_i^2)^{1/2}$ for that vector (Table 1). This step is unnecessary for the other routes because standardizing the vector lengths is implicit in the calculating r_n . Then calculate the Euclidean distance between the x_c and y_c and convert this 'chord distance' (Orlóci 1967; Beals 1984) to the arc distance = $2 \cdot \arcsin(\text{chord distance}/2)$, the arc distance generally measured in radians.

b. Omit previous centering, instead calculating e_d as a variant of r_n that incorporates centering on x_1 :

$$r_n = \frac{\sum_{i=1}^N (x_i - z_i)(y_i - z_i)}{\sqrt{\sum_{i=1}^N (x_i - z_i)^2 \sum_{i=1}^N (y_i - z_i)^2}} \quad (2)$$

where z_i is the abundance of species i at x_1 .

5b. *City-block method.* Because angles are foreign to city-block space, the error in direction is measured as a distance rather than an angle. Set vectors to unit length in city-block space by dividing each element x of the vectors x_c and y_c by $\sum |x_i|$ and $\sum |y_i|$, respectively (Fig. 1D). This relativization results in the city block distance being a true metric, satisfying the triangle inequality axiom. Then calculate the error in direction, e_d , as the city-block distance between x_c and y_c ; $e_d = \sum |y_i - x_i|$. This error can be expressed as a percentage of the maximum possible error as $e_d(\%) = 100 \cdot e_d/2$. Note that $0 \leq e_d \leq 2$ when $\sum |x_i| = 1$ and $\sum |y_i| = 1$, but that other standardizations of x_c and y_c will result in other maximum values for e_d .

6. If desired, one can calculate the percentage or absolute contributions of individual species to the overall error. These are calculated on the basis of the distance between the heads of the unstandardized vectors. Note that direction of change is a community-level property and that there can be no error in direction for an individual species. Thus, the distinction between error in direction and error in rate applies only to multispecies comparisons.

Table 2. Error in rate ($m^2 ha^{-1} yr^{-1}$): observed compositional change vs. predictions by the model TWIGS.

Woods	Yr	Euclidean distance			City-block distance		
		Pred.	Obs.	% error	Pred.	Obs.	% error
Baber	18	0.121	0.133	-8.7	0.284	0.336	-15.4
Donaldson	30	0.085	0.071	18.2	0.245	0.178	37.9
Kaskaskia	38	0.148	0.161	-8.0	0.298	0.369	-19.4
Noe	22	0.262	0.234	11.7	0.326	0.405	-19.4

In city-block space, the signed absolute error for species i (e_i) is $e_i = y_i - x_i$. As a percent of the total error, $e_i \% = 100 |e_i| / \sum |y_i - x_i|$. In Euclidean space, $e_i = (y_i - x_i)^2$ and $e_i \% = 100 e_i / \sum (y_i - x_i)^2$.

Example

To illustrate the method, predictions of compositional dynamics are compared to observed values from four published long-term studies of tree dynamics. The four long-term data sets were from Baber Woods in east-central Illinois (19 yr; McClain & Ebinger 1968; Newman & Ebinger 1985), Donaldson's Woods in southern Indiana (30 yr; Lindsey, Barton & Miles 1958; Barton & Schmelz 1987), Kaskaskia in southern Illinois (38 yr; Schlesinger 1976), and Noe Woods in southern Wisconsin (22 yr; McCune & Cottam 1985).

Based on the species composition at the earliest sampling date for each of the woods, predictions were made with the forest growth and yield model TWIGS ('The Woodsman's Ideal Growth Projection System') (Belcher 1982; Miner, Walters & Belli 1988), Central States Version 2.0, the microcomputer version of STEMS (Belcher 1981; Belcher, Holdaway & Brand 1982; Miner & Walters 1984), to project stand dynamics based on the earliest data set for each of the four woods. Coefficients in the model are based on analyses of forest inventory

Table 3. Error in direction between observed compositional change and that predicted by the model TWIGS, expressed as a non-standardized correlation coefficient (r_n) and as a percentage of the maximum possible Euclidean and City-block distances.

Woods	Yr	r_n	Euclidean distance	City-block distance
			(Chord)	distance
			% of max.	% of max.
Baber	18	0.626	45.0	43.8
Donaldson	30	0.341	57.4	56.6
Kaskaskia	38	0.930	18.8	22.4
Noe	22	0.888	23.7	33.4

data on over 4000 permanent plots. It is a complex, non-spatial model that simulates growth and mortality of individual trees. Input data required are: diameter distributions by species, stand age, and site index. Ingrowth is not generated automatically. Rather, the user enters new trees as they achieve a minimum diameter. Ingrowth in these simulations was taken from actual observations.

