Bernoulli **7**(4), 2001, 657–668

Density estimation for spatial linear processes

MARC HALLIN[1], ZUDI LU[2] and LANH TAT TRAN[3]

[1]*I.S.R.O., Université Libre de Bruxelles, B-1050 Bruxelles, Belgium. E-mail: mhallin@ulb.ac.be*
[2]*Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, China. E-mail: zdlu@iss04.iss.ac.cn*
[3]*Department of Mathematics, Indiana University, Bloomington, IN 47405-5701, USA. E-mail: tran@indiana.edu*

The problem of estimating the marginal densities of a spatial linear process, observed over a grid of $\mathbb{Z}^N$, is considered. Under general conditions, kernel density estimators computed at any k-tuple of sites are shown to be asymptotically multivariate normal. Their limiting covariance matrix is also computed. Despite the huge development of nonparametric estimation methods in the analysis of time series data, little has so far been done to introduce them into the context of random fields. The generalization is far from trivial since the points of $\mathbb{Z}^N$ do not have a natural ordering when $N > 1$. No mixing conditions are required, but linearity is assumed.

Keywords: bandwidth; density estimation; kernel; spatial process

1. Introduction

Data collected at different sites on the surface of the earth often have an associated two- or three-dimensional coordinate. Spatial data arise in various areas of research, including agricultural field trials, astronomy, econometrics, epidemiology, environmental science, geology, hydrology, image analysis, meteorology, neurology and oceanography. Numerous applications of spatial models and important developments in the general area of spatial statistics are found in Cressie (1991), Basawa (1996a; 1996b), Guyon (1995), and the references therein. Despite the attention devoted to models of this type, their statistical analysis seldom goes beyond the traditional second-order approach, which is somewhat surprising in view of the huge development of nonparametric estimation methods.

We assume a simple setting where these sites are points $\mathbf{s} = (s_1, \ldots, s_N) \in \mathbb{Z}^N$, $N \geqslant 1$. A spatial linear process $X_{\mathbf{n}}$, $\mathbf{n} \in \mathbb{Z}^N$, is defined by

$$X_{\mathbf{n}} := \sum_{\mathbf{s}\in\mathbb{Z}^N} \psi_{\mathbf{s}} Z_{\mathbf{n}-\mathbf{s}}. \tag{1.1}$$

The $\psi_{\mathbf{s}}$ are coefficients and the $Z_{\mathbf{n}}$ are real-valued independently and identically distributed (i.i.d.) random variables with zero mean and variance σ^2. Convergence, in (1.1) and below, is to be understood in the quadratic mean. An in-depth study of the theoretical properties of such models has been carried out by Tjøstheim (1978; 1983).

1350–7265

Assume that we observe $\{X_{\mathbf{i}}\}$ on $I_{\mathbf{n}}$, where $I_{\mathbf{n}}$ is a rectangular region given by

$$I_{\mathbf{n}} = \{\mathbf{i} : \mathbf{i} \in \mathbb{Z}^N, 1 \leqslant i_k \leqslant n_k, k = 1, \ldots, N\}.$$

We use $\mathbf{1}$ to denote the lattice point with all coordinates equal to one. The letter C will denote constants whose values are unimportant. We write $\mathbf{n} \to \infty$, where $\mathbf{n} := (n_1, \ldots, n_N)$, if $\min_{1 \leqslant k \leqslant N}\{n_k\} \to \infty$ with $|n_j/n_k| < C$ for some $0 < C < \infty$, $1 \leqslant j, k \leqslant N$, where C is a generic constant (independent of $\mathbf{n}$). Denote by $\hat{\mathbf{n}}$ the product $n_1 \ldots n_N$ and note that $\hat{\mathbf{n}} \leqslant C\overline{\mathbf{n}}^N$, where $\overline{\mathbf{n}} = (n_1 + n_2 + \ldots + n_N)/N$. All limits are taken as $\mathbf{n} \to \infty$ unless indicated otherwise. The integer part of a number z is denoted by $[z]$, the Euclidean norm $\sqrt{n_1^2 + \ldots + n_N^2}$ of $\mathbf{n}$ by $\|\mathbf{n}\|$.

Define the kernel density estimator $f_{\mathbf{n}}$ of f by

$$f_{\mathbf{n}} : x \mapsto f_{\mathbf{n}}(x) := (\hat{\mathbf{n}} b_{\mathbf{n}})^{-1} \sum_{\mathbf{i} \in I_{\mathbf{n}}} K((x - X_{\mathbf{i}})/b_{\mathbf{n}}), \qquad x \in \mathbb{R},$$

where K is some kernel function and $b_{\mathbf{n}}$ a bandwidth tending to zero as $\mathbf{n}$ tends to infinity.

Our objective in this paper is to investigate the limiting distribution of $f_{\mathbf{n}}$. Under general conditions, we show (Theorem 2.1) that $(f_{\mathbf{n}}(x_1), \ldots, (f_{\mathbf{n}}(x_k))$, where $x_1, \ldots, x_k$ are k arbitrary points in $\mathbb{R}$, is asymptotically multivariate normal. This result is useful for the construction of asymptotic confidence intervals for f, but also has other potential applications, such as adaptive estimation, optimal testing, and forecasting, which we do not investigate here. The asymptotic normality of $f_{\mathbf{n}}$ has been established by Hallin and Tran (1996) for the case $N = 1$. Tran (1990) and Tran and Yakowitz (1993) have investigated density estimators for strongly mixing random fields. Mixing assumptions, however, are particularly difficult, if not impossible, to verify for spatial processes, even for linear ones. The class of linear processes considered here is fairly general, and contains processes that are not strongly mixing.

The points of $\mathbb{Z}^N$ do not have a natural ordering. As a result, most techniques available for one-dimensional processes do not extend to random fields. Differences as well as similarities between spatial series ($N > 1$) and time series ($N = 1$) are highlighted in Tjøstheim (1987), who also discusses some of the difficulties inherent in the spatial context and provides some caveats. For background reading on random fields and spatial statistics, the reader is referred to Whittle (1954; 1963), Tjøstheim (1978; 1983; 1987), Ripley (1981), Possolo (1991), Anselin and Florax (1995), Guyon (1995), Dedecker (1998) and the references therein.

