The Annals of Statistics
1997, Vol. 25, No. 4, 1709–1724

ASYMPTOTIC INFERENCE FOR NEAR UNIT ROOTS IN SPATIAL AUTOREGRESSION

BY B. B. BHATTACHARYYA, G. D. RICHARDSON AND L. A. FRANKLIN

North Carolina State University, University of Central Florida and Indiana State University

Asymptotic inference for estimators of (α_n, β_n) in the spatial autoregressive model $Z_{ij}(n) = \alpha_n Z_{i-1,j}(n) + \beta_n Z_{i,j-1}(n) - \alpha_n \beta_n Z_{i-1,j-1}(n) + \varepsilon_{ij}$ is obtained when α_n and β_n are near unit roots. When α_n and β_n are reparameterized by $\alpha_n = e^{c/n}$ and $\beta_n = e^{d/n}$, it is shown that if the "one-step Gauss–Newton estimator" of $\lambda_1 \alpha_n + \lambda_2 \beta_n$ is properly normalized and embedded in the function space $D([0,1]^2)$, the limiting distribution is a Gaussian process. The key idea in the proof relies on a maximal inequality for a two-parameter martingale which may be of independent interest. A simulation study illustrates the speed of convergence and goodness-of-fit of these estimators for various sample sizes.

1. Introduction. Testing for the presence of a unit root in the order 1 time series model $y_t = \beta y_{t-1} + \varepsilon_t$, has received considerable attention. It is well known that the sequence $\{\hat{\beta}_n\}$ of least squares estimators of $\beta = 1$ satisfies the following asymptotic result: $2n(\hat{\beta}_n - 1) \to_{\mathfrak{D}} (W^2(1) - 1) \times (\int_0^1 W^2(t)\,dt)^{-1}$, where W denotes a one-parameter Brownian motion process [White (1958)]. Bobkoski (1983), Phillips (1987), and Chan and Wei (1987) have investigated the near unit root time series model $y_t(n) = \beta_n y_{t-1}(n) + \varepsilon_t$, $1 \le t \le n$, where $\beta_n = e^{c/n}$, c an unknown constant, and proved that the sequence $\{n(\hat{\beta}_n - \beta_n)\}$ converges in distribution to a quotient of stochastic integrals involving a standard Brownian motion process on $[0,1]$. Cox and Llatas (1991) proved similar results for M-estimators.

The doubly geometric model $Z_{ij} = \alpha Z_{i-1,j} + \beta Z_{i,j-1} + \alpha\beta Z_{i-1,j-1} + \varepsilon_{ij}$ introduced by Martin (1979) is an analogue in the spatial setting to the above AR(1) time series. In a subsequent paper Martin (1990) indicated that these models have wide practical applicability. The model has been used by Jain [(1981), page 514] in the study of image processing and in agriculture field trials by Martin (1990) and Cullis and Gleeson (1991). Tjostheim (1981) indicated that one-quadrant finite autoregressive spatial models are useful in studying digital filtering and system theory. In the stationary case when $|\alpha| < 1$ and $|\beta| < 1$, asymptotic normality of the limiting distribution of several estimators of (α, β) has been shown [e.g., Tjostheim (1978, 1983), Basu (1990), Khalil (1991)]; a recent treatment of these and further results can be found in Basu and Reinsel (1992, 1993). The normalizing sequence in

Received October 1995; revised September 1996.
AMS 1991 *subject classifications*. Primary 62F12, 62M30; secondary 60F17.
Key words and phrases. Spatial autoregressive process, near unit roots, Gauss–Newton estimation, central limit theory.

each of these asymptotic results is of order n. Contrary to the AR(1) time series case, the limiting distribution of the sequence of Gauss–Newton estimators of (α, β) has recently been shown to be a bivariate normal when $\alpha = \beta = 1$ [Bhattacharyya, Khalil and Richardson (1995)]. The normalizing sequence in this case is of order $n^{3/2}$.

Feasibility of the doubly geometric model being nearly nonstationary is illustrated by Basu and Reinsel (1994). A numerical example is presented showing that one of the best fits to wheat-yield data is obtained by using a linear regression model with an error structure at site (i, j) of the form $N_{ij} = Z_{ij} + U_{ij}$, where $Z_{ij} = \alpha Z_{i-1,j} + \beta Z_{i,j-1} - \alpha\beta Z_{i-1,j-1} + \varepsilon_{ij}$ and U_{ij} is an additional independent error component. They indicate that the residual values obtained from an ordinary least squares fit exhibit a trend behavior which suggest that the data is nonstationary and, indeed, their estimated value of β is 0.947 (a near unit root). Moreover, Cullis and Gleeson [(1991), page 1450, (4)] include the above model with $\alpha = \beta = 1$ in the class of models used to represent the error structure for the linear regression model in the analyses of field data.

The purpose of this paper is to estimate the values of the parameters (α_n, β_n) for the model

$$(1.1)\quad Z_{ij}(n) = \alpha_n Z_{i-1,j}(n) + \beta_n Z_{i,j-1}(n) - \alpha_n \beta_n Z_{i-1,j-1}(n) + \varepsilon_{ij},$$

$1 \le i, j \le n$, subject to the following assumptions.

(A.1) $\alpha_n = e^{c/n}$ and $\beta_n = e^{d/n}$, where c and d are nonzero but unknown constants.

(A.2) $Z_{ij}(n) = 0$ when either $i \le 0$ or $j \le 0$.

(A.3) $\{\varepsilon_{ij}\}$ are i.i.d., mean zero, variance σ^2 and each has a finite fourth moment.

(A.4) $\{\bar{\alpha}_n\}$ and $\{\bar{\beta}_n\}$ are initial estimators satisfying $\bar{\alpha}_n - \alpha_n = O_p(n^{-3/2})$ and $\bar{\beta}_n - \beta_n = O_p(n^{-3/2})$.

These results can be used to estimate the parameters (α_n, β_n) in the near unit roots case where the normalizing sequence changes from order n to order $n^{3/2}$. Some preliminary notation is needed before the preliminary result can be stated.

Let $K := [0, 1]^2$ denote the unit square and let $t \in K$. The four quadrants of K which have t as their origin are designated by $Q_1(\ge, \ge)$, $Q_2(<, \ge)$, $Q_3(<, <)$ and $Q_4(\ge, <)$. Define D_2 to be the set of all real-valued functions f on K for which $\lim_{s \to t} f(s)$ exists when s belongs to a single quadrant and $\lim_{s \to t} f(s) = f(t)$ when $s \in Q_1$.

Following Bickel and Wichura (1971), there is a metric on D_q which induces Skorohod's well-known topology when $q = 1$ and makes D_q separable, complete and whose Borel σ-field coincides with that generated by the coordinate mappings. Convergence in D_2 will be relative to the above-mentioned metric.

