# Testing for Simplification in Spatial Models

L. Scaccia[1] and R. J. Martin[2]

[1] Dipartimento di Scienze Statistiche, Università di Perugia, 06100 Perugia, Italy
[2] Department of Probability and Statistics, University of Sheffield, S3 7RH Sheffield, UK

**Abstract**. Data collected on a rectangular lattice occur frequently in many areas such as field trials, geostatistics, remotely sensed data, and image analysis. Models for the spatial process often make simplifying assumptions, including axial symmetry and separability. We consider methods for testing these assumptions and compare tests based on sample covariances, tests based on the sample spectrum, and model-based tests.

**Keywords**. Autoregressive process, axial symmetry, doubly-geometric process, lattice process, separability, spatial process.

## 1 Introduction

Data collected on a regular two-dimensional lattice arise in many areas. Most models used for such data assume axial symmetry, but can still be difficult to use. Many of the problems of two-dimensional modelling can be overcome by using separable processes. Under this assumption, time series methods can be used to analyze and model lattice data (Martin, 1990).

However, no tests for separability and axial symmetry are known apart from Guo and Billard (1998), who suggested the Wald test to test separability. This model-based test compares the fit of an AR(1)·AR(1) process and the more general Pickard process to the data, and actually tests for axial symmetry and separability together.

We propose and compare some different model-free and model-based tests for testing axial symmetry and separability. The paper is structured as follows. In section 2 we give some notation and definitions, illustrate the simulation method and present the different models used. The different tests for axial symmetry and separability are discussed in section 3. In section 4 we compare the performances of the tests.

## 2 Definitions, models and simulation method

### 2.1 Notation

We assume that data occur on an $n_1$ by $n_2$ rectangular lattice, with rows indexed by $i_1 = 1, \ldots, n_1$ and columns by $i_2 = 1, \ldots, n_2$. Row and column lags are $g_1$ and $g_2$, with $g_j = -(n_j - 1), \ldots, 0, \ldots, (n_j - 1)$, for $j = 1, 2$. 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 $Y_{i_1,i_2}$. Let the vector $Y$ contain the $Y_{i_1,i_2}$ in site order and assume a constant mean. We assume $Y \sim \mathrm{N}(0, V\sigma^2)$ where $V = V(\alpha)$ is an $n$ by $n$ positive definite matrix depending on the $q$-vector of parameters $\alpha$.

Assuming second-order stationarity, $C(g_1, g_2) = \mathrm{Cov}\,(Y_{i_1,i_2}, Y_{i_1+g_1,i_2+g_2})$ is the covariance at lags $g_1$, $g_2$, with $C(0,0) = \sigma^2$. The spectrum is $f(\omega_1, \omega_2) = \sum_{g_1=-\infty}^{\infty} \sum_{g_2=-\infty}^{\infty} C(g_1, g_2) \cos(g_1\omega_1 + g_2\omega_2)/(2\pi)^2$ (Priestley, 1981, section 9.7). Note that $C(g_1, g_2) = C(-g_1, -g_2)$ and $f(\omega_1, \omega_2) = f(-\omega_1, -\omega_2)$ always.

## 2.2 Axial symmetry and separability

For an *axially* or *reflection symmetric* process $C(g_1, g_2) = C(g_1, -g_2), \forall g_1, g_2$ and, equivalently, $f(\omega_1, \omega_2) = f(\omega_1, -\omega_2), \forall \omega_1, \omega_2$. This means that the covariances and the spectrum are both symmetric about the axes.

For a process to be *separable*, $C(g_1, g_2) \propto C(g_1, 0) \cdot C(0, g_2), \forall\ g_1, g_2$ and $f(\omega_1, \omega_2) \propto f(\omega_1, 0) \cdot f(0, \omega_2), \forall\ \omega_1, \omega_2$ are also required. Therefore, the covariances and the spectrum are determined, up to a multiplicative constant, by the margins. Clearly separability implies axial symmetry.

It is often desirable in practice that a process should be axially symmetric, and that the covariances should have a simple form. Separable processes have both these properties, and are still flexible enough to provide a reasonable representation of many planar structures (Martin, 1990).

For a separable process, the $V$ matrix can be expressed as $V = V_x \otimes V_z$, where $V_x$ and $V_z$ are two smaller matrices that arise from the two underlying one-dimensional processes, and this implies that its determinant and inverse (required for exact Gaussian maximum likelihood, generalised least-squares estimation and exact Gaussian simulation) are easily determinable, which is a major advantage of this subclass of processes (Martin, 1996).