Our paper is organized as follows. In Section 2, we deal with some preliminaries and state all our assumptions and the main theorem. In Section 3, we present in detail the blocking of spatial random variables which plays a crucial role in the proof of the main theorem. The proof of the theorem and preliminary lemmas are gathered in Section 4. Finally, the proofs of lemmas requiring complicated arguments are presented in the Appendix.

2. Assumptions and main result

The following assumptions are made on the kernel K, the linear model (1.1), and the bandwidth $b_{\mathbf{n}}$.

***Assumption* 1.** *The kernel function K is a density function with integrable radial majorant $Q(x) := \sup_{\{|y| \geqslant |x|\}} |K(y)|$, $x \in \mathbb{R}$, and satisfies the Lipschitz condition $|K(x) - K(y)| \leqslant C|x - y|$, for $x, y \in \mathbb{R}$.*

***Assumption* 2.** *The density of $X_{\mathbf{i}}$ is uniformly bounded. The coefficients of the linear representation* (1.1) *satisfy $|\psi_{\mathbf{s}}| \leqslant C\|\mathbf{s}\|^{-a}$ for some $a > \max\{N + 3, 2N + \frac{1}{2}\}$. In addition, $Z_{\mathbf{i}}$ has mean zero, finite variance σ^2 and absolutely integrable characteristic function.*

***Assumption* 3.** *The bandwidth $b_{\mathbf{n}}$ tends to zero slowly enough that $\hat{\mathbf{n}} b_{\mathbf{n}}^{(2a-1+6N)/(2a-1-4N)} \to \infty$.*

As an example, consider the two-dimensional autoregressive model

$$X_{(n_1,n_2)} - \alpha X_{(n_1-1,n_2)} - \gamma X_{(n_1,n_2-1)} = Z_{(n_1,n_2)}, \qquad \mathbf{n} = (n_1, n_2) \in \mathbb{Z}^2, \tag{2.1}$$

for some α, γ with $|\alpha| + |\gamma| < 1$, where $Z_{(n_1,n_2)}$ are i.i.d. random variables with mean 0 and variance σ^2. The stationary solution of (2.1) is given in Kulkarni (1992) as

$$X_{\mathbf{n}} = \sum_{s_1=0}^{\infty} \sum_{s_2=0}^{\infty} \binom{s_1 + s_2}{s_1} \alpha^{s_1} \gamma^{s_2} Z_{(n_1-s_1,n_2-s_2)} = \sum_{\mathbf{s} \geqslant \mathbf{0}} \binom{s_1 + s_2}{s_1} \alpha^{s_1} \gamma^{s_2} Z_{\mathbf{n}-\mathbf{s}}.$$

Thus $X_{\mathbf{n}}$ is a linear process with

$$\psi_{\mathbf{s}} = \binom{s_1 + s_2}{s_1} \alpha^{s_1} \gamma^{s_2},$$

and Assumption 2 is clearly satisfied.

Under Assumptions 1–3, we will prove the following asymptotic normality theorem, which is the main result of this paper.

Theorem 2.1. *Let $x_1, \ldots, x_k$ be k arbitrary fixed points in $\mathbb{R}$, $k \in \mathbb{N}$. Then,*

$$(\hat{\mathbf{n}} b_{\mathbf{n}})^{1/2} (f_{\mathbf{n}}(x_1) - \mathrm{E} f_{\mathbf{n}}(x_1), \ldots, f_{\mathbf{n}}(x_k) - \mathrm{E} f_{\mathbf{n}}(x_k))^{\mathrm{T}} \xrightarrow{\mathscr{L}} \mathscr{N}(\mathbf{0}, \mathbf{C}),$$

where $\mathbf{C}$ is a diagonal matrix with diagonal elements $C_{ii} = f(x_i) \int_{-\infty}^{\infty} K^2(u)\,\mathrm{d}u$, $i = 1, \ldots, k$.

The proof of Theorem 2.1 is based on the definition of a sequence $g_{\mathbf{n}}$ which is more tractable than $f_{\mathbf{n}}$, but asymptotically equivalent (see Lemma 2.1). The asymptotic distribution of $g_{\mathbf{n}}$, and hence that of $f_{\mathbf{n}}$, is obtained through a delicate blocking technique and a series of lemmas. The blocking technique is presented in Section 3, the lemmas and the proof of the theorem in Section 4.

Choose $h_1(\mathbf{n})$ and $h_2(\mathbf{n})$ to be arbitrary positive functions with $h_1(\mathbf{n}) \uparrow \infty$ and $h_2(\mathbf{n}) \downarrow 0$ as $\mathbf{n} \to \infty$, such that

$$\gamma(\mathbf{n}) := (\hat{\mathbf{n}} b_{\mathbf{n}}^{(2a-1+6N)/(2a-1-4N)})^{(2a-1+4N)/(4Na-2N)} (h_2(\mathbf{n})/h_1(\mathbf{n})) \to \infty.$$

Since $(2a - 1 - 4N)/(4Na - 2N) > 0$ by Assumption 2, such functions exist by Assumption 3. Defining $m := [(\hat{\mathbf{n}}^2 b_{\mathbf{n}}^{-3})^{1/(2a-1)} h_1(n)]$, put

$$X_{\mathbf{i}}^m := \sum_{|s_k| \leqslant m-1} \psi_{\mathbf{s}} Z_{\mathbf{i}-\mathbf{s}} \quad \text{and} \quad \tilde{X}_{\mathbf{i}} = X_{\mathbf{i}}^m + \Gamma_{\mathbf{i}},$$

where the $\Gamma_{\mathbf{i}}$ are independent random variables, with $\{Z_{\mathbf{i}}\}$ independent of $\{\Gamma_{\mathbf{i}}\}$, and (denoting by $\sim$ equality in distribution) $\Gamma_{\mathbf{i}} \sim X_{\mathbf{i}} - X_{\mathbf{i}}^m$. Here the summation $\sum_{|s_k| \leqslant m-1}$ runs over all lattice points $\mathbf{s} = (s_1, \ldots, s_N)$, with $|s_k| \leqslant m-1$ for all $1 \leqslant k \leqslant N$. Clearly, $X_{\mathbf{i}}$ and $\tilde{X}_{\mathbf{i}}$ have the same marginal distribution, but $\tilde{X}_{\mathbf{i}}$ has the characteristic feature of finite-order moving averages, that is $\tilde{X}_{\mathbf{i}}$ and $\tilde{X}_{\mathbf{j}}$ are independent if $|i_k - j_k| \geqslant 2m$ for some $1 \leqslant k \leqslant N$.

