

Statistical tool for soil biology. XI. Autocorrelogram and Mantel test

Jean-Pierre Rossi

Laboratoire d'Écologie des Sols Tropicaux, ORSTOM/Université Paris 6,
32, av. Varagnat, 93143 Bondy Cedex, France.
E-mail: rossijp@bondy.orstom.fr.

Received October 23, 1996; accepted March 24, 1997.

Abstract

Autocorrelation analysis by Moran's I and the Geary's c coefficients is described and illustrated by the analysis of the spatial pattern of the tropical earthworm *Chuniodrilus zielae* (Eudrilidae). Simple and partial Mantel tests are presented and illustrated through the analysis various data sets. The interest of these methods for soil ecology is discussed.

Keywords: Autocorrelation, correlogram, Mantel test, geostatistics, spatial distribution, earthworm, plant-parasitic nematode.

L'outil statistique en biologie du sol. XI. Autocorrélogramme et test de Mantel.

Résumé

L'analyse de l'autocorrélation par les indices I de Moran et c de Geary est décrite et illustrée à travers l'analyse de la distribution spatiale du ver de terre tropical *Chuniodrilus zielae* (Eudrilidae). Les tests simple et partiel de Mantel sont présentés et illustrés à travers l'analyse de jeux de données divers. L'intérêt de ces méthodes en écologie du sol est discuté.

Mots-clés : Autocorrélation, corrélogramme, test de Mantel, géostatistiques, distribution spatiale, vers de terre, nématodes phytoparasites.

INTRODUCTION

Spatial heterogeneity is an inherent feature of soil faunal communities with significant functional implications. In order to quantify that heterogeneity, many aggregation indices have been proposed and applied to various soil fauna taxa (Cancela Da Fonseca, 1966; Cancela Da Fonseca & Stamou, 1982; Campbell & Noe, 1985). Most of them are directly derived from the estimation of the population variance and mean from samples and are often sensitive to the average density of the population (Elliot, 1971). Thus, they may give varying index values for samples with different density while the true aggregation is the same. Among the available indices, the index of Taylor's Power Law (Taylor, 1961) is commonly used in soil biology (Boag & Topham, 1984; Ferris *et al.*, 1990; Boag *et al.*, 1994). The method is based on the empirical relationship between the mean and

the variance that appear to be related by a simple power law. The obvious interest of that method is that samples from various sites can be included in the analysis. It leads to a general and representative index that is considered as species-specific by Taylor (Taylor *et al.*, 1988).

Whatever the index used, it is obvious that one cannot restrict the study of the spatial distribution of a population to the computation of any of the available dispersion indices. Indeed, dispersion indices are limited to the description of the kind of distribution encountered and to a certain extent to the quantification of the degree of clustering (Nicot *et al.*, 1984). They do not take into account the spatial position of the sampling points, neither do they furnish information on the true pattern of the variable (Liebhold *et al.*, 1993; Rossi *et al.*, 1996). In a previous paper of this series (Rossi *et al.*, 1995) we introduced the use and



interest of the geostatistical tool in soil Ecology. We showed how geostatistics made it possible to look for the presence of spatial autocorrelation in the data and how variogram analysis (*viz.* structural analysis or variography) could quantify it. In addition, we introduced the kriging procedure, an optimal mapping method.

Ecologists often study complex ecological systems and thus face the problem of assessing species-environment relationships. Since most of the ecological variables are spatially structured at various scales, it is necessary to take into account the presence of spatial autocorrelation in the data sets. In that case, maps are particularly relevant to the problem. However, assessing relationships between spatially autocorrelated variables brings statistical problems as autocorrelation impairs the standard statistical tests *e.g.* correlation coefficient, ANOVA (Legendre *et al.*, 1990; Fortin & Gurevitch, 1993; Legendre, 1993). Moreover, imagine a significant common pattern is found shall we consider the variables to be correlated? In the case of spatially structured variables if we find a correlation, does it mean there is a true correlation between them or are they simply following the same gradient (or any other kind of pattern) due to unknown common driving factor(s)?

The aims of this paper are twofold. First methods alternative to the semi-variogram analysis are presented. They allow overall statistical testing for the presence of spatial autocorrelation. Second, matrix methods for assessing and testing for the presence of a common spatial pattern between two autocorrelated variables are showed. The different methods introduced in this paper have been used in various fields of the life sciences but are still sparsely applied in soil biology and ecology. This paper aims to introduce these approaches and provide relevant literature while emphasizing the potential interest in soil ecology with examples taken from various unpublished studies.

SPATIAL STRUCTURE AND AUTOCORRELATION TESTS

In a previous paper (Rossi *et al.*, 1995), we showed the use of the semi-variance and semi-variograms to identify spatial autocorrelation. In the variogram analysis however, semi-variance values and variograms are not tested for statistical significance. The presence of a consistent pattern is indicated by the shape of the variogram as well as the ratio of total heterogeneity that can be ascribed to the spatial structure. Correlogram analysis allows tests for the presence of autocorrelation in data.

