

Article

Not peer-reviewed version

Nonparametric Partial Linear Estimation for Spatial Functional Data with Missing At-Random

[Tawfik Benchikh](#)^{*}, [Ibrahim M. Almanjahie](#), Omar Fetitah, [Mohammed kadi Attouch](#)

Posted Date: 7 December 2023

doi: 10.20944/preprints202312.0496.v1

Keywords: Missing at random data; Functional data analysis; Asymptotic normality; spatial data; Kernel regression method



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Nonparametric Partial Linear Estimation for Spatial Functional Data with Missing At-Random

Tawfik Benchikh ^{1,2,†,*} , Ibrahim M. Almanjahie ^{3,†} , Omar Fetitah ^{1,†}  and Mohammed Kadi Attouch ^{1,†} 

¹ Laboratory of Statistics and Stochastic Processes University of Djillali Liabes BP 89, Sidi Bel Abbas 22000, Algeria; o.fetitah@esi-sba.dz ; attou_kadi@yahoo.fr

² Faculty of Medical Sciences, University of Djillali Liabes BP 89, Sidi Bel Abbas 22000, Algeria; benchikh.tawfik@gmail.com

³ Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia; imalmanjahi@kku.edu.sa

* Correspondence: benchikh.tawfik@gmail.com

† These authors contributed equally to this work.

Abstract: The aim of this paper is to study a semi-functional partial linear regression model (SFPLR) for spatial data with responses missing at random. The estimators are constructed by the kernel method, and some asymptotic properties such as probability convergence rates of the nonparametric component and asymptotic distribution of the parametric and nonparametric components are established under certain conditions. Next, the performances and the superiority of these estimators are presented and examined using a study on simulated data and on real data by carrying out a comparison between our semi-functional partially linear model with MAR estimator (SFPLRM), the semi-functional partially linear model with the full-case estimator (SFPLRC) and the nonparametric functional model estimator with MAR (FNPM). The results show that the proposed estimators outperform existing estimators as the number of random missing data increases.

Keywords: missing at random data; functional data analysis; asymptotic normality; spatial data; kernel regression method

MSC: 62H12; 62G07; 62G35; 62G20

1. Introduction

Nowadays, statistical techniques are necessary for the analysis of massive volumes of data with a spatial argument; this is the case of environmental sciences, climate, geography, econometrics, medicine, biology, and other applied fields. Thus, high-dimensional data becomes functional data which is then processed and analyzed using functional data analysis (FDA) methods and they will be called functional data with spatial correlation. The theoretical and practical aspects resulting from this new branch of statistics have given rise to beneficial advances in areas with geographic dependence (for recent reviews on the subject, see [1]). Indeed, functional data analysis (FDA) has seen rapid growth in recent years, moving from exploratory and descriptive data analysis to linear models and estimation techniques. This dynamic contributes to theoretical and methodological improvements, as well as to the diversification of fields of application (see, for example, [2–4] and through various recent bibliographic discussions such as [5,6]). One of the major subjects in this field of research is functional regression, which studies the influence of a functional random variable on a scalar variable. Among the typical questions that arise on this subject, we find, in particular semi-parametric functional regression models that they retain the flexibility of parametric regression models and free themselves from the sensitivity to the dimensional effects of the non-parametric approaches. For a complete discussion with a state of the art on semi-parametric modeling by functional regression, we refer the readers to the bibliographical surveys of [7].

In particular, one of the most important semi-parametric functional models is the partially linear regression model introduced by [8], which is called the semi-functional partial linear regression (SFPLR) model, and therefore, many works were carried out to estimate and apply this model. This model is expressed as:

$$Y = \mathbf{X}^T \boldsymbol{\beta} + m(Z) + \epsilon, \quad (1)$$

where Y is the scalar response variable, $\mathbf{X} = (X_1, X_2, \dots, X_p)$ is a p -vector of explanatory variables, $\boldsymbol{\beta}$ is unknown p -dimensional parameter vector, Z is functional explanatory variable, $m(\cdot)$ is an unknown smooth functional operator and ϵ are identically distributed random errors satisfying $\mathbb{E}(\epsilon) = 0$ and unknown variance $\sigma^2(\epsilon) < \infty$. We generally assume an additional condition of independence of the variable error ϵ with respect to the random vector (\mathbf{X}, Z) . The authors use the classical kernel method (Nadaraya-Watson type weights method) to prove the asymptotic normality of $\boldsymbol{\beta}$ and the convergence rate of m . This model was extended to dependent data by [9]. Furthermore, [10] proposed a bootstrap procedure to approximate the distribution of these estimators, while [11] proposes to generalize this model to the case where the linear component is also functional. More recently, [12] studied different bootstrapping procedures for this model under dependency structures. Additionally, a procedure for testing linearity in partially linear functional models is proposed in [13]. Other approaches have been proposed to estimate SFPLR model parameters; we cite, for example, the local linear approach used by [14], the robust procedures considered by [15], the k nearest neighbors (k NN) procedure used by [16] and Bayesian approaches proposed by [17]. For recent advances, we can consult the bibliographic reviews in [6,18].

