International Journal of Pure and Applied Mathematics
Volume 15 No. 4 2004, 447-460

PARAMETER ESTIMATION IN A NONSTATIONARY SPATIAL AUTOREGRESSION MODEL

B.B. Bhattacharyya[1], G. Richardson[2 §], J. Zhang[3]

[1]Department of Statistics
North Carolina State University
Raleigh, NC 27695-8203, USA

[2,3]Department of Mathematics
University of Central Florida
P.O. Box 161364, Orlando, FL 32816-1364, USA
[2]e-mail: garyr@pegasus.cc.ucf.edu

Abstract: The limiting distribution for Gauss-Newton estimators of (α,β) in the model $Z_{ij} = \alpha Z_{i-1,j} + \beta Z_{i,j-1} - \alpha\beta Z_{i-1,j-1} + \epsilon_{ij}$ is obtained for the case when $\alpha = 1$ and $|\beta| > 1$. The asymptotic distribution is shown to be a Gaussian process when the estimators are appropriately embedded in $D([0,1]^2)$. Results given here differ significantly from the earlier cases studied.

AMS Subject Classification: 62F12, 62M30, 60F17
Key Words: nonstationary processes, Gauss-Newton estimation, $D([0,1]^2)$

1. Introduction

The spatial model

$$Z_{ij} = \alpha Z_{i-1,j} + \beta Z_{i,j-1} - \alpha\beta Z_{i-1,j-1} + \epsilon_{ij} \tag{1.1}$$

has received considerable attention since being introduced by Martin [10]. For example, Martin [11], Cullis and Gleeson [9], Basu and Reinsel [4] used the model to analyze data in agriculture field trials. Properties of parameter estimators under stationarity assumptions can be found in Basu [1] and Basu and Reinsel [2], [3]. The limiting distribution of a sequence of Gauss-Newton

Received: June 14, 2004

§Correspondence author

estimators of $\theta' = (\alpha, \beta)$ under certain nonstationary conditions is given by Bhattacharyya [5], [6]. In particular, it is shown that when $\alpha = \beta = 1$, $n^{3/2}(\hat{\theta}_n - \theta) \xrightarrow{\mathfrak{D}} N(\underline{0}, \Gamma_1)$; whereas, $(n^{3/2}(\hat{\alpha}_n - \alpha), n(\hat{\beta}_n - \beta)) \xrightarrow{\mathfrak{D}} N(\underline{0}, \Gamma_2)$ provided $\alpha = 1$ and $|\beta| < 1$, where $\Gamma_1 = \text{diag}(2{,}2)$ and $\Gamma_2 = \text{diag}(2, 1 - \beta^2)$. It should be stressed that significant differences in the asymptotic results occur when one of the parameters exceeds one in absolute value. Indeed, the purpose of this work is to analyze the case when $\alpha = 1$ and $|\beta| > 1$.

It is assumed throughout that $\{Z_{ij}\}$ is a spatial process obeying model (1.1) and subject to the following constraints:

(A.1) $\alpha = 1$, $|\beta| > 1$.

(A.2) $Z_{i,j} = 0$ when $i \wedge j \leq 0$.

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

(A.4) $\{\overline{\alpha}_n\}$ and $\{\overline{\beta}_n\}$ are initial estimators satisfying $\overline{\alpha}_n - 1 = O_P(n^{-3/2})$ and $\overline{\beta}_n - \beta = O_P(n^{-1/2}\beta^{-n})$.

The existence of initial estimators obeying (A.4) is shown later. The sample path space $D([0,1])$, see Billingsley [8], of a process has been extended to the two parameter case $D_2 := D([0,1] \times [0,1])$ by Bickel and Wichura [7] and shown to be equipped with a metric which induces the Skorohod topology. The space D_2 is separable, complete and has Borel σ-field generated by the coordinate mappings. The Gauss-Newton estimator $\hat{\theta}_n$ of θ is the random variable defined in (2.7)-(2.8), and obeys equation (2.9). More generally, consider the random element $\Delta_n(\cdot)$ in D_2 which satisfies equation (2.10). Fix $0 < c \leq 1$, $\underline{\lambda}' = (\lambda_1, \lambda_2)\epsilon R^2$ and define for each $\underline{t}' = (t_1, t_2)\epsilon[0,1]^2$,

$$\Delta_n^c(\underline{t}) = \Delta_n(t_1, ct_2), \quad \psi_n^c(\underline{t}) = \underline{\lambda}'\text{diag}(n^{3/2}, n^{1/2}\beta^n)(\Delta_n^c(\underline{t}) - \theta)\,, \tag{1.2}$$

and denote $\psi_n^1(\underline{t})$ simply by $\psi_n(\underline{t})$. The primary result of this work is stated below.

Theorem 1.1. *Assume that model (1.1) and conditions (A.1)-(A.4) are satisfied. Let ψ_n and ψ_n^c be as defined in (1.2) and denote $c_1 = 2\lambda_1^2$ and $c_2 = (1 - \beta^2)^2\lambda_2^2$. Then:*

(i) the finite-dimensional distributions of $\{\psi_n\}$ converge in distribution to those of a mean zero Gaussian process ψ having $\text{cov}\,(\psi(\underline{s}), \psi(\underline{t})) = c_1(s_1 \wedge t_1)^2(s_2 \wedge t_2) + c_2(s_1 \wedge t_1)1_{\{1\}}(s_2 \wedge t_2)$

(ii) $\{\psi_n^c\}$ converges in distribution on D_2 to a mean zero Gaussian process ψ^c with $\text{cov}\,(\psi^c(\underline{s}), \psi^c(\underline{t})) = c_1(s_1 \wedge t_1)^2 \cdot (s_2 \wedge ct_2)$ *when $0 < c < 1$.*

In particular, $\text{diag}(n^{3/2}, n^{1/2}\beta^n)(\hat{\theta}_n - \theta) \xrightarrow{\mathfrak{D}} N(0, \Gamma)$, *where* $\Gamma = \text{diag}(2, (1 - \beta^2)^2)$.

