

Community surveys through space and time: testing the space–time interaction in the absence of replication

PIERRE LEGENDRE,¹ MIQUEL DE CÁCERES, AND DANIEL BORCARD

Département de Sciences Biologiques, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7 Canada

Abstract. In order to test hypotheses about changes in the environment induced by man, including climatic change, ecologists are sampling portions of the environment repeatedly across time. This paper describes a method for testing a space–time interaction in repeated ecological survey data, when there is no replication at the level of individual sampling units (sites). This methodological development is important for the analysis of long-term monitoring data, including systems under anthropogenic influence. In these systems, an interaction may indicate that the spatial structure of community composition has changed in the course of time or that the temporal evolution is not the same at all sites. This paper describes ANOVA models corresponding to the steps leading to a solution to the problem, which is based on the representation of space and time by principal coordinates of neighbor matrices (PCNM eigenfunctions) in the ANOVA. Numerical simulations showed that ANOVA Model 5 was the model of choice for the analysis of the space–time interaction because it always had correct rates of Type I error, and its power was always equal to or higher than those of other possible models of analysis. If the hypothesis of absence of interaction is not rejected, one cannot conclude that a change has occurred in the spatial structure of the response data across time; one should follow the ordinary rules of two-way ANOVA if testing the significance of the main factors is of interest. If the hypothesis of absence of interaction is rejected, one should model the spatial structure of each time period in a separate way. One can also conduct a single test involving a separate model of the spatial structure for each time period. This paper presents two applications to real ecological data.

Key words: ANOVA, analysis of variance; Barro Colorado Island permanent forest plot; canonical redundancy analysis; climate change; ecological surveys; numerical simulations; principal coordinates of neighbor matrices (PCNM); space–time interaction; Trichoptera.

INTRODUCTION

In order to test hypotheses about changes in ecosystems induced by man, including climatic change, ecologists sample portions of the landscape repeatedly across time. Selected portions of the landscape are set aside for long-term studies; these may be transects or surfaces of different sizes. Research of this kind is conducted, for example, at the 26 sites belonging to the Long Term Ecological Research (LTER) network in the United States (*available online*).² The mission of that network is to develop knowledge and predictive understanding to conserve, protect, and manage ecosystems, their biodiversity, and the services they provide. The best indicator variables available to monitor such changes are multispecies ecological communities; the data tables are thus multivariate. Different types of communities can be monitored; some types (e.g., zooplankton) are known to react to short-term changes, whereas other communities (e.g., forests) respond more

slowly and are studied in order to monitor longer term changes.

In long-term monitoring, the available resources (manpower, time, money) are used to cover as much ground as possible during each sampling campaign. Sampling designs for transects and surfaces are made to contain as many different sampling units (sites) as possible. Fine-scale replication is sacrificed in the process. Space–time sampling conducted in that way produces an unreplicated repeated-measures design, which is a type of two-way factorial design. According to classical statistical theory, the lack of replication means that one is able to test the variability along space and time if one can reasonably assume that the interaction is zero, but a test of the space–time (S–T) interaction is impossible (Zar 1999). Within the classical statistical testing procedure, one can circumvent the problem of lacking replicates by defining strata formed of neighboring sampling units; the strata thus contain replicates (Legendre and McArdle 1997). When an impact can be clearly located in space and time, the Before/After – Control/Impact (BACI) sampling designs, introduced in the seminal book of Roger Green (1979), can be implemented. In this type of design, a significant result for the test of the interaction between

Manuscript received 3 February 2009; revised 22 April 2009; accepted 23 April 2009. Corresponding Editor: A. M. Ellison.

¹ E-mail: Pierre.Legendre@umontreal.ca

² (<http://www.lternet.edu>)

the two crossed factors (before-after impact times and control-impact areas) is the statistical indication that a significant impact has occurred (Green 1979:42).

In the no-replication situation, ecologists would still like to test the space-time interaction for two reasons: (1) because a significant interaction would indicate that the temporal structures differ from site to site, and that the spatial structures differ from time to time; and (2) because estimating an interaction sum-of-squares allows one to partial out its effect in statistical tests of the main effects. A significant S-T interaction in the structure of communities is the first statistical indication ecologists should look for to detect changes in ecosystems resulting from natural or man-made causes: an interaction indicates that the spatial structure of the community has changed through time, and we can infer from that finding that some process must be at work that has caused the change. The implication for analysis is that if the interaction is significant, one should carry out separate analyses of the temporal variance for the different points in space, or separate analyses of the spatial variance for the different times. Overall tests of separate spatial or temporal structures are also possible. The absence of a significant interaction, on the contrary, would indicate either that the differences among times can be modeled in the same way at all points in space – and conversely, or that the analysis did not have enough power to detect a significant interaction. Hence, alternative ways to cope with space-time interaction are needed. A procedure was proposed by Tukey (1949) to test for a multiplicative effect in two-way ANOVA models. That procedure, however, only tests for multiplicative main effects using a single degree of freedom. Tukey developed it mostly to determine if data ought to be transformed prior to two-way ANOVA. The procedure described here tests for more general forms of interaction.

In this paper, we will first describe five mathematical models for ANOVA of space-time data. Three of these models involve coding the space and/or time factors using a different coding scheme (principal coordinates of neighbor matrices [PCNM] eigenfunctions), which is more economical in degrees of freedom than binary or Helmert coding, thus allowing a test of the interaction to be performed. That method only applies to spatially or temporally structured response variables. We will use numerical simulations to check if that new form of coding leads to tests of significance for the interaction that have correct levels of Type I error. We will also verify the power of these tests using simulated data sets designed to contain S-T interactions. Following that, we will address the question of how to test the main factors in different circumstances. Finally, we will apply the new method to real ecological examples.

MATERIALS AND METHODS

Models to test the space-time interaction

Consider a study design with s sampling units across space and t sampling campaigns at different times. Let

us start with the well-known classical two-factor ANOVA model with interaction, assuming that there are r replicates for each combination of space unit and sampling campaign. From that starting point, we will move to new models developed to solve the problem described in the *Introduction* (Fig. 1).