Let $\hat{\theta}_n$ denote the "one-step Gauss-Newton estimator" of (α, β) [Fuller (1976), page 213]. An expression for $\hat{\theta}_n$ in terms of the initial estimator $\bar{\theta}_n$

described in (A.4) is given by $(\hat{\theta}_n - \theta_n)' = n^{-3}(A_n(1,1))^{-1}\sum_{i=1}^n \sum_{j=1}^n G_{ij}(\bar{\theta}_n)$ [see (2.3), (2.5) and (2.7)]. Let θ_n^* denote the embedding of $\hat{\theta}_n$ in $D_2 \times D_2$ as follows:

$$n^{3/2}(\theta_n^*(t) - \theta_n)' = (A_n(1,1))^{-1} n^{-3/2} \sum_{i=1}^{[nt_1]} \sum_{j=1}^{[nt_2]} G_{ij}(\bar{\theta}_n), \tag{1.2}$$

$t = (t_1, t_2) \in K$. Moreover, define

$$g_\gamma(x) := (e^{2\gamma x} - 1 - 2\gamma x)/4\gamma^2 \quad \text{where } x \in [0,1] \tag{1.3}$$

and $\gamma \neq 0$. The principal result of this paper is now given.

THEOREM 1.1. *Assume that model* (1.1) *and conditions* (A.1)–(A.4) *are satisfied. Fix* $\lambda = (\lambda_1, \lambda_2) \in R^2$ *and let* θ_n^* *denote the random element in* $D_2 \times D_2$ *as defined in* (1.2). *Then the sequence* $\{n^{3/2}(\theta_n^* - \theta_n)\lambda'\}$ *of random elements in* D_2 *converges in distribution to a mean zero Gaussian process W having*

$$\operatorname{cov}(W(s), W(t)) = \frac{\lambda_1^2 (s_2 \wedge t_2) g_c(s_1 \wedge t_1)}{(g_c(1))^2} + \frac{\lambda_2^2 (s_1 \wedge t_1) g_d(s_2 \wedge t_2)}{(g_d(1))^2}.$$

In particular,

$$n^{3/2}(\hat{\theta}_n - \theta_n) \to_{\mathfrak{D}} N(0, \Gamma) \quad \textit{where } \Gamma = \operatorname{diag}\left(\frac{1}{g_c(1)}, \frac{1}{g_d(1)}\right).$$

REMARK 1.1. The maximal inequality given below is used to verify tightness in D_2 of the sequence $\{n^{3/2}(\theta_n^* - \theta_n)\lambda'\}$ given in Theorem 1.1. This inequality is a modification of the one given by Walsh [(1986), Theorem 2.6] which requires that the strong martingale vanish on the boundary. Moreover, observe that for Γ defined in Theorem 1.1, $\lim_{c,d \to 0} \Gamma = \operatorname{diag}(2,2)$, and this agrees with the limiting distribution of $\{n^{3/2}(\hat{\theta}_n - \theta_n)\}$ for the case when $\alpha = \beta = 1$ [Bhattacharyya, Khalil and Richardson (1995)].

LEMMA 1.1 (Maximal inequality). *Suppose that* $J = \{(i,j) : m \le i \le M$ and $n \le j \le N\} \subseteq I$. *Assume that* $\{Z_t, \mathfrak{F}_t, t \in J\}$ *is a square integrable, strong martingale* (*see Section* 2) *and denote* $\tau_{MN} = (E(Z_{MN}^2))^{1/2}$. *Fix* $\lambda > 0$. *Then there exist positive constants* α *and* A (*each independent of* λ, M *and* N) *for which*

$$P\{\max|Z_t| > \lambda, t \in J\} \le \frac{A\tau_{MN}}{\lambda}\left(P\left\{|Z_{MN}| > \frac{\lambda}{a}\right\}\right)^{1/2} + A\left(\frac{\tau_{MN}}{\lambda}\right)^{3/2}\left(P\left\{|Z_{MN}| > \frac{\lambda}{a}\right\}\right)^{1/4}.$$

A simulation study is given to asses the speed of convergence of $\{\hat{\theta}_n - \theta_n\}$ to zero as well as the goodness-of-fit of the normalized sequence.

2. Order properties and proof of maximal inequality. Throughout the remainder of this paper, model (1.1) and conditions (A.1)–(A.4) are assumed valid. For the sake of brevity, expressions such as $Z_{ij}(n)$ are denoted simply by $Z_{ij}, 1 \le i, j \le n$. First, an explicit expression for the Gauss–Newton estimator of (α_n, β_n) in model (1.1) is given. Denote $\theta_n = (\alpha_n, \beta_n)$, $\bar{\theta}_n = (\bar{\alpha}_n, \bar{\beta}_n)$, $f_{ij}(a, b) = aZ_{i-1,j} + bZ_{i,j-1} - abZ_{i-1,j-1}$ and $F_{ij}(a, b) = (\partial f_{ij}(a, b)/\partial a, \partial f_{ij}(a, b)/\partial b)$. Expanding $f_{ij}(\theta_n)$ about $\bar{\theta}_n$ in model (1.1) gives

$$Z_{ij} = f_{ij}(\bar{\theta}_n) + F_{ij}(\bar{\theta}_n)(\theta_n - \bar{\theta}_n)' + R_{ij}(\bar{\theta}_n) + \varepsilon_{ij}, \tag{2.1}$$

where $R_{ij}(\bar{\theta}_n) = -Z_{i-1,j-1}(\alpha_n - \bar{\alpha}_n)(\beta_n - \bar{\beta}_n)$. Define

$$\hat{\delta}_n' = \left(\sum_{i,j=1}^{n} F_{ij}'(\bar{\theta}_n) F_{ij}(\bar{\theta}_n)\right)^{-1} \sum_{i,j=1}^{n} F_{ij}'(\bar{\theta}_n)\left(Z_{ij} - f_{ij}(\bar{\theta}_n)\right). \tag{2.2}$$

Then $\hat{\theta}_n = \hat{\delta}_n + \bar{\theta}_n$ is called the "one-step Gauss–Newton estimator" of θ_n. Substituting (2.1) into (2.2) and simplifying gives

$$\left(\hat{\theta}_n - \theta_n\right)' = \left(\sum_{i,j=1}^{n} F_{ij}'(\bar{\theta}_n) F_{ij}(\bar{\theta}_n)\right)^{-1} \sum_{i,j=1}^{n} G_{ij}(\bar{\theta}_n), \tag{2.3}$$

where

$$G_{ij}(\bar{\theta}_n) = F_{ij}'(\bar{\theta}_n)\left(R_{ij}(\bar{\theta}_n) + \varepsilon_{ij}\right).$$

Define

$$X_{ij} := Z_{ij} - \beta_n Z_{i,j-1} \quad \text{and} \quad Y_{ij} := Z_{ij} - \alpha_n Z_{i-1,j} \tag{2.4}$$

and observe that

$$\begin{aligned} F_{ij}(\bar{\theta}_n) &= \left(Z_{i-1,j} - \bar{\beta}_n Z_{i-1,j-1}, Z_{i,j-1} - \bar{\alpha}_n Z_{i-1,j-1}\right) \\ &= \left(X_{i-1,j} + (\beta_n - \bar{\beta}_n) Z_{i-1,j-1}, Y_{i,j-1} + (\alpha_n - \bar{\alpha}_n) Z_{i-1,j-1}\right). \end{aligned} \tag{2.5}$$

Moreover, it follows from (1.1) and (2.4) that

$$X_{ij} = \alpha_n X_{i-1,j} + \varepsilon_{ij}, \qquad Y_{ij} = \beta_n Y_{i,j-1} + \varepsilon_{ij}$$

and thus by (A.1) and (A.2),

$$X_{ij} = \sum_{k=1}^{i} \alpha_n^{i-k} \varepsilon_{kj} \quad \text{and} \quad Y_{ij} = \sum_{l=1}^{j} \beta_n^{j-l} \varepsilon_{il}.$$

The following order properties are straightforward but somewhat tedious to verify.