Remark 1.1. It is shown in Lemma 3.1 (ii) that one of the terms of $\{\psi_n\}$ fails to be tight in D_2 and thus accounts for the restriction ψ_n^c of ψ_n in Theorem 1.1 (ii), $0 < c < 1$.

2. Order and Tightness Properties

Denote

$$X_{ij} = Z_{ij} - \beta Z_{i,j-1}, \quad Y_{ij} = Z_{ij} - \alpha Z_{i-1,j}\,, \tag{2.1}$$

and according to (1.1),

$$X_{ij} = \alpha X_{i-1,j} + \epsilon_{ij} \quad \text{and} \quad Y_{ij} = \beta Y_{i,j-1} + \epsilon_{ij}. \tag{2.2}$$

Employing (A.1)-(A.2),

$$X_{ij} = \sum_{k=1}^{i} \epsilon_{kj}, \quad Y_{ij} = \sum_{\ell=1}^{j} \beta^{j-\ell}\epsilon_{i\ell} \quad \text{and} \quad Z_{ij} = \sum_{k,\ell=1}^{i,j} \beta^{j-\ell}\epsilon_{k\ell}. \tag{2.3}$$

The following order properties are straightforward to verify. Here $\underline{t}' = (t_1, t_2)\epsilon[0,1]^2 = [0,1] \times [0,1]$ and $[n\underline{t}]' = ([nt_1], [nt_2])$.

Lemma 2.1.

(i) $n^{-3} \sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j}^2 = \frac{1}{2} t_1^2 t_2 \sigma^2 + O_P(n^{-1/2}),$

(ii) $n^{-1}\beta^{-2[nt_2]} \sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1}^2 = \frac{t_1\sigma^2}{(1-\beta^2)^2} + o_P(1),$

(iii) $n^{-2}\beta^{-2[nt_2]} \sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}^2 = \frac{t_1^2\sigma^2}{2(1-\beta^2)^2} + O_P(1),$

(iv) $\sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j}Y_{i,j-1} = O_P(n\beta^{[nt_2]}),$

(v) $\sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j}Z_{i-1,j-1} = O_P(n^2\beta^{[nt_2]}),$

(vi) $\sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1}Z_{i-1,j-1} = O_P(n\beta^{2[nt_2]}),$

(vii) $\sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j}\epsilon_{ij} = O_P(n^{3/2})$,

(viii) $\sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1}\epsilon_{ij} = O_P(n^{1/2}\beta^{[nt_2]})$,

(ix) $\sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}\epsilon_{ij} = O_P(n\beta^{[nt_2]})$,

(x) $\sup_{1\le i\le [nt_2]} E\left(\sum_{j=1}^{[nt_2]} X_{i-1,j}\epsilon_{ij}\right)^4 = O(n^4)$,

(xi) $\sup_{1\le i\le [nt_2]} E\left(\sum_{j=1}^{[nt_2]} Y_{i,j-1}\epsilon_{ij}\right)^4 = O(\beta^{4[nt_2]})$.

Bickel and Wichura (1971) give a sufficient condition in terms of moments in order to guarantee tightness of a sequence of random elements in D_2. Indeed, let $\underline{s}' = (s_1, s_2)$, $\underline{t}' = (t_1, t_2)\epsilon[0,1]^2$ and define $\underline{s} < \underline{t}$ when $s_i \le t_i$ and $\underline{s} \ll \underline{t}$ when $s_i < t_i$, $i = 1, 2$. Denote the rectangle set $(s_1, t_1] \times (s_2, t_2]$ by $(\underline{s}, \underline{t}]$. Let $T_n = \{(k/n, \ell/n) : k, \ell$ are integers satisfying $0 \le k, \ell \le n\}$. Then a sufficient condition for tightness of a sequence $\{V_n\}$ of random elements in D_2 is that there exist positive real numbers $\alpha_1, \alpha_2, \delta$ and M such that for each pair of disjoint rectangles $(\underline{s}, \underline{t}]$ and $(\underline{u}, \underline{v}]$ having vertices in T_n and either a common horizontal or vertical edge obeys:

$$E(|V_n(\underline{s}, \underline{t}]|^{\alpha_1}|V_n(\underline{u}, \underline{v}]|^{\alpha_2}) \le M(\mu(\underline{s}, \underline{t}]\mu(\underline{u}, \underline{v}])^{1+\delta}, \tag{2.4}$$

where μ is a finite measure on $[0,1]^2$ and $V_n(\underline{s}, \underline{t}]$ is defined below.

