geographical analysis

Geographical Analysis ISSN 0016-7363

A New Perspective about Moran's Coefficient: Spatial Autocorrelation as a Linear Regression Problem

Stéphane Dray[1]

[1]Université de Lyon, F-69000, Lyon; Université Lyon 1; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69622, Villeurbanne, France

The computation of Moran's index of spatial autocorrelation requires the definition of a spatial weighting matrix. The eigendecomposition of this doubly centered matrix (i.e., one that forces the sums of all rows and columns to equal zero) has interesting properties that have been exploited in various contexts: distribution properties of the Moran coefficient (MC), spatial filtering in linear models, generalized linear models, and multivariate analysis. In this article, this eigendecomposition is used to propose a new view of MC based on its interpretation in the simple context of linear regression. I use this interpretation to demonstrate the different properties of MC and also the inefficiency of this index in some situations involving simultaneous positive and negative spatial autocorrelation. I propose some new statistics and procedures for testing spatial autocorrelation, and conduct a simulation study to evaluate these new approaches.

Introduction

Identifying and estimating spatial autocorrelation are of prime interest in studies dealing with georeferenced data (Getis 2008). The most commonly used global measures of spatial autocorrelation are the Moran (1948) coefficient (MC) and Geary (1954) ratio. Moran's index of spatial autocorrelation has been used in many fields to identify spatial structures. It has also formed the basis of several methodological developments, including local indicators of spatial association (Anselin 1995), bivariate indices of spatial correlation (Lee 2001), and spatial multivariate analysis (Wartenberg 1985; Dray, Saïd, and Débias 2008). The computation of MC requires the definition of a spatial weighting matrix **W** that indicates the strength of the potential interaction between spatial units. The choice of **W** is a critical step in spatial analysis and can influence the significance of testing procedures

Correspondence: Stéphane Dray, Université de Lyon, F-69000 Lyon; Université Lyon 1, CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69622 Villeurbanne, France
e-mail: stephane.dray@univ-lyon1.fr

Submitted: July 4, 2008. Revised version accepted: October 6, 2009.

(Tiefelsdorf, Griffith, and Boots 1999). The eigendecomposition of this doubly centered matrix (i.e., one that forces the sums of all rows or columns to be equal to zero) has interesting properties that have been exploited in various contexts. For instance, the eigenvalues of **W** have been used to study the distribution of MC (de Jong, Sprenger, and van Veen 1984; Tiefelsdorf and Boots 1995) and to furnish its moment-generating functions (Cliff and Ord 1973, p. 34). Eigenvectors of this matrix can be used for spatial filtering purposes (Getis and Griffith 2002; Griffith 2003) and introduced as spatial predictors in linear models (Griffith 1996, 2000a), generalized linear models (Griffith 2002, 2004a), and multivariate analysis (Dray, Legendre, and Peres-Neto 2006; Jombart, Dray, and Dufour 2009).

In this article, I propose a new way of looking at MC based on its interpretation in the simple and classical context of linear regression. This particular regression considers the eigenvectors of the doubly centered spatial weighting matrix **W** as explanatory variables. Using this interpretation, I demonstrate various properties of MC and highlight its inefficiency in some situations involving a mixture of positive and negative spatial autocorrelation. I propose new statistics and procedures for testing spatial autocorrelation, and conduct a simulation study to evaluate these new approaches.

A classical approach using the MC

The MC is a very classical measure of spatial autocorrelation that has been popularized by Cliff and Ord (1973). In this part, I describe the mathematical formulation of the MC and focus on the definition of the spatial weighting matrix required in its computation. Then, I present a permutation procedure that is widely used to evaluate the significance of observed values of MC.

Definition

I consider the n-by-1 vector $\mathbf{x} = [x_1 \ldots x_n]^T$ containing measurements of a quantitative variable for n spatial units together with an n-by-n symmetric spatial weighting matrix **W** (Bavaud 1998). This matrix is a mathematical representation of the geographical layout of the region under study. The values w_{ij} stored in matrix **W** are spatial weights and provide a measure of the degree to which the units i and j are correlated with each other. The simplest weighting matrix is a connectivity matrix, so that $w_{ij} = 1$ if spatial units i and j are neighbors, and $w_{ij} = 0$ otherwise. More sophisticated weighting matrices (Getis and Aldstadt 2004) are able to take into account the distance between the spatial units or the length of the common boundary between the regions for areal data. The definition of this matrix should be relevant to the context of the applied problem and must be made before an analysis. This matrix is usually symmetric ($w_{ij} = w_{ji}$), but the results presented here can be generalized to a nonsymmetric matrix $\mathbf{W}^*$ using $\mathbf{W} = (\mathbf{W}^* + \mathbf{W}^{*T})/2$.

The definition of MC (Moran 1948; Cliff and Ord 1973) using matrix notation is

$$MC(\mathbf{x}) = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \frac{\mathbf{x}^T\mathbf{H}\mathbf{H}\mathbf{W}\mathbf{H}\mathbf{H}\mathbf{x}}{\mathbf{x}^T\mathbf{H}\mathbf{H}\mathbf{x}} = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \frac{\mathbf{x}^T\mathbf{H}\mathbf{W}\mathbf{H}\mathbf{x}}{\mathbf{x}^T\mathbf{H}\mathbf{x}}, \quad (1)$$

where $\mathbf{H} = (\mathbf{I} - \mathbf{1}\mathbf{1}^T/n)$ is a centering operator, $\mathbf{I}$ is the n-by-n identity matrix, and $\mathbf{1}$ is the n-by-1 vector of ones. $\mathbf{H}$ is an orthogonal projector, and hence $\mathbf{H} = \mathbf{H}^2$ (i.e., $\mathbf{H}$ is idempotent).