The method is based on the use of spatial autocorrelation coefficients, Moran's I (Moran, 1950) or Geary's c (Geary, 1954) coefficients, that are used to analyse quantitative variables. Note that Sokal & Oden (1978) proposed a special form of

spatial autocorrelation coefficient for qualitative data. A variable is said to be autocorrelated - or regionalized - when the measure made at one sampling site brings information on the values recorded at a point located a given distance apart. The autocorrelation coefficient measures the degree of autocorrelation *i.e.* likeness between couples of values recorded at sampling points separated by a given distance. The principle is the same as for the semi-variance analysis (Rossi *et al.*, 1995). Moran's I and Geary's c are respectively:

$$I(d) = \left[n \sum \sum w_{ij} (y_i - \bar{y})(y_j - \bar{y}) \right] / \left[W \sum (y_i - \bar{y})^2 \right] \quad \text{for } i \neq j$$

$$c(d) = \left[(n-1) \sum \sum w_{ij} (y_i - y_j)^2 \right] / \left[2W \sum (y_i - \bar{y})^2 \right] \quad \text{for } i \neq j$$

Data are grouped by distance classes (d) which are a function of the separating distance between sampling points, y_i and y_j are the values of the variables with i and j varying from 1 to n the number of data points. \bar{y} the mean of the y 's, w_{ij} is a weighting factor taking the value of 1 if the points belong to the same distance class and zero otherwise. W is the sum of the w 's *i.e.* the number of data pairs involved in the estimation of the coefficient for the distance class d . Positive values of Moran's I and value smaller than 1 for Geary's c coefficients correspond to positive autocorrelation. Notice that autocorrelation analysis requires at least 30 sampling localities to produce significant results (Legendre & Fortin, 1989).

The plot of the autocorrelation coefficient against the distance classes is called the correlogram. Each of the coefficient values can be tested for statistical significance. Formulas can be found in Sokal & Oden (1978), Cliff & Ord (1981) and Legendre & Legendre (1984). The test is based on the null hypothesis H_0 "There is no spatial autocorrelation" tested against the alternative hypothesis H_1 "There is a spatial autocorrelation". Under H_0 , the value of Moran's I coefficient is $E(I) = -(n-1)^{-1} \approx 0$ with $E(I)$ the expectation of I and the n number of data points. Geary's c coefficient equals $E(c) = 1$. Under H_1 , the value of Moran's I coefficient is significantly different from 0 and Geary's c coefficient is significantly different from 1.

However, the correlogram must be checked for global significance. A correction is thus to be used as we perform simultaneously k statistical tests (one for each coefficient). The test is made according to the Bonferroni method of correction. It consists in checking if at least one of the coefficients is significant

at the statistical level $\alpha' = \alpha/k$ with $\alpha = 5\%$ and k the number of distance classes (Oden, 1984).

The shape of the correlograms gives information on the type of spatial structure encountered. Some characteristic shapes are associated with specific patterns, for instance the alternation of significant positive and negative autocorrelation is typical of a patchy distribution. Sokal & Oden (1978) and Legendre & Fortin (1989) gave a series of typical correlograms and the corresponding patterns. However, the shape of the correlogram is not always specific of a given distribution type, e.g. data presenting a sharp step and a gradient, lead to quite the same correlogram shape. Therefore maps are necessary to fully describe the spatial patterns. Among the various mapping methods, the kriging procedure is particularly useful since it is optimal and unbiased (Burgess & Webster, 1980a, b; Isaaks & Srivastava, 1989; Webster & Oliver, 1990).

Example 1: Spatial autocorrelation of the earthworm *Chuniodrilus ziela* (Eudrilidae).

In order to illustrate the use of correlograms we shall apply the method to the earthworm *Chuniodrilus ziela* (Eudrilidae) data we presented in Rossi *et al.* (1995). The data set was collected in an African grass savanna in July 1994. 100 sampling points were located at regular intervals of 5 m on a 10×10 points grid. At each sample location a 25×25×10 cm soil monolith was taken and earthworms were handsorted. Data were \log_e transformed before analysis to reduce the asymmetry of the frequency distribution. Here the normalisation of data is obtained through the Box-Cox transformation (Sokal & Rohlf, 1995) which is: $y = (x^\delta - 1)/\delta$. The δ parameter was estimated as $\delta = 0.32606$ using the program VerNorm 3.0 from the "R package" developed by Legendre & Vaudor (1991). After transformation, the frequency distribution met normality as confirmed by a Kolmogorov-Smirnov test of normality (done, again, using the program VerNorm 3.0 from the "R package"). Figure 1 shows

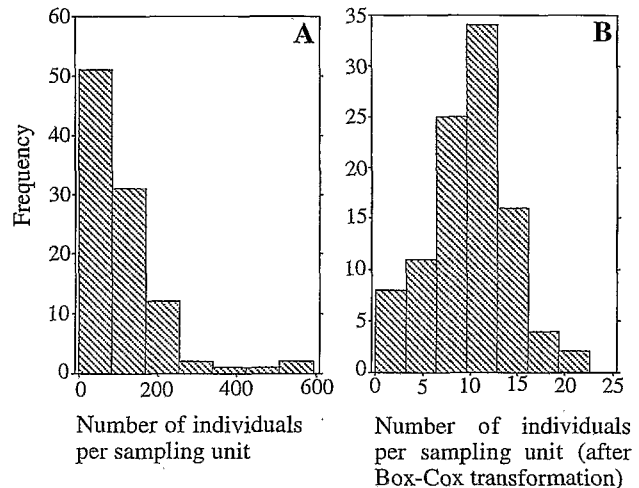


Figure 1. – Frequency distribution of the earthworm *Chuniodrilus ziela*: (A) Raw data and (B) Data transformed according to the Box-Cox transformation.

the frequency distribution of the raw and transformed data.

Data were allocated to 12 distance classes with 5.30 m width. Figure 2 shows the distribution of the point pairs among the distance classes. The number of data couples involved in the computation is small for the last 3 distance classes (table I), and autocorrelation coefficients cannot be interpreted because they are computed upon too few pairs of point. That is why dividing distances into classes with equal frequencies is sometimes preferred upon forming equal distance classes.

