Communications of the Korean Statistical Society
2012, Vol. 19, No. 1, 193–211
DOI: http://dx.doi.org/10.5351/CKSS.2012.19.1.193

Nonparametric M-Estimation for Functional Spatial Data

Mohammed Kadi Attouch[a], Benamar Chouaf[a], Ali Laksaci[1,a]

[a]Laboratoire de Mathématiques, University Djillali Liabès

Abstract

This paper deals with robust nonparametric regression analysis when the regressors are functional random fields. More precisely, we consider $Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}})$, $\mathbf{i} \in \mathbb{N}^N$ be a $\mathcal{F} \times \mathbb{R}$-valued measurable strictly stationary spatial process, where $\mathcal{F}$ is a semi-metric space and we study the spatial interaction of $X_{\mathbf{i}}$ and $Y_{\mathbf{i}}$ via the robust estimation for the regression function. We propose a family of robust nonparametric estimators for regression function based on the kernel method. The main result of this work is the establishment of the asymptotic normality of these estimators, under some general mixing and small ball probability conditions.

Keywords: Asymptotic distribution, spatial data, functional data, kernel estimate, nonparametric model, robust estimation, small balls probability.

1. Introduction

Denote the integer lattice points in the N-dimensional Euclidean space by $\mathbf{Z}^N$ and consider a strictly stationary functional random field $Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}})$, indexed by $\mathbf{Z}^N$ and defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Suppose that $Z_{\mathbf{i} \in \mathbf{Z}^N}$, takes values in $\mathcal{F} \times \mathbb{R}$. $\mathcal{F}$ is a semi-metric space which is endowed with a semi-metric $d(\cdot\,,\cdot)$. Our purpose is to study the spatial co-variation between ($X_{\mathbf{i}}$ and $Y_{\mathbf{i}}$) via the robust estimation of the regression function. This nonparametric model, denoted by θ_x, is implicitly defined as a zero with respect to (w.r.t.), t, of the equation

$$\Psi(x, t) := \mathbb{E}\left[\psi_x(Y_{\mathbf{i}}, t) \mid X_{\mathbf{i}} = x\right] = 0,$$

where ψ_x is a real-valued Borel function satisfying some regularity conditions to be stated below. We suppose that, for all $x \in \mathcal{F}$, θ_x exists and is unique (see, for instance, Boente and Fraiman (1989)).

Noting that, a robust regression is an important analysis tool in statistics. It is used to circumvent some limitations of a classical regression, namely, when the data are heteroscedastic or contain outliers. Moreover, in this area of functional spatial statistic, the data are collected in spatial order with the grid of the measurement fairly finer which makes it bulky and with a strong probability to be affected by the presence of outliers. Although, a modern technology and advanced computing environments have facilitated the collection and analysis of such data; however, the disinfection of the outliers is an important step to highlight the features of any data set. Therefore, there is a real necessity in this area to develop some procedures insensitive to deviations due to the presence of atypical or heteroskedatic observations. This is the main motivation of this work. It should be noted that, there is an increasing number of situations coming from different fields of applied sciences in which the data are of functional nature and show spatial interaction (such as soil science, geology,

This work is Supported by "Agence Nationale de Recherche Universitaire (A.N.D.R.U.)", Algérie.
[1] Corresponding author: Professor, Laboratoire de Mathématiques, University Djillali Liabès, Sidi Bel Abbès, BP 89, Sidi Bel Abbès 22000, Algérie. E-mail: alilak@yahoo.fr

oceanography, econometrics, epidemiology, environmental science, and forestry). In the finite dimensional framework, the statistic modelization of spatial data had received significant attention, see for example, Guyon (1995), Anselin and Florax (1995), Cressie (1991) or Ripley (1981). However, the nonparametric treatment of such data has only really been developed in the last two decades. Key references on this subject are Tran (1990), Biau and Cadre (2004), Dabo-Niang and Yao (2007), Carbon *et al.* (2007) and Li and Tran (2009). Recently, the spatial data has also modeled using some robust procedure see, for instance Xu and Wang (2008), Hallin *et al.* (2009) and Gheriballah *et al.* (2010).

Since the monographies by Ramsay and Silverman (2002, 2005), Bosq (2000) and Ferraty and Vieu (2006) a vast literature has been developed in functional data analysis. Nonparametric robust estimation for functional data had also received a great attention in this field. The first results in this topic were introduced by Cadre (2001). He studied the median estimation (without conditioning) of the distribution of a random variable taking its values in a Banach space. Azzedine *et al.* (2008) adapts to the functional data the local M-estimator of Collomb and Härdle (1986). They established the almost complete convergence of the adapted estimate in the *i.i.d.* case. The asymptotic normality of this latter has been studied by Attouch *et al.* (2009). The robust analysis in functional time series, which is a particular case of functional random filed ($N = 1$), has been investigated by many authors. We cite, for example, Crambes *et al.* (2008) for the convergence in L^q norm, Attouch *et al.* (2010) for the asymptotic normality and Chen and Zhang (2009) for the weak and strong consistency of the nonparametric functional conditional location estimate in the mixing case.

This paper studies the asymptotic behavior of a robust smoother estimate for functional spatial regression. This estimate is constructed by combining the ideas of robustness with those of smoothed regression that allows us to obtain reliable estimations when outlier observations are present in the responses. The main result of this work is that, under general mixing assumptions, the estimator considered is asymptotically normally distributed. Notice that, this work extends to spatial case the results given by Attouch *et al.* (2010) in functional time series data.

The paper is organized as follows: the next Section is dedicated to fixing notations and hypotheses. We state our main result and we list some preliminary results for its proof in Section 3. In Section 4 we discuss the impact of our results compared with those obtained in multivariate case or functional time series cases and we present some applications of the asymptotic normality property. All proofs are given in the appendix.

2. The Spatial Estimate

In the remainder of the paper, we suppose that $(Z_\mathbf{i})$ is observed over a rectangular domain $I_\mathbf{n} = \{\mathbf{i} = (i_1, \dots, i_N) \in \mathbb{Z}^N, 1 \le i_k \le n_k, k = 1, \dots, N\}$, $\mathbf{n} = (n_1, \dots, n_N) \in \mathbb{Z}^N$. A point $\mathbf{i}$ will be referred to as a *site*. We will write $\mathbf{n} \to \infty$ if $\min\{n_k\} \to \infty$ and $|n_j/n_k| < C$ for a constant C such that $0 < C < \infty$ for all j, k such that $1 \le j, k \le N$. For $\mathbf{n} = (n_1, \dots, n_N) \in \mathbb{Z}^N$, we set $\widehat{\mathbf{n}} = n_1 \times \cdots \times n_N$ and $\mathbf{1} = (1, \dots, 1) \in \mathbb{Z}^N$. We consider a spatial kernel estimate of $\Psi(x, t)$, denoted by $\widehat{\Psi}(x, t)$, defined by

$$\widehat{\Psi}(x, t) := \frac{\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K(h^{-1}d(x, X_\mathbf{i}))\psi(Y_\mathbf{i}, t)}{\sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} K(h^{-1}d(x, X_\mathbf{i}))}, \tag{2.1}$$

where K is a kernel and $h = h_{K,\mathbf{n}}$ is a sequence of positive real numbers. A natural estimator $\widehat{\theta_x}$ of θ_x, is a zero w.r.t. t of the equation

$$\widehat{\Psi}(x, t) = 0. \tag{2.2}$$

Our main goal is to study the asymptotic normality of the nonparametric estimate $\widehat{\theta_x}$ of θ_x when the random filed $(Z_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^N)$ satisfies the following mixing condition:

$$\begin{cases} \text{There exists a function } \varphi(t) \downarrow 0 \text{ as } t \to \infty, \text{ such that} \\ \forall\, E,\ E' \text{ subsets of } \mathbb{N}^N \text{ with finite cardinals} \\ \alpha\left(\mathcal{B}(E), \mathcal{B}\left(E'\right)\right) = \sup_{B\in\mathcal{B}(E),\, C\in\mathcal{B}(E')} \left|\mathbb{P}(B\cap C) - \mathbb{P}(B)\,\mathbb{P}(C)\right| \\ \qquad \leq s\left(\operatorname{Card}(E), \operatorname{Card}\left(E'\right)\right)\varphi\left(\operatorname{dist}\left(E, E'\right)\right), \end{cases} \tag{2.3}$$

where $\mathcal{B}(E)$(resp. $\mathcal{B}(E')$) denotes the Borel σ-field generated by $(Z_{\mathbf{i}}, \mathbf{i} \in E)$ (resp. $(Z_{\mathbf{i}}, \mathbf{i} \in E')$), Card$(E)$ (resp. Card(E')) the cardinality of E (resp. E'), dist(E, E') the Euclidean distance between E and E' and $s : \mathbb{N}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable such that:

$$s(n, m) \leq C \min(n, m), \quad \forall\, n, m \in \mathbb{N}. \tag{2.4}$$

We also assume that the process satisfies the following mixing condition:

$$\sum_{i=1}^{\infty} i^{\delta}\varphi(i) < \infty, \quad \delta > 0. \tag{2.5}$$

Noting that condition (2.4) and (2.5) are used in Tran (1990), Carbon *et al.* (1996) and are satisfied by many spatial models (see Guyon (1987) for some examples). It should be noted that if $N = 1$, then $Z_{\mathbf{i}}$ is called strongly mixing (see Doukhan *et al.* (1994) for discussion on mixing and examples).

