Robust quantile estimation and prediction for spatial processes

Sophie Dabo-Niang [*]   Baba Thiam [†]

January 26, 2010

Abstract

In this paper, we present a statistical framework for modeling conditional quantiles of spatial processes assumed to be strongly mixing in space. We establish the L_1 consistency and the asymptotic normality of the kernel conditional quantile estimator in the case of random fields. We also define a nonparametric spatial predictor and illustrate the methodology used with some simulations.

Keywords: Spatial processes ; Kernel estimate ; Conditional quantile ; Spatial prediction.

1 Introduction

Let (X, Y) be a pair of random variables with values in $\mathbb{R}^d \times \mathbb{R}$ and defined on a probability space $(\Omega, \mathcal{A}, P)$. Assume that the joint density of (X, Y) and the marginal density of X exist and are denoted respectively by $f(x, y)$ and $g(x)$. In the following, we suppose that $F(\cdot|x)$, the conditional distribution function of Y given $X = x$ exists and we denote by $f(\cdot|x)$ the density of Y given $X = x$. For $p \in]0, 1[$ and for fixed $x \in \mathbb{R}^d$, let $\mu_p(x)$ be the conditional quantile of order p of $F(\cdot|x)$, that can be seen as a solution of the equation $F(y|x) = p$. Another alternative characterization of the *pth* conditional quantile (see for example Gannoun et al. [6]) is $\mu_p(x) = \operatorname{argmin}_{\theta \in \mathbb{R}} \boldsymbol{E}\left[(2p - 1)(Y - \theta) + |Y - \theta| \,|X = x\right]$.

We are interested to the non-parametric estimation of $\mu_p(x)$ in the case of spatial dependent observations. Nonparametric conditional quantile estimation technics have already been developed for non spatial (independent or mixing) real valued processes. Such results have provided useful tools for solving for example some prediction problems of strictly stationary processes satisfying the α-mixing condition. The existing results in the non-spatial case include the works of Matzner-Løber [10], Collomb [4], Gannoun et al. [6], Laksaci et al. [8].

In nonparametric spatial estimation, the existing works concern mainly the estimation of a probability density and regression functions, see the key references: Tran [12], Biau and Cadre [2], Carbon et al. [3].

For the spatial quantile conditional estimation case, there exist only few results in our knowledge. Abdi et al. [1] considered the pointwise $p-$mean and almost complete consistencies of a double kernel quantile estimator for real-valued random fields. Hallin *et al.* [7] give a Bahadur representation and asymptotic normality results of the local linear quantile estimator. Laksaci and Fouzia [9] consider the case where the regressor take their values in a semi-metric space and show the strong and weak consistency of the conditional quantile.

In this paper, we will go beyond all these last spatial works and provide the L_1 consistency and an asymptotic normality of a kernel conditional quantile estimate of a strictly stationary spatial process satisfying the α-mixing condition. In addition, we employ our results to solve some nonparametric prediction problems. The organization of this paper is as follows. The estimation procedure is presented in Section 2. Section 3 gives some necessary conditions and then establishes the main asymptotic results. Section 4 is devoted to simulations results. Technical proofs are given in Section 5.

---

[*]Corresponding author: sophie.dabo@univ-lille3.fr

[†]Laboratoire EQUIPPE, Université Charles-De-Gaulle, Lille 3, Maison de la Recherche, domaine universitaire du Pont de Bois, BP 60149, 59653 Villeneuve d'Ascq cedex, France. baba.thiam@univ-lille3.fr

2 Nonparametric estimator of the conditional quantile

Let us consider a strictly stationary process $((X_\mathbf{i}, Y_\mathbf{i}), \mathbf{i} \in \mathcal{I_n})$ with values in $\mathbb{R}^d \times \mathbb{R}$ where $(X_\mathbf{i}, Y_\mathbf{i})$ has the same distribution as (X, Y). For $\mathbf{n} \in (\mathbb{N}^*)^N$, we define a rectangular region $\mathcal{I_n}$ by $\mathcal{I_n} = \{\mathbf{i} = (i_1, \ldots, i_N) \in (\mathbb{N}^*)^N, \ 1 \leq i_k \leq n_k, \ k = 1, \ldots, N\}$. We set $\widehat{\mathbf{n}} = n_1 \ldots n_N$, and we write $\mathbf{n} \to \infty$ if $\min_{k=1,\ldots,N} n_k \to \infty$. The well known kernel estimates of f and g are defined by

$$f_\mathbf{n}(x, y) = \frac{1}{\widehat{\mathbf{n}} h^{d+1}} \sum_{\mathbf{i} \in \mathcal{I_n}} K\left(\frac{x - X_\mathbf{i}}{h}\right) w\left(\frac{y - Y_\mathbf{i}}{h}\right) \text{ and } g_\mathbf{n}(x) = \frac{1}{\widehat{\mathbf{n}} h^d} \sum_{\mathbf{i} \in \mathcal{I_n}} K\left(\frac{x - X_\mathbf{i}}{h}\right), \text{ where } K \text{ and } w \text{ are}$$

two probability density functions, and the bandwidths $h = h(\mathbf{n})$ is a sequence of positive real numbers such that $h \to 0$ as $\mathbf{n} \to \infty$. The kernel estimate of the conditional density $f_\mathbf{n}(y|x)$ is naturally defined by the ratio $f_\mathbf{n}(x, y)$ over $g_\mathbf{n}(x)$ while the estimator of the conditional distribution function (see the one introduced by Roussas [11]) is defined by

$$
\begin{aligned}
F_\mathbf{n}(y|x) &= \frac{\psi_\mathbf{n}(x, y)}{g_\mathbf{n}(x)} \mathbb{1}_{\{g_\mathbf{n}(x) \neq 0\}}, \\
\psi_\mathbf{n}(x, y) &= \frac{1}{\widehat{\mathbf{n}} h^{d+1}} \sum_{\mathbf{i} \in \mathcal{I_n}} K\left(\frac{x - X_\mathbf{i}}{h}\right) \int_{-\infty}^{y} w\left(\frac{z - Y_\mathbf{i}}{h}\right) dz.
\end{aligned}
$$

For a fixed x, the estimator of the pth conditional quantile noted $\mu_{p,\mathbf{n}}(x)$ can be defined as the root of the equation $F_\mathbf{n}(z|x) = p$. Alternatively, one can consider the local constant estimator defined by

$$\nu_{p,\mathbf{n}}(x) = \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} \sum_{\mathbf{i} \in \mathcal{I_n}} \left(|Y_\mathbf{i} - \theta| + (2p - 1)(Y_\mathbf{i} - \theta)\right) K\left(\frac{x - X_\mathbf{i}}{h}\right).$$

In this paper, we will focus on the study of the asymptotic behavior of $\mu_{p,\mathbf{n}}$. For the study of $\nu_{p,\mathbf{n}}$, one can adapt the technics developed in Zhou [14].

3 Main results

To establish the asymptotic results, we will suppose that the sequence $(X_\mathbf{i}, Y_\mathbf{i}))_{\mathbf{i} \in (\mathbb{N}^*)^N}$ satisfies the following mixing condition: there exists a function $\chi : \mathbb{R}^+ \to \mathbb{R}^+$ with $\chi(t) \downarrow 0$ as $t \to \infty$, such that whenever $E, E' \subset (\mathbb{N}^*)^N$ with finite cardinals,

$$
\begin{aligned}
\alpha\left(\mathcal{B}(E), \mathcal{B}(E')\right) &:= \sup\left\{|\boldsymbol{P}(A \cap B) - \boldsymbol{P}(A)\boldsymbol{P}(B)|; \ A \in \mathcal{B}(E), B \in \mathcal{B}(E')\right\} \\
&\leq \phi(\operatorname{Card} E, \operatorname{Card} E') \chi(\operatorname{dist}(E, E')),
\end{aligned}
$$

where $\mathcal{B}(E)$ (resp. $\mathcal{B}(E')$) denotes the Borel σ-fields generated by $(X_\mathbf{i}, Y_\mathbf{i})_{\mathbf{i} \in E}$ (resp. $(X_\mathbf{i}, Y_\mathbf{i})_{\mathbf{i} \in E'}$), Card E (resp. Card E') the cardinality of E (resp. E'), $\operatorname{dist}(E, E')$ the Euclidean distance between E and E', and $\phi : \mathbb{N}^2 \to \mathbb{R}^+$ is a symmetric positive function which is non decreasing in each variable. Throughout this paper, we will assume that ϕ satisfies

