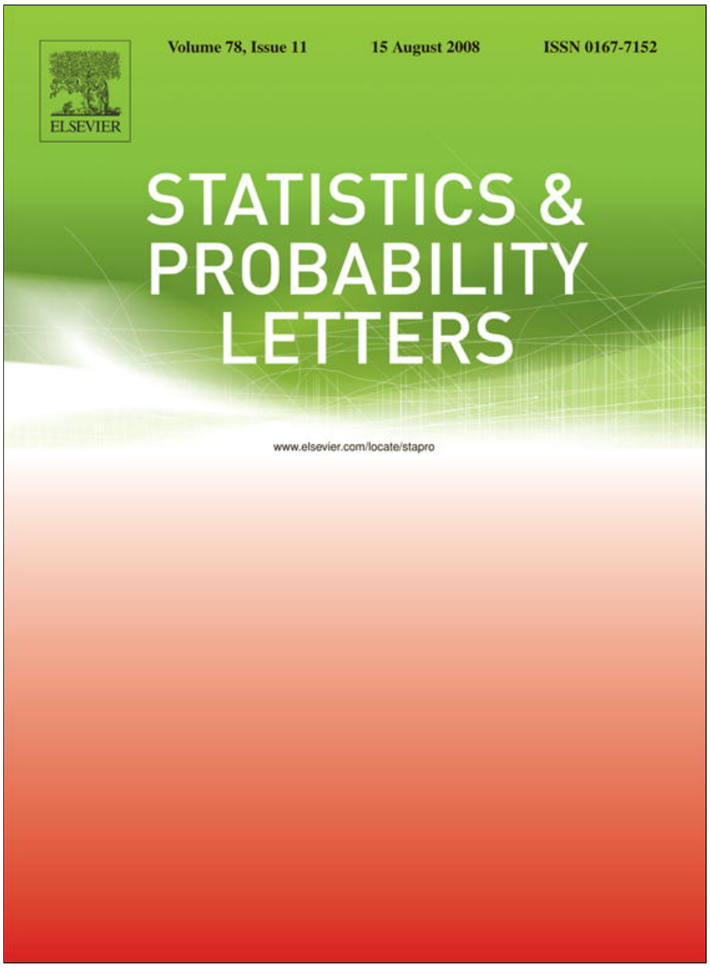
Volume 78, Issue 11
15 August 2008
ISSN 0167-7152
ELSEVIER
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Available online at www.sciencedirect.com
ScienceDirect
Statistics and Probability Letters 78 (2008) 1290–1300
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Codispersion coefficients for spatial and temporal series

Andrew L. Rukhin[a,b,*], Ronny Vallejos[c]

[a] Statistical Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, United States
[b] Department of Mathematics and Statistics, University of Maryland at Baltimore County, Baltimore, MD 21250, United States
[c] Departamento de Estadística-CIMFAV, Universidad de Valparaíso, Casilla 5030, Valparaíso, Chile

Received 14 March 2007; received in revised form 24 July 2007; accepted 6 November 2007
Available online 16 January 2008

Abstract

The paper gives explicit formulas for the first moments of a sample codispersion coefficient defined as a suitably normalized sum of products of increments for time or space sequences. Derived formulas allow for the optimal choice of the lag in several spatial and temporal models and lead to tests of independence and confidence intervals for correlation.

MSC: primary 62H20; secondary 62G60; 62M30

1. Introduction and definition

There has been considerable interest in spatial processes in diverse fields such as economics, environmental science, climatology, ecology, oceanography, soil science, real state markets, image modeling, etc. Many of the developments have been summarized in Ripley (1981), Cliff and Ord (1981), and Cressie (1993).

In times series it is of interest to study how well two time sequences move together. Nonparametric tests for the comovement of time series have appeared in the literature (see Yank and Schreckengost (1981)). Such tests have typically been oriented to providing the independence of two sequences. In the context of spatial sequences, there are several proposed measures of similarity between two spatial processes. A nonparametric technique has been developed by Tjostheim (1978a) to measure the association between spatial variables. This diagnostic was based on the ranks of the observations and on the location coordinates of the measurement points. Later it was generalized by Hubert and Golledge (1982). The problem of testing the association between autocorrelated variables can also be tackled by regression techniques when a relationship between a dependent variable and a set of independent variables is postulated (Cook and Pocock, 1983). The correlation coefficient also has been used as a measure of spatial association and there are several related tests (e.g. Richardson and Clifford (1991)).

In this paper we examine the quantitative assessment of the similarity between two spatial or temporal sequences via the so-called codispersion coefficient, suggested in the sixties by G. Matheron and discussed in Wackernagel

* Corresponding author at: Department of Mathematics and Statistics, University of Maryland at Baltimore County, Baltimore, MD 21250, United States. Fax: +1 410 455 1066.
E-mail address: rukhin@math.umbc.edu (A.L. Rukhin).

0167-7152/$ - see front matter © 2008 Published by Elsevier B.V.
doi:10.1016/j.spl.2007.11.017

(2003). This characteristic is based on the cross-variogram of two spatial series, a tool which can be employed to implement cokriging (Myers, 1991; Ver Hoef and Cressie, 1993) and which has been extensively used in multivariate spatial prediction (Xie et al., 1995).

The codispersion coefficient has been applied to financial and geostatistics data (e.g. Veronese Jr. et al. (2006)), but its sampling properties needed, for example, to test the hypothesis of no correlation, were not studied. Indeed the reviewer of the second edition of Wackernagel's book complains that in the context of hypothesis testing of intrinsic correlation on the basis of change in the codispersion coefficient "no guidance is given for how much of a change is enough to reject" (Haas, 1996).

Here explicit expressions for first order spatial autoregressive models and moving average models are derived in terms of the parameters of the models under several correlation structures between the errors. The asymptotic behavior of the sample codispersion coefficient is studied and its performance is examined via Monte Carlo simulation.