Furthermore, only a few research works have paid attention to estimation in the semi-functional partial linear regression model for spatially dependent observations. We cite the work of [19] on the autoregressive semi-functional linear regression (SAR) model, which proposed an estimator based on the method of quasi-maximum likelihood and local linear estimation while [20] obtained the asymptotic normality of the parametric component as well as the convergence in probability with the rate of the nonparametric component. The complete analysis of the data is the subject of all the works listed. On the other hand, in many applications, this is unfortunately not the case, and we are faced with missing data. There are several common causes of missed responses, including faulty equipment, sample contamination, manufacturing defects, clinical study dropouts, climatic circumstances, inaccurate data entry, and more. On this subject, filling in missing data and the uncertainty linked to such imputation has been the subject of in-depth studies in the multivariate case (see, for example, [21,22]). However, little research on nonparametric functional regression with missing data has been done. The first work was carried out by [23] to estimate the mean of a scalar response using an i.i.d. functional sample. The functional regressor is fully observed, with some responses missing at random (MAR). They generalize the result obtained in [24] in the multivariate case. While [25] took into account the estimation of the regression function and established these asymptotic properties under the condition of stationarity and ergodicity with MAR response. On the other hand, [26] used the k nearest neighbors method (k -NN) combined with the local linear method to estimate the regression function with a low number of randomly missing values responses (MAR). For semi-parametric partially linear multivariate models with randomly missing responses, we can cite the work of [27,28] while [29] is the first to study the SFPLR model for i.i.d data with a MAR response, and their results generalize those obtained in [27].

However, the handling of missing data for spatial data is very weak. We cite the work of [30] in the multivariate case to estimate the regression function with the kernel method. The performance of the estimator obtained is compared with the K nearest neighbors method. [31] considers multiple spatial regression models with missing data using the regression-based imputation method to predict missing values and adds a normally distributed residual term to each predicted value. It restores the loss of variability and biases associated with regression imputation. For functional data, we cite [32], which proposes to study the kernel estimation of the regression function when the data are spatially functional and MAR. They obtain the asymptotic properties of this estimator, notably the probability convergence (with rates) and the asymptotic normality of the estimator under certain weak conditions.

In this study, we propose to study SFPLR models for spatial data with missing responses. Our paper is organized as follows. In Section 2, the semi functional partial linear model for spatial data is presented with these estimators as well as the notations and hypotheses used in our work. whereas Section 3 is devoted to the asymptotic results. The proofs of the asymptotic results are postponed in Section 4. In section 5, a computer study on simulated and real data is carried out to demonstrate that the proposed estimators allow a nice improvement compared to the usual global approach.

2. The model and its estimators

Let \mathbb{Z}^N be the integer lattice points in the N -dimensional Euclidean space and let $\Lambda_i = \{(Y_i, X_i^{(1)}, X_i^{(2)}, \dots, X_i^{(p)}, Z_i)^T, i \in \mathbb{Z}^N\}$ be a measurable strictly stationary spatial process defined over a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ and identically distributed as $(Y, X^1, X^2, \dots, X^p, Z)^T$, where Y is a response variable that is depend on p -dimensional random vector $X_s (s = 1, \dots, p)$ in a linear way but is related to another independent functional variable Z in an indeterminate form by means of an SFPLR model

$$Y_i = \sum_{s=1}^p X_i^s \beta_s + g(Z_i) + \epsilon_i = \mathbf{X}_i^T \boldsymbol{\beta} + m(Z_i) + \epsilon_i \quad \mathbf{i} \in \mathbb{Z}^N, \quad (2)$$

where $\mathbf{X}_i = (X_i^{(1)}, X_i^{(2)}, \dots, X_i^{(p)})$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$, $m(\cdot)$ and ϵ are defined as before such that

$$\mathbb{E}(\epsilon_i | X_i^{(1)}, X_i^{(2)}, \dots, X_i^{(p)}, Z_i) = 0 \text{ and } \mathbb{E}(\epsilon_i^2 | X_i^{(1)}, X_i^{(2)}, \dots, X_i^{(p)}, Z_i) < \infty.$$

We assume that Z is valued in semi-metric space \mathcal{F} , where the associated semi-metric is denoted by $d(\cdot, \cdot)$ and we denote by $B(T, h) = \{Z' \in \mathcal{F} \text{ such that } d(Z, Z') \leq h\}$ the topological closed ball. We further assume that the process can be observed in the rectangular region $\mathcal{I}_n = \{\mathbf{i} = (i_1, \dots, i_N) \in \mathbb{Z}^N, 1 \leq i_k \leq n_k, k = 1, \dots, N\}$, with a sample size of $\hat{\mathbf{n}} = n_1 \times \dots \times n_N$ where $\mathbf{n} = (n_1, \dots, n_N)$. Suppose moreover that, for $l = 1, \dots, N$, n_l approaches infinity at the same rate: $C_1 < |n_j/n_k| < C_2$ for some $0 < C_1 < C_2 < \infty$ and we write that $\mathbf{n} \rightarrow \infty$ if $\min_{k=1, \dots, N}(n_k) \rightarrow \infty$. Remember that the term site is used to designate a point \mathbf{i} .

The kernel estimators of $\boldsymbol{\beta}_n$ and m_n (see [20]) are defined by:

$$\hat{\boldsymbol{\beta}}_n = \left(\sum_{\mathbf{i} \in \mathcal{I}_n} \check{X}_i (\check{X}_i)^T \right)^{-1} \left(\sum_{\mathbf{i} \in \mathcal{I}_n} \check{X}_i \check{Y}_i \right) \quad (3)$$

and

$$\hat{m}_n(t) = \sum_{\mathbf{i} \in \mathcal{I}_n} w_n(z, Z_i) \left(Y_i - X_i^T \hat{\boldsymbol{\beta}} \right), \quad (4)$$

where

$$\check{Y}_i = Y_i - \sum_{\mathbf{i} \in \mathcal{I}_n} Y_i w_n(Z_i, Z_j), \quad \check{X}_i = X_i - \sum_{\mathbf{i} \in \mathcal{I}_n} w_n(Z_i, Z_j) X_i$$

with

$$w_n(Z_i, Z_j) = \frac{K(d(Z_i, Z_j) b_n^{-1})}{\sum_{\mathbf{i} \in \mathcal{I}_n} K(d(Z_i, Z_j) b_n^{-1})}.$$

K denotes a real-valued kernel function and b_n a decreasing sequence of bandwidths which tends to zero as \mathbf{n} tends to infinity.

