

*Appendices and Supplement to:*

Legendre, P., M. De Cáceres, and D. Borcard. 2010. Community surveys through space and time: testing the space-time interaction in the absence of replication. *Ecology* 91(1): 262-272.

**APPENDIX A*****Ecological Archives E091-019-A1*****NUMERICAL SIMULATIONS: METHOD AND RESULTS**

Simulated numerical data with known properties were used to assess the type I error rate and power of the proposed test of S-T interaction. Univariate or multivariate data were generated on “maps” or “surfaces” whose abscissa represented a transect in space (S) while the ordinate represented sampling campaigns through time (T). Two main types of simulated data were generated:

*Random data.* — In the simplest simulations, done in order to assess the rate of type I error, random normal deviates were attributed to all sampling units.

*Autocorrelated data.* — Data fields were generated containing spatially and/or temporally autocorrelated data; this was accomplished by generating autocorrelated data with known variograms along the abscissa (spatial autocorrelation) and/or the ordinate of the map (temporal autocorrelation). The data generation program has been used in previous papers to study the consequences of spatial structures for the design of ecological field surveys (Legendre et al. 2002) and field experiments (Legendre et al. 2004), and to compare methods for partitioning the variation of a response data table with respect to spatial and environmental components (Legendre et al. 2005). The specific version of the program used for generation of autocorrelated space-time data in the present study is SimSSD4, which is available in ESA’s Electronic Data Archive: Ecological Archives M075-017-S1. It was used to generate files containing several thousands of fields of various sizes with known autocorrelation properties. All surfaces had 80 pixels along space and 80 along time. Two types of autocorrelated data were created using variogram ranges: (a) autocorrelated surfaces with ranges of 10 in both directions, and (b) autocorrelated surfaces with range of 20 in the spatial direction and 5 in the temporal direction.

500 replicate data surfaces of each type were generated for the univariate ( $p = 1$ ) and the multivariate ( $p = 5$ ) cases. Depending on the simulation scenario, the generated surfaces contained spatial or temporal structure, or both, as well as space-time interaction in some simulations, where the spatial structure changed along time. In order to study intermediate situations of simpler structures, the generated surfaces were modified as explained in the next paragraph. Four scenarios were studied:

- a) *Scenario 1*: Random data. Surfaces without any structure, either temporal or spatial.
- b) *Scenario 2*: Surfaces with no S-T interaction, no temporal structure, and a spatial autocorrelated structure common to all sampling times. This was obtained by choosing the first transect in a generated autocorrelated surface and duplicating it the stated number of times. Random noise was added to each value, drawn from a standard normal distribution with standard deviation equal to 0.1. An example surface is presented in Fig. A1a.
- c) *Scenario 3*: Surfaces with spatial and temporal structures but without space-time interaction. This was obtained as follows: for each autocorrelated surface, choose the first transect and duplicate it for all times. Then take the first time sequence and duplicate it for all spatial points. These two matrices were added up element-wise, and random noise was added to each value, drawn from a standard normal distribution with standard deviation equal to 0.1. An example surface is shown in Fig. A1b.
- d) *Scenario 4*: Surfaces were generated with spatial and temporal structures as well as space-time interaction. The autocorrelated data for these surfaces were generated by the program SimSSD4 without further modification. Example surfaces are shown in Fig. A2(a-b).

The surfaces from each scenario were sub-sampled to generate data sets of different sizes. The following combinations of space and sampling campaigns were studied:  $s = \{5, 10, 20, 40, 80\}$  and  $t = \{5, 10, 40\}$ . Smaller data sets were extracted from the original  $80 \times 80$  data sets: 2 data sets of  $80 \times 40$ , 4 sets of size  $40 \times 40$ , and so on, down to size  $5 \times 5$ . In this way, for each data set size, at least 1000 replicates were produced ( $500$  original data sets  $\times 2$ , for the largest subsets of size of  $80 \times 40$ ).