3. Notations and Hypotheses

All along the paper, when no confusion is possible, we will denote by C and C' some strictly positive generic constants, x is a fixed point in $\mathcal{F}$ and $\mathcal{N}_x$ denote a fixed neighborhood of x. Moreover, for all $\mathbf{i} \in \mathcal{I}_{\mathbf{n}}$, we put $K_{\mathbf{i}}(x) = K(h^{-1}d(x, X_{\mathbf{i}}))$ and we pose $\widehat{\Psi}(x,t) = \widehat{\Psi}_N(x,t)/\widehat{\Psi}_D(x)$ with

$$\widehat{\Psi}_D(x) = \frac{1}{\hat{\mathbf{n}}\mathbb{E}[K_{\mathbf{1}}(x)]}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}(x) \quad \text{and} \quad \widehat{\Psi}_N(x,t) = \frac{1}{\hat{\mathbf{n}}\mathbb{E}[K_{\mathbf{1}}(x)]}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} K_{\mathbf{i}}(x)\psi_x(Y_{\mathbf{i}}, t).$$

In order to establish our asymptotic results we need the following hypotheses:

(H1) $\forall\, r > 0, \mathbb{P}(X \in B(x, r)) =: \phi_x(r) > 0$, where $B(x,r) = \{x' \in \mathcal{F}/d(x,x') < r\}$.

(H2) $\forall\, \mathbf{i} \neq \mathbf{j}$,

$$0 < \sup_{\mathbf{i}\neq\mathbf{j}} \mathbb{P}\left[(X_{\mathbf{i}}, X_{\mathbf{j}}) \in B(x,h) \times B(x,h)\right] \leq C(\phi_x(h))^{\frac{a+1}{a}}, \quad \text{for some } 1 < a < \delta N^{-1}.$$

(H3) The function Ψ is such that:

(i) The function $\Psi(x, \cdot)$ is of class C^1 on $[\theta(x) - \delta, \theta(x) + \delta]$, $\delta > 0$,

(ii) For each fixed $t \in [\theta(x) - \delta, \theta(x) + \delta]$, the function $\Psi(\cdot, t)$ satisfies Hölder's condition w.r.t. the first one, that is: there exist strictly positives constants b such that $\forall\, x_1, x_2 \in \mathcal{N}_x$, $|\Psi(x_1, t) - \Psi(x_2, t)| \leq Cd^b(x_1, x_2)$ where $\mathcal{N}_x$ is a fixed neighborhood of x.

(H4) ψ_x is strictly monotone function and bounded.

(H5) The bandwidth h satisfies:

$$h \downarrow 0, \ \forall t \in [0,1] \ \lim_{h\to 0} \frac{\phi_x(th)}{\phi_x(h)} = \beta_x(t) \quad \text{and} \quad n\phi_x(h) \to \infty, \quad \text{as } n \to \infty.$$

(H6) The kernel K from $\mathbb{R}$ into $\mathbb{R}^+$ is a differentiable function supported on $[0,1]$. Its derivative $K^{'}$ exists and is such that there exist two constants C_3 and C_4 with $-\infty < C_3 < K^{'}(t) < C_4 < 0$ for $0 \le t \le 1$.

(H7) There exists $\eta_0 \in \left]\frac{1}{1+N+\delta}, \frac{1}{1+2N}\right[$, such that $\hat{\mathbf{n}}^{-\frac{1}{1+2N}+\eta_0} \le \phi_x(h)$.

Remarks on the assumptions: As all the asymptotic results in nonparametric statistics for functional variables are closely related to the concentration properties of the probability measure of the functional variable X. The same thing here, this property quantified by mean of the function $\phi_x(\cdot)$ defined in condition (H1). Such function can be explicated for several continuous processes (see Ferraty *et al.*, 2006). Condition (H2) measure the local dependence of the observations. As usually in nonparametric problems, the infinite dimension of the model is controlled by mean of a smoothness condition (H3). This condition is needed to evaluate the bias component of the rates of convergence. Condition (H4) controls the robustness properties of our model. More precisely, we consider the robustification given by Collomb and Härdle (1986) in the multivariate case, where the score function ψ_x is indexed by x. This permits is to include, for instance, the functional nonparametric regression model the scale of the error is assumed to be known, where $\psi_x(\cdot) = \psi(\cdot/\sigma(x))$, with $\sigma(\cdot)$ is measure of spread for the conditional distribution of Y given $X = x$. We point out that the boundedness hypotheses over $\psi_x(\cdot)$ can be dropped by using the truncation method as in Laïb and Ould-Saïd (2000). However the boundedness of the score function is fundamental constraint of the robustness properties of the M-estimators. Assumptions (H5)~(H7) are standard condition for obtaining the normality asymptotic in kernel estimate for functional statistic.

Theorem 1. *Assume that (H1)~(H7) hold and if the bandwidth parameter h satisfies* $\hat{\mathbf{n}}h^{2b_1}\phi_x(h) \to 0$ *as* $\mathbf{n} \to \infty$*, then we have for any* $x \in \mathcal{A}$*,*

$$\left(\frac{\hat{\mathbf{n}}\phi_x(h)}{\sigma^2(x,\theta(x))}\right)^{\frac{1}{2}} \left(\widehat{\theta(x)} - \theta(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \textit{as } \mathbf{n} \to \infty.$$

where

$$\sigma^2(x,\theta_x) = \frac{\beta_2 \mathbb{E}\left[\psi_x^2(Y,\theta_x)|X=x\right]}{\left(\beta_1 \frac{\partial}{\partial t}\Psi(x,\theta(x))\right)^2} \quad \left(\textit{with } \beta_j = -\int_0^1 (K^j)^{'}(s)\beta_x(s)ds, \textit{ for } j = 1, 2\right),$$

$$\mathcal{A} = \left\{x \in \mathcal{F}, \ \mathbb{E}\left[\psi_x^2(Y,\theta_x)|X=x\right] \frac{\partial}{\partial t}\Psi(x,\theta(x)) \neq 0\right\}$$

and $\xrightarrow{\mathcal{D}}$ *means the convergence in distribution.*

Proof: We give the proof for the case of a increasing ψ_x, decreasing case being obtained by considering $-\psi_x$. In this case, we define, for all $u \in \mathbb{R}$, $z = \theta(x) + u[\hat{\mathbf{n}}\phi_x(h)]^{-1/2}\sigma(x,\theta(x))$. Let us remark that, if $\widehat{\Psi}_D(x)$ is not equal to zero, the definition of the estimator by (2.2) is equivalent to

$$\widehat{\Psi}_N\left(x, \widehat{\theta(x)}\right) = 0.$$

Hence, if $\widehat{\Psi}_D(x) \neq 0$ we can write

$$\begin{aligned}\mathbb{P}\left\{\left(\frac{\hat{\mathbf{n}}\phi_x(h)}{\sigma^2(x,\theta(x))}\right)^{\frac{1}{2}}\left(\widehat{\theta(x)} - \theta(x)\right) < u\right\} &= \mathbb{P}\left\{\widehat{\theta(x)} < \theta(x) + u\left[\hat{\mathbf{n}}\phi_x(h)\right]^{-\frac{1}{2}}\sigma(x,\theta(x))\right\} \\ &= \mathbb{P}\left\{0 < \widehat{\Psi}_N(x,z)\right\} \\ &= \mathbb{P}\left\{\mathbb{E}\left[\widehat{\Psi}_N(x,z)\right] - \widehat{\Psi}_N(x,z) < \mathbb{E}\left[\widehat{\Psi}_N(x,z)\right]\right\}.\end{aligned}$$

Therefore, Theorem 1 is a consequence of the following intermediates results, where their proofs are postponed to the appendix. □

Lemma 1. *Under Hypotheses (H1)~(H3) and (H6) , we have*

$$\mathbb{P}\left\{\left(\widehat{\Psi}_D(x) = 0\right)\right\} \longrightarrow 0, \quad \text{as } \mathbf{n} \longrightarrow \infty.$$

Lemma 2. *Under the Hypotheses of Theorem 1, we have for any $x \in \mathcal{A}$*

$$\left(\frac{\hat{\mathbf{n}}\phi_x(h)}{\left(\frac{\partial}{\partial t}\Psi(x,\theta(x))\right)^2\sigma^2(x,\theta(x))}\right)^{\frac{1}{2}}\left(\widehat{\Psi}_N(x,z) - \mathbb{E}\left[\widehat{\Psi}_N(x,z)\right]\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } \mathbf{n} \to \infty.$$

Lemma 3. *Under hypotheses (H1′), (H3), (H6) and if the bandwidth parameter h satisfies $\hat{\mathbf{n}}h^{2b_1}\phi_x(h) \to 0$ as $\mathbf{n} \to \infty$, we have*

$$\left(\frac{\hat{\mathbf{n}}\phi_x(h)}{\left(\frac{\partial}{\partial t}\Psi(x,\theta(x))\right)^2\sigma^2(x,\theta(x))}\right)^{\frac{1}{2}}\mathbb{E}\left[\widehat{\Psi}_N(x,z)\right] = u + o(1), \quad \text{as } \mathbf{n} \to \infty.$$

