COENOSES 5 (2): 79-87, 1990

SPATIAL SCALE AND PATTERNS OF SPECIES-ENVIRONMENT RELATIONSHIPS IN HARDWOOD FOREST OF THE NORTH CAROLINA PIEDMONT

M.W. Palmer, Department of Botany, Duke University, Durham NC 27706 USA
Current address: Department of Botany and Microbiology, Oklahoma State University, Stillwater OK 74078-0289 USA

Keywords: Canonical correspondence analysis, Environmental variability, Geostatistics, Gradient analysis, Self-similarity, Spatial scale

Abstract. Plant-environment relationships in the Duke Forest, Durham and Orange Counties, North Carolina USA were examined within and among tenth hectare plots. Soil samples collected from subplots were analyzed for phosphorus, potassium, calcium, sodium, weight/volume, pH, buffered acidity, cation exchange capacity, manganese, copper, and humic acids. Canopy openness was estimated with a spherical densiometer. Canonical correspondence analysis reveals that the species composition of 0.1 ha plots is strongly related to soil calcium. All of the environmental variables demonstrate some spatial dependence within plots, indicating that within-plot gradient analysis must be interpreted cautiously. Species composition of 2 m^2 subplots is weakly related to soil calcium. Species positions along the calcium gradient within plots are weakly correlated with their positions along the among-plot calcium gradient.

Introduction

Direct gradient analysis, or the study of the relationship between plant community composition and environmental variation, is one of the cornerstones of quantitative plant community ecology. Direct gradient analysis contrasts with indirect gradient analysis, which includes most ordination techniques. Ter Braak and Prentice (1988), in their classification of the most widely used gradient analysis techniques, stated the main distinction between direct and indirect techniques (p. 304): ''Direct gradient analysis allows one to study the part (large or small) of the variation in community composition that can be explained by a particular set of environmental variables. In indirect analysis the attention is first focused on the major pattern of variation in community composition; the environmental basis of this pattern is to be established later.'' Recently, methodology for *direct* gradient analysis has greatly advanced with the development of Canonical Correspondence Analysis, or CCA (ter Braak 1986), which is an extension of Correspondence Analysis, a commonly used *indirect* gradient analysis.

Despite the importance of direct gradient analysis, little attention has been paid to dependence of plant-environment relationships on spatial scale. If there is an important, identifiable environmental gradient influencing species composition at one spatial scale, can an ecologist make reasonable predictions about this gradient at other scales? If the same patterns are observed on all spatial scales, then the system can be considered self-similar, or statistically self-similar. Ideas concerning self-similarity are closely related to fractal geometry (Mandelbrot 1983, Phillips 1985, Palmer 1988); strictly defined, a self-similar object has a constant fractal dimension. However, a less formal definition of self-similarity is perhaps of more interest to ecologists: a system is self-similar if important patterns and processes are observed and are dominant across spatial scales.

There are several factors which should be taken into account when considering spatial scale in the context of direct gradient analysis. Ecologists often assume that the environment is homogeneous at the scale of a plot. However, the scale dependence of homogeneity is rarely evaluated. For comparisons between spe cies composition and environment to be statistically valid, plots must be spatially independent replicates. Again, an evaluation of spatial independence is rarely performed. Direct gradient analysis is impossible if environmental variation within plots equals environmental variation among plots. This also is seldom tested.

Fortunately, scale dependence of homogeneity, spatial independence, and comparisons of within- and among-plot environmental variability can all be evaluated by inspecting the form of the semivariogram, a diagram which is essential to the discipline of geostatistics (Journel and Huijbregts 1978, Clark 1979, Webster 1985).

In this paper, I perform geostatistical analyses on environmental variables within 0.1 ha plots in hardwood stands of the Duke Forest. I also report direct gradient analyses on plant community variation among and within 0.1 ha plots, and contrast the results of large-scale (among-plot) plant-environment relationships to small-scale (within-plot) plant-environment relationships.

Methods

Vegetation. I estimated the percent cover of plant species in each of forty 2 m^2 subplots in each of thirty 0.1 ha plots in hardwood forest stands of the Duke Forest, Durham and Orange Counties, North Carolina, as described in the previous paper (Palmer 1990). As in Palmer (1990), the "floristic" data set is comprised of species presence and absence; the "vegetation" data set is comprised of species cover estimates.

Soil. In the summer of 1986, I collected soil samples from each subplot in each plot. Each sample consisted of four combined cores from the top 10 cm of the soil profile and was collected with a stainless steel soil auger. The four cores were located 50 cm from the center of each subplot, with one core in each of the four major compass directions. The soil samples were sifted and analyzed by the Soil Testing Service of the North Carolina Department of Agriculture for phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), sodium (Na), weight to volume ratio (W/V), pH, buffered acidity (BA), cation exchange capacity (CEC), manganese (Mn), copper (Cu), and humic acid (HA). P, K, Ca, Mg, Na, Mn and Cu were extracted by the Mehlich III method (Mehlich 1984a). Humic acid was extracted according to Mehlich (1984b). Soil pH and buffered acidity were measured in aqueous solution. Chemical concentrations are in parts per million, and weight/volume is in grams per cubic centimeter. Percent canopy openness (CO) was estimated with a hand-held spherical densiometer (Lemmon 1956).

