MEAN SQUARE PROPERTIES OF A CLASS OF KERNEL DENSITY ESTIMATES FOR SPATIAL FUNCTIONAL RANDOM VARIABLES

MOÏSE BASSE [§], SOPHIE DABO-NIANG [♯] AND ALIOU DIOP[§]

ABSTRACT. We investigate a kernel estimator of the probability density of a stationary random field $\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^N\}$ valued in a semi-metric space. We establish under mild mixing conditions the simple mean square consistency of the estimate when the field is observed over a rectangular domain of $\mathbb{N}^N$.

Key Words: Kernel density estimation; Spatial process; Functional random variables.

1. Introduction

Modelling spatial data has been the purpose of many investigations in the statistical literature. See for instance, Cressie (1991), Ripley (1981), Anselin and Florax (1995), Guyon (1995) and others for exposition, linear methods and applications. Spatial models are developed to take into account dependence structure observed between measurements at different positions for data collected from different spatial locations. Geology, soil science, image processing, ecology, atmospheric science are all examples of discipline where spatial data are collected.
The purpose of this paper is to study a kernel density estimator for random fields $X_{\mathbf{i}}$ which are valued in a space of eventually infinite dimensional, where the spatial index $\mathbf{i}$ varies discretely throughout the space $\mathbb{Z}^N$ $(N > 1)$.

Key references on spatial nonparametric estimation of the density of a real random field are Tran (1990), Tran and Yakowitz (1993), Carbon, Hallin and Tran (1996), Carbon, Tran and Wu (1997), Hallin, Lu and Tran (2004), Carbon (2006), Biau (2003), Bensaid and Dabo-Niang (2006), Dabo-Niang and Yao (2007), Carbon, Francq and Tran (2007), Lu *et al.* (2007) among others.

Problems involving density estimation for a spatial random variable taking values in an infinite dimensional space have received little attention until now. Dabo-Niang and Yao (2006) studied the convergence in probability and the strong convergence of the kernel density estimate for spatial functional random variables, under various types of asymptotics and mixing assumptions. In general, density estimation is a field for which

♯ Laboratoire GREMARS-EQUIPPE, Université Charles De Gaulle, Lille 3, maison de la recherche, domaine du pont de bois, BP 60149, 59653 Villeneuve d'ascq cedex, France. sophie.dabo@univ-lille3.fr.

§ Laboratoire LERSTAD, UFR de Sciences Appliquées et de Technologie, Université Gaston Berger, BP 234, Saint-Louis, Sénégal. bossmath@yahoo.fr, alioudiop52@yahoo.fr.

advanced theoretical studies in infinite dimension case was quite underdeveloped (see Dabo-Niang (2002-2004), Dabo-Niang, Ferraty and Vieu (2006) and Ferraty and Vieu (2006)).

In the setting of univariate or multivariate real random variables, the probability density of the sample is usually defined with respect to some reference measure (Lebesgue's measure or counting measure). When one studies a sample composed of functional data, for instance when the statistical unit is a curve (see Ramsay and Silverman (1997-2002), Bosq (2000), Ferraty and Vieu (2006)), the problem is much harder since there is no commonly accepted reference measure. Therefore, all our theoretical developments are given for any abstract measure making the probability distribution of the sample absolutely continuous.

The paper is organized as follows. The next section sets up the notations and the assumptions which will be considered in the sequel. Section 3 is devoted to some preliminaries results. In Section 4, we stated the mean square consistency results of the spatial kernel density estimate. Proofs of the main results and technical lemmas are postponed in an appendix in Section 5.

2. Spatial density estimator

Consider $\left(X_{\mathbf{i}},\ \mathbf{i} \in \mathbb{Z}^N\right)$, $N \geq 1$, a measurable strictly stationary spatial process defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$, having the same distribution as a variable X and valued in a separable semi-metric space $(\mathcal{E},\ d(.,.))$ which is eventually of infinite dimension. A point $\mathbf{i} = (i_1, ..., i_N) \in \mathbb{N}^N$ will be referred to as a site. We suppose that the $X_{\mathbf{i}}$'s have a common distribution with an unknown density f with respect to some given measure μ. As mentioned above, we will give our results without specifying this reference measure. Let $\mathcal{B}_{\mathcal{E}}$ be the Borel σ-field generated by the open sets of $\mathcal{E}$. We will assume that the measure μ is σ-finite and is such that $0 < \mu(A) < \infty$, for any open ball $A \subset \mathcal{E}$. We define a rectangular region $\mathcal{I}_{\mathbf{n}}$ by $\mathcal{I}_{\mathbf{n}} = \left\{\mathbf{i} \in \mathbb{N}^N : 1 \leq i_k \leq n_k,\ k = 1, ..., N\right\}$ where $\mathbf{n} = (n_1, ..., n_N)$. We write $\mathbf{n} \to +\infty$ if $\min_{k=1,...,N} n_k \to +\infty$ and $\left|\frac{n_j}{n_k}\right| < C$ for some constant $0 < C < \infty$ and $\forall\, j, k \in \{1, ..., N\}$. All the limits will be taken as $\mathbf{n} \to +\infty$. We set $\widehat{\mathbf{n}} = n_1 \times ... \times n_N$. Throughout the paper, C will denote an arbitrary constant.

We deal with the estimation of the density f. We suppose that a spatial sample is available on $\mathcal{I}_{\mathbf{n}}$. We define a general kernel estimate of f based on $(X_{\mathbf{i}},\ \mathbf{i} \in \mathcal{I}_{\mathbf{n}})$ as:

$$f_{\mathbf{n}}(x) = \frac{1}{\widehat{\mathbf{n}} a_{\mathbf{n}}^x} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_{\mathbf{n}}\left(d\left(X_{\mathbf{i}}, x\right)\right), \qquad x \in \mathcal{E},$$

where $(a_{\mathbf{n}}^x)$ is a sequence of strictly positive reals depending on x with $\lim_{\mathbf{n} \to +\infty} a_{\mathbf{n}}^x = 0$ and $K_{\mathbf{n}}$ is a real valued function which depends on $\mathbf{n}$.

