Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE @ DIRECT
Journal of Statistical Planning and Inference 131 (2005) 19–39

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Testing axial symmetry and separability of lattice processes

L. Scaccia[a], R.J. Martin[b,*]

[a]Dipartimento di Scienze Statistiche, Università di Perugia, Via A. Pascoli, Perugia 06100, Italy
[b]Wirksworth, Derbyshire DE4 4EB, UK

Received 15 March 2003; accepted 28 January 2004

Abstract

Data collected on a rectangular lattice are common in many areas, and models used often make simplifying assumptions. These assumptions include axial symmetry in the spatial process and separability. Some different methods for testing axial symmetry and separability are considered. Using the sample periodogram is shown to provide some simple satisfactory tests of both hypotheses, but tests for separability given axial symmetry have low power for small lattices.

Keywords: Autoregressive process; Axial symmetry; Doubly-geometric process; Lattice process; Separability; Spatial process

1. Introduction

Models for data collected on a rectangular lattice are widely used in applications such as field trials, geostatistics, remotely sensed data and image analysis (see, e.g., Cressie, 1993, Chapter 6). Most models used assume axial (or reflection) symmetry.

A continuous variable, observed at spatial locations on a regular two-dimensional lattice, can be considered a very close analogue to a time series observed at equally spaced time points. Although there are some differences, these do not preclude the use of time series techniques in modelling spatial data. Many of the problems of two-dimensional modelling can be overcome by using separable processes, and data analytic tools of time series analysis could be used to analyze and model lattice data —see Section 2.3.

* Corresponding author. Department of Probability and Statistics, University of Sheffield, Sheffield S3 7RH, UK. Tel.: 0441142223751; fax: 00441142223759.

E-mail address: r.j.martin@sheffield.ac.uk (R.J. Martin).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2004.01.003

However, few methods for testing for separability have been proposed. Modjeska and Rawlings (1983) discussed a two-stage procedure using the sample correlations, but did not develop it into a test. Shitan and Brockwell (1995) proposed a model-based asymptotic test for separability. This assumed a separable $AR(p_1) \cdot AR(1)$ model, and formulated the lattice data as a n_1-variate time series of length n_2 (for an $n_1 \times n_2$ lattice). The test for separability then reduces to testing whether the coefficient matrices are diagonal. Even for small lattices, the test requires very large matrices to be manipulated, and very large lattices appear to be needed for the size of the test to be correct. Guo and Billard (1998) also suggest a model-based test, based on the comparison, using the Wald test, between the fit to the data of an $AR(1) \cdot AR(1)$ process and the more general unilateral autoregressive Pickard process, and it actually tests for axial symmetry and separability together.

In this paper we develop tests that can be applied to a larger range of processes, and test separability in two stages: testing axial symmetry (which is a necessary condition for a process to be separable) first, and then, if the hypothesis of axial symmetry is not rejected, testing separability. Since most models used in practice assume axial symmetry, testing for it is also useful in its own right. The tests for axial symmetry developed here could be used, therefore, in an initial investigation of spatial data, in order to check for axial symmetry, or to test for separability in two stages.

The paper is structured as follows. In Section 2 we give some basic notation and definitions, together with some further motivation for this work. Section 3 is concerned with methods for testing axial symmetry and separability based on the sample periodogram. Section 4 presents some examples of spatial processes, which are then used in Section 5 for a simulation study to compare the tests proposed in Section 3. We conclude in Section 6 with a general discussion and outline possible future developments.

2. Notation and definitions

In this section we give the notation, some basic definitions used in time series and their extensions to spatial processes, and, finally, we define the hypotheses of axial symmetry and separability.

2.1. Spatial lattice processes

It is assumed that data are observed on an $n_1 \times n_2$ rectangular lattice, where rows are indexed by $i_1 = 1, \ldots, n_1$ and columns by $i_2 = 1, \ldots, n_2$. Row and column lags are represented by g_1 and g_2 respectively, with $g_j = -(n_j - 1), \ldots, -1, 0, 1, \ldots, (n_j - 1)$, for $j = 1, 2$. It is assumed that the sites are ordered lexicographically, so that (i_1, i_2) precedes $(i_1, i_2 + 1)$ for $i_2 < n_2$, and (i_1, n_2) precedes $(i_1 + 1, 1)$.

Data can be considered as a realization of random variables Y_{i_1,i_2}. Assuming that the vector Y contains the Y_{i_1,i_2} in site order, we can consider, as a model:

$$\mathrm{E}(Y) = A\vartheta \quad \text{and} \quad \mathrm{var}(Y) = V\sigma^2, \tag{1}$$

where A is an $n \times p$ matrix (with $n = n_1 n_2$), ϑ is a p-vector of parameters and the dispersion matrix $V = V(v)$ is an $n \times n$ positive-definite matrix depending on the q-vector of parameters v. We assume here that the process is Gaussian. We also assume that the study region considered is homogeneous or that the data have had obvious trends or other effects removed to permit us to make the assumption of stationarity with $\mathrm{E}(Y) = \mu 1_n$ so that $Y \sim \mathrm{N}(\mu 1_n, V\sigma^2)$, where 1_n is a column vector of ones.

Definition 1. For a second-order stationary spatial process in two dimensions, the covariance function (or covariogram) at lags g_1 and g_2 is

$$C(g_1, g_2) = \mathrm{cov}(Y_{i_1,i_2}, Y_{i_1+g_1,i_2+g_2}) = E[(Y_{i_1,i_2} - \mu)(Y_{i_1+g_1,i_2+g_2} - \mu)]. \tag{2}$$

Then the theoretical autocorrelation at lags g_1, g_2 is $\rho(g_1, g_2) = C(g_1, g_2)/C(0,0)$, where $C(0,0) = \sigma^2$, and $C(g_1, g_2) = C(-g_1, -g_2)\,\forall g_1,\ g_2$. Also the semi-variogram, defined as $S(g_1, g_2) = E[(Y_{i_1+g_1,i_2+g_2} - Y_{i_1,i_2})^2]/2$ can be expressed, under stationarity, in terms of the covariance function as $S(g_1, g_2) = C(0,0) - C(g_1, g_2)$.

We assume that the spectral density function exists—a sufficient condition is that $\sum_{g_1=-\infty}^{\infty} \sum_{g_2=-\infty}^{\infty} |C(g_1, g_2)| < \infty$.

Definition 2. For a second-order stationary spatial process in two dimensions, the non-normalised spectral density function (or spectrum) $f(\omega_1, \omega_2)$ at frequencies ω_1, ω_2, is defined (Priestley, 1981, Section 9.7) as the Fourier transform of the covariance function:

$$f(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \sum_{g_1=-\infty}^{\infty} \sum_{g_2=-\infty}^{\infty} C(g_1, g_2) \cos(g_1\omega_1 + g_2\omega_2). \tag{3}$$

Given the spectral density, the covariance function can be obtained as

$$C(g_1, g_2) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(\omega_1, \omega_2) \cos(g_1\omega_1 + g_2\omega_2)\, \mathrm{d}\omega_1\, \mathrm{d}\omega_2.$$

The spectrum is non-negative and is a periodic function in ω_1 and ω_2, with period 2π. It also inherits the symmetric property from the covariogram, so that: $f(\omega_1, \omega_2) = f(2\pi - \omega_1, 2\pi - \omega_2) = f(-\omega_1, -\omega_2)$, $\forall \omega_1, \omega_2$.

2.2. Axially symmetric processes

Definition 3. A second-order stationary two-dimensional process is axial or reflection symmetric if the following equivalent conditions are satisfied:

(i) $\rho(g_1, g_2) = \rho(g_1, -g_2)$ or $C(g_1, g_2) = C(g_1, -g_2)$ or $S(g_1, g_2) = S(g_1, -g_2)$, $\quad \forall g_1, g_2$,
(ii) $f(\omega_1, \omega_2) = f(\omega_1, -\omega_2)$, $\quad \forall \omega_1, \omega_2$.

These conditions mean that the correlation function, the covariance function, the semi-variogram and the spectrum are all symmetric about the axes.

Axial symmetry is only one example of the various symmetries that the correlations may satisfy. Another example, unrelated to axial symmetry, is *diagonal* or *lateral* *symmetry* under which dependence is the same in the two directions, i.e. $\rho(g_1,g_2)=\rho(g_2,g_1)$, $\forall g_1,g_2$, with obvious implications on the covariance function and spectrum. If the correlations are both axially and diagonally symmetric, they are called *completely symmetric*.

In this paper, the only symmetry tests considered are for axial symmetry. Baczkowski and Mardia (1990) examined several tests of lateral symmetry for the AR(1) · AR(1) process based upon the sample semi-variogram.