In order to assess the variation of soil parameters on a very small scale (< 1 m), I randomly selected five of the tenth hectare plots for more intensive sampling. One soil sample (consisting of two adjacent cores) was taken from the geometric center of each of the five plots. From this center, four 1 m transects were established; two were in opposite directions and parallel to the long axis of the plots, and two were in opposite directions and perpendicular to the long axis of the plots. A soil sample consisting of two adjacent cores was taken at every 10 cm interval along these transects.

The frequency distributions of several variables (P, K, Ca, Mg, CEC, Mn, Cu, HA) were found to be extremely right-skewed; these variables were therefore log-transformed before subsequent analysis. Sodium occurred as identical trace quantities within several plots; there was thus no recorded variance in Na in these plots. The sodium data were omitted because gradient analysis is impossible when the environmental value does not vary.

Geostatistics. Recently, ecologists have turned to geostatistics to describe spatial variation (Phillips 1985, Robertson 1987, Palmer 1988). The central concept of geostatistics is the semivariance, which describes the variation in a variable of interest (the "regionalized variable") as a function of spatial scale. The semivariance (γ) at a given distance h is defined as:

$$\gamma_h = \frac{1}{2\,N_h} \sum_{i=1}^{N_h} (z\,(i) - z\,(i+h))^2 \qquad \text{Eq. 1}$$

where z (i) is the value of the regionalized variable for location i, z (i+h) is the value of the regionalized variable in a location separated from location i by distance h, and N_h is the number of pairs of locations separated by distance h. A semivariogram is a plot of semivariance versus h. Although semivariograms are often used to display patterns in data from plots placed systematically along a transect (*e.g.*, Palmer 1988) they can be used for any sample of plots with known spatial coordinates (Webster 1985).

In this study, an overall semivariogram for each of the environmental variables was calculated by averaging the corresponding semivariograms of all 30 plots (*i.e.*, for each lag distance h, the sum of the semivariances of all plots was divided by thirty to produce an average). Smaller-scale semivariograms were calculated by averaging the semivariograms of the five plots randomly selected for the small-scale survey. One major drawback of averaging (and all other forms of data summary) is that some detail may be lost. For example, differences among plots will not be visible. However, averaging was deemed necessary here for two reasons: 1) 40 subplots per plot do not yield a large enough sample of pairs of points to adequately calculate semivariance, and 2) because I am interested in summarizing the general small-scale patterns, not emphasizing the particular patterns. Since there are roughly equal numbers of pairs of points per lag distance in all plots, the average semivariogram should be extremely close to a single semivariogram calculated by Eq. 1 using data from all plots (see Webster 1985, p. 12).

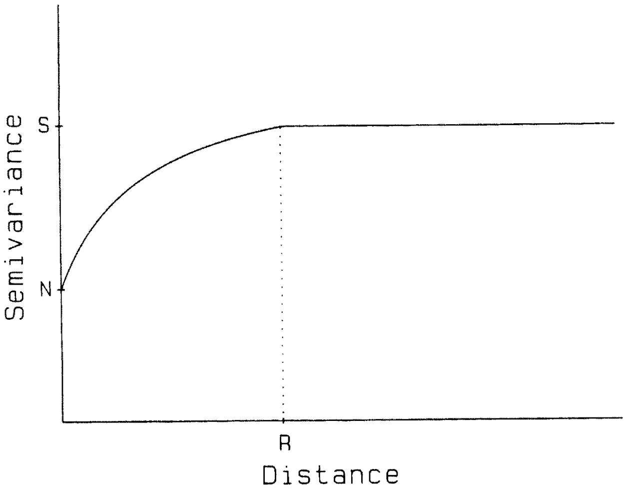

Fig. 1. An idealized semivariogram. S is the sill, N is the nugget, and R is the range.

The semivariogram can reveal whether there is spatial dependence in environmental variables. This is important because the existence of spatial dependence violates the standard statistical assumption that replicates are independent. Spatial dependence is one manifestation of pseudoreplication (Hurlbert 1984). Although statistical methods which account for spatial dependence have been developed (Fingleton 1986, Cook and Pocock 1983, Legendre *et al.* 1987, Wilkinson *et al.* 1983), these are not directly applicable to gradient analysis.

In the idealized semivariogram (Fig. 1), the semivariance increases as a function of distance until a sill is reached. If locations are separated by a distance less than the distance at which the sill is reached (the range), they cannot be considered statistically independent replicates.

The range of the semivariogram (Fig. 2) indicates the spatial scale above which the environment can be considered homogeneous. If the range is smaller than the dimensions of the plot, then the plot is homogeneous with respect to the environmental variable in question. Note that this concept of homogeneity does not mean lacking variability. In this case, homogeneity means lacking overall spatial trends. For example, a checkerboard is homogeneous with respect to color because one side has as much red and black as the opposite side. However, a section of the checkerboard consisting of two adjacent squares would not be homogeneous because it is entirely red on one side and entirely black on the other.

Analysis of the relationship between species composition and environmental variables within a plot is only possible if variability among subplots is greater than variability within subplots. For within-plot gradient analysis in this study, the semivariance at 1.60 m (the diameter of the subplots) should be substantially lower than the semivariance at 50 m (the length of the plot). If the semivariogram is completely horizontal, then subplots of any size would contain the same environmental variance as the entire plot and would be inappropriate for discerning intraplot plant-environment relations.