We will also consider the simple case of kernel naive estimator where $K_{\mathbf{n}}(d(t,x)) = \mathbb{1}_{B^x_{r_{\mathbf{n}}}}(t)$ is the indicator function and $a^x_{\mathbf{n}} = \mu(B^x_{r_{\mathbf{n}}})$, $B^x_{r_{\mathbf{n}}}$ is the closed ball of center x and radius $r_{\mathbf{n}} > 0$. This estimator is the following

$$\hat{f}^1_{\mathbf{n}}(x) = \frac{1}{\widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}})} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{i}}) \qquad x \in \mathcal{E}. \tag{1}$$

The general estimate $f_{\mathbf{n}}(x)$ is an extension of the well known Parzen (1962)-Rosenblatt (1956)'s estimate. A great attention has been paid to estimate the density (spatial and non-spatial cases) of a distribution of a random variable valued in a finite dimensional space by using a kernel method, see the references therein.

2.1. **General assumptions.** The following condition concerns the nonparametric statistical model and is the same as in the classical finite dimensional case:

F_1 - f is continuous at $x \in \mathcal{E}$.

We need the following assumptions on our estimates. We will not discuss now these technical hypotheses, but the special case treated in Theorem 1 of Section 4 shows that they are quite unrestrictive. Precisely, we need to assume:

H_1 - $\forall\delta, 0 < \delta \leq +\infty,\ \lim_{\mathbf{n}\to\infty} \left| \frac{1}{a^x_{\mathbf{n}}} \int_{d(y,x)<\delta} K_{\mathbf{n}}\left(d\left(y,x\right)\right) d\mu(y) - 1 \right| = 0;$

H_2 - $\forall x \in \mathcal{E}$:

$$S^x_{\mathbf{n}} = \sup_{y\in\mathcal{E}} K_{\mathbf{n}}\left(d(x,y)\right)/a^x_{\mathbf{n}} < \infty.$$

H_3 - $\|f\|_\infty = \sup_{y\in\mathcal{E}} |f(y)| < \infty$.

We set $\lim_{\mathbf{n}\to+\infty} S^x_{\mathbf{n}} = +\infty$, since $(a^x_{\mathbf{n}})$ is a sequence of strictly positive reals such that $\lim_{\mathbf{n}\to+\infty} a^x_{\mathbf{n}} = 0$.
For more discussion on these conditions, see Dabo-Niang, Ferraty and Vieu (2006).

2.2. **Dependency conditions.** In spatial dependent data analysis, the dependence of the observations has to be measured. Here we will consider the following two dependence measures:

2.2.1. *Mixing condition.* We will measure the spatial dependence of the process by means of α-mixing. We assume that $\left(X_{\mathbf{i}},\ \mathbf{i} \in \mathbb{N}^N\right)$ satisfies the following mixing condition: there exists a function $\varphi(t) \downarrow 0$ as $t \to \infty$, such that whenever $E,\ E'$ subsets of $\mathbb{N}^N$ with finite cardinals,

$$\alpha\left(\mathcal{B}(E),\mathcal{B}\left(E'\right)\right) = \sup_{B\in\mathcal{B}(E),\, C\in\mathcal{B}\left(E'\right)} \left|\mathbf{P}(B\cap C) - \mathbf{P}(B)\,\mathbf{P}(C)\right|$$

$$\le \psi\left(\operatorname{Card}(E),\operatorname{Card}\left(E'\right)\right)\varphi\left(\operatorname{dist}\left(E,E'\right)\right), \tag{2}$$

where $\mathcal{B}(E)$(*resp.* $\mathcal{B}\left(E'\right)$) denotes the Borel σ-field generated by $(X_{\mathbf{i}},\ \mathbf{i}\in E)$ (*resp.* $\left(X_{\mathbf{i}},\ \mathbf{i}\in E'\right)$), Card$(E)$ (*resp.* Card$\left(E'\right)$) the cardinality of E (*resp.* E'), dist$\left(E,E'\right)$ the Euclidean distance between E and E' and $\psi : \mathbb{N}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable. Throughout the paper, it will be assumed that ψ satisfies either

$$\psi(n,m) \le C\min(n,m), \qquad \forall n,m\in\mathbb{N} \tag{3}$$

or

$$\psi(n,m) \le C(n+m+1)^{\widetilde{\beta}}, \qquad \forall n,m\in\mathbb{N} \tag{4}$$

for some $\widetilde{\beta}\ge 1$ and some $C>0$. We assume also that the process satisfies a polynomial mixing condition:

$$\varphi(t) \le Ct^{-\theta},\ \ \theta>0,\ t\in\mathbb{R},\ C>0. \tag{5}$$

Another condition on $\varphi(i)$ is also often used in the literature: that is $\varphi(i) = C\ \exp(-si)$ tends to zero at an exponential rate, for some $s>0$.

These two exponential and polynomial conditions are linked (see Bosq (1998) for details).

If $\psi\equiv 1$, then $X_{\mathbf{i}}$ is called strongly mixing. Many stochastic processes, among them various useful time series models satisfy strong mixing properties, which are relatively easy to check. Conditions (3)-(4) are used in Tran (1990), Carbon *et al.* (1996-1997), Tran *et al.* (1993), Biau and Cadre (2004). See Doukhan (1994), Rio (2000) for discussion on mixing and examples.

2.2.2. *Local dependency condition.* We will assume that the joint probability $P\left[(X_{\mathbf{i}},X_{\mathbf{j}})\in B^x_{r_{\mathbf{n}}}\times B^x_{r_{\mathbf{n}}}\right]$ of $(X_{\mathbf{i}},X_{\mathbf{j}})$ satisfies the following local dependency condition which is necessary to reach the same rate of convergence as in the *i.i.d* case:

$$\forall \mathbf{i},\mathbf{j}\in\mathbb{Z}^N,\ P\left[(X_{\mathbf{i}},X_{\mathbf{j}})\in B^x_{r_{\mathbf{n}}}\times B^x_{r_{\mathbf{n}}}\right] = \left(\mu(B^x_{r_{\mathbf{n}}})\right)^{1+\varepsilon_1},\ \text{ for some } \varepsilon_1\in(0,1). \tag{6}$$

or

$$(7) \quad \forall \mathbf{i},\mathbf{j} \in \mathbb{Z}^N, \ \frac{1}{(a^x_{\mathbf{n}})^2} E\left(K_{\mathbf{n}}(d(x,X_{\mathbf{i}}))K_{\mathbf{n}}(d(x,X_{\mathbf{j}}))\right) = \left(S^x_{\mathbf{n}}\right)^{1-\varepsilon_1}, \ \text{ for some } \varepsilon_1 \in (0,1).$$

3. Preliminaries Results

The following two lemmas are due to Carbon *et al.* (1997). They are needed for the mean square convergence of our estimates. See the last reference for the proofs.