Let us consider two spatial processes, say X and Y, defined on a part of a region $D \subset \mathbb{R}^2$ (or on a rectangular lattice $D \subset \mathbb{Z}^2$.) The cross-variogram provides spatial information between $X(s)$ and $Y(s)$, $s \in D$. For intrinsic stationary processes the cross-variogram for $h = (h_1, h_2)$, $s + h \in D$, is defined (Cressie, 1993, pp. 67) as follows

$$\gamma_{XY}(h) = \mathbb{E}[X(s+h) - X(s)][Y(s+h) - Y(s)]. \tag{1}$$

The normalized version of (1) given by

$$\rho_{XY}(h) = \frac{\gamma_{XY}(h)}{\sqrt{\mathbb{E}[X(s+h) - X(s)]^2 \mathbb{E}[Y(s+h) - Y(s)]^2}} \tag{2}$$

is called *codispersion coefficient*. It is obvious that $|\rho_{XY}(h)| \le 1$.

In the formula for the codispersion coefficient one can employ the notion of the regional (cross-) variogram in the same spirit as in the context of change of support problems (Cressie, 1993, p. 284–286; Wackernagel, 2003, p. 50).

2. Codispersion coefficient for spatial ARMA models

Three broad categories of spatial models that have been studied are the simultaneous autoregressive (SAR) models (Whittle, 1954), the conditional autoregressive (CAR) models (Besag, 1974), and the moving average (MA) models (Haining, 1978). The simultaneous AR model is defined as

$$\Phi(B_1, B_2) X(i, j) = \epsilon(i, j), \tag{3}$$

where

$$\Phi(B_1, B_2) = \sum_k \sum_l \phi(k, l) B_1^k B_2^l, \tag{4}$$

with $B_1 X(i, j) = X(i-1, j)$ and $B_2 X(i, j) = X(i, j-1)$ and $\epsilon(i, j)$ are uncorrelated random variables with $\mathbb{E}(\epsilon(i, j)) = 0$ and $\mathrm{Var}(\epsilon(i, j)) = \sigma^2$.

Tjostheim (1978b) examines a special case of the unilateral AR models where the value at the site (i, j) is a finite autoregression on the values at the sites which are in the lower quadrant of (i, j). In two dimensions this model becomes

$$X(i, j) = \sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \phi(k, l) X(i-k, j-l) + \epsilon(i, j), \tag{5}$$

with $\phi(0, 0) = 0$. This model leads to the Wold-type representation

$$X(i, j) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \phi(k, l) \epsilon(i-k, j-l). \tag{6}$$

In this paper we consider models more general than (6), for which increments of the two-dimensional process $Z(s) = (X(s), Y(s))^{\mathrm{T}}$ admit the following structure,

$$Z(s_1 + h_1, s_2 + h_2) - Z(s_1, s_2) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} A(k, l) \epsilon(s_1 - k, s_2 - l), \tag{7}$$

where $A(k,l) = A_h(k,l)$ are 2×2 matrices defined for all integer k and l, such that

$$\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \|A(k,l)\|^2 < \infty. \tag{8}$$

Here $\|\cdot\|$ denotes any matrix norm, and uncorrelated two-dimensional random vectors $\epsilon(t)$ have zero mean and the covariance matrix Σ. These processes are characterized by the property of absolute continuity of the spectral measure of their stationary increments.

For two processes satisfying (7) we derive the formula for (2).

Proposition 1. *Let*

$$K = E[Z(s+h) - Z(s)][Z(s+h) - Z(s)]^{\mathrm{T}}$$

be the covariance matrix of the vector process for which (7) *holds. Then*

$$K = \sum_{k,l} A(k,l) \Sigma A^{\mathrm{T}}(k,l), \tag{9}$$

and

$$\rho_{XY}(h) = \frac{\kappa_{12}}{\sqrt{\kappa_{11}\kappa_{22}}},$$

where κ_{ij}, $i, j = 1, 2$, *denote elements of* K.

Proof. Because of (7),

$$\begin{aligned} K &= \mathbb{E} \sum_{k_1,k_2,l_1,l_2} A(k_1,l_1)\epsilon(s_1 - k_1, s_2 - l_1)\epsilon(s_2 - k_2, s_2 - l_2)^{\mathrm{T}} A^{\mathrm{T}}(k_2,l_2) \\ &= \sum_{k,l} A(k,l)\mathbb{E}\epsilon(s_1 - k, s_2 - l)\epsilon(s_2 - k, s_2 - l)^{\mathrm{T}} A^{\mathrm{T}}(k,l) = \sum_{k,l} A(k,l)\Sigma A^{\mathrm{T}}(k,l), \end{aligned}$$

and the formula for $\rho_{XY}(h)$ follows from its definition. $\square$

Assume now that both processes can be observed on the increasing part of the positive lattice $\{0 \le s_1 < M, 0 \le s_2 < M\}$ or of R_+^2. For convenience we will assume normality of the error vectors $\epsilon(t) = (\epsilon(t)_1, \epsilon(t)_2)^T$ in (7), although the main results hold other distributions with four first moments.

By the law of large numbers the following convergence in probability holds

$$\frac{1}{M^2} \sum_{0 \le s_i < M - h_i} (Z(s+h) - Z(s))(Z(s+h) - Z(s))^{\mathrm{T}} \to K,$$

where K is given in (9). The sample analogue of $\rho_{XY}(h)$ in Proposition 1, is given by

$$\hat{\rho}_{XY}(h) = \frac{\sum\limits_{0 \le s_i < M - h_i} (X(s_1 + h_1, s_2 + h_2) - X(s_1, s_2))(Y(s_1 + h_1, s_2 + h_2) - Y(s_1, s_2))}{\sqrt{\sum (X(s_1 + h_1, s_2 + h_2) - X(s_1, s_2))^2 \sum (Y(s_1 + h_1, s_2 + h_2) - Y(s_1, s_2))^2}}.$$

In Rukhin (2006) the following result is established.

Theorem 1. *If the process* Z *admits representation* (7) *with diagonal matrices* $A(s) = \mathrm{diag}[\alpha_s, \beta_s]$, $s = (k, \ell)$, *the limiting distribution of* $M[\hat{\rho}_{XY}(h) - \rho]$ *is normal with mean* 0 *and the variance*

$$v^2 = \left(1 - \frac{(\sum \alpha_s \beta_s)^2 \rho^2}{\sum \alpha_s^2 \sum \beta_s^2}\right)^2. \tag{10}$$

The value(s) of h minimizing (10) is optimal in the sense that it gives the smallest limiting variance.