## 2.3 Models simulated and simulation method

We compared the different tests proposed by simulation: separable (or axially symmetric) processes were used to simulate the null distribution of tests for separability (or axial symmetry), while non-separable (or non-symmetric) processes were used to simulate the distribution of the tests under the alternative hypothesis. We indicate tests for axial symmetry by $S$ and tests for separability by $R$, using different superscripts for their basis: ($c$) sample covariances, ($p$) sample spectrum, and ($m$) model-based.

The separable process we used is the AR(1)·AR(1) process (Martin, 1979):

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

where the $\epsilon_{i_1,i_2}$ are assumed to be independently distributed as $\mathrm{N}(0, \sigma_\epsilon^2)$.

As an axially symmetric, non-separable model we used a particular case of a second-order conditional autoregressive process, the $\mathrm{CAR}(2)_{SD}$ with symmetric diagonal term (Balram and Moura, 1993). It can be written as:

$$\begin{aligned} \mathrm{E}(Y_{i_1,i_2}|\cdot) = {} & \beta_1(Y_{i_1-1,i_2} + Y_{i_1+1,i_2}) + \beta_2(Y_{i_1,i_2-1} + Y_{i_1,i_2+1}) \\ & + \beta_3(Y_{i_1-1,i_2-1} + Y_{i_1+1,i_2+1} + Y_{i_1-1,i_2+1} + Y_{i_1+1,i_2-1}) \end{aligned}$$

with constant conditional variance. Conditioning is on $Y_{l_1,l_2}, \forall (l_1, l_2) \neq (i_1, i_2)$. A $\mathrm{CAR}(2)_{SD}$ with $\beta_3 = -\beta_1\beta_2$, $\beta_1 = \alpha_1/(1+\alpha_1^2)$ and $\beta_2 = \alpha_2/(1+\alpha_2^2)$ has the same covariance structure as an AR(1)·AR(1) with parameters $\alpha_1$, $\alpha_2$.

The non-separable, non-axially-symmetric model used, referred to as the Pickard process (Pickard, 1980, Tory and Pickard 1992) can be written as:

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

The AR(1)·AR(1) is a special case of the Pickard process with $\alpha_3 = -\alpha_1\alpha_2$.

Each process was simulated 1 000 times to estimate the distribution of the test statistics under the null and the alternative hypotheses. Choosing a matrix $T$ such that $V = TT'$, the observation vector $y$ can be simulated as $y = T\epsilon$ where $\epsilon$ is a random vector of $n$ independent $\mathrm{N}(0,1)$ observations.

# 3 Testing axial symmetry and separability

## 3.1 Tests based on sample covariances

Since axial symmetry and separability are usually defined in terms of the covariance structure, it seems natural to try and test these two hypotheses using the sample covariances. The sample estimator of the covariances we used is $c(g_1,g_2) = \sum_{i_1=1}^{n_1-g_1}\sum_{i_2=1}^{n_2-g_2}(Y_{i_1,i_2} - \overline{Y})(Y_{i_1+g_1,i_2+g_2} - \overline{Y})/n$ and $c(g_1,-g_2) = \sum_{i_1=1}^{n_1-g_1}\sum_{i_2=1}^{n_2-g_2}(Y_{i_1,i_2+g_2} - \overline{Y})(Y_{i_1+g_1,i_2} - \overline{Y})/n$, for $g_1,\, g_2 \geq 0$, where $\overline{Y} = \sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2} Y_{i_1,i_2}/n$. Writing $c(g_1,g_2)$ as a quadratic form in $Y$ and using standard results (Cressie, 1993 section 2.4.2), it is easy, given $V$, to calculate $\mathrm{cov}[c(g_1,g_2), c(g_1^\star, g_2^\star)]$, from which it is clear that the $c(g_1,g_2)$ are highly dependent and have non-constant variance.

### 3.1.1 Testing axial symmetry

We can consider basing tests for axial symmetry on the sample differences $F(g_1,g_2) = c(g_1,g_2) - c(g_1,-g_2)$ for $g_1,\, g_2 \geq 1$. Under the null hypothesis $\mathrm{E}[F(g_1,g_2)] = 0\ \forall\, g_1,\, g_2$. However the $F(g_1,g_2)$ have non-constant variance and they are correlated. For $1 \leq g_j \leq a_j$, let $F$ denote the vector containing the $F(g_1,g_2)$ in lexicographic order, and let $\mathrm{cov}(F) = \Sigma$. Since the $c(g_1,g_2)$ are asymptotically normally distributed, $F \sim \mathrm{N}(0,\Sigma)$ and a possible test uses

$$S^{(c)} = F'\hat{\Sigma}^{-1}F.$$

Under axial symmetry, $S^{(c)} \sim \chi^2_a$, asymptotically, with $a = a_1 \cdot a_2$. Notice that estimating $\Sigma$ requires estimating $V$ and, therefore, this test is not completely model-free. We fitted the AR(1)·AR(1) to estimate $V$.