Another approach used below to improve the normalizing factor for tightness of a sequence is use of the maximal inequality given by Bhattacharyya [6]. Here the sequence must form a strong martingale in the sense of Walsh [12]. In particular, assume that $\mathfrak{F}_{\underline{t}}, \underline{t} \in J$ is an increasing collection of sub-σ-fields on $(\Omega, \mathfrak{F}, P)$ in the sense that $\mathfrak{F}_{\underline{s}} \subseteq \mathfrak{F}_{\underline{t}}$ when $\underline{s} < \underline{t}$, where J denotes a subset of the set of all ordered pairs of positive integers. Suppose that each $V_{\underline{t}}$ is square integrable. Then $\{V_{\underline{t}}, \mathfrak{F}_{\underline{t}}, \underline{t} \in J\}$ is called a *strong martingale* provided that for each $\underline{s}, \underline{t} \in J$, $E(V_{\underline{t}}|\mathfrak{F}_{\underline{s}}) = V_{\underline{s}}$ when $\underline{s} < \underline{t}$ and, moreover, $E(V(\underline{s}, \underline{t}]|\mathfrak{F}^*_{\underline{s}}) = 0$, where $V(\underline{s}, \underline{t}] = V_{\underline{t}} - V_{s_1t_2} - V_{t_1s_2} + V_{\underline{s}}$ and $\mathfrak{F}^*_{\underline{s}}$ denotes the smallest σ-field containing each $\mathfrak{F}_{ij}$ with either $i \le s_1$ or $j \le s_2$.

Tightness of the sequences listed below are needed to show convergence in distribution on $D_2 = D([0,1]^2)$. Since the proofs of (i)-(vi) are similar, only verification of (vi) and (vii) are presented here. It should be mentioned that not all the normalizing factors given for tightness are necessarily the best possible but are sufficient for our purposes.

Lemma 2.2. *The following sequences in D_2 are tight:*

(i) $\left\{ n^{-3} \sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j}^2 \right\}$,

(ii) $\left\{ n^{-2} \beta^{-2[nt_2]} \sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1}^2 \right\}$,

(iii) $\left\{ n^{-3} \beta^{-2[nt_2]} \sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}^2 \right\}$,

(iv) $\left\{ n^{-5/2} \beta^{-[nt_2]} \sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j} Y_{i,j-1} \right\}$,

(v) $\left\{ n^{-5/2} \beta^{-[nt_2]} \sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j} Z_{i-1,j-1} \right\}$,

(vi) $\left\{ n^{-5/2} \beta^{-2[nt_2]} \sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1} Z_{i-1,j-1} \right\}$,

(vii) $\left\{ n^{-1-\rho} \beta^{-n} \sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1} \epsilon_{ij} \right\}, \quad \rho > 0.$

Proof. (vi) Denote $V_n(\underline{t}) = n^{-5/2} \beta^{-2[nt_2]} \sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1} Z_{i-1,j-1}$ and let $\underline{\lambda}$ denote the Lebesgue measure on $[0,1]^2$. Using (2.3),

$$E(Y_{i,j-1}^2) = \sum_{\ell=1}^{j-1} \beta^{2(j-1-\ell)} \sigma^2 = \frac{(1-\beta^{2(j-1)})\sigma^2}{1-\beta^2} = O(\beta^{2j})$$

and

$$E(Z_{i-1,j-1}^2) = \sum_{k,\ell=1}^{i-1,j-1} \beta^{2(j-1-\ell)} \sigma^2 = (i-1) \left(\frac{1-\beta^{2(j-1)}}{1-\beta^2} \right) \sigma^2 = O(i\beta^{2j}).$$

Verification is given when $(\underline{s}, \underline{t}]$ and $(\underline{u}, \underline{v}]$ have a common horizontal edge since the vertical case is analogous. Employing Cauchy's inequality, observe that

$$E(|V_n(\underline{s},\underline{t}]|\,|V_n(\underline{u},\underline{v}]|) \le Kn^{-5}\beta^{-2[nt_2]}\beta^{-2[nv_2]} \sum_{i,j=[n\underline{s}]+1}^{[n\underline{t}]} \sum_{i'=[ns_1]+1}^{[nt_1]} \sum_{j'=[nt_2]+1}^{[nv_2]} \beta^{2(j+j')}(ii')^{1/2} \le M\lambda(\underline{s},\underline{t}]\lambda(\underline{u},\underline{v}]$$

and thus by (2.4) $\{V_n\}$ is tight in D_2.

(vii) In order to obtain a sharper normalizing factor needed later, the moment criterion used to prove tightness in (vi) needs to be replaced by the technique used in Bhattacharyya [6], p. 1721. Define $W_n(\underline{t}) = n^{-1-\rho}\beta^{-n}\sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}\epsilon_{ij}$. Given $\delta > 0$, consider rectangles $R_{k\ell} = [k\delta,(k+1)\delta) \times [\ell\delta,(\ell+1)\delta)$ and for $\epsilon > 0$, define $A_{k\ell}^n = \{\sup_{\underline{t}\in R_{k\ell}} |W_n(\underline{t}) - W_n(k\delta,\ell\delta)| > \epsilon\}$. Tightness of $\{W_n\}$ in D_2 is shown by verifying that for each $\epsilon > 0$ and $\nu > 0$ there exists a $\delta > 0$ such that $\overline{\lim}_n \sum_{k\delta<1,\ell\delta<1} P(A_{k\ell}^n) < \nu$. Denote $\mu_0 = ([nk\delta],[n\ell\delta])$, $\mu_1 = ([n(k+1)\delta],[n(\ell+1)\delta])$ and let $J = \{(i,j) : \mu_0 < (i,j) < \mu_1\}$. Define for each $a = (a_1,a_2)\epsilon J$, $\mathfrak{F}_a(\mathfrak{F}_a^*)$ to be the smallest σ-field making each $\epsilon_{ij}, i \le a_1$ and $j \le a_2$ (either $i \le a_1$ or $j \le a_2$) measurable, respectively.