Lemma 1.
(i) Suppose that (2) holds. Denote by $\mathcal{L}_r(\mathcal{F})$ the class of $\mathcal{F}-$measurable r.v.'s X satisfying $\|X\|_r = (E|X|^r)^{1/r} < \infty$. Suppose $X \in \mathcal{L}_r(\mathcal{B}(E))$ and $Y \in \mathcal{L}_s(\mathcal{B}(E'))$. Assume also that $1 \leq r,\ s,\ t < \infty$ and $r^{-1}+s^{-1}+t^{-1} = 1$. Then

$$(8) \qquad |EXY - EXEY| \leq C\|X\|_r\|Y\|_s\{\psi(Card(E), Card(E'))\varphi(dist(E,E'))\}^{1/t}.$$

(ii) For r.v.'s bounded with probability 1, the right-hand side of (8) can be replaced by $C\psi(Card(E), Card(E'))\varphi(dist(E,E'))$.

Lemma 2.
If (5) holds for $\theta > 2N$, then

$$(9) \qquad \sum_{i=1}^{\infty} i^{N-1}(\varphi(i))^a < \infty,$$

for some $0 < a < 1/2$.

The following lemma is a direct consequence of Lemma 2 of Chapter 6 of Dabo-Niang (2002) or Theorem 1 of Dabo-Niang (2004), so its proof will be omitted.

Lemma 3. *Let $g \in L_1(\mathcal{E},\ \mu)$. If g is continuous at $x \in \mathcal{E}$ then, $lim_{\mathbf{n}\to\infty}\, r_{\mathbf{n}} = 0$ implies that:*

$$\lim_{\mathbf{n}\to\infty} \frac{1}{\mu(B^x_{r_{\mathbf{n}}})} \int_{B^x_{r_{\mathbf{n}}}} g(y)d\mu(y) = g(x).$$

The following two lemmas give respectively the convergence of the bias and asymptotic results of the variances of the two kernel estimates given above.

Lemma 4. *Let $x \in \mathcal{E}$, if F_1, H_1-H_3 and (7) hold then*

$$\lim_{\mathbf{n}\to+\infty} \frac{1}{a^x_{\mathbf{n}}} \int_{\mathcal{E}} K_{\mathbf{n}}(d(x,t))f(t)d\mu(t) = f(x).$$

Lemma 5.
Let $x \in \mathcal{E}$ and (3) or (4) be satisfied with $\sum_{i=1}^{\infty} i^{N-1}(\varphi(i))^a < \infty$, for some $0 < a < 1/2$.

(i) If F_1 and (6) hold, we get

$$\lim_{\mathbf{n}\to\infty} \widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}})V_f(\hat{f}^1_{\mathbf{n}}(x)) < \infty,$$

(ii) If F_1, H_1-H_3 and (7) hold then:

$$\lim_{\mathbf{n}\to\infty} (\widehat{\mathbf{n}}/S_{\mathbf{n}}^x) V_f(f_{\mathbf{n}}(x)) < \infty.$$

Remark 1. *As stated in Lemma 2, if (5) holds with $\theta > 2N$ then $\sum_{i=1}^{\infty} i^{N-1}(\varphi(i))^a < \infty$ for some $0 < a < 1/2$. It suffices to choose $(N/\theta) < a < 1/2$.*

4. Mean square consistency results

We are now able to state mean square consistency results of the two estimates. The following result gives the mean square consistency of the naive estimator. To prove that the conditions on $B_{r_{\mathbf{n}}}^x$ and $r_{\mathbf{n}}$ given in the following theorem are necessary and sufficient to the mean square convergence of the naive estimator, we assume that the measure μ is also diffuse.

Theorem 1. *Let $x \in \mathcal{E}$, if F_1, (6), (3) or (4) are satisfied, then*

$$\lim_{\mathbf{n}\to\infty} r_{\mathbf{n}} = 0 \quad \text{and} \quad \lim_{\mathbf{n}\to\infty} \widehat{\mathbf{n}}\mu(B_{r_{\mathbf{n}}}^x) = \infty,$$

are equivalent to

(10) $$\lim_{\mathbf{n}\to\infty} E_f(\hat{f}_{\mathbf{n}}^1(x) - f(x))^2 = 0,$$

when (5) holds with $\theta > 2N$.

The following result gives the mean square consistency of the general kernel estimator.

Theorem 2. *Let $x \in \mathcal{E}$, if F_1, H_1-H_3 and (7), $\lim_{\mathbf{n}\to\infty}(\widehat{\mathbf{n}}/S_{\mathbf{n}}^x) = \infty$, (3) or (4) are satisfied, then we have:*

(11) $$\lim_{\mathbf{n}\to\infty} E_f(f_{\mathbf{n}}(x) - f(x))^2 = 0,$$

when (5) holds with $\theta > 2N$.

4.1. Conclusion.

This study is motivated by a desire to provide a non-parametric estimate of the probability density for functional spatial processes. A first step is provided by Tran (1990) where an asymptotic result of the variance and asymptotic normality of the kernel density estimate are considered. The present work extend his result to the case where the random variables are valued in a metric space of eventually infinite dimension. This is a modest step to functional spatial data analysis but still does not master regression or others functional nonparametric spatial problems. However, the notion of spatial density estimate is an useful tool for studying regression and classification models among others.

5. Appendix

The proofs of the results are stated in this section..

Proof. of Lemma 3

We have

$$\frac{1}{\mu(B^x_{r_{\mathbf{n}}})}\int_{B^x_{r_{\mathbf{n}}}} g(t)d\mu(t) - g(x) \leq \frac{1}{\mu(B^x_{r_{\mathbf{n}}})}\int_{B^x_{r_{\mathbf{n}}}} |g(t)-g(x)|\, d\mu(t).$$

Since g is continuous at x, for all $\epsilon > 0$, there exists $\eta_\epsilon > 0$ such that for all $y \in \mathcal{E}$:

$$d(y,x) < \eta_\epsilon \Rightarrow |g(y)-g(x)| < \epsilon.$$

We also note that $r_{\mathbf{n}} \to 0$, as $\mathbf{n} \to +\infty$, so for $\eta_\epsilon > 0$: there exists $\mathbf{n}_{\eta_\epsilon}$ such that $r_{\mathbf{n}} < \eta_\epsilon$, for all $\mathbf{n} \geq \mathbf{n}_{\eta_\epsilon}$. Therefore

$$d(y,x) < r_{\mathbf{n}} < \eta_\epsilon \Rightarrow |g(y)-g(x)| < \epsilon.$$

Finally, we get for all $\epsilon > 0$ there exists $\mathbf{n}_\epsilon$ such that for all $\mathbf{n} \geq \mathbf{n}_\epsilon$,

$$\frac{1}{\mu(B^x_{r_{\mathbf{n}}})}\left|\int_{B^x_{r_{\mathbf{n}}}} g(t)d\mu(t) - g(x)\right| \leq \frac{1}{\mu(B^x_{r_{\mathbf{n}}})}\int_{B^x_{r_{\mathbf{n}}}} |g(t)-g(x)|\, d\mu(t) < \epsilon.$$

This yields the proof.