1000 simulation replicates were produced for all experimental conditions. Tests of significance were carried out by permutation for univariate and multivariate responses. Since the S and T factors are controlled and orthogonal, there is no point in using permutation of residuals of a null or full model (Anderson and Legendre 1999); so, permutation of the raw data was used in all cases. In a crossed ANOVA model, there is no exact permutation test for the interaction term. This is so because there are no possible permutations that would give an  $F$ -ratio different from the observed value when permutations are restricted to occur within levels of each of the main effects (Anderson and ter Braak 2003). If the interaction is not significant, exact permutation tests for the main factors are achieved by restricting the permutations to occur within the levels of the non-tested factor. When the interaction is significant, there is no exact test for the main factors either (Anderson and ter Braak 2003). Due to these considerations, all permutation tests for the S-T interaction in the present simulations were unrestricted and only approximate. The test for the main factors under Models 2, 3, 4, and 5 were performed by restricting permutations within levels of the non-tested factor. The rate of rejection of the null hypothesis after 1000 simulations, for significance level  $\alpha = 0.05$ , was reported in each case.

### ***Simulation results: test of the S-T interaction, Models 3, 4 and 5***

Scenario 1 (random normal structure) was used to measure the rates of type I error of the test of the S-T interaction under Models 3, 4 and 5. The results are presented in Appendix B, Tables B1a-c. In the absence of a spatial or temporal structure, all tests of the interaction had correct rates of type I error (i.e. the 95% confidence intervals contained the nominal 5% significance level used in the tests) for all values of  $s$  and  $t$  investigated.

In scenario 2, a spatial data transect had been replicated  $t$  times and random variation was added to the data. Hence, it contained a single common spatial structure and there was no interaction present. Analysis of the space-time interaction (STI) under Models 3 and 4 showed that the interaction was very rarely significant (Tables B2a-b).

Scenario 3 generated data containing common spatial and temporal structures, but again without interaction. The STI test corresponding to Models 3 and 4 never found a significant interaction under scenario 3 (Tables B3a-b). Why is it that the rate of rejection was not close to 5% in all these cases? The explanation is the following: consider scenario 2 and ANOVA Model 4 (Table B2b). Under this scenario, the explanatory S-PCNM variables could not “capture” all the common spatial structure. Hence the residual sum of squares ( $SS_{Res,4}$  in Fig. 1) contained two portions: the real residual sum-of-squares ( $SS_{Res1}$ ) corresponding to the random variation in the data, and a part of the non-modeled common spatial variation (light shading in Fig. 1). Since this sum-of-squares is used in the denominator of the  $F$ -statistic and it is larger than it should, this reduces the chances to reject the null hypothesis for the STI test, decreasing the rates below the nominal 5% level. The explanation is similar for the results in Tables B2a, B3a, and B3b.

When there is a spatial and a temporal structure (scenario 3), the rate of type I error for the STI test under Model 4 is too low. Similarly, when there is a spatial structure (scenario 2), the rate of type I error for the STI test under Model 3 is also too low. While the tests remain valid in these cases, this conservative behavior will reduce the power of the STI test under Models 3 and 4 to detect an interaction when it is present. Generally speaking, the use of a number of PCNM eigenfunctions to model space and time, which is smaller than  $(s-1)$  or  $(t-1)$ , leads to an overestimation of the residual sum-of-squares which is used in the denominator of the  $F$ -statistic, and that reduces the number of cases where the associated  $P$ -value is significant given the significance level  $\alpha$ . If that reasoning is correct, ANOVA Model 5 should resolve that problem and provide a test of STI that has a correct level of type I error, that is, a rejection rate approximately equal to the  $\alpha$  significance level. Here the results will differ slightly for the univariate and multivariate cases. Results of the simulations for the STI test using ANOVA Model 5 under scenarios 2 and 3 show that the rate of type I error was correct in all cases for univariate response data ( $p = 1$ , Tables B2c and B3c). For multivariate response data ( $p = 5$ ), the rate of rejection was lower than the  $\alpha$  significance level, but appeared to become asymptotically correct as  $s$  and  $t$  increased. That is, for scenario 2, the rates of rejection increased towards the  $\alpha$  significance level with increasing numbers of spatial points  $s$ . Similarly, for scenario 3, the rates

or rejection increase towards the  $\alpha$  significance level with increasing number of spatial points  $s$  or sampling campaigns  $t$ .