Let $U_a = \sum_{i,j=1}^{a} Z_{i-1,j-1}\epsilon_{ij}, a\epsilon J$. Then $\{U_a - U_{\mu_0}, \mathfrak{F}_a, a\epsilon J\}$ is a strong martingale in the sense of Walsh [12] as discussed above. According to the maximal inequality of Bhattacharyya [6], Lemma 1.1, there exist positive real numbers a_0 and A_0 for which

$$P(A_{k\ell}^n) \le A_0\tau_n\epsilon^{-1}n^{-1-\rho}\beta^{-n}(P\{n^{-1-\rho}\beta^{-n}|U_{\mu_1} - U_{\mu_0}| \ge \epsilon a_0^{-1}\})^{\frac{1}{2}} + A_0(\tau_n\epsilon^{-1}n^{-1-\rho}\beta^{-n})^{3/2}(P\{n^{-1-\rho}\beta^{-n}|U_{\mu_1} - U_{\mu_0}| \ge \epsilon a_0^{-1}\})^{\frac{1}{4}},$$

where $\tau_n = (E(U_{\mu_1} - U_{\mu_0})^2)^{\frac{1}{2}} = O(n\beta^n)$. Since $\tau_n n^{-1-\rho}\beta^{-n} \to 0$, it follows that $P(A_{k\ell}^n) \to 0$ as $n \to \infty$ and thus $\overline{\lim}_n \sum_{k\delta<1,\ell\delta<1} P(A_{k\ell}^n) = 0$. Hence $\{W_n\}$ is tight in D_2. □

Under (A.1)-(A.4), let $\theta' = (\alpha,\beta), \overline{\theta}_n' = (\overline{\alpha}_n, \overline{\beta}_n)$ and define

$$F_{ij}'(\overline{\theta}_n) = (X_{i-1,j} + (\beta - \overline{\beta}_n)Z_{i-1,j-1}, Y_{i,j-1} + (\alpha - \overline{\alpha}_n)Z_{i-1,j-1}),$$
$$G_n(\underline{t}) = \sum_{i,j=1}^{[n\underline{t}]} F_{ij}(\overline{\theta}_n)F_{ij}'(\overline{\theta}_n),$$
$$A_n(\underline{t}) = \mathrm{diag}(n^{-\frac{3}{2}}, n^{-\frac{1}{2}}\beta^{-[nt_2]}),$$
$$B(\underline{t}) = \sigma^2\mathrm{diag}\left(\frac{t_1^2 t_2}{2}, \frac{t_1}{(1-\beta^2)^2}\right),$$
$$R_{ij}(\overline{\theta}_n) = -(\alpha - \overline{\alpha}_n)(\beta - \overline{\beta}_n)Z_{i-1,j-1}. \tag{2.5}$$

Let

$$f_{ij}(a,b) = aZ_{i-1,j} + bZ_{i,j-1} - abZ_{i-1,j-1}$$

and

$$F'_{ij}(a,b) = (\partial f_{ij}(a,b)/\partial a, \partial f_{ij}(a,b)/\partial b).$$

Expanding $f_{ij}(\theta)$ about $\overline{\theta}_n$ in model (1.1) provides

$$Z_{ij} = f_{ij}(\overline{\theta}_n) + F'_{ij}(\overline{\theta}_n)(\theta - \overline{\theta}_n) + R_{ij}(\overline{\theta}_n) + \epsilon_{ij}. \quad (2.6)$$

Define

$$\hat{\delta}_n = G_n^{-1}(1,1) \sum_{i,j=1}^{n} F_{ij}(\overline{\theta}_n)(Z_{ij} - f_{ij}(\overline{\theta}_n)). \quad (2.7)$$

Then

$$\hat{\theta}_n = \hat{\delta}_n + \overline{\theta}_n \quad (2.8)$$

is called the "one step Gauss-Newton estimator" of θ. Substituting (2.6) into (2.7) shows that $\hat{\theta}_n$ satisfies:

$$A_n^{-1}(1,1)(\hat{\theta}_n - \theta) = (A_n(1,1)G_n(1,1)A_n(1,1))^{-1}A_n(1,1) \times \sum_{i,j=1}^{n} F_{ij}(\overline{\theta}_n)(R_{ij}(\overline{\theta}_n) + \epsilon_{ij}). \quad (2.9)$$

Moreover, let $\Delta_n(\cdot)$ denote the random element in D_2 obeying:

$$A_n^{-1}(1,1)(\Delta_n(\underline{t}) - \theta) = (A_n(1,1)G_n(1,1)A_n(1,1))^{-1}A_n(1,1) \times \sum_{i,j=1}^{[n\underline{t}]} F_{ij}(\overline{\theta}_n)(R_{ij}(\overline{\theta}_n) + \epsilon_{ij}). \quad (2.10)$$

Remark 2.1. Unfortunately addition in $D([0,1])$ equipped with the Skorohod metric is not a continuous operation (Billingsley [8], p. 137). However, the following results are valid and used without further mention:

(i) $U_n \xrightarrow{\mathfrak{D}} a$ in D_2, $V_n \xrightarrow{\mathfrak{D}} V$ in D_2 implies that $U_n + V_n \xrightarrow{\mathfrak{D}} a + V$ in D_2

(ii) $U_n \xrightarrow{\mathfrak{D}} a$ in R and $V_n \xrightarrow{\mathfrak{D}} V$ in D_2 implies that $U_nV_n \xrightarrow{\mathfrak{D}} aV$ in D_2, where "a" denotes the constant random variable.