4. Discussion

- *On the nonparametric robust analysis of functional spatial data:* As indicated in the introduction, the general framework of this paper is the robust modeling of functional data presenting spatial dependence. Noting that the modelization of this kind of data has been selected by Ramsay (2008) among the eight most interesting research subject in functional data analysis. This great consideration is motivated by the increasing number of situations coming from different fields of applied sciences for which the data are of functional nature and showing a spatial interaction. Typically, the continuously indexed random filed, the spatio-temporal process and the point process are the most important example of functional spatial data (see, Delicado *et al.*, 2010). Of course, these process can be analyzed through the use of the multivariate approach, where, we assume that the underlying process ($Z_{t\in\mathbb{R}^N}$) is observed on some discrete grid. We point out that there are various sampling designs can be employed; however, two kinds of these sampling are most useful: deterministic (points are chosen according to a deterministic rule, for example, periodic sampling) and random (points are chosen randomly, for example, Poisson sampling). In the nonparametric prediction context, this multivariate approach has been considered by many authors in the past (see, for instance, Biau (2003) and Gheriballah *et al.* (2010) for the robust case). However, this approach suffers from the curse of dimensionality, if the number of sampling location is large. Moreover, with this transformation we lose several characteristics of the original data (the correlation data, the functional nature of data, homoscedasticity or the heteroscedasticity of the data). Specifically, the variability of the transformed data is linked to the used sampling; subsequently, this transformation may generate ghosts outliers and hide the true outliers. These defects can seriously distort the use of the multivariate methodology in spatial continuously indexed process, where it is necessary to pay more attention to local differences among spatial neighborhood and to integrate the spatial properties to outlierness measurement (see, Lu *et al.*, 2003). Thus, in this area of functional spatial data, the multivariate approach is very limited and it is inadequate to analyze such type of process. So, we can say that the current work is not a simple adaptation, to the functional case, of the multivariate study of Gheriballah *et al.* (2010), but, it is a structural development allows us to avoid the influence of all these defects by keeping the functional feature of the original data. Recently, the functional spatial data has been modeled by two fundamentals regression analysis tools such the cokriging method (see, Nerini *et al.*, 2010) and the nonparametric regression (see, Dabo-Niang and Thiam, 2010) but, it is well known that our robust method has more advantages than the above cited methods by its outlier-resistance properties. In conclusion, we can say that the nonparametric robust analysis in functional spatial data is an important analysis tool and has great impact in practice. Such approach is more adapted than, the multivariate one, for the data collected with continuous monitoring in spatial order and it is better than the nonparametric regression for the functional spatial data affected by some outliers. It should be noted that most of these case can be treated as particular cases of our study. This generality of our model is discussed in the following Section by giving the formulation of our result for some special case.

- Some applications

  - *Application to continuous random fields:* As noticed above, the continuously indexed random filed is the most important example of functional spatial data. Indeed, let $(Z_t)_{t\in\mathbb{R}^N}$ be an $\mathbb{R}$-valued strictly stationary random spatial processes assumed to be bounded and observed over some subset $I \subset \mathbb{R}^N$. Our approach can be used to predict the value Z_{s_0} at an unobserved location $s_0 \notin I$ by taking into account, the observed part of $(Z_{t\in I})$ in its continuous form. For this, we suppose that, the value of Z_{s_0} depends only on the values of the process (Z_t) in a bounded neighborhood $\mathcal{V}_{s_0} \subset I$ of s_0. From Z_t we may construct m functional spatial random variables as follows: Consider some grid $\mathcal{G}_{\mathbf{n}} = \{t_i = (t_{i,1}, \dots t_{i,N}) \in I,\ 1 \le t_{i,j} \le n_j,\ j = 1, \dots, N,\ i = 1, \dots, m\}$ such that

$$\forall\, i = 1, \dots m \quad \min_{1\le j\le N-1}(t_{i,j+1} - t_{i,j}) \ge C > 0, \quad \text{for some constant } C$$

  and we define

$$\forall\, i = 1, \dots, m, \quad X_{t_i} = (Z_t,\ t \in \mathcal{V}_{t_i}),$$

  where $\mathcal{V}_{t_i} = \mathcal{V} + t_{t_i}$ with $\mathcal{V} = \mathcal{V}_{s_0} - s_0$. So, the predictor that we proposed (see Biau and Cadre (2004), Dabo-Niang and Yao (2007) for the finite dimension mean regression case), aims to evaluate a real characteristic denoted $Y_{s_0} = Z_{s_0}$, at a site s_0, given $X_{s_0} = (Z_t,\ t \in \mathcal{V}_{s_0})$. The random variable $\hat{\theta}(X_{s_0})$, is is the best approximation of the quantity Y_{s_0}, with respect to the loss function $\rho_x(t) = \int_0^t \psi_x(s)ds$. The latter is given by using the m pairs of r.v $(X_{t_i}, Y_{t_i})_{t_i}$ with $Y_{t_i} = Z_{t_i}$.

- *Application to conditional confidence curve:* An important application of the asymptotic normality result is the building of confidence intervals for the true value of θ given curve $X = x$. A plug-in estimate for the asymptotic standard deviation $\sigma(x, \theta(x))$ can be obtained using the estimators $\widehat{\lambda_2}(x, \widehat{\theta(x)})$ and $\widehat{\Gamma_1}(x, \widehat{\theta(x)})$ of $\lambda_2(x, \theta(x))$ and $\Gamma_1(x, \theta(x))$ respectively. We get $\hat{\sigma}(x, \widehat{\theta(x)}) := [\widehat{\beta_2}\widehat{\lambda_2}(x, \widehat{\theta(x)})/\{(\widehat{\beta_1})^2\widehat{\Gamma_1}^2(x, \widehat{\theta(x)})\}]^{1/2}$. Then $\hat{\sigma}(x, \widehat{\theta(x)})$ can be used to get the following approximate $(1-\zeta)$ confidence interval for $\theta(x)$

$$\widehat{\theta(x)} \pm t_{1-\frac{\zeta}{2}} \times \left(\frac{\hat{\sigma}^2\left(x, \widehat{\theta(x)}\right)}{n\phi_x(h)} \right)^{\frac{1}{2}},$$

where $t_{1-\zeta/2}$ denotes the $1-\zeta/2$ quantile of the standard normal law.

Here we point out that the estimators $\widehat{\lambda_2}(x, \widehat{\theta(x)})$ and $\widehat{\Gamma_1}(x, \widehat{\theta(x)})$ are calculated, for $x \in \mathcal{A}$, in the same way as in (2.2). We estimate β_1 and β_2 empirically by

$$\widehat{\beta_1} = \frac{1}{\hat{\mathbf{n}}\phi_x(h)} \sum_{\mathbf{i} \in \mathcal{I}_\mathbf{n}} K_\mathbf{i}(x) \quad \text{and} \quad \widehat{\beta_2} = \frac{1}{\hat{\mathbf{n}}\phi_x(h)} \sum_{\mathbf{i} \in \mathcal{I}_\mathbf{n}} K_\mathbf{i}^2(x),$$

where $K_i(x)$ are defined as before.

This last estimation is justified because under (H1), (H5) and (H6), we have

$$\frac{1}{\phi_x(h)} \mathbb{E}\left[K_1^j\right] \to \beta_j, \quad j = 1, 2.$$

Note that the function $\phi_x(\cdot)$ does not appear in the calculation of the confidence interval by simplification. Finally, the approximate $(1-\zeta/2)$ confidence interval, for any $x \in \mathcal{A}$, is

$$[a_-(x), a_+(x)], \quad \text{where } a_\pm(x) = \widehat{\theta(x)} \pm t_{1-\frac{\zeta}{2}} \times \left(\frac{\sum_{\mathbf{i} \in \mathcal{I}_\mathbf{n}} K_\mathbf{i}^2(x)\widehat{\lambda_2}\left(x, \widehat{\theta(x)}\right)}{\left(\sum_{\mathbf{i} \in \mathcal{I}_\mathbf{n}} K_\mathbf{i}(x)\right)^2 \widehat{\Gamma_1}^2\left(x, \widehat{\theta(x)}\right)} \right)^{\frac{1}{2}}.$$

- *Some particular cases:*

  - *The multivariate case:* In the vectorial case, when $\mathcal{F} = \mathbb{R}^p$, $p \geq 1$ and if the probability density of the random variable X (resp. the jointly density of (X_i, X_j)) denoted by f (resp. by $f_{i,j}$), is of C^1 class, then $\phi_x(h) = O(h^p)$. Then, the conditions (H1), (H2) and (H4) are trivially verified with $\beta_x(t) = f(x)t^p$. Thus, our Theorem leads systematically to the next Corollary,

Corollary 2. *Under assumptions (H3), (H5)~(H7), we have:*

$$\left(\frac{\hat{\mathbf{n}}h^p}{\sigma^2(x, \theta(x))} \right)^{\frac{1}{2}} \left(\widehat{\theta(x)} - \theta(x) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \quad \text{as } \mathbf{n} \to \infty.$$

We point out that this result is exactly what is obtained by Gheriballah *et al.* (2010)

- *The functional time series case:* In this situation ($N = 1$), the conditions (2.4) is automatically verified. Therefore, we obtain the following result which is exactly the same as obtained by Attouch *et al.* (2010)

Corollary 3. *Under assumptions (H1)~(H7), we have:*

$$\left(\frac{n\phi_x(h)}{\sigma^2(x,\theta(x))}\right)^{\frac{1}{2}}\left(\widehat{\theta(x)}-\theta(x)\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1),\quad as\ n\to\infty.$$

5. Appendix

Proof of results

We first state the following lemmas which can be found in Tran (1990). They are needed for the strong convergence of our estimates. There proofs will then be omitted.

