DOI 10.1007/s00477-009-0339-6

Spatial mode estimation for functional random fields with application to bioturbation problem

Sophie Dabo-Niang · Anne-Francoise Yao · Laura Pischedda · Philippe Cuny · Franck Gilbert

Abstract This work provides a useful tool to study the effects of bioturbation on the distribution of oxygen within sediments. We propose here heterogeneity measurements based on functional spatial mode. To obtain the mode, one usually needs to estimate the spatial probability density. The approach considered here consists in looking each observation as a curve that represents the history of the oxygen concentration at a fixed pixel.

Keywords Spatial mode · Density estimation · Functional data · Heterogeneity · Bioturbation

S. Dabo-Niang (✉)
Laboratory GREMARS-EQUIPPE, University Charles De Gaulle, Lille 3, Maison de la recherche, Domaine du pont de bois, BP 60149, 59653 Villeneuve d'ascq cedex, France
e-mail: sophie.dabo@univ-lille3.fr

A.-F. Yao · L. Pischedda · P. Cuny
Laboratoire LMGEM, University Aix-Marseille 2, Campus de Luminy, case 901, 13288 Marseille cedex 09, France
e-mail: anne-francoise.yao@univmed.fr

L. Pischedda
e-mail: laura.pischedda@univmed.fr

P. Cuny
e-mail: philippe.cuny@univmed.fr

F. Gilbert
Laboratoire ECOLAB (UMR 5245), CNRS, University Toulouse 3, INPT, 29 rue Jeanne Marvig, BP 24349, 31055 Toulouse, France

F. Gilbert
EcoLab, CNRS, 31055 Toulouse, France
e-mail: franck.gilbert@cict.fr

1 Introduction

In aquatic ecosystems a crucial control on carbon processing is exerted by the animals inhabiting the sediments, which extensively rework and oxygenate sediments while e.g. feeding and moving, a global process referred to as bioturbation. In particular, macrofauna, through their feeding, burrowing and ventilation activities, have an important influence on microbial activity and sediment metabolism in marine sediments (Aller and Aller 1998; Kristensen 2001). Subduction of oxygen into and removal of metabolites from otherwise anoxic sediments, as well as relocation of organic particles by the infauna, create a heterogeneity, which can result in several-fold enhanced rates of organic matter decomposition and solute fluxes (Aller and Aller 1998; Kristensen and Holmer 2001, Nielsen et al. 2003). Overall, bioturbation plays a key role in the ecosystem functioning. One way to investigate this role is to study the related oxygen heterogeneity (Pischedda et al. 2008). For this purpose the use of planar optodes makes possible high resolution measurement of two-dimensional vertical distribution and spatial heterogeneity of oxygen in the sedimentary column (Glud et al. 1996; Hulth et al. 2002).

In the frame of our study, we have used such an optical method to collect 2-D oxygen distribution on different times in order to quantify the spatio-temporal dynamics of the oxygen concentration. More precisely, the dataset is composed of 121 images (sequence of pixels) of the intermittently irrigated U-shaped burrow of the polychaete worm *Nereis diversicolor*, which were recorded every 2 min during 4 h (See Fig. 1).

Now that such complex spatio-temporal data can be produced, bioturbation quantification strongly requires tools that can take into account the specificity of the data

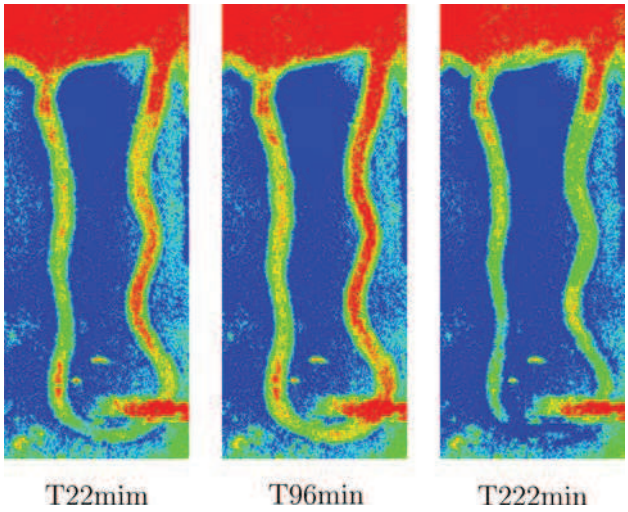

Fig. 1 Example of images on 3 dates

that (then) constitute a time series set of images. Such a tool could be very useful for example to compare the effects of bioturbation on oxygen distribution (i) in different regions within the sediments or (ii) as a function of various macrofauna species that differ in the way to provide oxygen into the sediments (e.g. Foster and Graf 1995).

In this paper, we propose to quantify spatial heterogeneity induced by a burrow-irrigating organism by means of a measure based on functional spatial mean, median and mode estimation. Indeed, the data can be seen as spatially dependant curves since for each pixel, $\mathbf{i}$, we can summary the evolution of the oxygen concentration on this pixel on a curve: $X_{\mathbf{i}} = (X_{\mathbf{i}}(t),\ 0 \leq t \leq 120)$; ($X_{\mathbf{i}}(t)$ being the oxygen concentration at the date t). If the estimation of the spatial mean and the median is relatively simple, the mode estimation is based on the estimation of a probability density for spatial functional random fields proposed by Dabo-Niang and Yao (2008). This last is a recent theoretical contribution on a crucial field, spatial modeling of functional data. Indeed, spatial modeling of functional or non-functional data has become one the most interesting and important areas of natural sciences, see for example Fernández de Castro and González Manteiga (2008), Angulo and Ruiz-Medina (2008), Porcu et al. (2008)... Note that the approach proposed here is complementary with the study done by Pischedda et al. (2008) that presented an estimation of oxygen heterogeneity in bioturbated sediments using a simple index based on oxygen horizontal variations. We will measure the spatial dependency by means of mixing condition and give theoretical results (weak convergence) concerning the mode estimate on Sect. 2. On Sect. 3 we give the heterogeneity measurements derived from the mode estimate. The last part is devoted to conclusion and discussion.

2 Spatial mode estimate for curves

We deal with a measurable strictly stationary spatial process $\left(X_{\mathbf{i}}, \mathbf{i} \in (\mathbb{N}^*)^N\right), N \geq 1$, defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ such that the $X_{\mathbf{i}}$'s have the same distribution as a variable X with values in an infinite dimensional separable semi-metric space $(\mathcal{E}, d)$ ($d(.,.)$ is the semi-metric). We assume that X has an unknown density f with respect to some given measure μ. We aim to estimate the spatial density from data, $X_{\mathbf{i}}$, observed on some rectangular region $\mathcal{I}_{\mathbf{n}} = \left\{\mathbf{i} \in \mathbb{N}^N : 1 \leq i_k \leq n_k, k = 1, \ldots, N\right\}$ where $\mathbf{n} = (n_1,\ldots,n_N)$. We write $\mathbf{n} \to +\infty$ if $\min_{k=1,\ldots,N} n_k \to +\infty$ and $\left|\frac{n_j}{n_k}\right| < C$ for some constant $0 < C < \infty$ and $\forall j,\ k \in \{1,\ldots,N\}$. We will set $\widehat{\mathbf{n}} = n_1 \times \cdots \times n_N$.

As in the *i.i.d.* case, in order to control the size of the set $\mathcal{C}$, in which we look for the mode, we choose this set such that $\mathcal{C} \subset \mathcal{C}_{\mathbf{n}} = \bigcup_{k=1}^{d_{\mathbf{n}}} B(x_k, r_{\mathbf{n}})$, where $d_{\mathbf{n}} > 0$ is some integer and for $k = 1, \ldots, d_{\mathbf{n}}$, $B(x_k, r_{\mathbf{n}})$ is the opened ball of center $x_k \in \mathcal{E}$ and radius $r_{\mathbf{n}} > 0$. Really, the set $\mathcal{C}_{\mathbf{n}}$, which can always be built, is here to ensure the existence of the set $\mathcal{C}$.

We assume that the mode of f:

$$\omega = \arg\sup_{\overset{\circ}{\mathcal{C}}} f \tag{1}$$

exists, where $\overset{\circ}{\mathcal{C}}$ is the interior of $\mathcal{C}$. We will set $\mathcal{A}_\epsilon = \{x \in \mathcal{C}, f(\omega) - f(x) < \epsilon\}$.

We are interested in this section to introduce an estimate of ω derived from an estimate of f. Hence, we need to study the uniform consistency over $\mathcal{C}$ of the estimate of the density.