2.3. Separable processes

Definition 4. A second-order stationary two-dimensional process is separable if the following equivalent conditions are satisfied:

(i) $\rho(g_1,g_2)=\rho(g_1,0)\rho(0,g_2)$ or $C(g_1,g_2)\propto C(g_1,0)C(0,g_2) \quad \forall g_1,g_2$,
(ii) $f(\omega_1,\omega_2)\propto f(\omega_1,0)f(0,\omega_2) \quad \forall \omega_1,\omega_2$.

For a separable process, therefore, the correlation and covariance structure, as well as the spectral density, are completely determined by the margins. Note that there is no equivalent form for separability in terms of the variogram. Separable covariances are also referred to as factorised covariances in Chilès and Delfiner (1999). A separable process is obviously also an axially symmetric process.

It is often convenient in practice that a process should be axially symmetric, and that the autocorrelations should have a simple form. Separable processes have both these properties, and can still be flexible enough to provide a reasonable representation of much planar autocorrelation.

For a separable process, the dispersion matrix for a realization on a regular rectangular lattice can be expressed as $V=V_1\otimes V_2$, where V_1 and V_2 are two matrices arising from the two underlying one-dimensional processes (Martin, 1979) and $\otimes$ indicates the Kronecker product. This implies that the determinant and inverse of V can be easily determined or calculated numerically, which is a major advantage of this subclass of processes. There are very few second-order stationary lattice processes for which it is possible to specify exactly in a simple form either V^{-1} or the correlations $\rho(g_1,g_2)$ and hence V. The inverse of the dispersion matrix and its determinant are required for exact Gaussian maximum likelihood estimation of the parameters of the spatial process. V^{-1} is also needed in the generalised least-squares estimation of the mean parameters, while exact Gaussian simulation essentially requires V or V^{-1} (Martin, 1996).

Despite their simplicity, separable models can often reasonably be used to represent data from a wide range of problems, especially if the lattice is not too large or the main interest is estimating ϑ in (1). Moreover, although this class of processes has been defined in terms of second-order stationary processes, generalizations to non-stationary processes are possible (Martin, 1990). Separable processes have been used for example in image processing (e.g., Habibi, 1972; Jain, 1981) and have been found to fit well remotely sensed data (e.g., Campbell, 1985). The use of separable processes has also been suggested in geostatistics and for agricultural field trials (e.g., Mardia, 1980; Cullis and Gleeson, 1991; Grondona et al., 1996).

3. Testing axial symmetry and separability using the periodogram

Since axial symmetry and separability are defined on the correlation structure or the spectrum of a spatial process, it seems natural to try and test these two hypotheses using sample versions of the covariogram, correlogram, semi-variogram or spectrum. However, simulation studies (Scaccia, 2000; Scaccia and Martin, 2002b) indicate that tests based on the sample covariogram require that the covariance structure of the sample covariances be taken into account. As a consequence, assumptions on the structure of the process need to be made, and parameters estimated, leading to model-dependent tests. Although a test considered for axial symmetry which assumed an AR(1)·AR(1) was very powerful for an AR(1)·AR(1) against a Pickard process, it was not useful for other processes. Tests considered for separability, including using the singular-value decomposition as suggested by Modjeska and Rawlings (1983), were found not to be useful. Tests based on the sample correlogram or semi-variogram would suffer from the same drawback.

As in time series analysis, an important advantage of the periodogram (the estimated spectrum), compared to the covariogram, is that the periodogram has a more satisfactory asymptotic sampling theory and, above all, it is asymptotically independent at different harmonic frequencies. The theoretical study of sampling variability and bias is much simpler in the case of the estimates of the spectrum than in the case of the estimates of the autocovariance function. Moreover, tests based on the periodogram are not model-based, and do not require estimation of the parameters of V. However, it is not obvious that tests based on the periodogram will be good in practice because large lattices may be needed for the asymptotic results to be reasonable.

Some possible tests for axial symmetry and separability based on the periodogram are justified and defined in this section. Section 3.1 discusses the distribution of the periodogram, while some possible tests using it are given in Section 3.2, which describes test statistics T_1 to T_3 (axial symmetry), and Section 3.3, which describes test statistic T_4 (separability). The performances of these statistics under the null and alternative hypotheses are considered, by simulation, in Section 5.

3.1. A sample estimator of the spectrum

In this paper we use the following sample version of the spectrum in (3):

$$I(\omega_1,\omega_2)=\frac{1}{(2\pi)^2}\sum_{g_1=-n_1+1}^{n_1-1}\sum_{g_2=-n_2+1}^{n_2-1}c(g_1,g_2)\cos(g_1\omega_1+g_2\omega_2),$$

where $c(g_1,g_2)$ is a sample estimator of the covariances. We use, for $g_1,g_2\geqslant 0$, the version

$$c(g_1,g_2)=\frac{1}{n}\sum_{i_1=1}^{n_1-g_1}\sum_{i_2=1}^{n_2-g_2}(Y_{i_1,i_2}-\overline{Y})(Y_{i_1+g_1,i_2+g_2}-\overline{Y})$$

$$c(g_1-g_2)=\frac{1}{n}\sum_{i_1=1}^{n_1-g_1}\sum_{i_2=1}^{n_2-g_2}(Y_{i_1,i_2+g_2}-\overline{Y})(Y_{i_1+g_1,i_2}-\overline{Y}),$$

with $\overline{Y}=\sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}Y_{i_1,i_2}/n$.

We say $\omega = (\omega_1, \omega_2)$ is a *Fourier* or *harmonic frequency* if ω_j is a multiple of $2\pi/n_j$. The main points in the asymptotic theory of the periodogram (sample spectrum) are given in Proposition 5.

Proposition 5 (see Priestley, 1981, Section 9.7). *The asymptotic distribution, expected value and variance of the sample spectrum, for harmonic frequencies $\omega_j \neq 0, \pi$, with $j = 1, 2$, are, as n_1 and $n_2 \to \infty$:*

$$\begin{aligned} &\frac{I(\omega_1,\omega_2)}{f(\omega_1,\omega_2)} && \to i.i.d.\ Exp(1) \\ &E[I(\omega_1,\omega_2)] && \to f(\omega_1,\omega_2) \\ &\mathrm{Var}[I(\omega_1,\omega_2)] && \to f^2(\omega_1,\omega_2). \end{aligned} \tag{4}$$

Note that $I(0,0) = 0$. Exact formulas for the mean and the variance of the one-dimensional periodogram can be found, for example, in Priestley (1981, Section 6.1.3). The two-dimensional results can be obtained by extending these. Note that the periodogram can be a seriously biased estimator of the spectrum for small n_1, n_2.

3.2. Testing axial symmetry

For an axially symmetric process we have $f(\omega_1,\omega_2) = f(\omega_1,-\omega_2)$. In view of (4), standard theory for generalized linear models could perhaps be used to test the null hypothesis of axial symmetry. An Exponential model can be fitted to the data and the deviance used to test $E[I(\omega_1,\omega_2)] = E[I(\omega_1,-\omega_2)]$ with $\omega_j = 2\pi k_j/n_j$, $k_j = 1, 2, \ldots, n_j^\star$, for some selected $n_j^\star$ provided its maximum value is $(n_j-1)/2$ if n_j is odd and $n_j/2-1$ if n_j is even. However, the deviance does not have the usual asymptotic χ^2 distribution on $n^\star$ degrees of freedom (where $n^\star = n_1^\star n_2^\star$). Assuming an Exponential model, the mean and variance of the deviance are given by the following lemma.

Lemma 6. *Assuming the asymptotic results* (4), *the exponential deviance*

$$\mathrm{Dev} = -2 \sum_{\omega_1,\omega_2} \left[\log \frac{I(\omega_1,\omega_2)}{\bar{I}(\omega_1,\omega_2)} + \log \frac{I(\omega_1,-\omega_2)}{\bar{I}(\omega_1,\omega_2)} \right],$$

with $\bar{I}(\omega_1,\omega_2) = [I(\omega_1,\omega_2) + I(\omega_1,-\omega_2)]/2$, has mean equal to $4[1-\log(2)]n^\star$, and variance equal to $8(2-\pi^2/6)n^\star$.

Proof. See appendix. □

However, even after adjusting for the mean and variance of the deviance, the test was not similar, and other methods were tried. Tests for axial symmetry can also be based on the differences $G(\omega_1,\omega_2) = I(\omega_1,\omega_2) - I(\omega_1,-\omega_2)$. Clearly $E[G(\omega_1,\omega_2)] = 0$ always, and from (4) it follows that $\mathrm{Var}[G(\omega_1,\omega_2)] \to 2f^2(\omega_1,\omega_2)$ as n_1 and $n_2 \to \infty$. Since the $G(\omega_1,\omega_2)$ do not have constant variance, we considered two modifications:

Modification 1. Take the logarithm of the $I(\omega_1, \omega_2)$ in order to approximately stabilize the variance and to reduce the non-normality. Asymptotically, the variance of the log periodogram is constant and the observations, at each frequency, are independently distributed, with a distribution nearly symmetric and closer to normality than the distribution of the periodogram.

