EXPONENTIAL INEQUALITIES FOR SPATIAL PROCESSES AND UNIFORM CONVERGENCE RATES FOR DENSITY ESTIMATION

QIWEI YAO
Department of Statistics
London School of Economics and Political Science
Houghton Street, London, WC2A 2AE, UK
E-mail: q.yao@lse.ac.uk

We establish an exponential type of inequality for α-mixing spatial processes. Based on it, an optimum convergence rate of a kernel density estimator for stationary spatial processes is obtained. Its asymptotic mean and variance are also derived.

1 Introduction

The classical asymptotic theory in statistics is built upon central limit theorems and laws of large numbers for the sequences of independent random variables. In the study of the asymptotic properties for linear time series which are sequences of dependent random variables, the conventional approach is to express a time series in terms of an moving average form in which the white noise is assumed to be independent. This idea also prevails for linear spatial processes; see Yao and Brockwell (2001) and Hallin, Lu and Tran (2002). Unfortunately the moving average representation becomes irrelevant in the context of nonlinear processes for which more complicated dependence structure is encountered and certain asymptotic independence, typically represented by a mixing condition, will be characterised. We refer to §2.6 of Fan and Yao (2003) for an introductory survey on the mixing conditions for time series. A mixing spatial process may be viewed as a process for which random variables from distant locations are *nearly* independent; see §1.3.1 of Doukhan (1994).

The goal of this paper is two-fold. First we establish an exponential type inequality for α-mixing spatial processes. Exponential type inequalities are powerful tools in estimating tail probabilities of partial sums of random sequences; see, for example, §1.4 of Bosq (1998). They imply weak laws of large numbers for mixing processes, and give sharp large deviation estimations. Note that the stationarity is not required. Further, based on an established exponential inequality, we derive an optimum uniform convergence rate for density estimators of stationary spatial processes. The application of those results in additive modelling for spatial processes will be reported elsewhere.

Nonparametric kernel estimation for spatial processes is still in its infant stage. A limited references include Diggle (1985), Diggle and Marron (1988), Hallin, Lu and Tran (2002), and Zhang, Yao, Tong and Stenseth (2002). Related to current work, Hallin *et al.* established the asymptotic normality for density estimators for linear spatial processes.

2 Exponential type inequalities

2.1 α-mixing processes

Let $\{X(\mathbf{s})\}$ be a real-valued spatial process indexed by $\mathbf{s} \equiv (u, v) \in Z^2$, where Z consists of all integers. Since the index $\mathbf{s}$ is two dimensional, $\{X(\mathbf{s})\}$ is also called a random field. For any $A \subset Z^2$, let $\mathcal{F}(A)$ denote the σ-algebra generated by $\{X(\mathbf{s}),\ \mathbf{s} \in A\}$. We write $|A|$ for the number of elements in A. For any $A, B \subset Z^2$, define

$$\alpha(A, B) = \sup_{U \in \mathcal{F}(A), V \in \mathcal{F}(B)} |P(UV) - P(U)P(V)|,$$

and

$$d(A, B) = \min\{||\mathbf{s}_1 - \mathbf{s}_2|| \mid \mathbf{s}_1 \in A,\ \mathbf{s}_2 \in B\}.$$

where $||\cdot||$ denotes the Euclidean norm. Now the α-mixing coefficient for the process $\{X(\mathbf{s})\}$ is defined as

$$\alpha(k; i, j) = \sup_{A, B \subset Z^2} \{\alpha(A, B) \mid |A| \le i,\ |B| \le j,\ d(A, B) \ge k\}, \tag{1}$$

where i, j, k are positive integers and i, j may take values of infinite. The process $\{X(\mathbf{s})\}$ is α-mixing (or strong mixing) if $\alpha(k; i, j) \to 0$ as $k \to \infty$. A typical choice is $i = j = \infty$. Note that $\alpha(k; i, j)$ is monotonically non-decreasing as a function of i or j. Hence $\alpha(k; \infty, \infty) \to 0$ implies $\alpha(k; i, j) \to 0$ for any values of i and j.