For a seek of simplicity, we look at some special case where the probability distribution of X satisfies some concentration condition (see Dabo-Niang et al. (2006)) and when the density estimate is of the simple usual Parzen-Rosenblatt form.

Let $\psi(.)$ be some increasing function taking values in $]0, +\infty[$ such that $\lim_{t \to 0} \psi(t) = 0$. Then, we propose the following kernel density estimator:

$$\forall x \in \mathcal{E},\ f_{\mathbf{n}}(x) = \frac{1}{\widehat{\mathbf{n}} C(K, \psi, h_{\mathbf{n}})} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K(d(X_{\mathbf{i}}, x)/h_{\mathbf{n}}), \tag{2}$$

where $C(K, \psi, h_{\mathbf{n}}) = -\int_0^1 K'(t)\psi(h_{\mathbf{n}}t)dt$ and where $h_{\mathbf{n}}$ is a sequence of positive numbers that converges to zero. Note that this constant $C(K, \psi, h_{\mathbf{n}})$ does not depend on x. So, the functional mode estimate $\hat{\omega}_n$ can be defined as any solution of the equation

$$\hat{\omega}_{\mathbf{n}} = \arg\sup_{\mathcal{C}} f_{\mathbf{n}}, \tag{3}$$

which is very easy to calculate in practice. To give an *almost sure convergence* result of $\hat{\omega}_{\mathbf{n}}$, we introduce the following assumptions.

2.1 Assumptions

For seek of simplicity, we consider only kernels satisfying the following classical assumptions:

$$H1 - supp(K) = (0,1), K(1) = 0 \quad \text{and} \quad -\infty < \tau_1 \le K' \le \tau_2 < 0.$$

$$H2 - d_{\mathbf{n}} = \widehat{\mathbf{n}}^{\beta} \quad \text{and} \quad r_{\mathbf{n}} \le \left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}(\psi(h_{\mathbf{n}}))^{\kappa}}\right)^{\frac{1}{2}}$$
where $\kappa < -2\beta_1 + 1, \beta > 0, \beta_1 > 1$.

The nonparametric model is defined by means of the next assumption on the density function f, which is basically the same as in the standard multivariate case (see Abraham et al. 2003):

$$H3 - f \text{ is uniformly continuous on } \mathcal{C},\ \inf_{x\in C} f(x) > 0 \text{ and } \lim_{\epsilon\to 0} diameter(\mathcal{A}_{\epsilon}) = \sup_{x\in\mathcal{A}_{\epsilon}, y\in\mathcal{A}_{\epsilon}} d(x,y) = 0.$$

We also need some standard assumptions on f and ψ used in nonparametric functional mode estimation model (see Dabo-Niang et al. 2006).

$$H4 - \exists c > 0, \quad \exists \varepsilon_0 > 0, \quad \forall \varepsilon < \varepsilon_0, \quad \int_0^{\varepsilon} \psi(z)dz > c\varepsilon\psi(\varepsilon).$$

$$H5 - \lim_{t\to 0} \sup_{x\in\mathcal{C}} \left| \frac{P(X \in B(x,t))}{\psi(t)} - f(x) \right| = 0.$$

2.2 Dependency conditions

As it often occurs in spatial dependent data analysis, one needs to defined the type of dependence. Here, we will consider the following two dependence measures.

2.2.1 Local dependence condition

We will assume that the joint probability density $f_{\mathbf{i},\mathbf{j}}(.,.)$ of $(X_{\mathbf{i}}, X_{\mathbf{j}})$ (with respect to $\mu \times \mu$) exists and satisfies

$$\left| f_{\mathbf{i},\mathbf{j}}(x,y) - f(x)f(y) \right| \le C, \tag{4}$$

for some constant C and for all $x, y \in \mathcal{E}$ and $\mathbf{i}, \mathbf{j} \in \mathbb{N}^N, \mathbf{i} \ne \mathbf{j}$, or

$$\exists \epsilon_1 \in (0,1], \lim_{t\to 0} \sup_{x\in\mathcal{C}} \left| \frac{\max_{\mathbf{i}\ne\mathbf{j}} P((X_{\mathbf{i}}, X_{\mathbf{j}}) \in B(x,t)\times B(x,t))}{\psi(t)^{1+\epsilon_1}} - \sup_{\mathbf{i},\mathbf{j}} f_{\mathbf{i},\mathbf{j}}(x,x) \right| = 0. \tag{5}$$

Such local dependency condition is necessary to reach the same rate of convergence as in the *i.i.d* case.

2.2.2 Mixing conditions

Another complementary dependency condition concerned the *mixing condition* which measures the dependency by means of α-mixing. We assume that $\left(X_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^N\right)$ satisfies the following mixing condition: there exists a function $\varphi(t) \downarrow 0$ as $t \to \infty$, such that for E, E' subsets of $\mathbb{N}^N$ with finite cardinals,

$$\alpha\left(\mathcal{B}(E), \mathcal{B}\left(E'\right)\right) = \sup_{B\in\mathcal{B}(E), C\in\mathcal{B}\left(E'\right)} |\mathbf{P}(B\cap C) - \mathbf{P}(B)\mathbf{P}(C)| \le \chi\left(\mathrm{Card}(E), \mathrm{Card}\left(E'\right)\right)\varphi\left(\mathrm{dist}\left(E, E'\right)\right), \tag{6}$$

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

$$\chi(n,m) \le C\min(n,m), \quad \forall n, m \in \mathbb{N} \tag{7}$$

or

$$\chi(n,m) \le C(n+m+1)^{\widetilde{\beta}}, \quad \forall n, m \in \mathbb{N} \tag{8}$$

for some $\widetilde{\beta} \ge 1$ and some $C > 0$. If $\chi \equiv 1$, then, the process $(X_{\mathbf{i}})$ is said to be strongly mixing. Many stochastic processes, among them various useful time series models satisfy strong mixing properties, which are relatively easy to check. Conditions (7–8) are weaker than strong mixing condition and have been used for finite dimensional variables in (for example) Tran (1990), Carbon et al. (1996, 1997) and Biau and Cadre(2004). We refer to Doukhan (1994) and Rio (2000) for discussion on mixing and examples.

Concerning the function $\varphi(.)$, as it is often done, two kind of conditions will be assumed: the case where $\varphi(i)$ tends to zero at a polynomial rate, i.e.

$$\varphi(i) \le Ci^{-\theta}, \quad \text{for some } \theta > 0 \tag{9}$$

or the case where $\varphi(i)$ tends to zero at an exponential rate:

$$\varphi(i) = C\exp(-si), \quad \text{for some } s > 0. \tag{10}$$

Remark 1 The two dependence measures are link (see Bosq 1998 for details).

2.3 A consistency result

Under unrestrictive hypotheses on the probability of small balls, on the kernel, the bandwidth, the local ((4) or (5)) and the global spatial dependence condition (6), the probability convergence of the mode estimate $\hat{\theta}_{\mathbf{n}}$ follows in the two cases of mixing (9) and (10).

Let

$$\theta_1 = \frac{\theta}{2N(\beta+1)-\theta}, \quad \theta_2 = \frac{\theta - 2N}{2N(\beta+1)-\theta},$$
$$\theta_3 = \frac{\theta+N}{N(1+2\beta+2\widetilde{\beta})-\theta}, \quad \theta_4 = \frac{\theta-N}{N(1+2\beta+2\widetilde{\beta})-\theta}.$$

We assume for the following two theorems that $\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})/(\log\widehat{\mathbf{n}}) \to \infty$.