LEMMA 2.1. *Fix $t = (t_1, t_2) \in K$ and let g_γ be as defined in* (1.3). *Then*:

(i) $$n^{-3} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} X_{i-1,j}^2 = t_2 g_c(t_1) \sigma^2 + O_p(n^{-1/2}),$$

(ii) $$n^{-3}\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} Y_{i,j-1}^2 = t_1 g_d(t_2)\sigma^2 + O_p(n^{-1/2}),$$

(iii) $$\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} Z_{i-1,j-1}^2 = O_p(n^4),$$

(iv) $$\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} X_{i-1,j}Y_{i,j-1} = O_p(n^2),$$

(v) $$\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} X_{i-1,j}Z_{i-1,j-1} = O_p(n^3),$$

(vi) $$\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} Y_{i,j-1}Z_{i-1,j-1} = O_p(n^3),$$

(vii) $$\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} X_{i-1,j}\varepsilon_{ij} = O_p(n^{3/2}),$$

(viii) $$\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} Y_{i,j-1}\varepsilon_{ij} = O_p(n^{3/2}),$$

(ix) $$\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} Z_{i-1,j-1}\varepsilon_{ij} = O_p(n^2),$$

(x) $$\sup_{1\le i,j\le n} E(X_{ij}^4) = O(n^2),$$

(xi) $$\sup_{1\le i,j\le n} E(Y_{ij}^4) = O(n^2),$$

(xii) $$\sup_{1\le i,j\le n} E(Z_{ij}^4) = O(n^4),$$

(xiii) $$\sup_{1\le i,j\le n} E(X_{i-1,j}Y_{i,j-1})^2 = O(n^2),$$

(xiv) $$\sup_{1\le i,j\le n} E(X_{i-1,j}Z_{i-1,j-1})^2 = O(n^3),$$

(xv) $$\sup_{1\le i,j\le n} E(Y_{i,j-1}Z_{i-1,j-1})^2 = O(n^3),$$

(xvi) $$\sup_{1\le i\le n} E\Big(\sum_{j=1}^{n} X_{i-1,j}\varepsilon_{ij}\Big)^4 = O(n^4),$$

(xvii) $$\sup_{1\le i\le n} E\Big(\sum_{j=1}^{n} Y_{i,j-1}\varepsilon_{ij}\Big)^4 = O(n^4).$$

REMARK 2.1. Assume that model (1.1) and (A.1)–(A.3) are valid. Then there exist initial estimators $\{\bar{\alpha}_n\}$ and $\{\bar{\beta}_n\}$ satisfying (A.4). Indeed, define $T_{ij} := Z_{ij} - Z_{i,j-1}$ and $U_{ij} := Z_{ij} - Z_{i-1,j}$. Recall from (2.4) that $X_{ij} = Z_{ij} - \beta_n Z_{i,j-1}$ and model (1.1) can be written as $X_{ij} = \alpha_n X_{i-1,j} + \varepsilon_{ij}$. It follows that $X_{ij} = T_{ij} - (\beta_n - 1)Z_{i,j-1}$ and hence $T_{ij} = \alpha_n T_{i-1,j} + (\beta_n - 1)Y_{i,j-1} + \varepsilon_{ij}$, where Y_{ij} is given in (2.4). Define $\bar{\alpha}_n := \sum_{i,j=1}^n T_{ij}T_{i-1,j} / \sum_{i,j=1}^n T_{i-1,j}^2$ and thus

$$\bar{\alpha}_n - \alpha_n = \left[(\beta_n - 1) \sum_{i,j=1}^n T_{i-1,j} Y_{i,j-1} + \sum_{i,j=1}^n T_{i-1,j}\varepsilon_{ij} \right] \Big/ \sum_{i,j=1}^n T_{i-1,j}^2 .$$

Expressing $T_{i-1,j}$ in terms of $X_{i-1,j}$ and $Z_{i-1,j-1}$, it follows from Lemma 2.1 that $n^{-3}\sum_{i,j=1}^n T_{i-1,j}^2 \to_P \sigma^2 g_c(1)$. Continuing in this manner, one shows that $\bar{\alpha}_n - \alpha_n = O_p(n^{-3/2})$. A similar result utilizing U_{ij} holds when $\bar{\beta}_n := \sum_{i,j=1}^n U_{ij}U_{i,j-1} / \sum_{i,j=1}^n U^2{}_{i,j-1}$.

Given $s = (s_1, s_2)$, $t = (t_1, t_2)$ in K for which $s_1 < t_1$ and $s_2 < t_2$, define the rectangle $(s, t] = (s_1, t_1] \times (s_2, t_2]$. If X is a random element in D_2, denote the increment of V over the above rectangle by $V(s,t] = V_t - V_{t_1 s_2} - V_{s_1 t_2} + V_s$ and let λ be the Lebesque measure on $(K, \mathfrak{B})$. Define $T_n = \{(k/n, l/n): k, l \text{ are integers satisfying } 0 \le k, l \le n\}$. Following Bickel and Wichura [(1971), page 1665], a sequence $\{V_n\}$ of random elements in D_2 is tight provided there exist positive real numbers α_1, α_2, δ and M such that if $(s, t]$ and $(u, v]$ are disjoint rectangles having corner points in T_n and a common edge:

$$E\left[|V_n(s,t]|^{\alpha_1} |V_n(u,v]|^{\alpha_2}\right] \le M\left(\lambda(s,t]\lambda(u,v]\right)^{(1+\delta)/2}. \tag{2.6}$$

LEMMA 2.2. *The following sequences of random elements in D_2 are tight*:

(i) $$\left\{ n^{-3} \sum_1^{[nt_1]} \sum_1^{[nt_2]} X_{i-1,j}^2 \right\},$$

(ii) $$\left\{ n^{-3} \sum_1^{[nt_1]} \sum_1^{[nt_2]} Y_{i,j-1}^2 \right\},$$

(iii) $$\left\{ n^{-4} \sum_1^{[nt_1]} \sum_1^{[nt_2]} Z_{i-1,j-1}^2 \right\},$$

(iv) $$\left\{ n^{-3} \sum_1^{[nt_1]} \sum_1^{[nt_2]} X_{i-1,j} Y_{i,j-1} \right\},$$

(v) $$\left\{ n^{-7/2} \sum_1^{[nt_1]} \sum_1^{[nt_2]} X_{i-1,j} Z_{i-1,j-1} \right\},$$

(vi) $$\left\{ n^{-7/2} \sum_1^{[nt_1]} \sum_1^{[nt_2]} Y_{i,j-1} Z_{i-1,j-1} \right\},$$

(vii) $$\left\{ n^{-3} \sum_1^{[nt_1]} \sum_1^{[nt_2]} Z_{i-1,j-1} \varepsilon_{ij} \right\}.$$

PROOF. Tightness will follow by verifying (2.6); only the proof of (ii) is given here since it is similar to the other parts. Assume that $(s, t]$ and $(u, v]$ are rectangles having corners in T_n and a common vertical edge connecting points $(t_1, s_2) = (u_1, u_2)$ and $(t_1, t_2) = (u_1, v_2)$.