As it contains space-time interaction, scenario 4 is the most complex, and so is the interpretation of the power study results. Rates of rejection of the null hypothesis for the STI test under ANOVA Models 3, 4, and 5 are shown in Tables B4a-c and illustrated in Fig. 2 of the paper. To achieve a balanced design in the analyses reported in the “Numerical simulations” section of the paper and in Fig. 2, some results from Table B4 were left out of the analysis; they correspond to the combinations  $(S=10, T=5)$ ,  $(S=40, T=5)$ , and  $(S=40, T=10)$  which duplicated the combinations  $(S=5, T=10)$ ,  $(S=5, T=40)$ , and  $(S=10, T=40)$  for the sum  $S+T$ . The rate of significant S-T interaction was only 6-13% for the small  $5\times 5$  data sets, but it increased with higher numbers of spatial units and/or sampling times. Multivariate response data provided higher power to the STI test than univariate data. Power of the STI test under ANOVA Model 5 was similar to that under Model 4 for equal amounts of SA and TA, but power was higher in Model 5 for unequal amounts of autocorrelation in the data. The STI test in Model 4 was handicapped by its high lack-of-fit, which produced a large residual sum-of-squares. However, that effect on Model 4 was counter-balanced in the presence of equal amounts of autocorrelation by the larger number of degrees of freedom in the denominator of the  $F$ -statistic. When SA and TA were identical, the STI test was less powerful under Model 3 than under Models 4 and 5. Model 3 was, however, equally or more powerful than Model 4 for unequal amounts of autocorrelation; because the variability across spatial units was small, the lack-of-fit of the S fraction in Model 3 may have been reduced. So, Model 5 is preferable because its power was always equal or superior to those of Models 3 and 4.

In summary, our simulation results indicate that in the absence of replication, the space-time interaction can safely be tested using ANOVA Model 5: 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 on “Models to test the space-time interaction”.

### ***Simulation results: testing the effect of factors S and T using Models 2 to 5***

If the hypothesis of no interaction is not rejected, spatial and temporal structures should be analyzed using the classical test of S and T without replication, which is our ANOVA Model 2. When the analysis is carried out by regression, the main factors should be described in the analysis using Helmert contrasts or ordinary dummy variables. Using Model 2 for testing the main factors is, however, imperiled by the possibility of a type II error during the test of the interaction, leading to an undetected interaction. We will now compare the simulation results obtained under Models 2, 3, 4 and 5.

Under scenario 1, the tests corresponding to Models 2-5 showed correct rates of type I error (Tables B5a-d). On theoretical grounds, one can assume that, in the absence of interaction, the

tests for main factors (S or T) analyzed using Model 3, 4 or 5 should be less powerful than using Model 2. We verified that prediction using scenarios 2 and 3. Indeed, under scenario 2 (presence of a common spatial structure), the test for S was more often significant for Model 2 than for the other models (Tables B6a-d). It was followed by Models 5 and 4. Model 3 was the least powerful when testing factor S in scenario 2, since it shares with Model 4 the lack of fit for the space fraction but it has fewer degrees of freedom for the residuals. Permutation tests for a main factor restricted within the levels of the non-tested factor are theoretically exact in the absence of interaction. Thus, the test for factor T had correct type I error for all models (Tables B6a-d), even for Models 3 and 4. Note that this would not be the same for parametric tests, where the lack of fit for the S fraction in Models 3 and 4 would lower the rate of rejection of the null hypothesis. Under scenario 3 (spatial and temporal structures present, Tables B7a-d), Models 2 and 5 almost always found significant effects for factors S and T, while Models 3 and 4 were less powerful, especially for small data sets. Here Model 3 outperformed Model 4 probably due to a better fitting of the time fraction. From these results, we can conclude that when there is no space-time interaction in the data, Model 2 is the most powerful, followed by Model 5 and, finally, Models 3 and 4.

