

DIRECT MULTI-SCALE ORDINATION WITH CANONICAL CORRESPONDENCE ANALYSIS

HELENE H. WAGNER¹

*Department of Biology, Colorado State University, Fort Collins, Colorado 80523 USA, and
WSL, Swiss Federal Research Institute for Forest, Snow and Landscape, 8903 Birmensdorf, Switzerland*

Abstract. Gradient analysis uses ordination methods to study the structure of biotic communities caused by biotic processes operating in a heterogeneous environment. This structure has two spatial components: spatial processes within the community create autocorrelation, and the spatial structure of environmental factors creates spatial dependence. Ordination methods, however, do not make use of spatial information. Spatial alternatives are available in multivariate geostatistics, but are not compatible with important ordination methods used in gradient analysis, correspondence analysis and canonical correspondence analysis (CA, CCA). This paper shows how CA and CCA can be partitioned by distance (indirect and direct multi-scale ordination) and integrated with geostatistics. A diagnostic tool enables ecologists to partition ordination results by distance, to distinguish between components of spatial dependence and of spatial autocorrelation, and to check assumptions of independent residuals, stationarity, and scale-invariant correlation. The application is illustrated with a well-known data set of oribatid mites. Empirical chi-square variograms of individual species, their pair-wise cross variograms, and the variogram of the total inertia are defined and summarized in a variogram matrix, which leads to a spatial partitioning of the eigenvalues. The empirical variogram matrix provides a link to coregionalization analysis that may be used to simultaneously model spatial dependence and spatial autocorrelation. This will be useful for answering questions about the organism-specific scale of response to the environment, the optimal spacing of sampling units, or the scale-dependent effect of environmental factors.

Key words: *canonical correspondence analysis; chi-square distance; gradient analysis; multi-scale ordination; multivariate geostatistics; residual analysis; scale; stationarity; variogram matrix.*

INTRODUCTION

Gradient analysis (Whittaker 1967, ter Braak and Prentice 1988) aims to explain the differences in species composition in a biotic community observed at different sampling locations. These differences are caused either by internal processes within the community, such as dispersal and competition (false gradients), or as a response to external factors, such as environmental variation (true gradients; Legendre 1993). Ordination is the primary quantitative method for studying both types of structure. Indirect ordination describes intrinsic gradients in species composition, while direct ordination identifies compositional gradients in a community as a response to measured environmental factors (De'ath 1999). The basic methods of indirect and direct ordination, principal component analysis (PCA) and redundancy analysis (RDA; see Legendre and Legendre 1998 for reviews of PCA and RDA), are used in other scientific areas as well, such as in soil science or remote sensing. Both PCA and RDA assume a linear relationship among variables. This assumption limits their ap-

plicability in gradient analysis to short gradients (i.e., to data sets with little species turnover between sampling units). Alternative methods based on correspondence analysis (CA; see Legendre and Legendre 1998 for a review) can accommodate a unimodal species response, and therefore are better suited for the analysis of longer gradients (e.g., Gauch 1982, Jongman et al. 1995, Legendre and Legendre 1998). As a consequence, community ecologists often perform indirect ordinations with CA or direct ordinations with canonical correspondence analysis (CCA; Table 1).

Legendre 1993 argued that the structure of biotic communities is inherently spatial for two reasons. First, population dynamics and interspecific interactions operate through individual organisms that exist and interact only within their immediate neighborhood (Tilman and Kareiva 1997). Such contagious biotic processes create "spatial autocorrelation" within the community (Legendre 1993). Second, physical processes create spatial structure in environmental factors, which in turn causes "spatial dependence" in biotic communities (Legendre 1993, Legendre and Legendre 1998). The size of the ecological neighborhood, wherein organisms interact with other organisms and with their physical environment, depends on the particular ecological process, the time period, and the organism's mobility or activity (Addicott et al. 1987, Wiens 1989).

Manuscript received 22 November 2002; revised 16 April 2003; accepted 7 May 2003; final version received 23 June 2003.
Corresponding Editor: G. M. Henebry.

¹ Direct correspondence to address in Switzerland. E-mail: helene.wagner@wsl.ch

TABLE 1. Overview of methods of indirect and direct ordination, depending on the type of species response to gradients, and their corresponding method of spatial partitioning (multi-scale ordination).

| Method | Short gradient (linear response model) | Long gradient (unimodal response model) |
|-------------------------------------|---|---|
| Indirect ordination | PCA | CA |
| Direct ordination | RDA | CCA |
| Preserved distances | Euclidean distance | chi-square distance |
| Multi-scale ordination using | $C(h)$ | $Q^2(h)$ |
| Variogram type | ordinary (Euclidean) variogram | chi-square variogram |
| Basic element of spatial covariance | $\gamma_{ij}(a, b) = (1/2)(y_{ia} - y_{ib})(y_{ja} - y_{jb})$ | $\gamma_{ij}(a, b) = (N/2)(q_{ia} - q_{ib})(q_{ja} - q_{jb})$ |

Therefore, different processes may create autocorrelation and spatial dependence at different scales, depending on the scale of an organism's response. The recognition of the spatial nature of biotic communities requires an extension of the research paradigm of gradient analysis (Fig. 1). The internal structure of the biotic community (B) and its response to the environment (E) cannot be fully understood without considering geographic space (S). Spatial effects in community structure may occur at more than one scale and can be direct (i.e., due to the spatial nature of biotic interactions [autocorrelation]), or indirect (i.e., due to the spatial structure of the environment [spatial dependence]). Ordination methods, however, are ill adapted to the spatial paradigm, as they do not make use of spatial information.