In the context of stochastic processes with single index, α-mixing is the weakest among the most frequently used mixing conditions (§2.6 of Fan and Yao 2003). Such a simple statement is no longer pertinent for spatial processes. For example if we let $i = j = \infty$, α-mixing is equivalent to ρ-mixing, and further, β-mixing reduces to simple m-dependence; see §1.3.1 of Doukhan (1994) and references within. Further in contrast to single-indexed processes, we need to choose appropriate values for i and j according to the nature of a spatial problem in hand. For example, the coefficient $\alpha(k; \infty, \infty)$ is not useful for Gibbs fields (Dobrushin 1968).

Remark 1. If $\{X(\mathbf{s})\}$ is a causal and invertible ARMA process (Whittle 1954) defined in terms an independent and identically distributed white noise sequence $\{\varepsilon(\mathbf{s})\}$, $E|\varepsilon(\mathbf{s})|^\delta < \infty$ for some $\delta > 2$, and density function p_ε of $\varepsilon(\mathbf{s})$ satisfies the condition that

$$\int |p_\varepsilon(x+z) - p_\varepsilon(x)|dx \le C|z|, \quad z \in R,$$

it follows from Corollary 1.7.3 of Guyon (1995) and Lemma 1 of Yao and Brockwell (2001) that

$$\alpha(k; i, j) \le Ci\rho^k, \quad i, j, k \ge 1, \tag{2}$$

where $C > 0, \rho \in (0,1)$ are constants independent of i, j and k.

2.2 Exponential inequalities

Suppose we have observations $\{X(u,v);\, u = 1, \cdots, N_1,\, v = 1, \cdots, N_2\}$. Let $N = N_1N_2$. Define the partial sum

$$S_N = \sum_{v=1}^{N_2}\sum_{u=1}^{N_1} X(u,v).$$

Theorem 1 below presents upper bounds for the tail probabilities of $|S_N|$, which resembles the exponential type inequalities for single-indexed stochastic processes; see Theorem 1.3 of Bosq (1998). We introduce some notation first. Define an auxiliary continuous-indexed process

$$\xi(t_1, t_2) = X([t_1], [t_2]), \quad (t_1, t_2) \in R^2,$$

where $[t]$ denotes the integer part of t. For an integer q between 1 and $N_1 \wedge N_1 \equiv \min\{N_1, N_2\}$, let $p_i = N_i/(2q)$ $(i = 1, 2)$. For $i, j = 1, \cdots, q$, define

$$V_{ij}^{(1)} = \int_{2(i-1)p_1}^{(2i-1)p_1} dt_1 \int_{2(j-1)p_2}^{(2j-1)p_2} \xi(t_1,t_2)dt_2,$$
$$V_{ij}^{(2)} = \int_{2(i-1)p_1}^{(2i-1)p_1} dt_1 \int_{(2j-1)p_2}^{2jp_2} \xi(t_1,t_2)dt_2,$$
$$V_{ij}^{(3)} = \int_{(2i-1)p_1}^{2ip_1} dt_1 \int_{2(j-1)p_2}^{(2j-1)p_2} \xi(t_1,t_2)dt_2,$$
$$V_{ij}^{(4)} = \int_{(2i-1)p_1}^{2ip_1} dt_1 \int_{(2j-1)p_2}^{2jp_2} \xi(t_1,t_2)dt_2.$$

It is easy to see that

$$S_N = \sum_{i,j=1}^{q} (V_{ij}^{(1)} + V_{ij}^{(2)} + V_{ij}^{(3)} + V_{ij}^{(4)}). \tag{3}$$

For some constant $\varepsilon, K > 0$, let

$$\nu(q)^2 = \frac{32q^4}{N^2} \max_{\substack{1\le i,j\le q \\ 1\le l\le 4}} E\left(V_{ij}^{(l)}\right)^2 + \frac{K\varepsilon}{2}.$$

Now we are ready to present the theorem.