It may happen that an interaction existed and was not detected using the STI test (i.e. a type II error could have been committed). The presence of an S-T interaction would then hamper testing the main factors under Models 2-5. Not all four models are equally affected, however. Differences are shown in the simulation results for these three models under scenario 4 (Tables B8a-d). In this case the power relationships described in the previous paragraph were inverted: if an undetected interaction existed, then Model 4 appears to be the most powerful to obtain the significance of the main factors, followed by Model 3, Model 5 and, finally, Model 2.

#### LITERATURE CITED

- Anderson, M. J., and P. Legendre. 1999. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *Journal of Statistical Computation and Simulation* 62: 271-303.
- 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 (2): 85-113.
- Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs* 75: 435-450.
- Legendre, P., M. R. T. Dale, M.-J. Fortin, P. Casgrain, and J. Gurevitch. 2004. Effects of spatial structures on the results of field experiments. *Ecology* 85: 3202-3214.
- Legendre, P., M. R. T. Dale, M.-J. Fortin, J. Gurevitch, M. Hohn, and D. Myers. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography* 25: 601-615.

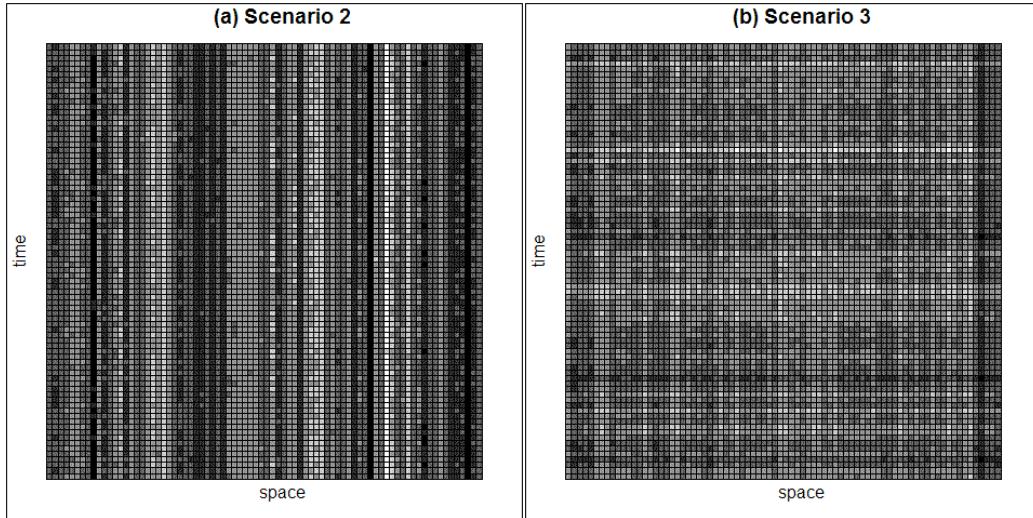


Fig. A1. Examples of surfaces generated for simulation scenarios 2 (no S-T interaction, no temporal structure, and a spatial autocorrelated structure common to all sampling time) and 3 (spatial and temporal structures present, no space-time interaction).