Gradient analysis. Direct gradient analysis is the process of comparing the species composition of ecological communities to characteristics of the environment. The measured environmental variables need not be direct causal agents of species composition. For example, the best known examples of gradient analysis involve species distributions along elevational gradients (*e.g.*, Whittaker 1956). It is unlikely that plant species are directly affected by elevation; indirect consequences of elevation such as temperature, moisture, and soil type are more likely to determine species distributions. Just as linear regression does not require an explanatory variable to explain most of the variance of a dependent variable, gradient analysis does not require that the most important environmental variables are included. For example, nitrogen availability may be an important determinant of understory species composition in the Duke Forest. However, excluding nitrogen from this study does not invalidate gradient analysis; it merely limits the number of environmental gradients we can expect to find.

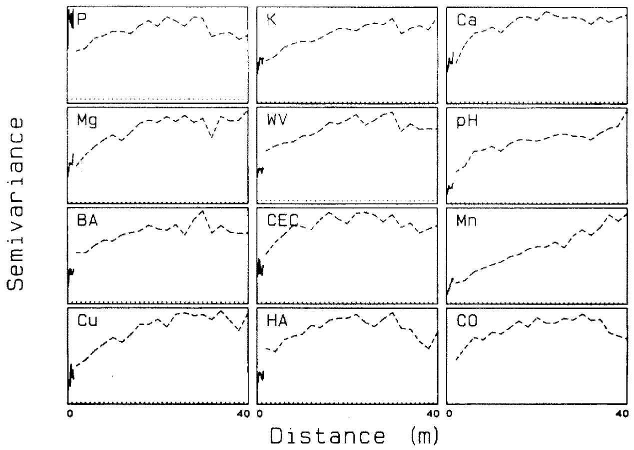

Fig. 2. Semivariograms for twelve environmental variables. Dashed lines are calculated from environmental data from subplots, and solid lines are calculated from the very small scale soil survey (not calculated for canopy openness). The horizontal dotted lines represent half the variance of 30 replicate laboratory analyses on the same soil sample. The minimum values on the semivariograms are zero; the maximum values are as follows: Phosphorus (P), 0.2; Potassium (K), 0.25; Calcium (Ca), 0.9; Magnesium (Mg), 0.3; Weight/Volume (WV), 0.011; pH, 0.33; Buffered acidity (BA), 1.25; Cation exchange capacity (CEC), 0.11; Manganese (Mn), 1.2; Copper (Cu), 0.065; Humic acid (HA), 0.042; Canopy openness (CO), 42.

Canonical correspondence analysis (CCA, ter Braak 1986, 1987a, b, c) is a new technique for multivariate direct gradient analysis. CCA is derived from the multivariate indirect gradient analysis (ordination) technique known as Correspondence Analysis (CA, Hill 1974). Both CA and CCA derive a set of ordination axis scores for species as well as sites. For the first ordination axis, species scores and site scores are assigned such that the correlation between the two is maximum. For subsequent axes, the site and species scores are also maximally correlated, but constrained to be uncorrelated to site scores and species scores along previous ordination axes. In CCA, the site scores are also constrained to be simple functions (*i.e.*, linear combinations) of the inputted environmental variables. The ordination diagram for CCA includes not only the species and site scores, but vectors representing environmental variables. From this diagram, one can determine which environmental variables explain the dispersion of species along the major ordination axes. The significance of the relationship

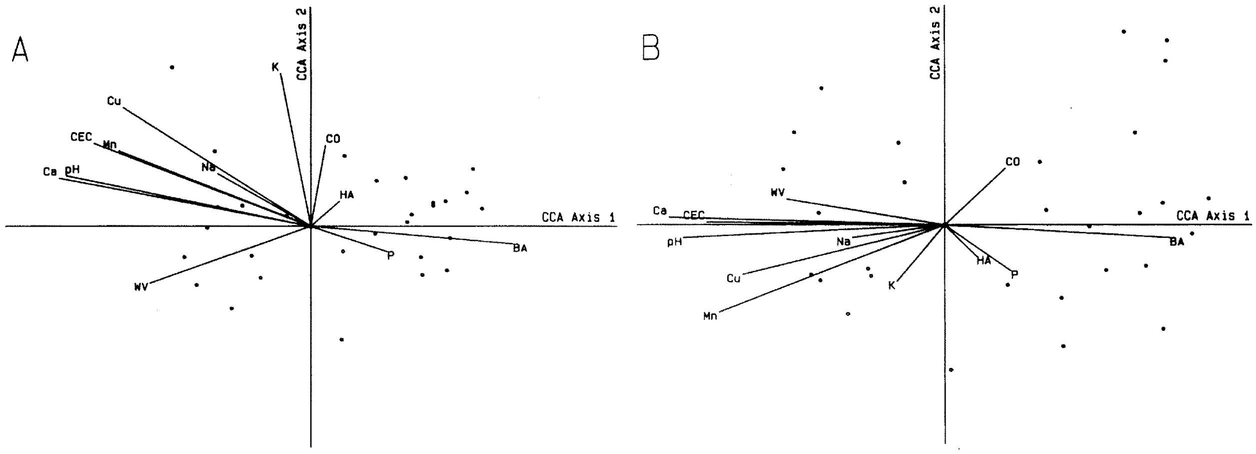

Fig. 3. DCCA biplot based on presence/absence (floristic) data (A) and percentage cover (vegetation) data (B). The arrows projecting from the origin represent environmental variables as described in text. Abbreviations are as in the caption to Fig. 2. Plots are represented by circles.

between species composition and the environmental CCA axes is determined by a Monte Carlo permutation test (ter Braak 1987a, b).