Theorem 1 *Suppose that H1–H5 are satisfied, the mixing coefficient satisfies (7) and (9) with* $\theta > 2N(\beta+1)$, $\widehat{\mathbf{n}}\,\psi(h_{\mathbf{n}})^{-\theta_1}(\log\widehat{\mathbf{n}})^{\theta_2} \to \infty$, *then we have*

$$\lim_{\mathbf{n}\to\infty} \hat{\omega}_{\mathbf{n}} = \omega, \text{ in probability.} \tag{11}$$

Proof We have from assumption *H*3 that:

$$\forall\varepsilon > 0, \exists\eta > 0;\ \forall x \in \mathcal{C}, d(\omega, x) \geq \varepsilon \Rightarrow |f(\omega) - f(x)| \geq \eta.$$

Hence, to get (11), it suffices to prove that $\lim_{\mathbf{n}\to\infty} \sup_{x\in\mathcal{C}} |f_{\mathbf{n}}(x) - f(x)| = 0$ a.s., since

$$\begin{aligned}|f(\hat{\omega}_{\mathbf{n}}) - f(\omega)| &\leq |f(\hat{\omega}_{\mathbf{n}}) - f_{\mathbf{n}}(\hat{\omega}_{\mathbf{n}})| + |f_{\mathbf{n}}(\hat{\omega}_{\mathbf{n}}) - f(\omega)| \\ &\leq \sup_{x\in\mathcal{C}} |f(x) - f_{\mathbf{n}}(x)| + \left|\sup_{x\in\mathcal{C}} f_{\mathbf{n}}(x) - \sup_{x\in\mathcal{C}} f(x)\right| \\ &\leq 2\sup_{x\in\mathcal{C}} |f_{\mathbf{n}}(x) - f(x)|.\end{aligned}$$

The consistency of the bias $\sup_{x\in\mathcal{C}} |E(f_{\mathbf{n}}(x)) - f(x)|$ is the same as in the i.i.d case of Dabo-Niang et al. (2006).

The main difference between this theoretical part and the i.i.d case of Dabo-Niang et al. (2006) comes from the proof of the consistency of the variance term $\sup_{x\in\mathcal{C}} |f_{\mathbf{n}}(x) - E(f_{\mathbf{n}}(x))|$. The extension of the variance term result of Dabo-Niang et al. (2006) to the spatial case is far from being trivial and is proved in the following. We set:

$$Q_{\mathbf{n}}(x) = f_{\mathbf{n}}(x) - E(f_{\mathbf{n}}(x)) = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} Z_{\mathbf{i},\mathbf{n},x}, \quad x \in \mathcal{C},$$

where

$$Z_{\mathbf{i},\mathbf{n},x} = \frac{1}{\widehat{\mathbf{n}}C(K,\psi,h_{\mathbf{n}})}\left(K(d(X_{\mathbf{i}},x)/h_{\mathbf{n}}) - EK(d(X_{\mathbf{i}},x)/h_{\mathbf{n}})\right).$$

It is proved in Dabo-Niang et al. (2006) that $\frac{K(d(X_{\mathbf{i}},x)/h_{\mathbf{n}})}{C(K,\psi,h_{\mathbf{n}})} \leq C/\psi(h_{\mathbf{n}})$.

Recall that $\mathcal{C}$ is covered by $d_{\mathbf{n}}$ balls $B_k = B(x_k, r_{\mathbf{n}})$ of radius $r_{\mathbf{n}}$ and center x_k.

Define

$$\begin{aligned}S_{1\mathbf{n}} &= \max_{1\leq k\leq d_{\mathbf{n}}} \sup_{x\in B_k} |f_{\mathbf{n}}(x) - f_{\mathbf{n}}(x_k)|, \\ S_{2\mathbf{n}} &= \max_{1\leq k\leq d_{\mathbf{n}}} \sup_{x\in B_k} |Ef_{\mathbf{n}}(x_k) - Ef_{\mathbf{n}}(x)|, \\ S_{3\mathbf{n}} &= \max_{1\leq k\leq d_{\mathbf{n}}} |f_{\mathbf{n}}(x_k) - Ef_{\mathbf{n}}(x_k)|.\end{aligned}$$

Then,

$$\sup_{x\in\mathcal{C}} |f_{\mathbf{n}}(x) - Ef_{\mathbf{n}}(x)| \leq S_{1\mathbf{n}} + S_{2\mathbf{n}} + S_{3\mathbf{n}}.$$

Using assumptions H1 and H2, one can easily show that $S_{1\mathbf{n}}$ and $S_{2\mathbf{n}}$ are equal to $o\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})}}\right)$ a.s., see proof of Theorem 3 of Dabo-Niang et al. (2006). It remains to study the consistency of $S_{3\mathbf{n}} = \max_{1\leq j\leq d_{\mathbf{n}}} |Q_{\mathbf{n}}(x_j)|$, by setting without loss of generality that for $1,\ldots,N$, $n_i = 2pt_i$ for some integers $p \geq 1$ and $t_1,\ldots,t_N$ and using the well known spatial block decomposition of Tran (1990). We group the random variables $Z_{\mathbf{i}}$, $\mathbf{n}$, x into small and large blocks of different sizes as follow:

$$\begin{aligned}U(1,\mathbf{n},x,\mathbf{j}) &= \sum_{i_k=2j_kp+1,1\leq k\leq N}^{(2j_k+1)p} Z_{\mathbf{i},\mathbf{n},x}, \\ U(2,\mathbf{n},x,\mathbf{j}) &= \sum_{i_k=2j_kp+1,1\leq k\leq N-1}^{(2j_k+1)p}\ \sum_{i_N=(2j_N+1)p+1}^{2(j_N+1)p} Z_{\mathbf{i},\mathbf{n},x}, \\ U(3,\mathbf{n},x,\mathbf{j}) &= \sum_{i_k=2j_kp+1,1\leq k\leq N-2}^{(2j_k+1)p}\ \sum_{i_{N-1}=(2j_{N-1}+1)p+1}^{2(j_{N-1}+1)p} \\ &\quad\times \sum_{i_N=2j_Np+1}^{(2j_N+1)p} Z_{\mathbf{i},\mathbf{n},x}, \\ U(4,\mathbf{n},x,\mathbf{j}) &= \sum_{i_k=2j_kp+1,1\leq k\leq N-2}^{(2j_k+1)p}\ \sum_{i_{N-1}=(2j_{N-1}+1)p+1}^{2(j_{N-1}+1)p} \\ &\quad\times \sum_{i_N=(2j_N+1)p+1}^{2(j_N+1)p} Z_{\mathbf{i},\mathbf{n},x}, \\ &\vdots \\ U(2^{N-1},\mathbf{n},x,\mathbf{j}) &= \sum_{i_k=(2j_k+1)p+1,1\leq k\leq N-1}^{2(j_k+1)p}\ \sum_{i_N=2j_Np+1}^{(2j_N+1)p} Z_{\mathbf{i},\mathbf{n},x},\end{aligned}$$

and

$$U(2^N,\mathbf{n},x,\mathbf{j}) = \sum_{i_k=(2j_k+1)p+1,1\leq k\leq N}^{2(j_k+1)p} Z_{\mathbf{i},\mathbf{n},x}.$$

Set $\mathcal{T} = \{0,\ldots,t_1-1\}\times\cdots\times\{0,\ldots,t_N-1\}$, and let for each integer $l = 1,\ldots,2^N$,

$$T(\mathbf{n},x,l) = \sum_{\mathbf{j}\in\mathcal{T}} U(l,\mathbf{n},x,\mathbf{j}).$$

Then, we obtain the following decomposition

$$Q_{\mathbf{n}}(x) = f_{\mathbf{n}}(x) - Ef_{\mathbf{n}}(x) = \sum_{l=1}^{2^N} T(\mathbf{n},x,l).$$

To prove that $S_{3\mathbf{n}} = O\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})}}\right)$ a.s., it is sufficient to show that, for a given arbitrary large positive constant c, there exists a positive constant C such that for any $\eta > 0$

$$P\left[\max_{1\leq j\leq d_{\mathbf{n}}} \left|T(\mathbf{n},x_j,1)\right| > \eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})}}\right] \leq Cd_{\mathbf{n}}(\widehat{\mathbf{n}}^{-c} + \beta_{1\widehat{\mathbf{n}}}).$$

Without loss of generality we will show this for $l = 1$.

Set $\epsilon_{\mathbf{n}} = \eta\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})}}$ (where $\eta > 0$ is a constant to be chosen later) and $\beta_{1\widehat{\mathbf{n}}} = \psi(h_{\mathbf{n}})^{-1}\chi(\widehat{\mathbf{n}}, p^N)\varphi(p)\epsilon_{\mathbf{n}}^{-1}$.