In this article, we develop an estimation approach in which the values of the independent variables (X and Z) are all observed, however, some observations of the response variable (Y values) are missing. We recall that we say that Y_i is missing if Y_i does not contain all of the required elements. For this we consider a real random variable denoted δ such that $\delta_i = 1$ if the value Y_i is known, and $\delta_i = 0$ otherwise. Thus, the study will be carried out on an incomplete sample of size n :

$$\{(Y_i, X_i, Z_i, \delta_i), \mathbf{i} \in \mathcal{I}_n\}.$$

It is assumed that the random missing data mechanism satisfies the saving condition

$$\mathbb{P}(\delta_i = 1/Y_i = y, X_i = x, Z_i = z) = \mathbb{P}(\delta_i = 1/X_i = x, Z_i = z) = p(x, z).$$

This conditional probability $p(x)$ is generally unknown.

First, we note that

$$\delta_i Y_i = \delta_i \mathbf{X}_i^T \boldsymbol{\beta} + \delta_i m(Z_i) + \delta_i \epsilon_i \quad \mathbf{i} \in \mathcal{I}_n, \quad (5)$$

From our assumption, by conditioning on $Z_i = z$, it follows that

$$\mathbb{E}(\delta_i Y_i / Z_i = Z) = \mathbb{E}(\delta_i \mathbf{X}_i / Z_i = Z)^T \boldsymbol{\beta} + \mathbb{E}(\delta_i / Z_i = Z) m(Z_i) \quad \mathbf{i} \in \mathcal{I}_n, \quad (6)$$

from where we have

$$m(Z) = \frac{(\mathbb{E}(\delta_i Y_i / Z_i = Z) - \mathbb{E}(\delta_i \mathbf{X}_i / Z_i = Z)^T \boldsymbol{\beta})}{\mathbb{E}(\delta_i / Z_i = Z)}. \quad (7)$$

In the following, we define $m_X(z) = \mathbb{E}(\delta \mathbf{X} / Z = z) / \mathbb{E}(\delta / Z = z)$, $m_Y(z) = \mathbb{E}(\delta Y / Z = z) / \mathbb{E}(\delta / Z = z)$ and we write $A^{\otimes 2} = AA^T$.

Then, by equations 5 and 6, we can write

$$\delta_i (Y_i - m_Y(Z_i)) = \delta_i (\mathbf{X}_i - m_X(Z_i))^T \boldsymbol{\beta} + \delta_i \epsilon_i.$$

Thus, if the functions $m_X(Z)$ and $m_Y(Z)$ are known, the least squares estimator of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}}_n = \left[\sum_{\mathbf{i} \in \mathcal{I}_n} \delta_i (\mathbf{X}_i - m_X(Z_i))^{\otimes 2} \right]^{-1} \left(\sum_{\mathbf{i} \in \mathcal{I}_n} \delta_i (\mathbf{X}_i - m_X(Z_i)) (Y_i - m_Y(Z_i)) \right). \quad (8)$$

However, $m_X(Z_i)$ and $m_Y(Z_i)$ are generally unknown and they must be estimated for applied 8. Assuming that m_X and m_Y are smooth functions of Z_i , these can be estimated by using the nonparametric Nadaraya-Watson kernel estimator noted by $\hat{m}_X(Z)$ and $\hat{m}_Y(Z)$ and expressed by

$$\hat{m}_{X,n}(Z) = \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_i w_n(Z, Z_i) \mathbf{X}_i,$$

$$\hat{m}_{Y,n}(Z) = \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_i w_n(Z, Z_i) Y_i$$

where

$$w_n(Z, Z_i) = \frac{K(d(Z, Z_i)/h_n)}{\sum_{\mathbf{j} \in \mathcal{I}_n} \delta_j K(d(Z, Z_j)/h_n)},$$

with K is a real-valued kernel function and h_n a decreasing sequence of bandwidths which tends to zero as n tends to infinity.

Hence, an estimator of $\boldsymbol{\beta}_n$ is given by

$$\hat{\boldsymbol{\beta}}_n = \left(\sum_{\mathbf{i} \in \mathcal{I}_n} \delta_i \tilde{X}_i (\tilde{X}_i)^T \right)^{-1} \left(\sum_{\mathbf{i} \in \mathcal{I}_n} \delta_i \tilde{X}_i \tilde{Y}_i \right), \quad (9)$$

where

$$\tilde{Y}_i = Y_i - \sum_{\mathbf{j} \in \mathcal{I}_n} \delta_j w_n(Z_i, Z_j) Y_j \quad \text{and} \quad \tilde{X}_i = X_i - \sum_{\mathbf{j} \in \mathcal{I}_n} \delta_j w_n(Z_i, Z_j) X_j.$$

We then insert $\hat{\boldsymbol{\beta}}_n$ in equation 7 to obtain

$$\hat{m}_n(Z) = \hat{m}_{Y,n}(Z) - \hat{m}_{X,n}^T(Z)\hat{\beta}_n. \quad (10)$$

3. Notations and Hypotheses

Our main objective is to obtain the asymptotic normality of our estimator under the condition that the process Λ_i is strictly stationary, which satisfies the following α -mixing condition (see [33]): there exists a real function $\varphi(t)$ which tends to 0 when t goes to ∞ , such that for finite cardinal subsets ($\text{Card}(\cdot) < \infty$) $E, E' \subset \mathbb{Z}^N$:

$$\begin{aligned} \alpha(\mathfrak{B}(E), \mathfrak{B}(E')) &= \sup_{\{(A,B) \in \mathfrak{B}(E) \times \mathfrak{B}(E')\}} \{|P(A \cap B) - P(A)P(B)|\} \\ &\leq \varphi(d'(E, E'))\psi(\text{Card}(E), \text{Card}(E')), \end{aligned} \quad (11)$$

where $d'(E, E')$ denotes the Euclidean distance in \mathbb{Z}^N , $\mathfrak{B}(S) = \mathfrak{B}(Z_i, i \in S)$, for $S \subset \mathbb{Z}^N$, the σ -fields generated by a random variable Z_i and $\psi : \mathbb{Z}^2 \rightarrow \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable such that:

$$\forall (r, s) \in \mathbb{Z}^2, \psi(r, s) \leq C \min(r, s), \text{ for some } C > 0, \quad (12)$$