Geostatistics are the standard method to deal with spatial autocorrelation and spatial dependence. Legendre and Fortin (1989) and Fortin (1999) grouped spatial statistics for ecological analysis by four objectives. Of these, three are commonly addressed using geostatistical methods: testing for the presence of spatial autocorrelation (e.g., Moran's I and Geary's c); describing spatial structure (e.g., variograms or correlograms); and estimation and mapping (e.g., interpolation by kriging). For the fourth objective of modeling and testing the correlation structure between autocorrelated data, ecologists are referred to methods developed outside of the geostatistical framework, such as the partial Mantel test or partial CCA (Legendre and Fortin 1989, Legendre and Legendre 1998, Fortin 1999). However, geostatistics is more than simple variography and or-

dinary kriging (Rossi et al. 1992), and offers methods for the modeling of the multivariate correlation of spatially structured variables (e.g., Goovaerts 1999). Particularly, I want to point out some insights and developments in multivariate geostatistics that are of high relevance for community and landscape ecologists but so far have gone largely unnoticed by the ecological literature. The first concerns the scale dependence of the species–environment correlation. In geostatistical terms, the multivariate correlation structure of a set of variables is said to be “intrinsic” if it is independent of the spatial correlation. In order to prevent confusion, I will use the term “scale invariant” for an intrinsic correlation structure (and “scale-dependent” correlation for its opposite), and reserve the term “intrinsic” to the context of the stationarity assumption, with which it is not related (Wackernagel 1998). Hence, if the species–environment correlation changes with scale (i.e., with geographic distance), the correlation is scale dependent, and a global parameter estimate of a correlation or regression coefficient is meaningless. Second, regionalized analysis regards an observed phenomenon as the sum of several independent subphenomena acting at different characteristic scales. In a regionalized PCA, for instance, a set of eigenvalues is extracted independently for each characteristic scale. Regionalized versions of PCA and RDA exist under the names of factorial kriging analysis or coregionalization analysis (Matheron 1982, Goovaerts 1992, 1994, Wackernagel 1998; for an application in genetic ecology see Monestiez and Goulard 1997). Regionalized multivariate methods can deal with scale-depen-

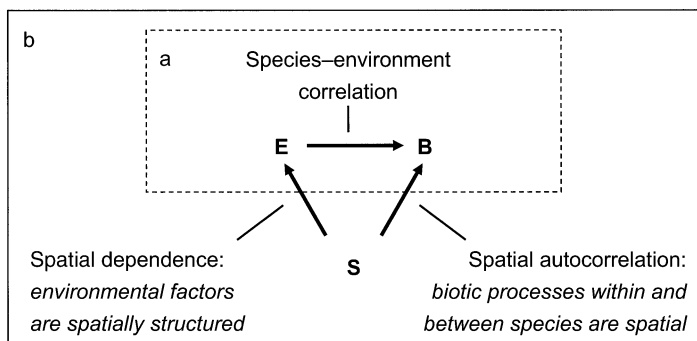


FIG. 1. (a) Nonspatial and (b) extended research paradigm of gradient analysis.

dent correlations, but this comes at a cost of increased model complexity (Wackernagel 1998). A matrix of coregionalization, containing fitted variogram and cross-variogram functions for all variables, needs to be derived separately for each characteristic scale. The problems of identifying characteristic scales and fitting appropriate models are beyond the scope of this paper and will be dealt with separately. Before the question of “how” should be addressed, however, one should answer the question “when.” There is a need for a diagnostic tool for assessing whether a regionalized analysis is necessary.

In a recent paper (Wagner 2003), I presented a geostatistical version of multi-scale ordination (MSO; originally defined for blocked-variance techniques by Noy-Meir and Anderson [1971], and Ver Hoef and Glenn-Lewin [1989]) and showed how the results of PCA can be partitioned by distance and summarized in an empirical variogram matrix. This facilitates the application of multivariate geostatistical methods in gradient analysis within a linear framework. An extension to the CA family of methods is needed to make geostatistics compatible with a unimodal response model.

This paper has two goals. The first is to present a spatial version of CA and CCA that allows their integration with multivariate geostatistics. This is achieved by extending indirect multiscale ordination with PCA to CA and to direct ordination with CCA. The second aim is to provide a diagnostic tool that enables ecologists to partition ordination results by distance, to distinguish between components of spatial dependence and of spatial autocorrelation, and to check assumptions of independent residuals, stationarity, and scale-invariant correlation. The application of the method is illustrated with a well-known data set of oribatid mites (Borcard et al. 1992, Borcard and Legendre 1994). A worked example in the Appendix demonstrates the calculations. See the Supplement for an R library with functions for the spatial partitioning of ordination results from PCA, RDA, CA, and CCA, including a diagnostic plot.

METHODS

Multi-scale ordination with PCA

Let \mathbf{Y} be a matrix with elements y_{ia} that describe the observed abundance of each species $i \in \{1, \dots, s\}$ in each sampling unit $a \in \{1, \dots, N\}$ of a spatially referenced sample of size N , and \mathbf{C} the empirical variance-covariance matrix of \mathbf{Y} . Then \mathbf{C} can be partitioned into an empirical variogram matrix (i.e., a series of distance-dependent matrices $\mathbf{C}(h)$, one for each distance class h [Wagner 2003]). Matrix \mathbf{C} is equal to the weighted sum of the component matrices $\mathbf{C}(h)$ of the variogram matrix, with weights proportional to n_h , the number of pairs of sampling units in each distance class h , where $n = \sum n_h = N^2$:

$$\mathbf{C} = \sum_h \frac{n_h}{n} \mathbf{C}(h). \quad (1)$$

The sum of the diagonal of \mathbf{C} is the total variance of \mathbf{Y} . PCA summarizes the variation in \mathbf{Y} by generating a set of s uncorrelated new variables that are linear combinations of the old variables. These new variables (the principal components) form orthogonal axes in the data space. The eigenvalue λ_f quantifies how much of the total variance is represented by PCA axis f . The eigenvector \mathbf{u}_f defines the translation between the old variables and the new variable f .

The geostatistical form of multi-scale ordination uses the variogram matrix to partition PCA results by distance (Wagner 2003). A plot of the distance-dependent variance along PCA axis f reveals the spatial structure of the multi-species pattern represented by axis f . In the case of PCA, the same result is obtained by plotting an empirical variogram of the scores of the sampling units on axis f .

In geostatistical terms, an empirical variogram should only be interpreted and a variogram model fitted to it if the assumption of intrinsic stationarity is met (see Henley 2001 for a nonmathematical review of stationarity assumptions). This assumption requires that the expected value of $[y_{ia} - y_{i(a+h)}]$ be zero for all vectors h separating any two points in the region of interest. Unfortunately, no tests of stationarity assumptions are available because the assumption concerns the underlying process and not its realized outcome, the observed data, and we only have one realization of the data (Myers 1989). However, Henley (2001) suggests that partial tests for specific components of stationarity, such as the absence of trend, might be constructed successfully.