□

Proof. of Lemma 4

The bias of $f_{\mathbf{n}}(x)$ for $x \in \mathcal{E}$ is

$$b(f_{\mathbf{n}}(x)) = \frac{1}{a^x_{\mathbf{n}}}\int_{\mathcal{E}} K_{\mathbf{n}}\left(d(x,t)\right) f(t)d\mu(t) - f(x).$$

Now, for some $\delta > 0$ and $x \in \mathcal{E}$, let

$$I_1(x) = \frac{1}{a^x_{\mathbf{n}}}\int_{\{t,\, d(x,t)<\delta\}} K_{\mathbf{n}}\left(d(x,t)\right)\left(f(t)-f(x)\right) d\mu(t),$$

and

$$I_2(x) = \frac{1}{a^x_{\mathbf{n}}}\int_{\{t,\, d(x,t)\geq\delta\}} K_{\mathbf{n}}\left(d(x,t)\right)\left(f(t)-f(x)\right) d\mu(t).$$

By H_1, we have

$$\lim_{\mathbf{n}\to\infty} |b(f_{\mathbf{n}}(x)) - I_1(x) - I_2(x)| = 0.$$

Since f is continuous at x, we have: $\forall\epsilon > 0$ there exits $\eta > 0$ such that for all $y \in \mathcal{E}$: $d(y,x) < \delta$ implies that $|f(y)-f(x)| < \epsilon$. Then we have

$$I_1(x) \leq \frac{1}{a^x_{\mathbf{n}}}\int_{\{t,\, d(x,t)<\delta\}} K_{\mathbf{n}}\left(d(x,t)\right)|f(t)-f(x)|\, d\mu(t) \leq \frac{1}{a^x_{\mathbf{n}}}\int_{\{t,\, d(x,t)<\delta\}} K_{\mathbf{n}}\left(d(x,t)\right)\epsilon d\mu(t) \leq \epsilon,$$

by H_1.
Moreover

$$I_2(x) \leq \frac{2\|f\|_\infty}{a_{\mathbf{n}}^x} \int_{\{t,\, d(x,t)\geq\delta\}} K_{\mathbf{n}}\left(d(x,t)\right) d\mu(t)$$

goes to zero by H_1 and H_3. Thus $I_1(x) \to 0$, $I_2(x) \to 0$. This yield the proof.

□

Proof. of Lemma 5

Let $G_{\mathbf{n}}(x) = f_{\mathbf{n}}(x)$ or $\hat{f}^1_{\mathbf{n}}(x)$. We have

$$V_f\left(G_{\mathbf{n}}(x)\right) = \widehat{\mathbf{n}}^2 \left[\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} V_f\left(Z_{\mathbf{i},\,x}\right) + \sum_{\mathbf{i}\neq\mathbf{j}} cov\left(Z_{\mathbf{i},\,x}, Z_{\mathbf{j},\,x}\right)\right]$$

where

$$Z_{\mathbf{i},\,x} = \frac{1}{a_{\mathbf{n}}^x} K_{\mathbf{n}}\left(d\left(X_{\mathbf{i}}, x\right)\right).$$

or

$$Z_{\mathbf{i},\,x} = \frac{1}{\mu(B^x_{r_{\mathbf{n}}})} 1\!\!1_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{i}}).$$

Then, we have

$$V_f\left(G_{\mathbf{n}}(x)\right) \leq \widehat{\mathbf{n}}^{-2} \left(\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} V_f(Z_{\mathbf{i},x}) + \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \sum_{\mathbf{j}\in\mathcal{I}_{\mathbf{n}},\, i_k\neq j_k \text{ for some } k} \left|cov\left(Z_{\mathbf{i},\,x}, Z_{\mathbf{j},\,x}\right)\right|\right) = \tilde{A}_1(x) + \tilde{A}_2(x).$$

Let us consider first the case of (i). We have

$$\widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}})\tilde{A}_1(x) = \frac{1}{\mu(B^x_{r_{\mathbf{n}}})} V_f(1\!\!1_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})) \leq \frac{1}{\mu(B^x_{r_{\mathbf{n}}})} \int_{B^x_{r_{\mathbf{n}}}} f(t)d\mu(t), \text{ for some } \mathbf{k}\in\mathcal{I}_{\mathbf{n}}.$$

Thus

$$\limsup_{\mathbf{n}\to+\infty} \widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}})\tilde{A}_1(x) \leq f(x),$$

by Lemma 3.
Let us split $\tilde{A}_2(x)$ into two parts:

$$\tilde{A}_2(x) = \widehat{\mathbf{n}}^{-2} \left(\sum \sum_{0<dist(\mathbf{i},\mathbf{j})\leq m_{\mathbf{n}}} cov(Z_{\mathbf{i},x}, Z_{\mathbf{j},x}) + \sum \sum_{dist(\mathbf{i},\mathbf{j})>m_{\mathbf{n}}} cov(Z_{\mathbf{i},x}, Z_{\mathbf{j},x})\right) = \tilde{I}_1(x) + \tilde{I}_2(x).$$

We have

$$E(Z_{\mathbf{i},x}Z_{\mathbf{j},x}) \leq \frac{1}{(\mu(B^x_{r_{\mathbf{n}}}))^2} P\left((X_{\mathbf{i}}, X_{\mathbf{j}}) \in B^x_{r_{\mathbf{n}}} \times B^x_{r_{\mathbf{n}}}\right)$$

$$\leq \frac{1}{(\mu(B^x_{r_{\mathbf{n}}}))^2} (\mu(B^x_{r_{\mathbf{n}}}))^{1+\varepsilon_1} = (\mu(B^x_{r_{\mathbf{n}}}))^{\varepsilon_1-1},$$

by assumption (6), where $0 < \varepsilon_1 < 1$. Then we get

$$\widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}})\tilde{I}_1(x) \leq C(\widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}}))\widehat{\mathbf{n}} m_{\mathbf{n}}^N \widehat{\mathbf{n}}^{-2} (\mu(B^x_{r_{\mathbf{n}}}))^{\varepsilon_1-1} = C m_{\mathbf{n}}^N (\mu(B^x_{r_{\mathbf{n}}}))^{\varepsilon_1}.$$

