International Association of Vegetation Science

An Application of Fuzzy Set Ordination to Determine Tree Habitat Suitability of Sites from a Regional Data Set
Author(s): Mark R. Fulton
Source: *Journal of Vegetation Science*, Vol. 7, No. 5 (Oct., 1996), pp. 739-746
Published by: Blackwell Publishing
Stable URL: http://www.jstor.org/stable/3236385
Accessed: 13/04/2011 17:41

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=black.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Blackwell Publishing and *International Association of Vegetation Science* are collaborating with JSTOR to digitize, preserve and extend access to *Journal of Vegetation Science.*

http://www.jstor.org

Journal of Vegetation Science 7: 739-746, 1996
© IAVS; Opulus Press Uppsala. Printed in Sweden

An application of fuzzy set ordination to determine tree habitat suitability of sites from a regional data set

Fulton, Mark R.

Department of Ecology and Evolutionary Biology, MS 170, Rice University, Houston, TX 77005-1892, USA;
Fax: +1 713 285 5232; E-mail: markf@lucybraun.rice.edu

Abstract. Fuzzy set ordination is employed to evaluate sites on the basis of their suitability for particular tree species. The technique orders sites along an axis defined by the presences and absences of a given species of interest. A rationale is given in terms of noise reduction; in many situations the overall vegetation of a site will reflect habitat conditions better than the presence, absence, or quantitative performance of any single species. A data set of tree presence/absence covering a large part of the southeastern United States was analyzed and habitat suitability scores were calculated for each species. Monte-Carlo tests were used to measure the statistical power of the data set with regard to habitat preferences; 38 of the 49 species have cumulative frequency distributions showing significant departures from random expectation. Most statistically significant habitat preferences seem to be related to geographic range limits located within the study area, but some species found throughout the area also show significant departures from random expectation. The method may find applications in autecological studies of species, selection of representative site conditions for simulation modeling, and the solution of certain technical problems in ordination.

Keywords: Fuzzy set theory; Noise reduction; Scale.

Nomenclature: Burns & Honkala (1990).

Abbreviations: FHM = Forest Health Monitoring program; USDA = United States Department of Agriculture; USEPA = United States Environmental Protection Agency; TVA = Tennessee Valley Authority.

Introduction

Noise reduction is a central problem of data analysis, appearing in the treatment of almost all kinds of data in almost every kind of science. However, the definition of noise depends on context; noise is information that is not relevant to the study of a particular problem (Cook et al. 1990). The general interest of vegetation ecologists is in coordinated, predictable patterns of species abundance. Noise in vegetation data arises because the presence, absence, or abundance of a species at a particular site may strongly reflect unpredictable historical factors such as small disturbances and accidents of dispersal (Gauch 1982; Minchin 1987), or spatial proximity (Shmida & Ellner 1984). Unpredictable factors seem to be more important when the sampling units are small (Reed et al. 1993). The methods of ordination and classification are partly intended to uncover meaningful vegetation patterns in the face of many sources of noise (Gauch 1982).

Noise can also arise in an autecological context. One of the interests of autecologists, and the main concern of this study, is the evaluation of the habitat preferences of individual species within a specified study area. Suppose a data set has been collected consisting of species abundance and environmental measurements at a set of sites. One approach to evaluating habitat preferences would be to simply summarize the environment of the sites at which the species of interest is abundant. However, because the performance of any individual species at a given site may be affected by many factors unrelated to habitat suitability, this approach is subject to a large component of noise. Existing knowledge of habitat requirements may be applied, but such knowledge is scattered among many different sources, often has large subjective elements, and can be difficult to quantify. Another approach uses one of the fundamental generalizations of vegetation science: the fact that the overall species composition of a site is often a better indicator of habitat conditions than the performance of any single species. This is the approach used in the present study.