Spatial definition of correspondence analysis

Correspondence analysis (CA) may be applied to any set of variables that are measured in the same units and contain only non-negative values (Legendre and Legendre 1998). Analysis is performed on the residuals of a model of the independence of species and observations, such as in a chi-square test of independence in a two-way contingency table. This relationship with contingency analysis leads to the fact that an ordination with CA rotates around a hybrid column-row centroid, which has sometimes been criticized for lack of ecological meaning. The chi-square distance, which is used to quantify the relationships among rows and columns, excludes double zeros so that the method is suitable for analyzing long gradients with a high degree of species turnover. The chi-square distance is appropriate for modeling symmetric, unimodal species distributions along an environmental gradient (ter Braak 1985). However, CA is not necessarily the best method for analyzing nonlinear relationships, as it has been shown that species response curves may often be asymmetric (Oksanen and Minchin 2002). Initially, some algo-

rithms for CA and related methods provided instable solutions (Tausch et al. 1995), but this problem has been solved for all major software packages used in community analysis (cf., Oksanen and Minchin 1997). The present description of CA adopts the (stable) matrix definition of Legendre and Legendre (1998).

CA is simply a PCA of transformed abundance values (see worked example in the Appendix). Instead of the matrix \mathbf{Y} , CA evaluates the matrix \mathbf{Q} that describes the deviance of the observed values from those expected under the assumption that the species and the sampling units are independent. The matrix \mathbf{Q} contains, for each species i and sampling unit a , the rescaled contribution to chi-square, q_{ia} (Legendre and Legendre 1998, where \mathbf{Q} is referred to as $\tilde{\mathbf{Q}}$):

$$q_{ia} = \frac{\chi_{ia}}{\sqrt{f_{++}}} = \frac{\text{observed} - \text{expected}}{\sqrt{f_{++}} \sqrt{\text{expected}}} \\ = \frac{f_{ia}f_{++} - f_{i+}f_{+a}}{f_{++}\sqrt{f_{i+}f_{+a}}} = \frac{p_{ia} - p_{i+}p_{+a}}{\sqrt{p_{i+}p_{+a}}} \quad (2)$$

where f_{i+} is the total frequency of species i , f_{+a} is the total frequency of all species in sample a , and f_{++} is the grand total. The character p denotes a relative frequency, which is the corresponding frequency f divided by the grand total f_{++} .

CA can be decomposed by distance in a similar way as PCA. Where PCA performs eigen analysis of the variance-covariance matrix \mathbf{C} , CA performs eigen analysis of the matrix $\mathbf{Q}^2 = \mathbf{Q}^T\mathbf{Q}$, the product of the transposed matrix \mathbf{Q}^T and \mathbf{Q} (Table 1; Legendre and Legendre 1998).

The sum of the diagonal of \mathbf{Q}^2 is equal to the sum of the squared elements of \mathbf{Q} , $\sum q_{ia}^2$, and is called the total inertia. The total inertia plays the same role in CA as the total variance in PCA. Technically, the inertia is the sum of squares of the matrix \mathbf{Q} , so that it differs from the total empirical variance of \mathbf{Q} by a factor N . CA returns $(s - 1)$ axes, as the last eigenvalue, λ_s , is always zero.

The matrix \mathbf{Q}^2 can be re-expressed in terms of the rescaled chi-square distance between observations. In Wagner (2003), I defined the basic unit of spatial covariance, $\gamma_{ij}(a, b)$, for the PCA case as half the product of the difference between two samples a and b , measured for two species i and j (Table 1). For the chi-square case, the basic unit is

$$\gamma_{ij}(a, b) = \frac{N}{2}(q_{ia} - q_{ib})(q_{ja} - q_{jb}). \quad (3)$$

The following equation defines an empirical chi-square variogram $\gamma_i(h)$ for a species i ($i = j$) or a chi-square cross variogram $\gamma_{ij}(h)$ for two species i and j ($i \neq j$) and distance class h :

$$\gamma_{ij}(h) = \frac{N}{2n_h} \sum_{a,b|h_{ab}=h} (q_{ia} - q_{ib})(q_{ja} - q_{jb}) \quad (4)$$

where n_h is the number of pairs of observations a and

b in distance class h . Eq. 4 can be used to construct the distance-dependent matrices $\mathbf{Q}^2(h)$ of a chi-square variogram matrix that contains a matrix of type \mathbf{Q}^2 for each distance class h .

The spatial definition of the total inertia is

$$\sum_i \sum_a q_{ia}^2 = \frac{1}{2N} \sum_i \sum_{a,b} (q_{ia} - q_{ib})^2 \\ = \frac{1}{N^2} \sum_i \sum_{a,b} \gamma_i(a, b) = \sum_i \sum_h w_h \gamma_i(h) \quad (5)$$

where the weight $w_h = n_h/n$ is the proportion of pairs of observations a and b that falls into distance class h . The variogram of the total inertia, $\gamma_Q(h)$, is

$$\gamma_Q(h) = \frac{N}{2n_h} \sum_i \sum_{a,b|h_{ab}=h} (q_{ia} - q_{ib})(q_{ja} - q_{jb}) \\ = \sum_i \gamma_i(h). \quad (6)$$

The eigenvalue λ_f of CA axis f is partitioned by distance h by multiplying $\mathbf{Q}^2(h)$ with the eigenvector \mathbf{u}_f of axis f :

$$\lambda_f = \mathbf{u}_f^T \mathbf{Q}^2 \mathbf{u}_f = \sum_h \frac{n_h}{n} \mathbf{u}_f^T \mathbf{Q}^2(h) \mathbf{u}_f = \sum_h w_h \lambda_f(h). \quad (7)$$

Contrary to the PCA case, this variance profile of eigenvalue λ_f or chi-square variogram of ordination axis f , is not equal to a variogram of the site scores on axis f . This is because the site scores contain a weighting by row totals. As in the PCA case, the geostatistical interpretation of the empirical variogram matrix $\mathbf{Q}(h)$ depends on the assumption of intrinsic stationarity.