Model 1: Two-way ANOVA crossed design with interaction.—In this model, the observations are replicated within each group (balanced design, r replicates per experimental condition):

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}. \tag{1}$$

In the equation, y_{ijk} is the response variate for the k th replicate in the i th spatial unit during the j th sampling campaign, μ is the overall mean, α_i is the effect of the i th spatial unit, β_j is the effect of the j th sampling campaign, $\alpha\beta_{ij}$ is the effect of the interaction between α_i and β_j , and ε_{ijk} is the error term. The total sum-of-squares, SS_{Tot} , and the corresponding degrees of freedom are partitioned in the following way:

$$SS_{Tot} = SS_{(\mu)} + SS_{Space} + SS_{Time} + SS_{Int1} + SS_{Res1} \tag{2a}$$

$$s \times t \times r = 1 + (s - 1) + (t - 1) + (s - 1) \times (t - 1) + s \times t \times (r - 1). \tag{2b}$$

There is one degree of freedom attached to the estimation of the overall mean using the data. The sum of squares corresponding to the interaction term, SS_{Int1} , captures all nonadditive effects of space and time. Contrary to the following models, there is no power decrease in Model 1 due to lack of fit since we estimate the residual SS from replicated values in the same experimental conditions (i.e., the expected value for the residual mean square is σ^2). However, if there are no replicates, $r = 1$; this leaves no degrees of freedom ($s \times t \times (r - 1) = 0$) to estimate the residual SS. That explains why, in the classical statistical framework, the S-T interaction cannot be tested in space-time studies without replication.

Model 2: Two-way crossed ANOVA design without interaction.—If we do not have replicates and we are willing to assume that the interaction is nonexistent, we may choose to move its SS to the residuals. This yields the following model:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk} \tag{3}$$

where the decompositions of SS and degrees of freedom are

$$SS_{Tot} = SS_{(\mu)} + SS_{Space} + SS_{Time} + SS_{Res2} \tag{4a}$$

and

$$s \times t = 1 + (s - 1) + (t - 1) + (s - 1) \times (t - 1). \tag{4b}$$

In this model, one cannot estimate a separate sum-of-squares for the interaction because there is a single observation in each group. If there were replicates, the

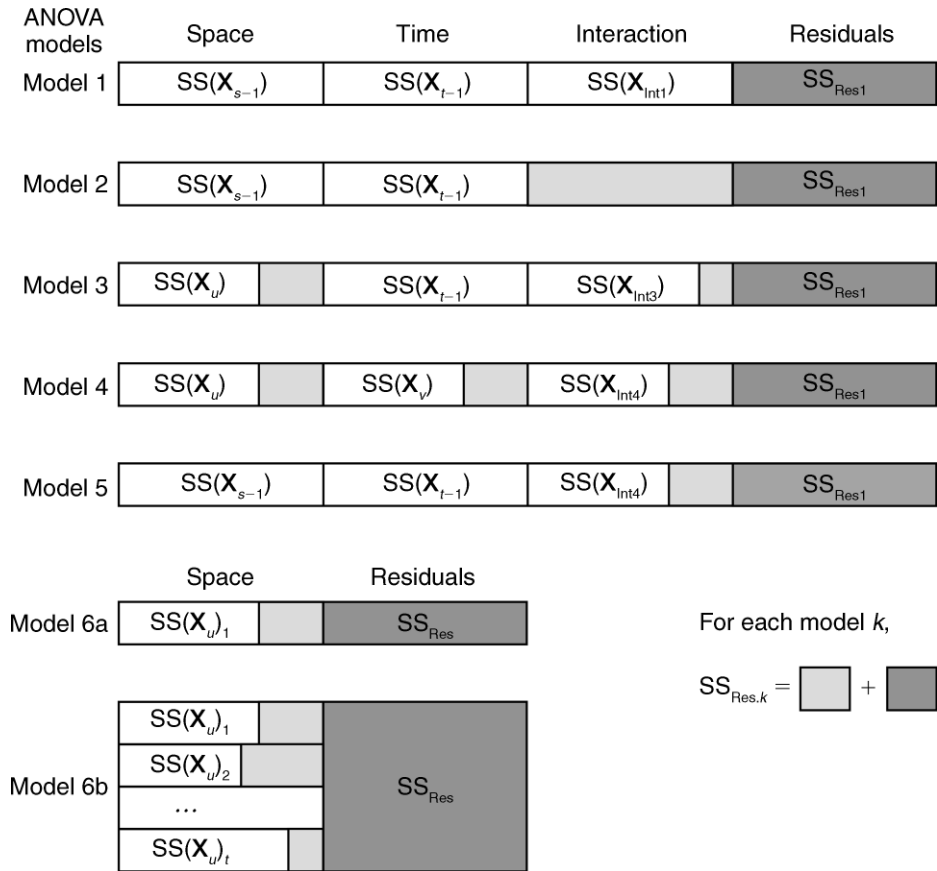


FIG. 1. Sum of squares partitioning for the six models of space–time analysis. Light shading corresponds to model lack of fit; these quantities are added to the residual sum of squares (dark shading). Models 6a and 6b (to test the significance of a spatial effect in the presence of a space–time interaction) are described in Appendix C.

residual sum-of-squares SS_{Res2} would actually be the sum of SS_{Int1} and SS_{Res1} of Model 1. If an interaction truly exists (i.e., if $\alpha\beta_{ij} \neq 0$ for some i and j combination), then there will be a lack of model fit, $SS_{LOF2} = SS_{Int1}$ where LOF designates lack of fit. The residual mean square will be overestimating σ^2 and the tests for significant overall spatial and temporal variation will lose power. In other words, using Model 2, the tests of significance for space and time will have correct Type I error only if the interaction is negligible (i.e., if $\alpha\beta_{ij} = 0$ for all i and j). Hence we see that analyzing the interaction is interesting both for testing its possible significant effect and for partialling its effect out in statistical tests of the main effects.

The approach adopted here for testing a space–time interaction considers the analysis of variance as a multiple regression problem using a design matrix, as described for example by Shaw and Mitchell-Olds (1993). Among the various forms of dummy variable coding, we select Helmert contrasts (Chambers and Hastie 1992, Venables and Ripley 2002), also called “orthogonal dummy variables” (see Appendix C in Legendre and Anderson 1999). In balanced crossed sampling designs (i.e., when the same number of sites are

surveyed during each sampling campaign), the Helmert contrast coding variables are orthogonal to one another (i.e., their scalar products are zero) and each one sums to zero. They are also orthogonal to the column of “ones” representing the overall mean effect. The s sites require $s - 1$ Helmert contrast variables for coding (design matrix X_{s-1}). Similarly, the t sampling campaigns require $t - 1$ Helmert contrast variables for coding (design matrix X_{t-1}). The interaction term is obtained by multiplying each of the $s - 1$ contrast variables coding for sites by each of the $t - 1$ contrast variables coding for times, producing $(s - 1) \times (t - 1)$ new variables (design matrix X_{Int}). These product variables are orthogonal to one another and have zero sums. They are also orthogonal to the main effect Helmert contrasts. As a consequence, the interaction can be tested as a term that is linearly independent of the variables coding for space and for time.