### 3.1.2 Testing separability

Modjeska and Rawlings (1983) suggested testing separability by using the singular value decomposition of the matrix $M$ of sample covariances $c(g_1,g_2)$ for $-(a_1-1) \leq g_1 \leq a_1 - 1$ and $0 \leq g_2 \leq a_2 - 1$. Since $C(g_1,g_2) \propto C(g_1,0)\cdot C(0,g_2)$ under the null hypothesis, the population $M$ has rank one for a separable process. If $\lambda_1$ is the largest singular value of $M$, in absolute value, a test for separability could use:

$$R^{(c)} = \lambda_1^2 / \sum \lambda_i^2.$$

Its distribution needs to be simulated. Note that $R^{(c)} \in [0,1]$, low values leading to rejection of separability.

## 3.2 Tests based on sample spectrum

The main advantage of the sample spectrum is that it has better asymptotic properties than the sample covariances, being asymptotically independent at different frequencies. For $\omega_j$ a multiple of $2\pi/n_j$, the sample spectrum we used is $I(\omega_1,\omega_2) = \sum_{g_1=-n_1+1}^{n_1-1}\sum_{g_2=-n_2+1}^{n_2-1} c(g_1,g_2)\cos(g_1\omega_1 + g_2\omega_2)/(2\pi)^2$.

Then, for $\omega_j \neq 0, \pi$, we have $I(\omega_1, \omega_2)/f(\omega_1, \omega_2) \to$ i.i.d. Exp(1), $\mathrm{E}[I(\omega_1, \omega_2)] \to f(\omega_1, \omega_2)$ and $\mathrm{Var}[I(\omega_1, \omega_2)] \to f^2(\omega_1, \omega_2)$, as $n_1, n_2 \to \infty$. Exact formula for the mean and the variance of the one-dimensional sample spectrum are in Priestley (1981, section 6.1.3).

#### 3.2.1 Testing axial symmetry

Consider the differences $G(\omega_1, \omega_2) = I(\omega_1, \omega_2) - I(\omega_1, -\omega_2)$ (with $\omega_j = 2\pi k_j/n_j$, $k_j = 1, 2, \ldots, b_j$, and $b_j = (n_j - 1)/2$ if $n_j$ is odd, or $b_j = n_j/2 - 1$ if $n_j$ is even). Clearly under axial symmetry $\mathrm{E}[G(\omega_1, \omega_2)] = 0$ and $\mathrm{Var}[G(\omega_1, \omega_2)] \to 2f^2(\omega_1, \omega_2)$ as $n_1, n_2 \to \infty$. Since the $G(\omega_1, \omega_2)$ do not have constant variance, we considered two modifications. Firstly, taking the logarithm of the $I(\omega_1, \omega_2)$ approximately stabilizes the variance and also reduces the non-normality. For $\omega_j \neq 0, \pi$ we then have, asymptotically $\log[I(\omega_1, \omega_2)] \sim \mathrm{Gumbel}\,(\log[f(\omega_1, \omega_2)], 1)$ from which $\mathrm{Var}\,\{\log[I(\omega_1, \omega_2)]\} \to \pi^2/6$ as $n_1, n_2 \to \infty$. Given $b = b_1 \cdot b_2$, a possible test uses:

$$S_1^{(p)} = \overline{D}\sqrt{b}/\sqrt{\pi^2/3},$$

where $\overline{D} = \sum_{\omega_1, \omega_2} D(\omega_1, \omega_2)/b$, $D(\omega_1, \omega_2) = \log[I(\omega_1, \omega_2)] - \log[I(\omega_1, -\omega_2)]$. Assuming $b$ large enough for the central limit theorem to hold, under axial symmetry $S_1^{(p)} \sim \mathrm{N}(0, 1)$. Other tests, for example, the sign test or the Wilcoxon test on the differences $D(\omega_1, \omega_2)$ or $G(\omega_1, \omega_2)$, can also be used.