Direct multi-scale ordination with CCA

Problems with nonstationarity are often related to the influence of some explanatory variable(s) exhibiting spatial structure (i.e., they are due to spatial dependence sensu Legendre [1993]). If this trend is removed (e.g., by fitting a regression model), many problems with spatial autocorrelation can be attenuated.

The heart of direct ordination by CCA is a weighted multivariate linear regression of the matrix \mathbf{Q} on the centered matrix of explanatory (environmental) variables, \mathbf{X} , with weights proportional to the row totals of \mathbf{Y} (Legendre and Legendre 1998; see worked example in the Appendix). The regression decomposes \mathbf{Q} into the matrix of fitted values, \mathbf{Q}_{fit} , and the matrix of residuals, \mathbf{Q}_{res} . Two separate ordinations by CA are performed on \mathbf{Q}_{fit} and \mathbf{Q}_{res} . The eigenvalues λ_f of the two ordinations sum up to the total inertia of \mathbf{Q} . The maximum number of nonzero eigenvalues, that is, the number of axes, is determined by the number of response variables, s , the number of explanatory variables, m , and the number of observations, N . The number of canonical axes (or CCA axes) corresponds to the minimum of $(s - 1)$, m , and $(N - 1)$, the number

of noncanonical axes (or CA axes) to the minimum of $(s - 1)$ and $(N - 1)$ (Legendre and Legendre 1998).

Partial canonical correspondence analysis (pCCA) can be used to remove the effect of a set of conditioning variables \mathbf{Z} . In a pCCA of \mathbf{Y} and \mathbf{X} conditioned for \mathbf{Z} , a CCA of \mathbf{Y} and \mathbf{X} is performed on the residuals of a CCA of \mathbf{Y} and \mathbf{Z} . This method has been proposed for partialling out the spatial component of variation based on trend surfaces (Borcard et al. 1992, Legendre and Legendre 1998).

For direct multi-scale ordination, two separate variogram matrices $\mathbf{Q}_{\text{fit}}^2(h)$ and $\mathbf{Q}_{\text{res}}^2(h)$ need to be calculated for the fitted values and for the residuals of a regression of \mathbf{Q} on \mathbf{X} . Hence, direct multi-scale ordination with CCA contains the following steps: (1) Perform CCA on the data. Matrix \mathbf{Q} is decomposed into the matrix of fitted values, \mathbf{Q}_{fit} , and the matrix of residuals, \mathbf{Q}_{res} . The matrix \mathbf{U}_{CCA} contains the canonical and matrix \mathbf{U}_{CA} the noncanonical eigenvectors. (2) Calculate a variogram matrix $\mathbf{Q}_{\text{fit}}^2(h)$ from \mathbf{Q}_{fit} and multiply it with \mathbf{U}_{CCA} to obtain variance profiles of the canonical eigenvalues (Eq. 7). (3) Calculate a variogram matrix $\mathbf{Q}_{\text{res}}^2(h)$ from \mathbf{Q}_{res} and multiply it with \mathbf{U}_{CA} to obtain variance profiles of the noncanonical eigenvalues (Eq. 7).

Diagnostic tool

The variance profiles of canonical and noncanonical eigenvalues, or the chi-square variograms of the ordination axes, partition the ordination results by distance. A plot of the variance profiles against distance can be used to identify the scales of overlapping and statistically uncorrelated multi-species patterns represented by the ordination axes (Wagner 2003).

The variogram of a matrix of type \mathbf{Q} can be derived directly from the squared Euclidean distance between observations (Eq. 6), thus avoiding the computationally intensive calculation of a variogram matrix of type $\mathbf{Q}^2(h)$. The R library "mso" in the Supplement calculates and plots variograms for different matrices of type \mathbf{Q} as returned by the function "cca" of the "vegan" library (information on R is available online).²

Often, it will suffice to characterize the overall spatial structure of the canonical and the noncanonical portions of variation. Under the assumption that all relevant factors have been included in the model, the variogram of the residual inertia \mathbf{Q}_{res} describes the total autocorrelation due to biotic processes, or false gradients, whereas the variogram of the constrained inertia \mathbf{Q}_{fit} contains the spatial dependence due to external factors, or true gradients. Plotting the variograms of the total, residual, explained, or conditioned inertia provides a simple yet powerful diagnostic tool for gradient analysis. It can be used to check one or more of the following assumptions related to geostatistical modeling or regression analysis.

Stationarity.—The spatial autocorrelation of a community, a species, or a group of similar species, may indicate the spatial scale of internal organization within the community, but only if the spatial dependence due to environmental factors is removed (Legendre 1993). This comes back to the assumption of intrinsic stationarity as a prerequisite for interpreting an empirical variogram. While there is no strict test of stationarity per se, there are some clear contra-indications. If the variogram of the residual inertia does not (asymptotically) reach a sill but continues to rise with distance, there probably is still a spatial trend in the data, which violates the assumption of (intrinsic) stationarity. In direct ordination, this may indicate that an important explanatory factor is missing from the model.

Autocorrelation of residuals.—CCA performs a multivariate regression analysis. The statistical tests commonly performed on regression coefficients are only valid if the residuals are (spatially) independent. If the empirical variogram of \mathbf{Q}_{res} depends on distance, the CCA residuals are spatially correlated. This means that the fitted model of the species–environment relationship should be interpreted with caution because its statistical significance is unknown.