A testing procedure

Normal approximations (Cliff and Ord 1973) to the distributions of MC can be based on the assumption that data are independent realizations either from a normal distribution (assumption N), or from an unknown distribution for which all random permutations can be considered (assumption R). An assessment of the significance of observed values of MC is performed either by employment of the assumption that the statistic follows a normal distribution or by simulation experiments that randomize the locations of a set of observed values. Assumptions for the normal approximation are rarely encountered in applications (e.g., Tiefelsdorf 2002); thus, I focus on a Hope (1968) type permutation test, which uses a randomly selected subset of the $n!$ possible permutations.[1] It involves the following steps:

(1) compute MC for a set of observed values $\mathbf{x}$;
(2) randomize the values within $\mathbf{x}$;
(3) compute MC for the permuted values;
(4) repeat steps (2) and (3) a large number of times (e.g., 4,999 times) to construct a distribution of MC according to the hypothesis of spatial independence (H_0) and,
(5) compare the observed statistic to the distribution obtained by permutation, and make the appropriate statistical decision.

The fifth step involves various different procedures corresponding to the specific alternative hypotheses (H_1). If the alternative hypothesis is the presence of positive autocorrelation, using a significance level equal to α, the P-value is estimated as (number of random values equal to or greater than the observed value+1)/(number of permutations+1). The null hypothesis is rejected if the P-value is less than the significance level α. I refer to this procedure hereafter as ψ^g (g for greater). If the alternative hypothesis is the presence of negative autocorrelation, using a significance level equal to α, the P-value is estimated as (number of random values equal to or less than the observed value+1)/(number of permutations+1). Again, the null hypothesis is rejected if the P-value is less than the significance level α. I refer to this procedure hereafter as ψ^l (l for less). Finally, if the alternative hypothesis is the presence of autocorrelation (positive or negative), a two-sided test can be used. The estimation of the P-value is equivalent to the ψ^g procedure, except that random and observed values are first centered (using the average of the random

values) and then transformed to their absolute values. I refer to this procedure hereafter as ψ^t (t for two-sided).

Eigenfunctions of a spatial weighting matrix

de Jong, Sprenger, and van Veen (1984) show that the upper and lower bounds of MC for a spatial matrix $\mathbf{W}$ are given by $\lambda_{max}(n/\mathbf{1}^T\mathbf{W}\mathbf{1})$ and $\lambda_{min}(n/\mathbf{1}^T\mathbf{W}\mathbf{1})$, where λ_{max} and λ_{min} are the extreme eigenvalues of $\mathbf{\Omega} = \mathbf{HWH}$. Hence, the eigenvectors of $\mathbf{\Omega}$ are vectors with a unit norm (for convenience, this reduces the Rayleigh quotient denominator to 1) maximizing MC. The eigenvalues of this matrix are equal to the MCs indexing spatial autocorrelation (postmultiplied by a constant) and can be either positive or negative. Eigenvectors associated with high positive (or negative) eigenvalues have high positive (or negative) autocorrelation. The eigenvectors associated with eigenvalues with extremely small absolute values correspond to low spatial autocorrelation and are not suitable for defining spatial structures (Dray, Legendre, and Peres-Neto 2006).

The diagonalization of a spatial weighting matrix consists of finding the normalized vectors $\mathbf{u}_i$, stored as columns in the matrix $\mathbf{U} = [u_{ij}] = [\mathbf{u}_1 \ldots \mathbf{u}_n]$, satisfying

$$\Omega = \mathbf{HWH} = \mathbf{U}\Lambda\mathbf{U}^T = \sum_{k=1}^{n} \lambda_k \mathbf{u}_k \mathbf{u}_k^T, \tag{2}$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1 \ldots \lambda_n)$, $\mathbf{u}_k^T\mathbf{u}_k = ||\mathbf{u}_k||^2 = 1$, and $\mathbf{u}_k^T\mathbf{u}_j = 0$ for $k \neq j$. Note that double centering of $\mathbf{\Omega}$ implies that the eigenvectors $\mathbf{u}_k$ are centered and that at least one eigenvalue is equal to zero.

Introducing these eigenvectors into the original formulation of Moran's index (1) leads to

$$MC(\mathbf{x}) = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \frac{\mathbf{x}^T\mathbf{HWHx}}{\mathbf{x}^T\mathbf{Hx}} = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \frac{\mathbf{x}^T\mathbf{U}\Lambda\mathbf{U}^T\mathbf{x}}{\mathbf{x}^T\mathbf{Hx}} = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \frac{\sum_{k=1}^{n} \lambda_k \mathbf{x}^T\mathbf{u}_k\mathbf{u}_k^T\mathbf{x}}{\mathbf{x}^T\mathbf{Hx}}. \tag{3}$$

Considering the vector $\mathbf{z}$ containing the z-score transformation of $\mathbf{x}$ (i.e., $z_i = (x_i - \bar{x})/\sqrt{\text{var}(\mathbf{x})}$, where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\text{var}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$) and using the properties of idempotence of $\mathbf{H}$, (3) is equivalent to

$$MC(\mathbf{x}) = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \frac{\sum_{k=1}^{n} \lambda_k \mathbf{z}^T\mathbf{u}_k\mathbf{u}_k^T\mathbf{z}}{\mathbf{z}^T\mathbf{z}} = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \frac{\sum_{k=1}^{n} \lambda_k ||\mathbf{u}_k^T\mathbf{z}||^2}{n}. \tag{4}$$

From an autocorrelation to a correlation coefficient

As the eigenvectors $\mathbf{u}_k$ have unit norm and are centered, whereas the vector $\mathbf{z}$ has unit variance and is centered, the equation (4) can be rewritten as

$$MC(\mathbf{x}) = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \frac{\sum_{k=1}^{n} \lambda_k \cdot \text{cor}^2(\mathbf{u}_k, \mathbf{z}) \cdot n}{n} = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \sum_{k=1}^{n} \lambda_k \text{cor}^2(\mathbf{u}_k, \mathbf{z}). \tag{5}$$