Theorem 7. *Under axial symmetry the differences*

$$D(\omega_1, \omega_2) = \log[I(\omega_1, \omega_2)] - \log[I(\omega_1, -\omega_2)]$$

have a symmetric (though non-normal) distribution with $E[D(\omega_1, \omega_2)] = 0$ *always and* $\mathrm{Var}[D(\omega_1, \omega_2)] \to \pi^2/3$ *as* n_1 *and* $n_2 \to \infty$.

Proof. The expected value of $D(\omega_1, \omega_2)$ follows immediately from Definition 3. Moreover, the logarithm of an exponential distribution with mean λ is distributed as a Gumbel distribution with location parameter $\log(\lambda)$ and scale parameter 1. Therefore, for $\omega_j \neq 0, \pi$, we have, asymptotically

$$\log[I(\omega_1, \omega_2)] \sim \mathrm{Gumbel}(\log[f(\omega_1, \omega_2)], 1)$$

from which $E\{\log[I(\omega_1, \omega_2)]\} \to \log[f(\omega_1, \omega_2)] - \gamma$ and $\mathrm{Var}\{\log[I(\omega_1, \omega_2)]\} \to \pi^2/6$ as n_1 and $n_2 \to \infty$, where γ is Euler's constant. Since $I(\omega_1, \omega_2)$ and $I(\omega_1, -\omega_2)$ are asymptotically independent, the variance of $D(\omega_1, \omega_2)$ is asymptotically equal to twice the variance of the log periodogram. □

If axial symmetry does not hold, then
$E[D(\omega_1, \omega_2)] \to \log[f(\omega_1, \omega_2)] - \log[f(\omega_1, -\omega_2)]$.

On the basis of Theorem 7, and letting $\bar{D} = \sum_{\omega_1, \omega_2} D(\omega_1, \omega_2)/n^\star$ denote the sample mean, and $S^2 = \sum_{\omega_1, \omega_2} [D(\omega_1, \omega_2) - \bar{D}]^2/(n^\star - 1)$ denote the sample variance, we considered the following test statistics.

Test statistic T_1. Let

$$T_1 = \bar{D}\sqrt{n^\star}/\sqrt{\pi^2/3}.$$

Assuming that $n^\star$ is large enough for the central limit theorem to hold, T_1 is approximately distributed as a $\mathrm{N}(0, 1)$ under axial symmetry. This suggests a test of size φ should reject the null hypothesis if $|T_1| > z_{\varphi/2}$. Assuming the size of this test is φ (see Section 5), the approximate power of the test for large n_1, n_2 is $1 - \Phi(z_{\varphi/2} - \mu_{T_1}) + \Phi(-z_{\varphi/2} - \mu_{T_1})$, where $\mu_{T_1} = \mathrm{E}(T_1)$ and Φ is the $\mathrm{N}(0, 1)$ distribution function.

Test statistic T_2. Replacing the asymptotic standard deviation $\sqrt{\pi^2/3}$ of $D(\omega_1, \omega_2)$ by the sample estimate S gives

$$T_2 = \bar{D}\sqrt{n^\star}/S$$

which approximately has a t distribution with $n^\star - 1$ degrees of freedom under axial symmetry. This alternative test might give better results for small lattices. However, it might depend more on the approximation to normality of the $D(\omega_1, \omega_2)$ than does the test using T_1.

Modification 2. Estimate the variance of the $G(\omega_1, \omega_2)$, and hence consider the standardized differences

$$G_s(\omega_1, \omega_2) = [I(\omega_1, \omega_2) - I(\omega_1, -\omega_2)]/[I(\omega_1, \omega_2) + I(\omega_1, -\omega_2)].$$

The distribution of $\bar{G}_s = \sum_{\omega_1, \omega_2} G_s(\omega_1, \omega_2)/n^\star$ rapidly converges to a Normal distribution:

Theorem 8. *Under axial symmetry,* $\bar{G}_s\sqrt{n^\star}$ *converges to a* $\mathrm{N}(0, 1/3)$ *as* n_1, $n_2 \to \infty$.

Proof. Asymptotically, the $I(\omega_1, \omega_2)$ are independent and exponentially distributed, so that each $G_s(\omega_1, \omega_2)$ has a uniform distribution on $[-1, 1]$, from which the result follows. □

If axial symmetry does not hold, we need the following result.

Proposition 9. *If* X_1, X_2 *have independent Exponential distributions with* $E(X_i) = \mu_i$, *and* $R = (X_1 - X_2)/(X_1 + X_2)$, *then, letting* $\theta = \mu_1/\mu_2$, *for* $\theta \neq 1$, $E(R) = \mu_R = -2\theta\log(\theta)/(1-\theta)^2 - (1+\theta)/(1-\theta)$ *and* $\mathrm{var}(R) = -1 - 2(1+\theta)\mu_R/(1-\theta) - \mu_R^2$. *If* $\theta = 1 - \epsilon$, *then for small* $|\epsilon|$, $E(R) \approx -\epsilon/3 - \epsilon^2/6$ *and* $\mathrm{var}(R) \approx 1/3 - 2\epsilon^2/45$.

Proof. These follow from the probability density function of R, which can be shown to be $f_R(r) = 2\theta/\{1 + \theta + (1-\theta)r\}^2$ for $|r| < 1$. □

Test statistic T_3. Let

$$T_3 = \bar{G}_s\sqrt{3n^\star}.$$

Under the null hypothesis of axial symmetry, from Theorem 8, T_3 is asymptotically distributed as a $\mathrm{N}(0, 1)$. This suggests a test of size φ should reject the null hypothesis if $|T_3| > z_{\varphi/2}$. Assuming the size of this test is φ (see Section 5), the approximate power of the test is $1 - \Phi[(z_{\varphi/2} - \mu_{T_3})/\sigma_{T_3}] + \Phi[(-z_{\varphi/2} - \mu_{T_3})/\sigma_{T_3}]$, where $\mu_{T_3} = E(T_3)$, and $\sigma^2_{T_3} = \mathrm{var}(T_3)$.

Additionally, we also considered non-parametric tests such as the Wilcoxon and the Sign tests. They can be carried out equivalently using either the $G(\omega_1, \omega_2)$ or the $D(\omega_1, \omega_2)$ differences and they do not require normality of the data. However, if normality of Y holds, they are expected to be less powerful than the tests using T_1, T_2, T_3. For further details see Scaccia (2000).

3.3. Testing separability

Test statistic T_4. Consider the log periodogram. Under separability, the log spectrum is additive, $\log[f(\omega_1, \omega_2)] = c + \log[f(\omega_1, 0)] + \log[f(0, \omega_2)]$ where c is a constant, and $\log[I(\omega_1,\omega_2)]$ and $\log[I(\omega_1,-\omega_2)]$ are two sample realizations of the same value. So testing for separability reduces to testing for lack of interaction in a two-way classification table with two realizations in each cell and this can be done in a standard way, using the analysis of variance, since $\text{var}(\log[I(\omega_1,\omega_2)])$ is approximately constant. The lack of interaction can be tested using the statistic

$$T_4 = \frac{MS_{rc}}{MS_e}, \tag{5}$$

where MS_{rc} is the interaction mean square, and MS_e is the residual mean square. Under the null hypothesis of separability, and assuming that the log differences $D(\omega_1,\omega_2)$ are approximately normal, this test should be asymptotically distributed as an $F_{(n_1^\star-1)(n_2^\star-1),n^\star}$.

We also investigated two other possible tests. One used the Tukey test for additivity, which tests for lack of interaction in a two-way classification with only one realization in each cell, using the average $\{I[(\omega_1,\omega_2)]+I[(\omega_1,-\omega_2)]\}/2$. The other used the singular value decomposition of a matrix of periodogram values for which the corresponding spectrum matrix has rank one under separability.

4. Spatial processes used in the simulation study

The different tests proposed were analyzed and compared by means of simulations: separable (and axially symmetric) processes were used to simulate the null distribution of tests for separability (and axial symmetry), while non-separable (and non-axially symmetric) processes were used to simulate the distribution of the tests under the alternative hypothesis. The simulation study has a double purpose: determining how large the lattice needs to be for the asymptotic distributions of the test statistics to hold and provide some idea about the power of the different tests. Here we briefly describe the spatial processes used in this paper and the simulation method. We then illustrate the results obtained through simulations in Section 5.