Moreover, as is often the case in spatial regression, we assume other condition on the function φ :

$$\sum_{i=1}^{\infty} i^\gamma (\varphi(i)) < \infty, \text{ for some } \gamma > 0. \quad (13)$$

In what follows, in order to simplify the writings, we introduce the following notations:

$$\begin{aligned} \theta_i^{(s)}(z) &= X_i^{(s)} - \mathbb{E}[\delta_i X_i^{(s)} | Z_i = z], \quad s = 1, \dots, p, \quad \theta_i^{(0)}(z) = Y_i - \mathbb{E}[Y_i | Z_i = z], \\ \theta_i &= (\theta_i^{(1)}, \dots, \theta_i^{(p)})^T, \quad \Delta_i^{(s)}(z) = \mathbb{E}[\delta_i X_i^{(s)} | Z_i = z] - \sum_{j \in \mathcal{I}_n} \delta_j w_n(Z_i, Z_j) X_j^{(s)}, \quad s = 1, \dots, p \\ \Delta_i^{(0)}(z) &= \mathbb{E}[\delta_i Y_i | Z_i = z] - \sum_{j \in \mathcal{I}_n} \delta_j w_n(Z_i, Z_j) Y_j, \quad \Delta_i = (\Delta_i^{(1)}, \dots, \Delta_i^{(p)})^T, \end{aligned}$$

and we need the following assumptions which are necessary to prove and obtain our results.

(H1): Let $v_{ij} = \sup_{i \neq j} \mathbb{P}[(Z_i, Z_j) \in (B(z, h) \times B(z, h))]$ the probabilistic joint distribution of Z_i and Z_j and let $\phi_z(h_n) = P(Z \in B(z, h_n)) = \mu(B(z, h_n))$ the small ball probability.

We suppose that, $\forall i \neq j \in \mathbb{Z}^N$ and $\forall z \in \mathcal{E}$, there exist $C > 0$ such that,

$$\text{for some } 1 < a < \gamma N^{-1}, \sup_{i \neq j} v_{ij} \leq C(\phi_z(h))^{a+1}.$$

(H2): The kernel function $K : \mathbb{R} \rightarrow \mathbb{R}^+$ is assumed to be differentiable with support in the interval $[0, 1]$ such that there exist two constants C_3 and C_4 with

$$-\infty < C_3 < K'(t) < C_4 < 0, \text{ for } t \in [0, 1].$$

(H3): We assume that any function $f \in \{m, \theta_i^{(1)}(z), \dots, \theta_i^{(p)}(z)\}$ is smooth i.e there exists $c > 0$ such that for $\kappa > 0$, we have:

$$|f(u) - f(v)| \leq c d(u, v)^\kappa, \quad \forall u, v \in \mathcal{F}$$

(H4) There exists a differentiable nonnegative functions τ and g such that:

$$\phi_z(h) = g(h) \times \tau(x) + o(g(h)).$$

(H5) For all $t \in [0, 1]$, we have:

$$\lim_{h_n \rightarrow 0} \frac{\phi_z(t h_n)}{\phi_z(h_n)} = t_z(t).$$

(H6) Let $R_i = \theta_i \varepsilon_i$ and $\Sigma = \mathbb{E}[p(X_1, Z_1) (\theta_1 (\theta_1)^T)]$, where $\mathbf{1}$ is the site spatial $(1, \dots, 1)$ and we denote by $\mathbf{0}$ the site spatial $(0, \dots, 0)$.

i) The matrix $\mathbf{B} = \sum_{i \in \mathcal{I}_n} \mathbb{E}[p(X_i, Z_i) (R_0(R_i)^T)]$ is assumed positive definite.

ii) We assume that Σ is an invertible matrix.

(H7): We suppose that:

i) $\mathbb{E} |\varepsilon_1|^\rho + \mathbb{E} |\theta_1^{(1)}|^\rho + \dots + \mathbb{E} |\theta_1^{(p)}|^\rho < \infty$ for some $\rho \geq 3$.

ii) For all $\mathbf{i} \neq \mathbf{j}$, $\mathbb{E} [Y_i Y_j | (Z_i, Z_j)] < \infty$

iii) For all $\mathbf{i} \neq \mathbf{j}$, $\max_{1 \leq s \leq p} \mathbb{E} [X_i^{(s)} X_j^{(s)} | (Z_i, Z_j)] < \infty$.

(H8)

i) Let $V_2(z) = \text{var}(Y_i - \mathbf{X}_i^T \boldsymbol{\beta} | Z_i = z)$ and $V_s(z') = \mathbb{E} [|Y_i - \mathbf{X}_i^T \boldsymbol{\beta} - m(z)|^s | Z_i = z']$, with $s > 2$. We suppose that the function $V_2(\cdot)$ and $V_s(\cdot)$ is continuous function near z , i.e. since h tends toward 0, we have

$$\sup_{\{z': d(z, z') \leq h\}} |V_k(z') - V_k(z)| = o(1), k \geq 2.$$

ii) Let $V(z', z'', z) = \mathbb{E} [(Y_i - \mathbf{X}_i^T \boldsymbol{\beta} - m(z))(Y_j - \mathbf{X}_j^T \boldsymbol{\beta} - m(z)) | Z_i = z', Z_j = z'']$, for $\mathbf{i} \neq \mathbf{j}$.