Lemma 4.

(i) Suppose that (2.3) holds. Denote by $\mathcal{L}_r(\mathcal{F})$ the class of $\mathcal{F}$-measurable r.v.'s X satisfying $\|X\|_r = (\mathbb{E}|X|^r)^{1/r} < \infty$. Suppose $X \in \mathcal{L}_r(\mathcal{B}(E))$ and $X \in \mathcal{L}_r(\mathcal{B}(E'))$. Assume also that $1 \leq r, s, t < \infty$ and $r^{-1} + s^{-1} + t^{-1} = 1$. Then

$$|\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]| \leq C\|X\|_r\|Y\|_s\left\{s(Card(E),Card(E'))\varphi(dist(E,E'))\right\}^{\frac{1}{t}}. \qquad (5.1)$$

(ii) For r.v.'s bounded with probability 1, the right-hand side of (5.1) can be replaced by $Cs(Card(E), Card(E'))\varphi(dist(E,E'))$.

We need, also, the following lemma.

Lemma 5. *Let $Z_1,\ldots,Z_n$ be a random vector such that $|\mathbb{E}\prod_{s=i}^{n} Z_s| < \infty$, $i = 1,\ldots,n-1$, $|Z_i| \leq C$, $i = 1,\ldots,n$. Then*

$$\left|\mathbb{E}\prod_{s=1}^{n}Z_s-\prod_{s=1}^{n}\mathbb{E}Z_s\right| \leq \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left|\mathbb{E}(Z_i-1)(Z_j-1)\prod_{s=j+1}^{n}Z_s-\mathbb{E}(Z_i-1)\mathbb{E}(Z_j-1)\prod_{s=j+1}^{n}Z_s\right|.$$

Proof: **(Lemma 1)** Clearly, for all $\varepsilon < 1$, we have

$$\mathbb{P}\left\{\widehat{\Psi}_D(x)=0\right\} \leq \mathbb{P}\left\{\widehat{\Psi}_D(x)\leq 1-\varepsilon\right\} \leq \mathbb{P}\left\{\left|\widehat{\Psi}_D(x)-1\right|\geq\varepsilon\right\}.$$

It now suffices to show that

$$\widehat{\Psi}_D(x)-1\to 0\quad\text{in probability.} \qquad (5.2)$$

Because $\mathbb{E}[\widehat{\Psi}_D(x)] = 1$, all it remains to show that the variance term tends to 0. For this, we write

$$\widehat{\Psi}_D(x)-\mathbb{E}\left[\widehat{\Psi}_D(x)\right]=\frac{1}{\hat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i}}(x)$$

and

$$\operatorname{Var}\left[\widehat{\Psi}_D(x)\right] = \operatorname{Var}\left[\frac{1}{\hat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i}}(x)\right] = \frac{1}{\hat{\mathbf{n}}^2\mathbb{E}^2\left[K_{\mathbf{1}}(x)\right]}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}\left|\operatorname{Cov}(\Delta_{\mathbf{i}}(x),\Delta_{\mathbf{j}}(x))\right|.$$

Let $Q_{\mathbf{n}} = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\operatorname{Var}[\Delta_{\mathbf{i}}(x)]$ and $R_{\mathbf{n}} = \sum_{\mathbf{i}\neq\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}|\operatorname{Cov}(\Delta_{\mathbf{i}}(x),\Delta_{\mathbf{j}}(x))|$. By Assumptions (H1) and (H2), we have

$$\operatorname{Var}[\Delta_{\mathbf{i}}(x)] \leq C\left(\phi_x(h) + (\phi_x(h))^2\right),$$

therefore

$$Q_{\mathbf{n}} = O\left(\hat{\mathbf{n}}\phi_x(h)\right).$$

Concerning $R_{\mathbf{n}}$ we introduce the following sets:

$$S_1 = \{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}} : 0 < \|\mathbf{i}-\mathbf{j}\| \leq c_{\mathbf{n}}\}, \qquad S_2 = \{\mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}} : \|\mathbf{i}-\mathbf{j}\| > c_{\mathbf{n}}\},$$

where $c_{\mathbf{n}}$ is a real sequence that converges to $+\infty$ and will be precise after. Split this sum into two separate summations over sites in S_1 and S_2

$$\begin{aligned} R_{\mathbf{n}} &= \sum_{(\mathbf{i},\mathbf{j})\in S_1}\left|\operatorname{Cov}\left(\Delta_{\mathbf{i}}(x),\Delta_{\mathbf{j}}(x)\right)\right| + \sum_{(\mathbf{i},\mathbf{j})\in S_2}\left|\operatorname{Cov}\left(\Delta_{\mathbf{i}}(x),\Delta_{\mathbf{j}}(x)\right)\right| \\ &= R_{\mathbf{n}}^1 + R_{\mathbf{n}}^2. \end{aligned}$$

On one hand, we have:

$$\begin{aligned} R_{\mathbf{n}}^1 &= \sum_{(\mathbf{i},\mathbf{j})\in S_1}\left|\mathbb{E}\left[K_{\mathbf{i}}K_{\mathbf{j}}\right] - \mathbb{E}\left[K_{\mathbf{i}}\right]\mathbb{E}\left[K_{\mathbf{j}}\right]\right| \\ &\leq C\hat{\mathbf{n}}c_{\mathbf{n}}^N\phi_x(h)\left((\phi_x(h))^{\frac{1}{a}} + \phi_x(h)\right) \\ &\leq C\hat{\mathbf{n}}c_{\mathbf{n}}^N\phi_x(h)^{\frac{a+1}{a}}. \end{aligned}$$

On the other hand, we have

$$R_{\mathbf{n}}^2 = \sum_{(\mathbf{i},\mathbf{j})\in S_2}\left|\operatorname{Cov}\left(\Delta\mathbf{i},\Delta\mathbf{j}\right)\right|.$$

As the random variables $K_{\mathbf{j}}$ are bounded, we deduce from Lemma 4 (ii) that

$$\left|\operatorname{Cov}\left(\Delta\mathbf{i},\Delta\mathbf{j}\right)\right| \leq C\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right),$$

thus

$$\begin{aligned} R_{\mathbf{n}}^2 &\leq C\sum_{(\mathbf{i},\mathbf{j})\in S_2}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right) \leq C\hat{\mathbf{n}}\sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}}\varphi\left(\|\mathbf{i}\|\right) \\ &\leq C\hat{\mathbf{n}}c_{\mathbf{n}}^{-Na}\sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}}\|\mathbf{i}\|^{Na}\varphi\left(\|\mathbf{i}\|\right). \end{aligned}$$

Let $c_{\mathbf{n}} = (\phi_x(h))^{-1/Na}$, then we have

$$R_{\mathbf{n}}^2 \leq C\hat{\mathbf{n}}c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na}\varphi(\|\mathbf{i}\|)$$
$$\leq C\hat{\mathbf{n}}\phi_x(h) \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na}\varphi(\|\mathbf{i}\|).$$

Because of (2.5) and (H2) we get

$$R_{\mathbf{n}}^2 \leq C\hat{\mathbf{n}}\phi_x(h).$$

Furthermore, under this choose of $c_{\mathbf{n}}$ we have

$$R_{\mathbf{n}}^1 \leq C\hat{\mathbf{n}}\phi_x(h).$$

Hence

$$\text{Var}\left[\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i}}(x)\right] = O(\hat{\mathbf{n}}\phi_x(h))$$

which imply that

$$\text{Var}\left[\widehat{\Psi}_D(x)\right] \to 0.$$

□

Proof: **(Lemma 2)**

It is easy to see that

$$\sqrt{\hat{\mathbf{n}}\phi_x(h)}\left(\widehat{\Psi}_N(x,z) - \mathbb{E}\left[\widehat{\Psi}_N(x,z)\right]\right) = \frac{1}{\sqrt{\hat{\mathbf{n}}}}\sum_{\mathbf{i}\in I_{\mathbf{n}}} \Lambda_{\mathbf{i}},$$

where

$$\Lambda_{\mathbf{i}} = \frac{\sqrt{\phi_x(h)}}{\mathbb{E}[K_{\mathbf{1}}]}\left\{K_{\mathbf{i}}\psi_x(Y_{\mathbf{i}},z) - \mathbb{E}\left[K_{\mathbf{i}}\psi_x(Y_{\mathbf{i}},z)\right]\right\}.$$