Let

$$T(\mathbf{n}, x, 1) = \sum_{\mathbf{j}\in\mathcal{T}} U(1, \mathbf{n}, x, \mathbf{j}),$$

be the sum of $\widehat{\mathbf{t}} = t_1 \times \cdots \times t_N$ of the $U(1, \mathbf{n}, x, \mathbf{j})$'s. Note that each $U(1, \mathbf{n}, x, \mathbf{j})$'s is measurable with respect to the σ-field generated by $X_{\mathbf{i}}$ with $\mathbf{i}$ belonging to the set of sites

$$\mathcal{I}_{\mathbf{i},\mathbf{j}} = \{\mathbf{i} : 2j_k p + 1 \le i_k \le (2j_k + 1)p, k = 1, \ldots, N\}.$$

These sets of sites are separated by a distance greater than p. Enumerate the random variables's $U(1, \mathbf{n}, x, \mathbf{j})$ and the corresponding σ-field with which they are measurable in an arbitrary manner and refer to them respectively as $V_1, \ldots, V_{\widehat{\mathbf{t}}}$ and $\mathcal{B}_1, \ldots, \mathcal{B}_{\widehat{\mathbf{t}}}$. Then, since $T(\mathbf{n}, x, 1) = \sum_{i=1}^{\widehat{\mathbf{t}}} V_i$ with

$$|V_i| = |U(1, \mathbf{n}, x, \mathbf{j})| < Cp^N(\widehat{\mathbf{n}}\psi(h_{\mathbf{n}}))^{-1}. \tag{12}$$

Lemma 4.5 of Carbon et al. (1997) allows us to approximate $V_1, \ldots, V_{\widehat{\mathbf{t}}}$ by $V_1^*, \ldots, V_{\widehat{\mathbf{t}}}^*$ such that:

$$P[|T(\mathbf{n}, x, 1)| > \epsilon_{\mathbf{n}}] \le P\left[\left|\sum_{i=1}^{\widehat{\mathbf{t}}} V_i^*\right| > \epsilon_{\mathbf{n}}/2\right] + P\left[\sum_{i=1}^{\widehat{\mathbf{t}}} |V_i - V_i^*| > \epsilon_{\mathbf{n}}/2\right]. \tag{13}$$

Now, using: Markov's inequality, (12), Lemma 4.4 of Carbon et al. (1997) and the fact that the sets of sites (with respect to which V_i's are measurable) are separated by a distance greater than p, we get:

$$P\left[\sum_{i=1}^{\widehat{\mathbf{t}}} |V_i - V_i^*| > \epsilon_{\mathbf{n}}\right] \le C\widehat{\mathbf{t}}p^N(\widehat{\mathbf{n}}\psi(h_{\mathbf{n}}))^{-1} \times \chi(\widehat{\mathbf{n}}, p^N)\varphi(p)\epsilon_{\mathbf{n}}^{-1} \sim \beta_{1\widehat{\mathbf{n}}}. \tag{14}$$

Let

$$\lambda_{\mathbf{n}} = (\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})\log\widehat{\mathbf{n}})^{1/2}, \tag{15}$$

then set,

$$p = \left[\left(\frac{\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})}{4\lambda_{\mathbf{n}}}\right)^{1/N}\right] \sim \left(\frac{\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})}{\log\widehat{\mathbf{n}}}\right)^{1/2N}, \tag{16}$$

and $\lambda_{\mathbf{n}}\epsilon_{\mathbf{n}} = \eta\log\widehat{\mathbf{n}}$.

If (9) holds for $\theta > 2N$, then one can prove (by following the same steps as those of the proof of Lemma 2.2 of Tran (1990)) by using our hypotheses H1, H4, H5 and Lemmas 4.2 and 4.3 of Carbon et al. (1997), that

$$\lambda_{\mathbf{n}}^2 \sum_{i=1}^{\widehat{\mathbf{t}}} E(V_i^*)^2 \le C\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})(U_{\mathbf{n}(x)} + R_{\mathbf{n}(x)})\log\widehat{\mathbf{n}} < C\log\widehat{\mathbf{n}}$$

where

$$U_{\mathbf{n}(x)} = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} E(Z_{\mathbf{i},\mathbf{n},x})^2$$

$$R_{\mathbf{n}(x)} = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \sum_{\mathbf{l}\in\mathcal{I}_{\mathbf{n}}, i_k\neq l_k \text{for some k}} |Cov(Z_{\mathbf{i},\mathbf{n},x}, Z_{\mathbf{l},\mathbf{n},x})|,$$

C is a constant independent of $x \in \mathcal{C}$. Using (12), we get $|\lambda_{\mathbf{n}}V_i^*| < 1/2$ for large $\widehat{\mathbf{n}}$ and deduce from Bernstein's inequality that

$$P\left[\left|\sum_{i=1}^{\widehat{\mathbf{t}}} V_i^*\right| > \epsilon_{\mathbf{n}}\right] \le 2\exp(-\lambda_{\mathbf{n}}\epsilon_{\mathbf{n}} + \lambda_{\mathbf{n}}^2 \sum_{i=1}^{\widehat{\mathbf{t}}} E(V_i^*)^2) \le 2\exp((-\eta + C)\log\widehat{\mathbf{n}}) \le \widehat{\mathbf{n}}^{-c} \tag{17}$$

for sufficiently large $\widehat{\mathbf{n}}$. We get from (13), (14) and (17) that

$$P[\max_{1\le j\le d_{\mathbf{n}}} |T(\mathbf{n}, x_j, 1)| > \epsilon_{\mathbf{n}}] \le Cd_{\mathbf{n}}(\widehat{\mathbf{n}}^{-c} + \beta_{1\widehat{\mathbf{n}}}).$$

To prove the consistency variance result it suffices to show that $d_{\mathbf{n}}\widehat{\mathbf{n}}^{-c} \to 0$ and $d_{\mathbf{n}}\beta_{1\widehat{\mathbf{n}}} \to 0$. We have $d_{\mathbf{n}}\widehat{\mathbf{n}}^{-c} \le C\widehat{\mathbf{n}}^{\beta-c}$, this goes to zero as soon as $c > \beta$, note that this inequality is possible since c is chosen as a large positive constant as stated above.

Remark that $\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})^{-\theta_1}(\log\widehat{\mathbf{n}})^{\theta_2} \to \infty$ is equivalent to $(d_{\mathbf{n}}\beta_{1\widehat{\mathbf{n}}})^{-1} \to \infty$ by assumption (7), since

$$d_{\mathbf{n}}\beta_{1\widehat{\mathbf{n}}} \le C\widehat{\mathbf{n}}^{\beta}\psi(h_{\mathbf{n}})^{-1}\chi(\widehat{\mathbf{n}}, p^N)p^{-\theta}\epsilon_{\mathbf{n}}^{-1} \le C(\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})^{-\theta_1}(\log\widehat{\mathbf{n}})^{\theta_2})^{\frac{-\theta+2N(\beta+1)}{2N}}.$$

This yields the proof. □

Theorem 2 *Under conditions H1–H5, if (8), (9) are satisfied with* $\theta > N(1 + 2\beta + 2\widetilde{\beta})$ *and if* $\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})^{-\theta_3}(\log\widehat{\mathbf{n}})^{\theta_4} \to \infty$, *then we have:*

$$\lim_{\mathbf{n}\to\infty} \hat{\omega}_{\mathbf{n}} = \omega, \text{ in probability.} \tag{18}$$

Proof Similarly as the previous proof, it suffices to remark that

$$d_{\mathbf{n}}\beta_{1\widehat{\mathbf{n}}} \le C\widehat{\mathbf{n}}^{\beta}\psi(h_{\mathbf{n}})^{-1}\chi(\widehat{\mathbf{n}}, p^N)p^{-\theta}\epsilon_{\mathbf{n}}^{-1} \le C\widehat{\mathbf{n}}^{\beta+\widetilde{\beta}}\psi(h_{\mathbf{n}})^{-1}\left(\frac{\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})}{\log\widehat{\mathbf{n}}}\right)^{-(\theta/2N)}((\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})\log\widehat{\mathbf{n}})^{\frac{1}{2}} = C(\widehat{\mathbf{n}}S_{\mathbf{n}}^{\theta_3}(\log\widehat{\mathbf{n}})^{\theta_4})^{\frac{-\theta+N(2\beta+2\widetilde{\beta}+1)}{2N}}.$$ □

It is worth to study the exponential mixing case since it includes the Geometrically Strong Mixing (GSM) case (with $\chi \equiv 1$) which is easier to check in practice.