Autocorrelation coefficients of Moran and Geary were estimated for each distance class and the probability for obtaining such values under the null hypothesis were computed using the program Autocorrelation from the "R package" written by Legendre & Vaudor (1991). Table 1 gives the coefficient values and associated probability. The

Table 1. – Moran and Geary autocorrelation coefficient values for the earthworm *Chuniodrilus ziela* density. The width of the distance classes is 5.3 m, $p(H_0)$ indicates the probability to obtain the coefficient value under the null hypothesis.

Distance classes	Lower limit (m)	Upper limit (m)	I (Moran)	$p(H_0)$	c (Geary)	$p(H_0)$
1	0	5.3	0.4870	0 [†]	0.4665	0 [†]
2	5.3	10.6	0.2579	0 [†]	0.6493	0 [†]
3	10.6	15.9	-0.0010	0.613	0.8790	0.016
4	15.9	21.2	-0.1412	0 [†]	1.0499	0.161
5	21.2	26.5	-0.1352	0 [†]	1.0665	0.069
6	26.5	31.8	-0.0397	0.193	0.9749	0.263
7	31.8	37.1	0.0898	0.003 [†]	0.9330	0.079
8	37.1	42.4	-0.0316	0.313	1.1727	0.014
9	42.4	47.7	0.0253	0.229	1.2257	0.028
10	47.7	53	-0.1910	0.035	1.5046	0.003 [†]
11	53	58.3	-0.5917	0 [†]	2.1275	0 [†]
12	58.3	63.6	-1.1141	0 [†]	2.7608	0.001 [†]

[†] Significant probabilities at the Bonferroni-corrected probability level of $0.05/12 = 0.00416$.

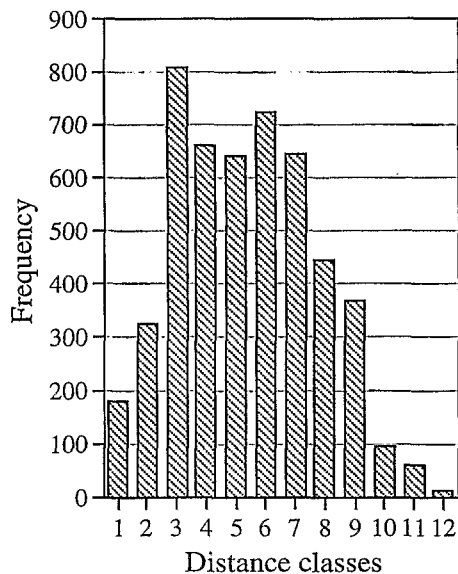


Figure 2. – Number of pairs of points in each distance class.

resulting correlograms are showed in figure 3 where open circles represent non-significant individual values of the coefficients at the statistical level $\alpha = 0.05$. Before examining the correlogram it is necessary to ensure that it is globally significant at the corrected probability level $\alpha' = 0.05/12 = 0.00416$. This condition is actually verified for several autocorrelation values in our example (table I). Correlograms indicate the presence of a positive autocorrelation at short distance classes, which reveals the presence of a contagious distribution. In the Moran correlogram, the autocorrelation coefficient is positive at distance class 1 and decreases up to distance class 4. From distance class 5 to 7 it increases again. Significant positive autocorrelation values for short distance classes indicates a contagious distribution. The similarity between samples decreases up to distance class 4 and then increases up to distance class 7. The distance class 7 ranges between 37.1 and 42.4 m which corresponds approximately to the interpatch distance in the map of the variable (Rossi *et al.* (1995), fig. 6a). Points separated by a distance falling within the bounds of distance class 7 are similar because the corresponding samples are taken in the patches. The analysis of Geary's c correlogram leads to the same conclusions. For small distance classes the coefficient is lower than 1 indicating a positive autocorrelation. From distance class 7 it becomes significantly greater than 1 which corresponds to a negative autocorrelation.

RELATIONSHIPS BETWEEN AUTOCORRELATED DATA

The assessment of the relationship between variables is generally made by computing their correlation

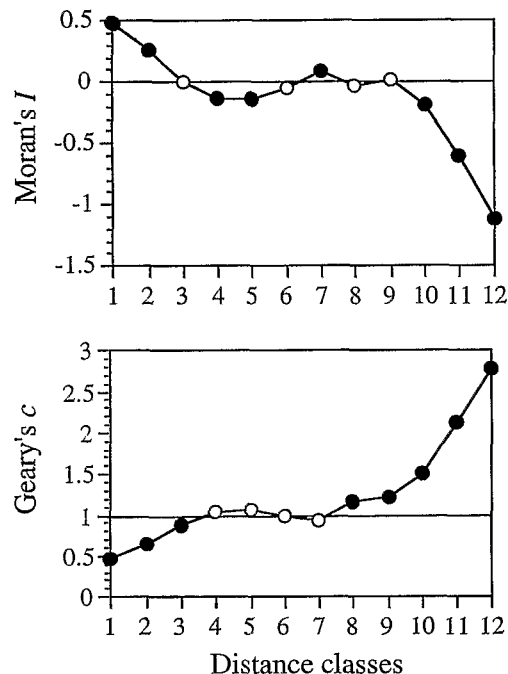


Figure 3. – Moran's I and Geary's c spatial correlograms for the earthworm *Chuniodrillus zizelae* density.

coefficient. This approach may be impaired if the variables are non-linearly related, and variables must be transformed or non-linear regression used. However, these approaches are based on the assumption that observations are independent. In that case, each of the observations brings one degree of freedom. In turn, the sum of the degrees of freedom is used to compute the statistic of the correlation coefficient and thus check for its significance.