We suppose that the function V is continuous in some neighborhood of (z, z) .

(H9) We also suppose that $p_1(\cdot) = \mathbb{P}(\delta = 1 | Z = z)$, is continuous function near z , i.e.

$$\sup_{\{u: d(z, z') \leq h\}} |p_1(z') - p_1(z)| = o(1), \text{ when } h \text{ go to } 0.$$

(H10) There exists $1 > \vartheta > \frac{N}{\gamma}$, in such that, $\phi_z(h) \geq \hat{n}^{\frac{\vartheta-1}{2N+1}}$.

Comments on the assumptions

The asymptotic results obtained subsequently are based on Theorem 4.1 [32] (see the proof section). It is, therefore, normal to impose the same hypotheses as in this theorem, as well as to hypotheses (H1)–(H5) and (H7)–H(10) (for comments on these hypotheses, see [32]). The additional hypothesis (H6) is usual in the context of the SFPLR.

4. Theoretical results

We are now in a position to give our asymptotic results. The first one gives the asymptotic distribution of the estimator for the parametric component of the model $(\hat{\boldsymbol{\beta}}_n)$.

Theorem 1. When the assumptions (H1)–(H10), 11, 12, 13 hold, if additionally the bandwidth parameter h_n and the function $\phi_z(h_n)$ satisfies $\hat{n} \frac{\phi_x(h_n)}{\log(\hat{n})} \rightarrow \infty$ and $\hat{n} h_n^k \frac{\phi_x(h_n)}{\log(\hat{n})} \rightarrow 0$, when $\mathbf{n} \rightarrow \infty$, we have:

$$(\hat{n})^{1/2} (\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \xrightarrow{D} \mathcal{N} \left(0, \Sigma^{-1} \mathbf{B} (\Sigma^{-1})^T \right) \quad (14)$$

The following results give the probability convergence and the asymptotic normality of the estimator of the non-parametric part.

Theorem 2. Based on hypotheses of the Theorem 1, we have:

$$|\hat{m}_{\mathbf{n}}(z) - m(z)| \xrightarrow{p} 0. \quad (15)$$

Theorem 3. Under the assumptions of 1, if in addition, $\hat{\mathbf{n}}h_{\mathbf{n}}^{2\kappa}(\phi_x(h_{\mathbf{n}})) \rightarrow 0$, as $\mathbf{n} \rightarrow \infty$, we have

$$(\hat{\mathbf{n}}\phi_z(h_{\mathbf{n}}))^{1/2} |\hat{m}_{\mathbf{n}}(z) - m(z)| \xrightarrow{D} \mathcal{N}(0, \sigma^2(z)). \quad (16)$$

where $\sigma^2(z) = \frac{\iota_2}{\iota_1^2} \frac{V_2(z)}{p_1(z)\tau(z)}$, with $\iota_k = -\int_0^1 \iota_z(t)(K^k)'(t)dt$, for $k = 1, 2$.

We note that these results obtained extend those which are established in the case of complete data (see [20]).

5. Proof

Throughout the rest of this paper, we set, $K_{\mathbf{i}} = K(h_{\mathbf{n}}^{-1}d(x, X_{\mathbf{i}}))$, for all $\mathbf{i} \in \mathcal{I}_{\mathbf{n}}$.

First, we recall the following results (Theorem 4.1 in [32]), which will be used in the demonstrations of our results.

Lemma 1. Under hypotheses (H1)-(H10), 11, 12, 13, if additionally the bandwidth parameter $h_{\mathbf{n}}$ and the function $\phi_z(h_{\mathbf{n}})$ satisfies $\frac{\hat{\mathbf{n}}\phi_x(h_{\mathbf{n}})}{\log(\hat{\mathbf{n}})} \rightarrow \infty$ and $\frac{\hat{\mathbf{n}}h_{\mathbf{n}}^{\kappa}\phi_x(h_{\mathbf{n}})}{\log(\hat{\mathbf{n}})} \rightarrow 0$, when \mathbf{n} tend to ∞ , then

$$\sqrt{\hat{\mathbf{n}}(\phi_x(h_{\mathbf{n}}) / \log(\hat{\mathbf{n}}))} |\hat{m}_{\mathbf{X}}(z) - m_{\mathbf{X}}(z)| \text{ converge in probability to } 0.$$

$$\sqrt{\hat{\mathbf{n}}(\phi_x(h_{\mathbf{n}}) / \log(\hat{\mathbf{n}}))} |\hat{m}_{\mathbf{Y}}(z) - m_{\mathbf{Y}}(z)| \text{ converge in probability to } 0.$$

Second, we need to state some preliminary results.

We denote by $\lambda_{\mathbf{i}}(z) = \frac{\sqrt{\phi_z(h)}}{\mathbb{E}(K_{\mathbf{i}})} [\delta_{\mathbf{i}}(Y_{\mathbf{i}} - \mathbf{X}_{\mathbf{i}}^T \boldsymbol{\beta} - m(z)) K_{\mathbf{i}}]$ and we define the random variable $W_{\mathbf{i}}(z)$ by $W_{\mathbf{i}}(z) = \lambda_{\mathbf{i}}(z) - \mathbb{E}(\lambda_{\mathbf{i}}(z))$.

Lemma 2. Based on hypotheses of the Theorem 1, we have:

$$(\hat{\mathbf{n}})^{-1} \text{Var} \left(\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} W_{\mathbf{i}}(z) \right) \rightarrow V(z) = p_1(z)V_2(z) \frac{\iota_2}{\iota_1^2 \tau(z)} \text{ as } \mathbf{n} \rightarrow \infty. \quad (17)$$

Proof:

Before going further, note that

$$\begin{aligned} \text{Var} \left(\sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} W_{\mathbf{i}}(z) \right) &= \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \text{Var} (W_{\mathbf{i}}(z)) + \sum_{\mathbf{i} \neq \mathbf{j}} \text{Cov} (W_{\mathbf{i}}(z), W_{\mathbf{j}}(z)) \\ &= I_{\mathbf{n}}(z) + R_{\mathbf{n}}(z), \end{aligned} \quad (18)$$

First, we have $\text{Var} (W_{\mathbf{i}}(z)) = \mathbb{E} (\lambda_{\mathbf{i}}^2(z)) - \mathbb{E}^2(\lambda_{\mathbf{i}}(z))$.