Suppose the species of interest is northern red oak (*Quercus rubra* L.). If the species composition of a site is similar to the species composition of sites that contain red oak and it is dissimilar from sites that do not contain red oak, then that site is likely to be a suitable habitat for red oak, whether or not red oak is actually present at the site. Quantifying this concept is an ordination problem, and fuzzy set ordination (Roberts 1986) provides a straightforward solution. Fuzzy set ordination allows the quantification of concepts like 'similar to sites with red oak present, while not similar to sites with red oak absent' in a mathematically rigorous way. The procedure is to evaluate the habitat suitability of a site for a given species by using the score along an ordination axis defined by the presences and absences of that species. The environment of sites with high habitat scores can then be taken as the typical habitat of the species of interest within the area sampled by the data set.

The detection and interpretation of 'pattern' when there is no pattern is a recurring problem in the use of ordination methods. The evaluation of habitat scores is no exception; there is nothing to prevent the calculation of a range of habitat scores for a species occurring at random within a data set. To account for this possibility, the distribution of habitat scores should be compared to distributions of scores from data in which occurrences of the species of interest are shuffled among the sites. The intention is not to identify species with no habitat preferences (such species are not widely believed to exist), but rather to ascertain the statistical power of a particular data set to resolve habitat preferences within a particular study area.

This paper demonstrates the use of fuzzy set ordination to evaluate habitat suitability of sites from a data set of forest vegetation covering a large area in the southeastern United States. The relation of actual species presence and absence in relation to the habitat scores is examined to check the internal consistency of the method. Randomization tests are used to identify statistically significant patterns. The results are discussed in relation to the geographic ranges of the species included.

Methods

Data compilation

The data used for this study were collected as a demonstration of the Forest Health Monitoring program (FHM) conducted by the TVA, the USDA Forest Service and the USEPA in the southeastern United States. Details of sampling procedures are given by Conkling & Byers (1993). A data set of tree species and dbh (diameter at breast height) by sites was extracted from the data base. Each site is located near the center of a 40 km^2 hexagon that is part of a grid extending over the entire United States. The area sampled by the data in this study includes the eastern Tennessee River Valley; the southern Appalachian mountains; and adjacent parts of the

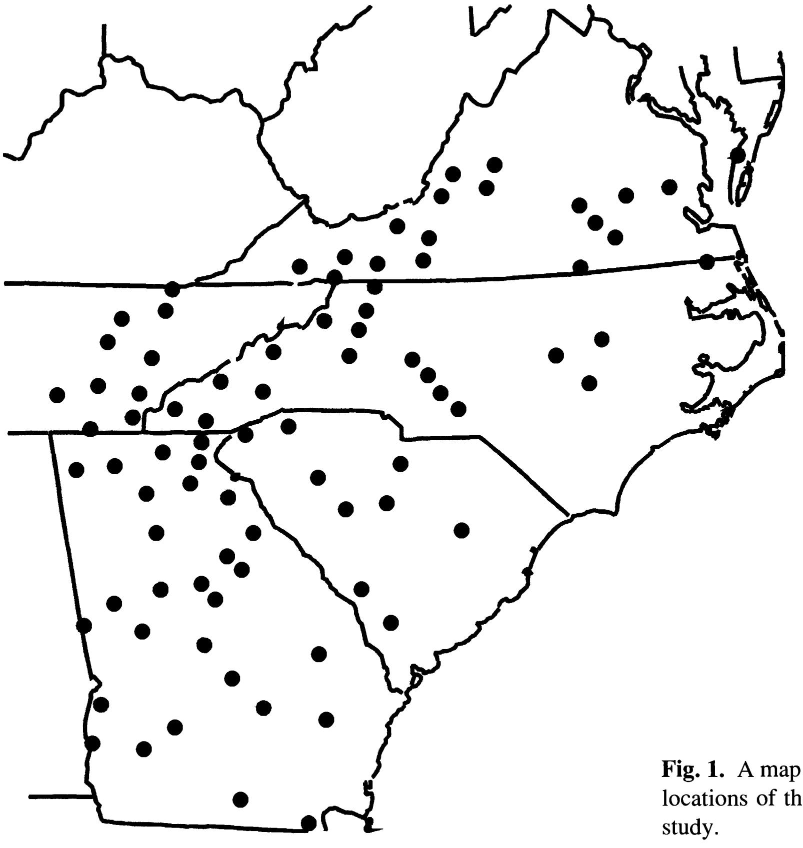

Fig. 1. A map of the southeastern United States showing the locations of the Forest Health Monitoring plots used in this study.