Considering the kernel estimator $g_{\mathbf{n}}(x)$ defined by

$$g_{\mathbf{n}}(x) := (\hat{\mathbf{n}} b_{\mathbf{n}})^{-1} \sum_{\mathbf{i} \in I_{\mathbf{n}}} K((x - \tilde{X}_{\mathbf{i}})/b_{\mathbf{n}}), \tag{2.2}$$

the following lemma, proved in the Appendix, establishes the asymptotic equivalence of $f_{\mathbf{n}}$ and $g_{\mathbf{n}}$.

Lemma 2.1. *For any* $x \in \mathbb{R}$, $|f_{\mathbf{n}}(x) - g_{\mathbf{n}}(x)| = o_{\mathrm{P}}(\hat{\mathbf{n}} b_{\mathbf{n}})^{-1/2}$ *as* $\mathbf{n} \to \infty$.

In view of this lemma, we may concentrate on $g_{\mathbf{n}}$. Define the average kernel $K_{\mathbf{n}}(x)$, and the functions $\Delta_{\mathbf{i}}(x)$ and $\tilde{\Delta}_{\mathbf{i}}(x)$, respectively, by

$$K_{\mathbf{n}}(x) := \frac{1}{b_{\mathbf{n}}} K\left(\frac{x}{b_{\mathbf{n}}}\right),$$

$$\Delta_{\mathbf{i}}(x) := K_{\mathbf{n}}(x - X_{\mathbf{i}}) - \mu_{\mathbf{n}} \quad \text{and} \quad \tilde{\Delta}_{\mathbf{i}}(x) := K_{\mathbf{n}}(x - \tilde{X}_{\mathbf{i}}) - \tilde{\mu}_{\mathbf{n}}, \tag{2.3}$$

where $\mu_{\mathbf{n}} := \mathrm{E} K_{\mathbf{n}}(x - X_{\mathbf{i}})$ and $\tilde{\mu}_{\mathbf{n}} := \mathrm{E} K_{\mathbf{n}}(x - \tilde{X}_{\mathbf{i}})$. Then $f_{\mathbf{n}}(x)$ and $g_{\mathbf{n}}(x)$ can be written as

$$f_{\mathbf{n}}(x) = \frac{1}{\hat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathrm{n}}} K_{\mathbf{n}}(x - X_{\mathbf{i}}) \quad \text{and} \quad g_{\mathbf{n}}(x) = \frac{1}{\hat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathrm{n}}} K_{\mathbf{n}}(x - \tilde{X}_{\mathbf{i}}),$$

respectively. Since $X_{\mathbf{i}}$ and $\tilde{X}_{\mathbf{i}}$ have the same distribution, we have $\mu_{\mathbf{n}} = \tilde{\mu}_{\mathbf{n}}$. Observe that

$$g_{\mathbf{n}}(x) - \mathrm{E} g_{\mathbf{n}}(x) = \frac{1}{\hat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathbf{n}}} \tilde{\Delta}_{\mathbf{i}}(x). \tag{2.4}$$

Without loss of generality, we consider the case $k = 2$ and, in order to avoid subscripts, we refer to x_1 and x_2 as x and y. By the Cramér–Wold device, it suffices to prove asymptotic normality of $c\xi_{\mathbf{n}}(x) + d\xi_{\mathbf{n}}(y)$ for arbitrary constants c and d, where

$$\xi_{\mathbf{n}}(x) = (\hat{\mathbf{n}} b_{\mathbf{n}})^{1/2}(g_{\mathbf{n}}(x) - \mathrm{E} g_{\mathbf{n}}(x)) \quad \text{and} \quad \xi_{\mathbf{n}}(y) = (\hat{\mathbf{n}} b_{\mathbf{n}})^{1/2}(g_{\mathbf{n}}(y) - \mathrm{E} g_{\mathbf{n}}(y)).$$

3. The blocking technique

We first describe the blocking technique. Clearly,

$$c\xi_{\mathbf{n}}(x) + d\xi_{\mathbf{n}}(y) = \sum_{\mathbf{i}\in I_{\mathbf{n}}} \varphi_{\mathbf{i}} = S_{\mathbf{n}}, \tag{3.1}$$

say, where $\varphi_{\mathbf{i}} = \hat{\mathbf{n}}^{-1/2} b_{\mathbf{n}}^{1/2}(c\tilde{\Delta}_{\mathbf{i}}(x) + d\tilde{\Delta}_{\mathbf{i}}(y))$. Let $l = l_{\mathbf{n}} := [(\hat{\mathbf{n}} b_{\mathbf{n}})^{1/(2N)} h_2(\mathbf{n})]$. The random variables $\varphi_{\mathbf{i}}$ are now set into large and small blocks. The large blocks have all sides equal to l and the small ones have at least one side equal to $2m$. Assume without loss of generality that for some integers $r_1, \ldots, r_N$, we have $n_1 = r_1(l+2m), \ldots, n_N = r_N(l+2m)$. Denote by

$$U(1, \mathbf{n}, \mathbf{j}) := \sum_{\substack{i_k = j_k(l+2m)+1\\ k=1,\ldots,N}}^{j_k(l+2m)+l} \varphi_{\mathbf{i}}$$

a sum of random variables running over a large block with all sides equal to l. Next, define

$$U(2, \mathbf{n}, \mathbf{j}) := \sum_{\substack{i_k = j_k(l+2m)+1\\ k=1,\ldots,N-1}}^{j_k(l+2m)+l} \; \sum_{i_N = j_N(l+2m)+l+1}^{(j_N+1)(l+2m)} \varphi_{\mathbf{i}},$$