The relative contributions of different environmental variables to the CCA ordination can be displayed with a biplot (ter Braak 1987c). A biplot is a graph of plots (represented by points) and environmental variables (represented as vectors) on the same ordination diagram (see Fig. 3). The coordinates of the environmental vectors are calculated from the CCA eigenvalues and the correlations between the environmental variables and the CCA eigenvalues. The strength of the relationship between an environmental variable and an ordination axis is proportional to length of the perpendicular projection of the environmental vector on the axis. The relative importance of the environmental variables (that is, the degree to which they correspond with species composition) is proportional to the relative lengths of the vectors.

In this study, I employ a modification of CCA known as detrended canonical correspondence analysis (DCCA) which corrects for distortions along ordination axes other than the first. While the utility of detrending has been debated (Wartenberg *et al.* 1987), it has been shown to be necessary and robust by others (Peet *et al.* 1987). However, most of the discussion here will focus on the first ordination axis, so the interpretation of DCCA is practically identical to that of CCA (the first axis of DCCA always has the same eigenvalue as that of CCA).

The eigenvalue is a measure of the strength or importance of a CCA or DCCA axis, and is equivalent to the squared correlation coefficient (r^2) between species scores and plot scores. The eigenvalue must be between zero and one. If the dominant gradient in plant species composition is largely related to the measured environmental variables, then the first eigenvalue will be close to the first eigenvalue of a detrended correspondence analysis (DCA) ordination (Hill and Gauch 1980, see Palmer 1990). In a DCA ordination, plot scores are not constrained to be simple functions of the environmental variables. If the dominant gradient in species composition is not reflected in the measured environmental variables, the DCCA eigenvalues will be much less than the DCA eigenvalues. Thus, comparison of DCA and DCCA eigenvalues can suggest whether important variables are being "left out" of the gradient analysis (ter Braak 1986).

The computer program CANOCO allows the user to "factor out" the effects of particular environmental variables by designating them as co-variables. This capability is used in two ways in this paper. The first use of covariables is to determine whether there are any significant residual relationships between species and the environment after the dominant, most influential environmental variable(s) is (are) accounted for. The second use of covariables is for examining within-plot species-environment relationships. In this case, each covariable corresponds to a particular plot. The covariable is known as a "dummy variable", and has the value 0 when a subplot is not in the corresponding plot, and the value 1 if it is in the plot. The net effect of using dummy variables in such an analysis is to eliminate the influence of among-plot variation when searching for within-plot patterns.

As in Palmer (1990), separate gradient analyses were done for floristic (presence/absence) and vegetation (species percent cover) data. For the large-scale (among plot) gradient analysis, plot means were used as DCCA input for each of the environmental variables. Species data were not transformed. Plot #23 was ordinated passively for the reasons described in Palmer (1990).

For the small-scale (within plot) gradient analysis, the thirty plots were divided into three groups (of ten plots each) based on the rank position of mean calcium concentration (Ca was found to be strongly related to among plot floristic and vegetation patterns, see Results section). The low Ca plots had mean Ca concentrations less than 150 ppm, intermediate Ca plots had mean concentrations between 150 and 380 ppm, and high Ca plots had mean concentrations greater than 380 ppm. Dummy variables were used as covariables to factor out the effects of among-plot variation on the DCCA results. DCCA was performed on the low Ca, intermediate Ca, and high Ca groups of plots in order to determine which of the environmental variables were most strongly related to within-plot species composition in low, intermediate, and high Ca conditions.

Detrended canonical correspondence analysis was also performed on all data sets (flora and vegetation, among-plot and within-plot) using calcium as the single inputted environmental variable. The resultant DCCA first axis species scores estimates the order of modal positions of each species along the calcium gradient (see ter Braak 1987a, c). The species scores for the among-plot comparisons were compared to the within-plot comparisons with linear regression in order to determine whether plant species are distributed along the calcium gradient in the same way within plots as they are among plots.

Results

Semivariograms. The semivariograms of the twelve measured environmental variables are given in Figure 2. The small irregularities in the semivariograms are probably due to statistical uncertainty, and the declines to the right of some semivariograms (*e.g.*, those of P, HA, and CO) are probably due an artifact of spurious periodicity known as the ''hole effect'' (Journel and Huijbregts 1978, Phillips 1985) caused by fewer replicates at greater distances. All of the variables exhibit some degree of spatial dependence (the semivariogram has some region with a positive slope). All of the environmental variables (except Mn and perhaps pH) have semivariograms with sills (horizontal asymptotes). The presence of a sill means that the plots, on average, are homogeneous with respect to the measured environmental variables.