Scale-dependent correlation structure.—Regression coefficients are only meaningful if the correlation structure is scale invariant (i.e., if the species–environment correlation does not depend on scale). If the PCA axes of two variables are correlated at any distance, the correlation between the two variables is scale dependent and the regression model should not be interpreted. The difference $\gamma_{\mathbf{Q}}(h) - (\gamma_{\mathbf{Q}_{\text{fit}}}(h) + \gamma_{\mathbf{Q}_{\text{res}}}(h))$ is twice the covariance $\gamma_{\mathbf{Q}_{\text{cross}}}(h)$ between \mathbf{Q}_{fit} and \mathbf{Q}_{res} for distance class h . Therefore, if the sum of the variograms of \mathbf{Q}_{fit} and of \mathbf{Q}_{res} deviates systematically from the variogram of \mathbf{Q} , the overall species–environment correlation depends on scale. The statistical significance of this deviation from the null hypothesis of a scale-invariant correlation structure can be tested by establishing a point-wise envelope for the variogram of \mathbf{Q} (i.e., by calculating the standard error for the mean of all $\gamma_{\mathbf{Q}}(a, b)$ within each distance class h):

$$\begin{aligned} \gamma_{\mathbf{Q}}(h) &\pm z_{\alpha} \text{SE}[\gamma_{\mathbf{Q}}(h)] \\ &= \gamma_{\mathbf{Q}}(h) \pm z_{\alpha} \sqrt{\frac{\text{Var}[\gamma_{\mathbf{Q}}(a, b | h_{ab} \approx h)]}{n_h}} \end{aligned} \quad (8)$$

where z_{α} is the critical value of the standard normal distribution $z \sim \text{norm}(0, 1)$ for a two-sided test with a significance level of α . A Bonferroni-type correction can be applied.

The oribatid mite example

To illustrate the use of the diagnostic tool, I repeated the analysis by Borcard et al. (1992) of a spatially referenced data set of oribatid mites (*Acari, Oribatei*) on the southern shore of a small bog lake, Lake Geai, on the territory of the Station de Biologie des Lauren-

² URL: <<http://www.r-project.org>>

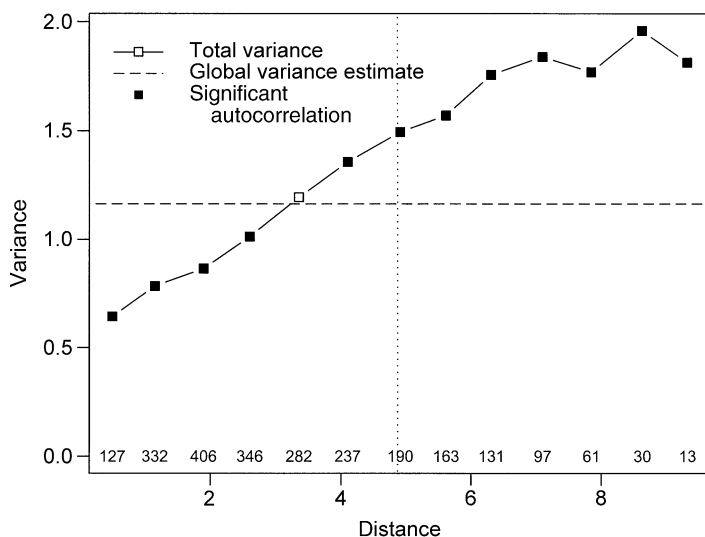


FIG. 2. Spatial partitioning of CA results. "Total variance" refers to the variogram of the total inertia of \mathbf{Y} (matrix \mathbf{Q}). Each point of the variogram represents the sum over all species of all basic units of spatial covariance (Eq. 3) averaged over all pairs of observations that fall into the respective distance class, plotted at the average geographic distance of the pairs. The dashed line indicates the global total inertia. A filled box indicates a significant autocorrelation of the total inertia for the respective distance class (Mantel permutation test). The dotted line corresponds to half the maximum distance between observations, beyond which the variograms should not be interpreted. The numbers above the x-axis indicate the number of unique pairs of observations that fall into each distance class.

tides of the Université de Montréal. Borcard et al. have used this data set extensively to illustrate methods for partialling out the spatial component of ecological variation (Borcard et al. 1992, Borcard and Legendre 1994, Meot et al. 1998). Readers are referred to these papers for a complete description of the sampling site and of the variables, as well as some biological information about the lake and its oribatid mite community. (The data set is available online.)³

A total of 70 cores, each 5 cm in diameter and 7 cm in depth, were extracted from a 10×2.6 m plot in the peat blanket. In the following analyses, I used these matrices. (1) Matrix \mathbf{Y} contained the Napierian logarithm [$y'_{ia} = \ln(y_{ia} + 1)$] of the abundances y_{ia} of 35 species of mites, with a total of 9800 individuals. (2) The matrix \mathbf{X} of environmental variables contained seven dummy variables for substratum (four species of *Sphagnum*, ligneous litter, bare peat, interface between *Sphagnum* species), three dummy variables for coverage density of the shrub cover (none, a few, many), two dummy variables for microtopography of the substratum (blanket and hummock), and two quantitative variables (density of the substratum in g/L of dry uncompressed matter, and water content in percent). (3) Matrix \mathbf{Z}_{xy} contained the spatial coordinates z_x and z_y ; (4) Matrix \mathbf{Z} , which defined the cubic trend surface fitted by Borcard et al. (1992), contained five transformations of the spatial coordinates (z_x , z_y , $z_x z_y$, z_x^2 , and z_y^2).

Using the function "cca" of the vegan library in R, I performed a CA of \mathbf{Y} , a CCA of \mathbf{Y} and \mathbf{X} , and a partial CCA (pCCA) with \mathbf{Y} , \mathbf{X} , and \mathbf{Z} .

For each matrix of type \mathbf{Q} (\mathbf{X} bar in the output of the "cca" function), I calculated an empirical variogram of the inertia (Eq. 6) with distance classes of 0.75 m

and breaks at 0.75, 1.5, and so forth. A point-wise envelope for the variogram of \mathbf{Q} was constructed with $z_\alpha = 2.9$, corresponding to an overall significance level of $\alpha = 0.05$ for 13 two-sided tests.

For CCA and pCCA, a significance test for spatial autocorrelation of the residual inertia was performed for each distance class using a Mantel test with 1000 permutations (Legendre and Legendre 1998). In order to account for the multiple tests for the 13 distance classes, I set the significance level for the two-sided test of the Mantel statistic to $\alpha = 0.05/13 = 0.00385$.

Calculations, significance tests, and plotting of the variograms were done with the functions "mso" and "plot.mso" in the R library "mso" in the Supplement, using the object created by the "cca" function and matrix \mathbf{Z}_{xy} as input. Please note that the functions "cca" and "rda" use \mathbf{X} to refer to the matrix of response variables (here: \mathbf{Y}) and \mathbf{Y} to refer to the matrix of explanatory variables (here: \mathbf{X}).