Model 1, re-expressed as a multiple regression model with design matrices, is the following using matrix notation:

$$y = \mathbf{1}\mu + X_{s-1}\alpha + X_{t-1}\beta + X_{Int}(\alpha\beta) + \varepsilon \quad (5)$$

where $\mathbf{1}$ is a column vector of ones; $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $(\boldsymbol{\alpha}\boldsymbol{\beta})$ are vectors of regression coefficients for the spatial, temporal, and interaction terms respectively, and $\boldsymbol{\varepsilon}$ is the vector of error terms. We will also change our notation for the sum-of-squares decomposition (Fig. 1, Model 1):

$$SS_{\text{Tot1}} = SS(\mathbf{1}) + SS(\mathbf{X}_{s-1}) + SS(\mathbf{X}_{t-1}) + SS(\mathbf{X}_{\text{Int1}}) + SS_{\text{Res1}}. \tag{6}$$

The degrees of freedom are the same as in Eq. 2b and are equal to the numbers of Helmert contrast variables used to express the various terms in the regression model.

In the case of Model 2 (Fig. 1, Model 2), our inability to test the interaction comes from the fact that $s - 1$ Helmert contrasts are necessary to code for space, and $t - 1$ for time. The same number of variables would be necessary using any other coding scheme (e.g., ordinary binary dummy variables) where the codes simply indicate that all s sites are different from one another, and likewise for time. As in the ANOVA model, when all explanatory variables are considered and there is no replication, one ends up with no degree of freedom for the error term. Therefore, we need a coding system for space and time that is more economical in the number of coding variables, in order to keep some degrees of freedom to estimate the residuals and their sum-of-squares. The trade-off is that the resulting model will produce a lack of fit, which will reduce power compared to an analysis with replication.

Model 3: Two-way ANOVA crossed design with space under-fitted.—Instead of not estimating a separate term for the interaction, as in Model 2, we may choose to risk underfitting other terms. For instance, if we fit the spatial structure using u variables, where $u < (s - 1)$, we end up with the following regression model:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_u\boldsymbol{\alpha} + \mathbf{X}_{t-1}\boldsymbol{\beta} + \mathbf{X}_{\text{Int3}}(\boldsymbol{\alpha}\boldsymbol{\beta}) + \boldsymbol{\varepsilon} \tag{7}$$

where design matrix \mathbf{X}_u codes now for the spatial fraction and \mathbf{X}_{Int3} contains $u \times (t - 1)$ variables now coding for the space-time interaction fraction. For a discussion of the nature of the u variables used to code for the spatial relationships see the *Encoding space and time using PCNM eigenfunctions* section. The total sum of squares partitioning is now (Fig. 1, Model 3)

$$SS_{\text{Tot}} = SS(\mathbf{1}) + SS(\mathbf{X}_u) + SS(\mathbf{X}_{t-1}) + SS(\mathbf{X}_{\text{Int3}}) + SS_{\text{Res3}}. \tag{8a}$$

Compared to Model 1, this model changes not only $SS(\mathbf{X}_{s-1})$ by $SS(\mathbf{X}_u)$, but also $SS(\mathbf{X}_{\text{Int1}})$ by $SS(\mathbf{X}_{\text{Int3}})$. Matrix \mathbf{X}_{Int3} is obtained by multiplying each of the variables in \mathbf{X}_u by each of the variables in \mathbf{X}_{t-1} . The partitioning of the degrees of freedom is

$$s \times t = 1 + u + (t - 1) + u \times (t - 1) + (s - u - 1) \times t. \tag{8b}$$

If we compare them to the degrees of freedom of Model 1, we see that, since $u < (s - 1)$ and $u \times (t - 1) < (s - 1) \times$

$(t - 1)$, this model saves some degrees of freedom that can be used to estimate the residual error. As a drawback, $SS(\mathbf{X}_u)$ can be lower than $SS(\mathbf{X}_{s-1})$, and $SS(\mathbf{X}_{\text{Int3}})$ can be lower than $SS(\mathbf{X}_{\text{Int1}})$. The lack of fit of the model, SS_{LOF3} , contains the lack of fit for the space and interaction terms with respect to Model 1. That is,

$$SS_{\text{LOF3}} = [SS(\mathbf{X}_{s-1}) - SS(\mathbf{X}_u)] + [SS(\mathbf{X}_{\text{Int1}}) - SS(\mathbf{X}_{\text{Int3}})]. \tag{9}$$

SS_{LOF3} is part of the residual sum-of-squares: $SS_{\text{Res3}} = SS_{\text{LOF3}} + SS_{\text{Res1}}$. The larger SS_{LOF3} is, the less power will the statistical tests have in Model 3. Furthermore, parametric tests will have a correct Type I error only when space and interaction effects are nonexistent (i.e., when $\alpha_i = 0$ for all i and $\alpha\beta_{ij} = 0$ for all ij). In the absence of an interaction effect, permutation tests for the time fraction can be made exact (i.e., with a correct Type I error) by restricting the permutations to within the same spatial units (Anderson and ter Braak 2003). That is, in permutation tests based on this model, we can handle the lack of fit of the spatial fraction but not that of the interaction.

Model 4: Two-way ANOVA crossed design with space and time under-fitted.—A similar situation arises when using fewer explanatory variables to express both the space and time factors:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_u\boldsymbol{\alpha} + \mathbf{X}_v\boldsymbol{\beta} + \mathbf{X}_{\text{Int4}}(\boldsymbol{\alpha}\boldsymbol{\beta}) + \boldsymbol{\varepsilon} \tag{10}$$

where \mathbf{X}_v codes now for the time fraction using $v < (t - 1)$ variables, and \mathbf{X}_{Int4} contains $u \times v$ variables now coding for the space-time interaction, obtained by multiplying each of the variables in \mathbf{X}_u by each of the variables in \mathbf{X}_v . The total sum-of-squares partitioning is as follows (Fig. 1, Model 4):

$$SS_{\text{Tot}} = SS(\mathbf{1}) + SS(\mathbf{X}_u) + SS(\mathbf{X}_v) + SS(\mathbf{X}_{\text{Int4}}) + SS_{\text{Res4}}. \tag{11a}$$

The corresponding partitioning of the degrees of freedom is

$$s \times t = 1 + u + v + u \times v + [s \times t - (u + v + u \times v) - 1]. \tag{11b}$$

Analogously to the previous models, $SS_{\text{Res4}} = SS_{\text{LOF4}} + SS_{\text{Res1}}$. Here, SS_{LOF4} contains the model lack of fit for the space, time, and interaction terms with respect to Model 1; that is,

$$SS_{\text{LOF4}} = [SS(\mathbf{X}_{s-1}) - SS(\mathbf{X}_u)] + [SS(\mathbf{X}_{t-1}) - SS(\mathbf{X}_v)] + [SS(\mathbf{X}_{\text{Int1}}) - SS(\mathbf{X}_{\text{Int4}})]. \tag{12}$$

Under this model, parametric statistical tests will have a correct level of Type I error only when all effects are nonexistent and the data only contains random error (i.e., when $\alpha_i = 0$ for all i , $\beta_j = 0$ for all j , and $\alpha\beta_{ij} = 0$ for all ij). Permutation tests for the main factors can be made exact only when the interaction is absent. Besides the fact that SS_{LOF4} will always be larger than SS_{LOF3} , more degrees of freedom are saved for the residuals in

Model 4 than in Model 3. As a consequence, significance tests are not necessarily less powerful for Model 4 than they are for Model 3.