Table 1. Eigenvalues and correlation coefficients between environmental variables and DCCA scores of plots.

	floristic data		vegetation data	
	Axis 1	Axis 2	Axis 1	Axis 2
DCCA Eigenvalue	0.30	0.14	0.43	0.24
DCA Eigenvalue	0.34	0.19	0.50	0.27
Correlation coefficients:				
Phosphorus	0.261	–0.141	0.210	–0.137
Potassium	–0.101	–0.689	–0.147	–0.331
Calcium	–0.861	–0.247	–0.856	–0.042
Magnesium	–0.880	–0.249	–0.873	–0.082
Weight/Volume	–0.550	0.250	–0.493	0.088
pH	–0.837	–0.234	–0.812	–0.163
Buffered acidity	0.684	0.137	0.697	0.062
CEC	–0.745	–0.274	–0.742	–0.087
Manganese	–0.656	–0.457	–0.699	–0.300
Copper	–0.641	–0.478	–0.628	–0.209
Humic acid	0.093	–0.317	0.099	0.072
Canopy openness	0.055	–0.146	0.192	0.302

The small-scale semivariograms (Fig. 2) reveal that for most of the variables, environmental variability on the scale of subplots is substantially smaller than environmental variability on the scale of the entire plot. The

Table 2. Eigenvalues and correlation coefficients between environmental variables and DCCA scores of subplots within plots.

	Low Calcium			
	floristic data		vegetation data	
	Axis 1	Axis 2	Axis 1	Axis 2
DCCA Eigenvalue	0.12	0.09	0.13	0.09
DCA Eigenvalue	0.36	0.33	0.68	0.57
Correlation coefficients:				
Phosphorus	0.202	0.165	0.255	0.133
Potassium	0.267	0.122	0.264	0.089
Calcium	0.417	–0.018	0.387	–0.096
Magnesium	0.037	–0.048	0.080	0.025
Weight/Volume	–0.248	–0.297	–0.205	–0.166
pH	0.422	–0.100	0.331	–0.011
Buffered acidity	–0.066	0.087	0.036	0.097
CEC	0.354	0.051	0.342	0.012
Manganese	0.293	–0.120	0.248	–0.175
Copper	0.244	–0.086	0.161	–0.093
Humic acid	0.210	–0.146	0.277	–0.009
Canopy openness	–0.042	0.104	–0.008	–0.041
	Intermediate Calcium			
	floristic data		**vegetation data**	
	Axis 1	**Axis 2**	**Axis 1**	**Axis 2**
DCCA Eigenvalue	0.14	0.09	0.24	0.15
DCA Eigenvalue	0.38	0.31	0.63	0.58
Correlation coefficients:				
Phosphorus	–0.056	–0.004	0.031	–0.084
Potassium	0.030	–0.005	–0.028	–0.233
Calcium	0.265	0.353	–0.201	0.054
Magnesium	0.442	0.159	–0.317	–0.077
Weight/Volume	–0.037	0.075	0.177	0.226
pH	0.250	0.556	–0.208	0.435
Buffered acidity	–0.164	–0.190	0.363	–0.431
CEC	0.255	0.241	0.014	–0.158
Manganese	–0.022	0.322	–0.145	0.088
Copper	0.645	0.073	–0.316	–0.211
Humic acid	0.281	–0.004	–0.224	–0.342
Canopy openness	–0.139	0.086	0.068	–0.216

Table 2 (continued)

	High Calcium			
	floristic data		vegetation data	
	Axis 1	Axis 2	Axis 1	Axis 2
DCCA Eigenvalue	0.07	0.05	0.11	0.10
DCA Eigenvalue	0.21	0.19	0.50	0.27
Correlation coefficients:				
Phosphorus	0.132	0.115	0.250	−0.112
Potassium	−0.080	−0.234	−0.057	−0.056
Calcium	0.100	−0.544	−0.146	−0.116
Magnesium	0.030	0.006	0.153	0.012
Weight/Volume	−0.044	−0.072	−0.205	0.124
pH	−0.005	−0.644	−0.193	−0.033
Buffered acidity	−0.144	0.506	0.261	0.076
CEC	0.050	−0.524	−0.117	−0.100
Manganese	−0.184	−0.438	−0.340	0.159
Copper	0.150	−0.203	−0.099	−0.142
Humic acid	0.021	0.145	−0.025	−0.054
Canopy openness	0.512	−0.062	−0.195	−0.434

number of soil samples used for determining these semivariograms is low, so the patterns within the semivariograms are irregular and not easily interpreted. However, it is clear that with the exception of phosphorus, the overall heights of the small-scale semivariograms are much lower than the overall heights of the larger-scale semivariograms.

Among-plot detrended canonical correspondence analysis. The first DCCA axis is strongly related to Ca and its correlates (Mg, pH, BA, CEC) for both the floristic and the vegetation comparisons (Fig. 3, Table 1). Monte Carlo permutation tests reveal that the eigenvalue of the first axis is significantly larger than the null expectation (that which we expect if the relationship between plant and environment is random) for both the floristic ($p<0.01$) and the vegetation ($p<0.01$) comparisons. Once Ca is "factored out" as a covariable, however, the first axis eigenvalue is neither significant for flora ($p=0.25$) nor vegetation ($p=0.17$). The first two DCCA eigenvalues are very close to the first two DCA eigenvalues for both vegetation and flora (Table 1), indicating that the dominant gradient in species composition is indeed related to the measured environmental variables (calcium and its correlates in particular), and not to some important but unmeasured factor.