Rate of change

The model's predictions were within $\pm 20\%$ of the observed value in both Euclidean and city-block spaces, except for a 38% overestimate for Donaldson Woods, as measured in city-block space (Table 2). Errors in predicted rate were somewhat higher in city-block than in Euclidean space. This results when errors are distributed across many species, rather than being concentrated in one or two species.

Direction of change

The direction of change was more difficult for TWIGS to predict than the rate of change. Typical errors ranged from 20-60% of the maximum possible error (Table 3). Predictions were generally similar or some-

Table 4. Species with largest errors (% of total error, measured in city-block space) between observed compositional change and that predicted by the model TWIGS. For a given woods, the sum of all percent errors equals 100. The 'Sign' column indicates whether the TWIGS prediction was too low (-) or too high (+).

Woods	Sign	%	Species	Sign	%	Species	Sign	%	Species
Baber	-	22	<i>Quercus borealis</i>	+	17	<i>Fraxinus americana</i>	-	16	<i>Carya glabra</i>
Donaldson	+	30	<i>Quercus alba</i>	-	17	<i>Quercus velutina</i>	-	12	<i>Liriodendron tulipifera</i>
Kaskaskia	-	37	<i>Liriodendron tulipifera</i>	+	14	<i>Sassafras albidum</i>	-	14	<i>Acer saccharum</i>
Noe	-	47	<i>Quercus alba</i>	-	26	<i>Quercus velutina</i>	-	10	<i>Prunus serotina</i>

what poorer using city-block measures than with Euclidean measures, again because errors tended to be distributed among many species, rather than large errors concentrated in a few species. Comparing the two Euclidean measures, percentage errors were consistently less as measured by the arc distance than by the chord distance. This will always be true in the usual case of $0 < e_d < e_{\max}$ because of the curvilinear relationship between the chord distance and the arc distance (see equation above).

Contributions of individual species

Partitioning the total error into components contributed by individual species has potential for giving insights into the sources of error. In most woods, errors were not strongly concentrated into single species (Table 4). In Noe Woods, however, almost half of the error was attributable to *Quercus alba* which increased from a basal area of 12.5 to 14.7 m²/ha as compared to the predicted drop to 12.4 m²/ha. The model did not appear to be consistently over- or underestimate any particular species in this small sample of woods.

Applications of community vectors

The method described here can be used to compare any two compositional vectors sharing a common origin. Some other possible applications are described below.

Vectors can be analyzed in a space with reduced dimensions (e.g. an ordination space of two or three dimensions) instead of the full n -dimensional space defined by abundances of n species. This analysis in an ordination space would shift the emphasis from a complete representation of compositional patterns to an analysis of dominant trends. This reduction in dimensionality has the advantages of suppressing some noise and subordinate trends (Gauch 1982b). However, one must be careful about distortions in the ordination space, since a straight vector through a curved ordination space may misrepresent the pace and direction of change.

The method would be well suited to describing the effects of experimental treatments on community dynamics given 'before' and 'after' data for treatment and control sample units. Geometrically, the problem is analyzing the divergence of treatment and control vectors from a common origin (pre-treatment). Three classes of hypothesis tests are possible: one can examine overall differences between treatment and control vectors, differences in direction only, or differences in rate only.

1. For an overall test incorporating both vector length

and direction, the hypothesis can be posed as no difference between two or more groups of vectors departing from a common origin. This is equivalent to a test of whether the heads of the vectors occupy the same region in n -dimensional species space. A nonparametric multivariate test that is usually appropriate for typical community data (non-normal, many zeros) is the multi-response permutation procedure (MRPP, see Mielke 1984; Biondini, Bonham & Redente 1985; and Zimmerman, Goetz & Mielke 1985). Parametric tests available with multivariate analysis of variance (MANOVA) or multiple discriminant analysis (DA) would be appropriate in some cases.

2. One can isolate and test for differences in direction between treatment and control vectors by first standardizing the vectors to unit length (see under steps 5a and 5b above), then applying one of the techniques just described (MRPP, DA, or MANOVA).

3. Differences in mean length of treatment and control vectors can be compared with a t -test, one-way ANOVA, or corresponding nonparametric techniques.

The method as described here compares only two vectors over one time step. We need methods for describing the trajectories of communities with multiple sequential samples (e.g. see Westman & O'Leary 1986). Such methods need to be developed for application to quantitative descriptors of community response to disturbance: resistance and the components of resilience (Westman 1978; Westman & O'Leary 1986).