Model 5: Two-way ANOVA crossed design with interaction under-fitted.—An interesting last possibility is to use Helmert contrasts for both space and time, as in Models 1 and 2, and use the u and v coding variables to generate the interaction variables and (under-) fit the interaction term as in Model 4, still saving degrees of freedom for the residuals:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_{s-1}\boldsymbol{\alpha} + \mathbf{X}_{t-1}\boldsymbol{\beta} + \mathbf{X}_{\text{Int4}}(\boldsymbol{\alpha}\boldsymbol{\beta}) + \boldsymbol{\varepsilon}. \quad (13)$$

The total sum of squares partitioning is as follows (Fig. 1, Model 5):

$$\begin{aligned} \text{SS}_{\text{Tot}} = & \text{SS}(\mathbf{1}) + \text{SS}(\mathbf{X}_{s-1}) + \text{SS}(\mathbf{X}_{t-1}) + \text{SS}(\mathbf{X}_{\text{Int4}}) \\ & + \text{SS}_{\text{Res5}} \end{aligned} \quad (14a)$$

where $\text{SS}_{\text{Res5}} = \text{SS}_{\text{LOF5}} + \text{SS}_{\text{Res1}} = [\text{SS}(\mathbf{X}_{\text{Int1}}) - \text{SS}(\mathbf{X}_{\text{Int4}})] + \text{SS}_{\text{Res1}}$. The corresponding decomposition of degrees of freedom is

$$\begin{aligned} s \times t = & 1 + (s - 1) + (t - 1) + u \times v \\ & + [(s - 1) \times (t - 1) - u \times v]. \end{aligned} \quad (14b)$$

For example, in the absence of replication ($r = 1$), if $t = s = 10$ and we choose a coding system such that $u = v = 5$, then the residual df will be $[(s - 1) \times (t - 1) - u \times v] = 81 - 25 = 56$. This model presents two advantages: (1) it would allow testing the interaction in the absence of replication, and (2) the test of the interaction would have a correct rate of Type I error, which is a rejection rate of H_0 equal to the significance level when the interaction is nonexistent (i.e., when $\alpha\beta_{ij} = 0$ for all ij); this is precisely what the null hypothesis states for the effect of the interaction term. Model 5 is thus probably the best model to test the interaction in the absence of replication. It may, however, be slightly less powerful than Model 4 because fewer degrees of freedom can be assigned to the residuals.

Summarizing this section, if one has replicates, Model 1 is the correct choice. If not, then space–time interaction can safely (in terms of Type I error rate) be tested using Model 5. If interaction turns out to be nonsignificant, one can test for the main effects using Model 2. The absence of a significant interaction indicates either (1) that the differences among times can be modeled in the same way at all points in space, and conversely, or (2) that an interaction was present but the number of observations was insufficient to obtain a significant result (n too small, lack of power; Type II error). In contrast, if the interaction is significant, one should perform separate spatial analyses for the different sampling campaigns and/or separate time series analyses for the different spatial units.

The F statistic

When the response is univariate, the F statistic is constructed in the usual way. Using the representation

of ANOVA as a multiple regression problem, as in the previous subsections, the equation for the F statistic is

$$F = \frac{\text{SS, factor of interest}/m}{\text{SS, residuals}/(n - 1 - m - q)} \quad (15)$$

where m is the number of variables coding for the factor of interest in the analysis, n is the number of observations, and q is the number of variables coding for any other factor or the interaction in the analysis, besides the factor of interest.

When the response table is multivariate with p columns, Miller (1975) has shown that the F statistic has the following form:

$$F = \frac{\text{SS, factor of interest}/pm}{\text{SS, residuals}/p(n - 1 - m - q)}. \quad (16)$$

The multiplication of the degrees of freedom by p in the numerator and denominator is due to the fact that each parameter is computed p times during the calculations when there are p response variables. Using numerical simulations, Miller (1975) also showed that in canonical redundancy analysis (RDA), in the special case where (1) the response variables are all standardized to mean = 0 and variance = 1, and (2) the error is normal, the F statistic constructed in that way can be tested using the Fisher–Snedecor F distribution with degrees of freedom pm for the numerator and $p(n - 1 - m - q)$ for the denominator. If these two conditions are not met, the F statistic must be tested by permutation.

We will use that same F statistic for our tests of significance in the multivariate case. We will not carry out parametric tests of significance, though, because typically community composition data are not standardized, and the error is certainly not normally distributed. We will be using permutation tests to alleviate the problems associated with the error distribution. Even though degrees of freedom do not really matter in permutation tests, it is still recommended to use an F statistic for such tests because F is pivotal (a pivotal statistic is a statistic whose distribution, under the null hypothesis, does not depend on the value of the measured effect or unknown parameters). For the calculation of F , Eq. 16 simplifies to Eq. 15 because p is present in both the numerator and denominator.

Eq. 15 for the F statistic can be used for the tests of the main factors and the interaction when space and time are considered to be fixed factors (Model I ANOVA; Zar 1999). This is not always the case. When the two factors are considered random (Model II ANOVA), the denominator of the F statistic contains the interaction mean square for the tests of the two main factors. In a mixed model with a fixed and a random factor (Model III ANOVA), the denominator of the F statistic contains the interaction mean square for the test of the fixed factor. The STI package provided in the Supplement provides options for space and time being treated as fixed or random factors.

Encoding space and time using PCNM eigenfunctions

In a previous section, different regression models have been presented for testing the significance of space, time, and interaction, along with their advantages and drawbacks. We have not yet described, however, a coding system for space and time that would be more economical in coding the main factors. The purpose here is to replace binary or Helmert contrast coding by continuous explanatory variables without losing too much model fit. We will make use of the extra information that we possess about the physical relationships among the sampling units in space and time. For example, if we are willing to assume that the sampling units through time form a linear gradient, we could replace the $t - 1$ variables coding for sampling times by a single continuous “time” variable containing dates (e.g., sampling years). That would be a rather strong assumption; it would limit the ability of the test to identify significant variation if the response is not linear with time.