Let $m_{\mathbf{n}} = (\mu(B^x_{r_{\mathbf{n}}}))^{-(1-\gamma)\varepsilon_1/\nu}$, where $\nu = -N-\varepsilon_1+(1-\gamma)\varepsilon_1 N a^{-1}$, γ and ε_1 (see hypothesis (6)) are small positive numbers such that $a^{-1}\varepsilon_1-(N+\varepsilon_1)(N(1-\gamma))^{-1} > 1$ (this is possible since $0 < a < 1/2$). Then, we have

$$(\widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}}))\tilde{I}_1(x) \le C(\mu(B^x_{r_{\mathbf{n}}}))^{\varepsilon_1(1-N(1-\gamma)/\nu)},$$

thus $\lim_{\mathbf{n}\to\infty}(\widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}}))\tilde{I}_1(x) = 0$ since $\nu > N(1-\gamma)$.

Let us turn to $\tilde{I}_2(x)$ and let $\gamma' = 1-(1-\gamma)\varepsilon_1$, $\delta = 2(1-\gamma')/\gamma'$. Notice that $\gamma' = 2/(2+\delta)$ and $1-\gamma' = \delta/(2+\delta)$. We apply Lemma 1 with $r = s = 2+\delta$; $t = (2+\delta)/\delta$ and get

$$|E\left(Z_{\mathbf{i},x}Z_{\mathbf{j},x}\right) - E\left(Z_{\mathbf{i},x}\right)E\left(Z_{\mathbf{j},x}\right)| \le C\left(\frac{1}{(\mu(B^x_{r_{\mathbf{n}}}))^{2+\delta}}E\left[\mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})\right]\right)^{\gamma'}(\varphi(\|\mathbf{i}-\mathbf{j}\|))^{1-\gamma'}\ ,\ .$$

for some $\mathbf{k}\in\mathcal{I}_{\mathbf{n}}$.

Employing the previous inequality, we get:

$$\widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}})\tilde{I}_2(x) \le C\widehat{\mathbf{n}}^{-1}\mu(B^x_{r_{\mathbf{n}}})\sum\sum_{\|\mathbf{i}-\mathbf{j}\|>m_{\mathbf{n}}}\left(\frac{1}{(\mu(B^x_{r_{\mathbf{n}}}))^{2+\delta}}E\left[\mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})\right]\right)^{\gamma'}(\varphi(\|\mathbf{i}-\mathbf{j}\|))^{1-\gamma'}$$

$$\le C\widehat{\mathbf{n}}^{-1}\mu(B^x_{r_{\mathbf{n}}})(\mu(B^x_{r_{\mathbf{n}}}))^{-\gamma'(1+\delta)}\left(\frac{1}{\mu(B^x_{r_{\mathbf{n}}})}E\left[\mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})\right]\right)^{\gamma'}\sum\sum_{\|\mathbf{i}-\mathbf{j}\|>m_{\mathbf{n}}}(\varphi(\|\mathbf{i}-\mathbf{j}\|))^{1-\gamma'}$$

$$\le C\widehat{\mathbf{n}}^{-1}\mu(B^x_{r_{\mathbf{n}}})(\mu(B^x_{r_{\mathbf{n}}}))^{-\gamma'(1+\delta)}\left(\frac{1}{\mu(B^x_{r_{\mathbf{n}}})}E\left[\mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})\right]\right)^{\gamma'}\widehat{\mathbf{n}}\sum_{\|\mathbf{i}\|>m_{\mathbf{n}}}(\varphi(\|\mathbf{i}\|))^{1-\gamma'}$$

$$\le C(\mu(B^x_{r_{\mathbf{n}}}))^{-1+\gamma'}\left(\frac{1}{\mu(B^x_{r_{\mathbf{n}}})}E\left[\mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})\right]\right)^{\gamma'}\sum_{\|\mathbf{i}\|>m_{\mathbf{n}}}(\varphi(\|\mathbf{i}\|))^{1-\gamma'}\ .$$

We deduce that $\varphi(i) = o(i^{-N/a})$ from assumption $\sum_{i=1}^{\infty} i^{N-1}(\varphi(i))^a < \infty$. Therefore

$$\|\mathbf{i}\|^{\nu}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'} = \|\mathbf{i}\|^{\nu}o(\|\mathbf{i}\|^{-N(1-\gamma')/a}) = o(\|\mathbf{i}\|^{-N-\varepsilon_1}),$$

since $\varphi(t) = o(t^{-N/a})$ as $t\to\infty$ (because φ is a decreasing function) and

$$\nu = -N-\varepsilon_1+(1-\gamma)\varepsilon_1 N a^{-1} > 0.$$

Then

$$\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\|\mathbf{i}\|^{\nu}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'} < \infty.$$

Since $(\mu(B^x_{r_{\mathbf{n}}}))^{-1+\gamma'}m_n^{-\nu} = 1$ and

$$\left(\frac{1}{\mu(B^x_{r_{\mathbf{n}}})}E\left[\mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})\right]\right)^{\gamma'} = \left(\frac{1}{\mu(B^x_{r_{\mathbf{n}}})}\int_{B^x_{r_{\mathbf{n}}}}f(t)d\mu(t)\right)^{\gamma'} \to (f(x))^{\gamma'},$$

we obtain

$$\limsup \widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}})\tilde{I}_2(x) \leq C \limsup \left(\mu(B^x_{r_{\mathbf{n}}})\right)^{-1+\gamma'} \left(\frac{1}{\mu(B^x_{r_{\mathbf{n}}})} E\left[\mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})\right]\right)^{\gamma'} \sum_{\|\mathbf{i}\|>m_{\mathbf{n}}} \left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}$$

$$\leq C \lim \mu(B^x_{r_{\mathbf{n}}})^{-1+\gamma'} m_n^{-\nu} \left(\frac{1}{\mu(B^x_{r_{\mathbf{n}}})} E\left[\mathbb{1}_{B^x_{r_{\mathbf{n}}}}(X_{\mathbf{k}})\right]\right)^{\gamma'} \sum_{\|\mathbf{i}\|>m_{\mathbf{n}}} \|\mathbf{i}\|^{\nu} \left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}$$

$$\leq C \lim \sum_{\|\mathbf{i}\|>m_{\mathbf{n}}} \|\mathbf{i}\|^{\nu} \left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}.$$

This last term tends to a finite value when $\mathbf{n}$ tends to infinity.