Within-plot detrended canonical correspondence analysis. Table 2 displays the correlations between the environmental variables and the first two DCCA axes for the within-plot floristic and vegetation data sets. In all cases, the plant-environment relationship along the first axis is "significant" at $p<0.05$. "Significant" is in quotes because of the problem of spatial dependence.

Table 3. Eigenvalues of DCCA using calcium as only inputted environmental variables. All of the eigenvalues are significantly greater than the null expectation at $p<0.01$.

	floristic data	vegetation data
Among-plot ordination	0.26	0.37
Within-plot ordination		
low calcium	0.08	0.09
intermediate calcium	0.07	0.07
high calcium	0.04	0.06

Ca and some of its correlates (pH and CEC but not Mg and BA) are the environmental variables with the highest correlations with the first DCCA axis for the low calcium group. This holds for both the floristic and the vegetation comparison. Weight/Volume is most strongly related to the second floristic DCCA axis, and Manganese is most strongly related to the second vegetation DCCA axis. For the floristic comparison of the intermediate Ca group, copper is most strongly correlated with the first axis, and pH is most strongly related to the second axis. In the vegetation comparison, BA and Mg (but not Ca, pH, or CEC) is highly related to the first axis, and BA and pH are related to the second axis. Canopy openness is strongly related to the first axis and calcium (and its correlates except Mg) are most strongly related to the second axis for the floristic comparison of the high calcium group. For the vegetation comparison, manganese is most strongly related to the first axis, and canopy openness is most strongly related to the second axis.

Species distributions along the calcium gradient.. The eigenvalues for the DCCA ordinations with calcium as the only environmental variable are given in Table 3. All of the ordinations were "significant" at $p<0.01$, suggesting that calcium or a related factor determines species composition to some degree. For the among-plot ordinations, the DCCA eigenvalues are reasonably clo-

Table 4. The correlation coefficient (r) between species positions along the among-plot calcium gradient and species positions along the within-plot calcium gradient. Abundant species include all species present in at least 10 subplots for the floristic data, and have a cover of at least 0.1 m^2 for the abundance data. The number of species used in each comparison is in parentheses.

	floristic data	vegetation data
All species		
low calcium	0.037 (104)	0.000 (104)
intermediate calcium	0.169 (183)*	0.148 (183)*
high calcium	0.124 (223)	0.100 (223)
Abundant species		
low calcium	0.278 (38)	0.120 (30)
intermediate calcium	0.397 (63)**	0.248 (62)*
high calcium	0.263 (106)**	0.240 (94)*

* $p<0.05$
** $p<0.01$

se to the DCA first axis eigenvalues (Table 1), implying that calcium is related to a dominant gradient in species composition. For the within-plot ordinations, however, the DCCA eigenvalues are substantially less than the DCA first eigenvalues (Table 2), implying that calcium is not strongly related to the dominant gradient(s) determining species composition.

The correlations between the species positions on the calcium gradient within and among plots are shown in Table 4. Some of these correlations are statistically significant. The significance and the strength of the correlations increase somewhat when one only considers the most abundant species.

Note, however, that all of the correlations in Table 4 are quite low ($r < 0.4$). The percentage variance in the within-plot data explained by the among-plot data ($100\ r^2$) is always less than 16%. These low correlations do not occur as a result of lack of variation in calcium concentration within a plot: calcium concentrations typically vary three-fold to five-fold within plots. The low calcium group is the only group which does not exhibit any significant relationships between the among-plot and within-plot distributions of species. However, it must be noted that this may be due in part to the lower sample size (number of species) in the group. The low calcium group was also the only group which had calcium as the environmental variable most strongly related to the first DCCA axis.

Discussion

Spatial dependence is a property of most of the environmental variables studied, as indicated by their semivariograms (Fig. 2). This implies that statistical analyses involving these variables are not entirely valid, because the assumption of independent replication is violated. Therefore, statistical inference regarding the within-plot gradient analyses must be considered inexact. It is likely that the relationships with manganese and copper (which are seldom implicated as causes of species distributions, but see Lieth 1980 for Mn) result from their strong patterns of spatial dependence (Fig. 2).

Spatial dependence is not likely to be a major problem for most variables because the slopes are generally very shallow, indicating that the spatial dependence which exists is weak. This is consistent with the results of Burrough (1981), who found that many kinds of environmental data have weak spatial dependence. The problem of spatial dependence is also ameliorated by the grouping of plots. This is because the species-environment relationship in a given plot is almost undoubtedly independent of the species-environment relationships in all other plots.

The horizontal sill on most of the semivariograms indicates that the plots, on the average, are homogeneous with respect to the environment. This is reassuring because plots were selected, in part, on the basis of an appearance of topographic and vegetational homogeneity.

The smaller-scale semivariograms in Fig. 2, while very noisy, are clearly much lower than the larger-scale semivariograms (with the exception of phosphorus). This lower variability on the scale of subplots greatly facilitates gradient analysis because it means that subplots are distinct from each other with respect to the environment. If subplots were not distinct, it would be impossible to examine plant-environment relationships within a plot.