$$U(3, \mathbf{n}, \mathbf{j}) := \sum_{\substack{i_k = j_k(l+2m)+1\\ k=1,\ldots,N-2}}^{j_k(l+2m)+l} \; \sum_{i_{N-1} = j_{N-1}(l+2m)+l+1}^{(j_{N-1}+1)(l+2m)} \; \sum_{i_N = j_N(l+2m)+1}^{j_N(l+2m)+l} \varphi_{\mathbf{i}},$$

$$U(4, \mathbf{n}, \mathbf{j}) := \sum_{\substack{i_k = j_k(l+2m)+1\\ k=1,\ldots,N-2}}^{j_k(l+2m)+l} \; \sum_{i_{N-1} = j_{N-1}(l+2m)+l+1}^{(j_{N-1}+1)(l+2m)} \; \sum_{i_N = j_N(l+2m)+l+1}^{(j_N+1)(l+2m)} \varphi_{\mathbf{i}},$$

and so on. Observe that $U(2, \mathbf{n}, \mathbf{j})$, $U(3, \mathbf{n}, \mathbf{j})$, $U(4, \mathbf{n}, \mathbf{j})$ contain random variables in blocks with one side equal to $2m$ and $N-1$ sides equal to l, two sides equal to $2m$ and $N-2$ equal to l, three sides equal to $2m$ and $N-3$ sides equal to l, respectively. More generally,

$$U(2^{N-1}, \mathbf{n}, \mathbf{j}) := \sum_{\substack{i_k = j_k(l+2m)+l+1\\ k=1,\ldots,N-1}}^{(j_k+1)(l+2m)} \; \sum_{i_N = j_N(l+2m)+1}^{j_N(l+2m)+l} \varphi_{\mathbf{i}}$$

contains random variables in a small block with one side equal to l and $N-1$ sides equal to $2m$. Finally,

$$U(2^N, \mathbf{n}, \mathbf{j}) := \sum_{\substack{i_k = j_k(l+2m)+l+1\\ k=1,\ldots,N}}^{(j_k+1)(l+2m)} \varphi_{\mathbf{i}}$$

contains random variables in a small block with all sides equal to $2m$.

For each integer $1 \leqslant i \leqslant 2^N$, define $T(\mathbf{n}, i) := \sum_{0\leqslant j_k \leqslant r_k - 1} U(i, \mathbf{n}, \mathbf{j})$. Then the sum of random variables $S_{\mathbf{n}}$ defined in (3.1) equals $\sum_{i=1}^{2^N} T(\mathbf{n}, i)$. Note that $T(\mathbf{n}, 1)$ is the sum of the random variables $\varphi_{\mathbf{i}}$ in large blocks with all sides equal to l. The statistics $T(\mathbf{n}, i)$, $2 \leqslant i \leqslant 2^N$, are sums of random variables in small blocks. If it is not the case that $n_1 = r_1(l+2m), \ldots, n_N = r_N(l+2m)$ for some integers $r_1, \ldots, r_N$, then a term,

$T(\mathbf{n}, 2^N + 1)$, say, containing all the $\varphi_{\mathbf{i}}$ at the ends which are not included in the large or in the small blocks, can be added. This term will not change the proof much.

4. Proof of Theorem 2.1

Clearly, $S_{\mathbf{n}} = \sum_{\mathbf{i}\in I_{\mathbf{n}}} \varphi_{\mathbf{i}} = T(\mathbf{n}, 1) + \sum_{i=2}^{2^N} T(\mathbf{n}, i)$. Theorem 2.1 will thus follow if we can show that

$$T(\mathbf{n}, 1) \xrightarrow{\mathscr{L}} \mathscr{N}(0, \tau^2) \quad \text{and} \quad \sum_{i=2}^{2^N} T(\mathbf{n}, i) = o_P(1). \tag{4.1}$$

The proof of these two statements relies on a series of lemmas.

Lemma 4.1. *Let*

$$q_{1\mathbf{n}} := \mathrm{E}[\{K_{\mathbf{n}}(x - \tilde{X}_{\mathbf{i}}) - \mu_{\mathbf{n}}\}\{K_{\mathbf{n}}(y - \tilde{X}_{\mathbf{j}}) - K_{\mathbf{n}}(y - (\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})))\}],$$

with

$$\mathscr{S}(\mathbf{i}, \mathbf{j}) := \{s : j_k - m + 1 \leqslant s_k \leqslant j_k + m - 1\} \cap \{\mathbf{s} : i_k - m + 1 \leqslant s_k \leqslant i_k + m - 1\}$$

and

$$R_{\mathbf{j}}(\mathbf{i}) := \sum_{\mathscr{S}(\mathbf{i},\mathbf{j})} \psi_{\mathbf{j}-\mathbf{s}} Z_{\mathbf{s}}. \tag{4.2}$$

Then, for all sites $\mathbf{i}$, $\mathbf{j}$ *and all* x *and* y *in* $\mathbb{R}$, $\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\} = q_{1\mathbf{n}}$.

For the proof of this lemma, see the Appendix.

Lemma 4.2. *For all sites* $\mathbf{i}$, $\mathbf{j}$ *and all* x *and* y *in* $\mathbb{R}$,

$$|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}| \leqslant C b_{\mathbf{n}}^{-3} \sum_{\|\mathbf{t}\| \geqslant \|\mathbf{i}-\mathbf{j}\|/\sqrt{N}} |\psi_{\mathbf{t}}|. \tag{4.3}$$

Proof. Since K is bounded, $|K_{\mathbf{n}}(x - \tilde{X}_{\mathbf{i}}) - \mu_{\mathbf{n}}| \leqslant C b_{\mathbf{n}}^{-1}$. Hence, in view of Lemma 4.1,

$$|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}| \leqslant C b_{\mathbf{n}}^{-2} \mathrm{E}\left| K\left(\frac{y - \tilde{X}_{\mathbf{j}}}{b_{\mathbf{n}}}\right) - K\left(\frac{y - \tilde{X}_{\mathbf{j}} + R_{\mathbf{j}}(\mathrm{i})}{b_{\mathbf{n}}}\right)\right|. \tag{4.4}$$