Theorem 1. Let $\{X(\mathbf{s})\}$ be a zero-mean spatial process with

$$P\left\{\sup_{\mathbf{s}\in Z^2} |X(\mathbf{s})| < K\right\} = 1.$$

Then for any integer q between 1 and $N_1 \wedge N_1$ and $\varepsilon > 0$, it holds that

$$P(|S_N| > N\varepsilon) \;\le\; 8\exp\left(-\frac{\varepsilon^2 q^2}{8K^2}\right) \tag{4}$$
$$+\, 44\left(1+\frac{4K}{\varepsilon}\right)^{1/2} q^2\,\alpha\left(\left[\frac{N_1}{2q}\right]\bigwedge\left[\frac{N_2}{2q}\right];\, \left[\frac{N}{4q^2}\right], N\right),$$

and

$$P(|S_N| > N\varepsilon) \;\le\; 8\exp\left(-\frac{\varepsilon^2 q^2}{8\nu(q)^2}\right) \tag{5}$$
$$+\, 44\left(1+\frac{4K}{\varepsilon}\right)^{1/2} q^2\,\alpha\left(\left[\frac{N_1}{2q}\right]\bigwedge\left[\frac{N_2}{2q}\right];\, \left[\frac{N}{4q^2}\right], N\right).$$

Proof. The key idea of this proof is to divide the rectangular $\{(u,v) : 1 \le u \le N_1,\ 1 \le v \le N_2\}$ into $4q^2$ small blocks (see (3)), and then apply Bradley's coupling lemma to replace the sums of random variables on those blocks by their independent counterparts. The required inequalities (4) and (5) then follow from Hoeffding's inequality and Bernstein's inequality respectively. We outline the main steps of the proof as follows.

Let $p = p_1p_2 = N/(4q^2)$, $\delta = 1 + \varepsilon/(2K)$, $\beta = \min\{N\varepsilon/(8q^2), (\delta-1)Kp\}$, and $c = \delta Kp$. Then $|V_{ij}^{(1)} + c| \ge c - |V_{ij}^{(1)}| \ge (\delta-1)Kp$ almost surely. Applying Bradley's coupling lemma (see, e.g. Lemma 1.2 of Bosq 1998) recursively, we may define a sequence of independent random variables $\{W_{ij}^{(1)}\}$ such that $W_{ij}^{(1)}$ and $V_{ij}^{(1)}$ share the same marginal distribution, and further

$$
\begin{aligned}
P(|W_{ij}^{(1)} - V_{ij}^{(1)}| > \beta) &\leq 11 \left(\frac{(\delta+1)Kp}{\min\{n\varepsilon/(8q^2), (\delta-1)Kp\}} \right)^{\frac{1}{2}} \alpha([p_1] \wedge [p_2]; [p], N) \\
&= 11 \left\{ \max\left(\frac{\delta+1}{\delta-1}, \frac{8(\delta+1)q^2pK}{N\varepsilon} \right) \right\}^{\frac{1}{2}} \alpha([p_1] \wedge [p_2]; [p], N) \\
&= 11 \left(1 + \frac{4K}{\varepsilon} \right)^{1/2} \alpha([p_1] \wedge [p_2]; [p], N).
\end{aligned}
$$

Since $\{W_{ij}^{(1)}\}$ are independent, it follows from Hoeffding's inequality (Theorem 1.2 of Bosq 1998) that

$$
P\left(\left| \sum_{i,j} W_{ij}^{(1)} \right| > \frac{N\varepsilon}{8} \right) \leq 2\exp\left(-\frac{N^2\varepsilon^2}{128q^2p^2K^2} \right) = 2\exp\left(-\frac{N\varepsilon^2}{32pK^2} \right). \quad (6)
$$