3. Examples

3.1. Moving average models

Consider a moving average process

$$Z(s) = \sum_{t \in \mathbb{T}} C_t \epsilon(s - t), \tag{11}$$

where $\mathbb{T} = \{t : t \in \mathbb{Z}^2, C_t \neq 0\}$, $\epsilon(t)$ are independent random vectors with mean 0 and the covariance matrix Σ. Then

$$Z(s+h) - Z(s) = \sum_{t \in (\mathbb{T}-h) \bigcap \mathbb{T}} (C_{t+h} - C_t)\epsilon(s-t) + \sum_{t \in (\mathbb{T}-h) \bigcap \mathbb{T}^c} C_{t+h}\epsilon(s-t) - \sum_{t \in (\mathbb{T}-h)^c \bigcap \mathbb{T}} C_t \epsilon(s-t),$$

and the representation (7) holds. If all matrices C_t, $t \in \mathbb{T}$, are diagonal, $C_t = \mathrm{diag}[\xi_t, \psi_t]$, the corresponding matrices A_s are also diagonal, and

$$\sum_s \alpha_s \beta_s = 2 \sum_{t \in \mathbb{T}} \xi_t \psi_t - \sum_{t \in (\mathbb{T}-h) \bigcap \mathbb{T}} (\xi_{t+h}\psi_t + \xi_t \psi_{t+h}),$$

$$\sum_s \alpha_s^2 = 2 \left(\sum_{t \in \mathbb{T}} \xi_t^2 - \sum_{t \in (\mathbb{T}-h) \bigcap \mathbb{T}} \xi_{t+h}\xi_t \right),$$

with a similar formula for $\sum_s \beta_s^2$.

If $(\mathbb{T} - h) \bigcap \mathbb{T} = \emptyset$,

$$Z(s+h) - Z(s) = \sum_{t \in (\mathbb{T}-h)^c \bigcap \mathbb{T}} C_t [\epsilon(s+h-t) - \epsilon(s-t)],$$

and the random variables, $\epsilon'(s) = \epsilon(s+h) - \epsilon(s)$, are independent. In this case one can use formula (10) with α_s, β_s replaced by ξ_s, ψ_s.

In the classical model of two one-dimensional time series,

$$X(t) = \phi_1 \upsilon(t-1) + \upsilon(t),$$
$$Y(t) = \phi_2 \zeta(t-1) + \zeta(t),$$

where $|\phi_1|, |\phi_2| < 1$, the errors $\epsilon(t) = (\upsilon(t), \zeta(t))^{\mathrm{T}}$, $t = 1, \ldots, M$, form i.i.d. random vectors with mean 0 and the covariance matrix Σ. In this case, for $h = 1$,

$$Z(s+1) - Z(s) = \epsilon(s+1) + (A - I)\epsilon(s) - A\epsilon(s-1)$$

with $A = \mathrm{diag}[\phi_1, \phi_2]$. One has

$$\sum_s \alpha_s^2 = 1 + \phi_1^2 + (1-\phi_1)^2, \qquad \sum_s \beta_s^2 = 1 + \phi_2^2 + (1-\phi_2)^2,$$
$$\sum_s \alpha_s \beta_s = 1 + \phi_1\phi_2 + (1-\phi_1)(1-\phi_2),$$

so that

$$\rho_{XY} = \frac{\rho(2 - \phi_1 - \phi_2 + 2\phi_1\phi_2)}{2\sqrt{(1-\phi_1+\phi_1^2)(1-\phi_2+\phi_2^2)}}.$$

Formula (10) shows that

$$v^2 = v_1^2 = \left(1 - \frac{\rho^2(2-\phi_1-\phi_2+2\phi_1\phi_2)^2}{4(1-\phi_1+\phi_1^2)(1-\phi_2+\phi_2^2)}\right)^2.$$

When $h \geq 2$,

$$Z(s+h) - Z(s) = \epsilon(s+h) - \epsilon(s) + A[\epsilon(s+h-1) - \epsilon(s-1)],$$

so that one can take $\mathbb{T} = \mathbb{S} = \{0, 1\}$ with new independent errors, $\epsilon(s+h) - \epsilon(s)$, whose covariance matrix is 2Σ. Therefore, for such h,

$$v^2 = v_2^2 = \left(1 - \frac{\rho^2(1+\phi_1\phi_2)^2}{(1+\phi_1^2)(1+\phi_2^2)}\right)^2.$$

It is easy to check that $v_2^2 \geq v_1^2$ with equality when $\phi_1 = \phi_2$.

3.2. First order autoregressive and related models

Basu and Reinsel (1993) investigated the correlation structure of a general first order autoregressive process of the form,

$$X(i,j) = \xi_1 X(i-1,j) + \xi_2 X(i,j-1) + \xi_3 X(i-1,j-1) + \epsilon_1(i,j), \tag{12}$$

by deriving its stationary representation as

$$X(i,j) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{r=0}^{\infty} \frac{(k+l+r)!}{k!l!r!}\xi_1^k \xi_2^l \xi_3^r \epsilon_1(i-k-r, j-l-r)$$

(which is different from (7).) They showed that the conditions, $|\xi_i| < 1$, $1 - \xi_2^2 > |\xi_1 + \xi_2\xi_3|$, $(1 + \xi_1^2 - \xi_2^2 - \xi_3)^2 > 4(\xi_1 + \xi_2\xi_3)^2$, guarantee convergence of the triple series.

