Statistica Sinica **18**(2008), 617-631

MINIMUM DISTANCE INFERENCE IN UNILATERAL AUTOREGRESSIVE LATTICE PROCESSES

Marc G. Genton[1,2] and Hira L. Koul[3]

[1]*University of Geneva,* [2]*Texas A&M University and* [3]*Michigan State University*

Abstract: This paper discusses two classes of minimum distance estimators of the underlying parameters and their robust variants in unilateral autoregressive lattice models. The paper also contains an asymptotically distribution free test for symmetry of the error distribution and a goodness-of-fit test for fitting an error distribution. A lack-of-fit test for the hypothesis that the given process is doubly geometric based on the least absolute deviation residuals is also briefly analyzed. A simulation study that investigates some small sample properties of the proposed estimators and their robustness is included. It shows that some of the proposed estimators are more efficient than the least squares estimator at non-normal error distributions. We also study the empirical level and power of the test of a doubly geometric process at various error distributions. The proposed methodology is then applied to a data set of yields from an agricultural experiment.

Key words and phrases: Doubly geometric process, efficiency, non-Gaussian spatial model, Pickard process, quadrant autoregressive process, robustness, weighted empirical processes.

1. Introduction

A unilateral autoregressive process in the plane consists of the observations $X_{i,j}$ satisfying

$$X_{i,j} = \alpha_{10}X_{i-1,j} + \alpha_{01}X_{i,j-1} + \alpha_{11}X_{i-1,j-1} + \varepsilon_{i,j}, \quad i, j = 0, \pm 1, \pm 2, \ldots. \quad (1.1)$$

Here the errors $\varepsilon_{i,j}$ are assumed to be i.i.d. according to a distribution function (d.f.) F having zero mean and finite and positive variance σ^2. Moreover, for each i, j, $\varepsilon_{i,j}$ is assumed to be independent of the past random variables (r.v.'s) $\{X_{r,s};\ r \le i,\ s \le j,\ (r,s) \ne (i,j)\}$. The process (1.1) is sometimes also referred to as a first-order quadrant autoregressive (QAR(1,1)) process (Tjøstheim (1978, 1983)) or a Pickard process (Pickard (1980) and Tory and Pickard (1992)). From now on we write QAR for the QAR(1,1) process of the form (1.1). When $\alpha_{11} = 0$, it is called a nearest neighbor (NN) process (Bartlett (1968, 1971) and Martin (1979, 1996, 1997)). The QAR process with $\alpha_{11} = -\alpha_{10}\alpha_{01}$ is called a doubly geometric (DG) process (Martin (1979)). It is the product of two autoregressive time series processes of order 1.

Tory and Pickard (1992) showed that the QAR process is stationary if

$$|\alpha_{10}+\alpha_{01}|<1-\alpha_{11}, \qquad |\alpha_{10}-\alpha_{01}|<1+\alpha_{11}. \tag{1.2}$$

These conditions reduce to $|\alpha_{10}|+|\alpha_{01}|<1$ for the NN process. The process (1.1) also has an infinite moving average representation (Martin (1997)) given by

$$X_{i,j}=\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\beta_{rs}\varepsilon_{i-r,j-s}, \quad \beta_{rs}=\sum_{k=0}^{\min(r,s)}\frac{(r+s-k)!}{(r-k)!(s-k)!k!}\alpha_{10}^{r-k}\alpha_{01}^{s-k}\alpha_{11}^{k}. \tag{1.3}$$

Unilateral autoregressive processes (1.1) are important for two main reasons. First, they are useful for practical modeling because they include a fairly flexible range of spatial correlation structures (see, e.g., Besag (1972) and Basu and Reinsel (1993)), while having some very attractive inferential properties due to their unilateral structure. Indeed, QAR processes are a natural generalization of autoregressive time series from $\mathbb{R}$ to $\mathbb{R}^d$, $d\geq 2$. In the plane, i.e., when $d=2$, these processes are of particular interest for agricultural and environmental applications, as well as digital filtering in image analysis. They are especially appropriate when there is evidence of a spatial movement over the plane in one direction, such as with environmental pollutants transported by winds or ocean currents, or with the spread of a disease. Moreover, they are simple alternatives to more complicated simultaneous autoregressive (SAR) and conditional autoregressive (CAR) models, see, e.g., Whittle (1954), Bartlett (1971) and Besag (1974). Secondly, QAR processes are the building blocks for inference in SAR models because they can be used as auxiliary models in an indirect inferential procedure, see de Luna and Genton (2002), as well as Genton and Ronchetti (2003). Therefore, it is important to construct good inferential procedures for QAR processes.

Some inference procedures based on the least squares (LS) methodology and/or maximum likelihood (ML) methodology under the additional assumption of normality of the errors $\varepsilon_{i,j}$ in the QAR model (1.1) are available in the literature, see, e.g., Guo and Billard (1998) and Basu and Reinsel (1993). But, as is well known, these inference procedures are non-robust to outliers in the errors. The development of inference for this model with non-Gaussian continuous error distributions seems to have lagged behind.

Here, we describe two classes of minimum distance estimators (M.D.E.'s) and their robust variants for the parameter vector $\alpha := (\alpha_{10},\, \alpha_{01},\, \alpha_{11})'$, based on the observations $\{X_{i,j}\}$, $i,j=0,1,\ldots,n$ on a regular square lattice, without assuming the knowledge of the error d.f. F. Extensions to $n_1\times n_2$ rectangular lattices are straightforward. Some members of these classes are asymptotically more efficient than the classical LSE at non-Gaussian errors, while also being robust against outliers in the innovations.