If the variables under study are positively spatially autocorrelated and if the sampling scale matches with the scale at which the spatial structure is expressed, the assumption of independence of observations is clearly violated. Since the information brought by a value is partly brought by another value as a sequel of autocorrelation, the degrees of freedom are overestimated and testing correlation coefficient becomes impossible. The overestimation of degrees of freedom leads to conclude that a correlation is significant when it is not, but conversely, if a correlation coefficient is declared non-significant the result is valid. This problem has been solved by Clifford *et al.* (1989) who proposed to estimate the autocorrelation of the variables and using these estimates obtain a reduced number of degrees of freedom which can then be used to test the significance of the correlation coefficient.

In geostatistics, the relationship between two autocorrelated variables is assessed by the cross-variogram (Isaaks & Srivastava, 1989; Rossi *et al.*, 1995). A structure function (the cross-variogram) describes the variation of the cross semi-variance in

function of the sample spacing. In soil biology, this method has been applied for example to study the relationships between the density of various plant-parasitic nematode species (Delaville *et al.*, 1996; Rossi *et al.*, 1996). We will now examine another method for assessing the relationship between two variables and checking its significance: the Mantel statistic (Mantel, 1967).

The Mantel test allows to look for correlation between two proximity or distance matrices (say, **A** and **B**) by computing the cross-product of the corresponding values in **A** and **B**. The matrices describe the relationships among the n sampling sites. **A** and **B** can be formed by using one of the many distance or dissimilarity indices available in the literature (Legendre & Legendre, 1984; Gower & Legendre, 1986). The method can be applied to different cases, the multivariate and the univariate approaches. In the multivariate case matrices are formed using many variables describing the sampling stations. One can then look for relationships between a distance matrix representing the environmental variables and a second matrix describing the community composition either in terms of density, biomass or presence/absence. In the univariate case, each matrix is formed from one variable with the Euclidean distance coefficient: the distance between two sampling stations is computed as the unsigned difference among values of the variable. One of the two matrices can be directly built from the geographic distances among sampling stations. In that case, comparing this matrix and another one derived from a given variable constitutes a way to look for a spatial trend in the data.

The Mantel statistic tests the null hypothesis H_0 "Distances among points in the matrix **A** are not linearly related to the corresponding distances in the matrix **B**" against the alternative hypothesis H_1 "Distances among points in matrix **B** are linearly correlated to the corresponding distances in the matrix **B**".

The Mantel statistic is:

$$Z = \sum_i \sum_j x_{ij} y_{ij} \text{ with } i \neq j \text{ and } i \text{ and } j \text{ the row and column indices.}$$

The Mantel statistic can be normalised to range between -1 and $+1$.

$$r = [1/(n-1)] \sum_i \sum_j [(x_{ij} - \bar{x})/s_x] [(y_{ij} - \bar{y})/s_y]$$

with $i \neq j$ and i and j the row and column indices and n the number of distances in one of the matrix without accounting for the diagonal.

The statistical significance of the Mantel coefficient can be tested in two ways i) by computing the expected value and variance under the null hypothesis and performing a z -test (Legendre & Fortin, 1989) provided the size of the matrix tested is large *i.e.* $n > 40$, and ii) application of a permutation test. The latter consists in simulating the realisations of

the null hypothesis by repeated permutations of the lines and columns in one of the matrix **A** and **B** and recomputing the Mantel statistic. The result is a sampling distribution of the Mantel statistic under the null hypothesis. If there is no relationship between the matrices, the observed r value is near the centre the sampling distribution, while if a relationship is present, one would expect the observed value to be more extreme than most of the values obtained by permutation (Legendre & Fortin, 1989).

The Mantel test is primarily used to search for linear trend in the data *e.g.* linear gradient corresponding to some kind of underlying diffusive processes. However, linear trends are not the most frequent spatial pattern encountered in soil ecology. Organisms frequently display complex spatial distributions in patches of different sizes (Wallace & Hawkins, 1994; Robertson & Freckman, 1995; Rossi *et al.*, 1996, 1997). Thus if the trend in the data is linear, the geographic distance is to be used. If some other relationship prevails, one may use some other function of the geographic distance D such as $1/D$ or $1/D^2$. It is recommended to use large sample size ($n > 20$) to detect significant spatial pattern (Fortin & Gurevitch, 1993).

Example 2: Application of the Mantel test to the *Chuniodrilus zielae* data.

The Mantel test was applied to look for the relationship between two dissimilarity matrices corresponding to the variables *Chuniodrilus zielae* density and the geographic distance between the sampling points. The data set is the same as in *example 1*. The matrix **A** was formed by taking the unsigned difference among values of *C. zielae* density for all possible pairs of station while the matrix **B** was formed by taking the geographic distances among the sampling localities. The distance matrix computations were carried out using the program Simil 3.01 from the "R package" (Legendre & Vaudor, 1991).

The Mantel statistic and permutation test were carried out with the program Mantel 3.0 from the "R package" (Legendre & Vaudor, 1991). The standardised Mantel r was $r=0.12230$ with a probability to reach such a value under the null hypothesis of $p=0.001$ (obtained with 1000 permutations). Since the Mantel statistic is significant we reject the null hypothesis that there is no relationships between the matrices **A** and **B**, in other words, the earthworm density displays a spatial pattern. It is important to notice that the Mantel statistic r is a correlation between two distance matrices and is not equivalent to the correlation between the variables used to form these matrices.

The example above shows that the Mantel test is able to detect the spatial pattern of *C. zielae* although it is not a simple gradient as there are two patches (see *fig. 6a* in Rossi *et al.* (1995)).

Example 3: Assessing the relationships between a plant-parasitic nematode and soil clay content.