Thus, the asymptotic normality of $(\sqrt{\hat{\mathbf{n}}}\sigma(x,\theta(x)))^{-1}\sum_{\mathbf{i}\in I_{\mathbf{n}}}\Lambda_{\mathbf{i}}$ is sufficient to show the proof a Lemma 2. This last is shown by the blocking method, where the random variables $\Lambda_{\mathbf{j}}$ are grouped into blocks of different sizes defined by

$$W(1,\mathbf{n},\mathbf{j}) = \sum_{\substack{i_k=j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+1,\\ k=1,\dots,N}}^{j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}} \Lambda_{\mathbf{i}},$$
$$W(2,\mathbf{n},\mathbf{j}) = \sum_{\substack{i_k=j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+1,\\ k=1,\dots,N-1}}^{j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}} \sum_{i_N=j_N(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}+1}^{(j_N+1)(p_{\mathbf{n}}+q_{\mathbf{n}})} \Lambda_{\mathbf{i}},$$
$$W(3,\mathbf{n},\mathbf{j}) = \sum_{\substack{i_k=j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+1,\\ k=1,\dots,N-2}}^{j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}} \sum_{i_{N-1}=j_{N-1}(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}+1}^{(j_{N-1}+1)(p_{\mathbf{n}}+q_{\mathbf{n}})} \sum_{i_N=j_N(p_{\mathbf{n}}+q_{\mathbf{n}})+1}^{j_N(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}} \Lambda_{\mathbf{i}},$$
$$W(4,\mathbf{n},\mathbf{j}) = \sum_{\substack{i_k=j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+1,\\ k=1,\dots,N-2}}^{j_k(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}} \sum_{i_{N-1}=j_{N-1}(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}+1}^{(j_{N-1}+1)(p_{\mathbf{n}}+q_{\mathbf{n}})} \sum_{i_N=j_N(p_{\mathbf{n}}+q_{\mathbf{n}})+p_{\mathbf{n}}+1}^{(j_N+1)(p_{\mathbf{n}}+q_{\mathbf{n}})} \Lambda_{\mathbf{i}},$$

and so on. The last two terms are

$$W\left(2^{N-1}, \mathbf{n}, \mathbf{j}\right) = \sum_{\substack{i_k=j_k(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}+1,\\ k=1,\ldots,N-1}}^{(j_k+1)(p_\mathbf{n}+q_\mathbf{n})} \sum_{i_N=j_N(p_\mathbf{n}+q_\mathbf{n})+1}^{j_N(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}} \Lambda_\mathbf{i},$$

$$W\left(2^{N}, \mathbf{n}, \mathbf{j}\right) = \sum_{\substack{i_k=j_k(p_\mathbf{n}+q_\mathbf{n})+p_\mathbf{n}+1,\\ k=1,\ldots,N}}^{(j_k+1)(p_\mathbf{n}+q_\mathbf{n})} \Lambda_\mathbf{i},$$

where $q_\mathbf{n} = o([\hat{\mathbf{n}}\phi_x(a_\mathbf{n})^{(1+2N)}]^{1/(2N)})$ and $p_\mathbf{n} = [(\hat{\mathbf{n}}\phi_x(h))^{1/(2N)}/s_\mathbf{n}]$ with $s_\mathbf{n} = o([\hat{\mathbf{n}}\phi_x(h)^{(1+2N)}]^{1/(2N)}q_\mathbf{n}^{-1})$. Noting that, by (H7) we can show all sequences $q_\mathbf{n}$, $p_\mathbf{n}$ and $s_\mathbf{n}$ tend to infinity.

Now, we define for each integer $i = 1, \ldots, 2^N$,

$$T(\mathbf{n}, i) = \sum_{\mathbf{j}\in\mathcal{J}} W(i, \mathbf{n}, \mathbf{j}).$$

where $\mathcal{J} = \{0, \ldots, r_1 - 1\} \times \cdots \times \{0, \ldots, r_N - 1\}$ with $r_k = n_k(p_\mathbf{n} + q_\mathbf{n})^{-N}$. Then, we have

$$\sqrt{\hat{\mathbf{n}}\phi_x(h)}\,[\sigma(x, \theta(x))]^{-1}\left(\widehat{\Psi}_N(x, z) - \mathbb{E}\left[\widehat{\Psi}_N(x, z)\right]\right) = \left[\sqrt{\hat{\mathbf{n}}}\sigma(x, \theta(x))\right]^{-1}\left(T(\mathbf{n}, 1) + \sum_{i=2}^{2N} T(\mathbf{n}, i)\right).$$

Therefore, it suffices to proof

$$\text{the asymptotic normality of : } \left[\sqrt{\hat{\mathbf{n}}}\sigma(x, \theta(x))\right]^{-1}(T(\mathbf{n}, 1)) \tag{5.3}$$

and

$$\text{the convergence in probability of : } \sqrt{\hat{\mathbf{n}}}^{-1}\left(\sum_{i=2}^{2N} T(\mathbf{n}, i)\right). \tag{5.4}$$

Firstly, we begin by proving (5.4). Clearly it is sufficient to show that

$$\hat{\mathbf{n}}^{-1}\mathbb{E}\left[\sum_{i=2}^{2^N} T(\mathbf{n}, i)\right]^2 \to 0.$$

We have

$$\hat{\mathbf{n}}^{-1}\mathbb{E}\left[\sum_{i=2}^{2^N} T(\mathbf{n}, i)\right]^2 = \hat{\mathbf{n}}^{-1}\left(\sum_{i=2}^{2^N}\mathbb{E}\,[T(\mathbf{n}, i)]^2 + \sum_{i,j=2,\ldots,2^N,\ i\neq j}\mathbb{E}\,[T(\mathbf{n}, i)T(\mathbf{n}, j)]\right).$$

By Cauchy-Schwartz inequality, we get:

$$\forall\ 2 \le i \le 2^N :\ \ \hat{\mathbf{n}}^{-1}\mathbb{E}\,[T(\mathbf{n}, i)T(\mathbf{n}, j)] \le \left(\hat{\mathbf{n}}^{-1}\mathbb{E}\,[T(\mathbf{n}, i)]^2\right)^{\frac{1}{2}}\left(\hat{\mathbf{n}}^{-1}\mathbb{E}\,[T(\mathbf{n}, j)]^2\right)^{\frac{1}{2}}.$$

Then, all what is left to be shown is to prove that

$$\hat{\mathbf{n}}^{-1}\mathbb{E}\,[T(\mathbf{n}, i)]^2 \to 0; \quad \forall\ 2 \le i \le 2^N. \tag{5.5}$$

We will only prove (5.5) for $i = 2$, the others case is very similar. Analogously to Lemma 1, we enumerate $W(2,\mathbf{n},\mathbf{j})$ in the arbitrary way $\widehat{W}_1,\dots,\widehat{W}_M$, and we write

$$\begin{aligned}E\left[T(\mathbf{n},2)\right]^2 &= \sum_{i=1}^{M}\operatorname{Var}\left[\widehat{W}_i\right] + \sum_{i=1}^{M}\sum_{j=1,\, i\neq j}^{M}\operatorname{Cov}\left(\widehat{W}_i,\widehat{W}_j\right)\\ &= A_1 + A_2.\end{aligned}$$

For the variance term we have

$$\begin{aligned}\operatorname{Var}\left[\widehat{W}_i\right] &= \operatorname{Var}\left[\sum_{\substack{i_k=1,\\ k=1,\dots,N-1}}^{p_\mathbf{n}}\sum_{i_N=1}^{q_\mathbf{n}}\Lambda_\mathbf{i}\right]\\ &= p_\mathbf{n}^{N-1}q_\mathbf{n}\operatorname{Var}\left[\Lambda_\mathbf{i}\right] + \sum_{\substack{i_k=1,\\ k=1,\dots,N-1}}^{p_\mathbf{n}}\sum_{i_N=1}^{q_\mathbf{n}}\sum_{\substack{j_k=1,\\ k=1,\dots,N-1,\ \mathbf{i}\neq\mathbf{j}}}^{p_\mathbf{n}}\sum_{j_N=1}^{q_\mathbf{n}}\mathbb{E}\left[\Lambda_\mathbf{i}\Lambda_\mathbf{j}\right].\end{aligned}$$

It is shown in Lemma 1 in Attouch *et al.* (2009) that

$$\operatorname{Var}[\Lambda_1] \to (\sigma(x,\theta(x)))^2. \tag{5.6}$$

Moreover, employing Lemma 4, to get, under (H4),

$$\left|E\Lambda_\mathbf{i}\Lambda_\mathbf{j}\right| \le C\phi_x(h)^{-1}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right). \tag{5.7}$$