A less restrictive form of coding is to represent the spatial and temporal relationships among the sampling units by principal coordinates of neighbor matrices (PCNM eigenfunctions; Borcard and Legendre 2002, Borcard et al. 2004), which are a form of Moran’s eigenvector maps (MEM; Dray et al. 2006). PCNM eigenfunctions represent a spectral decomposition of the spatial (or temporal) relationships among sampling sites (or times). They are orthogonal to one another, and may be used to fit many distinct nonlinear situations. PCNM eigenfunctions are a more parsimonious representation of the spatial and temporal relationships among observations than ordinary dummy variables or Helmert contrasts. How PCNM eigenfunctions are computed from the sampling locations and time positions (hence not from the response data), and used in ecological analysis, has been described in detail in recent papers cited above; these points will not be repeated here to save space. Note that these eigenfunctions are limited to modeling spatial and temporal relationships. Indeed, space and time are the continuous bases in which contagious spatial and temporal processes occur, which can create spatial and temporal structures in response data. PCNM eigenfunctions cannot be used to model regular factors encountered in experimental designs.

To model the space and time variation, we are proposing to use $s/2$ or $t/2$ PCNM functions; $s/2$ and $t/2$ are rounded down to the closest integer value. More eigenfunctions could be used, but this would leave fewer degrees of freedom for the estimation of the error mean square in the test of the interaction. For example, 10 equispaced sampling times can be modeled by five PCNM functions representing sine functions of decreasing periods. The references mentioned previously have shown that PCNM functions can model spatial or temporal variation having any shape, as long as the size of the structures to be modeled is larger than the period of the smallest PCNM variables used in the analysis.

Sampling each point in space (S) during each sampling campaign (T) creates an orthogonal design. For that reason, the PCNM functions, which are orthogonal within each set (S, T), are also orthogonal between sets. In space–time analysis, the S–T interaction can be modeled by creating variables that are the products of each S–PCNM by each T–PCNM. The S–T interaction variables are orthogonal to the S–PCNMs and the T–PCNMs. Orthogonality of the variables representing space, time, and interaction allows the analysis of each factor without interference (linear dependence) on the part of the other factors, just like in ordinary ANOVA.

NUMERICAL SIMULATIONS

Numerical simulations were conducted to assess the Type I error rate of ANOVA Models 3, 4, and 5 described in Fig. 1 and to determine which model had the highest power to test the S–T interaction in the absence of replication. The generation of the simulated data and the simulation scenarios are described in Appendix A. The results, reported in detail in Appendix B, are summarized here.

Simulations under Scenario 1 (random data) showed that in the absence of a spatial or temporal structure, all tests of the S–T interaction had correct rates of Type I error.

Simulations under Scenario 2 (surfaces with no S–T interaction, no temporal structure, and a spatial autocorrelated structure common to all sampling times) and Scenario 3 (surfaces with spatial and temporal structures but without space–time interaction) showed that the rejection rate of the S–T interaction was much lower than the nominal 5% significance level in ANOVA Models 3 and 4. It was, however, near the significance level for ANOVA Model 5.

We will now focus on the test of the space–time interaction under simulation Scenario 4 (surfaces with spatial and temporal structures as well as space–time interaction) using ANOVA Models 3, 4, and 5. This is the most interesting case in the present study. The rejection rates displayed in Appendix B (Table B4a–c) were first analyzed by ANOVA to detect effects of the simulation factors: the use of univariate or multivariate simulated response data, the amounts of spatial and temporal autocorrelation in the data, the number of spatial and temporal units in the simulated data, and the ANOVA model. The ANOVA results (not shown) indicated that interactions were present among the simulation factors. As a consequence, the univariate results were treated separately from the multivariate results, and the two combinations of spatial and temporal autocorrelation were also treated separately, resulting in four separate analyses. In the separate ANOVAs, the two factors “ANOVA model” and “sum of the numbers of spatial and temporal units” were significant in nearly every case (differences among ANOVA models were not significant only in univariate

data, spatial autocorrelation = 20 combined with temporal autocorrelation = 5). The results are illustrated in Fig. 2; they summarize the raw results shown in Appendix B (Table B4). To facilitate comparisons between methods, the rejection rates, on a scale from 0 to 1, are represented on a modified scale from -0.5 to 0.5, with negative values represented by white squares and positive values by black squares. The following observations can be made on Fig. 2. (1) Multivariate response data (b, d: more black squares) provided higher power to the S-T interaction test than univariate data (a, c: more white squares). (2) Power of the S-T interaction test under ANOVA Model 5 was similar to that of Model 4 for equal amounts of spatial and temporal autocorrelation (a, b: same number of + signs for Models 4 and 5), but power was higher in Model 5 for unequal amounts of autocorrelation in the data (c, d: more + signs for Model 5). (3) When spatial and temporal autocorrelation were identical (a, b), the S-T interaction test was less powerful under Model 3 (more - signs) than under Models 4 and 5 (more + signs). Model 3 was, however, equal to or more powerful than Model 4 for unequal amounts of autocorrelation (c, d), although Model 5 remained the most powerful (more + signs).

In summary, our simulation results indicate that in the absence of replication, ANOVA Model 5 is preferable to test S-T interactions: the rate of Type I error corresponds to the nominal significance level and that model provides maximum power in all situations. The simulation results are in agreement with and support the conclusion reached from theoretical considerations at the end of the section *Models to test the space-time interaction*.

When the hypothesis of no interaction is not rejected and one has strong reasons to believe that the interaction is zero, one possibility (but see the first recommendation in the *Conclusion* section) is to analyze the spatial and temporal structures using the classical test of S and T without replication, which is our ANOVA Model 2. We had to consider that model in our simulations because this is how most researchers have analyzed space-time data without replication until now. The simulation results discussed in Appendix A indicate that when there was truly no S-T interaction in the data, Model 2 was the most powerful, followed by Model 5 and, finally, Models 3 and 4. In Scenario 4, when an interaction was present in the data but was not detected (Type II error), Model 4 was the most powerful to detect significance of the main factors, followed by Models 3 and 5, and finally Model 2.

APPLICATIONS TO REAL ECOLOGICAL DATA

Trichoptera data

A transect of 22 “week” model emergence traps (LeSage and Harrison 1979), 0.25 m² in capture area, was laid along the outflow stream of Lac Cromwell (45°59'13" N, 73°59'47" W) on the territory of Station de Biologie des Laurentides (Université de Montréal). The

stream was surveyed during a period of 100 days; the traps were set on 24 May and visited daily from 25 May to 1 September 1984. The number of emerging adults belonging to 56 Trichoptera (Insecta) species was noted every day. Insect counts within each 10-day period were pooled in order to avoid an excess of zero values. The data set to be analyzed thus comprises 22 sites, 10 time periods (each one pooling 10 days of captures), and 56 species. Our aims in this application were to (1) check for the presence of space-time interaction in the data, and (2) if an interaction is present, describe it by identifying groups of points that were homogeneous in space-by-time space.