The following example is taken from a study carried out in the south-east of Martinique (Lesser Antilles: 14° 3' N and 62° 34' W) on a vertisol developed on volcanic ashes (Rossi and Quénéhervé, unpublished). In a 10-yr. old pasture regularly planted with the tropical grass *Digitaria decumbens* both plant-parasitic nematodes and some soil physico-chemical parameters were investigated by means of 60 sampling points randomly distributed within a 60 × 25 m plot. Soil samples with adhering roots for nematode analysis were removed from the 0-10 cm soil layer. The nematodes were extracted from the soil by the elutriation-sieving technique (Seinhorst, 1962) and from the shredded roots in a mist chamber (Seinhorst, 1950). Soil samples were removed from 0-10 cm soil layer and soil texture (clay, silt and sand contents) were determined by laser granulometry (Mastersizer E, Malvern).

The Mantel test was used to assess the relationships between density of the plant-parasitic nematode *Helicotylenchus retusus* (Siddiqui & Brown, 1964) and soil clay content (%). The clay content data met normality whereas the Box-Cox procedure was used to normalise the nematode data. Both nematode density and clay content displayed significant spatial pattern as shown by Moran's *I* correlogram (significant at the Bonferroni corrected statistical level). A distance matrix was formed for each variable (matrix CLAY for clay content and matrix NEM for *H. retusus* density) as explained above. A third matrix was formed by taking the geographic distance among sampling points (matrix SPACE). The Mantel test was used to test for the correlation between the three possible pairs of matrices (table 2). Since three tests are done simultaneously, it is necessary to use the Bonferroni-corrected probability level of $0.05/3 = 0.01667$ for an overall significance level of 0.05 (table 2). Table 2 shows that all the simple standardised Mantel tests are significant. Thus, the spatial patterns revealed by the correlograms analysis are confirmed by the Mantel tests. The Mantel test between CLAY and NEM indicates a significant correlation between the matrices. From this result, shall we conclude that this common pattern is due to a relationship between the variables or is it just a spurious correlation due to the fact that both variables are independently driven by a common cause? This topic is addressed in the next section.

Table 2. — Simple standardised Mantel statistics and associated probabilities. Tests of significance are one-tailed (data from Rossi and Quénéhervé, unpublished).

	Mantel's <i>r</i>	Probability
HRET.CLAY	0.13078	0.00599†
HRET.SPACE	0.16720	0.00200†
CLAY.SPACE	0.67063	0.00100†

† Simple Mantel test significant at the Bonferroni-corrected probability level of $(0.05/3 = 0.01667)$ for an overall significance level of 0.05 over three simultaneous tests.

PARTIALLING OUT SPATIAL EFFECT

Once the existence of a relationship between two variables has been demonstrated, one can wonder if it is a true correlation or if it is only a spurious correlation due to common spatial (or temporal) pattern. In other words, two variables may appear to be related while they are only independently driven by a third common cause. In ecological studies, space (sampling position) is likely to cause such spurious correlation and in that case determining whether the correlation is true or spurious requires partialling out the spatial component of ecological variation.

The partial Mantel test constitutes a way to assess the relationship between two dissimilarity matrices while controlling for the effect of a third one (Smouse *et al.*, 1986). The third matrix is formed with the geographic distances between sampling points if the "space" effect is to be partialled out. A computation method has been developed by Smouse *et al.* (1986) to test the partial Mantel statistic between two matrices **A** and **B** while controlling for the matrix **C**. The Mantel test is applied to the matrices **A'** and **B'** that respectively contain the residuals of the linear regression of the values of **A** and **B** on the values of **C**. The statistical test of significance is the same as explained above: either Mantel's normal approximation or the permutational test by permuting either **A'** or **B'**. The Mantel test between **A'** and **B'** is the partial Mantel test between **A** and **B** whilst controlling for the effect of **C**.

If a correlation is spurious because of the effect of unknown factors causing a common spatial pattern, the partial Mantel test is expected not to be significant while controlling for the effect of space. Conversely, in the presence of a true correlation, partialling out the effect of space still leads to a significant test (Legendre & Troussellier, 1988).

Example 4: Common pattern among a plant-parasitic nematodes and soil clay content.

The data are those presented in *example 3*. The partial Mantel test was used to assess correlation between the matrices NEM and CLAY while controlling for the effect of the matrix SPACE. Computations were done with the software Mantel 3.0 from the "R package" (Legendre & Vaudor, 1991). The statistical significance of the Mantel statistic was checked using the permutation test with 1000 permutations. The test is denoted (CLAY.NEM).SPACE. The *r* value was $r = 0.02550$ with an associated probability $p = 0.19780$. If the correlation were true, we would expect the test to be significant, a condition that is not met. So we can conclude that *H. retusus* and clay content display similar spatial pattern without being truly correlated. Another factor, not explicitly mentioned, independently drives the variables.

This example shows that caution is needed when analysing autocorrelated data, first because specific

tools are required and second because a common spatial pattern may lead to the observation of spurious correlation (Legendre & Troussellier, 1988).

The simple Mantel and partial Mantel tests permit segregation of different causal models (Legendre & Troussellier, 1988; Legendre & Fortin, 1989). In a causal model, the ecologist includes some hypotheses about the factors determining the process at hand. The second step consist in checking whether the field or laboratory data actually support predictions of the model. The investigations must not be restricted to the initial model but have to be extended to all the alternative possible models. Each model leads to a group of causal predictions that must be checked. Finally, a model is accepted if and only if all the predictions are supported by the data. In some cases, the ecologist can exclude some causal models that do not make sense, for instance those where space would be a dependent variable. A detailed study of the use of Mantel and partial Mantel tests in causal modelling can be found in Legendre & Troussellier (1988).