In the particular case $\xi_3 = -\xi_1\xi_2$, this multiplicative (separable) process simplifies to

$$X(i,j) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \xi_1^k \xi_2^l \epsilon_1(i-k, j-l).$$

This model has been investigated by Martin (1979, 1990) who argued that it has many practical uses. If a two-dimensional process has the corresponding form,

$$Z(i,j) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} C_1^k C_2^l \epsilon(i-k, j-l)$$

with diagonal matrices $C_\ell = \mathrm{diag}[\xi_\ell, \psi_\ell]$, $\ell = 1, 2$ and mean zero i.i.d. innovations $\epsilon(i,j)$ whose covariance matrix is given by Σ, then with $h = (h_1, h_2)$, $h_1, h_2 \geq 0$,

$$\rho_{XY} = \frac{\rho[1 - (\xi_1^{h_1}\xi_2^{h_2} + \psi_1^{h_1}\psi_2^{h_2})/2]\sqrt{(1-\xi_1^2)(1-\xi_2^2)(1-\psi_1^2)(1-\psi_2^2)}}{(1-\xi_1\psi_1)(1-\xi_2\psi_2)\sqrt{(1-\xi_1^{h_1}\xi_2^{h_2})(1-\psi_1^{h_1}\psi_2^{h_2})}}.$$

This formula demonstrates that for positive ξ_ℓ, ψ_ℓ, $\ell = 1, 2$, the optimal value of h minimizing (10) is $h = (1, 0)$, if

$$\frac{1 - (\xi_1 + \psi_1)/2}{\sqrt{(1-\xi_1)(1-\psi_1)}} > \frac{1 - (\xi_2 + \psi_2)/2}{\sqrt{(1-\xi_2)(1-\psi_2)}}$$

or $h = (0, 1)$, otherwise.

We also look at the models (12), when $\xi_3 = 0$, which were introduced by Whittle (1954). Then

$$X(i,j) = \xi_1 X(i-1,j) + \xi_2 X(i,j-1) + \epsilon_1(i,j), \tag{13}$$

$$Y(i,j) = \psi_1 Y(i-1,j) + \psi_2 Y(i,j-1) + \epsilon_2(i,j), \tag{14}$$

where $\epsilon(t) = (\epsilon_1(t), \epsilon_2(t))^{\mathrm{T}}$ form i.i.d. random vectors with mean 0 and covariance matrix Σ. Under conditions mentioned above on the parameters, $Z(i,j) = (X(i,j), Y(i,j))^{\mathrm{T}}$ has the convergent representation (11) with

$\mathbb{T} = \mathbb{Z}_+^2$, and

$$C_t = C_{kl} = \binom{k+l}{k} \begin{pmatrix} \xi_1^k \xi_2^l & 0 \\ 0 & \psi_1^k \psi_2^l \end{pmatrix}.$$

By setting for $h = (h_1, h_2)$ and x, y such that $1 + x^4 + y^4 - 2x^2 - 2y^2 - 2x^2y^2 > 0$,

$$D_h(x, y) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(k+l)!(k+l+h_1+h_2)!}{k!l!(k+h_1)!(l+h_2)!} x^{2k+h_1} y^{2l+h_2},$$

we get

$$\sum_s \alpha_s \beta_s = 2D_0(\sqrt{\xi_1\xi_2}, \sqrt{\psi_1\psi_2}) - D_h(\sqrt{\xi_1\xi_2}, \sqrt{\psi_1\psi_2}) \left[\left(\frac{\xi_1}{\xi_2}\right)^{h_1/2} \left(\frac{\psi_1}{\psi_2}\right)^{h_2/2} + \left(\frac{\xi_2}{\xi_1}\right)^{h_1/2} \left(\frac{\psi_2}{\psi_1}\right)^{h_2/2} \right],$$
$$\sum_s \alpha_s^2 = 2[D_0(\xi_1, \psi_1) - D_h(\xi_1, \psi_1)], \qquad \sum_s \beta_s^2 = 2[D_0(\xi_2, \psi_2) - D_h(\xi_2, \psi_2)].$$

Numerical evaluation of the codispersion coefficient is facilitated by noticing that

$$D_h(x, y) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \frac{\mathrm{e}^{-\mathrm{i}(\omega_1 h_1 + \omega_2 h_2)} \, \mathrm{d}\omega_1 \, \mathrm{d}\omega_2}{|1 - \mathrm{e}^{-\mathrm{i}\omega_1} x - \mathrm{e}^{-\mathrm{i}\omega_2} y|^2}.$$

In particular,

$$\begin{aligned} D_0(x, y) &= \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\omega_1 \, \mathrm{d}\omega_2}{|1 - \mathrm{e}^{-\mathrm{i}\omega_1} x - \mathrm{e}^{-\mathrm{i}\omega_2} y|^2} \\ &= \frac{1}{\sqrt{1 + x^4 + y^4 - 2x^2 - 2y^2 - 2x^2y^2}}, \end{aligned}$$
$$D_{(h,0)}(x, y) = D_0(x, y) \left[\frac{1 + x^4 - y^4 - \sqrt{1 + x^4 + y^4 - 2x^2 - 2y^2 - 2x^2y^2}}{2x} \right]^h,$$

and a similar formula holds for $D_{(0,h)}(x, y)$.

In the one dimensional case

$$X(t) = \phi_1 X(t-1) + \epsilon_1(t), \tag{15}$$
$$Y(t) = \phi_2 Y(t-1) + \epsilon_2(t), \tag{16}$$

are two classical AR(1) processes. We assume that $|\phi_1|, |\phi_2| < 1$, and the error vectors $\epsilon(t) = (\epsilon_1(t), \epsilon_2(t))^T$, $t = 1, \ldots, N$, are i.i.d. random vectors with mean 0 and covariance matrix Σ, so that $Z(t) = (X(t), Y(t))^{\mathrm{T}}$ has the form, $Z(t) = \sum_{k=0}^{\infty} A^k \epsilon(t-k)$ with a diagonal matrix $A = \mathrm{diag}[\phi_1, \phi_2]$.

The mean square successive difference, $\sum_{t=1}^{N-1} [X(t+1) - X(t)]^2/(N-1)$, was introduced by Von Neumann in the forties. It is known that the variance of this statistic when both sequences $X(t)$ and $Y(t)$ are i.i.d. normal random variables, $\phi_1 = \phi_2 = 0$, is $8N\mu_{20}^2/(N-1)^2$, which agrees with our formulas.