piedmont and coastal plain (Fig. 1). The FHM project is ongoing, so the coverage of this area by the current data set is not complete. In all, 103 sites were included in the data for this study. Each site consists of four 1/60-ha (7.3 m radius) circular subplots. There is one central subplot and the centers of the three remaining plots are each located 36.6 m from the central subplot in a fixed arrangement. Plots were established and measured in 1992 and 1993. The tree species and dbh data were transformed to a matrix of presence/absence. After deletion of species with fewer than three occurrences, the resulting data set contained 49 tree species by 103 sites. Habitat scores were calculated for all 49 species.

The area encompassed by the FHM data is extensive, and almost half of the species included in the data set have range limits within the study area. It was of interest to determine whether the range of habitat scores calculated simply reflect non-overlapping geographic distributions, or whether within-range habitat preferences can be resolved. Therefore, range limit data for the included species were summarized by comparing a map of the study area to maps of range distributions in Burns & Honkala (1990). Each species was scored according to whether the range overlapped the middle of the study area or was confined to the NW or SE parts, and how large a proportion (to the nearest 1/5th) of the study area was overlapped by the range.

Calculation of habitat scores

The equations and rationale for fuzzy set ordination are given by Roberts (1986). For each species k, two complementary fuzzy sets are defined: (1) the set of sites similar to sites containing species k, and (2) the set of sites similar to sites that do not contain species k. Similarity is calculated with Sørensen's index:

$$S_{ij} = \left(2\sum x_{ik}x_{jk}\right) / \left(\sum x_{ik} + \sum x_{jk}\right) \tag{1}$$

where S_{ij} is the similarity of site i to site j, x_{ik} is the presence or absence (1 or 0) of species k in site i, and the summations are over all species k. Both empirical experience (Beals 1984) and simulation results (Faith et al. 1987) favor the use of $1 - S_{ij}$ (usually called the Bray-Curtis measure when used with quantitative data) as a good general measure of ecological distance. Membership in the set 'similar to sites containing species k' is the mean of the similarities to sites at which the species k is present:

$$\mu_k(i) = \sum S_{ij} x_{jk} \Big/ \sum x_{jk} \tag{2}$$

where $\mu_k(i)$ is the membership function for site i and species k, and the summations are over all sites $j \neq i$. Membership in the set 'similar to sites not containing species k' is the mean of similarities to sites at which species k is absent ($\sim k$):

$$\mu_{\sim k}(i) = \sum S_{ij}\left(x_{jk} - 1\right) \Big/ \sum \left(x_{jk} - 1\right) \tag{4}$$

where the summations are over all sites $j \neq i$. The habitat score for a given site and species is the membership in the set "similar to sites containing the species, while not similar to sites not containing the species". Membership in this set is calculated using the anticommutative difference operator (Roberts 1986):

$$H_k(i) = \left[1 + \left(\mu_k(i)\right)^2 - \left(\mu_{\sim k}(i)\right)^2\right] \Big/ 2 \tag{5}$$

$H_k(i)$ will be referred to as the habitat score of site i for species k. As pointed out by Roberts (1986), this formula orders sites along a Bray-Curtis axis with synthetic weighted endpoints. Distributions of $H_k(I)$ were examined using plots of cumulative frequency distributions after the data were sorted from low to high values; the x-axis is the rank of the score and the y-axis is the score.