Theorem 3 (Exponential mixing case) Under the conditions *H*1–*H*5, (10), (7) or (8) and $\widehat{\mathbf{n}}\psi(h_{\mathbf{n}})(\log\widehat{\mathbf{n}})^{-2N-1}\to\infty$, *we have*

$$\lim_{\mathbf{n}\to\infty}\hat{\omega}_{\mathbf{n}}=\omega,\ \text{in probability.} \tag{19}$$

Proof The proof is obtained by sketching the proof of the first theorem and by using similar arguments as in the spatial non-functional case of Carbon et al. (1997). We then omit the proof. □

Remark 2 As usual in nonparametric setting, the choice of the smoothing factor is a crucial point to insure good behavior of the underlying procedure. This is done by estimating the small ball probabilities $P(X\in B(x,h))$, which play a key role in the theoretical properties of our mode estimate (for example in Sect. 3, we take $h=0.5$). Indeed, it has been shown by Dabo-Niang et al. (2006, Theorem 4) that the density estimate and the modal curve converge almost surely in the case of non-spatial general setting case. These results can be extended in the case of our spatial setting and deserve a paper on its own. Rates of convergence can be obtained by using additional assumptions linking the bandwidth h_n, the kernel K and the small ball probability functions of the variables $\{X_{\mathbf{i}}\}$. These rates are affected by the dependence condition introduced on the variables as shown in Sect. 3. For instance, one can obtain for some cases of fields a rate of order $(\log n)^{-\tau}$, where $\tau>0$. A special case of particular importance is when the distribution of the random field is absolutely continuous with respect to Wiener measure (diffusion processes), see for instance Li and Shao (2001), or Ferraty and Vieu (2006).

Let us now focus on the use of this consistent mode estimate to construct heterogeneity measurements.

3 Some applications of the mode

An application of the mode estimation is to build heterogeneity measurements. More precisely, we are interested with the spatial version of the heterogeneity measurements of Ferraty and Vieu (2006) or Dabo-Niang et al (2006). Recall that their heterogeneity measure is based on the centralities mean, median and mode curves. The idea is to look for some difference (in the sense of the considered semi-metric d) between the modal curve X_{modal} and one among the mean curve X_{mean} or the median curve X_{median}. As in the multivariate case, the modal curve is more useful than the mean or the median curve for detecting any structural differences between data. The heterogeneity measure of Ferraty and Vieu (2006) with respect to the median is:

$$HIR=\frac{d(X_{modal},X_{median})}{d(X_{median},0)+d(X_{modal},0)}.$$

One get a similar criterion based on the mean by replacing the X_{median} by X_{mean}. In order to get a more stable criterion, one can use a criterion obtained by splitting the initial dataset (of cardinal N) into D subgroups $g_1,\ldots,g_D$:

$$HIR(g_1,\ldots,g_D)=\frac{1}{N}\sum_{j=1}^{D}Card(g_j)HIR.$$

We are interested here with the spatial version of this heterogeneity index *HIR*. To do that we need to define the spatial version of centrality curves. Suppose that one deals with curves $X_{\mathbf{i}}=\{X_{\mathbf{i}}(t),t\in\Delta\},\mathbf{i}\in\mathcal{I}_{\mathbf{n}}$ with Δ some subset of $\mathbb{R}$ as it will be the case later on. Now, if the spatial mean, X_{mean}, can be obtained (and compute) exactly as in the *i.i.d.*:

$$\forall t\in\Delta,\ X_{mean}(t)=\frac{1}{\widehat{\mathbf{n}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}X_{\mathbf{i}}(t),$$

it is not the case of the spatial median and mode curves as we are going to see. We need to take into account the mixing condition.

3.1 The mixing condition in practice

Note that even if the estimator present here seems like in the i.i.d. case, one does take in mind that our results are obtained under a mixing condition. That means that we consider a non parametric spatial dependence measure defined by

$$\alpha(||\mathbf{i}-\mathbf{j}||)\le\chi(1,1)\varphi(||\mathbf{i}-\mathbf{j}||)$$

for any couple of sites $(\mathbf{i},\mathbf{j})$. For sake of simplicity, let us consider the strong mixing case (which corresponds to $\chi\equiv 1$) and suppose that we are in the case where $\varphi(||\mathbf{i}-\mathbf{j}||)$ tend to zero at a polynomial rate: $\varphi(||\mathbf{i}-\mathbf{j}||)\le C||\mathbf{i}-\mathbf{j}||^{-\theta}$, for some $\theta>0$. Then combining this assumption and the expression of the density:

$$\forall x_{\mathbf{j}},f_{\mathbf{n}}(x_{\mathbf{j}})=\frac{1}{\widehat{\mathbf{n}}C(K,\psi,h_{\mathbf{n}})}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}K\big(d(X_{\mathbf{i}},x_{\mathbf{j}})/h_{\mathbf{n}}\big),$$

we are dealing in practice with the estimator:

$$f_{\mathbf{n}}(x_{\mathbf{j}})=\frac{1}{\widehat{\mathbf{n}}C(K,\psi,h_{\mathbf{n}})}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}K\left(\frac{d(X_{\mathbf{i}},x_{\mathbf{j}})}{h_{\mathbf{n}}}\right)\mathbf{I}_{V_{\mathbf{j}}}(\mathbf{i})$$

where $\mathbf{I}_{v_{\mathbf{j}}}$ is the indicator function of the set $V_{\mathbf{j}}=\{\mathbf{i},\varphi(||\mathbf{i}-\mathbf{j}||)\le C\,||\mathbf{i}-\mathbf{j}||^{-\theta}\}$. Note that $V_{\mathbf{j}}$ is the set of Card($V_{\mathbf{j}}$) nearest neighbors sites of $\mathbf{j}$, a vicinity $V_{\mathbf{j}}$.

So, indeed, our density estimator at $x_{\mathbf{j}}$ concerns sum of terms $K(d(X_{\mathbf{i}},x_{\mathbf{j}})/h_{\mathbf{n}})$ for $X_{\mathbf{i}}$ where the $\mathbf{i}$ are closed neighbors of $\mathbf{j}$.

3.1.1 The spatial mode

Since the constant $C(K, \psi, h_{\mathbf{n}})$, does not depend on x, the modal curve of a sample of curves of a region $\mathcal{I}_{\mathbf{n}}$ can be defined as

$$X_{modal} = \arg \max_{m \in \{X_{\mathbf{i}}, \mathbf{i} \in \mathcal{I}_{\mathbf{n}}\}} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K\left(\frac{d(m, X_{\mathbf{i}})}{h}\right) \mathbf{I}_{V_{\mathbf{j}}}(\mathbf{i}).$$

In fact, K acts as a weight function: the larger is $d(m, X_{\mathbf{i}})$ and the smaller is $K\left(\frac{d(m,X_{\mathbf{i}})}{h}\right)$. The modal curve is obtained by using the following algorithm:

Algorithm for the spatial modal curve based on nearest neighbors. Let $k_{\mathbf{n}}$ be an integer.

(1) For each site $\mathbf{j}$ take the $k_{\mathbf{n}}$ nearest neighbors and compute the set $V_{\mathbf{j}}$ as defined above.
(2) Compute the sum of $k_{\mathbf{n}}$'s reals $K\left(\frac{d(X_{\mathbf{j}},X_{\mathbf{i}})}{h}\right), \mathbf{i} \in V_{\mathbf{j}}$ that is $\sum_{\mathbf{i} \in V_{\mathbf{j}}} K\left(\frac{d(X_{\mathbf{j}},X_{\mathbf{i}})}{h}\right)$.
(3) Take the maximum of over all the $\mathbf{j} \in \mathcal{I}_{\mathbf{n}}$.
(4) Take X_{modal} as the curve located at the site $\mathbf{j}_m$ where $\sum_{\mathbf{i} \in V_{\mathbf{j}_m}} K\left(\frac{d(X_{\mathbf{j}_m},X_{\mathbf{i}})}{h}\right)$ is maximum.

3.1.2 The median curve

Similarly, the notion of median curve can be extended to the functional framework. Here we will define the median curve as:

$$X_{median} = \arg \min_{m \in \{X_{\mathbf{i}}, \mathbf{i} \in \mathcal{I}_{\mathbf{n}}\}} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} d(X_{\mathbf{i}}, m).$$

We propose to use a similar algorithm by replacing $K\left(\frac{d(x,X_{\mathbf{i}})}{h}\right)$ with $d(x, X_{\mathbf{i}})$ and take the minimum instead of the maximum.