$$\phi(n, m) \leq C \min(n, m), \quad \forall n, m \in \mathbb{N} \tag{1}$$

or

$$\phi(n, m) \leq C (n + m + 1)^\kappa, \quad \forall n, m \in \mathbb{N} \tag{2}$$

for some $\kappa \geq 1$ and some $C > 0$. If $\phi \equiv 1$, then the field $(X_\mathbf{i}, Y_\mathbf{i})_{\mathbf{i} \in (\mathbb{N}^*)^\mathbb{N}}$ is called *strongly mixing*. In this paper, we consider the case where $\chi(i)$ tends to zero at a polynomial rate, that is,

$$\chi(i) = O(i^{-\beta}), \tag{3}$$

2

with $\beta > 0$. We fix a compact subset S of $\mathbb{R}^d$. Denote $a = \inf\{y : F(y|x) > 0\}$ and $b = \sup\{y : F(y|x) < 1\}$, we will suppose that $\mathcal{V} \subseteq [a, b]$ is a compact neighborhood of the unknown quantile $\mu(x)$. For mixing coefficients with polynomial decreasing rate (3), the constraints on the bandwidth will be related to β by means of

$$\theta_1 = \frac{N(d+1)(d+2) + (d+1)\beta}{\beta - N(d+5)}, \quad \theta_2 = \frac{N(d+2) - \beta}{\beta - N(d+5)},$$

$$\theta_3 = \frac{N(d^2 + 4d + 2) + (d+1)\beta}{\beta - N(d+4+2\kappa)}, \quad \theta_4 = \frac{N(d+1) - \beta}{\beta - N(d+4+2\kappa)}.$$

Denote $\Omega_{\mathbf{n}} = \sqrt{\frac{\log \hat{\mathbf{n}}}{\hat{\mathbf{n}} h^d}}$. Let ε be an arbitrary small positive number and set $u(\mathbf{n}) = \prod_{i=1}^{N} (\log n_i)(\log \log n_i)^{1+\varepsilon}$. It is clear that $\sum_{\mathbf{n} \in \mathbb{Z}^N} 1/(\hat{\mathbf{n}} u(\mathbf{n})) < \infty$. In the sequel, we use the following hypotheses.

(A1) f and g are respectively continuous on $\mathbb{R}^{d+1}$ and $\mathbb{R}^d$, g satisfies a Lipschitz condition, $g(x) > 0, \forall x \in S$.

(A2) There exists $D \geq 0$ such that the pairs $(X_{\mathbf{i}}, X_{\mathbf{j}})$ and $((X_{\mathbf{i}}, Y_{\mathbf{i}}), (X_{\mathbf{j}}, Y_{\mathbf{j}}))$ admit a density, say $g_{\mathbf{i},\mathbf{j}}$ and $f_{\mathbf{i},\mathbf{j}}$, as soon as $dist(\mathbf{i}, \mathbf{j}) > D$. Moreover, for some constant $c \geq 0$,

$$|f_{\mathbf{i},\mathbf{j}}(s, t) - f(s)f(t)| \leq c, \quad \forall s, t \in \mathbb{R}^{d+1}$$

and

$$|g_{\mathbf{i},\mathbf{j}}(u, v) - g(u)g(v)| \leq c, \quad \forall u, v \in \mathbb{R}^d.$$

(A3) i) $f^{(i,j)}(\cdot, \cdot) = \frac{\partial^{i+j} f}{\partial x_l^i \partial y^j}(\cdot, \cdot)$ exists, is bounded and integrable for $0 \leq i + j \leq 2$, and $0 \leq l \leq d$.
 ii) $F(y|x)$ has continuous second partial derivatives with respect to x.

(A4) $F(y|x)$ has continuous second derivative with respect to y.

(A5) The kernel K is integrable, symmetric and is a lipschitzian density function on $\mathbb{R}^d$ with compact support. Moreover $\int_{\mathbb{R}^d} \|s\|^2 K(s) ds < \infty$.

(A6) The kernel w is a symmetric and lipschitzian density function on $\mathbb{R}$ and has compact support.

(A7) $\lim_{\mathbf{n} \to \infty} \hat{\mathbf{n}} h^{d+2} (\log \hat{\mathbf{n}})^{-1} = 0$.

(A8) The function $\mu_p(x)$ satisfies a uniform uniqueness property on S:

$$\forall \varepsilon > 0, \exists \eta > 0, \forall r : S \to \mathbb{R}, \sup_{x \in S} |\mu_p(x) - r(x)| \geq \varepsilon \Rightarrow \sup_{x \in S} |F(\mu_p(x)|x) - F(r(x)|x)| \geq \eta.$$

(A9) $\hat{\mathbf{n}} h^{\theta_1} (\log \hat{\mathbf{n}})^{\theta_2} (u(\mathbf{n}))^{\frac{-2N}{\beta - N(d+5)}} \to \infty$.

(A10) $\hat{\mathbf{n}} h^{\theta_3} (\log \hat{\mathbf{n}})^{\theta_4} (u(\mathbf{n}))^{\frac{-2N}{\beta - N(d+4+2\kappa)}} \to \infty$.

Comments on the hypotheses:
Assumptions (A9) and (A10) imply conditions (3.7) and (3.8) of Theorem 3.3 in Carbon et al. [3] and they also imply the classical condition $\hat{\mathbf{n}} h^{d+1}/\log \hat{\mathbf{n}} \to \infty$.
Assumption (A8) is introduced for getting consistency results on the quantile from those of the conditional distribution.

In order to state the asymptotic results, we will suppose that (A9) and (1) or (A10) and (2) are satisfied. The following two theorems give uniform almost sure convergence results of respectively $F_{\mathbf{n}}(y|x)$ and $\mu_{p,\mathbf{n}}(x)$ and permit to establish the L_1 consistency of $\mu_{p,\mathbf{n}}(x)$ (see Corollary 1).

3

Theorem 1 *Assume (A1)-(A7) hold, then*

$$\sup_{y \in \mathcal{V}} \sup_{x \in S} |F_{\mathbf{n}}(y|x) - F(y|x)| \;=\; O\left(\Omega_{\mathbf{n}}\right) \quad a.s.$$

Theorem 2 *If (A1)-(A8) are satisfied, then we have*

$$\sup_{x \in S} |\mu_{p,\mathbf{n}}(x) - \mu_p(x)| \overset{a.s.}{\to} 0.$$

Corollary 1 *Assume (A1)-(A8) hold, then*

$$\boldsymbol{E}\left[\{\mu_{p,\mathbf{n}}(X_{\mathbf{n}}) - \mu_p(X_{\mathbf{n}})\}\mathbb{1}_{\{X_{\mathbf{n}} \in S\}}\right] \overset{a.s.}{\to} 0.$$

To establish the following asymptotic normality of $\mu_{p,\mathbf{n}}(x)$ (Theorem 3), we will suppose that for any $(x, y) \in S \times \mathcal{V}$, there exists $c > 0$ such that $f(y|x) > c$. Moreover, we will assume that the following additional conditions on the bandwidth hold for some $0 < \gamma < 1$.

(C1) $\widehat{\mathbf{n}} h^{d(1+2N(1-\gamma))} \to \infty$.

(C2) There exists a sequence of positive integers $q = q_{\mathbf{n}} \to \infty$ with $q = o\left((\widehat{\mathbf{n}} h^{d(1+2N(1-\gamma))})^{1/2N}\right)$ such that $\widehat{\mathbf{n}} \sum_{i=1}^{\infty} i^{N-1}\chi(iq) \to 0$ and $h^{-d(1-\gamma)} \sum_{i=q}^{\infty} i^{N-1}(\chi(i))^{1-\gamma} \to 0$.