For $h \geq 1$, in the representation (7) $A_k = A^{h+k}$, $-h \leq k \leq -1$, $A_k = (A^h - I)A^k$, $k \geq 0$. In particular, when $h = 1$,

$$\rho_{XY}(1) = \frac{\rho(2 - \phi_1 - \phi_2)\sqrt{(1+\phi_1)(1+\phi_2)}}{2(1 - \phi_1\phi_2)}$$

and when $h = 2$,

$$\rho_{XY}(2) = \frac{\rho(2 - \phi_1^2 - \phi_2^2)}{2(1 - \phi_1\phi_2)}.$$

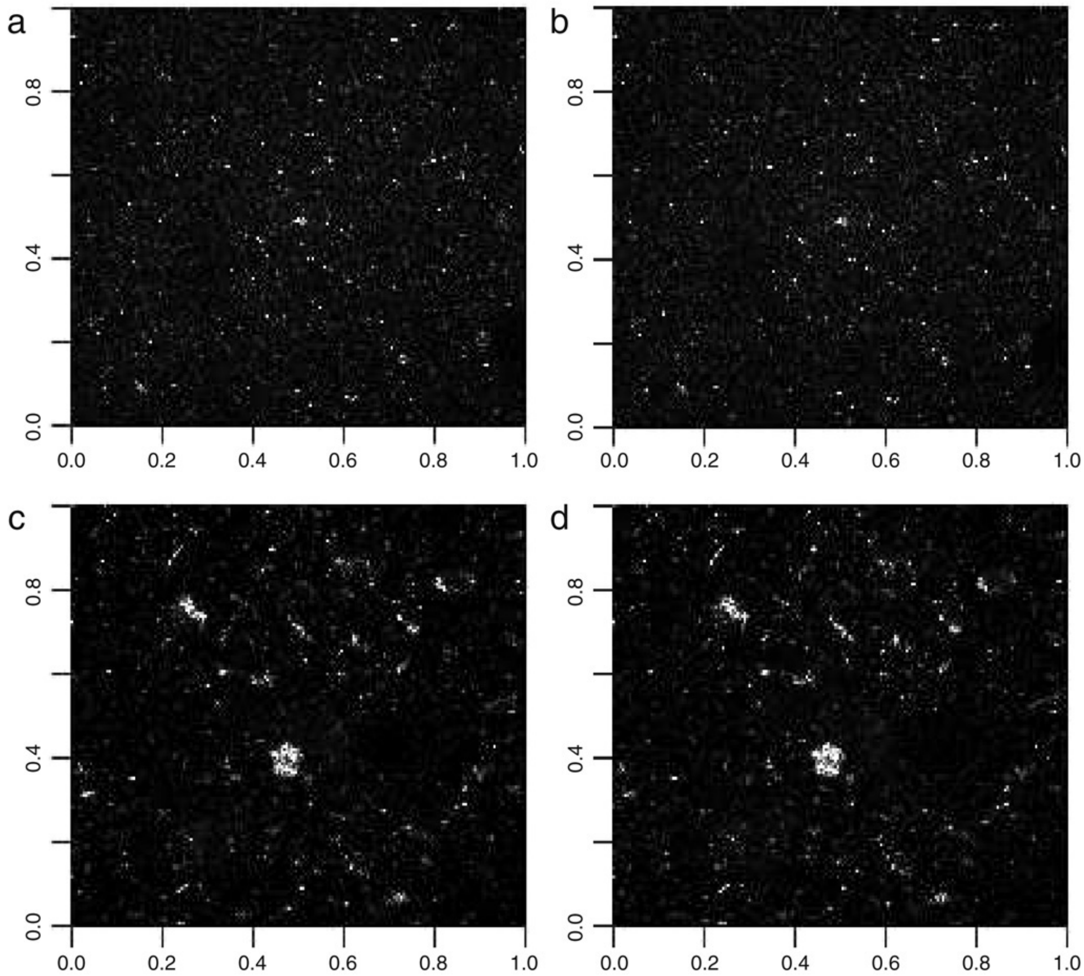

Fig. 1. Dispersion of nanotubes.

The limiting distribution of $\sqrt{N}[\hat{\rho} - \rho(h)]$ is normal with mean 0. Its variance according to (10) has the form

$$v_h^2 = \left[1 - \frac{\rho^2\left[1 - (\phi_1^h + \phi_2^h)/2\right]^2 (1 - \phi_1^2)(1 - \phi_2^2)}{(1 - \phi_1\phi_2)^2(1 - \phi_1^h)(1 - \phi_2^h)}\right]^2. \tag{17}$$

The comparison of the corresponding estimators (2) of ρ, $\hat{\rho}_{XY}(h)$, can be performed as before. The optimal value of h corresponds to the largest ratio, $(2 - \phi_1^h - \phi_2^h)^2/[(1 - \phi_1^h)(1 - \phi_2^h)]$. In particular, $\mathrm{Var}(\hat{\rho}_{XY}(1)) < \mathrm{Var}(\hat{\rho}_{XY}(2))$ if and only if $(\phi_1 + \phi_2)^2 < (1 + \phi_1)(1 + \phi_2)$.

3.3. Flammability of carbon nanotubes

To illustrate a practical application of the codispersion coefficient an example related to flammability properties of polymers is presented here. It is known that one weak aspect of polymers is that they are combustible under certain conditions. The flame retardant property of clay-polymer nanocomposites has been demonstrated to improve physical and flammability properties of polymers by Kashiwagi et al. (2005). To investigate the effects of the dispersion and concentration of one of these retardants, the so-called single-walled carbon nanotube on the flammability of polymers, the distribution of this nanotube was examined by optical microscopy. This distribution is believed to depend mainly on the distance from the top surface to the location in the polymer matrix (polymethyl methacrylate). In the study it was important to characterize the dispersion and the concentration of the flame retardant after an image of the polymer. In other words it was desirable to define a metric for closeness of two sample polymers in terms of their image characteristics.