Lemma 2.3. *Employing the notations of (2.5):*

(i) $A_n(\underline{t})G_n(\underline{t})A_n(\underline{t}) \xrightarrow{\mathfrak{D}} B(\underline{t})$ *in* D_2^4.

(ii) $A_n(\underline{t}) \sum_{i,j=1}^{[n\underline{t}]} F_{ij}(\overline{\theta}_n)R_{ij}(\overline{\theta}_n) \xrightarrow{\mathfrak{D}} 0$ *in* D_2^2.

$$(iii)\ \left(n^{-\frac{3}{2}}(\beta-\overline{\beta}_n)\sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}\epsilon_{ij},\, n^{-\frac{1}{2}}\beta^{-n}(\alpha-\overline{\alpha}_n)\sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}\epsilon_{ij}\right) \xrightarrow{\mathfrak{D}} 0.$$

Proof. Verification of (ii) and (iii) are supplied here.

(ii) Note that

$$\begin{aligned}
&A_n(\underline{t})\sum_{i,j=1}^{[n\underline{t}]} F_{ij}(\overline{\theta}_n)R(\overline{\theta}_n)\\
&\quad= -(\alpha-\overline{\alpha}_n)(\beta-\overline{\beta}_n)\Big(n^{-\frac{3}{2}}\sum_{i,j=1}^{[n\underline{t}]}(X_{i-1,j}Z_{i-1,j-1}+(\beta-\overline{\beta}_n)Z^2_{i-1,j-1}),\\
&\qquad n^{-\frac{1}{2}}\beta^{-[n\underline{t}_2]}\sum_{i,j=1}^{[n\underline{t}]}(Y_{i,j-1}Z_{i-1,j-1}+(\alpha-\overline{\alpha}_n)Z^2_{i-1,j-1})\Big)'\\
&\quad= (S_n(\underline{t})+T_n(\underline{t})+U_n(\underline{t})+V_n(\underline{t}))'.
\end{aligned}$$

It follows from (A.4), Lemma 2.1 and Lemma 2.2 that

$$\begin{aligned}
S_n(\underline{t}) &= -n^{\frac{3}{2}}(\alpha-\overline{\alpha}_n)(\beta-\overline{\beta}_n)\beta^n\cdot n^{-3}\beta^{-n}\\
&\quad\times\sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j}Z_{i-1,j-1},\, n^{\frac{3}{2}}(\alpha-\overline{\alpha}_n)(\beta-\overline{\beta}_n)\beta^n\\
&= O_P(n^{-\frac{1}{2}}),\{n^{-3}\beta^{-n}\sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j}Z_{i-1,j-1}\}
\end{aligned}$$

is tight in D_2 and converges to zero in probability for each fixed $\underline{t}\epsilon[0,1]^2$. Hence $S_n \xrightarrow{\mathfrak{D}} 0$ in D_2; likewise $T_n \xrightarrow{\mathfrak{D}} 0$ in D_2.

Observe that

$$\begin{aligned}
U_n(\underline{t}) &= -n^2(\alpha-\overline{\alpha}_n)(\beta-\overline{\beta}_n)\beta^n\cdot n^{-\frac{5}{2}}\beta^{-n}\beta^{-[nt_2]}\sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1}Z_{i-1,j-1},\\
&\qquad n^2(\alpha-\overline{\alpha}_n)(\beta-\overline{\beta}_n)\beta^n = O_P(1)
\end{aligned}$$

and according to Lemma 2.1 and Lemma 2.2,

$$\{n^{-5/2}\beta^{-n}\beta^{-[nt_2]}\sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1}Z_{i-1,j-1}\}$$

is tight in D_2 and converges to zero in probability for each fixed $\underline{t} \in [0,1]^2$. As in the finite-dimensional case, one can show that the product of a bounded in probability sequence of random variables with a sequence of random elements that converges to zero in D_2 also converges to zero in D_2. Hence $U_n \xrightarrow{\mathfrak{D}} 0$ in D_2; likewise, $V_n \xrightarrow{\mathfrak{D}} 0$ in D_2 and the desired conclusion follows.