3.1.3 The use of the criterion on classification context

Similar to the variance, in context of classification, one can decide to split a set of curves of a given area A_k, $k \in \mathbb{N}$ or not into subgroups either according to the experience of the user, or by using a statistics as:

$$\begin{aligned} GAIN(A_k) &= GAIN(A_k; g_1, \ldots, g_D) \\ &= \frac{|HIR(A_k) - HIR(A_k; g_1, \ldots, g_D)|}{HIR(A_k)}. \end{aligned}$$

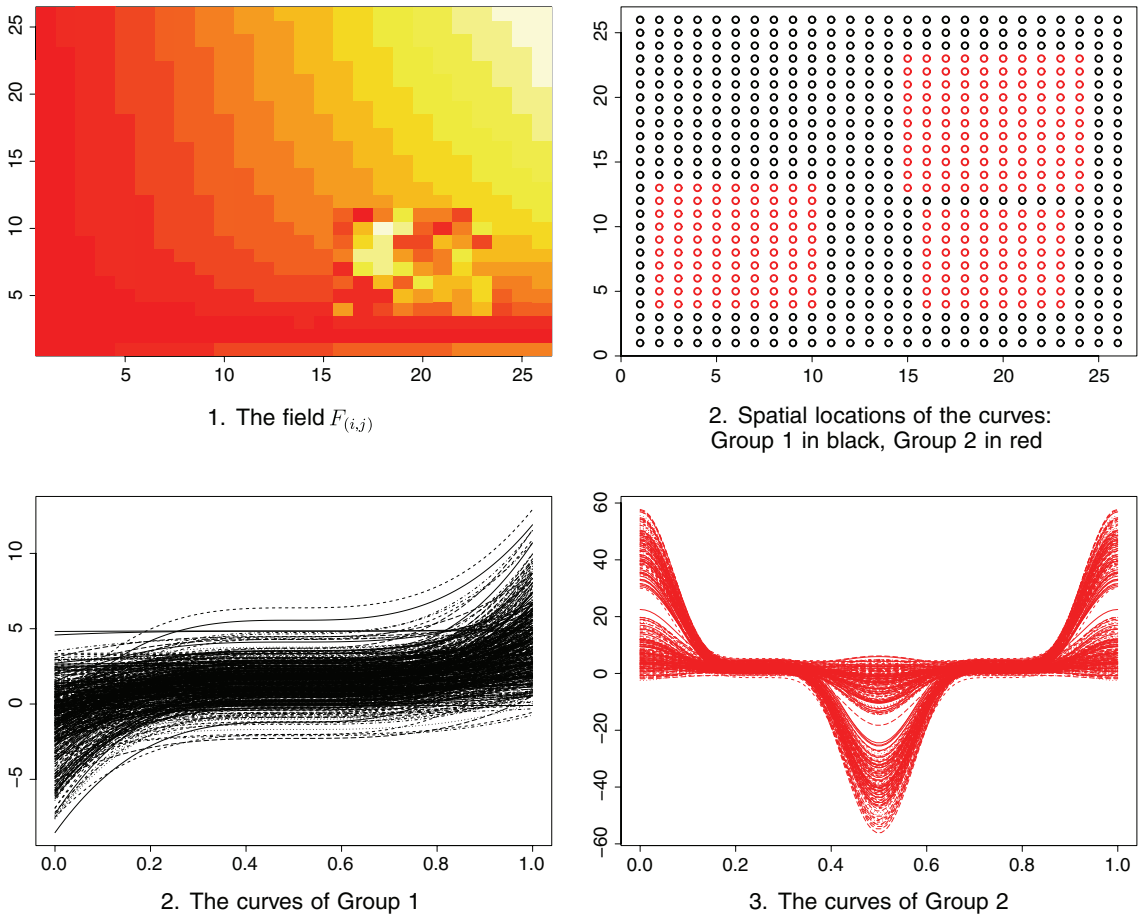

Fig. 2 The simulated dataset

Then the decision to split the dataset into $g_1, \ldots, g_K$ is taken if *GAIN* is greater than some fixed threshold $\gamma > 0$. Similar heterogeneity index can be computed with respect to the mean. One only needs to replace the median curve by the mean one in the expression of *HIR*. Then, we have computed the spatial version of the criterion of Dabo-Niang et al (2006) and compare it to that of the i.i.d. case in the following simulation results.

3.2 Applications

The spatial statistical modeling for treating spatial curves data consists in looking at them as a sample of dependent realizations of some functional variables $X_{\mathbf{i}}$, (observed on some rectangular region $\mathcal{I}_{\mathbf{n}}$) with the same distribution as a random field X taking values in some infinite dimensional semi-metric space $(\mathcal{E}, d)$. Here, we have considered a semi-metric based on the first $q = 8$ eigenfunctions of the Principal Components Analysis of the covariance operator $\Gamma(s, t) = cov(X(s), X(t))$; $s, t \in T$ (we refer for example to Ferraty and Vieu, 2006 for the theoretical setting of such a semi-metric).

3.2.1 *Simulations*

In the following let $N = 2$ and consider the following set S, of simulated curves in the area $\mathcal{I}_{(26,26)} = \{(i,j), 1 \leq i \leq 26; 1 \leq j \leq 26\}$. This set is built such that one have two different forms (groups) of curves.

The first group (Group 1) of spatial curves are built by simulating the field $X_{(i,j)}(t) = F_{(i,j)}.(t - 0.5)^3 + B_{(i,j)}$; $t \in T = [0,1]$ and $(i, j) \in R_1$ (spatial black locations of Fig. 2). While the second group (Group 2) of spatial curves are simulated using the field $X_{(i,j)}(t) = F_{(i,j)} \cos(2\pi t)^5 + B_{(i,j)}$ for $(i, j) \in R_2$ (spatial red locations of Fig. 2). B is a Gaussian random field with mean 2.5 and variance 1, the field $(F_{(i,j)}, (i,j) \in \mathcal{I}_{(26,26)})$ is the one presented in Fig. 2 (on top of the left).

The spatial locations of the curves are presented on Fig. 2 where the graphic at top-left represents the spatial location of the different groups (Group 1 in black and Group 2 in red). We want to compare the classification procedure of Ferraty and Vieu (2006) for independent curves and the spatial dependent version described above. The results are summaries in Figs. 3 and 4. They show that the two procedures do not give the same results. In fact, our cluster

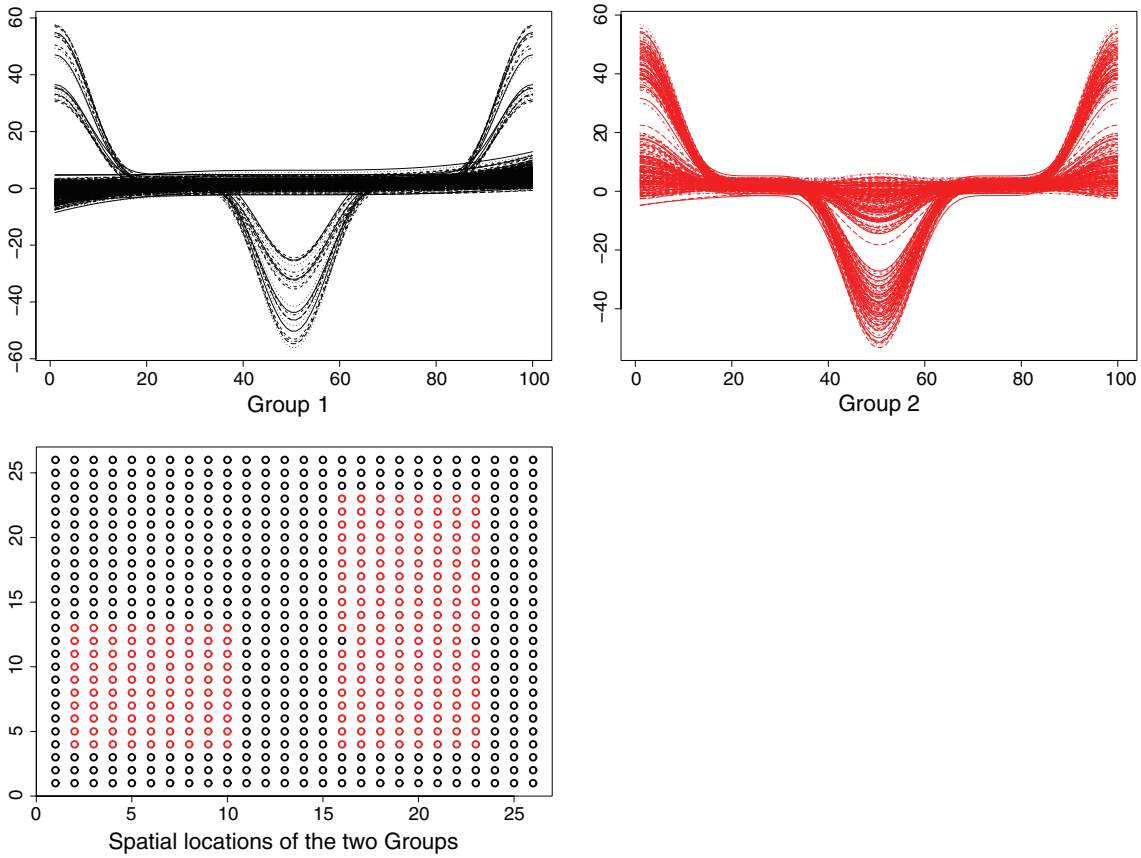

Fig. 3 Classification by taking into account spatial dependence

procedure that relies both on the structure and the spatial dependence of the curves is more likely to separate the two groups of spatial curves than those of Ferraty and Vieu clustering procedure which is based only on the closeness of curves on the functional space. Figure 3 gives the results of our cluster procedure which find that the set S of curves is heterogenous and retrieve the previous two groups with some curves misclassified (some of curves of Group 2 are allocated to Group 1). While the procedure of Ferraty and Vieu divides the set S into three groups of curves, see Fig. 4.