Let us prove (ii), so consider the case where $G_{\mathbf{n}}(x) = f_{\mathbf{n}}(x)$, then

$$V_f\left(f_{\mathbf{n}}(x)\right) \leq \widehat{\mathbf{n}}^{-2}\left(\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} V_f(Z_{\mathbf{i},x}) + \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \sum_{\mathbf{j}\in\mathcal{I}_{\mathbf{n}},\, i_k\neq j_k \text{ for some } k} \left|cov\left(Z_{\mathbf{i},x}, Z_{\mathbf{j},x}\right)\right|\right) = \tilde{A}_1(x)+\tilde{A}_2(x).$$

We have

$$(\widehat{\mathbf{n}}/S^x_{\mathbf{n}})\tilde{A}_1(x) \leq \frac{1}{a^x_{\mathbf{n}}}\int_{\mathcal{E}} K_{\mathbf{n}}(d(x,t))f(t)d\mu(t).$$

Thus by Lemma 4

$$\limsup_{\mathbf{n}\to+\infty}(\widehat{\mathbf{n}}/S^x_{\mathbf{n}}))\tilde{A}_1(x) \leq f(x).$$

Let us split $\tilde{A}_2(x)$ into two parts

$$\tilde{A}_2(x) = \widehat{\mathbf{n}}^{-2}\left(\sum \sum_{0<dist(\mathbf{i},\mathbf{j})\leq m_{\mathbf{n}}} cov(Z_{\mathbf{i},x}, Z_{\mathbf{j},x}) + \sum \sum_{dist(\mathbf{i},\mathbf{j})>m_{\mathbf{n}}} cov(Z_{\mathbf{i},x}, Z_{\mathbf{j},x})\right) = \tilde{I}_1(x)+\tilde{I}_2(x).$$

We have

$$E(Z_{\mathbf{i},x}Z_{\mathbf{j},x}) = (S^x_{\mathbf{n}})^{-\epsilon_1+1}$$

by assumption (7), where $0 < \epsilon_1 < 1$. Then we get:

$$(\widehat{\mathbf{n}}/S^x_{\mathbf{n}})\tilde{I}_1(x) \leq (\widehat{\mathbf{n}}/S^x_{\mathbf{n}})\widehat{\mathbf{n}}m^N_{\mathbf{n}}\widehat{\mathbf{n}}^{-2}(S^x_{\mathbf{n}})^{-\varepsilon_1+1} = m^N_{\mathbf{n}}(S^x_{\mathbf{n}})^{-\varepsilon_1}.$$

Let $m_{\mathbf{n}} = (1/S^x_{\mathbf{n}})^{-(1-\gamma)\varepsilon_1/\nu}$, where $\nu = -N - \varepsilon_1 + (1-\gamma)\varepsilon_1 N a^{-1}$ (see the proof of (i)). We deduce that:

$$(\widehat{\mathbf{n}}/S^x_{\mathbf{n}})\tilde{I}_1(x) \leq (1/S^x_{\mathbf{n}})^{\varepsilon_1(1-N(1-\gamma)/\nu)},$$

thus $\lim_{\mathbf{n}\to\infty}(\widehat{\mathbf{n}}/S^x_{\mathbf{n}})\tilde{I}_1(x) = 0$ since $\nu > N(1-\gamma)$.

Let us turn to $\tilde{I}_2(x)$ and notice that (as above):

$$\left|E\left(Z_{\mathbf{i},x}Z_{\mathbf{j},x}\right)-E\left(Z_{\mathbf{i},x}\right)E\left(Z_{\mathbf{j},x}\right)\right|\leq C\left(\frac{1}{\left(a_{\mathbf{n}}^{x}\right)^{2+\delta}}E\left[K_{\mathbf{n}}(d(x,X_{\mathbf{k}}))\right]^{2+\delta}\right)^{\gamma'}\left(\varphi(\|\mathbf{i}-\mathbf{j}\|)\right)^{1-\gamma'},$$

for some $\mathbf{k}\in\mathcal{I}_{\mathbf{n}}$. Then, we get

$$(\widehat{\mathbf{n}}/S_{\mathbf{n}}^{x})\tilde{I}_{2}(x)\leq C(\widehat{\mathbf{n}}S_{\mathbf{n}}^{x})^{-1}\sum\sum_{\|\mathbf{i}-\mathbf{j}\|>m_{\mathbf{n}}}\left(\frac{1}{\left(a_{\mathbf{n}}^{x}\right)^{2+\delta}}E\left[K_{\mathbf{n}}(d(x,X_{\mathbf{k}}))\right]^{2+\delta}\right)^{\gamma'}\left(\varphi(\|\mathbf{i}-\mathbf{j}\|)\right)^{1-\gamma'}.$$

By Assumption H_2, we have

$$\begin{aligned}(\widehat{\mathbf{n}}/S_{\mathbf{n}}^{x})\tilde{I}_{2}(x)&\leq C(\widehat{\mathbf{n}}S_{\mathbf{n}}^{x})^{-1}(1/S_{\mathbf{n}}^{x})^{-\gamma'(1+\delta)}\left(\frac{1}{a_{\mathbf{n}}^{x}}E\left[K_{\mathbf{n}}(d(x,X_{\mathbf{k}}))\right]\right)^{\gamma'}\sum\sum_{\|\mathbf{i}-\mathbf{j}\|>m_{\mathbf{n}}}\left(\varphi(\|\mathbf{i}-\mathbf{j}\|)\right)^{1-\gamma'}\\
&\leq C(\widehat{\mathbf{n}}S_{\mathbf{n}}^{x})^{-1}(1/S_{\mathbf{n}}^{x})^{-\gamma'(1+\delta)}\left(\frac{1}{a_{\mathbf{n}}^{x}}E\left[K_{\mathbf{n}}(d(x,X_{\mathbf{k}}))\right]\right)^{\gamma'}\widehat{\mathbf{n}}\sum_{\|\mathbf{i}\|>m_{\mathbf{n}}}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}\\
&\leq C(1/S_{\mathbf{n}}^{x})^{-1+\gamma'}\left(\frac{1}{a_{\mathbf{n}}^{x}}E\left[K_{\mathbf{n}}(d(x,X_{\mathbf{k}}))\right]\right)^{\gamma'}\sum_{\|\mathbf{i}\|>m_{\mathbf{n}}}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}.\end{aligned}$$