This paper is organized as follows. In Section 2.1 we describe M.D.E.'s and their robust modification for QAR processes that do not require the specification of the error distribution. Their asymptotic normality is stated in Propositions 2.1 and 2.2. An asymptotically distribution free test of the symmetry of the error distribution and a goodness-of-fit test pertaining to the error d.f. F is given in Section 2.2. A lack-of-fit test for fitting a DG process, i.e., for testing $H_0 : \alpha_{11} = -\alpha_{10}\alpha_{01}$, based on the least absolute deviation (LAD) residuals, appears in Section 2.3. In Section 3, we investigate numerically the small sample properties of the proposed estimators and their robustness, showing that they are more efficient than the LSE at non-normal error distributions. We also study the empirical level and power of the LAD test for the above H_0 at various error distributions. In Section 4, the proposed methodology is illustrated on a data set of yields from an agricultural experiment.

2. Main Results

This section contains some new estimation procedures for QAR processes based on minimum distance and their asymptotic normality, an asymptotically distribution free test for testing the symmetry of the errors, a goodness-of-fit test for fitting an error d.f. up to an unknown scale parameter, and a test of the hypothesis $H_0 : \alpha_{11} = -\alpha_{10}\alpha_{01}$ based on the LAD residuals.

2.1. Minimum distance estimation

In this sub-section we introduce the two classes of M.D.E.'s of α. For one of these classes, we need to assume the symmetry of the error d.f., while for the other no such assumption is needed, but the corresponding estimator is based on the residual ranks. To proceed further, define, for a $t := (t_1, t_2, t_3)' \in \mathbb{R}^3$,

$$Y_{i,j} := (X_{i-1,j}, X_{i,j-1}, X_{i-1,j-1})', \quad \varepsilon_{i,j}(t) := X_{i,j} - t'Y_{i,j}. \tag{2.1}$$

Now, suppose the error d.f. F is symmetric around zero. Then the r.v.'s $\{\varepsilon_{i,j}(\alpha);\, i,\, j = 1, \ldots, n\}$ and their reflections around the origin $\{-\varepsilon_{i,j}(\alpha);\, i, j = 1, \ldots, n\}$ have the same distribution. This motivates us to define an estimator of α as a value of t that minimizes some kind of a dispersion between $\{\varepsilon_{i,j}(t);\, i, j = 1, \ldots, n\}$ and $\{-\varepsilon_{i,j}(t);\, i, j = 1, \ldots, n\}$. A class of useful dispersions is obtained as follows. Let G be a nondecreasing right continuous function on $\mathbb{R}$, and define

$$K(x,t) := n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} Y_{i,j}\{I(\varepsilon_{i,j}(t) \le x) - I(-\varepsilon_{i,j}(t) \le x)\}, \quad x \in \mathbb{R},$$

$$\mathcal{M}(t) := \int \|K(x,t)\|^2 G(dx), \qquad t \in \mathbb{R}^3.$$

Because of the assumed symmetry of F, the conditional expectation of the (i, j)th summand of $K(x, \alpha)$ given $\{Y_{r,s};\ r \le i,\ s \le j,\ (r, s) \neq (i, j)\}$, is zero for all $x \in \mathbb{R}$, and hence $E(K(x, \alpha)) \equiv 0$. The dispersion $\mathcal{M}(t)$ is an analogue of the Cramér-von Mises distance between the weighted empirical processes of $\{\varepsilon_{i,j}(t)\}$ and $\{-\varepsilon_{i,j}(t)\}$. We are thus motivated to define a class of M.D.E.'s of α, one for each G, by the relation

$$\widehat{\alpha} := \mathrm{arginf}_t \mathcal{M}(t).$$

Write $\widehat{\alpha}_I$ for $\widehat{\alpha}$ when $G(x) \equiv x$. This estimator is an extension of the median of pairwise means estimator to the current set up. The $\widehat{\alpha}$ corresponding to the G degenerate at zero gives the LAD estimator.

Now suppose F is not symmetric but still unknown. In this case we use ranks of the residuals to estimate α as follows. Let L be a d.f. on the interval $[0, 1]$. Define

$$R_{i,j}(t) := \sum_{r=1}^{n}\sum_{s=1}^{n} I(\varepsilon_{r,s}(t) \le \varepsilon_{i,j}(t)), \quad 1 \le i, j \le n; \quad \bar{Y} := n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} Y_{i,j},$$

$$T(u, t) := n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}(Y_{i,j} - \bar{Y}) I(R_{i,j}(t) \le nu), \qquad 0 \le u \le 1,$$

$$\mathcal{R}(t) := \int_0^1 \|T(u, t)\|^2 dL(u), \quad t \in \mathbb{R}^3; \qquad \widetilde{\alpha} := \mathrm{arginf}_t \mathcal{R}(t).$$

Write $\widetilde{\alpha}_I$ for $\widetilde{\alpha}$ when $L(u) \equiv u$.

As argued in Koul (1986, 2002), in connection with the single time index autoregressive time series, the Pitman asymptotic relative efficiency (ARE) of $\widehat{\alpha}_I$ compared to the LSE is the same as that of the median of the pairwise means relative to the sample mean at a large class of error distributions. The estimator $\widetilde{\alpha}_I$ also has a very desirable ARE property. For example, with double exponential errors its ARE, relative to the LSE and $\widehat{\alpha}_I$, is 1.67 and 1.11, respectively. It is thus desirable to extend the applicability of these procedures to the current spatial set-up.