We first transformed the Trichoptera abundance values using the $y' = \log(y + 1)$ transformation. The multivariate log-transformed data was then used as the response table for testing for the presence of a space-time interaction using Model 5, which was shown to be the most powerful for that test in the simulation study. The space-time interaction was highly significant ($R^2 = 0.205$, $P = 0.0001$ after 9999 permutations).

After a Hellinger transformation of the raw abundance values (Legendre and Gallagher 2001), the 220 data points (22 sites, 10 time periods) were analyzed by *K*-means partitioning in order to visualize the interaction. The “cascadeKM” function of the “vegan” package (Oksanen et al. 2007) in the R statistical language was used to perform this analysis. The simple structure index (ssi) was called upon to identify the best partition; the index was maximum for five groups. A spatiotemporal map of the partition is presented in Fig. 3. Group 2 (triangles) dominated the emergences during the first two 10-day periods, but group 1 (circles) dominated the next three periods. During time periods 6–7, a well-defined spatial pattern appeared with blocks of traps belonging to different groups: groups 3 (plus), then 5 (diamonds), then 1 (circles), then 5 (diamonds), then 3 (plus), then 5 (diamonds), then 4 (×), then 2 (triangles), then 4 (×), then 5 (diamonds). Time period 8 (3–12 August) served as a transition to the last two time periods where groups 2 (triangles), 3 (plus), and 4 (×) became dominant.

Indicator species were identified for each group using the “duleg” function of the “labdsv” library (Roberts 2006), which implements the IndVal method of Dufrêne and Legendre (1997). The IndVal method is now part of a larger framework described by De Cáceres and Legendre (2009). After Holm correction for multiple testing (56 simultaneous tests had been performed), 17 species were identified as significant indicators of the five groups of space-time points. The results are discussed in Appendix D.

Barro Colorado Island permanent forest plot

We reanalyzed the Barro Colorado Island (BCI) permanent forest plot data in order to learn about the temporal stability of the spatial distribution of trees in that famous forest plot. We used the stem-based plot

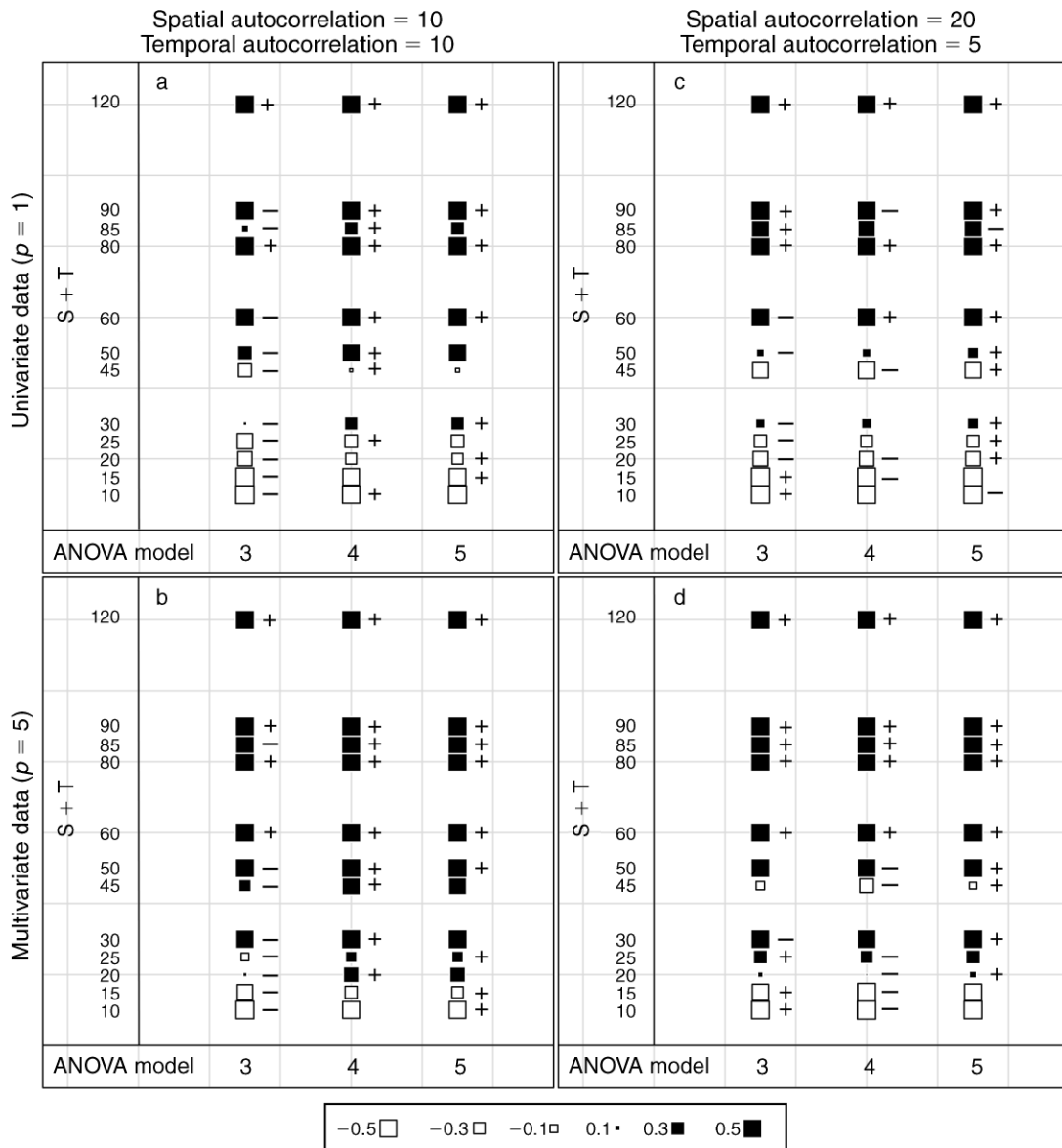


FIG. 2. Simulation results for three ANOVA models used to test space-time interaction. Values of rejection rates (see *Numerical simulations*) are represented on a scale from -0.5 (large white squares) to +0.5 (large black squares). The signs indicate which symbol in a row represents the smallest (-) and the largest (+) value; there are more than one “smallest” or “largest” value in some rows. The results are divided into four panels by the simulation factors: univariate data (a, c) or multivariate data (b, d), and importance of spatial and temporal autocorrelation (panels a, b vs. c, d). In each panel, three columns give the rejection rates for ANOVA Models 3, 4, and 5. The rows correspond to the sum of the numbers of spatial and temporal units in the simulated data. Each point is the result of 1000 independent simulations. On the ordinate, p equals the number of variables (columns) in the response data tables.