By the Lipschitz property of K, the last term of (4.4) is bounded by $C b_{\mathbf{n}}^{-3} \mathrm{E}|R_{\mathbf{j}}(\mathbf{i})|$. Using (4.2) and the fact that $|\mathrm{E} Z_{\mathbf{s}}|$ is finite, we finally obtain

$$|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}| \geqslant C b_{\mathbf{n}}^{-3} \sum_{\mathbf{s}\in\mathscr{S}(\mathbf{i},\mathbf{j})} |\psi_{\mathbf{i}-\mathbf{s}}|. \tag{4.5}$$

Since $|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}|$ and $|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{j}}(x), \tilde{\Delta}_{\mathbf{i}}(y)\}|$ are equal, (4.5) remains valid if $R_{\mathbf{j}}(\mathbf{i})$ is replaced by $R_{\mathbf{i}}(\mathbf{j})$. Therefore,

$$|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}| \leqslant Cb_{\mathbf{n}}^{-3} \min\left\{\sum_{\mathbf{s}\in\mathscr{S}(\mathbf{i},\mathbf{j})} |\psi_{\mathbf{i}-\mathbf{s}}| \sum_{\mathbf{s}\in\mathscr{S}(\mathbf{j},\mathbf{i})} |\psi_{\mathbf{j}-\mathbf{s}}|\right\}.$$

The lemma then follows directly from the geometry of N-dimensional Euclidean spaces. □

Lemma 4.3. *For any* $\eta \geqslant 0$, $\sum_{\|\mathbf{t}\|\geqslant\eta} |\psi_{\mathbf{t}}| \leqslant C\eta^{N-a}$.

Proof. In view of Assumption 2, we have

$$\sum_{\|\mathbf{t}\|\geqslant\eta} |\psi_{\mathbf{t}}| \leqslant \sum_{\|\mathbf{t}\|\leqslant\eta} C\|\mathbf{t}\|^{-a} \leqslant C \sum_{\eta<i} i^{N-1} i^{-a} \leqslant C\eta^{N-a}.$$

□

Lemma 4.4. *If X is a random variable with bounded density, then, for some constant C independent of w,* $|\mathrm{E}\{K_{\mathbf{n}}(w-X)\}| \leqslant C$.

Proof. The proof follows from Assumption 1 and the boundedness of the density of X. □

Lemma 4.5. *For all sites* $\mathbf{i}$, $\mathbf{j}$ *and all x and y in* $\mathbb{R}$, $|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}| \leqslant Cb_{\mathbf{n}}^{-3}\|\mathbf{i}-\mathbf{j}\|^{N-a}$.

Proof. Lemma 4.5 is a direct consequence of Lemmas 4.2 and 4.3. □

Lemma 4.6. *For all sites* $\mathbf{i}$ *and* $\mathbf{j}$, $\sup_{(x,y)\in\mathbb{R}^2} |\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}| \leqslant C$.

Proof. Note that $y - \tilde{X}_{\mathbf{j}} = y - R_{\mathbf{j}}(\mathbf{i}) - (\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i}))$. Hence,

$$|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}| = |\mathrm{E}[\{K_{\mathbf{n}}(x-\tilde{X}_{\mathbf{i}}) - \mu_{\mathbf{n}}\}\{K_{\mathbf{n}}(y - R_{\mathbf{j}}(\mathbf{i}) - (\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})))\}]|.$$

By Assumption 2, $\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})$ has a bounded density. Since $\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})$ is independent of $(\tilde{X}_{\mathbf{i}}, R_{\mathbf{j}}(\mathbf{i}))$, by Lemma 4.4, we obtain

$$\begin{aligned}
&|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\}| \\
&= \left|\int_{-\infty}^{\infty} \mathrm{E}\left[\{K_{\mathbf{n}}(x-\tilde{X}_{\mathbf{i}}) - \mu_{\mathbf{n}}\}\{K_{\mathbf{n}}(y - r - (\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})))\}]|R_{\mathbf{j}}(\mathbf{i}) = r\right] f_{R_{\mathbf{j}}(\mathbf{i})}(r)\mathrm{d}r\right| \\
&\leqslant C\mathrm{E}|[\{K_{\mathbf{n}}(x-\tilde{X}_{\mathbf{i}}) - \mu_{\mathbf{n}}\}]| \leqslant C.
\end{aligned}$$

Lemma 4.7. *If Assumption* 1 *holds, then*

$$\int_{-\infty}^{\infty} K_{\mathbf{n}}(x-u)f(u)\mathrm{d}u \to f(x) \tag{4.6}$$

and

$$\int_{-\infty}^{\infty} [K((x-u)/b_{\mathbf{n}})]^2 f(u)\mathrm{d}u \longrightarrow f(x)\int_{-\infty}^{\infty} [K(u)]^2\,\mathrm{d}u. \tag{4.7}$$

Proof. Relations (4.6) and (4.7) both follow from the Lebesgue density theorem (Devroye and Györfi, 1985, p. 7) by noting that $\int_{-\infty}^{\infty}[K(u)]^2 f(u)\mathrm{d}u < \infty$. □

Lemma 4.8. *Let* $g_{\mathbf{n}}$ *be the kernel density estimator defined in* (2.2). *Then,*

$$\lim_{\mathrm{n}\to\infty} \hat{\mathbf{n}} b_{\mathbf{n}} \mathrm{var}[g_{\mathbf{n}}(x)] = f(x)\int_{-\infty}^{\infty} K^2(y)\mathrm{d}y.$$

The proof of this lemma is given in the Appendix.

Lemma 4.9. *Let* $x \neq y \in \mathbb{R}$. *Then* $\lim_{\mathbf{n}\to\infty} \hat{\mathbf{n}} b_{\mathbf{n}} \mathrm{cov}\{g_{\mathbf{n}}(x), g_{\mathbf{n}}(y)\} = 0$.

The proof is similar to that of Lemma 4.8, and is omitted.

Lemma 4.10 *Under Assumption* 2, *each* $T(\mathbf{n}, i)$, $2 \leqslant i \leqslant 2N$, *tends to zero in probability.*

This final lemma is proved in the Appendix.