We consider four different spatial processes: two separable ones, the $\text{AR}(1)\cdot\text{AR}(1)$ and the more general $\text{ARMA}(a_1,m_1)\cdot\text{ARMA}(a_2,m_2)$ processes, an axially symmetric, non-separable one, the $\text{CAR}(2)_{\text{SD}}$ process, and a non-separable, non-axially symmetric one, the Pickard process. The $\text{CAR}(2)_{\text{SD}}$, the Pickard and the $\text{AR}(1)\cdot\text{AR}(1)$ processes are all nested in the more general CAR(2) process; also, the $\text{AR}(1)\cdot\text{AR}(1)$ process is a special case of the Pickard, the $\text{CAR}(2)_{\text{SD}}$ and the $\text{ARMA}(a_1,m_1)\cdot\text{ARMA}(a_2,m_2)$ processes.

4.1. The $AR(1)\cdot AR(1)$ processes

The simplest non-trivial unilateral autoregressive separable process is the $\text{AR}(1)\cdot\text{AR}(1)$ process, also called the doubly-geometric process (Martin, 1979). It can be written as

$$Y_{i_1,i_2} = \alpha_1 Y_{i_1-1,i_2} + \alpha_2 Y_{i_1,i_2-1} - \alpha_1\alpha_2 Y_{i_1-1,i_2-1} + \varepsilon_{i_1,i_2},$$

where the ε_{i_1,i_2} are assumed to be independently distributed as $N(0,\sigma^2)$. For stationarity the parameters need to satisfy $|\alpha_1| < 1$ and $|\alpha_2| < 1$. The correlations are $\rho(g_1,g_2) = \alpha_1^{|g_1|}\alpha_2^{|g_2|}$. Particular examples used here, labelled A1 to A5, are:

Label	α	$\rho(1,0)$	$\rho(0,1)$	$\rho(1,1)$	$\rho(1,-1)$
A1	(0, 0)	0	0	0	0
A2	(0.2,0.3)	0.2	0.3	0.06	0.06
A3	(0.4,0.5)	0.4	0.5	0.20	0.20
A4	(0.6,0.7)	0.6	0.7	0.42	0.42
A5	(0.8,0.9)	0.8	0.9	0.72	0.72

Note that A1 represents an independent (White Noise) process.

4.2. The $ARMA(a_1,m_1)\cdot ARMA(a_2,m_2)$ processes

These separable models, which include moving average processes, are examples of the models referred to as linear-by-linear in Martin (1979). We consider models with autoregressive lags $a_j \leqslant 2$ and moving average lags $m_j \leqslant 1$. A representation of the ARMA(2,1)·ARMA(2,1) model, is

$$\begin{aligned}Y_{i_1,i_2} = {} & \alpha_1 Y_{i_1-1,i_2} + \alpha_2 Y_{i_1,i_2-1} + \psi_1 Y_{i_1-2,i_2} + \psi_2 Y_{i_1,i_2-2}\\ & -\alpha_1\alpha_2 Y_{i_1-1,i_2-1} - \alpha_1\psi_2 Y_{i_1-1,i_2-2} - \psi_1\alpha_2 Y_{i_1-2,i_2-1}\\ & -\psi_1\psi_2 Y_{i_1-2,i_2-2} + \varepsilon_{i_1,i_2} + \theta_1\varepsilon_{i_1-1,i_2} + \theta_2\varepsilon_{i_1,i_2-1} - \theta_1\theta_2\varepsilon_{i_1-1,i_2-1},\end{aligned}$$

where the ε_{i_1,i_2} are assumed to be independently distributed as $N(0,\sigma^2)$. Conditions for stationarity and invertibility are, respectively, $|\alpha_j| < 1 - \psi_j$ and $\psi_j > -1$, and $|\theta_j| < 1$. It is evident that an ARMA(1,0)·ARMA(1,0) is simply an AR(1)·AR(1) process. Examples used here, labelled AM1 to AM5, are:

Label	(a_1,m_1)	(a_2,m_2)	α	ψ	θ	$\rho(1,0)$	$\rho(0,1)$	$\rho(1,1)$	$\rho(1,-1)$
AM1	(2,0)	(2,0)	(0.5,0.4)	(0.3,0.2)	(0,0)	0.833	0.375	0.313	0.313
AM2	(0,1)	(2,0)	(0,0)	(0.7,0.2)	(0.3,0)	0.275	0.875	0.241	0.241
AM3	(0,1)	(0,1)	(0,0)	(0,0)	(0.4,0.6)	0.345	0.441	0.152	0.152
AM4	(2,1)	(2,0)	(0.3,0.2)	(0.5,0.4)	(0.2,0)	0.601	0.833	0.501	0.501
AM5	(2,1)	(2,1)	(0.4,0.1)	(0.3,0.4)	(0.6,0.3)	0.586	0.770	0.451	0.451

4.3. The $CAR(2)_{SD}$ processes

The only axially symmetric, non-separable model considered in this paper is a particular case of a second-order conditional autoregressive process (Besag, 1974), generally referred to as CAR(2). The general CAR(2) is neither separable, nor axially symmetric.

Here we consider the CAR(2)$_{SD}$ with symmetric diagonal term (Balram and Moura, 1993), which is axially symmetric and can be expressed as

$$E[Y_{i_1,i_2}|Y_{l_1,l_2}:(l_1,l_2)\neq(i_1,i_2)]=\beta_1(Y_{i_1-1,i_2}+Y_{i_1+1,i_2})$$
$$+\beta_2(Y_{i_1,i_2-1}+Y_{i_1,i_2+1})+\beta_3(Y_{i_1-1,i_2-1}+Y_{i_1+1,i_2+1}+Y_{i_1-1,i_2+1}+Y_{i_1+1,i_2-1})$$

with constant conditional variance.

Stationarity requires $|\beta_1+\beta_2|+2\beta_3<1/2$, and $|\beta_1-\beta_2|-2\beta_3<1/2$, which implies $|\beta_1|<1/2$, $|\beta_2|<1/2$, $|\beta_3|<1/4$. In general, neither V nor V^{-1} can easily be specified precisely, although much of V^{-1} can be given exactly. The spectrum can be written as $f(\omega_1,\omega_2)\propto 1/h(\omega_1,\omega_2)$ where

$$\begin{aligned}h(\omega_1,\omega_2)&=1-2\beta_1\cos(\omega_1)-2\beta_2\cos(\omega_2)-4\beta_3\cos(\omega_1)\cos(\omega_2)\\&=[1-2\beta_1\cos(\omega_1)][1-2\beta_2\cos(\omega_2)]-4\eta_\beta\cos(\omega_1)\cos(\omega_2),\end{aligned}\qquad(6)$$

where η_β denotes $\beta_3+\beta_1\beta_2$. Note that if $\beta_1=\beta_2=0$, $\rho(g_1,g_2)=0$ for $|g_1|+|g_2|$ odd, and the non-zero correlations are related to those of the process with $\beta=(\beta_3,\beta_3,0)$: $\rho(g_1,g_2)$ for the latter is $\rho(g_1+g_2,g_1-g_2)$ for the former.

The AR(1) $\cdot$AR(1) is a special case of a CAR(2)$_{SD}$: a CAR(2)$_{SD}$ with $\beta_3=-\beta_1\beta_2$, and $\beta_j=\alpha_j/(1+\alpha_j^2)$, $j=1,2$ has the same correlation structure as an AR(1)$\cdot$AR(1) model with parameters α_1, α_2. Then η_β, with $|\eta_\beta|<1/4$, is a natural measure of how far the CAR(2)$_{SD}$ is from the AR(1)$\cdot$AR(1), since, for $|\eta_\beta|$ small, $\rho(g_1,g_2)$ can be approximated by

$$\alpha_1^{|g_1|}\alpha_2^{|g_2|}\times\left[1+\eta_\beta\,\frac{(1+\alpha_1^2)(1+\alpha_2^2)(|g_1|+4\alpha_1^2-|g_1|\alpha_1^4)(|g_2|+4\alpha_2^2-|g_2|\alpha_2^4)-16\alpha_1^2\alpha_2^2}{(1-\alpha_1^2)^2(1-\alpha_2^2)^2\alpha_1\alpha_2}\right].$$

Theorem 10. *For a CAR*(2)$_{SD}$ model, $\rho(g_1,g_2)=\rho_1(g_1,g_2)+o(\eta_\beta)$.

Proof. See Appendix. □

The following proposition shows how the log spectrum differs from additivity for small $|\eta_\beta|$.