Moran's index for a given eigenvector $\mathbf{u}_k$ is equal to $MC(\mathbf{u}_k) = (n/\mathbf{1}^T\mathbf{W}\mathbf{1})\lambda_k$. Therefore, (5) can be rewritten as

$$MC(\mathbf{x}) = \sum_{k=1}^{n} MC(\mathbf{u}_k)\text{cor}^2(\mathbf{u}_k, \mathbf{z}). \quad (6)$$

The term cor^2 ($\mathbf{u}_k$, $\mathbf{z}$) represents the part of the variance of $\mathbf{z}$ that is explained by $\mathbf{u}_k$ in the linear model $\mathbf{z} = b_k\mathbf{u}_k + \varepsilon_k$, where b_k is the regression coefficient and ε_k is the n-by-1 vector of error terms. This quantity is equal to b_k^2/n. By construction (equation [2]), the eigenvectors $\mathbf{u}_k$ are orthogonal and uncorrelated (Griffith 2000b), and therefore the regression coefficients of the linear models $\mathbf{z} = b_k\mathbf{u}_k + \varepsilon_k$ are those of the multiple regression model $\mathbf{z} = \mathbf{U}\mathbf{b} + \varepsilon = b_1\mathbf{u}_1 + \cdots + b_n\mathbf{u}_n + \varepsilon$, where $\mathbf{b} = [b_1 \ldots b_n]^T$ is the n-by-1 vector of regression coefficients and ε is the n-by-1 vector of error terms. Griffith (2004b) uses this property to formulate a new eigenvector stepwise selection procedure that minimizes sequentially the residual MC instead of maximizing the fit of the model.

The interpretation of MC in the context of multiple linear regression allows us to restate the following well-known properties of this measurement:

- The maximum value of MC is obtained if all of the variation of $\mathbf{z}$ is explained by the eigenvector $\mathbf{u}_1$ corresponding to the highest eigenvalue λ_1. In this case, cor^2 ($\mathbf{u}_1$, $\mathbf{z}$) = 1 (and cor^2 ($\mathbf{u}_k$, $\mathbf{z}$) = 0 for $k \neq 1$), and the maximum value of MC, which can be deduced from (6), is equal to $MC_{\max} = \lambda_1(n/\mathbf{1}^T\mathbf{W}\mathbf{1})$.
- The minimum value of MC is obtained if all of the variation of $\mathbf{z}$ is explained by the eigenvector $\mathbf{u}_n$, which corresponds to the most negative eigenvalue λ_n. This minimum value is equal to $MC_{\min} = \lambda_n(n/\mathbf{1}^T\mathbf{W}\mathbf{1})$.
- If the variable of interest is not spatialized (population parameter $R^2_{\mathbf{u}_k,\mathbf{z}} = 0$), the expected value of the variance explained by each eigenvector is equal to $E(R^2_{\mathbf{u}_k,\mathbf{z}} | R^2_{\mathbf{u}_k,\mathbf{z}} = 0) = E(\text{cor}^2(\mathbf{u}_k, \mathbf{z})) = 1/(n-1)$(Kendall and Stuart 1961, p. 341), where $R^2_{\mathbf{u}_k,\mathbf{z}}$ is the coefficient of determination associated to the regression model $\mathbf{z} = b_k\mathbf{u}_k + \varepsilon_k$ for a given sample ($R^2_{\mathbf{u}_k,\mathbf{z}}$ is the population parameter). For instance, if the values in $\mathbf{z}$ are randomly permuted, we should obtain this result on average. Combining the set of $n!$ random permutations (assumption R) with (6) implies that $E_R(MC) = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}(n-1)}\sum_{k=1}^{n}\lambda_k = \frac{n}{\mathbf{1}^T\mathbf{W}\mathbf{1}(n-1)}\text{trace}(\Omega)$. Accordingly, $\text{trace}(\Omega) = -\frac{\mathbf{1}^T\mathbf{W}\mathbf{1}}{n}$, and thus $E_R(MC) = -\frac{1}{n-1}$.

Spatial heterogeneity and the MC

The MC is a global statistic in the sense that it estimates the overall degree of spatial autocorrelation for an entire study area. Conceivably, different spatial patterns could lead to the same value of MC. Interpretating MC in the linear regression framework allows one to make some intuitive suppositions about the sign and the significance of MC. Different situations may be encountered, and I illustrate some of them in the following scenarios, also depicted in Fig. 1:

- Positive and significant autocorrelation can be obtained if $\mathbf{z}$ is highly correlated with an eigenvector associated with a large or moderate positive eigenvalue