Theorem 3 *Assume that (A1)-(A8), (C1) and (C2) hold. If there exists $c \geq 0$ such that $\widehat{\mathbf{n}} h^{d+4} \to c$, then*

$$\sqrt{\widehat{\mathbf{n}} h^d}\left(\mu_{p,\mathbf{n}}(x) - \mu_p(x)\right) \overset{\mathcal{L}}{\to} \mathcal{N}\left(c\frac{B(x, \mu_p(x))}{f(\mu_p(x)|x)}, \frac{\sigma^2(x, \mu_p(x))}{(f(\mu_p(x)|x))^2}\right),$$

where

$$
\begin{aligned}
B(x, y) &= \frac{1}{2}\Bigg\{ \sum_{i,j=1}^{d}\left[\frac{\partial^2 F(y|x)}{\partial x_i \partial x_j} + \frac{2}{g(x)}\frac{\partial g(x)}{\partial x_i}\frac{\partial F(y|x)}{\partial x_j}\right]\int_{\mathbb{R}^d}\|s\|^2 K(s)ds \\
&\quad + \frac{\partial^2 F(y|x)}{\partial y^2}\int_{\mathbb{R}} t^2 w(t)dt \Bigg\}.
\end{aligned}
\tag{4}
$$

$$\sigma^2(x, y) = \frac{F(y|x)\left[1 - F(y|x)\right]}{g(x)}\int_{\mathbb{R}^d} K^2(z)dz. \tag{5}$$

3.1 Prediction

Let $(\xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{I}_{\mathbf{n}})$ be a $\mathbb{R}-$valued strictly stationary random spatial process, assumed to be bounded, observable over a region $\mathcal{I}_{\mathbf{n}} \subset \mathbb{N}^N$ and observed over a subset $\mathcal{O}_{\mathbf{n}}$ of $\mathcal{I}_{\mathbf{n}}$, $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N$. The aim of this section is to predict $\xi_{\mathbf{i}_0}$, at a given fixed point $\mathbf{i}_0$ not in $\mathcal{O}_{\mathbf{n}} \subset \mathbb{N}^N$. In practice (e.g. for simplicity), we expect that $\xi_{\mathbf{i}_0}$ depends only on the values of the process on a bounded neighborhood $\mathcal{V}_{\mathbf{i}_0} \subset \mathcal{O}_{\mathbf{n}}$. In other words, we expect that the process $(\xi_{\mathbf{i}})$ satisfies a Markov property, see for example Biau and Cadre [2], Dabo-Niang and Yao [5]. Moreover, we assume that $\mathcal{V}_{\mathbf{i}_0} = \mathcal{V} + \mathbf{i}_0$, where $\mathcal{V}$ is a fixed bounded set of sites that does not contain 0. It is well known that the best predictor of $\xi_{\mathbf{i}_0}$ given the data in $\mathcal{V}_{\mathbf{i}_0}$ in the sense of *mean-square error* is

$$E(\xi_{\mathbf{i}_0}|\xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i}_0}).$$

Let $\mathcal{V}_{\mathbf{i}} = \mathcal{V} + \mathbf{i} = \{\mathbf{u} + \mathbf{i}, \mathbf{u} \in \mathcal{V}\}$ for each $\mathbf{i} \in \mathbb{N}^N$, and d be the cardinal of $\mathcal{V}$ (d is also the cardinal of each $\mathcal{V}_{\mathbf{i}}$). To define a predictor of $\xi_{\mathbf{i}_0}$, let us consider the $\mathbb{R}^d$-valued random variables $\tilde{\xi}_{\mathbf{i}} = \{\xi_{\mathbf{u}}, \mathbf{u} \in \mathcal{V}_{\mathbf{i}} \subset \mathcal{O}_{\mathbf{n}}\}$. The notation of the previous sections are used by setting $X_{\mathbf{i}} = \tilde{\xi}_{\mathbf{i}}, Y_{\mathbf{i}} = \xi_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^N$.

As a predictor of $\xi_{\mathbf{i_0}}$, we take the conditional quantile estimate $\widehat{\xi}_{\mathbf{i_0}} = \mu_{p,\mathbf{n}}(\tilde{\xi}_{\mathbf{i_0}})$ of order p, particularly the conditional median $p = 0.5$. We deduce from the previous consistency results, the following corollary that gives the convergence of the predictor $\widehat{\xi}_{\mathbf{i_0}}$.

4

Corollary 2 *i) Under the conditions of Corollary 1, we have*

$$\boldsymbol{E}\left[\{\mu_{p,\mathbf{n}}(\tilde{\xi}_{\mathbf{i_o}}) - \mu_p(\tilde{\xi}_{\mathbf{i_o}})\}\mathbb{1}_{\{\tilde{\xi}_{\mathbf{i_o}} \in S\}}\right] \overset{a.s.}{\to} 0.$$

ii) Under the conditions of Theorem 3, and if $\widehat{n}h^{d+4} \to 0$, then

$$\sqrt{\widehat{\mathbf{n}}h^d}\left(\mu_{p,\mathbf{n}}(\tilde{\xi}_{\mathbf{i_o}}) - \mu_p(\tilde{\xi}_{\mathbf{i_o}})\right) \overset{\mathcal{L}}{\to} \mathcal{N}\left(0, \frac{\widehat{\sigma}^2(\tilde{\xi}_{\mathbf{i_o}}, \mu_p(\tilde{\xi}_{\mathbf{i_o}}))}{(f(\mu_p(\tilde{\xi}_{\mathbf{i_o}})|\tilde{\xi}_{\mathbf{i_o}}))^2}\right),$$

where

$$\widehat{\sigma}^2(x, y) = \frac{F_{\mathbf{n}}(y|x)\left[1 - F_{\mathbf{n}}(y|x)\right]}{g_{\mathbf{n}}(x)}\int_{\mathbb{R}^d} K^2(z)dz.$$

These consistency results permit to have an approximation of an $1 - \alpha$ confidence interval of $\xi_{\mathbf{i_o}}$ given by $\widehat{I}_\alpha = [a_-(\tilde{\xi}_{\mathbf{i_o}}), a_+(\tilde{\xi}_{\mathbf{i_o}})]$, where

$$a_\pm(\tilde{\xi}_{\mathbf{i_o}}) = \mu_{p,\mathbf{n}}(\tilde{\xi}_{\mathbf{i_o}}) \pm Q_{1-\frac{\alpha}{2}}\frac{\sigma(\tilde{\xi}_{\mathbf{i_o}}, \mu_{p,\mathbf{n}}(\tilde{\xi}_{\mathbf{i_o}}))}{\sqrt{\widehat{\mathbf{n}}h^d}f_{\mathbf{n}}(\mu_{p,\mathbf{n}}(\tilde{\xi}_{\mathbf{i_o}})|\tilde{\xi}_{\mathbf{i_o}})}, \tag{6}$$

where Q_ζ denotes the ζ-quantile of the standard normal distribution, and the unknown parameters (of the asymptotic variance in Corollary 2) are replaced by kernel estimates.

Note also that the quantiles of order p_1 and p_2 ($p_1 < p_2$) can be used to construct a predictive interval that consists of the $(p_2 - p_1)100\%$ confidence interval with bounds $\mu_{p_1,\mathbf{n}}(\tilde{\xi}_{\mathbf{i_o}})$ and $\mu_{p_2,\mathbf{n}}(\tilde{\xi}_{\mathbf{i_o}})$.

4   A simulation study

In this section, we study the performance of the conditional quantile predictor introduced in the previous section towards some simulations. Let us denoted by $GRF(m, \sigma^2, s)$ a Gaussian random field with mean m and covariance function defined by

$$\vartheta(h) = \sigma^2 \exp\left\{-\left(\frac{\|h\|}{s}\right)^2\right\}, h \in \mathbb{R}^2.$$

Set

$$\begin{aligned}
\mathcal{I}_{\mathbf{n}} &= \{\mathbf{i} = (i, j) \in (\mathbb{N}^*)^2, \ 1 \leq i \leq 61, \ 1 \leq j \leq 61\} \tag{7}\\
\mathcal{O}_{\mathbf{n}} &= \{\mathbf{i} = (i, j) \in (\mathbb{N}^*)^2, \ 1 \leq i \leq 21, \ 1 \leq j \leq 21\} \cup \{\mathbf{i} = (22, j), \ 1 \leq j \leq 15\}. \tag{8}
\end{aligned}$$

We consider a random field $(\xi_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}}$ from the following model

$$\xi_{\mathbf{i}} = U_{\mathbf{i}} * \left(\sin(2X_{\mathbf{i}}) + 2\exp\{-(16X_{\mathbf{i}})^2\}\right) + Z_{\mathbf{i}}, \ \mathbf{i} \in \mathbb{N}^2 \tag{9}$$

where $X = (X_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}}$ is a $GRF(0, 5, 3)$, $Z = (Z_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}}$ is a $GRF(0, 0.1, 5)$ independent of X and $U_{\mathbf{i}} = \frac{1}{\widehat{\mathbf{n}}}\sum_{\mathbf{j} \in \mathcal{I}_{\mathbf{n}}} \exp\left(-\frac{\|\mathbf{i}-\mathbf{j}\|}{2}\right)$. The choice of $U_{\mathbf{i}}$ in the model (9) is motivated by a reinforcement of the spatial local dependency. The field $(\xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{I}_{\mathbf{n}})$ is observable over the rectangular region $\mathcal{I}_{\mathbf{n}}$ and observed over the subset $\mathcal{O}_{\mathbf{n}}$ defined in (7) and (8).