Subplots are not distinct from each other with respect to Phosphorus. The small-scale (< 1.6 m) semivariance of this variable is close to the large scale (> 10 m) semivariance. Within-plot gradient analyses with phosphorus may therefore be uninformative and any "significant" results may be spurious.

It appears that Ca or something related to Ca (perhaps unmeasured) is responsible for all of the detectable plant-environment relationships among plots (Table 1). The fact that the DCCA eigenvalue is very close to the DCA eigenvalue implies that the major gradient in species gradient is related to the measured environmental variables. In addition, there is no significant variation in species composition once calcium is factored out. Although a number of factors highly correlated with Ca (*i.e.*, Mg, pH, BA, CEC) appear to relate to species composition of 0.1 ha plots equally well, I concentrate on Ca here. This is because calcium has been widely used in the literature as an important environmental gradient (*e.g.*, Beals and Cope 1964, Monk 1967, Christensen and Peet 1984). However, patterns in plant communities correlated with calcium described here (and elsewhere) are not necessarily directly or indirectly caused by calcium. Another correlated parameter, measured or not, may be the determinant of the observed patterns.

For the small scale comparisons (Table 2), it is interesting that calcium has a strong relationship to the first DCCA axis only in the low calcium group. This implies that calcium (or a correlated resource) most strongly determines species composition where it is most limiting. In the intermediate and high calcium groups, this resource is less limiting, and other variables may become more important in determining species distributions. Alternatively, the scarcity of calcium or a correlate may limit the availability of another nutrient (*e.g.*, nitrogen) which influences species composition. For example, soil pH can control the availability (as well as the toxicity) of a number of elements (Etherington 1982).

In all cases, the within-plot DCCA eigenvalues were substantially lower than the corresponding DCA eigenvalues (Table 2, 3). Recall that in the among-plot comparisons, DCCA eigenvalues were close to DCA eigenvalues (Table 1). This indicates that the dominant gradients in species composition within plots are not strongly related to the inputted environmental variables. This agrees with the results of Palmer (1990), in which within-plot species compositional patterns were shown to be only weakly related to among-plot species compositional patterns. Even though the within-plot relationships between plants and the environment described here are "significant", they are not necessarily dominant or important. Thus, species-environment relationships are not strictly self-similar.

The high "nugget" variance (or y-intercept) of the semivariograms of the environmental variables (Fig. 2) may partially explain the low DCCA eigenvalues. Much of the variance in the environment is therefore *within* subplots. The extremely fine-scaled (within-subplot) plant-environment relationships will be hidden in the comparison among subplots. Reducing the size of the subplot will not substantially reduce this problem; the variance in the environment is not near zero when plot size is extremely small.

The general results for the vegetation DCCA ordinations are similar to the floristic DCCA ordinations. This contrasts with the indirect gradient analysis in Palmer (1990), in which the floristic data exhibit more prominent patterns than the vegetation data. The similarity of the results in this paper implies that direct gradient analysis is more robust to variations in the form of the data (*i.e.*, presence/absence vs. untransformed species abundances) than is indirect gradient analysis.

Species composition is related to calcium both among (Table 1, Fig. 3) and within (Table 3) plots, but this does not necessarily mean that the same species are behaving in similar manners at the two drastically different scales. The significant correlations in Table 4 indicate that there is indeed a barely detectable similarity of the small-scale distribution of species along the calcium gradient to the large-scale distribution of species along the calcium gradient. However, the weakness of the correlations suggests that small-scale species distributions along the calcium gradient differ to a great degree from large-scale distributions. For example, a species which is restricted to low calcium areas may be specialized to high calcium microsites. In other words, species distributions along gradients can potentially be scale-dependent. The calcium gradient cannot be considered self-similar.

Summary

Soil chemical characteristics and canopy openness demonstrate weak spatial dependence at scales smaller than one tenth hectare. This implies that statistical inference concerning gradient analysis on a small scale may not be entirely valid. Most environmental variables are homogeneous at the scale of one tenth hectare. Most variables also have considerably lower variance at the scale of subplots (1.6 m) than they do at the scale of plots (50 m). This means that subplots are distinct from each other with respect to the environment; gradient analysis is therefore possible (even if statistical inference is not).

Both the vegetation and floristic ordinations of plots display a significant relationship between soil calcium concentration and species composition. However, calcium was not the variable with the highest correlation with species composition among subplots, except within plots with low calcium. The low DCCA eigenvalues (compared to DCA eigenvalues) imply that some unmeasured gradient(s) may be important in determining within-plot compositional patterns. When only calcium was used in the DCCA ordination, calcium was "significantly" (but not strongly) related to species composition for both the vegetation and the floristic ordinations for the low, intermediate, and high calcium groups.

The order of plant species along the calcium gradient is somewhat different among plots and within plots. Although there is a significant correlation between species order within plots and species order among plots (for plots with intermediate and high calcium), the correlations are weak. This implies that the species are situated along environmental gradients in different ways at different spatial scales.