Example 5: Modelling the relationship between a plant-parasitic nematode and soil clay content.

If we come back to the nematode data (*examples 3 and 4*) and exclude all the models where space is a dependent variable as well as those where clay is dependent on *H. retusus*, only four possible models remain.

The first model states that the spatial structure in the nematode population is partly caused by the clay content gradient and partly by other factors not explicitly mentioned in the model and summarised under the term "space": $CLAY \rightarrow NEM \leftarrow SPACE$. If this model were supported by the data we would expect the simple Mantel test $SPACE.CLAY$ not to be significantly different from 0 which condition is not met in *table 2*.

The second model states that the spatial structure of the nematode population is caused by the clay pattern and also by the unknown factors (Space) whilst Space also causes the clay structure (*fig. 4*). According to that model we would expect the partial Mantel statistic $(CLAY.NEM).SPACE$ to be significantly different from 0 which condition is not met (see *example 4*).

The third model claims that the nematode structure is caused by clay spatial pattern which is in turn caused by unidentified factors (Space): $SPACE \rightarrow CLAY \rightarrow NEM$. If that model were true, we would expect the partial Mantel statistic $(SPACE.NEM).CLAY$ not to be significantly different

from 0, a condition that is not met as $r=0.1081$ with associated probability $p=0.005$.

Finally the fourth model states that both clay and nematode spatial pattern are independently caused by unknown factors (Space): $CLAY \leftarrow SPACE \rightarrow NEM$. If the model were supported by the data we would expect the partial Mantel test $(CLAY.NEM).SPACE$ not to be significantly different from 0 which condition was verified above (*example 4*). Before accepting the model, we must verify all the predictions that can be derived from it:

$SPACE.CLAY \neq 0$ $SPACE.NEM \neq 0$ $(SPACE.CLAY).NEM \neq 0$ $(SPACE.NEM).CLAY \neq 0$	$(CLAY.NEM).SPACE = 0$ $(SPACE.NEM).CLAY < SPACE.NEM$ $(SPACE.CLAY).NEM < SPACE.CLAY$ $(SPACE.CLAY) \times (SPACE.NEM) \approx CLAY.NEM$
---	---

Since the partial Mantel test $(SPACE.CLAY).NEM$ gives $r=0.66373$ with the significant value $p=0.001$ all the conditions required by the model 4 are verified and it is not rejected.

Notice that when using the partial Mantel statistics, the statistical significance level to be used is $0.005/3=0.01667$ as three partial Mantel tests are done simultaneously (Bonferroni correction).

CONCLUDING REMARKS

In soil ecology, most data from field work likely display spatial structure. It is increasingly recognised that soil living organisms and environmental variables exhibit various spatial patterns at different scales (Jackson & Caldwell, 1993; Robertson & Gross, 1994). Assessing the relationships between species distribution and environmental parameters is thus a ticklish problem. Correlogram analysis is a way to test for the presence of spatial autocorrelation (Cliff & Ord, 1981; Sokal & Oden, 1978). Testing for spatial autocorrelation may have two objectives. One may wish to ensure that there is no spatial dependence before applying standard statistics or, if autocorrelation is present, remove it before performing further data treatment. The other chief goal is to study the spatial structure in which case no removal is performed and specific statistical tools are used.

The simple and partial Mantel tests constitute very interesting alternative methods to classical correlation analysis. They allow to test for the presence of a spatial trend in the data or to compare two distance matrices formed by variables possibly related. The partial Mantel test is particularly useful in the framework of causal modelling as partialling out the spatial component of ecological variation allows to segregate between true and false correlation.

In this paper only the univariate approach was addressed but as distance matrices can be formed from an array of variables it is possible to use the

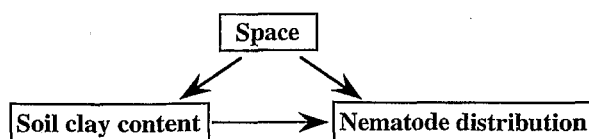


Figure 4. – Diagram of the interrelationships between soil clay content, *Helicotylenchus retusus* density and space.

Mantel test to assess the correlation between matrices corresponding to a complete set of environmental variables and a second formed with the abundance of the species making a community. In the data arrays, different kind of variables (quantitative, semi-quantitative, qualitative) can be mixed up and an adapted coefficient of association be used (*see* Gower & Legendre (1986) and Legendre & Legendre (1984) for a review of the coefficients of association and Legendre & Fortin (1989) for an example).

Although the Mantel test is devoted to the study of linear gradients, the examples presented in this paper show that the test is efficient in assessing patchy distribution at least if the patches are smooth and relatively large compared to the overall studied surface. The Mantel test may also be used to check for the goodness-of-fit of data to a model. In this case, the object is to compare a matrix formed by data to that formed from predictions of a given model.