Therefore

$$\|\mathbf{i}\|^{\nu}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}=\|\mathbf{i}\|^{\nu}o(\|\mathbf{i}\|^{-N(1-\gamma')/a})=o(\|\mathbf{i}\|^{-N-\varepsilon_{1}}),$$

and

$$\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\|\mathbf{i}\|^{\nu}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}<\infty.$$

Since $(1/S_{\mathbf{n}}^{x})^{-1+\gamma'}m_{n}^{-\nu}=1$, and

$$\left(\frac{1}{a_{\mathbf{n}}^{x}}E\left[K_{\mathbf{n}}(d(x,X_{\mathbf{k}}))\right]\right)^{\gamma'}=\left(\frac{1}{a_{\mathbf{n}}^{x}}\int_{\mathcal{E}}K_{\mathbf{n}}(d(x,t))f(t)d\mu(t)\right)^{\gamma'}\to(f(x))^{\gamma'}$$

(see Lemma 4) we obtain:

$$\begin{aligned}\limsup(\widehat{\mathbf{n}}/S_{\mathbf{n}}^{x})\tilde{I}_{2}(x)&\leq C\limsup\left(1/S_{\mathbf{n}}^{x}\right)^{-1+\gamma'}\left(\frac{1}{a_{\mathbf{n}}^{x}}E\left[K_{\mathbf{n}}(d(x,X_{\mathbf{k}}))\right]\right)^{\gamma'}\sum_{\|\mathbf{i}\|>m_{\mathbf{n}}}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}\\
&\leq C\limsup\left(1/S_{\mathbf{n}}^{x}\right)^{-1+\gamma'}m_{n}^{-\nu}\left(\frac{1}{a_{\mathbf{n}}^{x}}E\left[K_{\mathbf{n}}(d(x,X_{\mathbf{k}}))\right]\right)^{\gamma'}\sum_{\|\mathbf{i}\|>m_{\mathbf{n}}}\|\mathbf{i}\|^{\nu}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}\\
&\leq C\limsup\sum_{\|\mathbf{i}\|>m_{\mathbf{n}}}\|\mathbf{i}\|^{\nu}\left(\varphi(\|\mathbf{i}\|)\right)^{1-\gamma'}.\end{aligned}$$

This last term tends to a finite number when $m_{\mathbf{n}}$ tends to infinity. This yields the proof.

□

Proof. of Theorem 1

We have for all $x \in \mathcal{E}$

$$E_f(\hat{f}^1_{\mathbf{n}}(x) - f(x))^2 = V_f(\hat{f}^1_{\mathbf{n}}(x)) + b^2(\hat{f}^1_{\mathbf{n}}(x)), \tag{12}$$

where $b(\hat{f}^1_{\mathbf{n}}(x)) = E_f(\hat{f}^1_{\mathbf{n}}(x)) - f(x)$ is the bias of $\hat{f}^1_{\mathbf{n}}(x)$.

Let us first show the sufficient part of the theorem. When f is continuous at x and $r_{\mathbf{n}} \to 0$, we deduce from Lemma 3 that, $b(\hat{f}^1_{\mathbf{n}}(x)) \to 0$. The hypotheses (6), $\lim_{\mathbf{n}\to\infty} \widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}}) = \infty$, $r_{\mathbf{n}} \to 0$, $\theta > 2N$, the continuity of f at x and Lemma 5 imply that $V_f(\hat{f}^1_{\mathbf{n}}(x)) \to 0$.

Let us prove now that the conditions $\lim_{\mathbf{n}\to\infty} \widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}}) = \infty$ and $r_{\mathbf{n}} \to 0$ are necessary to mean square convergence of $\hat{f}^1_{\mathbf{n}}(x)$.

To this aim, let (as in chapter 4 of Dabo-Niang (2002)) $a \in \mathcal{E}$, $r > 0$, $C_1 > 0$, $C_2 > 0$, arbitrarily chosen, let us take the following density

$$\forall x \in \mathcal{E}, \;\; f(x) = C_1 1_{B^a_r}(x) + C_2 1_{(B^a_r)^c}(x),$$

where $(B^a_r)^c$ is the complementary of B^a_r and

$$C_1 \neq C_2\,,\; C_2 = \frac{1 - C_1\mu(B^a_r)}{\mu((B^a_r)^c)}.$$

Let us suppose that $\hat{f}^1_{\mathbf{n}}(x)$ converges to $f\,(x)$ in mean square, this is equivalent to, both the bias and the variance converge to zero.

Let us show by contradiction that $r_{\mathbf{n}} \to 0$. We know that $(r_{\mathbf{n}})_{\mathbf{n}\in(\mathbb{N}^*)^N}$ is a decreasing sequence, and more it is strictly positive, then its limit is finite. Let us suppose that, $r_{\mathbf{n}} \downarrow h > 0$, then we can deduce that, $B^x_{r_{\mathbf{n}}} \downarrow B^x_h$ as $\mathbf{n} \to \infty$, and we have that f is integrable with respect to μ, then we deduce that, (see Hoffmann-Jorgensen (1994), p168)

$$\int_{B^x_{r_{\mathbf{n}}}} f(y)d\mu(y) \underset{\mathbf{n}\to\infty}{\to} \int_{B^x_h} f(y)d\mu(y).$$

We can write now,

$$E_f\,(\hat{f}^1_{\mathbf{n}}(x)) = \frac{\int_{B^x_{r_{\mathbf{n}}}} f(y)d\mu(y)}{\mu(B^x_{r_{\mathbf{n}}})} \to \frac{\int_{B^x_h} f(y)d\mu(y)}{\mu(B^x_h)},$$

as $\mathbf{n} \to \infty$. Let $r = h$, and let us choose $x \in B^a_{h,o}$, where $B^a_{h,o}$ is the opened ball of center a and radius h, then f is continuous at x and $f(x) = C_1$, thus we write

$$\frac{\int_{B^x_h} f(y)d\mu(y)}{\mu(B^x_h)} = \frac{\int_{B^x_h \cap B^a_h} f(y)d\mu(y)}{\mu(B^x_h)} + \frac{\int_{B^x_h \cap (B^a_h)^c} f(y)d\mu(y)}{\mu(B^x_h)}.$$

Let, $A^x_{1,h} = B^x_h \cap B^a_h$ and $A^x_{2,h} = B^x_h \cap (B^a_h)^c$, then we get $\mu(B^x_h) = \mu(A^x_{1,h}) + \mu(A^x_{2,h})$,

$$\frac{\int_{B^x_h} f(y)d\mu(y)}{\mu(B^x_h)} = C_1 + \frac{\mu(A^x_{2,h})[1 - C_1.\mu(\mathcal{E})]}{\mu(B^x_h)\mu((B^a_r)^c)}.$$