Conditioning on $Z_{\mathbf{i}}$ and using MAR assumption, we have

$$\mathbb{E}(\lambda_{\mathbf{i}}(z)) = \sqrt{\phi_z(h)} (m(z_{\mathbf{i}}) - m(z)) \frac{\mathbb{E}(K_{\mathbf{i}})}{\mathbb{E}(K_{\mathbf{1}})},$$

then, (H3) and (H9) implies

$$\mathbb{E}(\lambda_{\mathbf{i}}(z)) \leq \sqrt{\phi_z(h)} h^{\kappa} (p_1(z) + o(1)) \frac{\mathbb{E}(K_{\mathbf{i}})}{\mathbb{E}(K_{\mathbf{1}})}.$$

Similarly, conditioning on Z_i and using MAR assumption we have

$$\mathbb{E} \left(\lambda_i^2(x) \right) = \phi_z(h) \left[\mathbb{E} \left(p(z_i) (m(z_i) - m(z))^2 \frac{K_i^2}{\mathbb{E}^2(K_1)} \right) \right] + \phi_x(h) \left[\mathbb{E} \left(p(x_i) V_2(x_i) \frac{K_i^2}{\mathbb{E}^2(K_1)} \right) \right].$$

It follows, by (H8)(i) and (H9), that:

$$\left[\mathbb{E} \left(p(z_i) V_2(z_i) \frac{K_i^2}{\mathbb{E}^2(K_1)} \right) \right] = (p_1(z) + o(1)) (V_2(z) + o(1)) \frac{\mathbb{E}(K_i^2)}{\mathbb{E}^2(K_1)},$$

and by (H3) and (H9) that:

$$\mathbb{E} \left(p_1(z_i) (m(z_i) - m(z))^2 \frac{K_i^2}{\mathbb{E}^2(K_1)} \right) \leq (p_1(z) + o(1)) h^{2\kappa} \frac{\mathbb{E}(K_i^2)}{\mathbb{E}^2(K_1)}.$$

From where, we have

$$\mathbb{E} \left(\lambda_i^2(z) \right) \leq \phi_z(h) (p_1(z) + o(1)) \left[(V_2(z) + o(1)) + h^{2\kappa} \right] \frac{\mathbb{E}(K_i^2)}{\mathbb{E}^2(K_1)},$$

which implies that

$$\begin{aligned} \mathbb{E}(W_i(z)^2) &\leq \phi_z(h) (p_1(z) + o(1)) (V_2(z) + o(1)) \frac{\mathbb{E}(K_i^2)}{\mathbb{E}^2(K_1)} + h^{2\kappa} \phi_z(h) (p_1(z) + o(1)) \frac{\mathbb{E}(K_i^2)}{\mathbb{E}^2(K_1)} \\ &\quad + h^{2\kappa} \phi_x(h) (p_1(z) + o(1)) \frac{\mathbb{E}^2(K_i)}{\mathbb{E}^2(K_1)}, \end{aligned} \quad (19)$$

and, thus

$$\begin{aligned} I_n(x) &\leq \widehat{\mathbf{n}} \phi_z(h) (p_1(z) + o(1)) (V_2(z) + o(1)) \frac{\mathbb{E}(K_i^2)}{\mathbb{E}^2(K_1)} + \widehat{\mathbf{n}} h^{2\kappa} (p_1(z) + o(1)) \frac{\mathbb{E}(K_i^2)}{\mathbb{E}^2(K_1)} \\ &\quad + \widehat{\mathbf{n}} h^{2\kappa} \phi_x(h) (p_1(z) + o(1)) \frac{\mathbb{E}^2(K_i)}{\mathbb{E}^2(K_1)}. \end{aligned}$$

On other hand, using (H1) – (H2), (H4) – (H5), we leads to the fact that

$$(\tau(z) \phi_z(h))^{-1} \mathbb{E}(K_i^j) \rightarrow \iota_j, j = 1, 2,$$

Consequently; we have

$$\frac{1}{\widehat{\mathbf{n}}} I_n(z) = \frac{1}{\widehat{\mathbf{n}}} \sum_{i \in \mathcal{I}_n} \mathbb{E}(W_i(z)^2) \rightarrow V(z) = \frac{\iota_2 p_1(z) V_2(z)}{\iota_1^2 \tau(z)} \text{ as } \mathbf{n} \rightarrow \infty. \quad (20)$$

For the covariance term $R_n(z)$, taking account that

$$\sum_{i \neq j} \text{Cov} (W_i(z), W_j(z)) = \sum_{i \neq j} \mathbb{E} (W_i(z) W_j(z)),$$

then, by some argument as above, we have

$$\sum_{i \neq j} \mathbb{E} (W_i(z) W_j(z)) \leq \frac{1}{\mathbb{E}^2(K_1)} h^{2\kappa} \phi_z(h) (p_1(z) + o(1))^2 \sum_{i \neq j} (\mathbb{E}(K_i K_j) - \mathbb{E}(K_i) \mathbb{E}(K_j)). \quad (21)$$