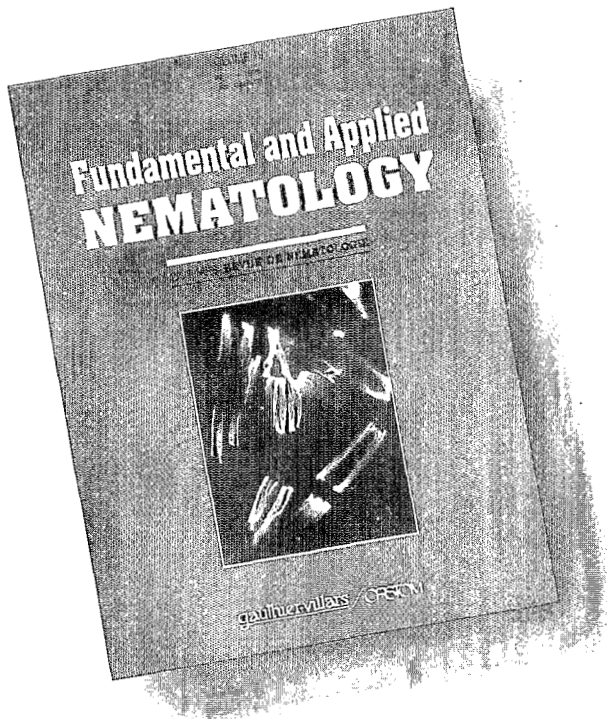
Acknowledgements

I wish to thank Dr J. P. Cancela Da Fonseca who initiated the "Statistical Tool for Soil Biology" series and allowed me to publish under this heading. I am greatly indebted to Prof. P. Legendre for providing me with his computer program the "R package" and giving me statistical guidance. I also wish to thank J. E. Tondoh and R. Zouzou Bi Danko for field assistance (it's been no bed of roses!). Special thanks to Prof. P. Lavelle, L. Rousseaux and P. Quénéhervé for stimulating discussions and two anonymous referees for perceptive and detailed comments on an earlier draft.

REFERENCES

- Boag B., Legg R. K., Neilson R., Palmer L. F. & Hackett C. A. (1994). – The use of Taylor's Power Law to describe the aggregated distribution of earthworms in permanent pasture and arable soil in Scotland. *Pedobiologia*, **38**, 303-306.
- Boag B. & Topham P. B. (1984). – Aggregation of plant parasitic nematodes and Taylor's power law. *Nematologica*, **30**, 348-357.
- Burgess T. M. & Webster R. (1980a). – Optimal interpolation and isarithmic mapping of soil properties. 1 The semi-variogram and punctual kriging. *Journal of Soil Science*, **31**, 315-331.
- Burgess T. M. & Webster R. (1980b). – Optimal interpolation and isarithmic mapping of soil properties. 2 Block kriging. *Journal of Soil Science*, **31**, 333-341.
- Campbell C. L. & Noe J. P. (1985). – The spatial analysis of soilborne pathogens and root diseases. *Annual review of Phytopathology*, **23**, 129-148.
- Cancela Da Fonseca J. P. (1966). – L'outil statistique en biologie du sol III. Indices d'intérêt écologique. *Revue d'écologie et de biologie du sol*, **3**, 381-407.
- Cancela Da Fonseca J. P. & Stamou G. P. (1982). – L'outil statistique en biologie du sol VII. L'indice d'agrégation de Strauss et son application aux populations édaphiques. *Revue d'écologie et de biologie du sol*, **19**, 465-484.
- Cliff A. D. & Ord J. K. (1981). – *Spatial processes: models and applications*. London: Pion limited.
- Clifford P., Richardson S. & Hémon D. (1989). – Assessing the significance of the correlation between two spatial processes. *Biometrics*, **45**, 123-134.
- Delaville L., Rossi J. P. & Quénéhervé P. (1996). – Row crop and soil factors incidence on the micro-spatial patterns of plant parasitic nematodes on sugarcane in martinique. *Fundamental and Applied Nematology*, **19**, 321-328.
- Elliot J. M. (1971). – Some methods for statistical analysis of samples of benthic invertebrates. Ambleside: Freshwater biological Association.
- Ferris H., Mullens T. A. & Foord K. E. (1990). – Stability and characteristics of spatial description parameters for nematode populations. *Journal of Nematology*, **22**, 427-439.
- Fortin M. J. & Gurevitch J. (1993). – Mantel tests: spatial structure in field experiments. In Scheiner S. M. and Gurevitch J. (Eds.) *Design and analysis of ecological experiments*. New York: Chapman & Hall, 342-359.
- Geary R. C. (1954). – The contiguity ratio and statistical mapping. *Incorporated Statistician*, **5**, 115-145.
- Gower J. C. & Legendre P. (1986). – Metric and euclidean properties of dissimilarity coefficients. *Journal of Classification*, **3**, 5-48.
- Isaaks E. H. & Srivastava R. M. (1989). – *Applied geostatistics*. Oxford: Oxford University Press.
- Jackson R. B. & Caldwell M. M. (1993). – Geostatistical patterns of soil heterogeneity around individual perennial plants. *Journal of Ecology*, **81**, 683-692.
- Legendre L. & Legendre P. (1984). – *Ecologie numérique*. 2^e éd. Tome 2: *La structure des données écologiques*. Paris: Masson et les Presses de l'Université du Québec.
- Legendre P. (1993). – Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**, 1659-1673.
- Legendre P. & Fortin M. J. (1989). – Spatial pattern and ecological analysis. *Vegetatio*, **80**, 107-138.
- Legendre P., Oden N. L., Sokal R. R., Vaudor A. & Kim J. (1990). – Approximate analysis of variance of spatially autocorrelated regional data. *Journal of Classification*, **7**, 53-75.