Observe that

$$V_n(s, t] = \sum_{i=[ns_1]+1}^{[nt_1]} \sum_{j=[ns_2]+1}^{[nt_2]} Y_{i,j-1}^2$$

and

$$V_n(u, v] = \sum_{i'=[nu_1]+1}^{[nv_1]} \sum_{j'=[nu_2]+1}^{[nv_2]} Y_{i',j'-1}^2.$$

Since $s_2 = u_2$ and $t_2 = v_2$, it follows that

$$E[|V_n(s,t]||V_n(u,v]|] = \sum_{i=[ns_1]+1}^{\{nt_1]} \sum_{i'=[nu_1]+1}^{[nv_1]} \sum_{j,j'=[nu_2]+1}^{[nv_2]} E\left(Y_{i,j-1}^2 Y_{i',j'-1}^2\right).$$

Hence, according to Lemma 2.1(xi) and Cauchy's inequality,

$$\begin{aligned}
&\frac{1}{n^6} E[|V_n(s,t]||V_n(u,v]|] \\
&\quad \le \frac{M}{n^4}([nt_1] - [ns_1])([nv_1] - [nu_1])([nv_2] - [nu_2])^2 \\
&\quad = M\lambda(s,t]\,\lambda(u,v]
\end{aligned}$$

since the corner points of both rectangles belong to T_n. A similar argument holds when the rectangles have a common horizontal edge. □

Since tightness and convergence of all finite-dimensional distributions characterizes convergence in distribution in D_2, the following results are implied by the Lemmas 2.1 and 2.2.

COROLLARY 2.1. *The following sequences converge in* D_2:

(i) $$n^{-3} \sum_1^{[nt_1]} \sum_1^{[nt_2]} X_{i-1,j}^2 \to_{\mathfrak{D}} t_2 g_c(t_1)\sigma^2,$$

(ii) $$n^{-3} \sum_1^{[nt_1]} \sum_1^{[nt_2]} Y_{i,j-1}^2 \to_{\mathfrak{D}} t_1 g_d(t_2)\sigma^2,$$

(iii) $$n^{-\gamma} \sum_1^{[nt_1]} \sum_1^{[nt_2]} Z_{i-1,j-1}^2 \to_{\mathfrak{D}} 0 \quad \textit{when} \quad \gamma > 4,$$

(iv) $$n^{-3} \sum_1^{[nt_1]} \sum_1^{[nt_2]} X_{i-1,j} Y_{i,j-1} \to_{\mathfrak{D}} 0,$$

(v) $$n^{-7/2} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} X_{i-1,j} Z_{i-1,j-1} \to_{\mathfrak{D}} 0,$$

(vi) $$n^{-7/2} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} Y_{i,j-1} Z_{i-1,j-1} \to_{\mathfrak{D}} 0,$$

(vii) $$n^{-3} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} Z_{i-1,j-1} \varepsilon_{ij} \to_{\mathfrak{D}} 0.$$

REMARK 2.2. Results given in Corollary 2.1(iii–vii) could be improved but verifications are more technical and do not follow from Lemmas 2.1 and 2.2 and, moreover, are not needed here.

Given the sequence $\{(\bar{\alpha}_n, \bar{\beta}_n)\}$ of initial estimators $\{(\alpha_n, \beta_n)\}$ described in (A.4), define the following random elements in D_2:

$$B_n(t) := n^{-3} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} \left(X_{i-1,j} + (\beta_n - \bar{\beta}_n) Z_{i-1,j-1}\right)^2,$$

$$C_n(t) := n^{-3} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} \left(X_{i-1,j} + (\beta_n - \bar{\beta}_n) Z_{i-1,j-1}\right) \times \left(Y_{i,j-1} + (\alpha_n - \bar{\alpha}_n) Z_{i-1,j-1}\right),$$

$$D_n(t) := n^{-3} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} \left(Y_{i,j-1} + (\alpha_n - \bar{\alpha}_n) Z_{i-1,j-1}\right)^2,$$

where $t = (t_1, t_2) \in K$ and $n \geq 1$. Moreover, denote

$$A_n(t) := \begin{bmatrix} B_n(t) & C_n(t) \\ C_n(t) & D_n(t) \end{bmatrix} \quad \text{and} \tag{2.7}$$

$$A(t) := \sigma^2 \operatorname{diag}(t_2 g_c(t_1), t_1 g_d(t_2)).$$

LEMMA 2.3. *Let $A_n(t)$ and $A(t)$ be considered as random elements in D_2^4 equipped with the product metric. Then $A_n(t) \to_{\mathfrak{D}} A(t)$ in D_2^4.*

PROOF. Since the entries of $A(t)$ are constant random elements of the separable metric space D_2, it suffices to show each component sequence converges in D_2 [Billingsley (1968), Theorem 4.4]. Let us show that $B_n(t) \to_{\mathfrak{D}} t_2 g_c(t_1)\sigma^2$ in D_2. Recall that $n^{-3}\sum_1^{[nt_1]} \sum_1^{[nt_2]} X_{i-1,j}^2 \to_{\mathfrak{D}} t_2 g_c(t_1)\sigma^2$ in D_2 according to Corollary 2.1(i). Also

$$\begin{aligned}(\beta_n - \bar{\beta}_n) n^{-3} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} X_{i-1,j} Z_{i-1,j-1} \\ = n^{1/2} (\beta_n - \bar{\beta}_n) n^{-7/2} \sum_{1}^{[nt_1]} \sum_{1}^{[nt_2]} X_{i-1,j} Z_{i-1,j-1} \to_{\mathfrak{D}} 0\end{aligned}$$

in D_2 by (A.4) and Corollary 2.1(v). Likewise,

$$\left(\beta_n - \bar{\beta}_n\right)^2 n^{-3} \sum_1^{[nt_1]} \sum_1^{[nt_2]} Z_{i-1,j-1} \to_{\mathfrak{D}} 0$$

in D_2 by Corollary 2.1(iii) and (A.4) and thus $B_n(t) \to_{\mathfrak{D}} t_2 g_c(t_1)\sigma^2$ in D_2. Verification of the remaining entries are shown similarly. □