In the following, we introduce the following sets $E_1 = \{\mathbf{i} \neq \mathbf{j} \in \mathcal{I}_n \text{ such that } \|\mathbf{i} - \mathbf{j}\| \leq c_n\}$, and $E_2 = \{\mathbf{i} \neq \mathbf{j} \in \mathcal{I}_n \text{ such that } \|\mathbf{i} - \mathbf{j}\| > c_n\}$, where c_n is a real sequence that tends to $+\infty$ as $n \rightarrow \infty$, which will be specified later, and we note in the following

$$R_n^1 = \sum_{E_1} |\mathbb{E}(K_i K_j) - \mathbb{E}(K_i) \mathbb{E}(K_j)| \text{ and } R_n^2 = \sum_{E_2} |\mathbb{E}(K_i K_j) - \mathbb{E}(K_i) \mathbb{E}(K_j)|.$$

On one hand, by assumption **(H1)**, we have:

$$R_n^1 \leq C \hat{n} c_n^N \frac{1}{\mathbb{E}^2(K_1)} h^{2\kappa} \phi_z(h) (p_1(z) + o(1)) \sum_{\mathbf{i} \neq \mathbf{j}} \phi_z(h)^{1+\frac{1}{a}}. \quad (22)$$

On the other hand, by Lemma (3.3) in [33] and using the fact that the random variables K_i are bounded, we have

$$|\mathbb{E}(K_i K_j) - \mathbb{E}(K_i) \mathbb{E}(K_j)| \leq C \varphi(\|\mathbf{i} - \mathbf{j}\|).$$

Consequently, we have

$$\begin{aligned} R_n^2 &\leq \frac{1}{\mathbb{E}^2(K_1)} C h^{2\kappa} \phi_z(h) (p_1(z) + o(1)) \sum_{\mathbf{i}, \mathbf{j} \in E_2} \varphi(\|\mathbf{i} - \mathbf{j}\|) \\ &\leq \frac{1}{\mathbb{E}^2(K_1)} C \hat{n} h^{2\kappa} \phi_z(h) (p_1(z) + o(1)) \sum_{\mathbf{i}: \|\mathbf{i}\| \geq c_n} \varphi(\|\mathbf{i}\|) \\ &\leq \frac{1}{\mathbb{E}^2(K_1)} C \hat{n} h^{2\kappa} \phi_z(h) (p_1(z) + o(1)) c_n^{-Na} \sum_{\mathbf{i} \neq \mathbf{j}: \|\mathbf{i}\| \geq c_n} \|\mathbf{i}\|^{Na} \varphi(\|\mathbf{i}\|). \end{aligned} \quad (23)$$

Then, by the condition (13); if we ask $c_n = (\phi_z(h))^{-\frac{1}{Na}}$, we have

$$R_n = R_n^1 + R_n^2 \leq C \hat{n} \frac{1}{\mathbb{E}^2(K_1)} h^{2\kappa} \phi_z(h) (p_1(z) + o(1)). \quad (24)$$

So the equations (21), (22) and (23) implies that

$$\sum_{\mathbf{i}, \mathbf{j}} \text{Cov}(W_{\mathbf{i}}(z), W_{\mathbf{j}}(z)) = o(\hat{n}). \quad (25)$$

The proof concludes from (18), (20), (25) ■

Using the lemma 2 and the same reasoning as in the proof of Theorem (4.2) in [32], with the same notations, we prove the following result.

Lemma 3. Under hypotheses **(H1)**-**(H10)**, if the bandwidth parameter h_n and if the function $\phi_z(h_n)$ satisfies $\hat{n}(\phi_z(h_n)) \rightarrow \infty$ and $\hat{n} h_n^{2\kappa} (\phi_x(h_n)) \rightarrow 0$, as $n \rightarrow +\infty$, then, we have

$$\sqrt{\hat{n} \phi_z(h_n)} |\hat{m}_n(z) - m(z)| \xrightarrow{D} \mathcal{N}(0, \sigma^2(z)). \quad (26)$$

with $\sigma^2(z) = \frac{t_2}{t_1^2} \frac{V_2(z)}{p_1(z) \tau(z)}$.

Lemma 4. For the hypotheses **(H1)** through **(H10)**, if additionally $\frac{\hat{n} \phi_x(h_n)}{\log(\hat{n})} \rightarrow \infty$ and $\frac{\hat{n} h_n^{\kappa} \phi_x(h_n)}{\log(\hat{n})} \rightarrow 0$, when n tend to ∞ , then we have:

$$\frac{1}{\hat{n}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \tilde{X}_{\mathbf{i}} (\tilde{X}_{\mathbf{i}})^T \longrightarrow \Sigma \text{ in probability.} \quad (27)$$

Proof:

Taking account that

$$\begin{aligned} \left(\frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \widetilde{X}_{\mathbf{i}} (\widetilde{X}_{\mathbf{i}})^T \right)_{rs} &= \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \theta_{\mathbf{i}}^{(r)} \theta_{\mathbf{i}}^{(s)} + \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}}^{(r)} \theta_{\mathbf{i}}^{(s)} \\ &+ \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}}^{(s)} \theta_{\mathbf{i}}^{(r)} + \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}}^{(r)} \Delta_{\mathbf{i}}^{(s)}, \end{aligned} \quad (28)$$

where

$$\Delta_{\mathbf{i}}^{(s)}(z) = \mathbb{E}[\delta_{\mathbf{i}} X_{\mathbf{i}}^{(s)} | Z_{\mathbf{i}} = z] - \sum_{\mathbf{j} \in \mathcal{I}_n} \delta_{\mathbf{j}} w_{\mathbf{n}}(Z_{\mathbf{i}}, Z_{\mathbf{j}}) X_{\mathbf{j}}^{(s)}, s = 1, \dots, p.$$

Then, by the strong law of large numbers and the MAR assumption, we have that

$$\frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \theta_{\mathbf{i}}^{(r)} \theta_{\mathbf{i}}^{(s)} \longrightarrow \mathbb{E} \left[p(X_{\mathbf{1}}, Z_{\mathbf{1}}) \theta_{\mathbf{1}}^{(r)} \theta_{\mathbf{1}}^{(s)} \right] = \Sigma_{rs} \text{ in probability.} \quad (29)$$