We now state additional assumptions needed for the asymptotic normality of the above M.D.E.'s. To keep their statements relatively transparent, assume G is continuous and symmetric around zero and that F has a density f. The symmetry of G is equivalent to

$$G(0) - G(-x) = G(x) - G(0), \quad \forall\ x > 0. \tag{2.2}$$

Furthermore, we need the following assumptions:

$$\int_0^\infty (1-F)dG < \infty; \tag{2.3}$$

$$0 < \int f^r dG < \infty, \quad r = 1, 2; \tag{2.4}$$

$$\sup_{|s|\le\delta} \int f^r(x+s)G(dx) < \infty, \text{ for some } \delta > 0,\ r = 1, 2; \tag{2.5}$$

$$\int (f(x+s) - f(x))^2 G(dx) \longrightarrow 0, \quad \text{as } |s| \to 0. \tag{2.6}$$

Now, let $\gamma(h,k) := E(X_{i,j}X_{i-h,j-k})$, and

$$\Gamma := E(Y_{i,j}Y'_{i,j}) = \begin{pmatrix} \gamma(0,0) & \gamma(-1,1) & \gamma(0,1) \\ \gamma(-1,1) & \gamma(0,0) & \gamma(1,0) \\ \gamma(0,1) & \gamma(1,0) & \gamma(0,0) \end{pmatrix}.$$

By the results established in Tjøstheim (1983), we obtain that

$$n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} Y_{i,j}Y'_{i,j} = \Gamma + o_p(1).$$

Let $\rho(x,y) := F(x \wedge y) - F(x)F(y)$, $x,\ y \in \mathbb{R}$, and $\Sigma := \Gamma - E(\bar{Y}\bar{Y}')$. Using the general methodology of Chapter 5 in Koul (2002), as used in Chapter 7 there, we obtain the following proposition.

Proposition 2.1. *Suppose the QAR model* (1.1) *holds. In addition, assume that* Γ *is positive definite, the error d.f.* F *is symmetric around zero, and has density* f *such that* (F,G) *together satisfy the conditions* (2.2)−(2.6), *and for any* $0 < c < \infty$,

$$\int E\{\|Y_{1,1}\|^2 \big| F(x + n^{-1/2}c\|Y_{1,1}\|) - F(x - n^{-1/2}c\|Y_{1,1}\|)\big|\} dG(x) \to 0. \tag{2.7}$$

Then, $n(\widehat{\alpha} - \alpha) \to_d \mathcal{N}_3(0, \Gamma^{-1}\tau^2)$, *where* $\tau^2 := \int\int \rho(x,y)f(x)f(y)dG(x)dG(y)/ (\int f^2 dG)^2$.

We now briefly discuss the assumptions (2.3)−(2.6) and (2.7). Let ε be a r.v. with d.f. F. The condition (2.3) is equivalent to requiring the finiteness of $E(|G(\varepsilon) - G(0)|)$. If G has a bounded Lebesgue density g, then $E(|\varepsilon|) < \infty$ and (2.4) with $r = 2$ imply (2.3)−(2.6). To see this use the fact that (2.4) with $r = 2$ implies that f is shift continuous in mean square with respect to the measure $g(x)dx$. In particular, in case $G(x) \equiv x$, $E(|\varepsilon|) < \infty$ and $\int f^2(x)dx < \infty$ imply these conditions.

If G is a finite measure, then by the Dominated Convergence Theorem, the continuity of F and $E(\|Y_{1,1}\|^2) < \infty$ imply (2.7). In particular, if F is known and one chooses $G = F$, then the latter two conditions imply (2.7). In case $G(x) \equiv x$ and F has a Lebesgue density then, by the Fubini Theorem, the left hand side of (2.7) equals $2n^{-1/2}c\,E(\|Y_{1,1}\|^3)$. Thus in this case $E(\|Y_{1,1}\|^3) < \infty$ implies (2.7).

These observations lead to the following result for the estimators $\widehat{\alpha}_I$ and $\widehat{\alpha}_F$.

Corollary 2.1. *Suppose the QAR model* (1.1) *holds,* Γ *is positive definite, the error d.f.* F *is symmetric around zero and has density* f*. Then the following hold.*

(a) *If* $E(\|Y_{1,1}\|^3) < \infty$ *and* $\int f^2(x)dx < \infty$, *then*

$$n(\widehat{\alpha}_I - \alpha) \to_d \mathcal{N}_3\Bigg(0, \frac{\Gamma^{-1}}{12\Big(\int f^2(x)dx\Big)^2}\Bigg).$$

(b) *If* F *is known and has a bounded density* f, *then*

$$n(\widehat{\alpha}_F - \alpha) \to_d \mathcal{N}_3(0, \Gamma^{-1}\tau_F^2), \quad \tau_F^2 := \frac{2\int_{-\infty}^{\infty}\int_{-\infty}^{y} F(x)(1-F(y))f^2(x)f^2(y)dxdy}{\Big(\int f^3(x)dx\Big)^2}.$$

Next, we state an asymptotic normality result about the rank based estimator $\widetilde{\alpha}$. Again, this is proved using the methods in Koul (2002).