The notation I will always be used to denote the set of all ordered pairs of positive integers. Let $s = (s_1, s_2)$ and $t = (t_1, t_2)$ be members of I. Define $s < t$ to mean $s_1 \le t_1$ and $s_2 \le t_2$ (product ordering); whereas, $s \ll t$ denotes $s_1 < t_1$ and $s_2 < t_2$. Moreover, given the underlying probability space $(\Omega, \mathfrak{F}, P)$ and sub-σ-field $\mathfrak{F}_t$, let Z_t denote an integrable, $\mathfrak{F}_t$-measurable random variable, where t belongs to the product set $J \subseteq I$. Assume that $\mathfrak{F}_s \subseteq \mathfrak{F}_t \subseteq \mathfrak{F}$ when $s < t$. Then $\{Z_t, \mathfrak{F}_t, t \in J\}$ [see Walsh (1986)] is called a *strong martingale*, provided, for each $s = (s_1, s_2)$ and $t = (t_1, t_2)$ in J,

(i) $$E(Z_t \mid \mathfrak{F}_s) = Z_s \quad \text{when } s < t$$

and

(ii) $$E(Z(s,t] \mid \mathfrak{F}_s^*) = 0 \quad \text{when } s \ll t,$$

where $Z(s,t] = Z_{t_1 t_2} - Z_{s_1 t_2} - Z_{t_1 s_2} + Z_{s_1 s_2}$ and $\mathfrak{F}_s^*$ denotes the smallest σ-field containing each $\mathfrak{F}_{ij}$ with either $i \le s_1$ or $j \le s_2$.

Following Walsh (1986), $\mathfrak{F}_s^*$ may be described as "everything which is not in the future of s."

LEMMA 2.4. *Fix a positive integer n and define* $J = \{(t_1, t_2) \in I : t_1, t_2 \le n\}$. *Given constants A and B, denote* $U_t = \sum_{i=1}^{T_1} \sum_{j=1}^{t_2} \xi_{ij}$ *for each* $t = (t_1, t_2) \in J$, *where* $\xi_{ij} := (AX_{i-1,j} + BY_{i,j-1})\varepsilon_{ij}$, *and let* $\mathfrak{F}_t$ *be the smallest σ-field making each* ε_{ij} *measurable, for* $1 \le i \le t_1$ *and* $1 \le j \le t_2$. *Then* $\{U_t, \mathfrak{F}_t, t \in J\}$ *is a strong martingale; that is,* $\{U_t(n), \mathfrak{F}_2(n), t \in J\}$ *is a strong martingale array.*

PROOF. (i) Let $s = (s_1, s_2) < t = (t_1, t_2)$ be elements in J. Note that

$$\begin{aligned} U_t &= U_s + \sum_{i=1}^{s_1} \sum_{j=s_2+1}^{t_2} \xi_{ij} + \sum_{i=s_1+1}^{t_1} \sum_{j=1}^{s_2} \xi_{ij} + \sum_{i=s_1+1}^{t_1} \sum_{j=s_2+1}^{t_2} \xi_{ij} \\ &:= U_s + F + G + H \end{aligned}$$

and thus $E(F \mid \mathfrak{F}_s) = E(G \mid \mathfrak{F}_s) = E(H \mid \mathfrak{F}_s) = 0$ since $X_{i-1,j} = \sum_{k=1}^{i-1} \alpha_n^{i-1-k} \varepsilon_{kj}$ and $Y_{i,j-1} = \sum_{l=1}^{j-1} \beta_n^{j-1-l} \varepsilon_{il}$ are independent of ε_{ij}. Hence $E(U_t \mid \mathfrak{F}_s) = U_s$.

(ii) Assume $s \ll t$ and observe that $U(s,t] = \sum_{i=s_1+1}^{t_1} \sum_{j=s_2+1}^{t_2} \xi_{ij}$. Therefore $E(U(s,t] \mid \mathfrak{F}_s^*) = 0$ since ε_{ij}'s are independent random variables. □

PROOF OF LEMMA 1.1. Define

(2.8) $$U_{ij} = Z_{ij} - Z_{in} - Z_{mj} + Z_{mn} \quad \text{for each } (i,j) \in J.$$

Then $\{U_t, \mathfrak{F}_t, t \in J\}$ is also a strong martingale with the property that

$$U_{mj} = 0 \quad \text{and} \quad U_{in} = 0 \quad \text{for} \quad m \le i \le M \quad \text{and} \quad n \le j \le N.$$

It follows from Brown's (1971) maximal inequality and Theorem 2.6 of Walsh (1986) that, for $\gamma > 0$,

$$P\{\max|U_t| > \gamma, t \in J\} \le \frac{C_1}{\gamma}\big(E(U_{MN}^2)\big)^{1/2}\left(P\left\{|U_{MN}| > \frac{\gamma}{6}\right\}\right)^{1/2}, \tag{2.9}$$

for some C_1. Moreover, since $\{Z_{Mn}, Z_{MN}\}$, $\{Z_{mN}, Z_{MN}\}$ and $\{Z_{mn}, Z_{MN}\}$ are each ordinary martingales, it follows by (2.8) and Brown's maximal inequality that

$$P\left\{|U_{MN}| > \frac{\gamma}{6}\right\} \le P\left\{|Z_{MN}| > \frac{\gamma}{24}\right\} + \frac{C_2\tau_{MN}}{\gamma}\left(P\left\{|Z_{MN}| > \frac{\gamma}{48}\right\}\right)^{1/2} \tag{2.10}$$

Substituting (2.10) into (2.9) gives

$$\begin{aligned} &P\{\max|U_t| > \gamma, t \in J\} \\ &\quad\le \frac{C_3}{\gamma}\big(E(U_{MN}^2)\big)^{1/2}\left[\left(P\left\{|Z_{MN}| > \frac{\gamma}{24}\right\}\right)^{1/2}\right. \\ &\qquad\qquad \left. + \frac{1}{\sqrt{\gamma}}\tau_{MN}^{1/2}\left(P\left\{|Z_{MN}| > \frac{\gamma}{48}\right\}\right)^{1/4}\right]. \end{aligned} \tag{2.11}$$

Rewriting (2.8) as

$$Z_{ij} = Z_{in} + Z_{mj} - Z_{mn} + U_{ij}$$

and noting that $\{Z_{mn}, Z_{m+1,n}, \dots, Z_{Mn}, Z_{MN}\}$, $\{Z_{mn}, Z_{m,n+1}, \dots, Z_{mN}, Z_{MN}\}$ and $\{Z_{mn}, Z_{MN}\}$ are ordinary martingales, one obtains

$$\begin{aligned} &P\{\max|Z_t| > \lambda, t \in J\} \\ &\quad\le \frac{C_4\tau_{MN}}{\lambda}\left(P\left\{|Z_{MN}| > \frac{\lambda}{8}\right\}\right)^{1/2} + P\left\{\max|U_t| > \frac{\lambda}{4}, t \in J\right\}. \end{aligned} \tag{2.12}$$