We now may proceed with the proof of Theorem 2.1. Without loss of generality, assume $k = 2$. Define $s_n^2 := \mathrm{var}[T(\mathbf{n}, 1)]$. By Lemma 4.10, $\mathrm{E}[T(\mathbf{n}, i)]^2 \to 0$ for each $2 \leqslant i \leqslant 2N$. In addition, Lemmas 4.8 and 4.9 imply that $\mathrm{E}S_{\mathbf{n}}^2 \to \tau^2$, where $S_{\mathbf{n}}$ is defined in (3.1). Following the proof of Lemma 3.2 in Hallin and Tran (1996), we have $s_{\mathbf{n}} \to \tau^2$. To complete the proof of the theorem, it is thus sufficient to show that

$$\frac{T(\mathbf{n}, 1)}{s_{\mathbf{n}}} \xrightarrow{\mathscr{L}} \mathscr{N}(0, 1). \tag{4.8}$$

Recall that $T(\mathbf{n}, 1)$ is the sum of $r_1 \times \cdots \times r_N$ independent random variables $U(1, \mathbf{n}, \mathbf{j})$. By the Lindeberg central limit theorem, (4.8) follows if, for every $\epsilon > 0$,

$$\sum_{0 \leqslant j_k \leqslant r_k - 1} \int_{(|x| \geqslant \epsilon)} x^2\,\mathrm{d}F_{(1,\mathbf{n},\mathbf{j})} \to 0, \tag{4.9}$$

where $F_{(1,\mathbf{n},\mathbf{j})}$ is the distribution function of $U(1, \mathbf{n}, \mathbf{j})$. A simple computation gives $|U(1, \mathbf{n}, \mathbf{j})| \leqslant Cl^N(\hat{\mathbf{n}} b_{\mathbf{n}})^{-1/2}$, which tends to zero by the definition of l. Thus the left-hand side of (4.9) is zero for $\hat{\mathbf{n}}$ sufficiently large.

Appendix

***Proof of Lemma* 2.1.** The proof of this lemma follows an argument similar to that of Lemma 2.9 in Hallin and Tran (1996). From the Lipschitz property of K in Assumption 1, we have

$$(\hat{\mathbf{n}}b_{\mathbf{n}})^{1/2}|f_{\mathbf{n}}(x) - g_{\mathbf{n}}(x)| \leqslant (\hat{\mathbf{n}}b_{\mathbf{n}})^{1/2}(\hat{\mathbf{n}}b_{\mathbf{n}})^{-1}\sum_{\mathbf{i}\in I_{\mathbf{n}}}\left|K\left(\frac{x-X_{\mathbf{i}}}{b_{\mathbf{n}}}\right) - K\left(\frac{x-\tilde{X}_{\mathbf{i}}}{b_{\mathbf{n}}}\right)\right|$$

$$\leqslant C\hat{\mathbf{n}}^{-1/2}b_{\mathbf{n}}^{-3/2}\sum_{\mathbf{i}\in I_{\mathbf{n}}}|X_{\mathbf{i}} - \tilde{X}_{\mathbf{i}}|.$$

Therefore, given any $\epsilon > 0$,

$$P[(\hat{\mathbf{n}}b_{\mathbf{n}})^{1/2}|f_{\mathbf{n}}(x) - g_{\mathbf{n}}(x)| > \epsilon] \leqslant P\left[\sum_{\mathbf{i}\in I_{\mathbf{n}}}|X_{\mathbf{i}} - \tilde{X}_{\mathbf{i}}| > C^{-1}\epsilon\hat{\mathbf{n}}^{1/2}b_{\mathbf{n}}^{3/2}\right]. \tag{A.1}$$

Since $I_{\mathbf{n}}$ contains $\hat{\mathbf{n}}$ elements, the event within brackets on the left-hand side of (A.1) occurs only if $|X_{\mathbf{i}} - \tilde{X}_{\mathbf{i}}| > C^{-1}\epsilon\hat{\mathbf{n}}^{1/2}b_{\mathbf{n}}^{3/2}/\hat{\mathbf{n}}$ for some $\mathbf{i} \in I_{\mathbf{n}}$. Thus, the right-hand side of (A.1) is bounded by $\sum_{\mathbf{i}\in I_{\mathbf{n}}} P[|X_{\mathbf{i}} - \tilde{X}_{\mathbf{i}}| > C^{-1}\epsilon\hat{\mathbf{n}}^{-1/2}b_{\mathbf{n}}^{3/2}]$ which in turn, by Chebyshev's inequality, is less than

$$C^2\epsilon^{-2}\hat{\mathbf{n}}^2 b_{\mathbf{n}}^{-3}\,\mathrm{var}(X_{\mathbf{i}} - \tilde{X}_{\mathbf{i}}). \tag{A.2}$$

Note that $\mathrm{var}(X_{\mathbf{i}} - \tilde{X}_{\mathbf{i}}) = \mathrm{var}(X_{\mathbf{i}} - X_{\mathbf{i}}^m - \Gamma_{\mathbf{i}})$, with

$$X_{\mathbf{i}} - X_{\mathbf{i}}^m = \sum_{\mathbf{s}\in\mathbb{Z}^N}\psi_{\mathbf{s}}Z_{\mathbf{i}-\mathbf{s}} - \sum_{|s_k|\leqslant m-1}\psi_{\mathbf{s}}Z_{\mathbf{i}-\mathbf{s}}.$$

Thus $X_{\mathbf{i}} - X_{\mathbf{i}}^m$ is the sum of random variables $\psi_{\mathbf{s}}Z_{\mathbf{i}-\mathbf{s}}$ over all lattice points $\mathbf{s}$ with $|s_k| \geqslant m$ for some $1 \leqslant k \leqslant N$. Using Assumption 2 and the fact that $X_{\mathbf{i}} - X_{\mathbf{i}}^m$ and $\Gamma_{\mathbf{i}}$ are i.i.d.,

$$\mathrm{var}(X_{\mathbf{i}} - X_{\mathbf{i}}^m - \Gamma_{\mathbf{i}}) \leqslant C\,\mathrm{E}Z_1^2\left(\sum_{\mathbf{s}\in\mathbb{Z}^N}\psi_{\mathbf{s}}^2 - \sum_{|s_k|\leqslant m-1}\psi_{\mathbf{s}}^2\right) \leqslant Cm^{-2a+1}.$$

A simple computation now shows that (A.2) is $O(\hat{\mathbf{n}}^2 b_{\mathbf{n}}^{-3}m^{-2a+1})$, hence $O((h_1(n))^{-2a+1})$, hence $o(1)$, which completes the proof. □