data covering four censuses (1982–1983, 1985, 1990, and 1995) and counted the trees with 10 mm diameter at breast height (dbh) or more, by species, in grid cells of 20×20 m. For each census, we obtained a data table containing 1250 rows (cells). The four data tables were merged into one large data table with 5000 rows and 315 species. We used a variant of Model 5 to test the space-time interaction. The interaction term was modeled using PCNM variables for space, whereas Helmert

contrasts were used to code for the four sampling campaigns. We first tested the space-time interaction on each of the 315 tree species separately, after log-transforming their abundances ($y' = \log(y + 1)$). We used $\alpha = 0.05$ and 999 permutations for the analysis of each species. According to our results, about 43% of the tree species had significantly changed their spatial structures across the four censuses. Changes in the spatial distribution of a species could result from an

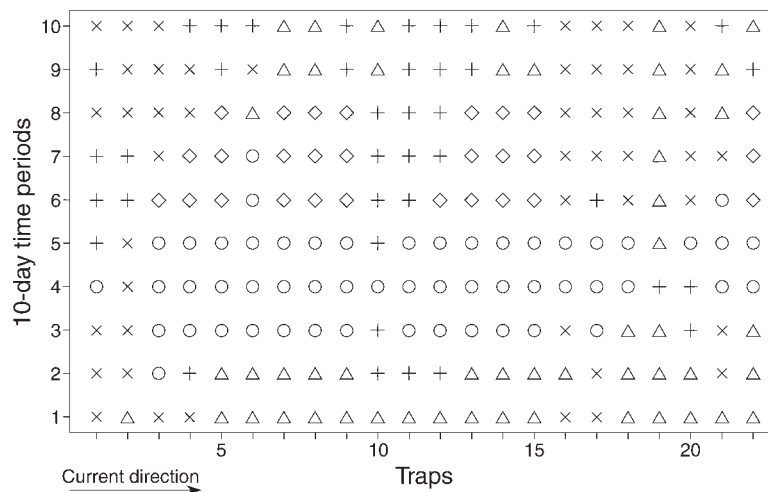


FIG. 3. Spatiotemporal map showing the *K*-means partition of the emergence trap observations into five groups. Abscissa: trap line with traps numbered 1 (upstream) to 22 (downstream). Ordinate: 10-day time periods numbered 1 to 10. Symbols for groups: 1, circle; 2, triangle; 3, “plus” sign (+); 4, “times” sign (×); 5, diamond.

invasive process. However, while *Drypetes standleyi* (Euphorbiaceae) is considered to be invading the plot from the east (Harms et al. 2001), we found that its spatial structure did not change significantly during the study period ($P = 0.498$). Other spatial changes may have originated due to strong perturbations or climatic change. Condit (1998) reported a decline in nearly all BCI tree and shrub species associated with moist microhabitats due to a reduction in rainfall and a lengthening of the dry seasons. We found that among the 31 species associated with slope habitats (Harms et al. 2001), which are moister than the plateaus in BCI, 23 (74%) had significantly changed their spatial distributions. Examples are the crash of dominant moisture specialists like *Poulsenia armata* (Moraceae, $P = 0.001$) and the displacement of the population of *Beilschmiedia pendula* (Lauraceae, $P = 0.001$) towards areas of higher slope (Fig. 4).

We also carried out a space–time interaction test on the multivariate Hellinger-transformed abundance data, which turned out to be highly significant ($P = 0.001$ after 999 permutations). Hence, the changes at the species level were strong enough to allow the detection of changes in the spatial distribution of species composition at the community level.

In another application, Laliberté et al. (2009) studied the space and time factors, as well as the space–time interaction, in a temperate forest understory where tree seedling abundances had been monitored during a 9-yr period at 40 permanent plots.

CONCLUSION

The testing procedure described in this paper allows ecologists to study the presence of an interaction in univariate or multivariate (e.g., community structure) response data without sampling replication. This was

not possible using classical analysis of variance for unreplicated space–time data.

Regarding the comparison of ANOVA models for testing the space–time interaction, we found that Model 5 had a better Type I error rate than Models 3 and 4: in the presence of main factor effects, the Type I error of Model 5 was always correct for univariate response data and asymptotically correct for multivariate data, while those of Models 3 and 4 were too low. Our results indicate that Model 5 is the model of choice for the analysis of the S–T interaction because its power was always equal to or higher than those of Models 3 and 4. Our recommendation is thus to perform first a test of the S–T interaction using ANOVA Model 5, and then proceed as follows.

(1) If the hypothesis of no interaction is not rejected, one cannot conclude that a change has occurred in the spatial structure of the response data across time. In that case, one should follow the ordinary rules of two-way ANOVA if testing the significance of the main factors is of interest. (a) In ANOVA Model II (two random factors) or III (mixed model with a fixed and a random factor), if the interaction is not significant at a very high significance level (e.g., $\alpha = 0.25$; Underwood 1997: section 9.7), the interaction and residual mean squares may be pooled following the rules stated in Underwood (1997: section 11.7). That decision is subject to the possibility that a Type II error occurred during the test of the interaction. Our ANOVA Model 2 is an unweighted version of post hoc pooling. This should never be done in Model I ANOVA (two fixed factors). (b) A more generally conservative attitude is to always stick to the ANOVA model chosen a priori. Always use that option in Model I ANOVA (two fixed factors).

(2) If the hypothesis of no interaction is rejected, the user is allowed to model the spatial structure of each time period in a separate way (Underwood 1997: section