Substituting (2.11) into (2.12) gives

$$\begin{aligned} P\{\max|Z_t| > \lambda, t \in J\} &\le \frac{C_5\tau_{MN}}{\lambda}\left(P\left\{|Z_{MN}| > \frac{\lambda}{192}\right\}\right)^{1/2} \\ &\quad + \frac{C_5\tau_{MN}^{3/2}}{\lambda^{3/2}}\left(P\left\{|Z_{MN}| > \frac{\lambda}{192}\right\}\right)^{1/4} \end{aligned}$$

and thus the desired conclusion follows. □

3. Proof of Theorem 1.1. Observe that

$$n^{-3/2}\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]} F'_{ij}(\bar{\theta}_n)R_{ij}(\bar{\theta}_n)$$
$$= -(\alpha_n-\bar{\alpha}_n)(\beta_n-\bar{\beta}_n)n^{-3/2}$$
$$\times\left(\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]}\left(X_{i-1,j}Z_{i-1,j-1}+(\beta_n-\bar{\beta}_n)Z^2_{i-1,j-1}\right),\right.$$
$$\left.\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]}\left(Y_{i,j-1}Z_{i-1,j-1}+(\alpha_n-\bar{\alpha}_n)Z^2_{i-1,j-1}\right)\right)\to_{\mathfrak{D}} 0$$

in D_2^2 according to Corollary 2.1 and (A.4). Likewise,

$$n^{-3/2}\left((\beta_n-\bar{\beta}_n)\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]}Z_{i-1,j-1}\varepsilon_{ij},(\alpha_n-\bar{\alpha}_n)\sum_{1}^{[nt_1]}\sum_{1}^{[nt_2]}Z_{i-1,j-1}\varepsilon_{ij}\right)$$
$$\to_{\mathfrak{D}} 0\quad\text{in } D_2^2$$

by Corollary 2.1(vii) and (A.4). It follows from Lemma 2.3 that

$$A_n^{-1}(1,1)\to_{\mathfrak{D}}\sigma^{-2}\operatorname{diag}\left((g_c(1))^{-1},(g_d(1))^{-1}\right)\quad\text{in } R^4$$

and thus by (1.2), $\{n^{3/2}(\theta_n^*(t)-\theta_n)\lambda'\}$ converges in distribution in D_2 for all $\lambda\in R^2$ when $\{n^{-3/2}\sum_1^{[nt_1]}\sum_1^{[nt_2]}(aX_{i-1,j}+bY_{i,j-1})\varepsilon_{ij}\}$ does (and has the same limit), where $a=\lambda_1/\sigma^2 g_c(1)$ and $b=\lambda_2/\sigma^2 g_d(1)$. It remains to show that the latter sequence is tight in D_2 and its finite-dimensional distributions converge to an appropriate multivariate normal.

Fix $s=(s_1,s_2)$ and $t=(t_1,t_2)\in K$. Let us consider that case when $s_1<t_1$ and $t_2<s_2$. Define $W_n(t):=n^{-3/2}\sum_1^{[nt_1]}\sum_1^{[nt_2]}\xi_{ij}$, where $\xi_{ij}=(aX_{i-1,j}+bY_{i,j-1})\varepsilon_{ij}$. Following Brown (1971) and McLeish (1974), a martingale central limit theorem is used to show that $(W_n(s),W_n(t))\to_{\mathfrak{D}}\sigma^2 N(0,\Sigma)$, where

$$\Sigma=\begin{bmatrix} a^2s_2g_c(s_1)+b^2s_1g_d(s_2) & a^2t_2g_c(s_1)+b^2s_1g_d(t_2)\\ a^2t_2g_c(s_1)+b^2s_1g_d(t_2) & a^2t_2g_c(t_1)+b^2t_1g_d(t_2)\end{bmatrix}.$$

Observe that for $u=(u_1,u_2)\in R^2$,

$$u_1W_n(s)+u_2W_n(t)=\frac{(u_1+u_2)}{n^{3/2}}\sum_{1}^{[ns_1]}\sum_{1}^{[nt_2]}\xi_{ij}+\frac{u_1}{n^{3/2}}\sum_{1}^{[ns_1]}\sum_{[nt_1]+1}^{[ns_2]}\xi_{ij}$$
$$+\frac{u_2}{n^{3/2}}\sum_{[ns_1]+1}^{[nt_1]}\sum_{1}^{[nt_2]}\xi_{ij};$$

define

$$V_{ni} := \begin{cases} \dfrac{(u_1 + u_2)}{n^{3/2}} \displaystyle\sum_{j=1}^{[nt_2]} \xi_{ij} + \dfrac{u_1}{n^{3/2}} \sum_{j=[nt_2]+1}^{[ns_2]} \xi_{ij}, & 1 \le i \le [ns_1], \\ \dfrac{u_2}{n^{3/2}} \displaystyle\sum_{j=1}^{[nt_2]} \xi_{ij}, & [ns_1] + 1 \le i \le [nt_1] \end{cases}$$

and note that $u_1 W_n(s) + u_2 W_n(t) = \sum_{i=1}^{[nt_1]} V_{ni}$. Denote the smallest σ-field for which each ε_{kl}, $1 \le k \le i$ and all $l \ge 1$, is measurable by $\mathfrak{F}_{ni} = \mathfrak{F}_i$ and let $\mathfrak{F}_{no}$ be the trival σ-field, $1 \le i \le [nt_1]$. Then $E(V_{ni} \mid \mathfrak{F}_{i-1}) = 0$ and thus $\{V_{ni}, \mathfrak{F}_i, 1 \le i \le [nt_1], n \ge 1\}$ is a mean zero martingale difference array. It can be verified that $E(\xi_{ij}\xi_{i'j'} \mid \mathfrak{F}_{i-1}) = 0$ unless $i = i'$ and $j = j'$ and, moreover, $V_n^2 := \sum_{i=1}^{[nt_1]} E(V_{ni}^2 \mid \mathfrak{F}_{i-1}) \to_P \sigma^2 u \Sigma u'$. Also $s_n^2 := E(V_n^2) \to \sigma^2 u \Sigma u'$ and thus $V_n^2 / s_n^2 \to_P 1$.