***Proof of Lemma* 4.1.** Note that $\tilde{X}_{\mathbf{i}} = \sum_{i_k-2m+1\leqslant s_k\leqslant i_k}\psi_{\mathbf{i}-\mathbf{s}}Z_{\mathbf{s}} + \Gamma_{\mathbf{i}}$, and that $\tilde{X}_{\mathbf{j}} = R_{\mathbf{j}}(\mathbf{i}) + (\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i}))$. Since $\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})$ is measurable with respect to the sigma-field generated by $\Gamma_{\mathbf{j}}$ and the random variables $Z_{\mathbf{j}}$ with $\mathbf{j}$ outside the set of sites $\{\mathbf{s} : i_k - m + 1 \leqslant s_k \leqslant i_k + m - 1\}$, the random variables $\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})$ and $\tilde{X}_{\mathbf{i}}$ are independent. Clearly,

$$\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(y)\} = q_{1\mathbf{n}} + \mathrm{E}[\{K_{\mathbf{n}}(x - \tilde{X}_{\mathbf{i}}) - \mu_{\mathbf{n}}\}K_{\mathbf{n}}(y - (\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})))],$$

where the last term equals zero by the independence of $\tilde{X}_{\mathbf{i}}$ and $\tilde{X}_{\mathbf{j}} - R_{\mathbf{j}}(\mathbf{i})$. □

***Proof of Lemma* 4.8.** Using (4.6) and (4.7), and noting that $b_{\mathbf{n}} \to 0$,

$$b_{\mathbf{n}}\,\mathrm{var}[\tilde{\Delta}_{\mathbf{i}}(x)] = \int_{-\infty}^{\infty} \frac{1}{b_{\mathbf{n}}}\left[K\left(\frac{x-y}{b_{\mathbf{n}}}\right)\right]^2 f(y)\mathrm{d}y - b_{\mathbf{n}}\left[\int_{-\infty}^{\infty}\frac{1}{b_{\mathbf{n}}}K\left(\frac{x-y}{b_{\mathbf{n}}}\right)f(y)\mathrm{d}y\right]^2$$

$$\to f(x)\int_{-\infty}^{\infty} K^2(y)\mathrm{d}y, \tag{A.3}$$

where $\tilde{\Delta}_{\mathbf{i}}(x)$ is defined in (2.3). Put

$$S = (\mathbf{i}, \mathbf{j}) \in I_{\mathbf{n}} \times I_{\mathbf{n}} : i_k \neq j_k \text{ for some } 1 \leqslant k \leqslant N\}. \tag{A.4}$$

Using (2.4),

$$\hat{\mathbf{n}} b_{\mathbf{n}}\,\mathrm{var}[g_{\mathbf{n}}(x)] = b_{\mathbf{n}}\,\mathrm{var}[\tilde{\Delta}_{\mathbf{1}}(x)] + (b_{\mathbf{n}}/\hat{\mathbf{n}})\sum_{S}|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(x)\}|. \tag{A.5}$$

Choosing a number θ that satisfies $0 < \theta < (a-N-3)/(a-N-1)$, set $\rho = b_{\mathbf{n}}^{(\theta-1)/N}$. Define S_1 to be the set containing all pairs $(\mathbf{i}, \mathbf{j})$ with $|i_k - j_k|$ no greater than ρ for all $1 \leqslant k \leqslant N$, that is, $S_1 = \{(\mathbf{i}, \mathbf{j}) \in S : |i_k - j_k| \leqslant \rho$ for all $1 \leqslant k \leqslant N\}$. Let $S_2 = \{(\mathbf{i}, \mathbf{j}) \in S \cap S_1^C\}$, where S_1^C denotes the complement of S_1. By Lemmas 4.6 and 4.5,

$$\sum_{S_1}|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(x)\}| \leqslant C\hat{\mathbf{n}}\rho^N \quad \text{and} \quad \sum_{S_2}|\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(x)\}| \leqslant C\hat{\mathbf{n}} b_{\mathbf{n}}^{-3}\rho^{N(N-a+1)}.$$

Lemmas 4.2–4.5 and Assumption 2 imply that the last term on the right-hand side of (A.5) is bounded by

$$C(b_{\mathbf{n}}/\hat{\mathbf{n}})(\hat{\mathbf{n}}\rho^N + \hat{\mathbf{n}} b_{\mathbf{n}}^{-3}\rho^{N(N-a+1)}) \leqslant C(b_{\mathbf{n}}\rho^N + b_{\mathbf{n}}^{-2}\rho^{N(N-a+1))}) \leqslant C(b_{\mathbf{n}}^{\theta} b_{\mathbf{n}}^{-2+(1-\theta)(a-N-1)}).$$

Since $a - N - 1 > 2$ by Assumption 2, we can choose θ sufficiently close to zero so that $(1-\theta)(a-N-1) > 2$. Thus, the last term of (A.6) tends to zero. □

***Proof of Lemma* 4.10.** The lemma will follow if we show that $\mathrm{E}[T^2(\mathbf{n}, i)] \to 0$ for each $2 \leqslant i \leqslant 2N$. Without loss of generality, we will restrict ourselves to showing that $\mathrm{E}[T^2(\mathbf{n}, 2)] \to 0$. Note that

$$T(\mathbf{n}, 2) = \sum_{0 \leqslant j_k \leqslant r_k - 1} U(2, \mathbf{n}, \mathbf{j}) = c\hat{\mathbf{n}}^{-1/2}\sum_{0 \leqslant j_k \leqslant r_k - 1} V(2, \mathbf{n}, \mathbf{j}) + d\hat{\mathbf{n}}^{-1/2}\sum_{0 \leqslant j_k \leqslant r_k - 1} W(2, \mathbf{n}, \mathbf{j}),$$

where

$$V(2, \mathbf{n}, \mathbf{j}) := \sum_{\substack{i_k = j_k(l+2m)+1 \\ k=1,\ldots,N-1}}^{j_k(l+2m)+l}\ \sum_{i_N = j_N(l+2m)+l+1}^{(j_N+1)(l+2m)} b_{\mathbf{n}}^{1/2}\tilde{\Delta}_{\mathbf{i}}(x)$$

and

$$W(2, \mathbf{n}, \mathbf{j}) := \sum_{\substack{i_k = j_k(l+2m)+1 \\ k=1,\ldots,N-1}}^{j_k(l+2m)+l}\ \sum_{i_N = j_N(l+2m)+l+1}^{(j_N+1)(l+2m)} b_{\mathbf{n}}^{1/2}\tilde{\Delta}_{\mathbf{i}}(y).$$