Table 1
MSE for models 1–4

	Model 1	Model 2	Model 3	Model 4
MSE	680.1253	685.2043	593.4743	572.1097

The available data set was represented as four 512×512 images plotted in Fig. 1. These images taken at the National Institute of Standards and Technology (NIST) were kindly provided by T. Kashiwagi. One has $\hat{\rho}_{ab}(1,0) = 0.804$, $\hat{\rho}_{ab}(1,1) = 0.841$, $\hat{\rho}_{ab}(0,1) = 0.807$. The same patterns are observed for the codispersion coefficient between images (c) and (d): $\hat{\rho}_{cd}(1,0) = 0.815$, $\hat{\rho}_{cd}(1,1) = 0.865$, $\hat{\rho}_{cd}(0,1) = 0.794$. However $\hat{\rho}_{ac}(1,0) = 0.002$, $\hat{\rho}_{ac}(1,1) = 0.003$, $\hat{\rho}_{ac}(1,0) = 0.001$. Moreover when both lags were used in the computation of the codispersion coefficient, we have found even stronger evidence of similarity, for example $\hat{\rho}_{ac}(2,2) = 0.875$. The codispersion coefficient is able to capture the association between images (a) and (b), and the dissimilarity between images (a) and (c).

To construct a confidence interval for ρ_{ab}, we need first to find a suitable model for images (a) and (b). Since the spatial autoregressive model has been satisfactorily used in many applications the following AR models

Model 1: $X(i,j) = \phi_1 X(i-1,j) + \epsilon(i,j)$,
Model 2: $X(i,j) = \phi_2 X(i,j-1) + \epsilon(i,j)$,
Model 3: $X(i,j) = \phi_1 X(i-1,j) + \phi_2 X(i,j-1) + \epsilon(i,j)$,
Model 4: $X(i,j) = \phi_1 X(i-1,j) + \phi_2 X(i,j-1) + \phi_3(i-1,j-1) + \epsilon(i,j)$

were considered. The least squares estimator was used to estimate the parameters of models 1–4. For each model the mean squared error (MSE) was computed over the observations located in the 30×30 region surrounding the 452×452 region. Table 1 shows that models 3 and 4 have smaller MSE values. In order to select the best model, we plotted generated images from models 3 and 4 to observe similarities between the simulated images and the original image (a), and concluded that model 3 seemed to be the most appropriate.

Estimators of $\hat{\rho}_{ab}$, $s.d(\hat{\rho_{ab}})$, a 95% confidence interval for ρ_{ab} and the length of the confidence interval were evaluated using the estimated parameters.

4. Monte Carlo simulations

In this section we report results of several Monte Carlo simulation studies.

Consider X and Y as in (13) and (14). The codispersion coefficient can be used as an estimator of the correlation between X and Y. To define Tjostheim's coefficient (1978a), let (z,w) denote spatial location, $F = F(z,w)$ and $G = G(z,w)$ be two spatial (regional) variables. Assume that n observations of F and G have been taken at locations $(z_1,w_1),\ldots,(z_n,w_n)$. Let $R_F(z_i,w_i)$ ($R_G(z_i,w_i)$) denote the rank of $F(G)$ at the point (z_i,w_i), Define $X_F(i)$ the z coordinate corresponding to the rank i of F with $Y_F(i)$, $X_G(i)$ and $Y_G(i)$ similarly defined. The Tjostheim coefficient is given by

$$\hat{\rho}_T = \frac{\sum\left[(X_F(i)-\overline{X}_F)(X_G(i)-\overline{X}_G)(Y_F(i)-\overline{Y}_F)(Y_G(i)-\overline{Y}_G)\right]}{\left(\sum\left[(X_F(i)-\overline{X}_F)^2+(Y_F(i)-\overline{Y}_F)^2\right]\sum\left[(X_G(i)-\overline{X}_G)^2+(Y_G(i)-\overline{Y}_G)^2\right]\right)^{1/2}},$$

where $\overline{X}_F = \sum X_F(i)/n$, and similarly for $\overline{X}_G$, $\overline{Y}_F$ and $\overline{Y}_G$.

In the Monte Carlo simulation, two sets of parameters and six values for ρ were used: $\xi_1 = 0.3$, $\xi_2 = 0.2$, $\psi_1 = 0.3$, $\psi_2 = 0.2$; $\xi_1 = 0.3$, $\xi_2 = 0.2$, $\psi_1 = 0.6$, $\psi_2 = 0.1$, $\rho = 0, 0.1, 0.3, 0.5, 0.7$, and 0.9. One thousand replicates were obtained for each case of the two sets of parameters. The average values and the standard deviations of the estimates over these replicates can be found in Table 2 and in Vallejos (2006). The simulations were performed for the grid size 30×30. One has $\mathbb{E}[\hat{\rho}_{XY}] = c\rho$, where the multiple c can be determined. Hence all simulation averages reported in Table 2 were corrected by the factor $\frac{1}{c}$.

From Table 2 we see that the codispersion coefficient and the correlation coefficient are comparable and less biased than Tjostheim's coefficient. The bias of Tjostheim's coefficient increases as ρ increases; the standard deviation of our coefficient is smaller than the standard deviation of Tjostheim's coefficient when $\rho \geq 0.5$. The standard deviations of the codispersion coefficient and of the correlation coefficient decrease as ρ increases. All estimators underestimate the

Table 2
The means and standard deviations of $\hat{\rho}_{XY}$, r and $\hat{\rho}_T$ for $\rho = 0, 0.1$ and 0.3