RESULTS

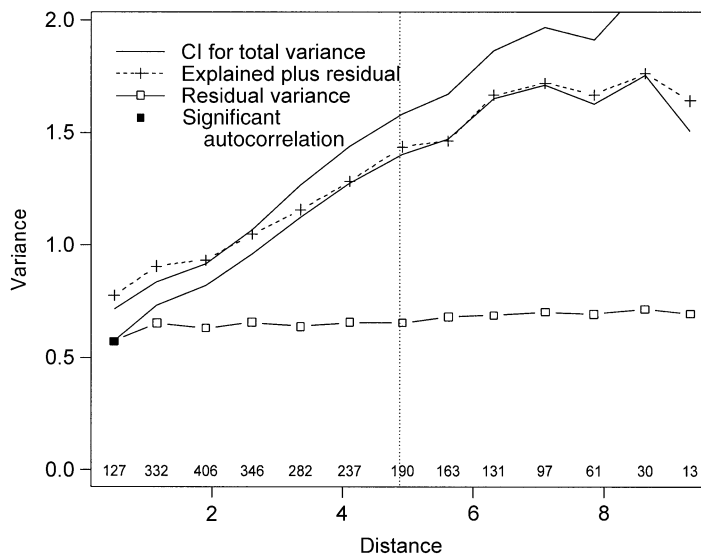
The global results of CCA and pCCA were identical with those reported by Borcard et al. (1992). Of the total inertia of 1.164, the environmental variables explained 44.8% in CCA, but only 13% in pCCA; that is, after accounting for the cubic trend surface represented by matrix \mathbf{Z} .

Stationarity

The variogram of the original matrix \mathbf{Q} (Fig. 2: "Total variance") showed a strong increase of inertia (variance) with distance, indicating the presence of a trend. This is a violation of the intrinsic stationarity assumption. After accounting for the environmental variables in \mathbf{X} , the variogram of the residual inertia \mathbf{Q}_{res} showed no dependence on distance, except for the smallest distance class with distances under 0.75 m (Fig. 3: "Residual variance"). Beyond 0.75 m, the residuals of the

³ URL: <http://www.fas.umontreal.ca/biol/casgrain/en/labo/oribates.html>

FIG. 3. Spatial partitioning of CCA results. "CI for total variance" is the point-wise envelope for the variogram of the total inertia of \mathbf{Y} (matrix \mathbf{Q}). "Residual variance" is the variogram of the residual inertia (matrix \mathbf{Q}_{res}). "Explained plus residual" is the sum of "Explained variance" (variogram of matrix \mathbf{Q}_{fit}) and "Residual variance." A filled box indicates a significant autocorrelation of the residual inertia for the respective distance class (Mantel permutation test; see Fig. 2 legend).



CCA of \mathbf{Y} and \mathbf{X} appeared to be spatially independent. There was no evidence that an important external factor could be missing from the model.

Spatial independence of regression residuals

The distance dependence of the total inertia (variogram of \mathbf{Q}) at larger distances was almost entirely explained by the environmental variables (Fig. 3); it was due to spatial dependence rather than spatial autocorrelation. The residual autocorrelation for the smallest distance class will cause a slight underestimation of the error variance. The variance estimate for the first distance class was 13% below the pooled estimate of the other distance classes and contributed 5% to the global estimate. The resulting underestimation of the error variance by 0.7% will inflate Type 1 errors in the regression analysis by very little.

Scale-dependent correlation structure

The sum of the variograms of \mathbf{Q}_{fit} and \mathbf{Q}_{res} (Fig. 3: "Explained plus residual") exceeded the point-wise envelope of the variogram of the original matrix \mathbf{Q} for the first three distance classes. This suggests that the species–environment correlation depends on scale and that a regionalized analysis should be performed. Simulations showed that the distance at which the curve intersects with the envelope (~ 2.5 m in this example) cannot be interpreted as the scale at which the correlation structure changes.

After accounting for the cubic trend surface defined by matrix \mathbf{Z} , neither the variogram of \mathbf{Q}_{fit} nor the variogram of \mathbf{Q}_{res} showed any striking distance dependence, although there was evidence of a slight negative autocorrelation of residuals at a distance around 1.25 m (Fig. 4). The sum of the explained and residual variance did not differ significantly from the variogram of the original matrix \mathbf{Q} . This suggests that the remaining

species–environment correlation is scale invariant and the regression parameters can be interpreted.

DISCUSSION

This paper shows how correspondence analysis and canonical correspondence analysis (and redundancy analysis, by analogy) can be partitioned by distance (multi-scale ordination) and integrated with geostatistics. The empirical chi-square variogram matrix $\mathbf{Q}^2(h)$ is a flexible framework for exploratory spatial analysis. It contains the empirical chi-square variograms of individual species (diagonal cells), their pair-wise cross variograms (off-diagonal cells), as well as the variogram of the total inertia (sum of the diagonal; cf., Wagner 2003). A spatial partitioning $\lambda_f(h)$ of each eigenvalue λ_f can be derived from the variogram matrix by multiplication with the eigenvector \mathbf{u}_f . Contrary to the principal component analysis case, the variance profile $\lambda_f(h)$ of an eigenvalue, or the chi-square variogram of an ordination axis, is not equal to an ordinary (Euclidean) variogram of the site scores on that axis, because the site scores contain a weighting by row totals. Further research is needed to answer the question of when to use which variogram of a CA or CCA axis.

The variograms of the total, residual, explained, and conditioned components of inertia, as they result from CA, CCA, and partial CCA, provide a simple yet powerful tool for characterizing the spatial autocorrelation and spatial dependence in community data and for checking important assumptions of geostatistical modeling (stationarity) and regression analysis (independence of residuals, scale-invariant correlation structure). The application to the oribatid mite data illustrated how apparent problems with nonstationarity or autocorrelation of the original community data can be attenuated by removing the effect of spatial dependence through direct ordination.