We want to predict the values $\xi_{\mathbf{i}_1}, \ldots, \xi_{\mathbf{i}_m}$ at given fixed sites $\mathbf{i}_1, \ldots, \mathbf{i}_m$ not in $\mathcal{O}_{\mathbf{n}}$, with $m = 10$. The sample obtained from model (9), observed in $\mathcal{O}_{\mathbf{n}}$ is plotted in Figure 1 below with the 10 non observable values of the field at $\mathbf{i}_1, \ldots, \mathbf{i}_m$.
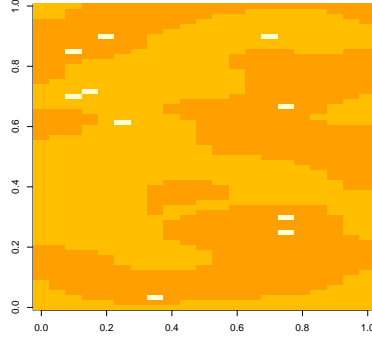
Figure 1: The random field $(\xi_{\mathbf{i}})$ with non observed values $\xi_{\mathbf{i}_1}, \ldots, \xi_{\mathbf{i}_{10}}$ in the white rectangular cases.

As explained in Section 3.1, for any $k \in \{1, \ldots, m\}$, we take the conditional quantile estimate $\widehat{\xi}_{\mathbf{i}_k} = \mu_{p,\mathbf{n}}(\tilde{\xi}_{\mathbf{i}_k})$ as a predictor of $\xi_{\mathbf{i}_k}$, where $\tilde{\xi}_{\mathbf{i}_k}$ are observed on $\mathcal{O}_{\mathbf{n}}$ and the vicinity $\mathcal{V} = \{-1, 1\} \times \{-1, 1\}$ or $\{-2, -1, 1, 2\} \times \{-2, -1, 1, 2\}$.

To compute $\mu_{p,\mathbf{n}}$, we select the standard normal density as kernel K and the Epanechnikov kernel as w. For the bandwidth selection, we use the rule developed in Yu and Jones [13],

$$h_{\mathbf{n}} = h_{\text{mean}} \left(\frac{p(1-p)}{\phi(\Phi^{-1}(p))^2} \right)^{1/5},$$

where h_{mean} is the bandwidth for kernel smoothing estimation of the regression mean, ϕ and Φ, are respectively, the standard normal density and distribution function.

To evaluate the performance of the predictor $\widehat{\xi}_{\mathbf{i}_k}$, we compute the mean absolute error (MAE):

$$MAE = \frac{1}{m} \sum_{k=1}^{m} |\widehat{\xi}_{\mathbf{i}_k} - \xi_{\mathbf{i}_k}|.$$

The following Table gives the predictors of $\xi_{\mathbf{i}_k}$, $k = 1, ..., m$ for $p \in \{0.05,\ 0.5,\ 0.95\}$, $\mathcal{V} = \{-1, 1\} \times \{-1, 1\}$ on the left, $\mathcal{V} = \{-2, -1, 1, 2\} \times \{-2, -1, 1, 2\}$ on the right and the prediction error.

Table 1: Predictive data for $\mathcal{V} = \{-1, 1\} \times \{-1, 1\}$ on the left and $\mathcal{V} = \{-2, -1, 1, 2\} \times \{-2, -1, 1, 2\}$ on the right.

	$p = 0.05$	True data	$p = 0.5$	$p = 0.95$	$p = 0.05$	True data	$p = 0.5$	$p = 0.95$
	0.1653	**0.2009**	0.1930	0.2192	0.1835	**0.2009**	0.2129	0.2362
	-0.2553	**−0.2315**	-0.2289	-0.1862	-0.2195	**−0.2315**	-0.1984	-0.1766
	0.1516	**0.1966**	0.1990	0.2362	0.1912	**0.1966**	0.2129	0.2362
	-0.5313	**−0.4906**	-0.5062	-0.4782	-0.5472	**−0.4906**	-0.5313	-0.5033
	0.2693	**0.2901**	0.2929	0.3168	0.2237	**0.2901**	0.2465	0.2676
	-0.2748	**−0.2535**	-0.2527	-0.2289	-0.2838	**−0.2535**	-0.2606	-0.2401
	0.3696	**0.3941**	0.4007	0.4269	0.3805	**0.3941**	0.3834	0.4269
	-0.5539	**−0.5177**	-0.5295	-0.5062	-0.5472	**−0.5177**	-0.5313	-0.5033
	-0.3678	**−0.3217**	-0.3463	-0.3193	-0.3637	**−0.3217**	-0.3487	-0.3231
	-0.2983	**−0.2843**	-0.2702	-0.2455	-0.3096	**−0.2843**	-0.2863	-0.2671
MAE	0.0308		0.0089	0.0252	0.0298		0.0206	0.0244

We derive from the results of Tables 1 a 90% predictive interval where the extremities are the 5% and 95% quantiles estimates, for each of the 10 prediction sites (see Section 3.1 for more details). Note that these 90% predictive intervals contain the true values. The average length for the 10 intervals is 0.05604 for $\mathcal{V} = \{-1, 1\} \times \{-1, 1\}$ and 0.04455 for $\mathcal{V} = \{-2, -1, 1, 2\} \times \{-2, -1, 1, 2\}$.

The numerical results show that our proposed predictor gives good results on the above simulated field. A next step would be to apply the predictor to a spatial real data and deserves futur investigations.

5   Appendix

The letter C will be used to denote constants whose values are unimportant. Before proving the main results, let us give the following notations:

$$\psi(x, y) = \int_{-\infty}^{y} f(x, z)dz, \quad K_1(y) = \int_{-\infty}^{y} w(z)dz.$$

Next, we need the following results.

Lemma 1 *Assume that (A3)-(A7) hold, then,*

$$\sup_{y \in \mathcal{V}} \sup_{x \in S} |\boldsymbol{E}\psi_{\mathbf{n}}(x, y) - \psi(x, y)| \quad = \quad O\left(\Omega_{\mathbf{n}}\right).$$

The proof of Lemma 1 is classical (see for example Matzner and Løber [10]).

Lemma 2 *Let Assumptions (A1)-(A6) hold. If (A9) and (1) or (A10) and (2) are satisfied, then,*

$$\sup_{y \in \mathcal{V}} \sup_{x \in S} |\psi_{\mathbf{n}}(x, y) - \boldsymbol{E}\psi_{\mathbf{n}}(x, y)| \quad = \quad O\left(\Omega_{\mathbf{n}}\right) \quad a.s.$$

Proof of Lemma 2. Define

$$
\begin{aligned}
Z_{\mathbf{i}}(x, y) &= \frac{1}{\widehat{\mathbf{n}}h^{d+1}}K\left(\frac{x - X_{\mathbf{i}}}{h}\right)\int_{-\infty}^{y} w\left(\frac{t - Y_{\mathbf{i}}}{h}\right)dt = \frac{1}{\widehat{\mathbf{n}}h^{d}}K\left(\frac{x - X_{\mathbf{i}}}{h}\right)K_1\left(\frac{y - Y_{\mathbf{i}}}{h}\right), \\
\Delta_{\mathbf{i}}(x, y) &= Z_{\mathbf{i}}(x, y) - \boldsymbol{E}Z_{\mathbf{i}}(x, y),
\end{aligned}
$$

then $S_{\mathbf{n}}(x, y) = \psi_{\mathbf{n}}(x, y) - \boldsymbol{E}\psi_{\mathbf{n}}(x, y) = \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i}}(x, y).$

Lemma 3 *Under Assumptions (A1), (A2), (A5) and (A6), we have*

$$\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \boldsymbol{E}(\Delta_{\mathbf{i}}(x, y))^2 + \sum_{\substack{\mathbf{i} \neq \mathbf{j} \\ \mathbf{i}, \mathbf{j} \in \mathcal{I}_{\mathbf{n}}}} \boldsymbol{E}\Delta_{\mathbf{i}}(x, y)\Delta_{\mathbf{j}}(x, y) = O\left(\frac{1}{\widehat{\mathbf{n}}h^d}\right).$$

The proof of Lemma 3 is similar to the proof of Lemma 2.2 in Tran [12].