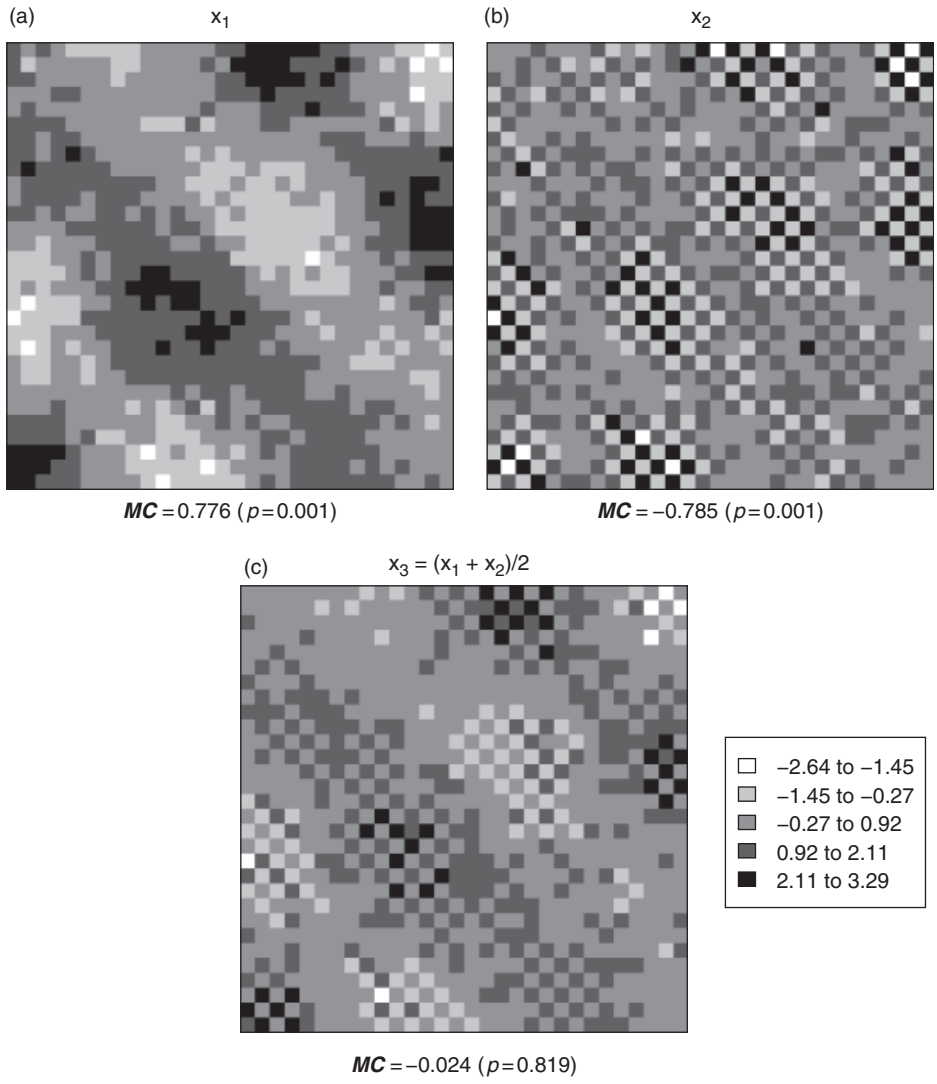

Figure 1. A simple example portraying the mapping of the values for three sample vectors and the computation of MC for a regular 30-by-30 square grid (the spatial weighting matrix was constructed with a rook neighbor definition and was row standardized). Testing procedures ψ^g, ψ^l, and ψ^t were used for vectors (a) $\mathbf{x}_1$, (b) $\mathbf{x}_2$, and (c) $\mathbf{x}_3$, respectively.

(Fig. 1a). It can also be obtained if **z** is moderately correlated with a set of eigenvectors associated with large or moderate positive eigenvalues.

- Negative and significant autocorrelation is obtained if **z** is highly correlated with an eigenvector associated with a high or moderate negative eigenvalue (Fig. 1b). It can also be obtained if **z** is moderately correlated with a set of eigenvectors associated with high or moderate negative eigenvalues.

- Null and nonsignificant autocorrelation is obtained if $\mathbf{z}$ is only slightly correlated with eigenvectors. Null autocorrelation also can result when $\mathbf{z}$ satisfies simultaneously the conditions for positive and negative autocorrelation. In this case, positive and negative elements in $\sum_{k=1}^{n} MC(\mathbf{u}_k)\text{cor}^2(\mathbf{u}_k, \mathbf{z})$ cancel each other out, resulting in a null and nonsignificant autocorrelation (Fig. 1c).

In other words, varying the coefficients and the magnitude of noise in the linear combination $\sum_{k=1}^{n} b_k \mathbf{u}_k + \varepsilon$ results in these different scenarios.

The third situation can also be interpreted in the context of nonstationarity. Due to spatial heterogeneity, the degree of autocorrelation may vary significantly across space. Local statistics (Anselin 1995), such as local MCs (MC_i), provide estimates of the dependency relationships across space. The decomposition of the global MC into a sum of local MC_i allows estimation of the contribution of each observation:

$$MC(\mathbf{x}) = \frac{1}{\mathbf{1}^T\mathbf{W}\mathbf{1}} \sum_{i=1}^{n} MC_i(\mathbf{x}), \quad \text{where } MC_i(\mathbf{x}) = z_i \sum_{j=1}^{n} w_{ij} z_j. \tag{7}$$

Usually significance tests of local MC_is are based on the normal approximation; see Tiefelsdorf (2002) or Lee (2009) for alternatives. The three preceding scenarios can be interpreted using these local statistics. A positive and significant autocorrelation corresponds to the situation in which many local MC_i values are both positive and significant (Fig. 2a). Negative and significant autocorrelation corresponds to the situation in which many local MC_i values are both negative and significant (Fig. 2b). The third situation (null and nonsignificant autocorrelation) could lead to several different interpretations. For instance, if many local MC_is are positive and significant, whereas many others are negative and significant, their sum (i.e., the global MC) would be null. This case corresponds to pockets of spatial nonstationarity. However, the example presented in Fig. 1c leads to a different situation in which the global MC is nonsignificant because the local MC_is are null and nonsignificant (Fig. 2c).

This third situation is quite problematic and highlights the inadequacy of a global MC and its local MC_is for detecting spatial structures in some situations involving mixtures of positive and negative spatial autocorrelation. Structures of this kind can occur in various fields, usually as the consequence of several processes acting simultaneously that influence the level of autocorrelation. The spatial distribution of species abundance is a variable of interest in ecology (Legendre 1993). Environmental factors that influence this distribution are usually spatially structured at a broad scale and thus induce positive autocorrelation. Biotic processes, such as competition between individuals or species, act at the community level and can create negative autocorrelation. In most situations, the spatial heterogeneity of a species distribution is due to the simultaneous impacts of these two processes and thus could simultaneously include positive and negative spatial autocorrelation.