Proposition 11. *For small* $|\eta_\beta|$, $\log[f(\omega_1,\omega_2)]$ *can be approximated by a constant plus*

$$-\log[1-2\beta_1\cos(\omega_1)]-\log[1-2\beta_2\cos(\omega_2)]$$
$$+\frac{4\eta_\beta\cos(\omega_1)\cos(\omega_2)}{[1-2\beta_1\cos(\omega_1)][1-2\beta_2\cos(\omega_2)]}.$$

Particular examples of the CAR$(2)_{SD}$ used here, labelled C1 to C5, are listed below. Given the difficulty of getting V for the general CAR$(2)_{SD}$, we inverted V^{-1} for a boundary-corrected version of the CAR$(2)_{SD}$ (which sets values outside the lattice to μ) using the so called Dirichlet or free boundary condition in Balram and Moura (1993). The correlations shown are from the mid 11×11 array of a larger 29×29 array. Although C1–C4 are strictly non-stationary, values of the β_j close to the stationarity limit can be needed for even moderate correlations, and these examples are only being used here for power calculations.

Label	β	η_β	$\rho(1,0)$	$\rho(0,1)$	$\rho(1,1)$	$\rho(1,-1)$
C1	(0,0,0.25)	0.25	0	0	0.599	0.599
C2	(0.11,0.07,0.16)	0.1677	0.491	0.476	0.488	0.488
C3	(0.2,0.2,0.05)	0.09	0.593	0.593	0.514	0.514
C4	(0.25,0.25,0)	0.0625	0.631	0.631	0.538	0.538
C5	(0.3,0.2,−0.03)	0.03	0.391	0.290	0.165	0.165

4.4. The Pickard processes

The non-separable, non-axially-symmetric model considered in this paper, and referred to as the Pickard process, was considered by Pickard (1980) and Tory and Pickard (1992). It is a unilateral autoregressive model which can be expressed as

$$Y_{i_1,i_2} = \alpha_1 Y_{i_1-1,i_2} + \alpha_2 Y_{i_1,i_2-1} + \alpha_3 Y_{i_1-1,i_2-1} + \varepsilon_{i_1,i_2}.$$

Stationarity requires $|\alpha_1 + \alpha_2| < 1 - \alpha_3$ and $|\alpha_1 - \alpha_2| < 1 + \alpha_3$. Note that the correlation function is separable in half of the plane: $\rho(g_1, -g_2) = \rho(g_1, 0)\rho(0, g_2)$ for $g_1 g_2 \geqslant 0$. The spectrum can be written as $f(\omega_1, \omega_2) \propto 1/h(\omega_1, \omega_2)$ where

$$h(\omega_1,\omega_2) = (1 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2) - 2\alpha_3 \cos(\omega_1 + \omega_2) + 2\alpha_1\alpha_2 \cos(\omega_1 - \omega_2)$$
$$-2(\alpha_1 - \alpha_2\alpha_3)\cos(\omega_1) - 2(\alpha_2 - \alpha_1\alpha_3)\cos(\omega_2).$$

The AR(1)·AR(1) is a special case of a Pickard process with $\alpha_3 = -\alpha_1\alpha_2$, and this is the only axially symmetric (and separable) Pickard process (Martin, 1979). Let η_α denote $\alpha_3 + \alpha_1\alpha_2$. Then η_α, with $|\eta_\alpha| < 1/2$, is a natural measure of how far the Pickard process is from the AR(1)·AR(1), since, for $|\eta_\alpha|$ small, $\rho(g_1, g_2)$ can be approximated by

$$\rho_2(g_1,g_2) = \alpha_1^{|g_1|}\alpha_2^{|g_2|}\left[1 + \eta_\alpha\left(\frac{|g_1||g_2|}{\alpha_1\alpha_2} + \frac{|g_1|\delta_2}{\alpha_1} + \frac{|g_2|\delta_1}{\alpha_2}\right)\right], \quad \text{for } g_1 g_2 \geqslant 0$$

and

$$\rho_2(g_1,g_2) = \alpha_1^{|g_1|}\alpha_2^{|g_2|}\left[1 + \eta_\alpha\left(\frac{|g_1|\delta_2}{\alpha_1} + \frac{|g_2|\delta_1}{\alpha_2}\right)\right], \quad \text{for } g_1 g_2 < 0,$$

where $\delta_j = \alpha_j/(1 - \alpha_j^2)$.

Theorem 12. *For a Pickard process,* $\rho(g_1,g_2)=\rho_2(g_1,g_2)+o(\eta_\alpha)$.

Proof. See Appendix. □

Moreover, for $|\eta_\alpha|$ small, the differences $R(g_1,g_2)=\rho(g_1,g_2)-\rho(g_1,-g_2)$ can be approximated by $R_1(g_1,g_2)=g_1g_2\,\alpha_1^{(g_1-1)}\alpha_2^{(g_2-1)}\eta_\alpha$, for $g_1,g_2>0$:

Lemma 13. *For a Pickard process,* $R(g_1,g_2)=R_1(g_1,g_2)+o(\eta_\alpha)$, *for* $g_1,g_2>0$.

Proof. Follows from Theorem 12. □

As far as the periodogram is concerned, the differences $f(\omega_1,\omega_2)-f(\omega_1,-\omega_2)$ are proportional to $-4\eta_\alpha\sin(\omega_1)\sin(\omega_2)/[h(\omega_1,\omega_2)h(\omega_1,-\omega_2)]$, and, for small $|\eta_\alpha|$, the differences $d(\omega_1,\omega_2)=\log[f(\omega_1,\omega_2)]-\log[f(\omega_1,-\omega_2)]$ and the standardized differences $g(\omega_1,\omega_2)=[f(\omega_1,\omega_2)-f(\omega_1,-\omega_2)]/[f(\omega_1,\omega_2)+f(\omega_1,-\omega_2)]$ can be approximated, respectively, by $d_1(\omega_1,\omega_2)=-4\eta_\alpha h_1(\omega_1)h_2(\omega_2)$ and $g_1(\omega_1,\omega_2)=-2\eta_\alpha h_1(\omega_1)h_2(\omega_2)$ with $h_j(\omega_j)=\sin(\omega_j)/[1+\alpha_j^2-2\alpha_j\cos(\omega_j)]$.

Lemma 14. *For a Pickard process,* $d(\omega_1,\omega_2)=d_1(\omega_1,\omega_2)+o(\eta_\alpha)$ *and* $g(\omega_1,\omega_2)=g_1(\omega_1,\omega_2)+o(\eta_\alpha)$.

Proof. We have

$$d(\omega_1,\omega_2)=\log\left[1+\frac{h(\omega_1,-\omega_2)-h(\omega_1,\omega_2)}{h(\omega_1,\omega_2)}\right]=\log\left[1-\frac{4\eta_\alpha\sin(\omega_1)\sin(\omega_2)}{h(\omega_1,\omega_2)}\right]$$

so, for small $|\eta_\alpha|$, $d(\omega_1,\omega_2)=-\dfrac{4\eta_\alpha\sin(\omega_1)\sin(\omega_2)}{h(\omega_1,\omega_2)}+o(\eta_\alpha)$, which is equal to $-4\eta_\alpha h_1(\omega_1)h_2(\omega_2)+o(\eta_\alpha)$ since

$$h(\omega_1,\omega_2)=[1+\alpha_1^2-2\alpha_1\cos(\omega_1)][1+\alpha_2^2-2\alpha_2\cos(\omega_2)]+o(\eta_\alpha).$$

Similarly, $g(\omega_1,\omega_2)=-\frac{4\eta_\alpha\sin(\omega_1)\sin(\omega_2)}{[h(\omega_1,\omega_2)+h(\omega_1,-\omega_2)]}+o(\eta_\alpha)$ which is equal to $-2\eta_\alpha h_1(\omega_1)h_2(\omega_2)+o(\eta_\alpha)$. □

Note that, from Lemma 14, for the Pickard process with small $|\eta_\alpha|$, $E(\bar{D})$ is approximately $-4\eta_\alpha\bar{h}_1\bar{h}_2$ and $E(\bar{G})$ is approximately $-(4/3)\eta_\alpha\bar{h}_1\bar{h}_2$. Hence the choice of $n_j^\star$ will affect the power of those tests using T_1 to T_3 (see Section 3.2). Since for small $|\eta_\alpha|$, $\mu_{T_1}\approx-(4\eta_\alpha/\pi)\sqrt{3n^\star}\bar{h}_1\bar{h}_2$, and $\mu_{T_3}\approx-(4\eta_\alpha/3)\sqrt{3n^\star}\bar{h}_1\bar{h}_2$, the test using T_3 should then be slightly more powerful than the one using T_1.