Let us introduce a spatial block decomposition that has been used by Tran [12] and Carbon et al. [3]. Without loss of generality, assume that $n_i = 2pq_i$ for $1 \leq i \leq N$. The random variables $\Delta_{\mathbf{i}}(x, y)$ can be

grouped into $2^N q_1 \times q_2 \times \cdots \times q_N$ cubic blocks of side p. Denote

$$U^x(1,\mathbf{n},\mathbf{j},y) = \sum_{\substack{i_k=2j_kp+1 \\ k=1,\ldots,N}}^{(2j_k+1)p} \Delta_{\mathbf{i}}(x,y), \ U^x(2,\mathbf{n},\mathbf{j},y) = \sum_{\substack{i_k=2j_kp+1 \\ k=1,\ldots,N-1}}^{(2j_k+1)p} \sum_{i_N=(2j_N+1)p+1}^{2(j_N+1)p} \Delta_{\mathbf{i}}(x,y),$$

$$U^x(3,\mathbf{n},\mathbf{j},y) = \sum_{\substack{i_k=2j_kp+1 \\ k=1,\ldots,N-2}}^{(2j_k+1)p} \sum_{i_{N-1}=(2j_{N-1}+1)p+1}^{2(j_{N-1}+1)p} \sum_{i_N=2j_Np+1}^{(2j_N+1)p} \Delta_{\mathbf{i}}(x,y),$$

$$U^x(4,\mathbf{n},\mathbf{j},y) = \sum_{\substack{i_k=2j_kp+1 \\ k=1,\ldots,N-2}}^{(2j_k+1)p} \sum_{i_{N-1}=(2j_{N-1}+1)p+1}^{2(j_{N-1}+1)p} \sum_{i_N=(2j_N+1)p+1}^{2(j_N+1)p} \Delta_{\mathbf{i}}(x,y),$$

and so on. Finally, note that

$$U^x(2^{N-1},\mathbf{n},\mathbf{j},y) = \sum_{i_k=2j_kp+1}^{(2j_k+1)p} \sum_{i_N=2j_Np+1}^{(2j_N+1)p} \Delta_{\mathbf{i}}(x,y), \ \ U^x(2^N,\mathbf{n},\mathbf{j},y) = \sum_{\substack{i_k=2j_kp+1 \\ k=1,\ldots,N}}^{(2j_k+1)p} \Delta_{\mathbf{i}}(x,y).$$

For each integer $1 \le i \le 2^N$, let

$$T^x(\mathbf{n},i,y) = \sum_{\substack{j_k=0 \\ k=1,\ldots,N}}^{q_k-1} U^x(i,\mathbf{n},\mathbf{j},y), \ \ S_{\mathbf{n}}(x,y) = \sum_{i=1}^{2^N} T^x(\mathbf{n},i,y).$$

Observe that, for any $\varepsilon > 0$

$$\boldsymbol{P}\left(|S_{\mathbf{n}}(x,y)| > \varepsilon\right) = \boldsymbol{P}\left(\left|\sum_{i=1}^{2^N} T^x(\mathbf{n},i,y)\right| > \varepsilon\right) \le 2^N \boldsymbol{P}\left(|T^x(\mathbf{n},i,y)| > \varepsilon/2^N\right).$$

Without loss of generality, we consider just the case where $i = 1$ and we enumerate in an arbitrary way the $\widehat{q} = q_1 \ldots q_N$ terms $U^x(1,\mathbf{n},\mathbf{j},y)$ of the sum $T^x(\mathbf{n},1,y)$ that we call $W_1,\ldots,W_{\widehat{q}}$. Note that $U^x(i,\mathbf{n},\mathbf{j},y)$ is measurable with respect to the σ-field generated by $Z_{\mathbf{i}}$, with $\mathbf{i}$ such that $2j_kp + 1 \le i_k \le (2j_k + 1)p$, $k = 1,\ldots,N$.

These sets of sites are separeted by a distance at least p and since the $Z_{\mathbf{i}}$ are bounded, then we have for all $i = 1,\ldots,\widehat{q}$, $|W_i| \le C(\widehat{\mathbf{n}}h^d)^{-1}p^N\|K\|_\infty$. Lemma 4.4 in Carbon et al. [3] ensures that there exist independent random variables $W_1^*,\ldots,W_{\widehat{q}}^*$ such that for all $i = 1,\ldots,\widehat{q}$,

$$\boldsymbol{E}|W_i - W_i^*| \le C(\widehat{\mathbf{n}}h^d)^{-1}p^N\|K\|_\infty\phi(\widehat{\mathbf{n}},p^N)\chi(p).$$

Markov's inequality leads to

$$\boldsymbol{P}\left(\sum_{i=1}^{\widehat{q}}|W_i - W_i^*| > \varepsilon/2^{N+1}\right) \le C2^{N+1}(\widehat{\mathbf{n}}h^d)^{-1}p^N\widehat{q}\|K\|_\infty\phi(\widehat{\mathbf{n}},p^N)\varepsilon^{-1}\chi(p). \tag{10}$$

By Bernstein's inequality, we have

$$\boldsymbol{P}\left(|\sum_{i=1}^{\widehat{q}}W_i^*| > \varepsilon/2^{N+1}\right) \le 2\exp\left\{\frac{-\varepsilon^2/(2^{N+1})^2}{4\sum_{i=1}^{\widehat{q}}\boldsymbol{E}W_i^{*2} + 2C(\widehat{\mathbf{n}}h^d)^{-1}p^N\|K\|_\infty\varepsilon/2^{N+1}}\right\}. \tag{11}$$

8

Combining (10) and (11), we have

$$
\begin{aligned}
\boldsymbol{P}\left(|S_{\mathbf{n}}(x,y)| > \varepsilon\right) &\leq 2^N \boldsymbol{P}\left(\sum_{i=1}^{\widehat{q}} |W_i - W_i^*| > \varepsilon/2^{N+1}\right) + 2^N \boldsymbol{P}\left(|\sum_{i=1}^{\widehat{q}} W_i^*| > \varepsilon/2^{N+1}\right) \\
&\leq 2^{N+1} \exp\left\{\frac{-\varepsilon^2/(2^{N+1})^2}{4\sum_{i=1}^{\widehat{q}} \boldsymbol{E}W_i^{*2} + 2C(\widehat{\mathbf{n}}h^d)^{-1}p^N\|K\|_\infty \varepsilon/2^{N+1}}\right\} \\
&\quad + C2^{2N+1}\phi(\widehat{\mathbf{n}},p^N)(\widehat{\mathbf{n}}h^d)^{-1}p^N\widehat{q}\|K\|_\infty \varepsilon^{-1}\chi(p).
\end{aligned}
$$

By Lemma 3, one has $\sum_{i=1}^{\widehat{q}} \boldsymbol{E}W_i^{*2} = O(\frac{1}{\widehat{\mathbf{n}}h^d})$ and since $\widehat{\mathbf{n}} = 2^N p^N \widehat{q}$, we have

$$
\boldsymbol{P}\left(|S_{\mathbf{n}}(x,y)| > \varepsilon\right) \leq 2^{N+1} \exp\left\{\frac{-\varepsilon^2 \widehat{\mathbf{n}}h^d}{2^{2N+4}C + 2^{N+2}Cp^N\varepsilon}\right\} + C2^{N+1}\phi(\widehat{\mathbf{n}},p^N)h^{-d}\varepsilon^{-1}\chi(p).
$$