(iii) It follows from (A.4) that $n^{\frac{1}{2}}(\beta - \overline{\beta}_n)\beta^n = O_P(1)$ and by Lemma 2.1 and Lemma 2.2 the sequence $\{n^{-2}\beta^{-n}\sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}\epsilon_{ij}\}$ converges to zero in probability for each fixed $\underline{t} \in [0,1]^2$ and is tight in D_2. Hence $n^{-\frac{3}{2}}(\beta - \overline{\beta}_n)\sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}\epsilon_{ij} \xrightarrow{\mathfrak{D}} 0$ in D_2. Similarly, $n^{-\frac{1}{2}}\beta^{-n}(\alpha-\overline{\alpha}_n)\sum_{i,j=1}^{[n\underline{t}]} Z_{i-1,j-1}\epsilon_{ij} \xrightarrow{\mathfrak{D}} 0$ in D_2 and (iii) follows. □

3. Proof of Theorem 1.1

Given $\underline{\lambda}\epsilon R^2$ and using the notations given in (2.5), define $(a_n, b_n) = \underline{\lambda}'[A_n(1,1)G_n(1,1)A_n(1,1)]^{-1}$. Employing (2.10), $\underline{\lambda}'A_n^{-1}(1,1)(\Delta_n(\underline{t})-\theta) = (a_n,b_n)A_n(1,1)\sum_{i,j=1}^{[n\underline{t}]} F_{ij}(\overline{\theta}_n)(R_{ij}(\overline{\theta}_n)+\epsilon_{ij})$. According to Lemma 2.3 (i),

$$(a_n, b_n) \xrightarrow{\mathfrak{D}} (a,b) = \sigma^{-2}(2\lambda_1, (1-\beta^2)^2\lambda_2) \tag{3.1}$$

in R^2 and thus it follows from Lemma 2.3 (ii) that the k-dimensional distributions of $\{\underline{\lambda}'A_n^{-1}(1,1)(\Delta_n(\underline{t}) - \theta)\}$ converge in distribution provided the corresponding k-dimensional distributions of

$$\{(a,b)A_n(1,1)\sum_{i,j=1}^{[n\underline{t}]} F_{ij}(\overline{\theta}_n)\epsilon_{ij}\}$$

converge, and the limits coincide.

Some preliminary notations and results are needed. According to Bickel and Wichura [7], a necessary condition for a sequence $\{V_n\}$ in D_2 to be tight is for each $\epsilon > 0$, $\lim_{\delta\downarrow 0}\overline{\lim}\, P\{w''_\delta(V_n) \geq \epsilon\} = 0$, where the notation is described below. Let $x \in D_2$ and $t \in [0,1]$. Then $x_t^{(1)} : [0,1] \to R$ is defined by $x_t^{(1)}(t_2) = x(t,t_2)$

and likewise $x_t^{(2)}(t_1) = x(t_1, t)$. Moreover, $\|x_t^{(1)} - x_s^{(1)}\| = \sup_{t_2 \in [0,1]} |x(t, t_2) - x(s, t_2)|$ and $\|x_t^{(2)} - x_s^{(2)}\| = \sup_{t_1 \in [0,1]} |x(t_1, t) - x(t_1, s)|$. Given $\delta > 0$ and $x \in D_2$, denote $w_\delta''^{(1)}(x) = \sup\{\|x_t^{(1)} - x_s^{(1)}\| \wedge \|x_u^{(1)} - x_t^{(1)}\| : s \le t \le u, u - s \le \delta\}$ and similarly for $w_\delta''^{(2)}$. Finally, $w_\delta''(x) = w_\delta''^{(1)}(x) \vee w_\delta''^{(2)}(x)$.

Recall the definitions of X_{ij} and Y_{ij} defined in (2.3), where $|\beta| > 1$. The following notations are used:

$$V_n(\underline{t}) = n^{-\frac{1}{2}} \beta^{-n} \sum_{i,j=1}^{[n\underline{t}]} Y_{i,j-1} \epsilon_{ij} \,,$$

$$V_n^c(\underline{t}) = n^{-\frac{1}{2}} \beta^{-n} \sum_{i,j=1}^{[nt_1],[nct_2]} Y_{i,j-1} \epsilon_{ij}, \quad 0 < c < 1 \,,$$

$$W_n(\underline{t}) = \sum_{i,j=1}^{[n\underline{t}]} \left(a n^{-\frac{3}{2}} X_{i-1,j} + b n^{-\frac{1}{2}} \beta^{-n} Y_{i,j-1} \right) \epsilon_{ij}. \tag{3.2}$$

Lemma 3.1. *Let V_n, V_n^c and W_n be random elements in D_2 as defined in (3.1) and (3.2). Then:*

(i) the finite-dimensional distributions of $\{W_n\}$ converge in distribution to a mean zero normal random vector W with

$$\operatorname{cov}(W(\underline{s}), W(\underline{t})) = \sigma^4 \Big(\frac{a^2}{2} (s_1 \wedge t_1)^2 (s_2 \wedge t_2) + \frac{b^2 (s_1 \wedge t_1)}{(1-\beta^2)^2} 1_{\{1\}}(s_2 \wedge t_2) \Big) \,.$$

(ii) $\{V_n\}$ fails to be tight in D_2.

(iii) $V_n^c \xrightarrow{\mathfrak{D}} 0$ in D_2 when $0 < c < 1$.

Proof. Verification of (i) is omitted since it follows the steps given by Bhattacharyya [6], p. 1719 by employing a Martingale Central Limit Theorem.