A second modification is to standardize the differences $G(\omega_1, \omega_2)$, estimating $f(\omega_1, \omega_2)$ by $[I(\omega_1, \omega_2) + I(\omega_1, -\omega_2)]/2$. Thus, if $H(\omega_1, \omega_2)$ denotes $G(\omega_1, \omega_2)/[I(\omega_1, \omega_2) + I(\omega_1, -\omega_2)]$, then another possible test uses

$$S_2^{(p)} = \sqrt{3b}\ \overline{H}.$$

If the $I(\omega_1, \omega_2)$ are independent and exponentially distributed, $H(\omega_1, \omega_2)$ has a uniform distribution in the interval $[-1, 1]$, and the sum converges rapidly to a $\mathrm{N}(0, b/3)$, so $S_2^{(p)} \sim \mathrm{N}(0, 1)$ asymptotically under axial symmetry.

#### 3.2.2 Testing separability

The same idea in section 3.1.2 can be applied to a matrix of $I(\omega_1, \omega_2)$ to test separability (we use $R_1^{(p)}$ to denote this test statistic).

As an alternative, under separability, $\log[I(\omega_1, \omega_2)]$ and $\log[I(\omega_1, -\omega_2)]$ can be regarded as two sample realizations of the same value of the log spectrum $\log[f(\omega_1, \omega_2)] = d + \log[f(\omega_1, 0)] + \log[f(0, \omega_2)]$ where $d$ is a constant. So testing for separability reduces to testing for lack of interaction in a two-way classification table with two realizations in each cell and, $\mathrm{var}(\log[I(\omega_1, \omega_2)])$ being approximately constant, this can be done using the statistic:

$$R_2^{(p)} = SS_{rc}/SS_e$$

where $SS_{rc}$ is the mean of squares due to the interaction, and $SS_e$ is the residual mean of squares. Under separability, and assuming the log differences are approximately normal, $R_2^{(p)} \sim F_{(b_1-1)(b_2-1), b}$, asymptotically.

### 3.3 Model-based tests

Assuming a specific model behind the data allows the use of model-based tests for axial symmetry and separability. The idea is to fit a particular unrestricted model (non-separable or non-symmetric) to the data, then to restrict it to be axially symmetric or separable, imposing constraints on its parameters and to test which model is more appropriate. The comparison can use the generalised likelihood ratio test (GLRT) or tests which are asymptotically equivalent, such as the Wald test or the Score test.

#### 3.3.1 Testing axial symmetry and separability together

Under the assumption $Y \sim \mathrm{N}(0, \sigma^2)$, the log-likelihood can be written as $\ell(\alpha, \sigma^2; y) = -(n/2)\log(2\pi) - (n/2)\log(\sigma^2) - \log \mid V \mid /2 - y'V^{-1}y/(2\sigma^2)$. The GLRT statistic is then

$$S^{(m)} = 2[\ell(\hat{\alpha}, \hat{\sigma}^2; y) - \ell(\hat{\alpha}_0, \hat{\sigma}_0^2; y)].$$

Now, to test for axial symmetry and separability together we can consider the Pickard as the unrestricted model and the AR(1)·AR(1) as the restricted one, with the constraint $\alpha_3 = -\alpha_1\alpha_2$ (see section 2.3). Alternatively, the Wald (Guo and Billard, 1998) and the Score tests could be used, although we found that GLRT is preferred for smaller lattice sizes. Under the null text hypothesis, the three test statistics are asymptotically distributed as a $\chi^2_1$.

#### 3.3.2 Testing separability

Separability can be tested comparing the fit of a $\mathrm{CAR}(2)_{SD}$ with the fit of an AR(1)·AR(1), applying the restriction $\beta_3 = -\beta_1\beta_2$ (see section 2.3) to the parameters of the $\mathrm{CAR}(2)_{SD}$. The GLRT can be expressed this time as:

$$R^{(m)} = 2[\ell(\hat{\beta}, \hat{\sigma}^2; y) - \ell(\hat{\beta}_0, \hat{\sigma}_0^2; y)].$$

## 4 Results and discussion

For the tests on axial symmetry, a good correspondence between the simulated and theoretical distributions was found for $S_1^{(p)}$, $S_2^{(p)}$, $S^{(m)}$ on 11 by 11 or larger lattices. For $S^{(c)}$ with $a_1 = a_2 = 4$, a 15 by 15 lattice was required. Tests $S_1^{(p)}$, $S_2^{(p)}$ are completely general, the main assumption required being stationarity, while $S^{(c)}$, $S^{(m)}$ require partial or complete specification of a model. On the other hand, $S^{(m)}$ and $S^{(c)}$ are more powerful (see Table 1). Note that the power partially depends on how far the Pickard process is from the AR(1)·AR(1), measured by $|\alpha_3 + \alpha_1\alpha_2|$. Also note that $S^{(m)}$ is testing for separability at the same time.