Combining the above two inequalities, we have that

$$
\begin{aligned}
& P\left(\left| \sum_{i,j} V_{ij}^{(1)} \right| > \frac{N\varepsilon}{4} \right) \\
\leq\ & P\left(\left| \sum_{i,j} V_{ij}^{(1)} \right| > \frac{N\varepsilon}{4},\ |V_{ij}^{(1)} - W_{ij}^{(1)}| \leq \beta \text{ for all } i,j \right) + q^2 P(|V_{ij}^{(1)} - W_{ij}^{(1)}| > \beta) \\
\leq\ & P\left(\left| \sum_{i,j} W_{ij}^{(1)} \right| > \frac{N\varepsilon}{4} - q^2\beta \right) + q^2\, P(|V_{ij}^{(1)} - W_{ij}^{(1)}| > \beta) \\
\leq\ & P\left(\left| \sum_{i,j} W_{ij}^{(1)} \right| > \frac{N\varepsilon}{8} \right) + q^2\, P(|V_{ij}^{(1)} - W_{ij}^{(1)}| > \beta) \\
\leq\ & 2\exp\left(-\frac{q^2\varepsilon^2}{8K^2} \right) + 11 \left(1 + \frac{4K}{\varepsilon} \right)^{1/2} q^2\, \alpha([p_1] \wedge [p_2]; [p], N).
\end{aligned}
$$

It is easy to see that the above inequality also holds for $\{V_{ij}^{(l)}\}$ for $l = 2, 3, 4$. Now (4) follows immediately from the relation

$$
P(|S_N| > N\varepsilon) \leq \sum_{l=1}^{4} P\left(\left| \sum_{i,j} V_{ij}^{(l)} \right| > \frac{N\varepsilon}{4} \right),
$$

which is implied by (3).

Inequality (5) may be proved in the same manner with (6) replaced by

$$P\left(\left|\sum_{i,j} W_{ij}^{(1)}\right| > \frac{N\varepsilon}{8}\right) \le 2\exp\left(-\frac{\varepsilon^2 N^2/64}{4\sum_{i,j} E(W_{ij}^{(1)})^2 + 2pKN\varepsilon/8}\right)$$
$$\le 2\exp\left(-\frac{\varepsilon^2 q^2}{8\nu(q)^2}\right),$$

which is guaranteed by Bernstein's inequality (Theorem 1.2 of Bosq 1998). □

3 Density estimation for spatial processes

3.1 Estimators and regularity conditions

We assume now that the process $\{X(\mathbf{s})\}$ is strictly stationary with marginal density function $f(\cdot)$. The kernel estimator for f is defined as

$$\widehat{f}(x) = \frac{1}{N}\sum_{u=1}^{N_1}\sum_{v=1}^{N_2} W_h\{X(u,v) - x\}, \tag{7}$$

where $W_h(\cdot) = h^{-1}W(\cdot/h)$, $W(\cdot)$ is a probability density function defined in R and $h > 0$ is a bandwidth. We introduce some regularity conditions first. We use C to denote some positive generic constant which may be different at different places.

(C1) As $N = N_1 N_2 \to \infty$, N_1/N_2 converges to a positive and finite constant.

(C2) As $N \to \infty$, $h \to 0$ and $N^{\beta-5}h^{\beta+5}(\log N)^{-(\beta+1)} \to \infty$, where $\beta > 5$ is a constant.

(C3) The kernel function $W(\cdot)$ is bounded, symmetric and Lipschitz continuous.

(C4) The density function $f(\cdot)$ has continuous second derivative $\ddot{f}(\cdot)$. Further the joint density function $\{X(u,v), X(u+i, v+j)\}$ is bounded by a constant independent of (i,j).

(C5) It holds that $\alpha(k;\, k', j) \le Ck^{-\beta}$ for any k, j and $k' = O(k^2)$.

Conditions (C2) – (C4) are standard in kernel estimation. For optimum bandwidth $h = O(N^{-1/5})$, (C2) requires $\beta > 7.5$. For causal and invertible ARMA processes satisfying conditions in Remark 1, $\alpha(k; k', \infty)$ decays at an exponential rate as $k \to \infty$. Therefore, condition (C5) fulfils for any $\beta > 0$.