Particular examples of the Pickard process used here, labelled P1–P5, are:

Label	α	η_α	$\rho(1,0)$	$\rho(0,1)$	$\rho(1,1)$	$\rho(1,-1)$
P1	(0.1, 0.2, 0.6)	0.62	0.426	0.476	0.733	0.203
P2	(0.6, 0.7, −0.8)	−0.38	0.180	0.624	−0.300	0.112
P3	(0.3, 0.4, 0.26)	0.38	0.694	0.733	0.757	0.509
P4	(0.1, 0.6, 0.2)	0.26	0.426	0.716	0.527	0.305
P5	(0.3, 0.6, −0.15)	0.03	0.329	0.611	0.231	0.201

5. Simulation study results

Simulations were used to estimate the distribution of the different test statistics under the null and the alternative hypotheses. Each process was simulated 1000 times to estimate the distribution of the test statistics. The vector of observations y was simulated as $y = T\varepsilon$ where ε is a random vector of n independent $\mathrm{N}(0,1)$ observations and T is a matrix such that $V = TT'$. The random vector ε was generated using the computer software Matlab (MathWorks, 2000), and randomly permuting the result to guard against possible serial correlation. The matrix T can be chosen in any convenient way. We usually used the Cholesky decomposition that specifies T as a lower triangular matrix, and if the observations are ordered lexicographically, the method corresponds to a finite unilateral moving average representation of the process. For the Pickard process, a slightly different method, corresponding to a finite unilateral autoregressive representation, was used (Martin, 1996).

Simulations were carried out on lattices of increasing dimension, starting from an 11×11 lattice. Processes were simulated under the null hypothesis (axial symmetry: A, AM, C, or separability: A, AM) and the distribution of the test statistic was considered. Firstly, for each test, we checked whether the test was similar. If so, and there was a theoretical (possibly asymptotic) distribution, the simulated distribution of the test statistic was compared with this. We used the mean, standard deviation, and the standardised measures of skewness and kurtosis for this, as well as X^2 tests of homogeneity and goodness-of-fit. However, principally we considered the 5% and 1% points of the simulated distribution, as these are of most concern for using the test. We then simulated under the alternative hypothesis (axial symmetry: P or separability: C, P) so that the power of the test could be estimated. If the power estimate is p%, then its estimated standard error is $\sqrt{p(100-p)/1000}$%.

5.1. Test statistics for axial symmetry

For test statistics T_1, T_2 and T_3 a good correspondence between expected and simulated distributions was confirmed by a χ^2 test for goodness of fit, even for 11×11 lattices, with $n_j^* = 5$ (corresponding to the maximum range $[0, 10\pi/11]$ for ω_j).

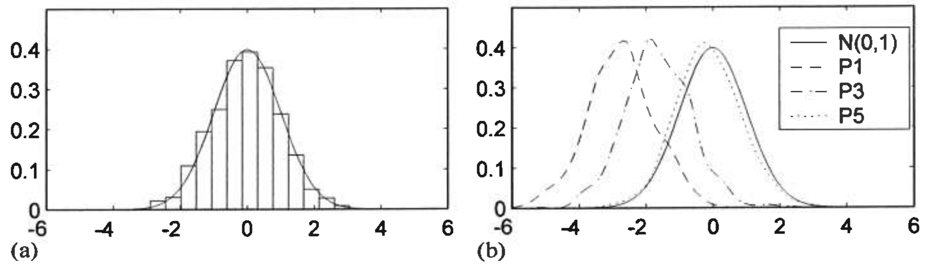

Fig. 1. (a) Comparison between the theoretical and the simulated distribution of test statistic T_1 (simulation of AM5 on an 11×11 lattice with $n_j^{\star} = 5$); (b) Theoretical distribution, under the null hypothesis, of test statistic T_1 and smoothed simulated distribution under different alternative hypotheses (simulation on an 11×11 lattice with $n_j^{\star} = 5$).

Fig. 1(a) shows, for example, a very good approximation of the simulated distribution of the test statistic T_1, for process AM5, to the standard normal distribution. Test statistics T_1, T_2, T_3 show a good correspondence between expected and simulated distributions for 11×11 lattices, also when restricting the range of ω_j. We considered $n_j^{\star} = 2,3,4,5$ for the 11×11 lattice and $n_j^{\star} = 4,5,6,7$ for the 15×15 lattice.

Now consider the powers of the tests. An example of the smoothed simulated distribution of T_1 for processes P1, P3, P5 is given in Fig. 1(b). A shift to the left is evident as η_α increases (see Section 4.4). The estimated power for the five different Pickard processes, on an 11×11 lattice, and on a 15×15 lattice, is given in Table 1 for two different $n_j^{\star}$. The powers of the tests using T_1 and T_3 are alike and, as expected, that using T_2 is slightly worse. It was also found for P1 to P5 that the simulated power of test statistics T_1, T_2 and T_3 is usually maximum when $n_j^{\star} = 4$ for the 11×11 lattice and when $n_j^{\star} = 5$ for the 15×15 lattice, i.e. when ω_j is approximately in the range $[0, 2\pi/3]$.

Assuming the asymptotic Exponential distribution is reasonable for $I(\omega_1, \omega_2)$, but calculating the exact value of $E[I(\omega_1, \omega_2)]$ and using the results in Section 3.2, gives very good indications of the power for T_1 and T_3. For example, for P1, the 11×11 lattice and $n_j^{\star} = 4$, this gives $\mu_{T_1} = -3.015$, resulting in a power of respectively 85.4% and 67.0% for the 5% and the 1% test based on T_1; and $\mu_{T_3} = -2.897$, with $\sigma_{T_3} = 0.877$, resulting in a power of respectively 85.7% and 64.3% for the 5% and the 1% test based on T_3. This suggests that for a particular alternative the power can be well estimated directly, and that a good value for $n_j^{\star}$ can be determined.

The estimated power of the Wilcoxon test is quite close to those of the other test statistics (at significance level 5%, the simulated power for the alternative P2, was equal to 43.9% on an 11×11 lattice with $n_j^{\star} = 5$, and equal to 81.1% on a 15×15 lattice with $n_j^{\star} = 7$), while the estimated power of the Sign test is considerably smaller (at significance level 5%, the estimated power for the alternative P2, was equal to 31.3% on an 11×11 lattice with $n_j^{\star} = 5$, and equal to 62.8% on a 15×15 lattice with $n_j^{\star} = 7$).

Table 1
Simulated power (%) of tests for axial symmetry for five different Pickard processes, on an 11×11 lattice, and on a 15×15 lattice, when the level of the tests is equal to 5% and 1%.

Process	Test	11 × 11 lattice				15 × 15 lattice			
		5% level		1% level		5% level		1% level	
		$n_j^\star = 4$	$n_j^\star = 5$	$n_j^\star = 4$	$n_j^\star = 5$	$n_j^\star = 5$	$n_j^\star = 7$	$n_j^\star = 5$	$n_j^\star = 7$
P1	T_1	87.0	78.1	66.2	58.4	99.1	97.4	94.4	91.6
	T_2	80.0	71.6	54.5	44.7	97.0	95.6	85.7	83.6
	T_3	87.4	78.3	66.5	55.4	98.8	97.1	94.0	91.1
P2	T_1	62.8	50.2	38.9	25.0	94.4	87.0	86.0	67.2
	T_2	56.7	46.2	30.1	22.2	91.1	82.5	73.5	57.9
	T_3	65.0	50.0	37.6	25.0	95.0	86.3	83.9	65.0
P3	T_1	47.1	41.3	24.8	20.5	81.7	67.8	59.1	47.5
	T_2	45.6	37.0	20.5	17.8	78.3	64.9	50.8	40.0
	T_3	49.4	44.0	26.2	20.8	82.3	71.3	61.6	46.0
P4	T_1	26.8	22.0	10.2	7.9	51.3	44.6	27.4	20.9
	T_2	24.6	22.1	8.6	7.3	48.4	42.1	26.1	20.2
	T_3	28.3	22.3	10.2	7.4	55.6	45.8	28.7	23.1
P5	T_1	4.3	4.9	1.0	0.7	4.8	5.1	1.3	1.2
	T_2	4.3	5.4	0.9	0.6	4.8	5.2	0.8	1.4
	T_3	4.7	5.2	0.8	0.8	5.1	5.3	1.2	0.9

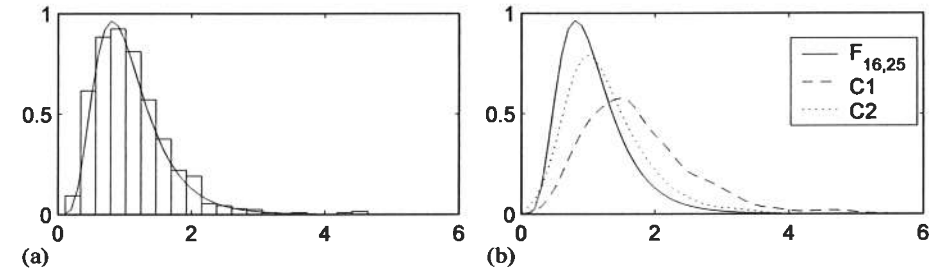

Fig. 2. (a) Comparison between the theoretical and the simulated distribution of test statistic T_4 (simulation of AM1 on an 11×11 lattice with $n_j^\star = 5$); (b) Theoretical distribution, under the null hypothesis, of test statistic T_4 and smoothed simulated distribution under different alternative hypotheses (simulation on an 11×11 lattice with $n_j^\star = 5$).