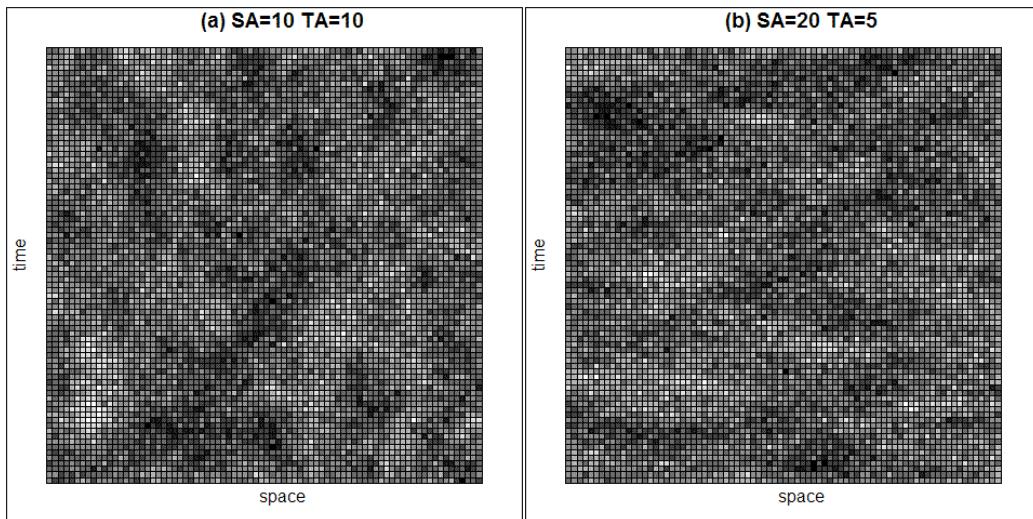


Fig. A2. Example of simulated autocorrelated surfaces with spatial and temporal structures as well as space-time interaction generated for simulation scenario 4. SA: range of the spatial autocorrelation (along abscissa), given in grid units; TA: range of the temporal autocorrelation (along ordinate). (a) SA = TA, so the light and dark patches are isotropic (no preferred direction). (b) SA > TA, so that the light and dark patches are elongated horizontally (along the spatial axis).

**APPENDIX B*****Ecological Archives E091-019-A2*****TABLES OF SIMULATION RESULTS**

Tables B1 to B9 present the simulation results as rates of rejection of the null hypothesis after 1000 simulations (values between 0 and 1).  $s$  is the number of spatial units,  $t$  the number of time units in the simulated surface.  $p$  is the number of response variables:  $p = 1$  for univariate data,  $p = 5$  for multivariate data. Columns **S** contain the results of the tests for spatial structures, columns **T** the results of the tests for temporal structures.  $SA$  indicates the range of the variogram used to generate autocorrelation along space,  $TA$  the range of the variogram used to generate autocorrelation along time. In the left-hand portions of tables B2 to B4 and B6 to B9, these two amounts are equal ( $SA10, TA10$ ) whereas they are unequal in the right-hand portions of these tables ( $SA20, TA5$ ).































where  $SS_{Res6b} = \sum_{j=1 \dots t} (SS_{Res(j)} + LOF_j)$  and  $LOF_j$  is the lack of fit for time  $j$ . The partitioning of the degrees of freedom is:

$$s \times t = t + ut + (st - ut - t) .$$

The implementation of Model 6b, using PCNM eigenfunctions for the  $u$  variables describing space, is described in Fig. C1. The arrangement of the PCNM variables representing spatial structure (S-PCNMs) is shown for a test of the hypothesis ( $H_0$ ) that there is no spatial structure at any one of the different times.

Model 6b can also be seen as a nested model derived from Model 3. The equivalence is:

$$\begin{aligned} \sum_{j=1 \dots t} SS(\mathbf{1}_{(j)}) &= SS(\mathbf{X}_{t-1}) \\ \sum_{j=1 \dots t} SS(\mathbf{X}_{u(j)}) &= SS(\text{space nested within time}) = SS(\mathbf{X}_u) + SS(\mathbf{X}_{Int3}) \\ \text{and } SS_{Res6b} &= \sum_{j=1 \dots t} (SS_{Res(j)} + LOF_j) = SS_{Res3} \end{aligned}$$

The advantage of a global test (Model 6b) is that it does not require any correction for multiple testing. We have to run simulations to see which test is the more powerful. Consider  $t = 5$ . The situations would be: create a spatial structure in 1, 2, ... 5 of the times with the other times having a random spatial structure. Compare the two tests.