Fix $\delta > 0$. It must be shown that $\sum_{i=1}^{[nt_1]} E(V_{ni}^2 1_{\{|V_{ni}| \ge \delta\}}) \to 0$ as $n \to \infty$. Observe from Lemma 2.1(xvi) and (xvii) that

$$\begin{aligned} E\left(n^{1/2} V_{ni}\right)^4 &\le C n^{-4} E\left(\sum_{j=1}^{[ns_2]} \xi_{ij}\right)^4 \\ &\le D n^{-4} \left(E\left(\sum_{j=1}^{[ns_2]} X_{i-1,j}\varepsilon_{ij}\right)^4 + E\left(\sum_{j=1}^{[ns_2]} Y_{i,j-1}\varepsilon_{ij}\right)^4 \right) = 0(1). \end{aligned}$$

It follows that $\{nV_{ni}^2 : 1 \le i \le [nt_1]\}$ is uniformly integrable and thus Lindeberg's condition follows. This shows that $(W_n(s), W_n(t)) \to_{\mathfrak{D}} \sigma^2 N(0, \Sigma)$. Likewise, other cases can be justified and thus the finite-dimensional distributions of $\{n^{-3/2}(\theta_n^* - \theta_n)\lambda'\}$ converge to a Gaussian process $W(t), t \in K$, satisfying

$$\begin{aligned} \operatorname{cov}(W(s), W(t)) &= \frac{\lambda_1^2}{(g_c(1))^2} (s_2 \wedge t_2) g_c(s_1 \wedge t_1) \\ &\quad + \frac{\lambda_2^2}{(g_d^2(1))} (s_1 \wedge t_1) g_d(s_2 \wedge t_2). \end{aligned}$$

In order to prove tightness of the sequence $\{W_n(t)\}$ in D_2 consider, for $\delta > 0$, rectangles of the form

$$R_{kl} = [k\delta, (k+1)\delta) \times [l\delta, (l+1)\delta)$$

and, for $\varepsilon > 0$, denote

$$A_{kl}^n = \left\{ \sup_t |W_n(t) - W_n(k\delta, l\delta)| > \varepsilon, t \in R_{kl} \right\}.$$

Given $\varepsilon > 0$ and $\eta > 0$, it suffices to show there exists a $\delta > 0$ such that

$$\limsup_n \sum_{k\delta < 1} \sum_{l\delta < 1} P(A_{kl}^n) < \eta. \tag{3.1}$$

Denote $u_0 = ([nk\delta], [nl\delta])$ $u_1 = ([n(k+1)\delta], [n(l+1)\delta])$ and let $U_u = \sum_{k=1}^{i} \sum_{l=1}^{j} \xi_{kl}$, where $u = (i, j)$. It follows from Lemma 2.4 that $\{U_u - U_{u_0}, \mathfrak{F}_u, u_0 \le u \le u_1\}$ is a strong martingale and thus by Lemma 1.1, for $\varepsilon > 0$, there exist positive constants a_0 and A_0 for which

$$
\begin{aligned}
P(A_{kl}^n) > {} & \frac{A_0 \tau_n}{b_n \varepsilon} \left(P\left\{ |U_{u_1} - U_{u_0}| > \frac{b_n \varepsilon}{a_0} \right\} \right)^{1/2} \\
& + A_0 \left(\frac{\tau_n}{b_n \varepsilon} \right)^{3/2} \left(P\left\{ |U_{u_1} - U_{u_0}| > \frac{b_n \varepsilon}{a_0} \right\} \right)^{1/4},
\end{aligned} \tag{3.2}
$$

where $\tau_n = (E(U_{u_1} - U_{u_0})^2)^{1/2}$ and $b_n = n^{3/2}$. Since $\tau_n b_n^{-1} = O(1)$, it follows that

$$
P(A_{kl}^n) \le M_\varepsilon \left(P\left\{ b_n^{-1} |U_{u_1} - U_{u_0}| > a_0^{-1} \varepsilon \right\} \right)^{1/4},
$$

for some M_ε. However, from the two-dimensional convergence in distribution of $\{W_n\}$,

$$
\begin{aligned}
b_n^{-1}(U_{u_1} - U_{u_0}) &= W_n((k+1)\delta, (l+1)\delta) - W_n(k\delta, l\delta) \\
& \to_{\mathfrak{D}} N(0, \rho),
\end{aligned} \tag{3.3}
$$

where

$$
\begin{aligned}
\rho = {} & \frac{\lambda_1^2 (l+1)\delta g_c((k+1)\delta)}{(g_c(1))^2} + \frac{\lambda_2^2 (k+1)\delta g_d((l+1)\delta)}{(g_d(1))^2} \\
& - \frac{\lambda_1^2 l\delta g_c(k\delta)}{(g_c(1))^2} - \frac{\lambda_2^2 k\delta g_d(l\delta)}{(g_d(1))^2} \\
= {} & \frac{\lambda_1^2}{(g_c(1))^2} ((l+1)\delta g_c((k+1)\delta) - l\delta g_c(k\delta)) \\
& + \frac{\lambda_2^2}{(g_d(1))^2} ((k+1)\delta g_d((l+1)\delta) - k\delta g_d(l\delta)).
\end{aligned}
$$

The law of the mean applied to $R(x, y) = xg_c(y)$ and $s(x, y) = xg_d(y)$ over K shows that for some $M_1 > 0$, $\rho \le M_1 \delta$. It follows from (3.2) and (3.3) that

$$
\limsup_n \sum_{k\delta < 1} \sum_{l\delta < 1} P(A_{kl}^n) \le M_\varepsilon \delta^{-2} \left[2\left(1 - \Phi\left(\varepsilon / a_0 \sqrt{M_1 \delta}\right)\right) \right]^{1/4}, \tag{3.4}
$$

where Φ denotes the standard normal distribution function. However, the right-hand side of (3.4) approaches zero as $\delta \downarrow 0$ and thus (3.1) is valid. □

4. A simulation investigation. A simulation study was undertaken to determine the speed of convergence of $n^{3/2}(\hat{\alpha}_n - \alpha_n)$ and $n^{3/2}(\hat{\beta}_n - \beta_n)$ to normality as given in Theorem 1.1. Fixing c, d and n values of α_n and β_n were calculated using (A.1) and a sample of n^2, $N(0, 4)$ deviates was drawn using IMSL subroutines and recorded as ε_{ij}, $1 \le i, j \le n$. Since Z_{ij} can be expressed in terms of ε_{kl}, $1 \le k, l \le n$, and known parameters using (1.1), Remark 2.1 was utilized to find initial estimates $\bar{\alpha}_n$ and $\bar{\beta}_n$. Finally, $\hat{\alpha}_n$ and $\hat{\beta}_n$ were calculated using (2.1)–(2.3) and this single simulation was repeated $M = 1{,}000$ times for fixed values of c, d and n.

Three features of the convergence were examined: (1) the speed at which $\hat{\alpha}_n - \alpha_n$ and $\hat{\beta}_n - \beta_n$ converge to zero (determined by finding the sample average of $M = 1{,}000$ values), (2) the speed of convergence of the standard errors of $\hat{\alpha}_n$ and $\hat{\beta}_n$ to their respective asymptotic standard errors $(1/n^{3/2}\sqrt{g_c(1)}\,, 1/n^{3/2}\sqrt{g_d(1)}\,)$ as given in Theorem 1.1 and denoted by ASE (determined by finding the sample standard deviation of $M = 1{,}000$ values) and (3) the speed for which $n^{3/2}(\hat{\alpha}_n - \alpha_n, \hat{\beta}_n - \beta_n)$ approaches a bivariate normal with uncorrelated marginals (determined from a chi-square test).

The chi-square test was formed by dividing a univariate normal into the following three segments: $(-\infty, -1$ standard deviation), $(-1$ standard deviation, 1 standard deviation] and (1 standard deviation, ∞) having probability 0.3085375, 0.3829050 and 0.3085375, respectively. The cross product partitions the plane into nine cells and $n^{3/2}(\hat{\alpha}_n - \alpha_n)$ and $n^{3/2}(\hat{\beta}_n - \beta_n)$ independent means that the joint probability of falling into one of the nine cells can be obtained by multiplying the two marginal probabilities.