Remark 2. We assume in this paper that the observations were taken from a rectangular. This assumption can be relaxed. In fact Proposition 1 and Theorem 2 below still hold if the observations were taken over a connected region in Z^2, and both minimal length of the side of the squares containing the region and the maximal length of side of the squares contained in the region converge to infinite at the same rate. For general discussion on the condition of sampling sets, we refer to Perera (2001).

3.2 Asymptotic means and variances

Proposition 1. Let conditions (C1) and (C3) — (C5) hold with $\beta > 4$. Then for $h \to 0$ and $Nh \to \infty$ as $N \to \infty$, it holds that

$$E\{\widehat{f}(x)\} = f(x) + \frac{1}{2}h^2 \ddot{f}(x) \int u^2 W(u)du + o(h^2), \tag{8}$$

and

$$\mathrm{Var}\{\widehat{f}(x)\} = \frac{1}{Nh} f(x) \int W(u)^2 du + o(N^{-1}h^{-1}). \tag{9}$$

Proof. Equation (8) follows from simple algebraic manipulation. Put $Z_{uv} = W_h\{X(u,v) - x\}$. Then,

$$\mathrm{Var}\{\widehat{f}(x)\} = \frac{1}{N}\mathrm{Var}(Z_{11}) + \frac{1}{N^2} \sum_{(u,v)\neq(i,j)} \mathrm{Cov}(Z_{uv}, Z_{ij}).$$

The first term on the RHS of the above expression is equal to the RHS of (9). We only need to prove the second term is of the order $o(\frac{1}{Nh})$. To this end, note that for $(u,v) \neq (i,j)$,

$$|\mathrm{Cov}(Z_{uv}, Z_{ij})| \leq C_1,$$

where $C_1 > 0$ is a constant independent of u, v, i, j. Define a unilateral order in Z^2 as follows: $(u,v) > 0$ if either $u > 0$ or $u = 0$ and $v > 0$, and further $(u,v) > (i,j)$ if $(u-i, v-j) > 0$. Let $\mathcal{S}_n = \{(u,v) : \ 1 \leq u \leq N_1, \ 1 \leq v \leq N_2\}$. Then

$$\frac{1}{N^2} \sum_{(u,v)\neq(i,j)} |\mathrm{Cov}(Z_{uv}, Z_{ij})| = \frac{2}{N^2} \sum_{(u,v)<(i,j)} |\mathrm{Cov}(Z_{uv}, Z_{ij})|$$

$$= \frac{2}{N^2} \sum_{(u,v)\in\mathcal{S}_N} \sum_{\substack{(i,j)>0 \\ (u+i,v+j)\in\mathcal{S}_N}} |\mathrm{Cov}(Z_{uv}, Z_{u+i,v+j})|$$

$$
\begin{aligned}
&\leq \frac{2}{N^2} \sum_{(u,v)\in\mathcal{S}_N} \left\{ \sum_{\substack{(i,j)>0,\, i^2+j^2\leq T^2 \\ (u+i,v+j)\in\mathcal{S}_N}} + \sum_{i^2+j^2>T^2} \right\} |\mathrm{Cov}(Z_{uv}, Z_{u+i,v+j})| \\
&= O(T^2/N) \; + \frac{2}{N^2} \sum_{(u,v)\in\mathcal{S}_N} \sum_{i^2+j^2>T^2} |\mathrm{Cov}(Z_{uv}, Z_{u+i,v+j})|, \qquad (10)
\end{aligned}
$$

where $T = T(N) > 0$ is a constant. By Billingsley's inequality (see, e.g. Corollary 1.1 of Bosq 1998) and condition (C5),

$$
|\mathrm{Cov}(Z_{uv}, Z_{u+i,v+j})| \leq C\alpha\{(i^2+j^2)^{\frac{1}{2}}; 1, 1\}/h^2 \leq C(i^2+j^2)^{-\frac{\beta}{2}}/h^2.
$$