ρ	ξ_1	ξ_2	ψ_1	ψ_2	h_1	h_2	$\hat{\rho}_{XY}$	$s.d.(\hat{\rho}_{XY})$	r	$s.d.(r)$	$\hat{\rho}_T$	$s.d.(\hat{\rho}_T)$
0	0.3	0.2	0.3	0.2	1	0	−0.0006	0.0384	0.0002	0.0389	0.0028	0.0391
0	0.3	0.2	0.3	0.2	0	1	0.0008	0.0416	0.0004	0.0388	0.0027	0.0391
0	0.3	0.2	0.3	0.2	1	1	0.0012	0.0405	0.0006	0.0394	0.0028	0.0391
0	0.3	0.2	0.6	0.1	1	0	−0.0008	0.0365	0.0003	0.0421	0.0054	0.0397
0	0.3	0.2	0.6	0.1	0	1	0.0006	0.0454	0.0002	0.0421	0.0054	0.0397
0	0.3	0.2	0.6	0.1	1	1	0.0012	0.0434	0.0005	0.0426	0.0054	0.0397
0.1	0.3	0.2	0.3	0.2	1	0	0.0990	0.0411	0.1008	0.0423	0.0466	0.0406
0.1	0.3	0.2	0.3	0.2	0	1	0.1009	0.0424	0.1006	0.0417	0.0466	0.0406
0.1	0.3	0.2	0.3	0.2	1	1	0.0992	0.0443	0.1010	0.0427	0.0466	0.0406
0.1	0.3	0.2	0.6	0.1	1	0	0.0955	0.0382	0.0941	0.0449	0.0453	0.0421
0.1	0.3	0.2	0.6	0.1	0	1	0.0941	0.0462	0.0941	0.0446	0.0453	0.0421
0.1	0.3	0.2	0.6	0.1	1	1	0.0929	0.0469	0.0944	0.0389	0.0453	0.0421
0.3	0.3	0.2	0.3	0.2	1	0	0.2989	0.0378	0.3006	0.0389	0.1343	0.1343
0.3	0.3	0.2	0.3	0.2	0	1	0.3006	0.0390	0.3004	0.0383	0.1343	0.0405
0.3	0.3	0.2	0.3	0.2	1	1	0.2990	0.0408	0.3007	0.0393	0.1343	0.0405
0.3	0.3	0.2	0.6	0.1	1	0	0.2878	0.0352	0.2793	0.0416	0.1246	0.0401
0.3	0.3	0.2	0.6	0.1	0	1	0.2806	0.0428	0.2792	0.0413	0.1246	0.0401
0.3	0.3	0.2	0.6	0.1	1	1	0.2796	0.0434	0.2795	0.0421	0.1246	0.0401

true value. In general, the performance of codispersion and correlation coefficients is better than that of Tjostheim's coefficient for $\rho \neq 0$.

4.1. *Simulation of spatial moving average models*

Here we consider the models

$$X(i,j) = \epsilon_1(i,j) - \theta_1\epsilon_1(i-1,j) - \theta_2\epsilon_1(i,j-1), \tag{18}$$

$$Y(i,j) = \epsilon_2(i,j) - \eta_1\epsilon_2(i-1,j) - \eta_2\epsilon_2(i,j-1). \tag{19}$$

It follows from (2) that

$$\rho_{XY}(h_1,h_2) = \begin{cases} \dfrac{\rho\,[2(\theta_1\eta_1+\theta_2\eta_2)-\theta_1-\eta_1]}{2\sqrt{(1+\theta_1^2+\theta_2^2+\theta_1)(1+\eta_1^2+\eta_2^2+\eta_1)}}, & h_1=1, h_2=0, \\ \dfrac{\rho\,[2(\theta_1\eta_1+\theta_2\eta_2)-\theta_2-\eta_2]}{2\sqrt{(1+\theta_1^2+\theta_2^2+\theta_2)(1+\eta_1^2+\eta_2^2+\eta_2)}}, & h_1=0, h_2=1, \\ \dfrac{\rho(\theta_1\eta_1+\theta_2\eta_2)}{\sqrt{(1+\theta_1^2+\theta_2^2)(1+\eta_1^2+\eta_2^2)}}, & h_1\geq 1, h_2\geq 1. \end{cases}$$

We carried out a simulation study to observe the performance of $\hat{\rho}_{XY}$, r and $\hat{\rho}_T$ for several values of ρ, and three sets of parameters: (i) $\theta_1 = 0.3, \theta_2 = 0.5, \eta_1 = 0.3, \eta_2 = 0.5$, (ii) $\theta_1 = 0.3, \theta_2 = 0.5, \eta_1 = 0.2, \eta_2 = 0.4$, and (iii) $\theta_1 = 0.3, \theta_2 = 0.5, \eta_1 = 0.1, \eta_2 = 0.6$. The simulation results show patterns similar to the autoregressive case discussed in the previous section. Tjostheim's coefficient is strongly biased; the performance of the codispersion coefficient is similar to that of the correlation coefficient. However the sample codispersion coefficient appears to do a good job at estimating its theoretical value. In all cases the standard deviation of the codispersion coefficient is smaller than the standard deviation of the correlation coefficient. For large values of ρ ($\rho \geq 0.7$) the standard deviation of the codispersion coefficient is the smallest. When $\rho = 0.9$ the standard deviation of the codispersion coefficient reported for smaller values of ρ exceeds the standard deviation of the codispersion coefficient. In all cases the coefficients are biased downward.

4.2. Coverage Probability

To construct an approximate confidence interval for ρ_{XY} we use Theorem 1,

$$\hat{\rho}_{XY} \pm z_{\alpha/2} \frac{v_h}{\sqrt{M}},$$

where $z_{\alpha/2}$ is the $(1-\alpha/2)$-th percentile of the standard normal distribution.

To study the performance of this interval we carried out a simulation study for several values of ρ with 10 000 simulation runs from models (15) and (16). The estimated coverage probability (ECP) was compared with the theoretical (known) coverage probability written as CP. For $\rho \le 0.3$, the ECP is between 0.88 and 0.94. However for larger values of ρ, the ECP is between 0.08 and 0.96. The worse case found corresponds to $\phi_1 = -0.5$ and $\phi_2 = 0.3$, and $\rho = 0.7$. Indeed the codispersion coefficient is biased for large values of ρ, especially when $|\phi_1 - \phi_2|$ is large. For the pairs of parameters $\phi_1 = 0.8, \phi_2 = 0.5$ and $\phi_1 = 0.4, \phi_2 = 0.2$, we have $0.90 < \text{ECP} < 0.96$. Hence, in these cases all estimated coverage probabilities look very accurate. We did not perceive any clear pattern when M increases; $M = 50$ was enough to get accurate estimates.

The same procedure was repeated for the spatial models (13) and (14). We studied the coverage probability of a confidence interval for ρ_{XY} of the form

$$\hat{\rho}_{XY} \pm z_{\alpha/2} \frac{v_h}{M}, \tag{20}$$

where v_h is the variance of the sample codispersion coefficient. We observed similar behavior of the ECP.