3.2.2 *Heterogeneity measurements for oxygen spatial distribution*

The dataset was collected in laboratory by the co-authors (Gilbert, Pischeda and Cuny) specialist on bioturbation (i.e. perturbation of some area by macro-organisms). Five strategic regions (see Fig. 5) were chosen in order to handle the spatial and temporal oxygen distribution in the *Nereis diversicolor* (a worm) burrow: the burrow entrance and exit (A1 and A5, respectively) were defined based on the direction of water circulation which, depends of the worm orientation in the burrow (peristaltic movement from head to tail). During the whole experiment (4h), the worm was located at the bottom of the burrow. We have also chosen to follow 2 intermediate regions (A2 and A4) and a bottom area (A3) (See Fig. 5).

The aim is to compare the heterogeneity induced by this organism on these regions during the experience. To do that one needs a measure of the heterogeneity. Note that such a tool interests researchers since it is a way of quantifying the bioturbation (the perturbation of the area by the concerned organism). That motivate this application where we apply the functional spatial heterogeneity index HIR. In order to test the capacity of our heterogeneity index to detect homogenous region, we have also considered on each area A_k, $k \in \{1, 2, 3, 4, 5\}$, a sub-region that represents a part of A_k which falls on the water (which should be more homogenous than A_k).

As mentioned on Sect. 1, on each site $\mathbf{i}$, we consider that we have a curve, $X_\mathbf{i}$ (some examples are given on Fig. 6). Then, for a fixed area A_k, $k \in \{1, 2, 3, 4, 5\}$, we have a set of curves:

$$\forall \mathbf{i} \in A_k,\ X_\mathbf{i} = \{X_\mathbf{i}(t), t \in [0, 120]\}.$$

Each region A_k has N_k number of sites, so N_k oxygen concentration (μmol L^{-1}) curves. The number of curves on

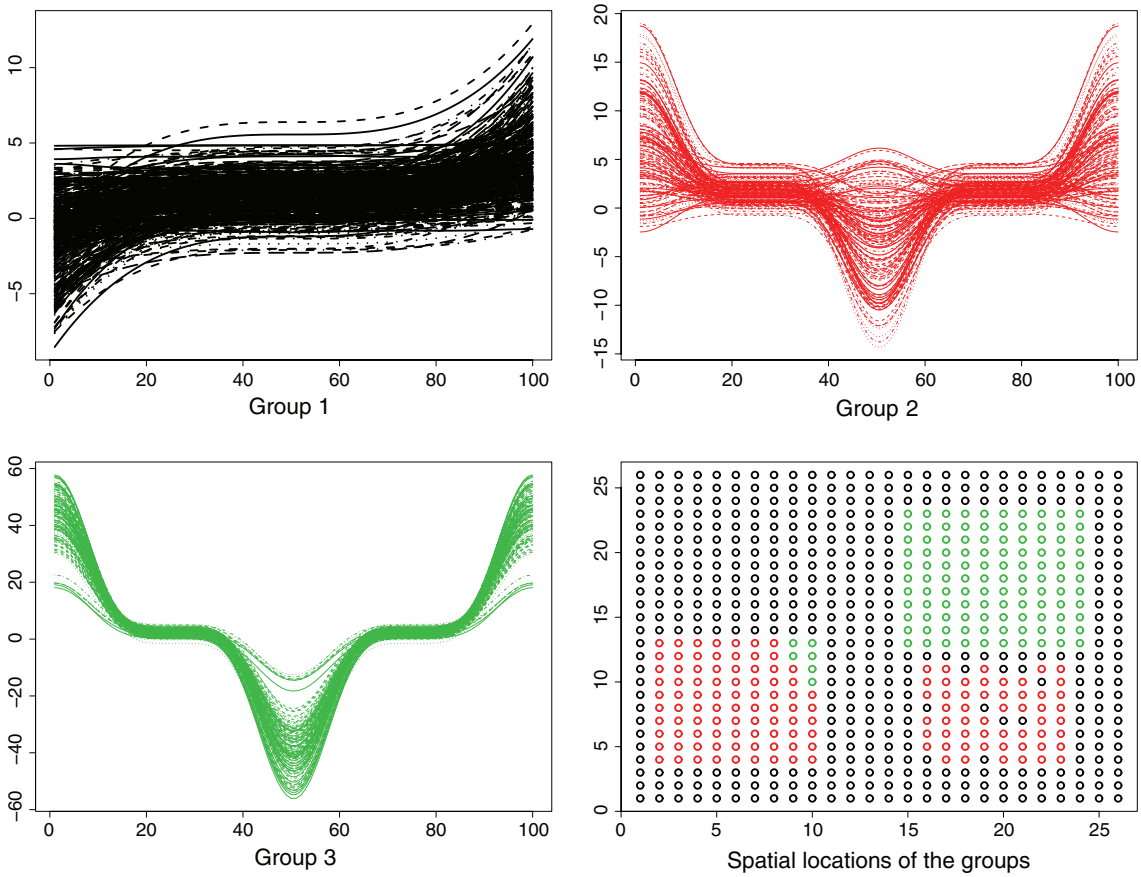

Fig. 4 Classification without taking into account the spatial dependence

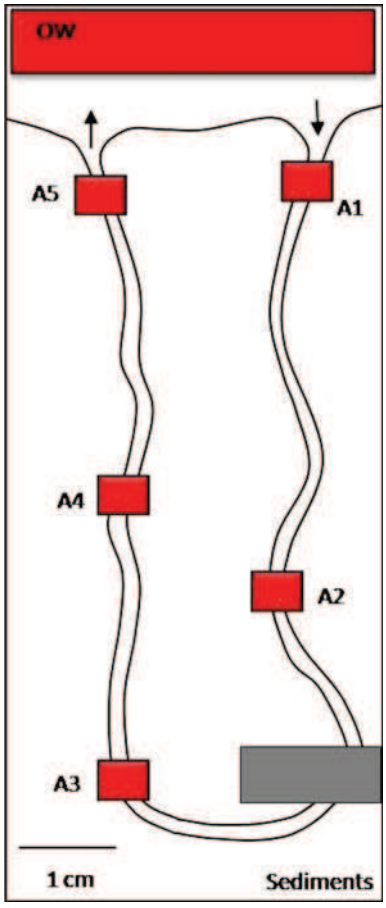

Fig. 5 The regions of interest

each zone is given in Table 1. The mean, median and mode curves for each region A_k are denoted respectively by: X^k_{mean}, X^k_{median}, X^k_{modal}.

The idea that we use is to look for some difference (in the sense of the considered semi-metric) between the modal curve X^k_{modal} and one among X^k_{mean} or X^k_{median}. Here the subgroups of each region are three neighboring sub-regions, keeping the spatial dependency of the observations.

Results. The results concerning the obtained index are summarized in Table 2. Then, we can say that in the lumen water, A1 and A5 regions are more homogeneous than intermediates regions. This could be explained by a "buffer effect": the close proximity with overlying water allowing constant direct oxygen exchanges between surface and lumen waters. On the other hand, if we consider each region as a whole (i.e. lumen water + burrow wall), A5 presents the highest heterogeneity indexes. Irregularities of the burrow structure (e.g. biofilm development, mucus lining) and related bacterial densities and respiratory activity could be sources for this heterogeneity.