It will be noted that the data on the species of interest is included in the calculation of similarities used to define the fuzzy sets. This is not circular; the goal is a scalar representing the suitability of a particular site as habitat for a particular species, as such, the presence or absence of that species is certainly relevant. The fact that a random assignment of presences and absences of a given species can also produce a range of habitat scores is addressed by the randomization tests described below.

Randomization tests

The critical assumption behind this method is that the overall species composition of a site contributes information about the likely success of a given species. If a data set resolves meaningful information about the habitat preferences of a species, then the distribution of habitat scores should differ from the distribution of scores calculated from data sets in which the species is allocated to the appropriate number of sites at random. The randomization procedure and the comparisons made are described below.

If the data are in a sites (rows) X species (columns) matrix, a randomization is achieved by shuffling the column of data representing the species of interest. The intention is to sample (with replacement) from the set of all possible permutations of data within the column of the species of interest, while holding the other data constant. For most vegetation data sets the number of possible permutations is too large to enumerate exhaustively, so sampling by Monte-Carlo randomization is used. The dissimilarity matrix and habitat scores are recalculated for each randomization. One thousand randomizations were carried out for each of the 49 species, giving a total of 49 049 ordinations calculated for the 103 sites.

To indicate how the presences and absences of species k are sorted out along the axis defined by $H_k(i)$, the following index was calculated:

$$D_k = \left(1/N_p\right)\sum H_k(i) * x_{ik} - \left(1/\left(N - N_p\right)\right)\sum H_k(i) * \left(1 - x_{ik}\right) \quad (6)$$

where N is the number of sites, N_p is the number of sites with species k present, and the summations are over all sites $i = 1$ to N. This is the difference in mean habitat scores between sites with species k present and sites with species k absent. The index D_k(random) was calculated for each data permutation. The number of D_k(random) greater than or equal to D_k was tabulated. The ratio of this number to the number of permutations is an estimate of the statistical significance of D_k under the null hypothesis that D_k is no greater than expected for a random permutation of the data for species k.

One expected feature of vegetation data is that a species is likely to be found in a relatively restricted part of the overall data space; there should be a preponderance of high and/or low habitat scores compared to data in which the species is allocated to the appropriate number of plots at random. This can be tested directly by comparing the cumulative frequency distribution of habitat scores with the cumulative frequency distributions of scores calculated from partially randomized data. A Monte-Carlo 95 % confidence envelope was calculated from the 1000 randomizations for each species. The habitat scores for each randomization were ranked from smallest to largest. The 95 % confidence envelope was constructed from the 25th largest and 25th smallest scores from each rank; for a 99 % confidence envelope the 5th largest and 5th smallest scores would be used.

Calculations were carried out with a FORTRAN program written by the author. Randomization tests are computer intensive; the analysis for this study took about 2 hours on a fast IBM-compatible PC equipped with a Pentium processor running at 120 MHz.

Results

Habitat scores ranged between 0.24 and 0.81. The method is internally consistent; the mean score difference (D_k) was always positive (Table 1), indicating that a species always tends to be found in sites with high habitat scores. In 18 of the 49 species, D_k was significantly larger than expected under a null hypothesis of random assignment of species occurrences. The cumulative distributions of habitat scores seem to be more sensitive than D_k; only 11 of the 49 species failed to show some departure from their 95% confidence envelopes. The species showing no departure from their confidence envelopes had low frequencies (<12), except for *Prunus serotina* (frequency = 22); the width of the confidence envelope was a function of the frequency of the species (mean C.I. width = 0.302/frequency + 0.05, $R^2 = 0.98$). With the exception of *Ilex opaca*, habitat score distributions depart from their confidence envelopes by having low scores lower than the envelope and/or high scores higher than the envelope (Fig. 2).

Most species are completely absent from sites with low habitat scores and are present at sites with high habitat scores (Fig. 2). There is usually little overlap between site scores with a species present and site scores with a species absent. Some of the exceptions are considered in the discussion.