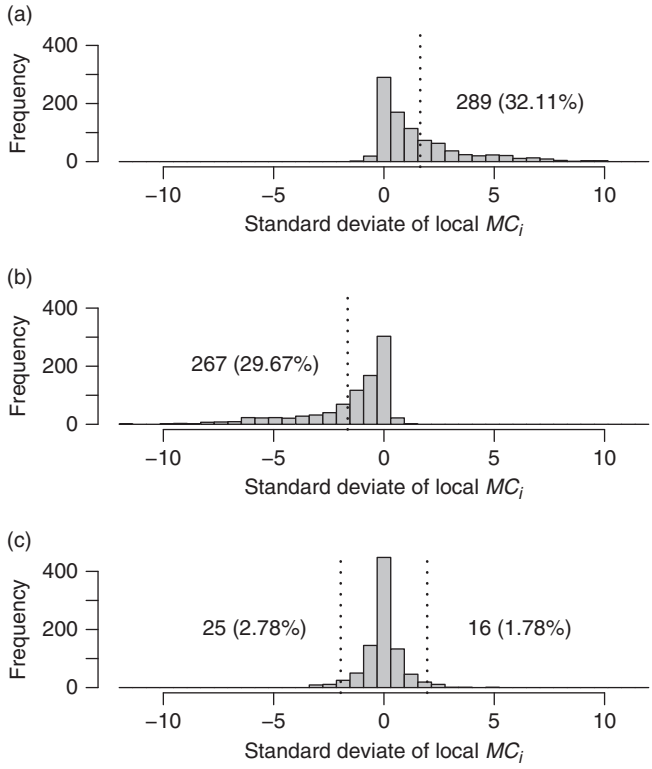

Figure 2. The distribution of the standard deviate of the local MC_i for the three sample vectors mapped in Fig. 1. In (a), $\mathbf{x}_1$, 32.11% of the local MC_is are significant (i.e., greater than the 0.95 normal quantile), whereas 29.67% are significant for (b) $\mathbf{x}_2$. For (c), $\mathbf{x}_3$, 2.78% of the negative and 1.78% of the positive local MC_is are significant, using a bilateral test. Dotted lines represent the critical values of the test: 0.95 (a), 0.05 (b), and 0.025 and 0.0975 (c) normal quantiles.

Classical tools are not suitable for identifying such complex structures, and I propose new statistics and testing procedures based on rewriting the MC in the context of linear regression.

New testing procedures against the alternate hypotheses of spatial autocorrelation
As indicated by (6), an MC is the sum of positive and negative autocorrelation:

$$MC(\mathbf{x}) = \sum_{k=1}^{n} MC(\mathbf{u}_k)\mathrm{cor}^2(\mathbf{u}_k, \mathbf{z})$$
$$= \underbrace{\sum_{\substack{k=1,\\ MC(\mathbf{u}_k) < E_R(MC)}}^{n} MC(\mathbf{u}_k)\mathrm{cor}^2(\mathbf{u}_k, \mathbf{z})}_{\text{negative autocorrelation}} + \underbrace{\sum_{\substack{i=1,\\ MC(\mathbf{u}_k) > E_R(MC)}}^{n} MC(\mathbf{u}_k)\mathrm{cor}^2(\mathbf{u}_k, \mathbf{z})}_{\text{positive autocorrelation}}. \quad (8)$$

Hence, I propose two new statistics:

$$S^{-}(\mathbf{x}) = \sum_{\substack{k=1,\\ MC(\mathbf{u}_k) < E_R(MC)}}^{n} MC(\mathbf{u}_k)\mathrm{cor}^2(\mathbf{u}_k, \mathbf{z}) \text{ and} \tag{9}$$

$$S^{+}(\mathbf{x}) = \sum_{\substack{k=1,\\ MC(\mathbf{u}_k) > E_R(MC)}}^{n} MC(\mathbf{u}_k)\mathrm{cor}^2(\mathbf{u}_k, \mathbf{z}), \tag{10}$$

which are strict measurements of negative and positive spatial autocorrelation, respectively. By definition, we have $MC(\mathbf{x}) = S^{-}(\mathbf{x}) + S^{+}(\mathbf{x})$.

Using these new statistics, a permutation procedure to test against the alternative hypothesis of negative spatial autocorrelation is exactly equivalent to the previously described procedure ψ^{l}, except that the $S^{-}(\mathbf{x})$ statistic is used instead of $MC(\mathbf{x})$. I refer to this procedure hereafter as ψ^{-}. The new ψ^{+} procedure to test against the alternative hypothesis of positive spatial autocorrelation is exactly equivalent to the previously described ψ^{g} procedure, except that the $S^{+}(\mathbf{x})$ statistic is used instead of $MC(\mathbf{x})$. Finally, I propose a two-sided test that considers the alternative hypothesis of positive or negative autocorrelation. The procedure involves several steps:

(1) The ψ^{-} procedure is used to test against the alternative hypothesis of negative autocorrelation. The *P*-value P^{-} (i.e., the probability of obtaining a given value of $S^{-}(\mathbf{x})$ equal to or less than the observed value by chance alone) is estimated.

(2) The ψ^{+} procedure is used to test against the alternative hypothesis of positive autocorrelation. The *P*-value P^{+} (i.e., the probability of obtaining a given value of $S^{+}(\mathbf{x})$ equal to or greater than the observed value by chance alone) is estimated.

(3) These two *P*-values are then combined to compute a global *P*-value for the two-sided test. This *P*-value is given by $2 \times \min(P^{-}, P^{+})$ (Manly 1997, p. 72). The null hypothesis is rejected if the *P*-value is less than the significance level α.

I refer to this two-sided testing procedure hereafter as ψ^{+-}.

Illustrations

I present a simulation study and an empirical example to illustrate the use of the new procedures. The simulation study was carried out to evaluate and compare the relative merits of the standard and new procedures, in terms of power and type I error under various conditions. The empirical example involves the analysis of mosquito counts for 120 sites sampled in Kenya. This real data set illustrates nicely the simultaneous presence of positive and negative autocorelation, and I show how this mixture can be identified and interpreted.