Proposition 2.2. *Suppose the QAR model* (1.1) *holds. In addition, assume that* Σ *is positive definite,* F *is strictly increasing on* $\mathbb{R}$ *and has a uniformly continuous density* f*. Then,*

$$n(\widetilde{\alpha} - \alpha) \to_d \mathcal{N}_3(0, \Sigma^{-1}\tau_r^2), \quad \tau_r^2 := \frac{\int\int \rho(x,y)f(x)f(y)dL(F(x))dL(F(y))}{\Big(\int f^2 dL(F)\Big)^2}.$$

We now discuss a robust variant of $\widehat{\alpha}$. Because the weights in K are unbounded, unlike in the regression set-up, the corresponding M.D.E.'s are not robust against innovative or additive outliers in autoregressive lattice processes. One way to overcome this deficiency in $\widehat{\alpha}$ is to use bounded weights that are concordant or discordant with the $Y_{i,j}$'s. A natural choice is to replace the weights $Y_{i,j}$ in K by $h(Y_{i,j})$, where

$$h(y) = yI(\|y\| \le k) + k\frac{y}{\|y\|}I(\|y\| > k), \quad y \in \mathbb{R}^3,$$

with $k > 0$ a known constant. More precisely, let

$$K_h(x,t) := n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} h(Y_{i,j})\{I(\varepsilon_{i,j}(t) \le x) - I(-\varepsilon_{i,j}(t) \le x)\}, \qquad x \in \mathbb{R},$$

$$\mathcal{M}_h(t) := \int \|K_h(x,t)\|^2 G(dx), \quad t \in \mathbb{R}^3; \qquad \mathcal{B}_n := n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} h(Y_{i,j})Y'_{i,j}.$$

Then $\widehat{a}_h := \text{argmin}_t \mathcal{M}_h(t)$ is a robust M.D.E.'s of α. Observe that by the Ergodic Theorem, $\mathcal{B}_n \to \mathcal{B} := E(h(Y_{1,1})Y'_{1,1})$, a.s. Also, by the definition of h, $\mathcal{B}$ is positive definite. Under the same conditions on (F, G) as in Proposition 2.1, we obtain $n(\widehat{a}_h - \alpha) \to_d \mathcal{N}_3(0, \mathcal{B}^{-1}\Gamma\mathcal{B}^{-1}\tau^2)$. A robust variant of $\widetilde{\alpha}$ can be obtained similarly.

2.2. Tests of symmetry and goodness-of-fit

Consider the problem of testing $H_s : F$ is symmetric around 0, against the alternative that it is not. For $t \in \mathbb{R}^3$, let $H_n(y,t) := \sum_{i=1}^n \sum_{j=1}^n I(|\varepsilon_{i,j}(t)| \le y)/n^2$, $\hat{H}_n(y) := H_n(y, \hat{\alpha})$,

$$S_n(y,t) := n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[I(\varepsilon_{i,j}(t) \le y) + I(\varepsilon_{i,j}(t) \le -y) - 1\right],$$

$$\hat{S}_n(y) := S_n(y, \hat{\alpha}), \; y > 0,$$

$$\mathcal{S}_n := \sup_{y>0} |\hat{S}_n(y)|, \qquad \mathcal{C}_n := \int_0^\infty \hat{S}_n^2(y) dL(\hat{H}_n(y)),$$

where $\hat{\alpha}$ is an estimator of α and L is a continuous d.f. on $[0,1]$, symmetric around $1/2$. The statistics $\mathcal{S}_n$ and $\mathcal{C}_n$ are, respectively, the analogues of the Smirnov (1947) and Cramér - von Mises tests for testing H_s in the current set-up. From the general theory developed in Chapters 6 and 7 in Koul (2002), we obtain that if $E(\|Y_{1,1}\|^2 + \varepsilon_{1,1}^2) < \infty$, $n\|\hat{\alpha} - \alpha\| = O_p(1)$ under H_s, F has a uniformly continuous density f and, if H_s holds, $\mathcal{S}_n \to_d \sup_{0 \le u \le 1} |W(u)|$, and $\mathcal{C}_n \to_d \int_0^1 W^2(u) dL(u)$, where W is the standard Brownian motion. It is interesting to note that, unlike in the regression model, the estimation of α has little effect on the asymptotic null distribution of these statistics.

Next, we describe a test of goodness-of-fit of an error distribution. Let F_0 be a known d.f. with mean zero and variance one. Consider the problem of testing $\tilde{H}_0 : F(x) = F_0(x/\sigma)$, for all $x \in \mathbb{R}$ and some $\sigma > 0$. Let $\hat{\sigma}$ denote an n-consistent estimator of σ and $\hat{F}_n$ denote the empirical d.f. of the standardized residuals $\widehat{\varepsilon}_{i,j} := \varepsilon_{i,j}(\hat{\alpha})/\hat{\sigma}$. The proposed test of $\tilde{H}_0$ is based on the process $Z_n(x) := n[\hat{F}_n(x) - F_0(x)]$; $x \in \mathbb{R}$. Arguing as in Koul (2002), one can verify that if F_0 has a uniformly continuous density f_0 with $\sup_x |x f_0(x)| < \infty$, if f_0 has finite and positive Fisher information for a scale parameter, and if $\hat{\sigma}$ is an asymptotically efficient estimator of σ at F_0, then under $\tilde{H}_0$, Z_n converges weakly to a mean zero Gaussian process with the covariance function $F_0(x)\wedge F_0(y)-F_0(x)F_0(y)-b^{-1}xyf_0(x)f_0(y)$, for x, y. Consequently, the asymptotic null distribution of $\sup_x|Z_n(x)|$ is difficult to determine even when one knows F_0. One possible approach is to use the Khmaladze (1981) transformation whose asymptotic null distribution is the same as that of $W(F_0)$; see also Khmaladze and Koul (2004). Alternatively, we propose the use of a parametric bootstrap to construct the distribution of $\sup_x|Z_n(x)|$ under $\tilde{H}_0$. We note that the case of a Gaussian F_0 is of particular interest in practice.