Models 6a and 6b both require that the number of variables coding for space be smaller than  $(s - 1)$ . Otherwise, the degrees of freedom associated with the residuals would be zero.

### ***Simulation results***

Under scenario 1, the tests for factors space (S) and time (T) using Model 6b had correct rates of type I error (Appendix B, Table B9a). Under scenario 4 (presence of S and T effects and of an S-T interaction, Table B9b), the tests, which were expected to be significant, had increased rejection rates with increase in the number of points along the spatial structure and the number of sampling events through time. Power was also higher for multivariate data than it was for univariate data.

### LITERATURE CITED

Zar, J. H. 1999. Biostatistical analysis. Fourth edition. Prentice Hall, Upper Saddle River, New Jersey, USA.

		<b>X</b> Model of independent spatial structures at the various times					
		<b>Y</b> Response variables					
Sites	1 2 3 4 • • s	Time 1					
		S-PCNM	0	0	0	0	
Sites	1 2 3 4 • • s	Time 2	0	S-PCNM	0	0	0
		0	0	S-PCNM	0	0	
Sites	1 2 3 4 • • s	Time 3	0	0	S-PCNM	0	0
		0	0	0	0	0	
• • •		• •					
Sites	1 2 3 4 • • s	Time <i>t</i>	0	0	0	0	S-PCNM
		0	0	0	0	S-PCNM	

Fig. C1. Implementation of Model 6b, using PCNM eigenfunctions for the variables describing space in the analysis. The response data (matrix **Y**) are centered on their means computed for each time group separately. A large “0” in a cell of matrix **X** indicates that the cell is filled with zeros. Since the S-PCNMs are centered on zero by construction, the blocks of zeros contain values corresponding to the means of their columns.

**APPENDIX D*****Ecological Archives E091-019-A4*****INDICATOR SPECIES OF THE TRICHOPTERA EXAMPLE**

Trichoptera indicator species were identified by indicator species analysis for the 5 space-time groups shown in Fig. 3. Combining indicator values and numerical dominance in each of the 5 groups, one obtains the following scenario:

- Time periods 1-2: group 2 rules along most of the transect. This group is dominated by the shredder *Ceraclea diluta*, the filterer *Hydropsyche sparna*, and the algae piercer *Hydroptila delineata*. The shredder *Lepidostoma pictile* characterizes this group thanks to a localized emergence during the first 10 days in a group of sites near the middle of the transect.
- Time periods 3-5: group 1 occupies most of the transect. Numerically, this group is dominated by the filterer *Cheumatopsyche minuscula*, with high numbers of the algae piercer *Hydroptila valhalla*. Together they form almost 60% of the total number of individuals in this group. The former species emerges during the whole period, while the latter is more concentrated around time period 4. Other species have been found to be significant indicators of this group due to high values of either specificity or fidelity.
- The 5 remaining time periods are characterized by a parcelling out of the transect, which becomes more finely partitioned, mainly among groups 3, 4, and 5. Group 3 is characterized by the algae piercer *Oxyethira grisea*. Group 4 is dominated by the filterer *Cheumatopsyche pettiti* and the algae piercer *Oxyethira grisea*. Group 5 is dominated by the grazer *Neotrichia okopa*, emerging mostly during periods 6 and 7. Another numerically important species is the filterer *Cheumatopsyche minuscula*, dominating group 1.

**SUPPLEMENT*****Ecological Archives E091-019-S1*****STI PACKAGE FOR THE ANALYSIS OF SPACE, TIME, AND INTERACTION  
IN SPACE-TIME STUDIES**

The STI package, an R-language library for the analysis of the main factors space and time and the interaction in space-time studies using permutation tests, is available on the ESA Web page <http://esapubs.org/archive/ecol/E091/019/> and on the Software page of Miquel De Cáceres <http://sites.google.com/site/miqueldecaceres/software> (source code and compiled libraries for Windows and Mac OS X). The main functions are called STIMODELS and QUICKSTI. Function STIMODELS offers ANOVA models 2 to 6 described in Fig. 1 of the paper.