22 species are found throughout the FHM study area (Table 1), of these, three show a significant D_k and 16 show departures from the 95% confidence envelope. Of the remaining 27 species with more restricted distributions, 15 show a significant D_k and 23 show departures from the 95% confidence envelope.

Discussion

Fuzzy set ordination provides a way to apply knowledge of vegetation patterns to quantitatively evaluate sites in terms of their suitability as habitat for particular species; it is a synecological approach to an autecological problem. When used in this way, fuzzy set ordination is a data re-scaling operation (Allen & Starr 1982; Allen et al. 1984); the grain (resolution) of the abundance data is coarsened to presence/absence, and the grain of the presence/absence data is coarsened to better resolve large-scale patterns of habitat preference. As in any data re-scaling operation, there is information lost and information gained. The information lost is knowledge of the presence or absence of a species in any particular site. The information gained is an indication of the suitability of the site based on the entire vegetation composition.

The usual pattern of habitat scores seen in this study, high scores higher than random expectation and/or low scores lower than random expectation, is consistent with the usual way a species is expected to be distributed among sites. In a regionally extensive data set such as

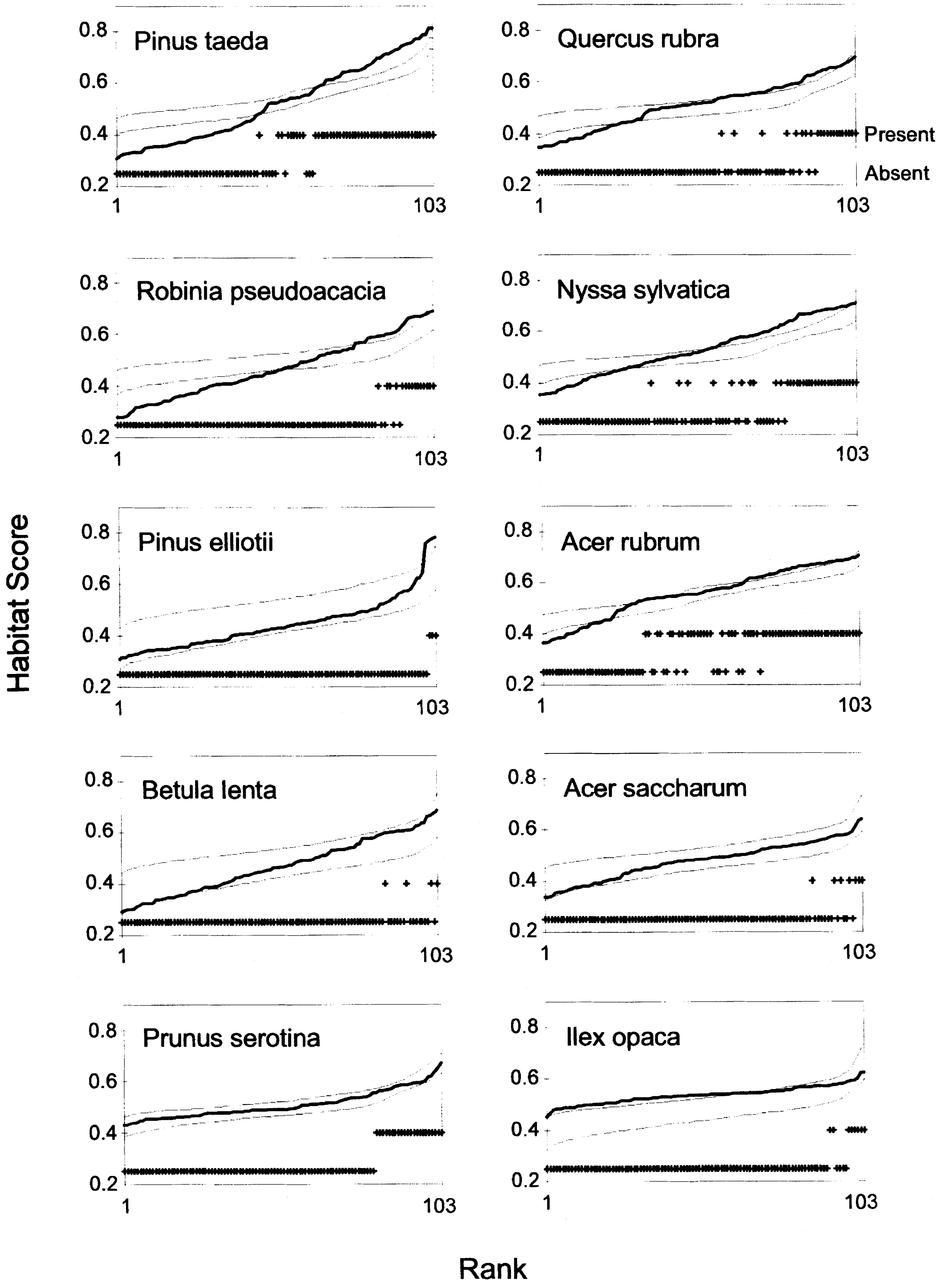

Fig. 2. Cumulative distributions of habitat scores for selected species from the original data (thick lines) and the Monte-Carlo confidence envelopes (thin lines). '+' symbols indicate the presence or absence of the species along the gradient of $H_k(i)$.

Table 1. Results of randomization tests for each species in the Forest Health Monitoring data set, along with the number of occurrences within the data (frequency), and geographic range information. The column 'envelope departure' indicates the nature of the departure of the habitat scores $H_k(i)$ from the Monte-Carlo confidence envelope; none indicates no departure (the cumulative distribution is entirely within the confidence envelope), high indicates high scores for a given rank, low indicates low scores for a given rank. Bold values in column P-value indicate significance at $P = 0.05$.