2.3. LAD test of a doubly geometric process

Under the hypothesis $H_0 : \alpha_{11} = -\alpha_{10}\alpha_{01}$, the process (1.1) is known as a doubly geometric (DG) process. In the case of normal errors, Guo and Billard (1998) provide an asymptotic test of this hypothesis using the LS residuals; see also Scaccia and Martin (2002, 2005). Here, we describe a test of H_0 based on the LAD residuals when the error d.f. F is not necessarily known but has a continuous density f in an open neighborhood of 0 with $f(0) > 0$. To proceed further, let

$$\mu_{i,j}(s) := s_1X_{i-1,j} + s_2X_{i,j-1} - s_1s_2X_{i-1,j-1}, \quad s = (s_1, s_2)' \in \mathbb{R}^2,$$
$$M_0(s) := \sum_{i=1}^{n}\sum_{j=1}^{n}|X_{i,j} - \mu_{i,j}(s)|, \qquad M(t) := \sum_{i=1}^{n}\sum_{j=1}^{n}|\varepsilon_{i,j}(t)|, \quad t \in \mathbb{R}^3.$$

The proposed test of H_0 is based on the difference $\mathcal{D} := \inf_{s\in\mathbb{R}^2} M_0(s) - \inf_{t\in\mathbb{R}^3} M(t)$. To state its asymptotic null distribution, let $\alpha_0 = (\alpha_{01}, \alpha_{10})'$,

$$\dot{\mu}_{i,j} := (X_{i-1,j} - \alpha_{10}X_{i-1,j-1},\, X_{i,j-1} - \alpha_{01}X_{i-1,j-1})', \quad \Gamma_0 := E_0(\dot{\mu}_{1,1}\dot{\mu}_{1,1}'),$$
$$T_0 := n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\dot{\mu}_{i,j}\mathrm{sign}(\varepsilon_{i,j}), \qquad T := n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}Y_{i,j}\mathrm{sign}(\varepsilon_{i,j}).$$

Then, using the methods developed in Koul (2002, Chap. 5, Chap. 8), one obtains that under H_0,

$$\mathcal{D} = \frac{1}{4f(0)}[T'\Gamma^{-1}T - T_0'\Gamma_0^{-1}T_0] + o_p(1).$$

This in turn implies that $4f(0)\mathcal{D} \to_d \chi_1^2$, under H_0. Thus to implement this test, at least for large samples, one needs a consistent estimator of $f(0)$. One such estimator is $\hat{f}_n := [F_n(h_n) - F_n(-h_n)]/(2h_n)$, where h_n is a sequence of window widths such that $h_n \to 0$, $nh_n \to \infty$, and F_n is the empirical d.f. of the residuals $X_{i,j} - \mu_{i,j}(\hat{\alpha}_0)$, with $\hat{\alpha}_0$ being any n-consistent estimator of α_0 under H_0. One may use the M.D.E. $\widehat{\alpha}_0$. Consequently, the LAD test that rejects H_0 whenever $\widehat{\mathcal{D}} := 4\hat{f}_n\mathcal{D}$ exceeds a critical value obtained from the χ_1^2 distribution is asymptotically distribution free.

3. Simulations

In this section we present some finite sample simulation results. For computational purposes, we first describe an alternative expression for the dispersion $\mathcal{M}$. Using the continuity of G and arguing as in Koul (2002, (5.3.12)), one can write

$$\mathcal{M}(t) = \frac{1}{n^2} \sum_{i,j=1}^{n} \sum_{k,l=1}^{n} Y_{i,j}' Y_{k,l} \Big[|G(\varepsilon_{i,j}(t)) - G(-\varepsilon_{k,l}(t))| - |G(\varepsilon_{i,j}(t)) - G(\varepsilon_{k,l}(t))| \Big].$$

The asymptotic distribution of $\widehat{\alpha}$ corresponding to a given G is the same as that of the estimator $\widetilde{\alpha}$ with $L(F) \equiv G$. Also note that the $\widehat{\alpha}$ with G the empirical d.f. of $\{\varepsilon_{i,j}(t)\}$ is equivalent to $\widetilde{\alpha}_I$. Thus its asymptotic behavior will be similar to that of $\widehat{\alpha}$ with $G = F$, assuming F is known. For this reason, in the simulations below, we only simulated $\widehat{\alpha}$ for the two integrating measures G, viz, $G(x) \equiv x$ and $G = F$.

3.1. Nearest neighbor process

Consider the setting of an NN process, that is (1.1) with $\alpha_{11} = 0$, and further assume $\alpha_{10} = \alpha_{01} = \gamma$. The condition (1.2) for stationarity reduces to $|\gamma| < 1/2$. The LSE of γ in this case is $\hat{\gamma}_{LS} = \sum_{i=1}^n \sum_{j=1}^n X_{i,j}(X_{i-1,j} + X_{i,j-1}) \big/ \sum_{i=1}^n \sum_{j=1}^n (X_{i-1,j} + X_{i,j-1})^2$, which is also asymptotically equivalent to the MLE of γ when the error d.f. F is normal, see Ord (1975).

We investigate the M.D.E.'s of γ when $G(x) = x$ (denoted by $\hat{\gamma}_I$) and $G(x) = F(x)$ (denoted by $\hat{\gamma}_F$), and when F is a normal, logistic, Laplace, and Cauchy distribution, respectively. We set $n = 10$, $\gamma = 0.1$ and $\sigma^2 = 1$. We generate 500 simulation runs of this experiment using the fast and exact algorithm proposed by Martin (1996, p.400) for general QAR processes. We observe that all the estimators are approximately unbiased. The empirical variances of $\hat{\gamma}_I$ and $\hat{\gamma}_F$ are slightly larger than the empirical variance of the LSE at the normal distribution, but smaller at all other heavier tail error distributions considered here. Table 1 reports the relative efficiencies (RE's) of $\hat{\gamma}_I$ and $\hat{\gamma}_F$ to $\hat{\gamma}_{LS}$.