Therefore, we deduce that

$$\begin{aligned}\operatorname{Var}\left[\widehat{W}_i\right] &\le Cp_\mathbf{n}^{N-1}q_\mathbf{n}\left(1+\phi_x(h)^{-1}\sum_{\substack{i_k=1,\\ k=1,\dots,N-1}}^{p_\mathbf{n}}\sum_{i_N=1}^{q_\mathbf{n}}\left(\varphi\left(\|\mathbf{i}\|\right)\right)\right)\\ &\le Cp_\mathbf{n}^{N-1}q_\mathbf{n}\phi_x(h)^{-1}\sum_{\substack{i_k=1,\\ k=1,\dots,N-1}}^{p_\mathbf{n}}\sum_{i_N=1}^{q_\mathbf{n}}\left(\varphi(\|\mathbf{i}\|\right).\end{aligned}$$

Therefore

$$\hat{\mathbf{n}}^{-1}A_1 \le CMp_\mathbf{n}^{N-1}q_\mathbf{n}\hat{\mathbf{n}}^{-1}\phi_x(h)^{-1}\sum_{i=q_\mathbf{n}}^{\infty}i^{N-1}\varphi(i)\ .$$

The definitions of M and $p_\mathbf{n}$ permit to get

$$\begin{aligned}CMp_\mathbf{n}^{N-1}q_\mathbf{n}\hat{\mathbf{n}}^{-1}\phi_x(h)^{-1} &= \hat{\mathbf{n}}\left(p_\mathbf{n}+q_\mathbf{n}\right)^{-N}p_\mathbf{n}^{N-1}q_\mathbf{n}\hat{\mathbf{n}}^{-1}\phi_x(h)\\ &\le \left(\frac{q_\mathbf{n}}{p_\mathbf{n}}\right)\phi_x(h)^{-1}\\ &= q_\mathbf{n}s_\mathbf{n}\left(\hat{\mathbf{n}}\phi_x(h)\right)^{\frac{-1}{2N}}\phi_x(h)^{-1}\\ &= q_\mathbf{n}s_\mathbf{n}\left(\hat{\mathbf{n}}\phi_x(h)^{(1+2N)}\right)^{\frac{-1}{2N}}.\end{aligned}$$

Using the fact that $s_{\mathbf{n}} = o([\hat{\mathbf{n}}\phi_x(h)^{(1+2N)}]^{1/(2N)}q_{\mathbf{n}}^{-1})$ it is easy to see that the last term of (5.8) converges to $\to 0$. Furthermore, by (2.5) with $\delta > N$ (see, hypothesis (H2)) we show also that

$$\sum_{i=1}^{\infty} i^{N-1}\varphi(i) < \infty.$$

Finally, we deduce that

$$\hat{\mathbf{n}}^{-1}A_1 \to 0.$$

We now proceed to evaluate A_2. A simple computation shows that the sites of random variables $\Lambda_{\mathbf{i}}$ involved in two variables $\widehat{W}_i$ and $\widehat{W}_j$ with $i \neq j$ are far apart by distant of $q_{\mathbf{n}}$ at least. So, by covariance inequality in a spatial mixing variables (see, Lemma 4) we get

$$\begin{aligned} A_2 &\leq \sum_{\substack{j_k=1,\\ k=1,\dots,N}}^{n_k} \sum_{\substack{i_k=1,\\ k=1,\dots,N \|\mathbf{i}-\mathbf{j}\|>q_{\mathbf{n}}}}^{n_k} E\Lambda_{\mathbf{i}}\Lambda_{\mathbf{j}} \\ &\leq C\phi_x(h)^{-1}\hat{\mathbf{n}} \sum_{\substack{i_k=1,\\ k=1,\dots,N \|\mathbf{i}\|>q_{\mathbf{n}}}}^{n_k} \varphi(\|\mathbf{i}\|) \end{aligned}$$

and

$$\hat{\mathbf{n}}^{-1}A_2 \leq C\phi_x(h)^{-1} \sum_{i=q_{\mathbf{n}}}^{\infty} i^{N-1}\varphi(i).$$

Observe that

$$\phi_x(h)^{-1} \sum_{i=q_{\mathbf{n}}}^{\infty} i^{N-1}\varphi(i) \leq \phi_x(h)^{-1} \sum_{i=q_{\mathbf{n}}}^{\infty} i^{N-1-\delta} \leq \phi_x(h)^{-1} \int_{q_{\mathbf{n}}}^{\infty} t^{N-1-\delta}dt = C\phi_x(h)^{-1} q_{\mathbf{n}}^{N-\delta}.$$

This latter goes to 0 by means of (H7) and the definition of $q_{\mathbf{n}}$. So, we get

$$\hat{\mathbf{n}}^{-1}A_2 \to 0.$$

This completes the proof of (5.4).

Secondly, to prove the asymptotic normality (5.3) it is sufficient to show the three claim

$$Q_1 \equiv \left| \mathbb{E}\left[\exp\left[iuT(\mathbf{n},1)\right]\right] - \prod_{j_k=0,\ k=1,\dots,N}^{r_k-1} \mathbb{E}\left[\exp\left[iuW(1,\mathbf{n},\mathbf{j})\right]\right] \right| \to 0, \tag{5.8}$$

$$Q_2 \equiv \hat{\mathbf{n}}^{-1} \sum_{\mathbf{j}\in\mathcal{J}} \mathbb{E}\Big[W(1,\mathbf{n},\mathbf{j})\Big]^2 \to (\sigma(x,\theta(x)))^2 \tag{5.9}$$

and

$$Q_3 \equiv \hat{\mathbf{n}}^{-1} \sum_{\mathbf{j}\in\mathcal{J}} \mathbb{E}\left[(W(1,\mathbf{n},\mathbf{j}))^2 \mathbf{1}_{\left\{|W(1,\mathbf{n},\mathbf{j})|>\epsilon\left((\sigma(x,\theta(x)))^2\hat{\mathbf{n}}\right)^{\frac{1}{2}}\right\}}\right] \to 0, \quad \text{for all } \epsilon > 0. \tag{5.10}$$

□

Proof: **(Equation (5.8))** The proof of (5.8) is based on the Lemma 5 to the variable $(\exp(iu\tilde{W}_1), \ldots, \exp(iu\tilde{W}_M))$ where $\tilde{W}_1, \ldots, \tilde{W}_M$ are the random variables $W(1, \mathbf{n}, \mathbf{j})_{\mathbf{j}\in\mathcal{J}}$ enumerated in the arbitrary way. As $|\prod_{s=j+1}^{M} \exp[iu\tilde{W}_s]| \leq 1$, then

$$\begin{aligned}
Q_1 &= \left| \mathbb{E}\left[\exp\left[iuT(\mathbf{n}, 1)\right]\right] - \prod_{j_k=0,\, k=1,\ldots,N}^{r_k-1} \mathbb{E}\left[\exp\left[iuW(1, \mathbf{n}, \mathbf{j})\right]\right] \right| \\
&= \left| \mathbb{E}\left[\prod_{j_k=0,\, k=1,\ldots,N}^{r_k-1} \exp\left[iuW(1, \mathbf{n}, \mathbf{j})\right] \right] - \prod_{j_k=0,\, k=1,\ldots,N}^{r_k-1} \mathbb{E}\left[\exp\left[iuW(1, \mathbf{n}, \mathbf{j})\right]\right] \right| \\
&\leq \sum_{k=1}^{M-1} \sum_{j=k+1}^{M} \left| \mathbb{E}\left(\exp\left[iu\tilde{W}_k\right] - 1\right)\left(\exp\left[iu\tilde{W}_j\right] - 1\right) \prod_{s=j+1}^{M} \exp\left[iu\tilde{W}_s\right] \right. \\
&\quad \left. - \mathbb{E}\left(\exp\left[iu\tilde{W}_k\right] - 1\right) \mathbb{E}\left(\exp\left[iu\tilde{W}_j\right] - 1\right) \prod_{s=j+1}^{M} \exp\left[iu\tilde{W}_s\right] \right| \\
&= \sum_{k=1}^{M-1} \sum_{j=k+1}^{M} \left| \mathbb{E}\left(\exp\left[iu\tilde{W}_k\right] - 1\right)\left(\exp\left[iu\tilde{W}_j\right] - 1\right) - \mathbb{E}\left(\exp\left[iu\tilde{W}_k\right] - 1\right) \mathbb{E}\left(\exp\left[iu\tilde{W}_j\right] - 1\right) \right| \\
&\quad \times \left| \prod_{s=j+1}^{M} \exp\left[iu\tilde{W}_s\right] \right| \\
&\leq \sum_{k=1}^{M-1} \sum_{j=k+1}^{M} \left| \mathbb{E}\left(\exp\left[iu\tilde{W}_k\right] - 1\right)\left(\exp\left[iu\tilde{W}_j\right] - 1\right) - \mathbb{E}\left(\exp\left[iu\tilde{W}_k\right] - 1\right) \mathbb{E}\left(\exp\left[iu\tilde{W}_j\right] - 1\right) \right|.
\end{aligned}$$

Let $\tilde{I}_j$ be the set of sites such that $\tilde{W}_j = \sum_{\mathbf{i}\in\tilde{I}(1,\mathbf{n},\mathbf{j})} \Lambda_{\mathbf{i}}$. Since the sets $\tilde{I}_{1\leq j\leq M}$ contains $p_{\mathbf{n}}^N$ sites, we have by Lemma 4, under (2.3)

$$\left| \mathbb{E}\left(\exp\left[iu\tilde{W}_k\right] - 1\right)\left(\exp\left[iu\tilde{W}_j\right] - 1\right) - \mathbb{E}\left(\exp\left[iu\tilde{W}_k\right] - 1\right) \mathbb{E}\left(\exp\left[iu\tilde{W}_j\right] - 1\right) \right| \leq C\varphi\left(d\left(\tilde{I}_j, \tilde{I}_k\right)\right) p_{\mathbf{n}}^N.$$