5. Hypothesis testing

In this section we study the hypothesis testing problem

$$H_0 : \rho_{XY} = 0 \quad \text{versus} \quad H_A : \rho_{XY} > 0. \tag{21}$$

This problem is relevant because in the normal case it leads to independence between X and Y. Statistic $Q = v^{-1}(M(\hat{\rho}_{XY} - \rho_{XY}))$ is asymptotically standard normal because of Theorem 1. In particular, under H_0, $Q_0 = v^{-1}M\hat{\rho}_{XY}$ is approximately standard normal, and one can reject H_0 if $Q_0 > z_\alpha$.

Let r^2 be the sample correlation coefficient computed over M^2 pixels. Its distribution is well known for $i.i.d.$ random variables X and Y. We denote the power function of the corresponding test as β_ρ. To compare the power functions β and β_ρ of two tests of $H_0 : \rho_{XY} = \rho_0 > 0$, we carried out a simulation study, which showed that $\beta > \beta_\rho$ for $\rho_0 > 0.28$. The curves intersect at a small ρ_0, and this pattern is common in all studied cases. When $\alpha = 0.05$, $\beta > \beta_\rho$ for $\rho_0 > 0.15$. When M increases, the power functions grow faster. Thus, one can conclude that the β is uniformly greater than β_ρ for $\rho_0 > 0.25$.

We have observed before that the performance of the codispersion coefficient is comparable to Tjostheim's coefficient for ρ close to zero. One important disadvantage of Tjostheim's coefficient is that under independence its asymptotic bias and variance depend in a rather complicated manner on the coordinates transformation. This makes it difficult to examine the dependence of the variance on specific population coordinate layouts.

To explore the power of tests for spatial models (13) and (14) in the simulation study we took $h_1 = 1, h_2 = 0$, $\phi_1 = \psi_1 = 0.3$, $\phi_2 = \psi_2 = 0.2$, so that $\sum_s \alpha_s = 0.0207$, $\sum_s \alpha_s^2 = 0.0549$, $\sum \beta_s^2 = 0.7305$. Then for $\rho = 0.05$, $v = 0.9971$, for $\rho = 0.10$, $v = 0.9904$ and for $\rho = 0.15$, $v = 0.9784$. In all simulations β dominated β_ρ for small values of ρ_0.

6. Conclusions and acknowledgments

This paper gives explicit formulas for the first moments of a sample codispersion coefficient which lead to the optimal choice of the lag in several spatial and temporal models. Our results provide tests of independence and confidence intervals for the correlation coefficient in several moving average and autoregression models.

The first author (whose research was supported by NSA grant no. H98230-06-1-0068) wants to thank the editor R. Johnson for his constant support.

References

Basu, S., Reinsel, G., 1993. Properties of the spatial unilateral first-order ARMA model. Adv. Appl. Prob. 25, 631–648.

Besag, J.E., 1974. Spatial interaction and the statistical analysis of lattice systems. J. R. Statist. Soc. B 36, 192–209.

Cliff, A., Ord, J., 1981. Spatial Processes: Models and Applications. Pion Ltd., London.

Cook, D.G., Pocock, S.J., 1983. Multiple regression in geographical mortality studies, with allowance for spatially correlated errors. Biometrics 39, 361–371.

Cressie, N., 1993. Statistics for Spatial Data. Wiley, New York.

Haas, T., 1996. Review of multivariate geostatistics: An introduction with applications by H. Wackernagel. J. Amer. Statist. Assoc. 91, 1375–1376.

Haining, R.P., 1978. Estimating spatial-interaction models. Env. Planning A 10, 305–320.

Hubert, L.J., Golledge, R.G., 1982. Measuring association between spatially defined variables: Tjostheim index and some extensions. Geogr. Anal. 4, 273–278.

Kashiwagi, T., Du, F., Winey, K.I., Groth, K.M., Shields, J.R., Bellayer, S.P., Kim, S., Douglas, J.F., 2005. Flammability properties of polymer nanocomposites with single-walled carbon nanotubes: Effects of nanotube dispersion and concentration. Polymer 46, 471–481.

Martin, R.J., 1979. A subclass of lattice processes applied to a problem in planar sampling. Biometrika 66, 209–217.

Martin, R.J., 1990. The use of time series models and methods in the analysis of agricultural trials. Commun. Statist. Theory Methods A19, 55–81.

Myers, D., 1991. Matrix formulation of co-kriging. Math. Geol. 14, 249–257.

Richardson, S., Clifford, P., 1991. Testing association between spatial processes. In: Possolo, A. (Ed.), Spatial Statistics and Imaging. In: IMS Lecture Notes, vol. 20. pp. 295–308.

Ripley, B., 1981. Spatial Statistics. Wiley, New York.

Rukhin, A., 2006. Association characteristics in spatial statistics. Math. Methods Statist. 15, 309–326.

Tjostheim, D., 1978a. A measure of association for spatial variables. Biometrika 65, 109–114.

Tjostheim, D., 1978b. Statistical spatial series modelling. Adv. Appl. Prob. 10, 130–154.

Vallejos, R., 2006. A similarity coefficient for spatial and temporal sequences. Ph.D. Thesis, University of Maryland Baltimore County.

Ver Hoef, J., Cressie, N., 1993. Multivariate spatial prediction. Math. Geol. 25, 219–240.

Veronese Jr., V., Carvalho, M.P., Dafonte, J., Freddi, O.S., Vidal Vasquez, E., Ingaramo, O.E., 2006. Spatial variability of soil water content and mechanical resistance of Brazilian ferrosoil. Soil Till. Res. 85, 166–177.

Wackernagel, H., 2003. Multivariate Geostatistics: An Introduction with Applications, 3rd ed. Springer.

Whittle, P., 1954. On stationary processes in the plane. Biometrika 41, 434–449.

Xie, T., Myers, D.E., Long, A.E., 1995. Fitting matrix- valued variogram models by simultaneous diagonalization (Part I: Theory). Math. Geol. 27, 867–876.

Yank, M.C.K., Schreckengost, J.F., 1981. Difference sign test for comovements between two time series. Commun. Statist. Theory Methods A10, 355–369.