5.2. Test statistics for separability

The simulations of the test statistic T_4 under the null hypothesis showed a reasonable correspondence between the estimated and expected distributions of the test statistic for an 11×11 lattice, as shown, for example, in Fig. 2(a). The simulated power for the different processes, on an 11×11 lattice, and on a 15×15 lattice, is given in Table 2. An example of the smoothed simulated distribution of T_4, for processes C1 and C2, is given in Fig. 2(b). The test has low estimated power for all the $\mathrm{CAR}(2)_{\mathrm{SD}}$ processes

Table 2
Simulated power (%) of test statistic T_4 for different CAR(2)$_{SD}$, on an 11×11 lattice, and on a 15×15 lattice, when the level of the tests is equal to 5% and 1%

Process	11×11 lattice				15×15 lattice			
	5% level		1% level		5% level		1% level	
	$n_j^\star = 4$	$n_j^\star = 5$	$n_j^\star = 4$	$n_j^\star = 5$	$n_j^\star = 5$	$n_j^\star = 7$	$n_j^\star = 5$	$n_j^\star = 7$
C1	10.8	27.4	2.7	10.3	11.4	48.8	4.1	21.5
C2	6.3	9.1	2.2	1.8	9.8	14.3	3.3	4.3
C3	7.0	6.8	1.7	1.8	8.2	9.4	2.0	2.4
C4	6.0	6.2	1.8	1.8	7.6	6.4	1.6	1.2
C5	5.8	5.9	1.2	1.3	7.5	6.2	1.6	0.9

considered, except for C1. The simulated power is generally maximum when $n_j^\star = n_j$. However, T_4 can be misleadingly small when axial symmetry does not hold, since the two values in each cell do not then have the same mean, i.e. $E[D(\omega_1, \omega_2)] \neq 0$, which causes MS_e to be biased upwards. Thus if axial symmetry is in doubt, it is necessary to test for it before using this test for separability.

For the Tukey test, the correspondence between the theoretical and simulated distributions was not completely satisfactory, while tests using the singular value decomposition were not similar and had poor discriminating power.

6. Conclusions and further developments

Different ways for testing the hypotheses of axial symmetry and separability, for processes on a regular lattice, were proposed and compared. The aim was to find tests which can be applied to a large range of processes and to test separability in two stages: testing axial symmetry first, and then if this hypothesis is not rejected, testing separability. We have not attempted, however, to control the joint size of the sequential tests of axial symmetry followed by separability. A simple method for this would use the Bonferroni inequality, with the two test sizes adding to the required overall size.

Sample covariances are correlated, and tests considered which used the sample covariogram were either unsatisfactory, or model-dependent. The periodogram is asymptotically independent at different frequencies, and some tests based on it were found to be satisfactory, even for quite small lattices. These used T_1, T_2, T_3 for axial symmetry, and T_4 for separability given axial symmetry. These tests are very general, and can also easily be used on large lattices. The main assumption required is the stationarity of the process.

In future work, other tests for separability could be considered, based for example on the log differences

$$\{\log[I(\omega_1, \omega_2)] + \log[\hat{f}(0,0)]\} - \{\log[I(\omega_1, 0)] + \log[I(0, \omega_2)]\},$$

where $\hat{f}(0,0)$ is an estimate of $f(0,0)$. Also, the robustness of the tests developed in this paper to having a non-constant parameterized mean, and to non-normality, could be studied. If data are not normal, a possible solution is, obviously, to transform them in some appropriate way in order to achieve approximate normality. Otherwise ad hoc tests could be developed. Also most of the tests based on the periodogram are expected to be robust to the lack of normality unless the n_j are small, particularly the non-parametric ones.

We have also, by analogy with the Durbin–Watson test for serial correlation in regression models, examined generalised likelihood ratio tests of axial symmetry and separability by comparing AR(1)·AR(1) processes with Pickard processes, and CAR(2)$_{SD}$ processes. These will be considered in a forthcoming manuscript (see also Scaccia, 2000 and Scaccia and Martin, 2002a).

Further extensions could encompass testing separability between space and time in spatio-temporal processes which leads to a significant simplification in the analysis of the process and model fitting, and is generally assumed without testing for it. For two spatial dimensions and time, tests of $f(\omega_1,\omega_2,\omega_3)=f(\omega_1,\omega_2,-\omega_3)$, for all $\omega_1,\omega_2,\omega_3$, would be similar to those using T_1 to T_3; and a test of $f(\omega_1,\omega_2,\omega_3) \propto f(\omega_1,\omega_2,0)f(0,0,\omega_3)$ would, similarly to T_4, test that all interactions except that between the spatial dimensions are 0. A related test using the smoothed coherence has recently been proposed by Fuentes (2003).

Finally, existing tests for lateral symmetry, based on the semi-variogram for the AR(1)·AR(1) model, have only a limited extension to other covariance models. It is believed that simpler tests, based on the periodogram, could be found, and that these would probably also have the advantage of being of more general application and implementable even on relatively small lattices. One possibility has been suggested by Lu and Zimmerman (2002).

Matlab code for carrying out these tests is available from the authors.

Acknowledgements

This work is based on the University of Perugia Ph.D. Thesis of L. Scaccia. We are grateful to T. Subba Rao and J. Biggins for suggestions on the periodogram tests, and to the University of Sheffield for support for L. Scaccia. We are also grateful for the referees' comments, and for some anonymous comments on an earlier version of the paper.

Appendix.

Proof of Lemma 6. Let us consider the variables $X=\{X_i : i=1,\dots,n\} \sim$ i.i.d. Exp(λ) and $Y=\{Y_i : i=1,\dots,n\} \sim$ i.i.d. Exp(λ). We want to determine the mean and the variance of the deviance $\mathrm{Dev} = -2\sum_i \{\log[2X_i/(X_i+Y_i)] + \log[2Y_i/(X_i+Y_i)]\}$.

Put $U_1 = X/(X+Y)$, $U_2 = Y/(X+Y)$, $W_j = 2U_j$ and $Z_j = -\log(U_j)$, for $j = 1,2$. It follows that $U_j \sim \mathrm{Unif}(0,1)$ and $Z_j \sim \mathrm{Exp}(1)$, with $E(Z_j) = \mathrm{var}(Z_j) = 1$. Moreover $\log(W_j) = \log(2) - Z_j$ with $E[\log(W_j)] = \log(2) - 1$ and $\mathrm{var}[\log(W_j)] = \mathrm{var}(Z_j) = 1$. Then

$$E(\mathrm{Dev}) = -2n\{E[\log(W_1)] + E[\log(W_2)]\} = 4n[1 - \log(2)]$$

$$\mathrm{var}(\mathrm{Dev}) = 4n\mathrm{var}[\log(W_1) + \log(W_2)] = 8n[\mathrm{var}(Z_1) + E(Z_1Z_2) - E(Z_1)E(Z_2)].$$

Now $E(Z_1Z_2) = \int_0^1 \log(U_1)\log(1-U_1)\,\mathrm{d}U_1$, which using integration by parts and substitution equals $\int_0^1 [-2\log(U_1) + \log(U_1)/(1-U_1)]\mathrm{d}U_1$, which is $2 - \pi^2/6$ (Abramowitz and Stegun, 1965, Eq. 4.1.55) from which $\mathrm{var}(\mathrm{Dev}) = 8(2 - \pi^2/6)n$. □

Proof of Theorem 10. The autocovariance generating function (acgf) of a stationary process is:

$$\Gamma(z_1, z_2) = \sum_{g_1=-\infty}^{\infty} \sum_{g_2=-\infty}^{\infty} C(g_1, g_2) z_1^{g_1} z_2^{g_2}.$$

It is related to the spectrum by $f(\omega_1, \omega_2) = \Gamma(\mathrm{e}^{\mathrm{i}\omega_1}, \mathrm{e}^{\mathrm{i}\omega_2})/(2\pi)^2$. For ARMA and CAR processes, $\Gamma(z_1, z_2)$ can be obtained directly from the generating equation.