Hence

$$\begin{aligned}
Q_1 &\leq C p_{\mathbf{n}}^N \sum_{k=1}^{M-1} \sum_{j=k+1}^{M} \varphi\left(d\left(\tilde{I}_j, \tilde{I}_k\right)\right) \\
&\leq C p_{\mathbf{n}}^N M \sum_{k=2}^{M} \varphi\left(d\left(\tilde{I}_1, \tilde{I}_k\right)\right) \\
&\leq C p_{\mathbf{n}}^N M \sum_{i=1}^{\infty} \sum_{k: iq_{\mathbf{n}} \leq d(\tilde{I}_1, \tilde{I}_k) < (i+1)q_{\mathbf{n}}} \varphi\left(d\left(\tilde{I}_1, \tilde{I}_k\right)\right) \\
&\leq C p_{\mathbf{n}}^N M \sum_{i=1}^{\infty} i^{N-1} \varphi(iq_{\mathbf{n}}).
\end{aligned}$$

It follows from (2.5) that

$$Q_1 \leq C\hat{\mathbf{n}}q_{\mathbf{n}}^{-\delta} \sum_{i=1}^{\infty} i^{N-1-\delta}.$$

The convergence result (5.8) is consequence of (H7) and the definition of $q_{\mathbf{n}}$. □

Proof: **(Equation (5.9))** On the one hand

$$\hat{\mathbf{n}}^{-1}\mathbb{E}\left[T(\mathbf{n},1)\right]^2 = \hat{\mathbf{n}}^{-1} \sum_{j_k=0,\ k=1,\ldots,N}^{r_k-1} \mathbb{E}\left[W(1,\mathbf{n},\mathbf{j})\right]^2$$
$$+ \hat{\mathbf{n}}^{-1} \sum_{j_k=0,\ k=1,\ldots,N}^{r_k-1} \sum_{\substack{i_k=0,\ k=1,\ldots,N \\ i_k\neq j_k \text{ for some } k}}^{r_k-1} \mathrm{Cov}[W(1,\mathbf{n},\mathbf{j}), W(1,\mathbf{n},\mathbf{i})].$$

By the same arguments as those used for A_2, this last term tend to zero. Hence, the limit in (5.9) is equal to the limit of $\hat{\mathbf{n}}^{-1}E\left(T(\mathbf{n},1)\right)^2$. However, recall that

$$S_{\mathbf{n}} := \sum_{i=1}^{2^N} T(\mathbf{n},i) = T(\mathbf{n},1) + S''_{\mathbf{n}},$$

where $S''_{\mathbf{n}} = \sum_{i=2}^{2^N} T(\mathbf{n},i)$. Therefore

$$\hat{\mathbf{n}}^{-1}\mathbb{E}\left[T(\mathbf{n},1)\right]^2 = \hat{\mathbf{n}}^{-1}\mathbb{E}\left[S_{\mathbf{n}}^2\right] + \hat{\mathbf{n}}^{-1}\mathbb{E}\left[S''_{\mathbf{n}}\right]^2 - 2\hat{\mathbf{n}}^{-1}\mathbb{E}\left[S_{\mathbf{n}}S''_{\mathbf{n}}\right].$$

It is shown in (5.5) that $\hat{\mathbf{n}}^{-1}\mathbb{E}[S''^2_{\mathbf{n}}] \to 0$. Moreover, by Cauchy-Schwartz's inequality, we can write:

$$\left|\hat{\mathbf{n}}^{-1}\mathbb{E}\left[S_{\mathbf{n}}S''_{\mathbf{n}}\right]\right| \leq \hat{\mathbf{n}}^{-1}\mathbb{E}\left|S_{\mathbf{n}}S''_{\mathbf{n}}\right| \leq \left(\hat{\mathbf{n}}^{-1}\mathbb{E}\left[S_{\mathbf{n}}\right]^2\right)^{\frac{1}{2}} \left(\hat{\mathbf{n}}^{-1}\mathbb{E}\left[S''_{\mathbf{n}}\right]^2\right)^{\frac{1}{2}}.$$

Thus, all what it remains to compute is the limit of $\hat{\mathbf{n}}^{-1}\mathbb{E}\left[S_{\mathbf{n}}\right]^2$ which can be writhen

$$\hat{\mathbf{n}}^{-1}\mathbb{E}\left[S_{\mathbf{n}}\right]^2 = \hat{\mathbf{n}}^{-1}\mathrm{Var}\left[S_{\mathbf{n}}^2\right] = \hat{\mathbf{n}}^{-1}\left(\sum_{\mathbf{i}} \mathrm{Var}\left[\Lambda_{\mathbf{i}}\right] + \sum_{\mathbf{i}\neq\mathbf{j}} \mathrm{Cov}\left[\Lambda_{\mathbf{i}},\Lambda_{\mathbf{j}}\right]\right).$$

As indicated in (5.6) the variance term is

$$\mathrm{Var}[\Lambda_1] \to \sigma(x,\theta(x))^2.$$

Let us evaluate the covariance term. Reasoning as in Lemma 1 we consider

$$E_1 = \left\{\mathbf{i},\mathbf{j} \in \mathcal{I}_{\mathbf{n}} : 0 < \|\mathbf{i}-\mathbf{j}\| \leq c_{\mathbf{n}}\right\},$$
$$E_2 = \left\{\mathbf{i},\mathbf{j} \in \mathcal{I}_{\mathbf{n}} : \|\mathbf{i}-\mathbf{j}\| > c_{\mathbf{n}}\right\},$$

where $c_{\mathbf{n}}$ is a sequence of integers that converges to infinite and that will be precise after.

Now, we write

$$\sum_{\mathbf{i}\neq\mathbf{j}} \mathrm{Cov}\left[\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right] = \sum_{(\mathbf{i},\mathbf{j})\in E_1} \mathrm{Cov}\left[\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right] + \sum_{(\mathbf{i},\mathbf{j})\in E_2} \mathrm{Cov}\left[\Lambda_1, \Lambda_{\mathbf{j}}\right].$$

For the first sum on E_1, proceeding as in $(R^1_{\mathbf{n}})$ in Lemma 1, we get from the definition of $\Lambda_{\mathbf{i}}$

$$\left|\mathrm{Cov}\left[\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right]\right| \leq C\left(\phi_x(h) + (\phi_x(h))^{\frac{1}{a}}\right) \leq C(\phi_x(h))^{\frac{1}{a}}.$$

It follows that,

$$\sum_{E_1} \mathrm{Cov}\left(\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right) \leq C\hat{\mathbf{n}}c_{\mathbf{n}}^N \phi_x(h)^{\frac{1}{a}}.$$

Next, on E_2 we apply Lemma 4, once again similarly to (5.7), we write that:

$$\left|\mathrm{Cov}\left(\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right)\right| \leq C\phi_x(h)^{-1}\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right)$$

and

$$\begin{aligned}
\sum_{E_2} \mathrm{Cov}\left(\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right) &\leq C\phi_x(h)^{-1} \sum_{(\mathbf{i},\mathbf{j})\in E_2} \varphi\left(\|\mathbf{i}-\mathbf{j}\|\right) \\
&\leq C\hat{\mathbf{n}}\phi_x(h)^{-1} \sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}} \varphi\left(\|\mathbf{i}\|\right) \\
&\leq C\hat{\mathbf{n}}\phi_x(h)^{-1} c_{\mathbf{n}}^{-\delta} \sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}} \|\mathbf{i}\|^{\delta}\varphi\left(\|\mathbf{i}\|\right).
\end{aligned}$$

Finally, we have:

$$\sum \mathrm{Cov}\left(\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right) \leq \left(C\hat{\mathbf{n}}c_{\mathbf{n}}^N \phi_x(h)^{\frac{1}{a}} + C\hat{\mathbf{n}}\phi_x(h)^{-1} c_{\mathbf{n}}^{-\delta} \sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}} \|\mathbf{i}\|^{\delta}\varphi\left(\|\mathbf{i}\|\right)\right).$$

Let $c_{\mathbf{n}} = \phi_x(h)^{-\alpha}$ for some $(\delta)^{-1} < \alpha < (Na)^{-1}$, then we have:

$$\sum \mathrm{Cov}\left(\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right) \leq \left(C\hat{\mathbf{n}}\phi_x(h)^{-\alpha N+\frac{1}{a}} + C\hat{\mathbf{n}}\phi_x(h)^{\alpha\delta-1} \sum_{\mathbf{i}:\|\mathbf{i}\|>c_{\mathbf{n}}} \|\mathbf{i}\|^{\delta}\varphi\left(\|\mathbf{i}\|\right)\right).$$

Hence, we obtain that

$$\sum \mathrm{Cov}\left(\Lambda_{\mathbf{i}}, \Lambda_{\mathbf{j}}\right) = o\left(\hat{\mathbf{n}}\right).$$