Comparing more complex trajectories will require the development of additional descriptors of the relative movements of different samples through species space. For example, the methods given here can be extended to describe differences between vectors not sharing an origin. An additional parameter is needed to describe the distance between the starting points of the vectors. The rate and angles of divergence can be measured by applying the formulas in this paper, after translating the vectors to a common origin.

There is a proliferation of data sets based on permanent plots, models of community dynamics, and long-term field experiments. The growth of data and theory is creating a growing need for a logical and precise language for comparing and describing trajectories in species space. The concept of communities as points moving in species space provides an intuitive basis for further development of that language.

Acknowledgements. Thanks are due to P. Dixon, P. S. Muir, R. K. Peet, and anonymous reviewers for comments and suggestions.

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.
- Austin, M. P. 1977. Use of ordination and other multivariate descriptive methods to study succession. *Vegetatio* 35: 165-175.
- Austin, M. P. 1981. Permanent quadrats: An interface for theory and practice. *Vegetatio* 46: 1-10.
- Barton, J. D. & Schmelz, D. V. 1987. Thirty years of growth records in Donaldson's Woods. *Proc. Indiana Acad. Sci.* 96: 209-214.
- Beals, E. W. 1973. Ordination: mathematical elegance and ecological naivete. *Vegetatio* 13: 65-87.
- Beals, E. W. 1984. Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. *Adv. Ecol. Res.* 14: 1-55.
- Belcher, D. M. 1981. *The user's guide to STEMS*. USDA Forest Service General Technical Report NC-70.
- Belcher, D. M. 1982. TWIGS: a description paper. In: *Microcomputers, a new tool for foresters*. Proc. conf. Purdue University, West Lafayette, Indiana May 18-20, 1982.
- Belcher, D. M., Holdaway, M. R. & Brand, G. J. 1982. *A description of STEMS, the Stand and Tree Evaluation and Modeling System*. USDA Forest Service General Technical Report NC-79.
- Bergeron, Y. & Dubuc, M. 1989. Succession in the southern part of the Canadian boreal forest. *Vegetatio* 59: 51-63.
- Biondini, M. E., Bonham, C. D. & Redente, E. F. 1985. Secondary successional patterns in a sagebrush (*Artemisia tridentata*) community as they relate to soil disturbance and soil biological activity. *Vegetatio* 60: 25-36.
- Blondel, J. & Farré, H. 1988. The convergent trajectories of bird communities along ecological successions in European forests. *Oecologia* 75: 83-93.
- Brisson, J., Bergeron, Y. & Bouchard, A. 1988. Les successions secondaires sur sites mésiques dans la MRC du Haut-Saint-Laurent, Québec. *Can. J. Bot.* 66: 1192-1203.
- Carleton, T. J. & Maycock, P. F. 1978. Dynamics of the boreal forest south of James Bay. *Can. J. Bot.* 56: 1157-1173.
- del Moral, R. & Wood, D. M. 1988. Dynamics of herbaceous vegetation recovery on Mount St. Helens, Washington, USA, after a volcanic eruption. *Vegetatio* 74: 11-27.
- Domon, G., Bouchard, A., Bergeron, Y. & Gauvin, C. 1986. La répartition et la dynamique des principales espèces arborescentes du Bois-de-Saraguay, Montréal (Québec). *Can. J. Bot.* 64: 1027-1038.
- Enright, N. J. 1982. Recognition of successional pathways in forest communities using size-class ordination. *Vegetatio* 48: 133-140.
- Faith, D. P., Minchin, P. R. & Belbin, L. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* 69: 57-68.
- Gauch, H. G. 1982a. *Multivariate Analysis in Community Ecology*. Cambridge University Press, New York.
- Gauch, H. G. 1982b. Noise reduction by eigenvector ordination. *Ecology* 63: 1643-1649.
- Goff, F. G. & Zedler, P. H. 1972. Derivation of species succession vectors. *Am. Midl. Nat.* 87: 397-412.
- Greig-Smith, P. 1983. *Quantitative Plant Ecology*, 3d ed. Blackwell Scientific Publications, Oxford.
- Legendre, L. & Legendre, P. 1983. *Numerical Ecology*. Elsevier, Amsterdam.
- Lepš, J. 1987. Vegetation dynamics in early old-field succession: a quantitative approach. *Vegetatio* 72: 95-102.
- Lindsey, A. A., Barton Jr., J. D. & Miles, S. R. 1958. Field efficiencies of forest sampling methods. *Ecology* 39: 428-444.
- Lippe, E., de Smidt, J. T. & Glenn-Lewin, D. C. 1985. Markov models and succession: a test from a heathland in The Netherlands. *J. Ecol.* 73: 775-791.
- Malanson, G. P. & Trabaud, L. 1987. Ordination analysis of components of resilience of *Quercus coccifera* garrigue. *Ecology* 68: 463-472.
- McClain, W. E. & Ebinger, J. E. 1968. Woody vegetation of Baber Woods, Edgar County, Illinois. *Am. Midl. Nat.* 79: 419-428.
- McCune, B. & Allen, T. F. H. 1985. Forest dynamics in the Bitterroot Canyons, Montana. *Can. J. Bot.* 63: 377-383.
- McCune, B. & Cottam, G. 1985. The successional status of a southern Wisconsin oak woods. *Ecology* 66: 1270-1278.
- McCune, B. & Menges, E. 1986. Historical data quality in Midwestern old-growth forests. *Am. Midl. Nat.* 116: 163-172.
- Mielke Jr., P. W. 1984. Meteorological applications of permutation techniques based on distance functions. In: Krishnaiah, P. R. & Sen, P. K. (eds.) *Handbook of Statistics, Vol. 4*, pp. 813-830. Elsevier, Amsterdam.
- Miner, C. L. & Walters, N. R. 1984. *STEMS: a nontechnical description for foresters*. USDA Forest Service Research Paper NC-252.
- Miner, C. L., Walters, N. R. & Belli, M. L. 1988. *A guide to the TWIGS program for the north central United States*. USDA Forest Service General Technical Report NC-125.
- Newman, J. A. & Ebinger, J. E. 1985. *Woody vegetation of Baber Woods: composition and change since 1965*. Central Hardwoods Conference, Proceedings, April 15-17, 1985, Urbana, Illinois.
- Noy-Meir, I. & Whittaker, R. H. 1977. Continuous multivariate methods in community analysis: some problems and developments. *Vegetatio* 33: 79-98.
- Økland, R. H. 1986. Rescaling of ecological gradients. I. Calculation of ecological distance between vegetation stands by means of their floristic composition. *Nord. J. Bot.* 6: 651-660.
- Orlói, L. 1967. An agglomerative method for classification of plant communities. *J. Ecol.* 55: 193-205.
- Rauscher, H. M. 1986. *The microcomputer scientific software series 4: testing prediction accuracy*. USDA Forest Service, General Technical Report NC-107.
- Reynolds, M. R. & Chung, J. 1986. Regression methodology for estimating model prediction error. *Can. J. For. Res.* 16: 931-938.
- Schlesinger, R. C. 1976. Hard maples increasing in an upland