Now, let $a(\alpha_j, z_j)$ denote $(1 - \alpha_j z_j)(1 - \alpha_j z_j^{-1})$. Then for the $\mathrm{CAR}(2)_{\mathrm{SD}}$,

$$1/\Gamma(z_1, z_2) \propto 1 - \beta_1(z_1 + z_1^{-1}) - \beta_2(z_2 + z_2^{-1}) - \beta_3(z_1 + z_1^{-1})(z_2 + z_2^{-1}),$$

which can be written as $[1 - \beta_1(z_1 + z_1^{-1})][1 - \beta_2(z_2 + z_2^{-1})] - \eta_\beta(z_1 + z_1^{-1})(z_2 + z_2^{-1})$. Putting $\beta_j = \alpha_j/(1 + \alpha_j^2)$, gives

$$1/\Gamma(z_1, z_2) \propto a(\alpha_1, z_1)a(\alpha_2, z_2) - \eta_\beta(1 + \alpha_1^2)(1 + \alpha_2^2)b(z_1)b(z_2),$$

where $b(z_j) = z_j + z_j^{-1}$. Hence, $\Gamma(z_1, z_2)$ is proportional to

$$\frac{1}{a(\alpha_1, z_1)a(\alpha_2, z_2)} + \eta_\beta(1 + \alpha_1^2)(1 + \alpha_2^2)\frac{b(z_1)b(z_2)}{[a(\alpha_1, z_1)a(\alpha_2, z_2)]^2} + o(\eta_\beta).$$

Now $1/a(\alpha_j, z_j) = \sum_{g_j=-\infty}^{\infty} \alpha_j^{|g_j|} z^{g_j}/(1 - \alpha_j^2)$, and

$$b(z_j)/[a(\alpha_j, z_j)]^2 = \sum_{g_j=-\infty}^{\infty} (|g_j| + 4\alpha_j^2 - |g_j|\alpha_j^4)\alpha_j^{|g_j|-1} z^{g_j}/(1 - \alpha_j^2)^3,$$

so that $C(g_1, g_2)$ is proportional to

$$\alpha_1^{|g_1|}\alpha_2^{|g_2|}\left[1 + \eta_\beta \frac{(1 + \alpha_1^2)(1 + \alpha_2^2)(|g_1| + 4\alpha_1^2 - |g_1|\alpha_1^4)(|g_2| + 4\alpha_2^2 - |g_2|\alpha_2^4)}{(1 - \alpha_1^2)^2(1 - \alpha_2^2)^2\alpha_1\alpha_2}\right] + o(\eta_\beta).$$

The result on $\rho(g_1, g_2) = C(g_1, g_2)/C(0,0)$ then follows. □

Proof of Theorem 12. For the Pickard process,

$$1/\Gamma(z_1, z_2) \propto (1 - \alpha_1 z_1 - \alpha_2 z_2 - \alpha_3 z_1 z_2)(1 - \alpha_1 z_1^{-1} - \alpha_2 z_2^{-1} - \alpha_3 z_1^{-1} z_2^{-1}),$$

which can be written as

$$a(\alpha_1,z_1)a(\alpha_2,z_2) - \eta_\alpha[k(\alpha_1,z_1)k(\alpha_2,z_2) + k(\alpha_1,z_1^{-1})k(\alpha_2,z_2^{-1})] + \eta_\alpha^2,$$

where $k(\alpha_j,z_j) = z_j - \alpha_j = z_j(1-\alpha_j z_j^{-1})$. Hence, $\Gamma(z_1,z_2)$ is proportional to

$$\frac{1}{a(\alpha_1,z_1)a(\alpha_2,z_2)} + \eta_\alpha \frac{k(\alpha_1,z_1)k(\alpha_2,z_2) + k(\alpha_1,z_1^{-1})k(\alpha_2,z_2^{-1})}{[a(\alpha_1,z_1)a(\alpha_2,z_2)]^2} + o(\eta_\alpha).$$

Now, $k(\alpha_j,z_j)/[a(\alpha_j,z_j)]^2 = \left(\sum_{g_j=1}^{\infty} g_j\alpha_j^{g_j-1}z_j^{g_j} + \delta_j \sum_{g_j=-\infty}^{\infty} \alpha_j^{|g_j|}z_j^{g_j}\right)/(1-\alpha_j^2)$ where $\delta_j = \alpha_j/(1-\alpha_j^2)$, so that $C(g_1,g_2)$ is proportional to

$$\alpha_1^{|g_1|}\alpha_2^{|g_2|}\left[1+\eta_\alpha\left(\frac{|g_1||g_2|}{\alpha_1\alpha_2} + \frac{|g_1|\delta_2}{\alpha_1} + \frac{|g_2|\delta_1}{\alpha_2} + 2\delta_1\delta_2\right)\right] + o(\eta_\alpha) \quad \text{for } g_1g_2 \geqslant 0$$

and to

$$\alpha_1^{|g_1|}\alpha_2^{|g_2|}\left[1+\eta_\alpha\left(\frac{|g_1|\delta_2}{\alpha_1} + \frac{|g_2|\delta_1}{\alpha_2} + 2\delta_1\delta_2\right)\right] + o(\eta_\alpha) \quad \text{for } g_1g_2 < 0.$$

The result on $\rho(g_1,g_2) = C(g_1,g_2)/C(0,0)$ then follows. □

References

Abramowitz, M., Stegun, I.A., 1965. Handbook of Mathematical Functions. Dover, New York.

Baczkowski, A.J., Mardia, K.V., 1990. A test of spatial symmetry with general application. Commun. Statist. Theory Methods 19, 555–572.

Balram, N., Moura, J.M.F., 1993. Noncausal Gauss Markov random fields: parameter structure and estimation. IEEE Trans. Inf. Theory 39, 1333–1355.

Besag, J.E., 1974. Spatial interaction and the statistical analysis of lattice systems (with discussion). J. Roy. Statist. Soc. B 36, 192–236.

Campbell, N.A., 1985. Remote sensing in practice. Paper presented at the Open Forum on Statistics and Pattern Recognition, Edinburgh, July 1985.

Chilès, J.P., Delfiner, P., 1999. Geostatistics; Modeling Spatial Uncertainty. Wiley, New York.

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

Cullis, B.R., Gleeson, A.C., 1991. Spatial analysis of field experiments—an extension to two dimensions. Biometrics 47, 1449–1460.

Fuentes, M., 2003. Testing for separability of spatial–temporal covariance functions. Proceedings of the ISI Conference on Environmental Statistics and Health, Universidade de Santiago de Compostela, pp. 67–75.

Grondona, M.O., Crossa, J., Fox, P.N., Pfeiffer, W.H., 1996. Analysis of variety yield trials using two-dimensional separable ARIMA processes. Biometrics 52, 763–770.

Guo, J.H., Billard, L., 1998. Some inference results for causal autoregressive processes on a plane. J. Time Ser. Anal. 19, 681–691.

Habibi, A., 1972. Two-dimensional Bayesian estimate of images. Proc. IEEE 60, 878–883.

Jain, A.K., 1981. Advances in mathematical models for image processing. Proc. IEEE 69, 502–528.

Lu, N., Zimmerman, D.L., 2002. Testing for directional symmetry using the periodogram. Research Report #317, Department of Statistics and Actuarial Science, University of Iowa.

Mardia, K.V., 1980. Some statistical inference problems in Kriging II: theory. Proceedings of the 26th International Geological Congress, Paris. Sciences de la Terre, Serie Informatique Geologique, pp. 113–131.

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 field trials. Commun. Statist. Theory Methods 19, 55–81.

Martin, R.J., 1996. Some results on unilateral ARMA lattice processes. J. Statist. Plann. Inference 50, 395–411.

MathWorks, 2000. Matlab version 6.0. Release 12. Natick, Massachusetts.

Modjeska, J.S., Rawlings, J.O., 1983. Spatial correlation analysis of uniformity data. Biometrics 39, 373–384.

Pickard, D.K., 1980. Unilateral Markov fields. Adv. Appl. Probab. 12, 655–671.

Priestley, M.B., 1981. Spectral Analysis and Time Series. Academic Press, London.

Scaccia, L., 2000. Testing for simplification in spatial models. Ph.D. Thesis, University of Perugia, Italy.

Scaccia, L., Martin, R.J., 2002a. Testing for simplification in spatial models. In: Härdle, W., Rönz, B. (Eds.), COMPSTAT 2002, Proceedings in Computational Statistics. Physica-Verlag, Heidelberg, pp. 581–586.

Scaccia, L., Martin, R.J., 2002b. Testing axial symmetry and separability of lattice processes. Research Report No. 530/02, Department of Probability and Statistics, University of Sheffield.

Shitan, M., Brockwell, P.J., 1995. An asymptotic test for separability of a spatial autoregressive model. Commun. Statist. Theory Methods 24, 2027–2040.

Tory, E.M., Pickard, D.K., 1992. Unilateral Gaussian fields. Adv. Appl. Probab. 24, 95–112.