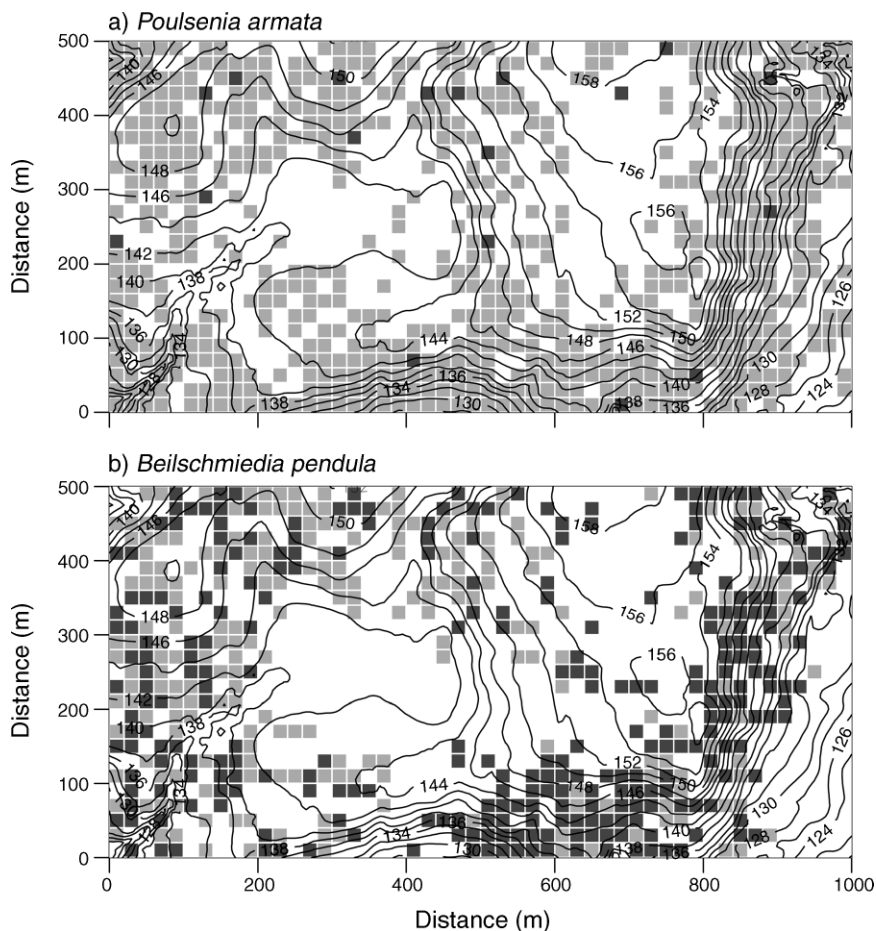


FIG. 4. Changes in numbers of individuals in the Barro Colorado Island (BCI), Panama, permanent forest plot for two species associated with slopes that significantly changed their distributions between the 1982–1983 and 1995 censuses. Light gray squares indicate loss, whereas dark gray squares indicate gain of individuals.

10.8). The test can be carried out separately for each sampling time (Appendix C: Model 6a: one-factor ANOVA model). One can also conduct a single test involving a separate model of the spatial structure for each time period (Appendix C: Model 6b: stacked one-factor ANOVA model). The temporal structure of each sampling point can be analyzed in the same way by interchanging space and time.

A test of an interaction without replication has, of course, less power than a test conducted using replicated data. The problem challenged in this paper is, however: What can one do to test an interaction in the absence of replication? This question is of interest to all researchers who conduct spatial ecological surveys repeated across time. Identifying an interaction in the community composition data is a clear signal that the community has reacted to changing environmental conditions (or other causes) by modifying its spatial structure. The new test should prove useful to detect natural or man-made changes in ecological communities and ecosystems.

ACKNOWLEDGMENTS

We are grateful to Pierre-Paul Harper, Université de Montréal, for permission to use the Trichoptera emergence data as an example in this paper. This research was supported by NSERC Grant no. 7738 to P. Legendre.

LITERATURE CITED

- Anderson, M. J., and C. J. F. ter Braak. 2003. Permutation tests for multi-factorial analysis of variance. *Journal of Statistical Computation and Simulation* 73:85–113.
- Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* 153:51–68.
- Borcard, D., P. Legendre, C. Avois-Jacquet, and H. Tuomisto. 2004. Dissecting the spatial structure of ecological data at multiple scales. *Ecology* 85:1826–1832.
- Chambers, J. M., and T. J. Hastie. 1992. Statistical models. Chapter 2 in J. M. Chambers and T. J. Hastie, editors. *Statistical Models in S*. Wadsworth and Brooks/Cole, Pacific Grove, California, USA.
- Condit, R. 1998. Ecological implications of changes in drought patterns: shifts in forest composition in Panama. *Climatic Change* 39:413–427.
- De Cáceres, M., and P. Legendre. 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology* 90:3566–3574.

- Dray, S., P. Legendre, and P. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* 196:483–493.
- Dufrène, M., and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs* 67:345–366.
- Green, R. H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley and Sons, New York, New York, USA.
- Harms, K. E., R. Condit, S. P. Hubbell, and R. B. Foster. 2001. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *Journal of Ecology* 89:947–959.
- Laliberté, E., A. Paquette, P. Legendre, and A. Bouchard. 2009. Assessing the scale-specific importance of niches and other spatial processes on beta diversity: a case study from a temperate forest. *Oecologia* 159:377–388.
- Legendre, P., and M. J. Anderson. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs* 69:1–24.
- Legendre, P., and E. D. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129:271–280.
- Legendre, P., and B. H. McArdle. 1997. Comparison of surfaces. *Oceanologia Acta* 20:27–41.
- LeSage, L., and A. D. Harrison. 1979. Improved traps and techniques for the study of aquatic emerging insects. *Entomological News* 90:65–78.
- Miller, J. K. 1975. The sampling distribution and a test for the significance of the bivariate redundancy statistic: a Monte Carlo study. *Multivariate Behavioral Research* 10: 233–244.
- Oksanen, J., R. Kindt, P. Legendre, and R. B. O'Hara. 2007. *vegan: community ecology package*. R package version 1.9-25. (<http://cran.r-project.org/>)
- Roberts, D. W. 2006. *labdsv: Laboratory for Dynamic Synthetic Vegetation Phenomenology*. R package version 1.2-2. (<http://cran.r-project.org/>)
- Shaw, R. G., and T. Mitchell-Olds. 1993. ANOVA for unbalanced data: an overview. *Ecology* 74:1638–1645.
- Tukey, J. W. 1949. One degree of freedom for nonadditivity. *Biometrics* 5:232–242.
- Underwood, A. J. 1997. *Experiments in ecology: their logical design and interpretation using analysis of variance*. Cambridge University Press, Cambridge, UK.
- Venables, W. N., and B. D. Ripley. 2002. *Modern applied statistics with S*. Fourth edition. Springer-Verlag, Berlin, Germany.
- Zar, J. H. 1999. *Biostatistical analysis*. Fourth edition. Prentice Hall, Upper Saddle River, New Jersey, USA.

APPENDIX A

Numerical simulations for Type I error and power of the five ANOVA models: method and results (*Ecological Archives* E091-019-A1).

APPENDIX B

Tables of simulation results (*Ecological Archives* E091-019-A2).

APPENDIX C

Tests of S and T in the presence of a significant interaction (*Ecological Archives* E091-019-A3).

APPENDIX D

Indicator species of the Trichoptera example (*Ecological Archives* E091-019-A4).

SUPPLEMENT

The STI package (source code and compiled libraries for Windows and Mac OS X), an R language library for the analysis of the main factors space and time and the interaction in space–time studies using permutation tests (*Ecological Archives* E091-019-S1).