Simulations were undertaken for the four combinations of $c = 1, -1$ and $d = 1, -1$ and for $n = 5$, 10, 40, 60, 80 and 100. The program was written in FORTRAN using double precision whenever possible and run on a VAX supermini computer. However, only the two combinations $c = -1$, $d = -1$ and $c = -1$, $d = 1$ are presented in Tables 1 and 2, respectively, since the other simulations produced similar results. A replication size of $M = 1{,}000$ was deemed adequate by first examining the stability of the estimates, their standard errors and the values of the associated chi-square goodness-of-fit test for various replication sizes.

TABLE 1

Averages, standard errors and goodness-of-fit tests for $\hat{\alpha}_n$ and $\hat{\beta}_n$, $c = -1$ and $d = -1$

n	AVE($\hat{\alpha}_n - \alpha_n$) STD($\hat{\alpha}_n$)	α_n ASE($\hat{\alpha}_n$)	$\chi^2(8)$ p-value
5	−0.0362	0.8187	34.4
	0.1863	0.1679	0.000
10	−0.0144	0.9048	52.2
	0.0648	0.0594	0.000
20	−0.0040	0.9512	22.6
	0.0225	0.0210	0.004
40	−0.0009	0.9753	29.8
	0.0076	0.0074	0.000
60	−0.0002	0.9835	12.0
	0.0041	0.0040	0.149
80	−0.0003	0.9876	8.3
	0.0027	0.0026	0.403
100	−0.0003	0.9900	11.7
	0.0019	0.0019	0.165

TABLE 2

Averages, standard errors and goodness-of-fit tests for $\hat{\alpha}_n$ and $\hat{\beta}_n$, $c = -1$ and $d = 1$

n	AVE($\hat{\alpha}_n - \alpha_n$) STD($\hat{\alpha}_n$)	α_n ASE($\hat{\alpha}_n$)	$\chi^2(8)$ p-value
5	−0.0448	0.8187	116.7
	0.1893	0.1679	0.000
10	−0.0138	0.9048	40.6
	0.0667	0.0594	0.000
20	−0.0043	0.9512	30.3
	0.0218	0.0210	0.000
40	−0.0010	0.9753	22.3
	0.0073	0.0074	0.004
60	−0.0004	0.9835	12.9
	0.0042	0.0040	0.117
80	−0.0002	0.9876	9.7
	0.0027	0.0026	0.289
100	−0.0002	0.9900	10.5
	0.0019	0.0019	0.231

Listed in Tables 1 and 2 for selected values of n are the following: (1) the second column shows the sample average and standard deviation of $(\hat{\alpha}_n - \alpha_n)$, (2) column three gives the value of α_n determined from (A.1) and the ASE of $\hat{\alpha}_n$ as calculated from Theorem 1.1. Entries in column four are the values of the chi-square goodness-of-fit test statistic and the corresponding p-values for eight degrees of freedom. Only averages and standard deviations of $\hat{\alpha}_n$ are given in the tables since $\hat{\alpha}_n$ and $\hat{\beta}_n$ have the same asymptotic properties.

Both tables show that convergence of estimates $\hat{\alpha}_n - \alpha_n$ to zero occurs quickly; indeed, $\hat{\alpha}_n$ is within about 1% of α_n when $n = 10$ to 20. However, $n = 60$ to 80 was needed before standard deviations of α_n were near 1% of the corresponding ASE. Similarly, convergence of $n^{3/2}(\hat{\alpha}_n - \alpha_n, \hat{\beta}_n - \beta_n)$ to the desired bivariate normal was rather slow and $n = 60$ was needed before the p-values for our chi-square goodness-of-fit test exceeded 0.05.

REFERENCES

BASU, S. (1990). Analysis of first-order spatial bilateral ARMA models. Ph.D. dissertation, Univ. Wisconsin, Madison.

BASU, S. and REINSEL, G. C. (1992). A note on properties of spatial Yule–Walker estimators. *J. Statist. Comput. Simulation* **41** 243–255.

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

BASU, S. and REINSEL, G. C. (1994). Regression models with spatially correlated errors. *J. Amer. Statist. Assoc.* **89** 88–99.

BHATTACHARYYA, B. B., KHALIL, T. M. and RICHARDSON, G. D. (1996). Gauss–Newton estimation of parameters for a spatial autoregression model. *Statist. Probab. Lett.* **28** 173–179.

BICKEL, P. J. and WICHURA, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42** 1656–1670.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BOBKOSKI, M. J. (1983). Hypothesis testing in nonstationary time series. Ph.D. dissertation, Univ. Wisconsin, Madison.

BROWN, B. M. (1971). Martingale central limit theorems. *Ann. Math. Statist.* **42** 59–66.
CHAN, N. H. and WEI, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.* **15** 1050–1063.
COX, D. D. and LLATAS, I. (1991). Maximum likelihood type estimation for nearly nonstationary autoregressive time series. *Ann. Statist.* **19** 1109–1128.
CULLIS, B. R. and GLEESON, A. C. (1991). Spatial analysis of field experiments—an extension to two dimensions. *Biometrics* **47** 1449–1460.
FULLER, W. A. (1976). *Introduction to Statistical Time Series*. Wiley, New York.
JAIN, A. K. (1981). Advances in mathematical models for image processing. *Proc. IEEE* **69** 502–528.
KHALIL, T. M. (1991). A study of the doubly geometric process, stationary cases and a nonstationary case. Ph.D. dissertation, North Carolina State Univ., Raleigh.
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. *Comm. Statist. Theory Methods* **19** 55–81.
MCLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probab.* **2** 620–628.
PHILLIPS, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika* **74** 535–547.
TJOSTHEIM, D. (1978). Statistical spatial series modeling. *Adv. in Appl. Probab.* **10** 130–154.
TJOSTHEIM, D. (1981). Autoregressive modeling and spectral analysis of array data in the plane. *IEEE Trans. on Geoscience and Remote Sensing* **19** 15–24.
TJOSTHEIM, D. (1983). Statistical spatial series modelling II: some further results on unilateral processes. *Adv. in Appl. Probab.* **15** 562–584.
WALSH, J. B. (1986). Martingales with a multidimensional parameter and stochastic integrals in the plane. *Lectures in Probability and Statistics. Lecture Notes in Math.* **1215** 329–491. Springer, New York.
WHITE, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.* **29** 1188–1197.

B. B. BHATTACHARYYA
DEPARTMENT OF STATISTICS
NORTH CAROLINA STATE UNIVERSITY
RALEIGH, NORTH CAROLINA 27695

G. D. RICHARDSON
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CENTRAL FLORIDA
ORLANDO, FLORIDA 32816
E-MAIL: garyr@pegasus.cc.ucf.edu

L. A. FRANKLIN
COLLEGE OF BUSINESS
INDIANA STATE UNIVERSITY
TERRE HAUTE, INDIANA 47809