By Minkowski's inequality,

$$(\mathrm{E}[T^2(\mathbf{n}, 2)])^{1/2} \leqslant |c| \left[\frac{1}{\hat{\mathbf{n}}} \mathrm{E} \left(\sum_{0 \leqslant j_k \leqslant r_k - 1} V(2, \mathbf{n}, \mathbf{j}) \right)^2 \right]^{1/2} + |d| \left[\frac{1}{\hat{\mathbf{n}}} \mathrm{E} \left(\sum_{0 \leqslant j_k \leqslant r_k - 1} W(2, \mathbf{n}, \mathbf{j}) \right)^2 \right]^{1/2}. \tag{A.6}$$

It is sufficient to show that the two terms on the right-hand side of (A.6) tend to zero. Since the $V(2, \mathbf{n}, \mathbf{j})$ are independent random variables with zero means,

$$\begin{aligned} \frac{1}{\hat{\mathbf{n}}} \mathrm{E} \left(\sum_{0 \leqslant j_k \leqslant r_k - 1} V(2, \mathbf{n}, \mathbf{j}) \right)^2 &= \frac{1}{\hat{\mathbf{n}}} \sum_{0 \leqslant j_k \leqslant r_k - 1} \mathrm{var}[V(2, \mathbf{n}, \mathbf{j})] \\ &\leqslant \frac{1}{\hat{\mathbf{n}}} \sum_{0 \leqslant j_k \leqslant r_k - 1} \sum_{\substack{i_k = j_k(l+2m)+1 \\ k=1,\ldots,N-1}}^{j_k(l+2m)+l} \sum_{i_N = j_N(l+2m)+l+1}^{(j_N+1)(l+2m)} \mathrm{var}[b_{\mathbf{n}}^{1/2} \tilde{\Delta}_{\mathbf{i}}(x)] \qquad \text{(A.7)} \\ &\quad + \frac{b_{\mathbf{n}}}{\hat{\mathbf{n}}} \sum_S |\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(x)\}|. \end{aligned}$$

Recall from Lemma 4.8 that $(b_{\mathbf{n}}/\hat{\mathbf{n}}) \sum_S |\mathrm{cov}\{\tilde{\Delta}_{\mathbf{i}}(x), \tilde{\Delta}_{\mathbf{j}}(x)\}| \to 0$, where S is defined in (A.4). By (A.3) and (A.7),

$$\begin{aligned} \frac{1}{\hat{\mathbf{n}}} \mathrm{E} \left(\sum_{0 \leqslant j_k \leqslant r_k - 1} V(2, \mathbf{n}, \mathbf{j}) \right)^2 &\leqslant C \hat{\mathbf{n}}^{-1} r_1 \cdots r_N l^{N-1} m + o(1) \\ &= C \frac{r_1 l}{n_1} \times \cdots \times \frac{r_{N-1} l}{n_{N-1}} \times \frac{m}{l + 2m} + o(1) \leqslant C \frac{m}{l + 2m} + o(1), \end{aligned}$$

which tends to zero since m/l tends to zero. The proof for the second term on the right-hand side of (A.7) is entirely similar.

Acknowledgements

The first and third named authors were supported by an Action de Recherche Concertée of the Communauté Française de Belgique, and the third named author was partially supported by NSF grant DMS-9403718. The second named author was partially supported by the Natural Science Foundation of China. We would like to thank the editors and referees for their useful comments.

References

Anselin, L. and Florax, R.J.G.M. (1995) *New Directions in Spatial Econometrics*. Berlin: Springer-Verlag.

Basawa, I.V. (1996a) Special issue on spatial statistics. Part I. *J. Statist. Plann. Inference*, **50**, 311–411.

Basawa, I.V. (1996b) Special issue on spatial statistics. Part II. *J. Statist. Plann. Inference*, **51**, 1–97.

Cressie, N.A.C. (1991) *Statistics for Spatial Data*. New York: Wiley.

Dedecker, J. (1998) A central limit theorem for stationary random fields. *Probab. Theory Related Fields*, **110**, 397–426.

Devroye, L. and Györfi, L. (1985) *Nonparametric Density Estimation. The L_1 View.* New York: Wiley.

Guyon, X. (1995) *Random Fields on a Network: Modelling, Statistics, and Applications*. New York: Springer-Verlag.

Hallin, M. and Tran, L.T. (1996) Density estimation for linear time series: asymptotic properties and optimal bandwidth selection. *Ann. Inst. Statist. Math.*, **48**, 429–449.

Kulkarni, P.M. (1992) Estimation of parameters of a two-dimensional spatial autoregressive model with regression. *Statist. Probab. Lett.*, **15**, 157–162.

Possolo, A. (1991) *Spatial Statistics and Imaging*, IMS Lecture Notes Monogr. Ser. Hayward, CA: Institute of Mathematical Statistics.

Ripley, B. (1981) *Spatial Statistics*. New York: Wiley.

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

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

Tjøstheim, D. (1987) Spatial series and time series: similarities and differences. In F. Droesbeke (ed.), *Spatial Processes and Spatial Time Series Analysis, Proceedings of the 6th Franco-Belgian Meeting of Statisticians*, pp. 217–228. Brussels: Publications des Facultés Universitaires Saint-Louis.

Tran, L.T. (1990) Kernel density estimation on random fields. *J. Multivariate Anal.*, **34**, 37–53.

Tran, L.T. and Yakowitz, S. (1993) Density estimation on random fields. *J. Multivariate Anal.*, **44**, 23–46.

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

Whittle, P. (1963) Stochastic processes in several dimensions. In *Bulletin of the International Statistical Institute: Proceedings of the 34th Session (Ottawa)*, pp. 974–994. The Hague: ISI.

Received November 1999 and revised January 2001