In conclusion, we have

$$\hat{\mathbf{n}}^{-1} \sum_{\mathbf{j}\in\mathcal{J}} E\left[W(1,\mathbf{n},\mathbf{j})\right]^2 \to \sigma^2(x,\theta(x)), \quad \text{when } \mathbf{n}\to\infty.$$

□

Proof: **(Equation (5.10))** Because $|\Lambda_{\mathbf{i}}| \leq C\phi_x(h)^{-1/2}$, then $|W(1,\mathbf{n},\mathbf{j})| \leq Cp_{\mathbf{n}}^N\phi_x(h)^{-1/2}$. Thus

$$Q_4 \leq Cp_{\mathbf{n}}^{2N}\phi_x(h)^{-1}\hat{\mathbf{n}}^{-1}\sum_{j_k=0,\ k=1,\dots,N}^{r_k-1}\mathbb{P}\left[|W(1,\mathbf{n},\mathbf{j})| > \epsilon\left((\sigma(x,\theta(x)))^2\,\hat{\mathbf{n}}\right)^{\frac{1}{2}}\right].$$

Since $p_{\mathbf{n}} = [(\hat{\mathbf{n}}\phi_x(h))^{1/(2N)}/s_{\mathbf{n}}]$ and $s_{\mathbf{n}} \to \infty$ then

$$\begin{aligned}|W(1,\mathbf{n},\mathbf{j})| \,/ \left(\left((\sigma(x,\theta(x)))^2\hat{\mathbf{n}}\right)^{\frac{1}{2}}\right) &\leq Cp_{\mathbf{n}}^N\,(\hat{\mathbf{n}}\phi_x(h))^{-\frac{1}{2}}\\ &= C\,(s_{\mathbf{n}})^{-N} \to 0.\end{aligned}$$

So, for all ϵ and $\mathbf{j} \in \mathcal{J}$; if $\mathbf{n}$ is great enough, we have
$\mathbb{P}[W(1,\mathbf{n},\mathbf{j}) > \epsilon(\sigma(x,\theta(x)))^2\hat{\mathbf{n}})^{1/2}] = 0$. Then $Q_4 = 0$ for $\mathbf{n}$ great enough. This yields the proof. □

Proof: **(Lemma 3)** Clearly by the equiprobability of the couples (X_i, Y_i) we have

$$\begin{aligned}&\mathbb{E}\left[\widehat{\Psi}_N(x,z)\right]\\ &= \frac{1}{\mathbb{E}[K_1(x)]}\mathbb{E}\left[K_1(x)\psi_x(Y,z)\right]\\ &= \frac{\mathbb{E}\left\{K_1(x)\left[\mathbb{E}[\psi_x(Y,z)|X_1] - \mathbb{E}[\psi_x(Y,\theta(x))|X=x]\right]\right\}}{\mathbb{E}[K_1(x)]}\\ &= \frac{\mathbb{E}\left\{K_1(x)\left[\mathbb{E}[\psi_x(Y,z)|X_1] - \mathbb{E}[\psi_x(Y,z)|X=x]\right]\right\}}{\mathbb{E}[K_1(x)]} + \mathbb{E}[\psi_x(Y,z)|X=x] - \mathbb{E}[\psi_x(Y,\theta(x))|X=x]\\ &=: I_1 + I_2. \qquad (5.11)\end{aligned}$$

For $I_1(x)$ we use (H2) to write,

$$K_1(x)\,|\mathbb{E}[\psi_x(Y,z)|X_1] - \mathbb{E}[\psi_x(Y,z)|X=x]| \leq Ch^bK_1(x)$$

which gives

$$I_1 = O\left(h^b\right).$$

Concerning I_2 we use a Taylor expansion to get, under (H1′) and (H4′)

$$I_2 = u\left[n\phi_x(h)\right]^{-\frac{1}{2}}\sigma(x,\theta(x))\frac{\partial}{\partial t}\Psi(x,\theta(x)) + o\left(\left[n\phi_x(h)\right]^{-\frac{1}{2}}\right).$$

The result is then a consequence of (5.11). □

Acknowledgement

The authors wish to thank the referee for helpful comments.

References

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

Attouch, M. K., Laksaci, A. and Ould Saïd, E. (2009). Asymptotic distribution of robust estimator for functional nonparametric models, *Communications in Statistics - Theory and Methods*, **38**, 1317–1335.

Attouch, M. K., Laksaci, A. and Ould-Saïd, E. (2010). Asymptotic normality of a robust estimator of the regression function for functional time series, *Journal of the Korean Statistical Society*, **39**, 489–500.

Azzedine, N., Laksaci, A. and Ould Saïd, E. (2008). On the robust nonparametric regression estimation for functional regressor, *Statistics & Probability Letters*, **78**, 3216–3221.

Biau, G. (2003). Spatial Kernel Density Estimation, *Mathematical Methods of Statistics*, **12**, 371–390.

Biau, G. and Cadre, B. (2004). Nonparametric spatial prediction, *Statistical Inference for Stochastic Processes*, **7**, 327–349.

Boente, G. and Fraiman, R. (1989). Nonparametric regression estimation, *Journal of Multivariate Analysis*, **29**, 180–198.

Bosq, D. (2000). *Linear Processes in Function Spaces. Theory and Application*, Lectures Notes in Statistics. Springer Verlag, New York.

Cadre, B. (2001). Convergent estimators for the L_1-median of a Banach-valued random variable, *Statistics*, **35**, 509–521.

Carbon, M., Francq, C. and Tran, L. T. (2007). Kernel regression estimation for random fields, *Journal of Statistical Planning and Inference*, **137**, 778–798.

Carbon, M., Hallin, M. and Tran, L. T. (1996). Kernel density estimation for random fields: The L^1 theory, *Journal of Nonparametric Statistics*, **6**, 157–170.

Chen, J. and Zhang, L. (2009). Asymptotic properties of nonparametric M-estimation for mixing functional data, *Journal of Statistical Planning and Inference*, **139**, 533–546.

Collomb, G. and Härdle, W. (1986). Strong uniform convergence rates in robust nonparametric time series analysis and prediction: Kernel regression estimation from dependent observations, *Stochastic Processes and their Applications*, **23**, 77–89.

Crambes, C., Delsol, L. and Laksaci, A. (2008). Robust nonparametric estimation for functional data, *Journal of Nonparametric Statistics*, **20**, 573–598.

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

Dabo-Niang, S. and Thiam, B. (2010). Robust quantile estimation and prediction for spatial processes, *Statistics & Probability Letters*, **80**, 1447–1458.

Dabo-Niang, S. and Yao, A. F. (2007). Spatial kernel regression estimation, *Mathematical Methods of Statistics*, **16**, 1–20.

Delicado, P., Giraldo, R., Comas, C. and Mateu, J. (2010). Statistics for spatial functional data: Some recent contributions, *Environmetrics*, **21**, 224–239.

Doukhan, P., Massart, P. and Rio, E. (1994). The functional central limit theorem for strongly mixing processes, *Annales de l'Institut Henri Poincaré*, **30**, 63–82.

Ferraty, F., Laksaci, A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models, *Statistical Inference for Stochastic Processes*, **9**, 47–76.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis. Theory and Practice*, Springer-Verlag, New York.

Gheriballah, A., Laksaci, A. and Rouane, R. (2010). Robust nonparametric estimation for spatial regression, *Journal of Statistical Planning and Inference*, **140**, 1656–1670.

Guyon, X. (1987). Estimation d'un champ par pseudo-vraisemblance conditionnelle: Etude asymptotique et application au cas Markovien, *Proceedings of the Sixth Franco-Belgian Meeting of Statisticia.*

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

Hallin, M., Lu, Z. and Yu, K. (2009). Local linear spatial quantile regression, *Bernoulli*, **15**, 659–686.

Laïb, N. and Ould Saïd, E. (2000). A robust nonparametric estimation of the autoregression function under ergodic hypothesis, *Canadian Journal of Statistics*, **28**, 817–828.

Li, J. and Tran, L. T. (2009). Nonparametric estimation of conditional expectation, *Journal of Statistical Planning and Inference*, **139**, 164–175.

Lu, C., Chen, D. and Kou, Y. (2003). Detecting spatial outliers with multiple attributes, *IEEE*, **2-3**, 122–128.

Nerini, D., Monestiez, P. and Manté, C. (2010). A cokriging method for spatial functional, *Journal of Multivariate Analysis*, **101**, 409–418.

Ramsay, J. (2008). FDA problems that I like to talk about, *Personal Communication*.

Ramsay, J. O. and Silverman, B. W. (2002). *Applied Functional Data Analysis. Methods and Case Studies*, Springer-Verlag, New York.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*, Second edition, Springer-Verlag, New York.

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

Tran, L. T. (1990). Kernel density estimation under dependence, *Statistics* & *Probability Letters*, **10**, 193–201.

Xu, R. and Wang, J. (2008). L_1- estimation for spatial nonparametric regression, *Journal of Nonparametric Statistics*, **20**, 523–537.

Received November 2011; Accepted December 2011