Let $\lambda > 0$ and set $\varepsilon = \lambda\Omega_{\mathbf{n}}$, $p = \Omega_{\mathbf{n}}^{-1/N}$. For the first part of Lemma 2, a simple computation shows that for sufficiently large $\widehat{\mathbf{n}}$,

$$
\begin{aligned}
\boldsymbol{P}\left(|S_{\mathbf{n}}(x,y)| > \lambda\Omega_{\mathbf{n}}\right) &\leq 2^{N+1} \exp\left\{\frac{-\lambda^2 \log\widehat{\mathbf{n}}}{2^{2N+4}C + 2^{N+2}C\lambda}\right\} + C2^{N+1}p^N h^{-d}\lambda^{-1}\Omega_{\mathbf{n}}^{-1}\chi(p) \\
&\leq C\widehat{\mathbf{n}}^{-b} + C\lambda^{-1}h^{-d}\Omega_{\mathbf{n}}^{\frac{\beta-2N}{N}}, \tag{12}
\end{aligned}
$$

with $b > 0$. Analogously, for the second part, as above we have

$$
\begin{aligned}
\boldsymbol{P}\left(|S_{\mathbf{n}}(x,y)| > \lambda\Omega_{\mathbf{n}}\right) &\leq 2^{N+1} \exp\left\{\frac{-\lambda^2 \log\widehat{\mathbf{n}}}{2^{2N+4}C + 2^{N+2}C\lambda}\right\} + C2^{N+1}\widehat{\mathbf{n}}^\kappa h^{-d}\lambda^{-1}\Omega_{\mathbf{n}}^{-1}\chi(p) \\
&\leq C\widehat{\mathbf{n}}^{-b} + C\lambda^{-1}\widehat{\mathbf{n}}^\kappa h^{-d}\Omega_{\mathbf{n}}^{\frac{\beta-N}{N}}. \tag{13}
\end{aligned}
$$

Now, set $R_{\mathbf{n}} = h^{d+1}\Omega_{\mathbf{n}}$ and $r_{\mathbf{n}} = \left(\frac{h^{d+2}}{\mathbf{n}}\right)^{1/2}$. Since S is a compact, it can be covered with $d_{\mathbf{n}}$ cubes B_k having sides of length $R_{\mathbf{n}}$ and center at x_k with $d_{\mathbf{n}} \leq CR_{\mathbf{n}}^{-d}$. The compact set $\mathcal{V}$ can be covered with $l_{\mathbf{n}}$ intervalls I_l having length $r_{\mathbf{n}}$ and center at y_l, with $l_{\mathbf{n}} \leq Cr_{\mathbf{n}}^{-1}$. We have

$$
d_{\mathbf{n}}l_{\mathbf{n}} \leq C\mathbf{n}^{1/2}h^{-\frac{d+2}{2}}\left(h^{d+1}\Omega_{\mathbf{n}}\right)^{-d-1}. \tag{14}
$$

Define $S_{1\mathbf{n}} = \sup_{y\in\mathcal{V}} \max_k \sup_{x\in B_k} |\psi_{\mathbf{n}}(x,y) - \psi_{\mathbf{n}}(x_k,y)|$

$$
\begin{aligned}
S_{2\mathbf{n}} &= \max_k \max_l \sup_{y\in I_l} |\psi_{\mathbf{n}}(x_k,y) - \psi_{\mathbf{n}}(x_k,y_l)|, \quad S_{3\mathbf{n}} = \max_k \max_l \sup_{y\in I_l} |\boldsymbol{E}\psi_{\mathbf{n}}(x_k,y_l) - \boldsymbol{E}\psi_{\mathbf{n}}(x_k,y)| \\
S_{4\mathbf{n}} &= \sup_{y\in\mathcal{V}} \max_k \sup_{x\in B_k} |\boldsymbol{E}\psi_{\mathbf{n}}(x_k,y) - \boldsymbol{E}\psi_{\mathbf{n}}(x,y)|, \quad S_{5\mathbf{n}} = \max_k \max_l |\psi_{\mathbf{n}}(x_k,y_l) - \boldsymbol{E}\psi_{\mathbf{n}}(x_k,y_l)|.
\end{aligned}
$$

Then, we can write $\sup_{y\in\mathcal{V}} \sup_{x\in S} |\psi_{\mathbf{n}}(x,y) - \boldsymbol{E}\psi_{\mathbf{n}}(x,y)| \leq S_{1\mathbf{n}} + S_{2\mathbf{n}} + S_{3\mathbf{n}} + S_{4\mathbf{n}} + S_{5\mathbf{n}}$. The proof of Lemma 2 follows easily from the combination of the two following lemmas.

Lemma 4 *Under Assumptions (A5) and (A6),*

$$
S_{i\mathbf{n}} = O\left(\Omega_{\mathbf{n}}\right) \quad a.s., \ i = 1,2,3,4.
$$

Lemma 5 *Assume that (A1), (A2), (A5) and (A6) hold. If (A9) and (1) or A(10) and (2) are satisfied, then*

$$
S_{5\mathbf{n}} = O\left(\Omega_{\mathbf{n}}\right) \quad a.s.
$$

Proof of Lemma 4. On one hand, since the kernel K satisfies the Lipschitz condition, we have clearly

$$|\psi_{\mathbf{n}}(x,y) - \psi_{\mathbf{n}}(x_k,y)| \leq C\widehat{\mathbf{n}}^{-1}h^{-d-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\|x-x_k\| \leq Ch^{-d-1}R_{\mathbf{n}} = O\left(\Omega_{\mathbf{n}}\right).$$

On the other hand, observe that $h^{-d-1}r_{\mathbf{n}} = \left(\widehat{\mathbf{n}}h^d\right)^{-1/2} = O\left(\Omega_{\mathbf{n}}\right)$. Since w satisfies the Lipschitz condition,

$$|\psi_{\mathbf{n}}(x_k,y) - \psi_{\mathbf{n}}(x_k,y_l)| \leq C\widehat{\mathbf{n}}^{-1}h^{-d-1}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}|y-y_l| \leq Ch^{-d-1}r_{\mathbf{n}} = O\left(\Omega_{\mathbf{n}}\right),$$

which gives the proof of Lemma 4.

Proof of Lemma 5. For $\varepsilon > 0$, we have

$$\boldsymbol{P}\left(\max_l\max_k|\psi_{\mathbf{n}}(x_k,y_l) - \boldsymbol{E}\psi_{\mathbf{n}}(x_k,y_l)| > \varepsilon\right) \leq l_{\mathbf{n}}d_{\mathbf{n}}\boldsymbol{P}\left(|\psi_{\mathbf{n}}(x_k,y_l) - \boldsymbol{E}\psi_{\mathbf{n}}(x_k,y_l)| > \varepsilon\right).$$

Setting $\varepsilon = \lambda\Omega_{\mathbf{n}}$ with $\lambda > 0$ and taking into account (12) and (13), it suffices to show that $l_{\mathbf{n}}d_{\mathbf{n}}\widehat{\mathbf{n}}^{-b}\widehat{\mathbf{n}}u(\mathbf{n}) \to 0$ and $l_{\mathbf{n}}d_{\mathbf{n}}h^{-d}\Omega_{\mathbf{n}}^{\frac{\beta-2N}{N}}\widehat{\mathbf{n}}u(\mathbf{n}) \to 0$, or $l_{\mathbf{n}}d_{\mathbf{n}}\widehat{\mathbf{n}}^{\kappa}h^{-d}\Omega_{\mathbf{n}}^{\frac{\beta-N}{N}}\widehat{\mathbf{n}}u(\mathbf{n}) \to 0$.
First, observe that condition $\widehat{\mathbf{n}}h^d \to \infty$ implies that $\widehat{\mathbf{n}} > Ch^{-d}$, so that

$$\widehat{\mathbf{n}}^{(d+1)(d+2)/2d} > Ch^{-(d+1)(d+2)/2}. \tag{15}$$

Using (14) and (15), we have

$$\begin{aligned}
l_{\mathbf{n}}d_{\mathbf{n}}\widehat{\mathbf{n}}^{-b}\widehat{\mathbf{n}}u(\mathbf{n}) &\leq Ch^{-(d+1)(d+2)/2}(\log\widehat{\mathbf{n}})^{-d/2}\widehat{\mathbf{n}}^{1/2}\widehat{\mathbf{n}}^{-b}\widehat{\mathbf{n}}^{d/2}\widehat{\mathbf{n}}u(\mathbf{n}) \\
&\leq C\widehat{\mathbf{n}}^{(d^2+3d+1)/d-b}(\log\widehat{\mathbf{n}})^{-d/2}u(\mathbf{n}),
\end{aligned}$$

which goes to 0 if $b > (d^2 + 3d + 1)/d$.
Next, again (14) and a computation show that

$$l_{\mathbf{n}}d_{\mathbf{n}}h^{-d}\Omega_{\mathbf{n}}^{\frac{\beta-2N}{N}}\widehat{\mathbf{n}}u(\mathbf{n}) \leq C\left[\widehat{\mathbf{n}}(\log\widehat{\mathbf{n}})^{\frac{\beta-N(d+2)}{N(d+5)-\beta}}u(\mathbf{n})^{\frac{2N}{N(d+5)-\beta}}h^{\frac{-N(d+1)(d+2)-d\beta}{N(d+5)-\beta}}\right]^{\frac{N(d+5)-\beta}{2N}},$$

which goes to 0 by Assumption (A9) and $\beta > N(d+5)$.
Analogously, (14) and a computation show that

$$l_{\mathbf{n}}d_{\mathbf{n}}\widehat{\mathbf{n}}^{\kappa}h^{-d}\Omega_{\mathbf{n}}^{\frac{\beta-N}{N}}\widehat{\mathbf{n}}u(\mathbf{n}) \leq C\left[\widehat{\mathbf{n}}(\log\widehat{\mathbf{n}})^{\frac{\beta-N(d+1)}{N(d+4+2\kappa)-\beta}}u(\mathbf{n})^{\frac{2N}{N(d+4+2\kappa)-\beta}}h^{\frac{-N(d^2+4d+2)-d\beta}{N(d+4+2\kappa)-\beta}}\right]^{\frac{N(d+4+2\kappa)-\beta}{2N}},$$

which goes to 0 by Assumption (A10) and $\beta > N(d+4+2\kappa)$. The conclusion of Lemma 5 follows from the Borel Cantelli's Lemma.