- hardwood stand. In: Fralish, J. S., Weaver, G. T. & Schlesinger, R. C. (eds.) *Central Hardwoods Conference, proceedings*, pp. 177-185. Carbondale, Illinois.
- Smith, N. J. & Iles, K. 1988. A graphical depiction of multivariate similarity among sample plots. *Can. J. For. Res.* 18: 467-472.
- Smith, T. M. & Urban, D. L. 1988. Scale and resolution of forest structural pattern. *Vegetatio* 74: 143-150.
- Sneath, P. H. A. & Sokal, R. R. 1973. *Numerical Taxonomy*. W. H. Freeman, San Francisco.
- Tessier, L. 1989. Spatio-temporal analysis of climate-tree ring relationships. *New Phytol.* 111: 517-529.
- van der Maarel, E. 1969. On the use of ordination models in phytosociology. *Vegetatio* 19: 21-46.
- Warwick, J. J. & Cale, W. G. 1988. Estimating model reliability using data with uncertainty. *Ecol. Model.* 41: 169-181.
- Webb, R. H., Steiger, J. W. & Turner, R. M. 1987. Dynamics of Mojave desert shrub assemblages in the Panamint Mountains, California. *Ecology* 68: 478-490.
- Westman, W. E. 1978. Measuring the inertia and resilience of ecosystems. *BioScience* 28: 705-710.
- Westman, W. E. & O'Leary, J. F. 1986. Measures of resilience: the response of coastal sage scrub to fire. *Vegetatio* 65: 179-189.
- Whittaker, R. J. 1991. The vegetation of the Storbreen gletschervorfeld, Jotunheimen, Norway. IV. Short-term vegetation change. *J. Biogeogr.* 18: 41-52.
- Zimmerman, G. M., Goetz, H. & Mielke Jr., P. W. 1985. Use of an improved statistical method for group comparisons to study effects of prairie fire. *Ecology* 66: 606-611.

Received 3 March 1991;
Revision received 24 September 1991;
Accepted 30 September 1991.