Species	Frequency	D_k	P-value of D_k	Envelope Departure	Geographic Range	FHM Area Covered
Juniperus virginiana	10	0.135	0.112	None	Central	1
Pinus echinata	32	0.105	0.640	High	Central	1
P. elliotii	3	0.339	**0.000**	High	SE	0.2
P. palustris	5	0.250	**0.006**	High, Low	SE	0.4
P. pungens	3	0.275	**0.018**	Low	NW	0.2
P. rigida	3	0.213	0.057	Low	NW	0.2
P. strobus	15	0.197	**0.000**	High, Low	NW	0.2
P. taeda	48	0.254	**0.000**	High, Low	SE	0.6
P. virginiana	14	0.100	0.598	High	NW	0.4
Tsuga canadensis	4	0.208	0.052	High	NW	0.2
Acer pensylvanicum	3	0.300	**0.014**	High, Low	NW	0.2
A. rubrum	57	0.147	**0.000**	High, Low	Central	1
A. saccharum	7	0.126	0.242	None	NW	0.2
Betula lenta	4	0.177	0.103	Low	NW	0.2
Carya spp.	15	0.118	0.217	High, Low	Central	1
C. glabra	12	0.126	0.152	High, Low	Central	1
C. ovata	3	0.078	0.613	None	NW	0.6
C. tomentosa	17	0.118	0.212	High, Low	Central	1
Castanea dentata	4	0.162	0.139	None	NW	0.4
Cercis canadensis	3	0.144	0.240	High	NW	0.8
Cornus florida	31	0.116	0.219	High	Central	1
Diospyros virginiana	5	0.132	0.206	None	Central	1
Fagus grandifolia	7	0.119	0.269	None	Central	1
Fraxinus pennsylvanica	5	0.077	0.731	None	Central	1
Ilex opaca	8	0.064	0.981	High	Central	1
Liquidambar styraciflua	36	0.157	**0.001**	High, Low	SE	0.8
Liriodendron tulipifera	33	0.108	0.478	High	Central	1
Magnolia spp.	3	0.162	0.185	High	Central	1
M. acuminata	4	0.179	0.075	High, Low	NW	0.4
M. virginiana	4	0.280	**0.007**	High	SE	0.4
Nyssa sylvatica	33	0.164	**0.001**	High, Low	Central	1
N. sylvatica var. *biflora*	4	0.334	**0.001**	High	SE	0.4
Oxydendrum arboreum	30	0.144	**0.002**	High, Low	NW	0.6
Platanus occidentalis	5	0.154	0.110	High	Central	1
Prunus serotina	22	0.107	0.465	None	Central	1
Quercus alba	41	0.149	**0.001**	High, Low	Central	1
Q. coccinea	28	0.208	**0.000**	High, Low	NW	0.6
Q. falcata	28	0.116	0.238	High	SE	0.8
Q. marilandica	4	0.103	0.470	None	Central	1
Q. nigra	17	0.158	**0.007**	High, Low	SE	0.8
Q. phellos	7	0.104	0.473	High	Central	0.4
Q. prinus	24	0.214	**0.000**	High, Low	NW	0.4
Q. rubra	21	0.151	**0.010**	High, Low	NW	0.6
Q. stellata	11	0.099	0.577	None	Central	1
Q. velutina	21	0.113	0.267	High	Central	1
Robinia pseudoacacia	15	0.212	**0.000**	High, Low	NW	0.4
Sassafras albidum	10	0.117	0.277	High, Low	Central	1
Ulmus alata	6	0.110	0.369	None	Central	0.4
U. americana	4	0.102	0.487	High	Central	1