Separability proved to be harder to test. Tests $R^{(c)}$ and $R_1^{(p)}$ are not invariant, their distributions strongly depending on the parameters of the process considered, and also can have values very close to 1 under the alternative hypothesis. Simulations show $R_2^{(p)}$ to be distributed as expected on 11 by 11 lattices, although this test does not seem to be very powerful. Also, it assumes axial symmetry, which therefore should be tested first. Test $R^{(m)}$ is very powerful (see Table 2), although it requires specification of a model.

As well as the tests mentioned, we have also considered some other tests, and considered variants of those mentioned - for example using different $a_j$. We have evaluated the tests on other parameter values than those mentioned, and on some other processes - for example some separable ARMA processes. We have some theoretical results to help explain the power for different parameters under the alternative hypothesis. We have tried the tests on a real data set - a SAR (synthetic aperture radar) image for which the postulated form of the point-spread function implies separability. The theoretical distribution of the data is Exponential, so further work is needed to see how the tests perform for theoretical SAR data. For further details, see Scaccia (2000) and Scaccia & Martin (2002).

| Parameters | $S_1^{(p)}$ | $S_2^{(p)}$ | $S^{(c)}$ | $S^{(m)}$ | $\lvert\alpha_3 + \alpha_1\alpha_2\rvert$ |
|---|---|---|---|---|---|
| $\alpha = (0.1, 0.2, 0.6)$ | 97.4% | 97.1% | 99.9% | 100.0% | 0.62 |
| $\alpha = (0.6, 0.7, -0.8)$ | 87.0% | 86.3% | 99.5% | 100.0% | 0.38 |
| $\alpha = (0.3, 0.4, 0.2)$ | 55.2% | 59.0% | 78.1% | 99.8% | 0.32 |
| $\alpha = (0.1, 0.6, 0.2)$ | 44.6% | 45.8% | 61.2% | 99.3% | 0.26 |
| $\alpha = (0.3, 0.6, -0.15)$ | 5.1% | 5.3% | 5.8% | 6.8% | 0.03 |

**Table 1.** *Simulated power of tests for axial symmetry for different Pickard processes, on a 15 by 15 lattice when the level of the tests is equal to 5%.*

| Parameters | $R_2^{(p)}$ | $R^{(m)}$ | $\lvert\beta_3 + \beta_1\beta_2\rvert$ |
|---|---|---|---|
| $\beta = (0, 0, 0.25)$ | 88.8% | 100.0% | 0.25 |
| $\beta = (0.11, 0.07, 0.16)$ | 13.6% | 98.7% | 0.17 |
| $\beta = (0.2, 0.2, 0.05)$ | 9.1% | 92.7% | 0.09 |
| $\beta = (0.25, 0.25, 0)$ | 8.3% | 86.8% | 0.06 |
| $\beta = (0.3, 0.2, -0.03)$ | 5.3% | 12.3% | 0.03 |

**Table 2.** *Simulated power of tests for separability for different CAR(2)$_{SD}$ processes, on a 15 by 15 lattice when the level of the tests is equal to 5%.*

## References


Balram, N. and Moura, J.M.F. (1993). Noncausal Gauss Markov random fields: parameter structure and estimation. *IEEE Trans. Inform. Theory*, **39**, 1333-1355.

Cressie, N.A.C. (1993). *Statistics for Spatial Data.* New York: Wiley.

Guo, J.H. and Billard, L. (1998). Some inference results for causal autoregressive processes on a plane. *J. of Time Ser. An.*, **19**, 681-691.

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 Meth.*. **19**, 55-81.

Martin, R.J. (1996). Some results on unilateral ARMA lattice processes. *J. Statist. Plann. Inf.*. **50**, 395-411.

Modjeska, J.S. and Rawlings, J.O. (1983). Spatial correlation analysis of uniformity data. *Biometrics*, **39**, 373-384.

Pickard, D.K. (1980). Unilateral Markov fields. *Adv. Appl. Prob.*, **12**, 655-671.

Priestley, M.B. (1981). *Spectral Analysis and Time Series.* London: Academic Press.

Scaccia, L. (2000). *Testing for Simplification in Spatial Models.* Ph.D. thesis, University of Perugia, Italy.

Scaccia, L. and Martin R.J. (2002). Testing axial symmetry and separability of lattice processes. Preprint.

Tory, E.M. and Pickard, D.K. (1992). Unilateral Gaussian fields. *Adv. Appl. Prob.*, **24**, 95-112.