- Legendre P. & Troussellier M. (1988). – Aquatic heterotrophic bacteria: Modeling in the presence of spatial autocorrelation. *Limnology and Oceanography*, **33**, 1055-1067.
- Legendre P. & Vaudor A. (1991). – *The R Package: Multidimensional analysis, spatial analysis*. Montréal: Département des Sciences Biologiques, Université de Montréal.
- Liebholt A. M., Rossi R. E. & Kemp W. P. (1993). – Geostatistics and geographic information systems in applied insect ecology. *Annual Review of Entomology*, **38**, 303-327.
- Mantel N. (1967). – The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209-220.
- Moran P. A. P. (1950). – Notes on continuous stochastic phenomena. *Biometrika*, **37**, 17-23.
- Nicot P. C., Rouse D. I. & Yandell B. S. (1984). – Comparison of statistical methods for studying spatial patterns of soilborne plant pathogens in the field. *Phytopatology*, **74**, 1399-1402.
- Oden N. L. (1984). – Assessing the significance of spatial correlograms. *Geographical Analysis*, **16**, 1-16.
- Robertson G. P. & Freckman D. (1995). – The spatial distribution of nematode trophic groups across a cultivated ecosystem. *Ecology*, **76**, 1425-1432.
- Robertson G. P. & Gross K. L. (1994). – Assessing the heterogeneity of belowground resources: Quantifying pattern and scale. In: Caldwell M. and Pearcy R. (Eds.) *Plant exploitation and environmental heterogeneity*. New York: Academic Press, 237-253.
- Rossi J. P., Delaville L. & Quénéhervé P. (1996). – Microspatial structure of a plant-parasitic nematode community in a sugarcane field in Martinique. *Applied Soil Ecology*, **3**, 17-26.
- Rossi J. P., Lavelle P. & Albrecht A. (1997). – Relationships between spatial pattern of the endogeic earthworm *Polypheretima elongata* and soil heterogeneity. *Soil Biology and Biochemistry* (in press).
- Rossi J. P., Lavelle P. & Tondoh J. E. (1995). – Statistical tool for soil biology X. Geostatistical analysis. *European Journal of Soil Biology*, **31**, 173-181.
- Seinhorst J. W. (1950). – De betekenis van de toestand van de grond voor het optreden van aanstasting door het stengelaaltje (*Ditylenchus dipsaci* (Khün) Filipjev). *Tidschrift over Plantenziekten*, **56**, 289-348.
- Seinhorst J. W. (1962). – Modifications of the elutriation method for extracting nematodes from soil. *Nematologica*, **8**, 117-128.
- Siddiqui M. R. & Brown K. F. (1964). – *Helicotylenchus retusus* n.sp. (Nematoda: Hoplolaiminae) found around sugar cane roots in Negros Oriental, Philippines. *Proceeding of the Helminthological Society of Washington*, **31**, 209-211.
- Smouse P. E., Long J. C. & Sokal R. R. (1986). – Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, **35**, 627-632.
- Sokal R. R. & Oden N. L. (1978). – Spatial autocorrelation in biology. 1. Methodology. *Biological Journal of the Linnean Society*, **10**, 199-228.
- Sokal R. R. & Rohlf F. J. (1995). – *Biometry: the principles and practice of statistics in biological research*. 3rd Ed. New York: W. H. Freeman and company.
- Taylor L. R. (1961). – Aggregation, variance and the mean. *Nature*, **189**, 732-735.
- Taylor L. R., Perry J. N., Woiwod I. P. & Taylor R. A. J. (1988). – Specificity of the spatial power-law exponent in ecology and agriculture. *Nature*, **332**, 721-722.
- Wallace M. K. & Hawkins D. M. (1994). – Applications of geostatistics in plant nematology. *Journal of Nematology*, **26**, 626-634.
- Webster R. & Oliver M. A. (1990). – *Statistical methods in soil and land resource survey*. Oxford: Oxford University Press.



Fundamental and Applied NEMATODOLOGY

A leading publication in Nematology
co-edited by Gauthier-Villars and Orstom

From ultrastructure to field work

- *No page charge for regular publication in **Fundamental and Applied Nematology**.*
- *50 free reprints*
- *If you are interested in shorten your delays of publication, do not hesitate to ask for the "Rush publication" instructions*
- A worldwide distribution over 50 countries
- FAN publishes:
 - Original articles
 - Short notes
 - Book reviews
 - Forum articles
 - Symposium papers

Honorary editor: M. Luc
Editor in chief: P. Baujard
Associate Editor: T. C. Vrain
Editing advisor: G. Merny
Editorial board: I. Andrassy,
 J. Bridge, P. De Ley, D. G. De Waele,
 A. A. F. Evans, C. Griffin, N. G. M. Hague,
 B. Jaffee, D. Mugniery, C. E. Opperman,
 J. Pinochet, G. O. Poinar Jr., T. O. Powers,
 G. Reversat, R. Rodriguez-Kabana,
 J. L. Starr, D. L. Trudgill, B. Weischer,
 U. Wyss, G. W. Yeates.

Indexed in

Current Contents (Agriculture, Biology and Environmental Sciences), Nematological Abstracts: Agricola, Agris, Biosis, CABI, INIST, Horizon.

ORDER FORM

Return this form to your usual supplier or to Gauthier-Villars Publishers - Promotion dept. - 120, boulevard Saint-Germain - 75006 Paris - France
 Tel. (33) 01 40 46 62 00 - Fax (33) 01 40 46 62 31 - e-mail: gauthier.villars.publisher@mail.sgip.fr

Please enter my subscription to **FUNDAMENTAL AND APPLIED NEMATODOLOGY** for 1997 - 6 issues - ISSN 0012-9593 at the rate of 1 450 FF (France only 1 100 FF)
 Special reduced rates for members of ESN, SON, ONTA.
 Export 1 080 FF (France only 825 FF)

Payment enclosed by cheque payable to SPES

Please charge my credit card (VISA/EUROCARD/MASTERCARD)

N° _____

Expiration date _____

Signature _____

Please send me a *proforma* invoice

Please send me a free sample copy

Name

Address

Zip code Town

Country

You are Researcher Professor Librarian
 Other

Field of activity

Affiliation University Industry

Other

INFORMATION ON LINE

<http://www.gauthier-villars.fr>

Gauthier-Villars abides by French law 78.17.

ORSTOM
gauthier-villars