Acknowledgements. I wish to thank Norm Christensen, Robert Peet, Boyd Strain, Brent Mishler, Donald Burdick, and Suzanne McAlister for advice and criticism related to this research. Norm Kenkel and an anonymous reviewer offered valuable comments on the manuscript. I am indebted to the North Carolina Department of Agriculture's Soil Testing Service for evaluating the soil samples. Financial support was provided by a National Science Foundation Graduate Fellowship, a James B. Duke Fellowship from Duke University, and a teaching assistantship with the Department of Botany, Duke University.

REFERENCES

BEALS, E.W. and J.B. COPE. 1964. Vegetation and soils in an eastern Indiana woods. Ecology 45: 777-792.

BURROUGH, P.A. 1981. Fractal dimensions of landscapes and other environmental data. Nature 294: 240-242.

CHRISTENSEN, N.L. and R.K. PEET. 1984. Convergence during secondary forest succession. J. Ecol. 72: 25-36.

CLARK, I. 1979. Practical Geostatistics. Applied Science Publishers, London.

COOK, D.G. and S.J. POCOCK. 1983. Multiple regression in geographical mortality studies, with allowance for spatially correlated errors. Biometrics 39: 361-371.

ETHERINGTON, J.R. 1982. Environment and Plant Ecology. Second ed. Wiley, New York.

FINGLETON, B. 1986. Analyzing cross-classified data with inherent spatial dependence. Geogr. Anal. 18: 48-61.

HILL, M.O. 1974. Correspondence analysis: a neglected multivariate method. Appl. Stat. 23: 340-354.

HILL, M.O. and H.G. GAUCH, Jr. 1980. Detrended correspondence analysis: an improved ordination technique. Vegetatio 42: 47-58.

HURLBERT, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Mon. 54: 187-211.

JOURNEL, A.G. and C. HUIJBREGTS. 1978. Mining Geostatistics. Academic Press, London. 600 pp.

LEGENDRE, P., R.R. SOKAL, N.L. ODEN and A. VAUDOR. 1987. Analysis of variance with autocorrelation in both the variable and the classification criterion. Proceedings, First Conference of the International Federation of Classification Societies, Aachen. p. 110.

LIETH, H. 1980. Untersuchungen über den Mineralstoffspiegel im De la Howe Forest, Mc Cormick County, South Carolina. Veröff. Geobot. Inst. ETH, Stiftung Rübel, Zürich. 69: 162-170.

LEMMON, P.E. 1956. A spherical densiometer for estimating forest overstory density. For. Sci. 2: 314-320.

MANDELBROT, B.B. 1983. The fractal geometry of nature. Freeman, San Francisco. 468 p.

MEHLICH, A. 1984a. Mehlich 3 soil test extract: a modification of Mehlich 2 extractant. Comm. Soil Sci. Plant Anal. 15: 1409-1416.

MEHLICH, A. 1984b. Photometric determination of humic matter in soils, a proposed method. Comm. Soil Sci. Plant Anal. 15: 1417-1422.

MONK, C.D. 1967. Tree species diversity in the eastern deciduous forest with particular reference to north central Florida. Am. Nat. 101: 173-187.

PALMER, M.W. 1988. Fractal Geometry: a tool for describing spatial patterns of plant communities. Vegetatio 75: 91-102.

PALMER, M.W. 1990. Spatial scale and patterns of vegetation, flora and species richness in hardwood forests of the North Carolina piedmont. Coenoses 5: 79-86.

PEET, R.K. and N.L. CHRISTENSEN. 1980. Hardwood forest vegetation of the North Carolina piedmont. Veröff. Geobot. Inst. ETH, Stiftung Rübel, Zürich. 69: 14-39.

PEET, R.K., R.G. KNOX, J.S. CASE and R.B. ALLEN. 1988. Putting things in order: the advantages of detrended correspondence analysis. Am. Nat. 131: 924-934.

PHILLIPS, J.D. 1985. Measuring complexity of environmental gradients. Vegetatio 64: 95-102.

ROBERTSON, G.P. 1987. Geostatistics in ecology: interpolating with known variance. Ecology 68: 744-748.

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. 1987a. The analysis of vegetation environment relationships by canonical correspondence analysis. Vegetatio 69: 69-77.

TER BRAAK, C.J.F. 1987b. CANOCO - a FORTRAN program for community ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis (version 2.1) ITI-TNO, Wageningen, The Netherlands, 95 pp.

TER BRAAK, C.J.F. 1987c. Ordination, p. 91-173 in Jongman, R.H., C.J.F. ter Braak, and O.F.R. van Tongeren, eds. Data analysis in community ecology. Pudoc, Wageningen, The Netherlands.

TER BRAAK, C.J.F. and I.C. PRENTICE. 1988. A theory of gradient analysis. Adv. Ecol. Res. 18: 271-313.

WARTENBERG, D., S. FERSON and F.J. ROHLF. 1987. Putting things in order: a critique of detrended correspondence analysis. Am. Nat. 129: 434-448.

WEBSTER, R. 1985. Quantitative spatial analysis of soil in the field. Adv. Soil Sci. 3: 1-70.

WHITTAKER, R.H. 1956. Vegetation of the Great Smoky Mountains. Ecol. Mon. 26: 1-80.

WILKINSON, G.N., S.R. ECKERT, T.W. HANCOCK and O. MAYO. 1983. Nearest Neighbor (NN) Analysis of Field experiments. J. Royal Stat. Soc. Ser. B: 151-211.

Manuscript received: May 1989