Designed simulation studies

I considered a regular 10-by-10 grid. The spatial weighting matrix $\mathbf{W}$ was constructed with a rook neighbor definition and was row standardized. Spatially autocorrelated data were generated as simultaneous autoregressive random variables (e.g., Haining 1990, p. 117). I used the following steps to obtain a sample with a given autocorrelation level ρ: (1) generate a vector $\mathbf{y}$ containing 100 iid normally distributed random values, (2) compute the inverse matrix $(\mathbf{I} - \rho\mathbf{W})^{-1}$, and (3) premultiply the vector $\mathbf{y}$ by the matrix obtained in (2) to obtain autocorrelated data in the vector $\mathbf{x}$ (i.e., $\mathbf{x} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{y}$). Note that for this particular specification of the spatial weighting matrix, the range of feasible values for ρ is $]-1; 1[$.

In a first study, simple spatial structures were generated for 13 different values of ρ (in sequence, with successive increments of 0.15, between -0.9 and 0.9). In a second study, complex spatial structures were generated by adding two autocorrelated vectors (one with positive autocorrelation level ρ_1 and one with negative autocorrelation level ρ_2): $\mathbf{x} = (\mathbf{I} - \rho_1\mathbf{W})^{-1}\mathbf{y}_1 + (\mathbf{I} - \rho_2\mathbf{W})^{-1}\mathbf{y}_2$. For the first vector, I took six values of ρ_1 between 0.15 and 0.9, with successive increments of 0.15. For the second vector, three values of ρ_2 were used (-0.3, -0.6, -0.9). Hence, $3 \times 6 = 18$ combinations were investigated. For each value (or combination of values) of the autocorrelation level(s), 5,000 samples were generated. For each sample, the testing procedures (ψ^I, ψ^g, ψ^t, ψ^-, ψ^+, ψ^{+-} for the first study and ψ^t, ψ^{+-} for the second one) were conducted with 4,999 random permutations.

For each value (or combination of values) of the autocorrelation level(s), I report the rate of rejection of the null hypothesis at significance level $\alpha = 0.05$ over the 5,000 generated samples for each testing procedure.

An empirical example

This real data set concerns the analysis of mosquito counts (*Anopheles gambiae* s.l.) in the urban area of Karima (Kenya). These data were kindly provided by Dr. B. Jacob (University of Alabama, Birmingham). One hundred and twenty sites were sampled, and a Thiessen polygon tessellation was superimposed on the area to construct the spatial weighting matrix. Two sites were considered to be neighbors if they shared a common Thiessen polygon boundary. The neighborhood structure was then edited to remove unlikely geographic neighbors, and lastly the spatial weighting matrix was row standardized. Autocorrelation statistics (MC, S^+, S^-) were computed for log-transformed counts and tested with 4,999 permutations.

Results

Results for the two simulation studies are presented in Figs. 3 and 4. In the case of simple structures, the new procedures are as powerful as the standard approaches based on MC. The second simulation study and the empirical example show the inefficiency of the standard two-sided procedure ψ^t to detect the presence of a

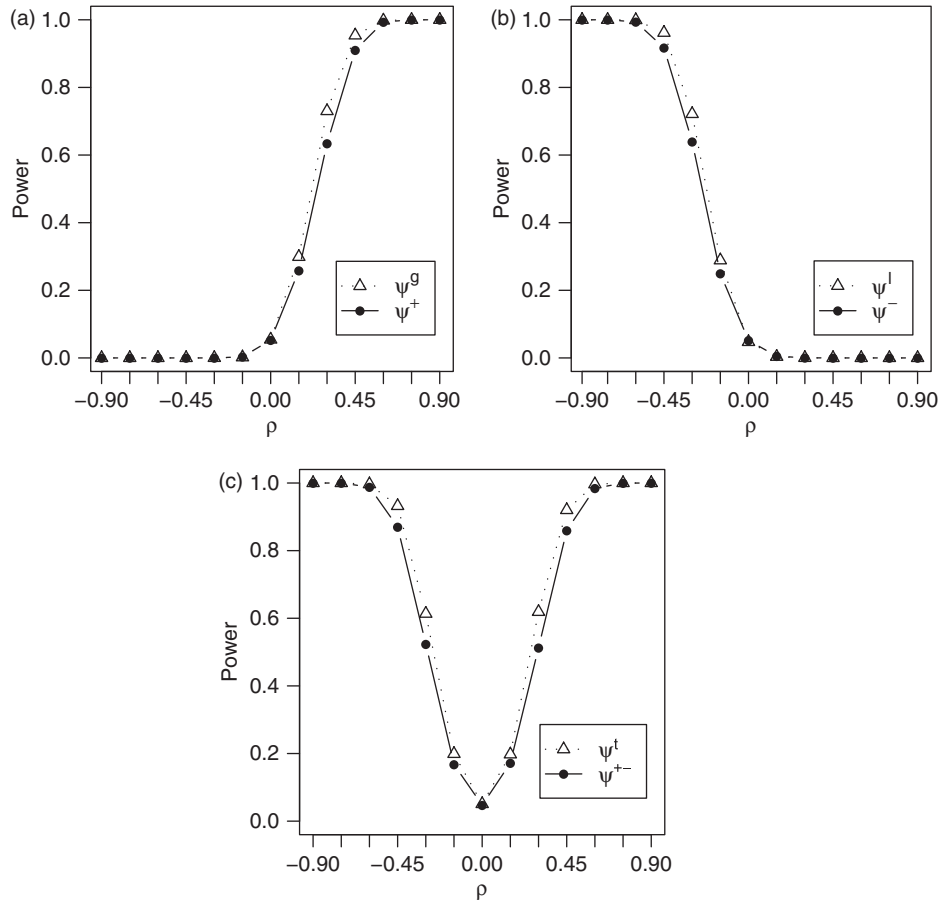

Figure 3. Results of the simulation study for simple spatial structures. Influence of the autocorrelation level ρ on the rejection rates of the null hypothesis at significance level $\alpha = 0.05$ based on 5,000 independent simulations. Results are reported for the ψ^g, ψ^+ (a),ψ^l, ψ^- (b) and the ψ^t, ψ^{+-} (c) procedures.

mixture of positive and negative autocorrelation. In this context, the new procedure ψ^{+-} is much more powerful and demonstrates its usefulness.