Next, we briefly explore the robustness properties of $\hat{\gamma}_I$, its robust version $R\hat{\gamma}_I$ with $k = 2.5$ defined above, and the LSE. The latter is well-known to have a lack of robustness reflected through a breakdown-point of zero, as described by Genton (2003) in the setting of QAR processes. Indeed, a single outlying

Table 1. RE's of $\hat{\gamma}_I$ and $\hat{\gamma}_F$ to $\hat{\gamma}_{LS}$.

	Normal	Logistic	Laplace	Cauchy
$\hat{\gamma}_I$	0.93	1.12	1.38	3.02
$\hat{\gamma}_F$	0.88	1.11	1.49	9.71

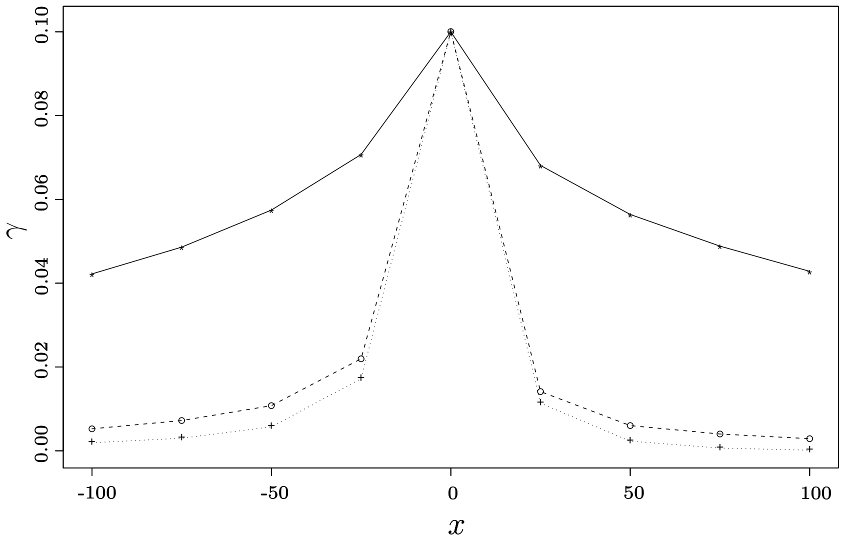

Figure 1. Average effect over 200 replicates of letting $X_{5,5} = x$, $x \in [-100, 100]$, on $\hat{\gamma}_I$ (circles), $R\hat{\gamma}_I$ (stars) and $\hat{\gamma}_{LS}$ (pluses) of γ in a NN process on a grid with $n = 10$, $\gamma = 0.1$, and $N(0, 1)$ errors.

observation becoming arbitrarily large pushes the estimator to zero. Note that this effect is typical of time series and spatial problems where the estimator $\hat{\gamma}_{LS}$ is pushed toward the center of the parameter space rather than to the edge, $\pm 1/2$ in our particular setting of the NN process with $\alpha_{10} = \alpha_{01} = \gamma$. In the setting described in the previous paragraph with a normal d.f. F, we consider 200 simulation runs and select $X_{5,5}$ as an outlying value. Figure 1 depicts the average effect of letting $X_{5,5} = x$, where x varies in the interval $[-100, 100]$, on the $\hat{\gamma}_I$ (circles and dashed lines), the $R\hat{\gamma}_I$ (stars and solid lines), and $\hat{\gamma}_{LS}$ (pluses and dotted lines). All estimators are approximately unbiased when there is no outlier. When the magnitude of the outlier increases, $\hat{\gamma}_I$ and $\hat{\gamma}_{LS}$ are pushed toward zero, even faster so for $\hat{\gamma}_{LS}$, whereas $R\hat{\gamma}_I$ is robust to the outlier. A smaller value of k produces more robustness of $R\hat{\gamma}_I$. The choice $k = 2.5$, suggested by Hampel, Ronchetti, Rousseeuw and Stahel (1986), offers a balance between robustness and efficiency.

3.2. Quadrant autoregressive process

We investigate the performance of the M.D.E. of a QAR process by means of a simulation study similar to the one used by Basu and Reinsel (1993, p.641, Table 1) for the normal ML estimator. Specifically, we simulate 500 realizations of a QAR process with normal errors on a grid with $n = 8$. Three sets of α values are considered: (i) $\alpha = (0.3, 0.2, 0.2)'$; (ii) $\alpha = (0.5, 0.3, 0.1)'$; (iii) $\alpha = (0.8, 0.7, -0.6)'$.

Table 2 reports average values (mean) and standard deviations (sd) of the QAR model parameter estimates over the 500 simulations. The average values of the 500 estimated standard deviations (est. sd) based on asymptotics are also provided. The estimation methods are ML (from Basu and Reinsel (1993)), and $\widehat{\alpha}_I$ and its asymptotic properties derived in Corollary 2.1. We see that the average values of the estimated standard deviations are in good agreement with the standard deviations of the QAR model parameter estimates. The performances of ML and $\widehat{\alpha}_I$ are very similar in this setting of a normal error d.f., but the efficiency of $\widehat{\alpha}_I$ will be much better in the case of error d.f.'s with heavier tails, as suggested by Table 1. ML estimators for QAR models with heavy tailed error d.f.'s would also be very difficult to implement.