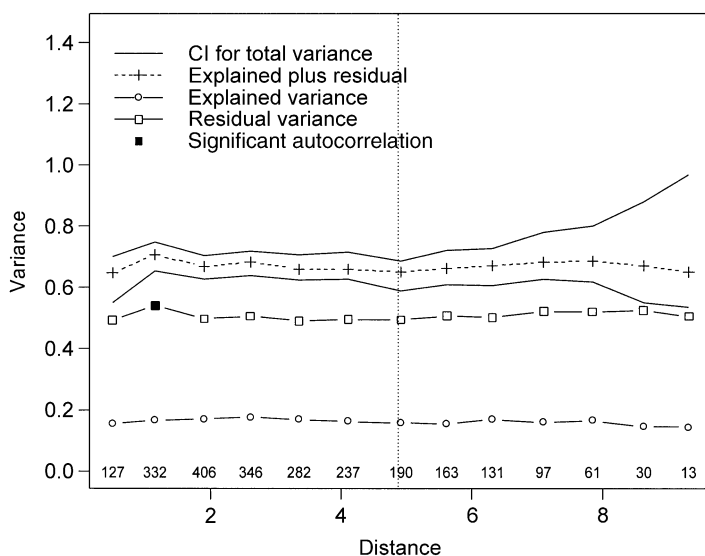


FIG. 4. Spatial partitioning of pCCA results. "CI for total variance" is the point-wise envelope for the variogram of the total inertia of \mathbf{Y} conditioned on \mathbf{Z} , "Explained variance" is the variogram of the inertia of \mathbf{Y} constrained by \mathbf{X} after accounting for \mathbf{Z} (matrix \mathbf{Q}_{fit}), and "Residual variance" is the variogram of the residual inertia (matrix \mathbf{Q}_{res}). "Explained plus residual" is the sum of "Explained variance" and "Residual variance." A filled box indicates a significant autocorrelation of the residual inertia for the respective distance class (Mantel permutation test; see Fig. 2 legend).

In the oribatid mite example, there was an important inconsistency between the results presented here and those by Borcard et al. (1992). Borcard found that 12.2% of the total inertia was spatially structured but could not be explained by the environmental variables. In the spatial partitioning of CCA results by multi-scale ordination (MSO), however, spatial autocorrelation appeared to be limited to distances smaller than 0.75 m, and there was no evidence of any cyclic pattern that could account for such a large portion of inertia. The large portion of nonenvironmental spatial structure identified by Borcard et al. (1992) may partly be due to a confounding of the effects of space and environment (Meot et al. 1998).

Multi-scale ordination showed a clear indication of a scale-dependent correlation structure for the oribatid mite data. After conditioning on the cubic trend surface, however, the species–environment correlation appeared to be scale invariant. This suggests that Borcard's method may be effective in removing problems of scale-dependent correlation, although the conditions under which this is true will need to be specified. A further question is how to interpret a residual correlation (i.e., what remains and what is lost by partialling out the spatial variation in both the environmental and the biotic data).

The spatial information that is summarized in the empirical chi-square variogram matrix can be used for answering questions about the organism-specific scale of response to the environment, the optimal spacing of sampling units, or the scale-dependent effect of environmental factors. The first two questions concern the residual inertia in \mathbf{Q}_{res} , whereas the last question focuses on \mathbf{Q}_{fit} and the underlying regression model.

The rephrasing of RDA and CCA based on regression (Legendre and Legendre 1998) confronts gradient analysis with the question of the validity of the fitted

multivariate regression model. This concerns residual analysis, but also the limitation to a linear regression model (Makarek and Legendre 2002). Even for a nonspatial analysis, existing software tools that perform RDA and CCA should be extended to include a choice of link functions (which would allow, for instance, logistic or Gaussian regression models) and diagnostic plots as customary for an analysis within the framework of the generalized linear model, GLM. In addition, the option of a spatial partitioning of regression or ordination results (MSO) and a diagnostic plot as presented in this paper would enable ecologists to evaluate the assumptions of spatial independence of residuals and a scale-invariant correlation structure and gain further insights into the data.

One goal of direct MSO may be to remove spatial dependence, or trend, in order to interpret the residual spatial pattern in a community as caused by biotic processes. This could be done by fitting local (e.g., spherical or exponential models) or periodic (e.g., hole-effect model) variogram models and combinations thereof to the residual empirical variogram matrix (see Legendre and Legendre 1998 for an overview of these models). There is an apparent contradiction of assumptions if one uses regression analysis for trend removal in order to study the spatial autocorrelation of the residual variance, as regression results are compromised by the presence of spatial autocorrelation in the residuals. The main effect of autocorrelation in the residuals is that the error variance is underestimated and the significance tests of the regression coefficients become too liberal. For an analysis of spatial autocorrelation, however, one would not be concerned too much about the significance of the regression coefficients used to remove trend, as long as one is not going to interpret the model and its parameters in a statistical sense. In addition, MSO provides an assessment of the

underestimation of the error variance. On the other hand, if the species–environment correlation is scale dependent and this is not incorporated in the regression model, the problems may be more severe. Not only will the parameter estimates be meaningless, but the residuals will contain a nonrandom error component that is not due to the interactions between organisms but to their scale-dependent response to the environment. This may be a more severe threat to the interpretation of residual autocorrelation than an inflated Type I error.

If the main aim is to describe the species–environment relationship, one might restrict CCA to the larger distances between observations. This requires identifying a range, or threshold distance, beyond which the residuals are considered spatially independent. A variogram matrix is then constructed that contains two matrices, one for distances up to the range and one for larger distances, and CCA is performed on the second one. However, the results still need to be checked for scale dependence in the correlation structure.

True spatial gradient analysis, however, needs to go further and incorporate both spatial dependence and spatial autocorrelation into the model. Even more importantly, ordination methods are needed that incorporate scale-dependent species–environment correlations. This may be attempted in the framework of factorial kriging analysis, or coregionalization analysis, where the spatial covariance is modeled as an additive set of uncorrelated variogram functions defined at characteristic scales. More research on methods, applications, and software for such a regionalized gradient analysis is needed.