(ii) Using the notations above, it suffices to show that

$$\lim_{\delta \downarrow 0} \overline{\lim}_n \, P\{w_\delta''(V_n) \ge \epsilon\} = 0$$

fails to hold. Suppose that $\epsilon > 0$ is any arbitrary positive number. Then for

$$n \ge 2\delta^{-1}, P\{w_\delta''(V_n) \ge \epsilon\} \ge P\Bigg\{ \Big| V_n \left(1, 1 - \frac{1}{n}\right) - V_n \left(1, 1 - \frac{2}{n}\right) \Big| \wedge \Big| V_n(1,1) - V_n \left(1, 1 - \frac{1}{n}\right) \Big| \ge \epsilon \Bigg\}. \tag{3.3}$$

Observe that $V_n\left(1,1-\frac{1}{n}\right) - V_n\left(1,1-\frac{2}{n}\right) = n^{-\frac{1}{2}}\beta^{-n}\sum_{i=1}^{n} Y_{i,n-2}\epsilon_{i,n-1}$ and $V_n(1,1) - V_n\left(1,1-\frac{1}{n}\right) = n^{-\frac{1}{2}}\beta^{-n}\sum_{i=1}^{n} Y_{i,n-1}\epsilon_{in}$. Note that $Y_{i,n-1}\epsilon_{in}$ and $Y_{i',n-1}\epsilon_{i'n}$ are independent random variables when $i \neq i'$. One can use the Lindeberg-Feller Central Limit Theorem to show that the triangular array

$$n^{-\frac{1}{2}}\beta^{-n}\sum_{i=1}^{n} Y_{i,n-1}\epsilon_{in} \xrightarrow{\mathfrak{D}} N(0,s_1^2)$$

and

$$n^{-\frac{1}{2}}\beta^{-n}\sum_{i=1}^{n} Y_{i,n-2}\epsilon_{i,n-1} \xrightarrow{\mathfrak{D}} N(0,s_2^2)\,,$$

where $s_1^2 = \sigma^4/\beta^2(\beta^2-1)$ and $s_2^2 = \sigma^4/\beta^4(\beta^2-1)$. More generally, an application of the Crámer-Wold device shows that

$$\left(V_n\left(1,1-\frac{1}{n}\right) - V_n\left(1,1-\frac{2}{n}\right), V_n(1,1) - V_n\left(1,1-\frac{1}{n}\right)\right) \xrightarrow{\mathfrak{D}} N(\underline{0},\textstyle\sum),$$

where $\sum = \frac{\sigma^4}{\beta^2(\beta^2-1)}\mathrm{diag}(1,\beta^{-2})$. It follows from (3.3) that

$$\overline{\lim}_n P\{w''_\delta(V_n) \geq \epsilon\} \geq P\{|N(0,s_1^2)| \wedge |N(0,s_2^2)| \geq \epsilon\} > 0$$

and thus $\{V_n\}$ fails to be tight in D_2.

(iii) Let $\|\cdot\|$ denote the sup-norm and d the Skorohod metric on D_2. Since $d(x,y) \leq \|x-y\|$ on D_2, it follows that for $\delta > 0, P\{d(V_n^c,0) \geq \delta\} \leq P\{\|V_n^c\| \geq \delta\} \leq E\|V_n^c\|/\delta$ and thus it suffices to show that $E\|V_n^c\| \to 0$ as $n \to \infty$. However,

$$\|V_n^c\| = \sup_{\underline{t}\epsilon[0,1]^2} |V_n^c(\underline{t})| \leq |\beta|^{-n}\sum_{i,j=1}^{n,[nc]} |Y_{i,j-1}\epsilon_{ij}|$$

and thus $E\|V_n^c\| = O(n|\beta|^{[nc]-n})$. Since $0 < c < 1$ and $|\beta| > 1$, it follows that $V_n^c \xrightarrow{\mathfrak{D}} 0$ in D_2. □

Proof of Theorem 1.1 (i) As mentioned at the beginning of Section 3, the finite-dimensional distributions of $\{\psi_n(\underline{t})\}$ converge in distribution whenever those of $\left\{(a,b)A_n(1,1)\sum_{i,j=1}^{[n\underline{t}]} F_{ij}(\overline{\theta}_n)\epsilon_{ij}\right\}$ converge, and the two limits coincide. Employing Lemma 2.3 (iii), the finite-dimensional distributions of the latter sequence converge in distribution exactly when those of the sequence

$\{W_n\}$ defined in (3.2) converge. This fact combined with Lemma 3.1 (i) and (3.1) shows that the finite-dimensional distributions of $\{\psi_n(\underline{t})\}$ converge in distribution to the desired limit.

(ii) Since $0 < c < 1$, it follows from (i) that the finite-dimensional distributions of $\{\psi_n^c\}$ converge in distribution to those of a mean zero Gaussian process ψ^c having $\operatorname{cov}(\psi^c(\underline{s}), \psi^c(\underline{t})) = c_1(s, \wedge t_1)^2(s_2 \wedge ct_2)$. Moreover, according to (3.1) and Lemma 2.3 and Lemma 3.1, $\psi_n^c \xrightarrow{\mathfrak{D}} \psi^c$ in D_2 provided $a_n n^{-\frac{3}{2}} \sum_{i,j=1}^{[nt_1],[nct_2]} X_{i-1,j}\epsilon_{ij} \xrightarrow{\mathfrak{D}} \psi^c(\underline{t})$ in D_2. An argument similar to that given by Bhattacharyya [6], p. 1720, shows that $T_n(\underline{t}) = an^{-\frac{3}{2}} \sum_{i,j=1}^{[n\underline{t}]} X_{i-1,j}\epsilon_{ij} \xrightarrow{\mathfrak{D}} T(\underline{t})$ in D_2, where T is a Gaussian process having $\operatorname{cov}(T(\underline{s}), T(\underline{t})) = \frac{a^2}{2}(s_1 \wedge t_1)^2(s_2 \wedge t_2) = 2\lambda_1^2(s_1 \wedge t_1)^2(s_2 \wedge t_2)$ and the desired conclusion follows. □