this one, most species will be found in a relatively restricted set of sites with a relatively restricted set of associates. If the vegetation data are visualized with an ordination diagram, most species are restricted to some part of the data space, and most species respond to approximately the same set of site factors. If every species responded to a unique set of environmental variables, ordination methods could not successfully reduce the dimensionality of vegetation data and this method of calculating habitat scores would not work. In the FHM data, the habitat scores appear to resolve largely geographic range differences among the species; most of the species with restricted distributions show departures from random expectation. However, the fact that some species found throughout the region also show non-random patterns in the distribution of $H_k(i)$ shows that the method also uncovers more local habitat preferences.

The distribution of $H_k(i)$ for *Ilex opaca* is an exception to the rule discussed above; the low scores are higher than the confidence envelope, D_k is less than 98.1% of D_k(random), and the scores are very tightly clustered around an intermediate value. This indicates that *I. opaca* is more widely dispersed through the data space than would be expected even for a random allocation to sites. The answer to the puzzle may lie in the autecology of the species. *I. opaca* is capable of growing in the understory of a wide range of forests and has very few site restrictions within the study area other than intolerance of extensive flooding (Burns & Honkala 1990). It is, however, extremely sensitive to fire and its seeds are animal dispersed. Its occurrence may be more related to recent fire history and accidents of dispersal than to more stable site factors, such as soil and climate, to which the other species respond. A species like *I. opaca*, which responds to different habitat characteristics from most of the other taxa in a data set, is not well suited to this application of fuzzy set ordination.

Two species, *Nyssa sylvatica* and *Acer rubrum*, show some evidence of a bimodal distribution of presences along the gradients defined by $H_k(i)$. This may be an indication of bimodality in the data space; both species are notable for having a broad ecological amplitude. *Acer rubrum* is found in both uplands and lowlands in the northern part of its range, and is exclusively a bottomland species in the southern part of its range (Burns & Honkala 1990). The FHM region is extensive enough to incorporate both of these associations. *Nyssa sylvatica* is also found in a wide range of moisture conditions. Keeley (1979) showed that seedlings of *N. sylvatica* from different habitats have very different physiological mechanisms of flood tolerance, and that these differences are genetic. The question of what is an ecological species, as opposed to a taxonomic or genetic species, is still open to meaningful debate.