3.3. Doubly geometric process

Here our goal is to illustrate some finite sample properties of the LAD test for testing $H_0 : \alpha_{11} = -\alpha_{10}\alpha_{01}$, i.e., (1.1) is a DG process. In this simulation we chose $\alpha_{10} = \alpha_{01} = 0.5$, and $n = 10$. When the process is not DG, the condition (1.2) for stationarity implies that $-1 < \alpha_{11} < 0$, and we study the power of the LAD test under the alternatives $\alpha_{11} = -0.9, -0.8, \ldots, -0.1$. The error d.f. F is taken to be normal, logistic, Laplace, and Cauchy. The nominal level is set

Table 2. Averages values (mean) and standard deviations (sd) of QAR model parameter estimates and the average values of the estimated standard deviations (est. sd) over 500 simulations in the settings (i)-(iii): ML (from Basu and Reinsel (1993)) and $\widehat{\alpha}_I$.

Estimator and setting	α_{10} mean	α_{10} sd	α_{10} est. sd	α_{01} mean	α_{01} sd	α_{01} est. sd	α_{11} mean	α_{11} sd	α_{11} est. sd
ML (i)	0.285	0.127	0.118	0.187	0.123	0.122	0.184	0.145	0.137
ML (ii)	0.471	0.116	0.104	0.281	0.123	0.118	0.098	0.144	0.140
ML (iii)	0.774	0.084	0.069	0.672	0.095	0.088	−0.576	0.106	0.105
$\widehat{\alpha}_I$ (i)	0.288	0.115	0.121	0.183	0.114	0.124	0.196	0.129	0.129
$\widehat{\alpha}_I$ (ii)	0.471	0.104	0.108	0.278	0.120	0.120	0.114	0.139	0.131
$\widehat{\alpha}_I$ (iii)	0.772	0.091	0.082	0.682	0.106	0.095	−0.570	0.132	0.108

to 5%. The value $f(0)$ is estimated by the slope of the linear regression of F_n, the empirical d.f. of the residuals under H_0, on the residuals with magnitude smaller than 0.1.

Empirical values (in %) of significance level and power are calculated over 1,000 simulations, and reported in Table 3. We see that the empirical level of our test is close to the 5% nominal level across the d.f.'s F considered in this simulation, except for the Cauchy. The empirical power reveals that departures from the null hypothesis of the order of 0.01-0.02 can be detected in this setting. Note that the test developed by Guo and Billard (1998) is based on the assumption of normal errors. In contrast, the LAD test does not assume the knowledge of the error distribution and is seen to have higher power when the errors have heavier tails than the Gaussian tails in this simulation study.

Table 3. Empirical values (in %) of significance level (in bold) and power for the LAD test of a DG process at a 5% nominal level in the setting described in Section 3.3.

error d.f. F	$\alpha_{11} = -0.1$	$\alpha_{11} = -0.2$	$\alpha_{11} = -0.25$	$\alpha_{11} = -0.3$	$\alpha_{11} = -0.4$	$\alpha_{11} = -0.5$	$\alpha_{11} = -0.6$	$\alpha_{11} = -0.7$	$\alpha_{11} = -0.8$	$\alpha_{11} = -0.9$
Normal	43.6	7.1	**4.7**	8.6	33.3	74.3	96.7	99.8	100.0	100.0
Logistic	50.1	7.6	**4.2**	9.1	41.1	84.3	97.9	99.7	100.0	100.0
Laplace	72.5	10.6	**5.4**	12.1	62.5	95.4	99.7	100.0	100.0	100.0
Cauchy	99.6	83.2	**6.8**	84.5	99.2	99.8	99.8	99.8	99.3	100.0

4. Application to an Agricultural Experiment

We consider a data set on a 7×28 regular grid of the yield of barley (in kg) from an agricultural experiment in the UK. The data were first analyzed by Kempton and Howes (1981), and then by Basu and Reinsel (1993, Sec. 5) in the context of unilateral spatial models. They argued that a model of the form $W_{i,j} = \beta_0 + X_{i,j}$, where $W_{i,j}$ denotes the yield of barley at location (i, j) and $X_{i,j}$ follows a QAR model (1.1), possibly doubly geometric, would be appropriate. We apply our M.D.E.'s to this data set.

Table 4 summarizes the estimates of the QAR and DG models fitted to the yield of barley data by ML assuming a normal error d.f. (values from Basu and Reinsel (1993)) and by the M.D. method when $G(x) = x$, assuming only the symmetry of the errors. The constant β_0 is estimated iteratively by the sample mean of $W_{i,j} - \hat{\alpha}' Y_{i,j}$, where $\hat{\alpha}$ is either MLE or $\widehat{\alpha}_I$. The MLE and $\widehat{\alpha}_I$ for the QAR model parameters are observed to be rather similar. The test of symmetry of the error d.f. F described in Section 2.2 yields $\mathcal{S}_n = 0.864$ which, compared to the asymptotic 5% critical value of 2.245, suggests that the null hypothesis of F being symmetric around zero cannot be rejected. When we applied the test of normality

Table 4. QAR and DG models fitted to the yield of barley data: by ML (values from Basu and Reinsel (1993)) and by M.D.