Proof of Theorems 1 and 2. First, from Carbon et al. [3], we have $\sup_{x\in S}|g_{\mathbf{n}}(x) - g(x)| = O\left(\Omega_{\mathbf{n}}\right)$ a.s. Now, a standard decomposition gives

$$\sup_{y\in\mathcal{V}}\sup_{x\in S}|F_{\mathbf{n}}(y|x) - F(y|x)|$$

$$\leq \frac{1}{\inf_{x\in S}g_{\mathbf{n}}(x)}\left\{\sup_{y\in\mathcal{V}}\sup_{x\in S}|\psi_{\mathbf{n}}(x,y) - \psi(x,y)| + \sup_{y\in\mathcal{V}}\sup_{x\in S}\left(F(y|x)|g(x) - g_{\mathbf{n}}(x)|\right)\right\}$$

$$\leq \frac{1}{\inf_{x\in S}g_{\mathbf{n}}(x)}\left\{\sup_{y\in\mathcal{V}}\sup_{x\in S}|\psi_{\mathbf{n}}(x,y) - \boldsymbol{E}\psi_{\mathbf{n}}(x,y)| + \sup_{y\in\mathcal{V}}\sup_{x\in S}|\boldsymbol{E}\psi_{\mathbf{n}}(x,y) - \psi(x,y)| + \sup_{x\in S}|g(x) - g_{\mathbf{n}}(x)|\right\}.$$

10

Since by (A1), $g_{\mathbf{n}}(x)$ is bounded away from 0, Theorem 1 follows from the preceding inequality, Lemmas 1 and 2. Next, from (A8), to prove Theorem 2, it suffices to show that $\sup_{x \in S} |F(\mu_{p,\mathbf{n}}(x)|x) - F(\mu_p(x)|x)| \to 0$. We have

$$
\begin{aligned}
|F(\mu_{p,\mathbf{n}}(x)|x) - F(\mu_p(x)|x)| &\leq |F(\mu_{p,\mathbf{n}}(x)|x) - F_{\mathbf{n}}(\mu_{p,\mathbf{n}}(x)|x)| + |F_{\mathbf{n}}(\mu_{p,\mathbf{n}}(x)|x) - F(\mu_p(x)|x)| \\
&\leq |F(\mu_{p,\mathbf{n}}(x)|x) - F_{\mathbf{n}}(\mu_{p,\mathbf{n}}(x)|x)| \leq \sup_{y \in \mathcal{V}} |F_{\mathbf{n}}(y|x) - F(y|x)|.
\end{aligned}
$$

Thus, $\sup_{x \in S} |F(\mu_{p,\mathbf{n}}(x)|x) - F(\mu_p(x)|x)| \leq \sup_{x \in S} \sup_{y \in \mathcal{V}} |F_{\mathbf{n}}(y|x) - F(y|x)|$, so that Theorem 2 follows from an application of Theorem 1.

Proof of Corollary 1. First, by Lyapounov's inequality, we have

$$
\boldsymbol{E}\left[\mu_{p,\mathbf{n}}(X_{\mathbf{n}}) - \mu_p(X_{\mathbf{n}})\right] \leq \boldsymbol{E}\left[|\mu_{p,\mathbf{n}}(X_{\mathbf{n}}) - \mu_p(X_{\mathbf{n}})|\right] \leq \left(\boldsymbol{E}\left[(\mu_{p,\mathbf{n}}(X_{\mathbf{n}}) - \mu_p(X_{\mathbf{n}}))^2\right]\right)^{1/2},
$$

so, we can write

$$
\boldsymbol{E}\left[(\mu_{p,\mathbf{n}}(X_{\mathbf{n}}) - \mu_p(X_{\mathbf{n}})) \, \mathbb{1}_{\{X_{\mathbf{n}} \in S\}}\right] \leq \left(\boldsymbol{E}\left[(\mu_{p,\mathbf{n}}(X_{\mathbf{n}}) - \mu_p(X_{\mathbf{n}}))^2 \mathbb{1}_{\{X_{\mathbf{n}} \in S\}}\right]\right)^{1/2}.
$$

Then we also have

$$
\boldsymbol{E}\left[(\mu_{p,\mathbf{n}}(X_{\mathbf{n}}) - \mu_p(X_{\mathbf{n}}))^2 \mathbb{1}_{\{X_{\mathbf{n}} \in S\}}\right] \leq \boldsymbol{E}\left[\sup_{x \in S} (\mu_{p,\mathbf{n}}(x) - \mu_p(x))^2\right].
$$

An integration by parts gives

$$
\boldsymbol{E}\left[\sup_{x \in S} (\mu_{p,\mathbf{n}}(x) - \mu_p(x))^2\right] \leq 2 \int_0^{\infty} v \boldsymbol{P}\left(\sup_{x \in S} |\mu_{p,\mathbf{n}}(x) - \mu_p(x)| > v\right) dv.
$$

Using Assumption (A8), we have that for $\mathbf{n}$ large enough,

$$
\begin{aligned}
\int_0^{\infty} v \boldsymbol{P}\left(\sup_{x \in S} |\mu_{p,\mathbf{n}}(x) - \mu_p(x)| > v\right) dv &\leq 2 \int_0^{\infty} t \boldsymbol{P}\left(\sup_{x \in S} |F(\mu_p(x)|x) - F_{\mathbf{n}}(\mu_p(x)|x)| > t\right) dt \\
&\leq 2 \int_0^{\infty} t \boldsymbol{P}\left(\sup_{x \in S} \sup_{y \in \mathcal{V}} |F(y)|x) - F_{\mathbf{n}}(y|x)| > t\right) dt,
\end{aligned}
$$

and Corollary 1 follows from Theorem 1.

Proof of Theorem 3. Since

$$
F(\mu_p(x)|x) = p = F_{\mathbf{n}}(\mu_{p,\mathbf{n}}(x)|x),
$$

by Taylor's expansion, we have

$$
\mu_{p,\mathbf{n}}(x) - \mu_p(x) = \frac{1}{f_{\mathbf{n}}(\mu_{p,\mathbf{n}}^*(x)|x)} \left[F(\mu_p(x)|x) - F_{\mathbf{n}}(\mu_p(x)|x)\right],
$$

where $\mu_{p,\mathbf{n}}^*(x)$ lies between $\mu_{p,\mathbf{n}}(x)$ and $\mu_p(x)$. To prove the asymptotic normality, it suffices to show that the numerator is normally distributed, and the denominator converges to $f(\mu_p(x)|x)$ in probability. We have the following propositions.

Proposition 1 *Under (A1)-(A6), (C1) and (C2), if there exists $c \geq 0$ such that $\widehat{\mathbf{n}} h^{d+4} \to c$, then*

$$
\sqrt{\widehat{\mathbf{n}} h^d} \, (F_{\mathbf{n}}(y|x) - F(y|x)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(cB(x,y), \sigma^2(x,y)\right),
$$

where $B(x,y)$ and $\sigma^2(x,y)$ are defined in (4) and (5).

11

Proposition 2 *Assume that (A1), (A3)-(A7) hold. If (A9) and (1) or (A10) and (2) are satisfied, then*

$$\sqrt{\frac{\widehat{\mathbf{n}}h^{d+1}}{\log\widehat{\mathbf{n}}}}\sup_{y\in\mathcal{V}}|f_{\mathbf{n}}(y|x)-f(y|x)|\overset{a.s.}{\rightarrow}0.$$

Proof of Proposition 1. Proposition 1 is a consequence of the two following lemmas.