Furthermore, the results obtained for the region in the water are smaller than the others one (i.e. lumen water + burrow wall) . This show that our heterogeneity index can handle more homogenous areas. In fact, in the water area, there is generally less variation of the oxygen than on elsewhere. Note that the fact that regions A2–A5 have closed heterogeneity indexes (with respect to the median) can be explain by the fact that they are smaller rectangular regions that overlap the hole nearly in the same way.

Remark

(1) Other studies based on more sophisticated tools are under investigation where we first smooth the rug dataset and use a semi-metric based on the derivative of the curves. Then, we propose a method to choose the vicinities $V_{\mathbf{j}}$.

Fig. 6 Examples of curves $X_{\mathbf{i}}$

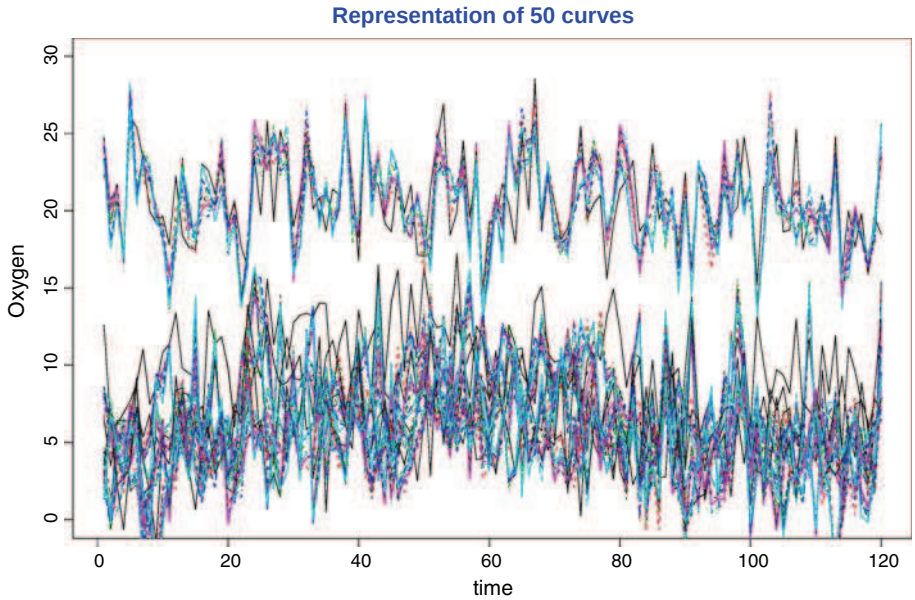

Table 1 Number of observations

Regions	A_1	A_2	A_3	A_4	A_5
Number of observations	21 × 101	36 × 84	37 × 75	43 × 93	30 × 101
Sub-samples (subregion)	21 × 16	21 × 16	37 × 15	21 × 16	21 × 11

Table 2 Heterogeneity measured

Heterogeneity index	A_1	A_2	A_3	A_4	A_5
HIR Median	0.758	0.896	0.900	0.981	0.994
HIR Mean	0.316	0.557	0.475	0.568	0.893
Sub-region in the water					
HIR Median	0.072	0.634	0.566	0.663	0.128
HIR Mean	0.054	0.052	0.053	0.057	0.057

(2) One can use other measure such as the coefficient of variation of all the observations on a region. But, as it is well known, such a measure is influenced by extreme or aberrant values of the observations.

4 Conclusion and discussions

The statistical method used here works well in the biological case presented above. It allows to quantify the heterogeneity of the biogeochemical parameter considered (oxygen) in each region. Indeed, it provides a helpful index to compare and understand the impact of an irrigated burrow in different regions of the sedimentary column.

In this paper, we apply a spatial discrete tool to our dataset when in fact we deals with a spatial continuous random field. Actually, our approach is well appropriated because the considered regions are relatively small. Nevertheless, as soon as we will need to work on a larger region for example to consider all the image (for example to compare the heterogeneity induce by two macro-organisms), we will need to take into account the spatial continuous aspect of the problem. Such a method for functional spatial continuous random fields is the subject of ongoing work. This later work in progress is an extension of the work of Biau (2003) and Dabo and Yao (2007).

References

Abraham C, Biau G, Cadre B (2003) Simple estimation of the mode of a multivariate density. Can J Stat 31:23–34

Aller RC (2001) Transport and reactions in the biorrigated zone. In: Boudreau BP, Jorgensen BB (eds) The benthic boundary layer: transport processes and biogeochemistry. Oxford Press, Oxford, pp 269–301

Aller RC, Aller JY (1998) The effect of biogenic irrigation intensity and solute exchanges on diagenetic reaction rates in marine sediments. J Mar Res 56:905–936

Angulo JM, Ruiz-Medina MD (2008) Spatio-temporal modeling of environmental and health processes. Stoch Environ Res Risk Asses 22(supp 1):1–2

Biau G (2003) Spatial kernel density estimation. Math Methods Stat 12:371–390

Biau G, Cadre B (2004) Nonparametric spatial prediction. Stat Inference Stochastic Proc 7:327–349

Bosq D (1998) Nonparametric statistics for stochastic processes—estimation and prediction, 2nd edn. Lecture Notes in Statistics, Springer-Verlag, New York

Carbon M, Hallin M, Tran LT (1996) Kernel density estimation for random fields. Stat Probab Lett 36:115–125

Carbon M, Tran LT, Wu B (1997) Kernel density estimation for random fields: the L_1 theory. Nonparam Stat, 6:157–170

Dabo-Niang S, Yao A-F (2007) Spatial kernel regression estimation. Math Methods Stat 16(4):1–20

Dabo-Niang S, Yao A-F (2008) Spatial kernel density estimation for functional random fields (Preprint submitted)

Dabo-Niang S, Ferraty F, Vieu P (2006) Mode estimation for functional random variable and its application for curves classification. Far East J Theor Stat 18:93–119

Doukhan P (1994) Mixing—properties and examples. Lecture Notes in Statistics, Springer-Verlag, New York

Fernández de Castro BM, González Manteiga W (2008) Boosting for real and functional samples: an application to an environmental problem. Environ Res Risk Asses 22(supp 1):27–37

Ferraty F, Vieu Ph (2006) Nonparametric functional data analysis. Springer-Verlag, New York

Foster S, Graf G (1995) Impact of irrigation on oxygen flux into the sediment: intermittent pumping by Callianassa subterranea and "piston-pumping" by Lanice conchilega. Mar Biol 123:335–346

Glud RB, Ramsing NB, Gundersen JK, Klimant I (1996) Planar optrodes: a new tool for fine scale measurements two-dimensional O_2 distributions in benthic communities. Mar Ecol Prog Ser 140:217–226

Hulth S, Aller RC, Engstrom P, Selander E (2002) A pH plate fluorosensor (optode) for early diagenetic studies of marine sediments. Limnol Oceanogr 47:212–220

Kristensen E (2001) Organic matter diagenesis at the oxic/anoxic interface in coastal marine sediments, with emphasis on the role of burrowing animals. Hydrobiologia 426:1–24

Kristensen E, Holmer M (2001) Decomposition of plant materials in marine sediments exposed to different electron acceptors (O_2, NO^{3-}, and SO_4^{2-}), with emphasis on substrate origin, degradation kinetics, and the role of bioturbation. Geochim Cosmochim Acta 65(3):419–433

Li WV, Shao QM (2001) Gaussian processes: inequalities, small ball probabilities and applications. In: Rao CR, Shanbhag D (ed.) Stochastic processes: theory and methods. Handbook of statistics, vol 19. North-Holland, Amsterdam

Nielsen OI, Kristensen E, Macintosh DJ (2003) Impact of fiddler crabs (Uca spp.) on rates and pathways of benthic mineralization in deposited mangrove shrimp pond waste. J Exp Mar Biol Ecol 289:59–81

Pischedda L, Poggiale J-C, Cuny P, Gilbert F (2008) Spatial and temporal oxygen heterogeneity in a Nereis diversicolor irrigated burrow (Submitted)

Porcu E, Mateu J, Saura F (2008) New classes of covariance and spectral density functions for spatio-temporal modelling. Res Risk Asses 22(supp 1):65–79

Rio E (2000) Thoerie Asymptotic des Processus Aleatoires Faiblement Dependants. Springer-Verlag, Berlin

Tran LT (1990) Kernel density estimation on random fields. J Multivariate Anal 34:37–53