Model	Estimator	$\hat{\beta}_0$	$\hat{\alpha}_{10}$	$\hat{\alpha}_{01}$	$\hat{\alpha}_{11}$	$\hat{\sigma}^2$
QAR	MLE	2.663	0.240	0.796	-0.108	0.032
QAR	$\widehat{\alpha}_I$	2.739	0.247	0.783	-0.090	0.031
DG	MLE	2.627	0.241	0.812	—	0.032
DG	$\widehat{\alpha}_I$	2.658	0.453	0.820	—	0.034

of the error d.f. F described in Section 2.2, we found $\sup_x |Z_n(x)| = 9.778$ and an associated p-value of 0.39 based on 100 parametric bootstrap samples. Therefore, the null hypothesis of the error d.f. F being normal cannot be rejected.

The MLE and $\widehat{\alpha}_I$ of the DG model are somewhat different. Basu and Reinsel (1993) argued that a DG model provided an adequate fit because $\hat{\alpha}_{11} \approx -\hat{\alpha}_{10}\hat{\alpha}_{01} = -0.191$ for the ML estimates of the QAR model. However, they did not perform a statistical test. We investigate the claim of Basu and Reinsel (1993) by means of the LAD test of $H_0 : \alpha_{11} = -\alpha_{10}\alpha_{01}$ described in Section 2.3. We find $\mathcal{D} = 0.390$ and estimate $f(0)$ by the slope of the linear regression of F_n on the residuals with magnitude smaller than 0.1, yielding $\hat{f}_n = 2.391$ and thus $\widehat{\mathcal{D}} = 3.730$. Because $\widehat{\mathcal{D}} < \chi^2_1(95\%) = 3.841$, the null hypothesis of a DG process cannot be rejected at the 5% level, although this is not strong evidence. For the DG model, tests on the error d.f. F reject neither symmetry nor normality.

Acknowledgements

The authors would like to thank Richard Martin and a diligent referee for constructive and thoughtful comments. This project was initiated at North Carolina State University, Raleigh, in February-March 2004, while the second author was on sabbatical leave from the Michigan State University. The work of Genton was partially supported by NSF grant DMS-0504896.

References

Bartlett, M. S. (1968). A further note on nearest neighbour models. *J. Roy. Statist. Soc. Ser. A* **131**, 579-580.

Bartlett, M. S. (1971). Physical nearest neighbour models and non-linear time series. *J. Appl. Probab.* **8**, 222-232.

Basu, S. and Reinsel, G. C. (1993). Properties of the spatial unilateral first-order ARMA model. *Adv. Appl. Probab.* **25**, 631-648.

Besag, J. (1972). On the correlation structure of some two dimensional stationary processes. *Biometrika* **59**, 43-48.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **36**, 192-236.

de Luna, X. and Genton, M. G. (2002). Simulation-based inference for simultaneous processes on regular lattices. *Stat. Comput.* **12**, 125-134.

Genton, M. G. (2003). Breakdown-point for spatially and temporally correlated observations. In *Developments in Robust Statistics, International Conference on Robust Statistics* 2001 (Edited by R. Dutter, P. Filzmoser, U. Gather and P. J. Rousseeuw), 148-159. Springer, Heidelberg.

Genton, M. G. and Ronchetti, E. (2003). Robust indirect inference. *J. Amer. Statist. Assoc.* **98**, 67-76.

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

Hampel, F., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.

Kempton, R. A. and Howes, C. W. (1981). The use of neighbouring plot values in the analysis of variety trials. *Appl. Statist.* **30**, 59-70.

Khmaladze, È. V. (1981). A martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.* **26**, 246-265.

Khmaladze, È. V. and Koul, H. L. (2004). Martingale transforms goodness-of-fit tests in regression models. *Ann. Statist.* **32**, 995-1034.

Koul, H. L. (1986). Minimum distance estimation and goodness-of-fit tests in first-order autoregression. *Ann. Statist.* **14**, 1194-1213.

Koul, H. L. (2002). *Weighted Empirical Processes in Dynamic Non-Linear Models*. Lecture Notes in Statistics, **166**. Springer, New York.

Martin, R. J. (1979). A subclass of lattice process applied to a problem in planar sampling. *Biometrika* **66**, 209-217.

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

Martin, R. J. (1997). A three-dimensional unilateral autoregressive lattice process. *J. Statist. Plann. Inference* **59**, 1-18.

Ord, K. (1975). Estimation methods for models of spatial interaction. *J. Am. Statist. Assoc.* **70**, 120-126.

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

Scaccia, L. and Martin, R. J. (2002). Testing for simplification in spatial models. In *COMPSTAT 2002, Proceedings in Computational Statistics* (Edited by W. Härdle and B. Rönz), 581-586. Physica-Verlag, Heidelberg.

Scaccia, L. and Martin, R. J. (2005). Testing axial symmetry and separability of lattice processes. *J. Statist. Plann. Inference* **131**, 19-39.

Smirnov, N. V. (1947). Sur un critère de symétrie de la loi de distribution d'une variable aléatoire. *Compte Rendu de l'Académie des Sciences de l'URSS* **56**, 11-14.

Tjøstheim, D. (1978). Statistical spatial series modelling. *Adv. Appl. Probab.* **10**, 130-154.

Tjøstheim, D. (1983). Statistical spatial series modelling II: Some further results on unilateral processes. *Adv. Appl. Probab.* **15**, 562-584.

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

Whittle, P. (1954). On stationary processes in the plane. *Biometrika* **41**, 434-449.

Department of Econometrics, University of Geneva, Bd du Pont-d'Arve 40, CH-1211 Geneva 4, Switzerland.

E-mail: Marc.Genton@metri.unige.ch

Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA.

E-mail: genton@stat.tamu.edu

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA.

E-mail: koul@stt.msu.edu

(Received May 2006; accepted September 2006)