Finally, it is established that initial estimators $\{\overline{\alpha}_n\}$ and $\{\overline{\beta}_n\}$ obeying (A.4) exist. According to (2.2), $Y_{ij} = \beta Y_{i,j-1} + \epsilon_{ij}$; define the least squares estimator $\overline{\beta}_n = \sum_{i,j=1}^n Y_{ij}Y_{i,j-1} \big/ \sum_{i,j=1}^n Y_{i,j-1}^2$ and note that $Y_{ij} = Z_{ij} - \alpha Z_{i-1,j} = Z_{ij} - Z_{i-1,j}$ is observable since $\alpha = 1$. It follows that $n^{\frac{1}{2}}\beta^n(\overline{\beta}_n - \beta) = n^{-\frac{1}{2}}\beta^{-n} \sum_{i,j=1}^n Y_{i,j-1}\epsilon_{ij} / n^{-1}\beta^{-2n} \sum_{i,j=1}^n Y_{i,j-1}^2$ and, moreover, the denominator converges to $\sigma^2/(1-\beta^2)^2$ in probability and the numerator is $O_P(1)$ by Lemma 2.1. Hence $\overline{\beta}_n - \beta = O_P\left(n^{-\frac{1}{2}}\beta^{-n}\right)$.

Again by (2.2), $X_{ij} = \alpha X_{i-1,j} + \epsilon_{ij}$; however,

$$\frac{\sum_{i,j=1}^n X_{ij}X_{i-1,j}}{\sum_{i,j=1}^n X_{i-1,j}^2}$$

is not a valid estimator because $X_{ij} = Z_{ij} - \beta Z_{i,j-1}$ is not observable since β is unknown. This leads to the estimator

$$\overline{\alpha}_n = \frac{\sum_{i,j=1}^n (Z_{ij} - \overline{\beta}_n Z_{i,j-1})(Z_{i-1,j} - \overline{\beta}_n Z_{i-1,j-1})}{\sum_{i,j=1}^n (Z_{i-1,j} - \overline{\beta}_n Z_{i-1,j-1})^2}. \tag{3.4}$$

Note that $Z_{ij} - \overline{\beta}_n Z_{i,j-1} = Z_{ij} - \beta Z_{i,j-1} + (\beta - \overline{\beta}_n)Z_{i,j-1} = X_{ij} + (\beta - \overline{\beta})Z_{i,j-1}$. Substituting $X_{ij} - X_{i-1,j} = \epsilon_{ij}$ and $Y_{i,j-1} = Z_{i,j-1} - Z_{i-1,j-1}$ into (3.4) results in

$$n^{\frac{3}{2}}(\overline{\alpha}_n - 1) = n^{-\frac{3}{2}} \frac{\sum_{i,j=1}^n (X_{i-1,j} + (\beta - \overline{\beta}_n)Z_{i-1,j-1})(\epsilon_{ij} + (\beta - \overline{\beta}_n)Y_{i,j-1})}{n^{-3}\sum_{i,j=1}^n (X_{i-1,j} + (\beta - \overline{\beta}_n)Z_{i-1,j-1})^2}.$$

Expanding the products and applying Lemma 2.1 shows that the denominator converges to $\sigma^2/2$ in probability and the numerator is $O_P(1)$. Hence $\overline{\alpha}_n - 1 = O_P(n^{-\frac{3}{2}})$ and (A.4) is satisfied. $\square$

References

[1] S. Basu, *Analysis of First-Order Spatial Bilateral ARMA Models*, Ph.D. Dissertation, Univ. Wisconsin, Madison (1990).

[2] S. Basu, G.C. Reinsel, A note on properties of spatial Yule-Walker estimators, *J. Statist. Comput. Simulation*, **41** (1992), 243-255.

[3] S. Basu, G.C. Reinsel, Properties of the spatial unilateral first-order ARMA model, *Adv. in Appl. Probabl.*, **25** (1993), 631-648.

[4] S. Basu, G.C. Reinsel, Regression models with spatially correlated errors, *J. Amer. Statist. Assoc.*, **89** (1994), 88-99.

[5] B.B. Bhattacharyya, T.M. Khalil, G.D. Richardson, Gauss-Newton estimation of parameters for a spatial autoregression model, *Statist. Probab. Lett.*, **28** (1996), 173-179.

[6] B.B. Bhattacharyya, G.D. Richardson, L.A. Franklin, Asymptotic inference for near unit roots in spatial autoregression, *Ann. Statist.*, **25** (1997), 1709-1724.

[7] P.J. Bickel, M.J. Wichura, Convergence criteria for multiparameter stochastic processes and some applications, *Ann. Math. Statist.*, **42** (1971), 1656-1670.

[8] P. Billingsley, *Convergence of Probability Measures*, Second Edition, Wiley, New York (1999).

[9] B.R. Cullis, A.C. Gleeson, Spatial analysis of field experiments — an extension to two dimensions, *Biometrics*, **47** (1991), 1449-1460.

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

[11] R.J. Martin, The use of time-series models and methods in the analysis of agricultural field trials, *Comm. Statist. Theory Methods*, **19** (1990), 55-81.

[12] J.B. Walsh, Martingales with a multidimensional parameter and stochastic integrals in the plane, *Lectures in Probability and Statistics Lecture Notes in Math.*, **1215**, Springer, New York (1986), 329-491.