Designed simulation studies

The results of the first simulation study are portrayed in Fig. 3. Rejection rates of the null hypothesis correspond to type I error when $\rho = 0$, or to power ($\rho \neq 0$ for ψ^{+-} and ψ^t; $\rho < 0$ for ψ^- and ψ^l; $\rho > 0$ for ψ^+ and ψ^g). All procedures have an acceptable type I error (the range of variation is 0.0462–0.0504). Results obtained with the new procedures ψ^-, ψ^+, and ψ^{+-} are consistent with those obtained using the standard procedures ψ^l, ψ^g, and ψ^t. The highest observed difference is equal to 0.1002 (ψ^t versus ψ^{+-} for $\rho = 0.3$). Power increases substantially with the level of autocorrelation (> 0.98 for $|\rho| \geq 0.6$).

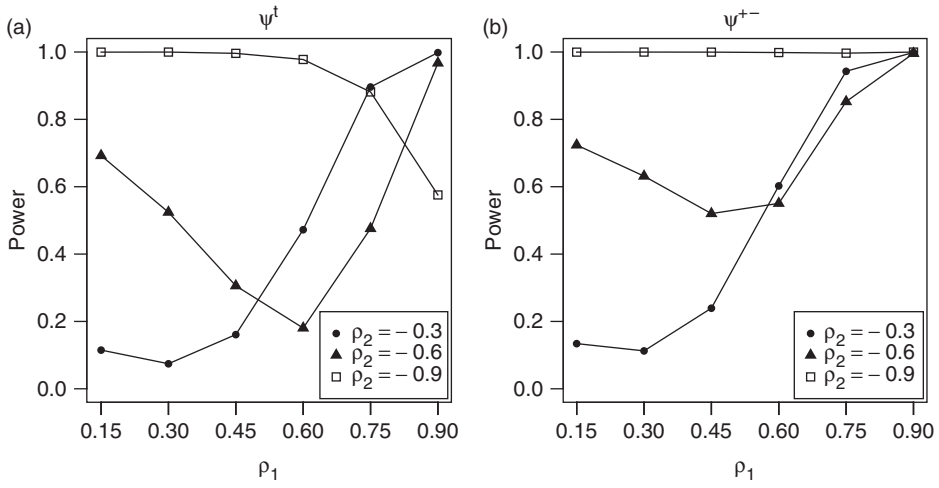

Figure 4. Results of the simulation study for complex spatial structures (i.e., the sum of two autocorrelated vectors). Influence of the autocorrelation levels ρ_1 and ρ_2 on the rejection rates of the null hypothesis at significance level $\alpha = 0.05$ based on 5,000 independent simulations. Results are reported for the ψ^t (a) and the ψ^{+-} (b) procedures.

For the second simulation study (Fig. 4), only the two-sided testing procedures are considered, and the rejection rates of the null hypothesis correspond to power. For a given value of ρ_2, the power of the two procedures is lowest when $|\rho_2| = |\rho_1|$ (except for ψ^{+-} and $\rho_2 = -0.6$, where the power is 0.5072 for $\rho_1 = 0.45$ and 0.5218 for $\rho_1 = 0.6$). For all combinations of autocorrelation parameters, the power of the ψ^{+-} procedure is always greater than the power of the standard procedure ψ^t. The average power difference equals 0.103 (the maximum difference is 0.4158 for $\rho_1 = 0.9$ and $\rho_1 = -0.9$).

An empirical example

The standard global test is not significant ($MC = 0.0794$, $P = 0.1074$ for the ψ^t procedure), which implies the absence of spatial autocorrelation. The new procedures reveal a significant positive autocorrelation ($S^+ = 0.2607$, $P = 0.0238$ for the ψ^+ procedure) and a nonsignificant negative autocorrelation ($S^- = -0.1813$, $P = 0.644$ for the ψ^- procedure) component. The two-sided procedure ψ^{+-} leads to a significant autocorrelation ($MC = 0.0794$, $P = 2 \times 0.0238 = 0.0476$). These results could be explained by two counteracting spatial effects (Jacob, Griffith, and Novak 2008): environmental factors indicating suitable aquatic habitats induce positive spatial autocorrelation, whereas competitive factors result in negative spatial autocorrelation. These two factors tend to have equal impacts leading to their nondetection by classical autocorrelation procedures.

Discussion and conclusions

The simulation studies show that the new procedures have acceptable type I errors, and that they are as powerful as standard procedures for simple spatial structures. The new procedures are more powerful than standard procedures in a context of simultaneous positive and negative autocorrelation. In the past, such situations may have led researchers to consider their data to be independent (as they could not reject the hypothesis of the absence of spatial autocorrelation using standard procedures) when both positive and negative autocorrelation were present simultaneously. The consequences of this incorrect inference affect the choice of appropriate statistical methods, or model misspecification, and need to be handled better in future analyses. Scientists are usually concerned with the question of positive spatial autocorrelation, and so negative autocorrelation is often ignored. Griffith (2006) identifies this problem of ''hidden negative spatial autocorrelation'' and shows how spatial filtering techniques could help in this context. The new procedures proposed in this article could be useful in these situations because they provide an effective way to measure and test complex spatial structures in which positive and negative spatial autocorrelation are present simultaneously. The $S^{+}(\mathbf{x})$ and $S^{-}(\mathbf{x})$ statistics could also be considered as possible alternatives to the criteria proposed by Tiefelsdorf and Griffith (2007) to select eigenvectors in the context of spatial filtering.