Habitat scores have several possible practical applications. The most obvious application is for the objective evaluation of the habitat preferences of a tree species. Correlations or regressions of $H_k(i)$ against particular environmental factors may provide a useful evaluation of the realized niche of a species in situations where the raw species performance is subject to large amounts of noise. Habitat scores can be used to identify sites with habitat characteristics that are typical of a species within the area sampled by a data set, thus providing an objective way to choose values of environmental variables appropriate for simulation modeling of a particular species of interest.

Habitat scores may also be a solution to a common technical problem in community ordination. The problem is that a zero in site X species data matrix conveys no information about the suitability of that site for that species; in data sets spanning long gradients the abundance of these zeros leads to distorted patterns in ordinations (Swan 1970). Beals (1984) refers to this as the zero truncation problem, and advocates replacing the raw species performance data with the 'sociological favorability index':

$$b_{ik} = (1/S)\sum N_{kj}/N_j \qquad (7)$$

where b_{ik} is the value of the index for site i and species k, S is the number of species at site i, N_{kj} is the number of sites with both species k and j, N_j is the number of sites with species j, and the summation is over all species j. Informal tests of this index for habitat suitability evaluation showed it to be somewhat less suitable than fuzzy set ordination; presences and absences of species k were less well separated along the gradient defined by b_{ik} than along the gradient defined by $H_k(i)$. It is not known whether $H_k(i)$ would perform better than b_{ik} for correcting distortion in ordinations over long gradients.

Randomization tests provide a check of whether ordinations are contributing useful information. Most scientists have had the experience of seizing upon an interesting pattern that later proved to be meaningless. The tests may also be seen as a measure of the statistical power of a vegetation data set with regard to the habitat preferences of particular species. Checks of statistical power and significance are under-used in vegetation ecology. Formal hypothesis testing is not the only use of such tests; they can allow the ecologist to distinguish potentially meaningful patterns from noise.

Acknowledgements. Mike Kelly and the Tennessee Valley Authority provided a congenial working environment. Betsy Smith, Brian Cordova, and Dennis Yankee provided data and technical assistance. Comments by and discussions with Jari Oksanen, David Roberts, Rosine Hall, and an anonymous reviewer considerably improved an earlier version of this manuscript. This research was supported in part by an appointment to the Postgraduate Research Program at the Tennessee Valley Authority administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and TVA.

References

Allen, T.F.H. & Starr, T.B. 1982. *Hierarchy; perspectives for ecological complexity.* University of Chicago, Chicago, IL.

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

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

Burns, R.M. & Honkala, B.H. (tech. coords.) 1990. *Silvics of North America. Agriculture Handbook 654.* U.S. Department of Agriculture, Forest Service, Washington, DC.

Conkling, B.L. & Byers, G.E. (eds.) 1993. *Forest health monitoring field methods guide.* Internal Report. U.S. Environmental Protection Agency, Las Vegas, NV.

Cook, E.R., Briffa, K., Shiyatov, S., Mazepa, V. & Jones, P.D. 1990. Data analysis. In: Cook, E.R. & Kairiukstis, L.A. (eds.) *Methods of dendrochronology*, pp. 97-162. Kluwer, Dordrecht.

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. 1982. *Multivariate analysis in community ecology.* Cambridge University Press, Cambridge.

Keeley, J.E. 1979. Population differentiation along a flood frequency gradient: physiological adaptations to flooding in *Nyssa sylvatica*. *Ecol. Monogr.* 49(1): 89-108.

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

Reed, R.A., Peet, R.K., Palmer, M.W. & White, P.S. 1993. Scale dependence of vegetation-environment correlations: A case study of a North Carolina piedmont woodland. *J. Veg. Sci.* 4: 329-340.

Roberts, D.W. 1986. Ordination on the basis of fuzzy set theory. *Vegetatio* 66: 123-131.

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

Swan, J.M.A. 1970. An examination of some ordination problems by use of simulated vegetational data. *Ecology* 51(1): 89-102.

Received 11 April 1995;
Revision received 17 March 1996;
Final revision received 18 May 1996;
Accepted 15 April 1996.