But,

$$C_1 \neq \frac{1}{\mu(\mathcal{E})},$$

because $C_1 \neq C_2$. Finally we have, for all $x \in B^a_{h,o}$:

$f(x) = C_1 \neq \lim_{r_{\mathbf{n}} \to h} E_f\left(\hat{f}^1_{\mathbf{n}}(x)\right)$, and this is a contradiction with the fact that the bias converges to zero, thus

$$\lim_{\mathbf{n} \to +\infty} r_{\mathbf{n}} = 0.$$

Concerning the variance, let us recall that $\lim_{\mathbf{n} \to +\infty} \widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}})V_f\left(\hat{f}^1_{\mathbf{n}}(x)\right) < +\infty$ (see Lemma 5). Moreover, we have by hypothesis $V_f\left(\hat{f}^1_{\mathbf{n}}(x)\right) \to 0$ as $\mathbf{n} \to \infty$, this imply that $\lim_{\mathbf{n} \to \infty} \widehat{\mathbf{n}}\mu(B^x_{r_{\mathbf{n}}}) = \infty$. The proof is therefore complete.♦

□

Proof. of Theorem 2

It follows directly from Lemmas 4 and 5.

□

References

[1] Anselin, L. and Florax, R.J.G.M. (1995). *New Directions in Spatial Econometrics*, Springer, Berlin.

[2] Bensaid, N. and Dabo-Niang, S. (2006). Frequency polygone for continuously indexed spatial processes. *Preprint*. Submitted

[3] Biau, G. and Cadre, B. (2004). Nonparametric Spatial Prediction, *Stat. Inference Stochastic Proc*, **Vol. 7**, pp. 327-349.

[4] Biau, G. (2003). Spatial Kernel Density Estimation, *Mathematical Methods of Statistics*, **Vol. 12**, pp. 371-390.

[5] Billingsley P. (1968). *Convergence of Probability Measures*, Wiley Series In Probability And Mathematical Statistics.

[6] Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes - Estimation and Prediction - 2nd Edition*, Lecture Notes in Statistics, Springer-Verlag, New York.

[7] Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and applications.* Lecture Notes in Statistics, 149, Springer.

[8] Carbon, M. (2006). Polygon des fréquences pour des champs aléatoires. *Comptes. Rendus. Acad. Sci. Paris I,* **Vol. 342**, **Issue 9**, pp. 693-696.

[9] Carbon, M., Hallin, M. and Tran, L.T. (1996). Kernel density estimation for random fields, *Statistics and Probability Letters*, **Vol. 36**, pp. 115-125.

[10] Carbon, M., Tran, L.T. and Wu, B. (1997). Kernel density estimation for random fields: the L_1 theory, *J. Nonparametric Statistics*, **Vol. 6**. pp. 157-170.

[11] Carbon, M., Francq, C. and Tran, L.T. (2007). Kernel regression estimation for random fields. *Journal of Statistical Planning and Inference*, **Vol. 137**, **Issue 3**, pp 778-798.

[12] Cressie, N.A.C. (1991). *Statistics for Spatial Data*, Wiley Series in Probability and Mathematical Statistics, New York.

[13] Dabo-Niang, S (2002). Estimation de la densité dans un espace de dimension infinie: Applications aux diffusions, *C. R. Acad. Sci.* Paris I, 334, 213-216.

[14] Dabo-Niang, S (2002). *Sur l'estimation fonctionnelle en dimension finie. Application aux diffusions.* Thèse de Doctorat. Université Paris VI.

[15] Dabo-Niang, S (2004). Kernel density estimator in an infinite dimensional with a rate of convergence in the case of diffusion processes, *Applied Mathematic Letters*, **17**, 4, 381-386.

[16] Dabo-Niang, S and Yao, A. F. (2006). Density estimation for spatial functional random variables. *Preprint.*

[17] Dabo-Niang, S.; Ferraty, F. and Vieu, P. (2006). Mode estimation for functional random variable and its application for curves classification. *Far East Journal of Theoretical Statistics.* **Vol 18**, No. 1, pp 93 - 119.

[18] Dabo-Niang, S. and Yao, A.F. (2007). Kernel regression estimation for continuous spatial processes, *Mathematical Methods of Statistics,* **Vol 16, No 4**, 1-20.

[19] Doukhan, P. (1994). *Mixing - Properties and examples, Lecture Notes in Statistics*, Springer-Verlag, New-York.

[20] Ferraty, F. and Vieu, Ph., (2006), *Nonparametric functional data analysis.* Springer-Verlag, New York.

[21] Guyon, X. (1995). *Random Fields on a Network - Modelings, Statistics ans Applications*, Springer-Verlag, New York.

[22] Hallin, M., Lu, Z. and Tran, L.T. (2004). Kernel density estimation for spatial processes: the L_1 theory, *J. of Mult. Anal.*, **Vol. 88**, **No. 1**, pp. 61-75.

[23] Hoffmann-Jorgensen J. (1994). *Probability with a view Toward Statistics.* Chapman and Hall. New-York.

[24] Lu, Z., Lundervold, A., Tjostheim, D. and Yao, Q. (2007). Exploring spatial nonlineariry using additive approximation, *Bernoulli,* **13**, 447-472.

[25] Parzen, E. (1962). On Estimation of a Probability Density Function and Mode, *Ann. Math. Statist.* 33, 1065-1076.

[26] Ramsay, J.O. and Silverman, B.W. (1997). *Functional data analysis*, Springer Series in Statistics.

[27] Ramsay, J.O. and Silverman, B.W. (2002). *Applied functional data analysis; Methods and case studies.* Springer-Verlag, New York.

[28] Rio, E. (2000). *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants,* Springer-Verlag, Belin.

[29] Ripley, B.D. (1981). *Spatial Statistics*, Wiley Series in Probability and Mathematical Statistics, New York.

[30] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function, *Ann. Math. Statist.* 27, 832-837.

[31] Tran, L. T. (1990). Kernel Density Estimation on Random Fields, *Journal of Multivariate Analysis,* **Vol. 34**, pp. 37-53.

[32] Tran, L. T. and Yakowitz, S. (1993). Nearest neighbor estimators for random fields, *Journal of Multivariate Analysis,* **Vol. 44**, pp. 23-46.