Acknowledgements

I warmly thank Dr. B. Jacob for providing the mosquito data. This article benefited from the helpful comments of three anonymous referees and the editor of this journal (Daniel A. Griffith).

Notes

1 This assumes that each value in a set of n observed values is unique. Otherwise, a multinomial coefficient counts the number of possible permutations.

References

Anselin, L. (1995). ''Local Indicators of Spatial Association.'' *Geographical Analysis* 27, 93–115.

Bavaud, F. (1998). ''Models for Spatial Weights: A Systematic Look.'' *Geographical Analysis* 30, 153–71.

Cliff, A., and J. Ord. (1973). *Spatial Autocorrelation*. London: Pion.

Dray, S., P. Legendre, and P. Peres-Neto. (2006). ''Spatial Modeling: A Comprehensive Framework for Principal Coordinate Analysis of Neighbor Matrices (PCNM).'' *Ecological Modelling* 196, 483–93.

Dray, S., S. Saïd, and F. Débias. (2008). ''Spatial Ordination of Vegetation Data Using a Generalization of Wartenberg's Multivariate Spatial Correlation.'' *Journal of Vegetation Science* 19, 45–56.

Geary, R. (1954). ''The Contiguity Ratio and Statistical Mapping.'' *The Incorporated Statistician* 5, 115–45.
Getis, A. (2008). ''A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective.'' *Geographical Analysis* 40, 297–309.
Getis, A., and J. Aldstadt. (2004). ''Constructing the Spatial Weights Matrix Using a Local Statistic.'' *Geographical Analysis* 36, 90–104.
Getis, A., and D. A. Griffith. (2002). ''Comparative Spatial Filtering in Regression Analysis.'' *Geographical Analysis* 34, 130–40.
Griffith, D. A. (1996). ''Spatial Autocorrelation and Eigenfunctions of the Geographic Weights Matrix Accompanying Geo-Referenced Data.'' *Canadian Geographer* 40, 351–67.
Griffith, D. A. (2000a). ''A Linear Regression Solution to the Spatial Autocorrelation Problem.'' *Journal of Geographical Systems* 2, 141–56.
Griffith, D. A. (2000b). ''Eigenfunction Properties and Approximations of Selected Incidence Matrices Employed in Spatial Analyses.'' *Linear Algebra and its Applications* 321, 95–112.
Griffith, D. A. (2002). ''A spatial Filtering Specification for the Auto-Poisson Model.'' *Statistics and Probability Letters* 58, 245–51.
Griffith, D. A. (2003). *Spatial Autocorrelation and Spatial Filtering: Gaining Understanding Through Theory and Scientific Visualization*. Berlin: Springer-Verlag.
Griffith, D. A. (2004a). ''A Spatial Filtering Specification for the Autologistic Model.'' *Environment and Planning A* 36, 1791–811.
Griffith, D. A. (2004b). ''Distributional Properties of Georeferenced Random Variables Based on the Eigenfunction Spatial Filter.'' *Journal of Geographical Systems* 6, 263–88.
Griffith, D. A. (2006). ''Hidden Negative Spatial Autocorrelation.'' *Journal of Geographical Systems* 8, 335–55.
Haining, R. (1990). *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge: University Press.
Hope, A. (1968). ''A simplified Monte Carlo Significance Test Procedure.'' *Journal of the Royal Statistical Society Series B-Methodological* 30, 582–98.
Jacob, B., D. Griffith, and R. Novak. (2008). ''Decomposing Malaria Mosquito Aquatic Habitat Data into Spatial Autocorrelation Eigenvectors in a SAS/GIS Module.'' *Transactions in GIS* 12, 341–64.
Jombart, T., S. Dray, and A. Dufour. (2009). ''Finding Essential Scales of Spatial Variation in Ecological Data: A Multivariate Approach.'' *Ecography* 32, 161–68.
de Jong, P., C. Sprenger, and F. van Veen. (1984). ''On Extreme Values of Moran's I and Geary's c.'' *Geographical Analysis* 16, 17–24.
Kendall, M., and A. Stuart. (1961). *The Advanced Theory of Statistics: Inference and Relationships*. London: Charles Griffin.
Lee, S. (2001). ''Developing a Bivariate Spatial Association Measure: An Integration of Pearson's r and Moran's I.'' *Journal of Geographical Systems* 3, 369–85.
Lee, S. (2009). ''A Generalized Randomization Approach to Local Measures of Spatial Association.'' *Geographical Analysis* 41, 221–48.
Legendre, P. (1993). ''Spatial Autocorrelation: Trouble or New Paradigm?'' *Ecology* 74, 1659–73.
Manly, B. (1997). *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 2nd ed. London: Chapman and Hall.

Moran, P. (1948). ''The Interpretation of Statistical Maps.'' *Journal of the Royal Statistical Society Series B-Methodological* 10, 243–51.

Tiefelsdorf, M. (2002). ''The Saddlepoint Approximation of Moran's I and local Moran's Ii Reference Distributions and Their Numerical Evaluation.'' *Geographical Analysis* 34, 187–206.

Tiefelsdorf, M., and B. Boots. (1995). ''The Exact Distribution of Moran's I.'' *Environment and Planning A* 27, 985–99.

Tiefelsdorf, M., and D. A. Griffith. (2007). ''Semi-Parametric Filtering of Spatial Autocorrelation: The Eigenvector Approach.'' *Environment and Planning A* 39, 1193–221.

Tiefelsdorf, M., D. A. Griffith, and B. Boots. (1999). ''A Variance-Stabilizing Coding Scheme for Spatial Link Matrices.'' *Environment and Planning A* 31, 165–80.

Wartenberg, D. (1985). ''Multivariate Spatial Correlation: A Method for Exploratory Geographical Analysis.'' *Geographical Analysis* 17, 263–83.