Lemma 6 *If Assumptions (A1) and (A3)-(A6) are satisfied, then*

$$\boldsymbol{E}F_{\mathbf{n}}(y|x)-F(y|x)=\frac{h^2}{2}B(x,y)+o(h^2)+O\left(\frac{1}{\widehat{\mathbf{n}}h^d}\right),$$

where $B(x,y)$ is defined in (4).

The proof of Lemma 6 is classical and therefore is omitted (see Matzner-Løber [10]).

Lemma 7 *Assume that Assumptions (A1)-(A6), (C1) and (C2) hold, then*

$$\sqrt{\widehat{\mathbf{n}}h^d}\left(F_{\mathbf{n}}(y|x)-\boldsymbol{E}F_{\mathbf{n}}(y|x)\right)\overset{\mathcal{L}}{\rightarrow}\mathcal{N}\left(0,\sigma^2(x,y)\right),$$

where $\sigma^2(x,y)$ is defined in (5).

Proof of Lemma 7. Assume for the moment that for any pair $(c_1,c_2)\in\mathbb{R}^2$ with $c_1^2+c_2^2\neq0$,

$$\sqrt{\widehat{\mathbf{n}}h^d}\left[c_1\left(g_{\mathbf{n}}(x)-\boldsymbol{E}g_{\mathbf{n}}(x)\right)+c_2\left(\psi_{\mathbf{n}}(x,y)-\boldsymbol{E}\psi_{\mathbf{n}}(x,y)\right)\right]\overset{\mathcal{L}}{\rightarrow}\mathcal{N}\left(0,\sigma^{*2}\right),\qquad(16)$$

where $\sigma^{*2}=\left[c_1^2g(x)+c_2^2\psi(x,y)+2c_1c_2\psi(x,y)\right]\int_{\mathbb{R}^d}K^2(z)dz$. Now, set

$$W_{\mathbf{n}}\ =\ \frac{1}{g_{\mathbf{n}}(x)}\left[\psi_{\mathbf{n}}(x,y)-\boldsymbol{E}\psi_{\mathbf{n}}(x,y)\right]-\frac{\boldsymbol{E}\psi_{\mathbf{n}}(x,y)}{g_{\mathbf{n}}(x)\boldsymbol{E}g_{\mathbf{n}}(x)}\left[g_{\mathbf{n}}(x)-\boldsymbol{E}g_{\mathbf{n}}(x)\right],$$

$$W_{\mathbf{n}}^*\ =\ \frac{1}{g(x)}\left[\psi_{\mathbf{n}}(x,y)-\boldsymbol{E}\psi_{\mathbf{n}}(x,y)\right]-\frac{\psi(x,y)}{g^2(x)}\left[g_{\mathbf{n}}(x)-\boldsymbol{E}g_{\mathbf{n}}(x)\right].$$

On one hand, according to (16), we have $\sqrt{\widehat{\mathbf{n}}h^d}W_{\mathbf{n}}^*\overset{\mathcal{L}}{\rightarrow}\mathcal{N}\left(0,\sigma^2\right)$, with σ^2 being defined in (5).
On the other hand,

$$\sqrt{\widehat{\mathbf{n}}h^d}\left[W_{\mathbf{n}}-W_{\mathbf{n}}^*\right]$$
$$=\ \sqrt{\widehat{\mathbf{n}}h^d}\left\{\left[\frac{1}{g_{\mathbf{n}}(x)}-\frac{1}{g(x)}\right]\left[\psi_{\mathbf{n}}(x,y)-\boldsymbol{E}\psi_{\mathbf{n}}(x,y)\right]+\left[\frac{\psi(x,y)}{g^2(x)}-\frac{\boldsymbol{E}\psi_{\mathbf{n}}(x,y)}{g_{\mathbf{n}}(x)\boldsymbol{E}g_{\mathbf{n}}(x)}\right]\left[g_{\mathbf{n}}(x)-\boldsymbol{E}g_{\mathbf{n}}(x)\right]\right\}$$

which goes to 0 in probability.
Finally, we conclude with the decomposition

$$\sqrt{\widehat{\mathbf{n}}h^d}\left(F_{\mathbf{n}}(y|x)-\boldsymbol{E}F_{\mathbf{n}}(y|x)\right)=\sqrt{\widehat{\mathbf{n}}h^d}W_{\mathbf{n}}^*+\sqrt{\widehat{\mathbf{n}}h^d}\left[W_{\mathbf{n}}-W_{\mathbf{n}}^*\right]+O\left(\frac{1}{\sqrt{\widehat{\mathbf{n}}h^d}}\right).$$

Now, (16) is proved following the same lines as the proof of Theorem 3.1 in Tran [12].

Proof of Proposition 2. We have

$$\sup_{y\in\mathcal{V}}|f_{\mathbf{n}}(y|x)-f(y|x)|\ \leq\ \frac{1}{g_{\mathbf{n}}(x)}\sup_{y\in\mathcal{V}}|f_{\mathbf{n}}(x,y)-f(x,y)|+\frac{1}{g_{\mathbf{n}}(x)}\sup_{y\in\mathcal{V}}f(y|x)|g_{\mathbf{n}}(x)-g(x)|.$$

Hence to prove Proposition 2, it suffices to show that $\sup_{y\in\mathcal{V}}|f_{\mathbf{n}}(x,y)-f(x,y)|\overset{P}{\rightarrow}0$, this is proved by following the same lines as in the proof of Theorem 3.3 in Carbon et al. [3].

References

[1] Abdi, S., Dabo-Niang, S., Diop, A. and Abdi, A. (2009). Consistency of a nonparametric conditional quantile estimator for random fields. To appear in *Mathematical Methods of Statistics.*

[2] Biau, G. and Cadre, B. (2004). Nonparametric spatial regression, *Statistical Inference for Stochastic Processes,* **vol. 7**, pp. 327-349.

[3] Carbon, M., Tran, L.T. and Wu, B. (1997). Kernel density estimation for random fields, *Statistics and Probability Letters,* **vol. 36**, pp. 115-125.

[4] Collomb, G. (1980). Estimation non paramétrique de probabilités conditionnelles, *C. R. Acad. Sci. Paris Sér I Math.,* **vol. 291**, pp. 427-430.

[5] Dabo-Niang, S. and Yao, A. F. (2007). Spatial kernel regression estimation, *Mathematical Methods of Statistics,* **vol. 16**, Number 4, pp. 1-20.

[6] Gannoun, A., Saracco, J. and Yu, K. (2003). Nonparametric prediction by conditional median and quantiles, *Journal of Statistical Planning and Inference,* **vol. 117**, pp. 207-223.

[7] Hallin, M., Lu, Z. and Yu, K. (2009). Local linear spatial quantile regression, *Bernouilli,* **vol. 15**, Number 3, pp. 659-686.

[8] Laksaci, A. Lemdani, M. and Ould-Said E. (2009). A generalized L^1- approach for a kernel estimator of conditional quantile with functionnal regressors: Consistency and asymptotic normality. *Statistics and Probability Letters,* **vol. 79**, pp. 1065-1073.

[9] Laksaci, A. and Fouzia, M. (2009). Estimation non paramétrique de quantiles conditionnels pour des variables fonctionnelles spatialement dépendantes, *Comptes Rendus Mathematique,* **vol. 347**, pp. 1075-1080.

[10] Matzner-Løber, E. (1997). Prévison nonparamétrique des processus stochastiques. Thèse de l'Université Montpellier II.

[11] Roussas, G.G. (1969). Nonparametric estimation of the transition distribution function of a Markov process, *Ann. Math. Statist.,* **vol. 40**, pp. 1386-1400.

[12] Tran, L. T. (1990). Kernel density estimation on random fields, *Journal of Multivariate Analysis,* **vol. 34**, pp. 37-53.

[13] Yu, K. and Jones, M.C. (1998). Local linear quantile regression. *J. Amer. Statist. Assoc.* **vol. 93**, pp. 228-237.

[14] Zhou, Y. and Liang, H. (2000). Asymptotic Normality for L_1 Norm Kernel estimator of Conditional Median under $\alpha-$Mixing Dependence, *Journal of Multivariate Analysis,* **vol. 73**, pp. 136-154.