ACKNOWLEDGMENTS

This research is part of a project funded by the Environmental Protection Agency (EPA) under STAR grant #R826764-01-0. Additional funding came from the Swiss National Science Foundation (SNF), NCCR Plant Survival, grant #5005-65748. John Wiens, Beatrice Van Horne, and Jonathan Bossenbroek from the Department of Biology, Colorado State University, and Otto Wildi, Christoph Scheidegger, and Beatrice Miranda-Gut from the Swiss Federal Research Institute for Forest, Snow and Landscape (WSL) in Birmensdorf, Switzerland, provided important support and valuable comments on the manuscript. I would like to thank Geoffrey M. Heebry, Kevin McGarigal, and an anonymous reviewer for their valuable comments.

LITERATURE CITED

- Addicott, J. F., J. M. Aho, M. F. Antolin, D. K. Padilla, J. S. Richardson, and D. A. Soluk. 1987. Ecological neighborhoods—scaling environmental patterns. *Oikos* **49**:340–346.
- Borcard, D., and P. Legendre. 1994. Environmental control and spatial structure in ecological communities: an example using Oribatid mites (Acari, Oribatei). *Environmental and Ecological Statistics* **1**:37–53.
- Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. *Ecology* **73**:1045–1055.
- De'ath, G. 1999. Principal curves: a new technique for indirect and direct gradient analysis. *Ecology* **80**:2237–2253.
- Fortin, M.-J. 1999. Spatial statistics in landscape ecology. Pages 253–279 in J. M. Klopatek and R. H. Gardner, editors. *Landscape ecological analysis: issues and applications*. Springer-Verlag, New York, New York, USA.
- Gauch, H. G. 1982. *Multivariate analysis in community ecology*. Cambridge University Press, Cambridge, UK.
- Goovaerts, P. 1992. Factorial kriging analysis—a useful tool for exploring the structure of multivariate spatial soil information. *Journal of Soil Science* **43**:597–619.
- Goovaerts, P. 1994. Study of spatial relationships between two sets of variables using multivariate geostatistics. *Geoderma* **62**:93–107.
- Goovaerts, P. 1999. Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma* **89**:1–45.
- Henley, S. 2001. The importance of being stationary. *Earth Science Computer Applications* **16**:1–3.
- Jongman, R. H. G., C. J. F. ter Braak, and O. F. R. van Tongeren. 1995. *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge, UK.
- Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* **74**:1659–1673.
- Legendre, P., and M.-J. Fortin. 1989. Spatial pattern and ecological analysis. *Vegetatio* **80**:107–138.
- Legendre, P., and L. Legendre. 1998. *Numerical ecology*. Elsevier, Amsterdam, The Netherlands.
- Makarek, V., and P. Legendre. 2002. Nonlinear redundancy analysis and canonical correspondence analysis based on polynomial regression. *Ecology* **83**:1146–1161.
- Matheron, G. 1982. *Pour une analyse krigéante des données régionalisées*. Centre de Géostatistique, Fontainebleau, France.
- Meot, A., P. Legendre, and D. Borcard. 1998. Partialling out the spatial component of ecological variation: questions and propositions in the linear modelling framework. *Environmental and Ecological Statistics* **5**:1–27.
- Monestiez, P., and M. Goulard. 1997. Analysing spatial genetic structures by multivariate geostatistics: study of wild populations of perennial ryegrass (*Lolium perenne*). Pages 1197–1208 in E. Y. Baafi and N. A. Schofield, editors. *Geostatistics Wollongong '96*. Kluwer, Dordrecht, The Netherlands.
- Myers, D. E. 1989. “To be or not to be . . . stationary—That is the question.” *Mathematical Geology* **21**(3):347–362.
- Noy-Meir, I., and D. Anderson. 1971. Multiple pattern analysis or multiscale ordination: towards a vegetation hologram. Pages 207–232 in G. P. Patil, E. C. Pielou, and E. W. Water, editors. *Statistical ecology: populations, ecosystems, and systems analysis*. Pennsylvania State University Press, University Park, Pennsylvania, USA.
- Oksanen, J., and P. R. Minchin. 1997. Instability of ordination results under changes in input data order: explanations and remedies. *Journal of Vegetation Science* **8**:447–454.
- Oksanen, J., and P. R. Minchin. 2002. Continuum theory revisited: what shape are species responses along ecological gradients? *Ecological Modelling* **157**:119–129.
- Rossi, R. E., D. J. Mulla, A. G. Journel, and E. H. Franz. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs* **62**:277–314.
- Tausch, R. J., D. A. Charlet, D. A. Weixelman, and D. C. Zamudio. 1995. Patterns of ordination and classification instability resulting from changes in input data order. *Journal of Vegetation Science* **6**:897–902.
- ter Braak, C. J. F. 1985. Correspondence-analysis of incidence and abundance data—properties in terms of a unimodal response model. *Biometrics* **41**:859–873.
- ter Braak, C. J. F., and I. C. Prentice. 1988. A theory of gradient analysis. *Advances in Ecological Research* **18**:271–317.
- Tilman, D., and P. Kareiva. 1997. *Spatial ecology: the role*

- of space in population dynamics and interspecific interactions. Princeton University Press, Princeton, New Jersey, USA.
- Ver Hoef, J. M., and D. C. Glenn-Lewin. 1989. Multiscale ordination—a method for detecting pattern at several scales. *Vegetatio* **82**:59–67.
- Wackernagel, H. 1998. Multivariate geostatistics. Second, completely revised edition. Springer, Berlin, Germany.
- Wagner, H. H. 2003. Spatial covariance in plant communities: integrating ordination, variogram modeling, and variance testing. *Ecology* **84**:1045–1057.
- Whittaker, R. H. 1967. Gradient analysis of vegetation. *Biological Reviews of the Cambridge Philosophical Society* **42**:207–264.
- Wiens, J. A. 1989. Spatial scaling in ecology. *Functional Ecology* **3**:385–397.

APPENDIX

A worked example that demonstrates the calculations is available in ESA's Electronic Data Archive: *Ecological Archives* E085-006-A1.

SUPPLEMENT

An R library “mso” for the spatial partitioning of ordination results from PCA, RDA, CA, and CCA, including a diagnostic plot, is available in ESA's Electronic Data Archive: *Ecological Archives* E085-006-S1.