Thus the second term on the RHS of (10) is bounded from the above by

$$
\frac{2C}{Nh^2T^{\beta-2}} \sum_{i^2+j^2>T^2} \left(\frac{i^2+j^2}{T^2}\right)^{-\frac{\beta}{2}} \frac{1}{T^2} = O\left(\frac{1}{Nh^2T^{\beta-2}}\right) = o\left(\frac{1}{Nh}\right),
$$

provided $T = h^{-2/\beta}$ for $\beta > 4$. This also ensures $O(\frac{T^2}{N}) = o(\frac{1}{Nh})$; see (10). The proof is completed now. □

3.3 Optimum uniform convergence rates

Theorem 2. Under conditions (C1) – (C5), it holds that for any finite $a < b$,

$$
\sup_{x\in[a,b]} \left|\widehat{f}(x) - E\widehat{f}(x)\right| = O_p\left\{\left(\frac{\log N}{Nh}\right)^{1/2}\right\}. \qquad (11)
$$

Further,

$$
\sup_{x\in[a,b]} \left|\widehat{f}(x) - f(x)\right| = O_p\left\{\left(\frac{\log N}{Nh}\right)^{1/2} + h^2\right\}. \qquad (12)
$$

Remark 3. Theorem 2 presents the uniform convergence rates for the kernel estimator $\widehat{f}(\cdot)$. For $h = O\{(\log N/N)^{1/5}\}$, (12) admits the form

$$
\sup_{x\in[a,b]} \left|\widehat{f}(x) - f(x)\right| = O_p\left\{\left(\frac{\log N}{N}\right)^{2/5}\right\},
$$

which is the optimal convergence rate according to Hasminskii (1978). The similar results for single-indexed processes may be found in, for example, Masry (1996) and Theorem 5.3 of Fan and Yao (2003).

Proof of Theorem 2. Write the N observations as $X(\mathbf{s}_1), \cdots, X(\mathbf{s}_N)$. Partition $[a,b]$ into L subintervals $\{I_j\}$ of equal length. Let x_j be the centre of I_j. Since

$$|\widehat{f}(x) - \widehat{f}(x')| \le \frac{1}{N}\sum_{j=1}^{N} |W_h\{X(\mathbf{s}_j) - x\} - W_h\{X(\mathbf{s}_j) - x'\}| \le \frac{C}{h}|x - x'|,$$

it holds that $|E\widehat{f}(x) - E\widehat{f}(x')| \le \frac{C}{h}|x - x'|$. Hence

$$\sup_{z \in I_j} |\widehat{f}(x) - E\widehat{f}(x)| \le |\widehat{f}(x_j) - E\widehat{f}(x_j)| + \frac{C}{Lh}.$$

Therefore,

$$\sup_{x\in[a,b]} |\widehat{f}(x) - E\widehat{f}(x)| \le \max_{1\le j\le L} |\widehat{f}(x_j) - E\widehat{f}(x_j)| + \frac{C}{Lh}. \tag{13}$$

Since $|W_h(\cdot)| \le Ch^{-1}$, it follows from (5) that

$$\begin{aligned} P\{|\widehat{f}(x) - E\widehat{f}(x)| > \varepsilon\} \ &\le \ 8\exp\left(-\frac{\varepsilon^2 q^2}{8\nu(q)^2}\right) \\ + 44\left(1 + \frac{4C}{\varepsilon h}\right)^{1/2} & q^2\alpha([p_1]\wedge[p_2]; [p_1 p_2], N), \end{aligned} \tag{14}$$

where $p_i = N_i/(2q)$. Let $q = [\varepsilon^{1/2}(N_1 \wedge N_2)]$. By (9), $\nu(q)^2 \le C/(p_1p_2h) + C\varepsilon/h \le C\varepsilon/h$. Now let $\varepsilon^2 = \frac{8aC\log N}{(N_1\wedge N_2)^2 h}$ for some large constant $a > 0$. It is easy to see that

$$\exp\left(-\frac{\varepsilon^2 q^2}{8\nu(q)^2}\right) \le \exp\left(-\frac{\varepsilon^2 (N_1\wedge N_2)^2 h}{8C}\right) = N^{-a}. \tag{15}$$

On the other hand, condition (C5) entails that

$$\begin{aligned} (\varepsilon h)^{-1/2} q^2 \alpha([p_1]\wedge[p_2]; [p_1p_2], N) \ &\le \ C(\varepsilon h)^{-1/2} q^2 (p_1 \wedge p_2)^{-\beta} \\ = O(\varepsilon^{(\beta+1)/2} N h^{-1/2}) \ &= \ O\{N^{-\beta/4+3/4} h^{-\beta/4-3/4} (\log N)^{\beta/4+1/4}\} \end{aligned} \tag{16}$$

Let $L = (N/h)^{1/2}$. It follows from (14) – (16) and condition (C2) that

$$\begin{aligned} & P\{\max_{1\le j\le L} |\widehat{f}(x_j) - E\widehat{f}(x_j)| > \varepsilon\} \\ & \le L\{N^{-a} + CN^{-\beta/4+3/4} h^{-\beta/4-3/4} (\log N)^{\beta/4+1/4}\} \to 0. \end{aligned}$$

Note that $\varepsilon = O\{(\frac{\log N}{Nh})^{1/2}\}$. Now (11) follows from (13) immediately.

Note that

$$\sup_{x\in[a,b]} \left|\widehat{f}(x) - f(x)\right| \leq \sup_{x\in[a,b]} \left|\widehat{f}(x) - E\widehat{f}(x)\right| + \sup_{x\in[a,b]} \left|E\widehat{f}(x) - f(x)\right|,$$

and the second term on the RHS of the above expression is non-random. Simple algebraic manipulation shows that it is of the order h^2 under the condition that W is symmetric and f has two continuous derivatives. Now (12) follows from (11). □

Acknowledgments

The research was partially supported by a BBSRC/EPSRC joint research grant.

References

1. Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes* (2nd edition). Springer, New York.
2. Diggle, P.J. (1985). A kernel method for smoothing point process data. *Appl. Statist.* **34**, 138-147.
3. Diggle, P.J. and Marron, J.S. (1988). Equivalence of smoothing parameter selectors in density and intensity estimation. *J. Amer. Statist. Assoc.* **83**, 793-800.
4. Dobrushin, R.L. (1968). The description of the random field by its conditional distribution. *Theory Probab. Appl.* **13**, 201-229.
5. Doukhan, P. (1994). *Mixing*. Springer-Verlag, New York.
6. Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer, New York.
7. Guyon, X. (1995). *Random Fields on a Network: Modeling, Statistics, and Application*. Springer-Verlag, New York.
8. Hallin, M., Lu, Z. and Tran, L.T. (2002). Density estimation for linear processes. *Bernoulli*, to appear.
9. Hasminskii, R.Z. (1978). A lower bound on the risks of nonparametric estimates densities in the uniform metric. *Theory Prob. Appl.*, **23**, 794–798.
10. Masry, E. (1996). Multivariate local polynomial regression for time series: uniform strong consistency and rates. *J. Time Series Analy.* **17**, 571-559.
11. Perera, G. (2001). Random fields on Z^d: limit theorems and irregular sets. In *Spatial Statistics: Methodological Aspects and Applications*, M. Moore (edit.), p.57-82. Springer, New York.

12. Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, **41**, 434-449.
13. Yao, Q. and Brockwell, P.J. (2001). Gaussian maximum likelihood estimation for ARMA models II: spatial processes. *A preprint.*
14. Zhang, W., Yao, Q., Tong, H. and Stenseth, N.C. (2002). Smoothing for spatio-temporal models and its application in modelling muskrat-mink interaction. *A preprint.*