Second, by the lemma 1 and Cauchy-Schwartz inequality, it is easy to check that

$$\frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}}^{(r)} \theta_{\mathbf{i}}^{(s)} \text{ converge in probability to } 0, \quad (30)$$

and

$$\frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}}^{(s)} \theta_{\mathbf{i}}^{(r)} \text{ converge in probability to } 0. \quad (31)$$

So, we conclude the proof of Lemma 4 by using (29)-(31). ■

Proof of Theorem 1 Let us use the following decomposition:

$$\begin{aligned} (\widehat{\mathbf{n}})^{1/2} (\widehat{\boldsymbol{\beta}}_{\mathbf{n}} - \boldsymbol{\beta}) &= \left(\frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \widetilde{X}_{\mathbf{i}} (\widetilde{X}_{\mathbf{i}})^T \right)^{-1} \frac{1}{\sqrt{\widehat{\mathbf{n}}}} \left(\sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} R_{\mathbf{i}} + \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \theta_{\mathbf{i}} \left(\Delta_{\mathbf{i}}^{(0)} - \Delta_{\mathbf{i}}^T \boldsymbol{\beta} \right) \right. \\ &\left. + \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}} \varepsilon_{\mathbf{i}} + \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}} \left(\Delta_{\mathbf{i}}^{(0)} - \Delta_{\mathbf{i}}^T \boldsymbol{\beta} \right) \right) \end{aligned} \quad (32)$$

where

$$\Delta_{\mathbf{i}}^{(0)} = \mathbb{E}[\delta_{\mathbf{i}} Y_{\mathbf{i}} | Z_{\mathbf{i}}] - \sum_{\mathbf{j} \in \mathcal{I}_n} \delta_{\mathbf{j}} w_{\mathbf{n}}(Z_{\mathbf{i}}, Z_{\mathbf{j}}) Y_{\mathbf{j}}$$

So, by using (30)-(31) and the Cauchy-Schwartz inequality, we can see that

$$\frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \theta_{\mathbf{i}} \left(\Delta_{\mathbf{i}}^{(0)} - \Delta_{\mathbf{i}}^T \boldsymbol{\beta} \right) + \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}} \varepsilon_{\mathbf{i}} + \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} \Delta_{\mathbf{i}} \left(\Delta_{\mathbf{i}}^{(0)} - \Delta_{\mathbf{i}}^T \boldsymbol{\beta} \right) = o_p(1). \quad (33)$$

Which implies, by using Lemma 4, that

$$\widehat{\boldsymbol{\beta}}_{\mathbf{n}} - \boldsymbol{\beta} \longrightarrow \left(\Sigma^{-1} + o(1) \right) \left(\frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} R_{\mathbf{i}} + o(1) \right) \text{ in probability.} \quad (34)$$

Then, according to the condition **(H6)(ii)** and by applying **theorem 6.1.1** in [35] on $\{\delta_{\mathbf{i}} R_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^2\}$, we have:

$$(\widehat{\mathbf{n}})^{-1/2} \sum_{\mathbf{i} \in \mathcal{I}_n} \delta_{\mathbf{i}} R_{\mathbf{i}} \xrightarrow{D} \mathcal{N}(0, \mathbf{B}). \quad (35)$$

The proof concludes from (30), (34), (35) and Lemma 4. ■

Proof of Theorem 2.

It suffices to see that

$$\begin{aligned}\hat{m}_{\mathbf{n}}(z) - m(t) &= (\hat{m}_{y,\mathbf{n}}(Z) - m_{y,\mathbf{n}}(Z)) - (\hat{m}_{X,\mathbf{n}}(Z) (\hat{\beta}_{\mathbf{n}} - \beta)) \\ &= S_1 - S_2,\end{aligned}\quad (36)$$

Then, we have

$$\sqrt{\hat{\mathbf{n}}(\phi_x(h_{\mathbf{n}})/\log(\hat{\mathbf{n}}))} |\hat{m}_{\mathbf{n}}(z) - m(z)| \leq \sqrt{\hat{\mathbf{n}}(\phi_x(h_{\mathbf{n}})/\log(\hat{\mathbf{n}}))} |S_1| + \sqrt{\hat{\mathbf{n}}(\phi_x(h_{\mathbf{n}})/\log(\hat{\mathbf{n}}))} |S_2| \quad (37)$$

By the lemma 1, we have

$$\sqrt{\hat{\mathbf{n}}(\phi_x(h_{\mathbf{n}})/\log(\hat{\mathbf{n}}))} |S_1| \xrightarrow{p} 0 \quad (38)$$

On the other hand, we can write

$$|S_2| \leq |\hat{\beta} - \beta| \left[|\mathbb{E}(\delta_i X_i | Z_i = z)| + \left| \sum_{i \in \mathcal{I}_{\mathbf{n}}} \delta_i w_{\mathbf{n}}(t, Z_i) X_i - \mathbb{E}(\delta_i X_i | Z_i = z) \right| \right].$$

Then, by the theorem 1, $|\hat{\beta}_{\mathbf{n}} - \beta|$ converge in probability to 0 and according to (H7), we have $|\mathbb{E}(\delta_i X_i | Z_i = z)| < \infty$. Moreover, by the lemma 1, we have

$$\sqrt{\hat{\mathbf{n}}(\phi_x(h_{\mathbf{n}})/\log(\hat{\mathbf{n}}))} \left| \sum_{i \in \mathcal{I}_{\mathbf{n}}} \delta_i w_{\mathbf{n}}(z, Z_i) X_i - \mathbb{E}(\delta_i X_i | Z_i = z) \right| \xrightarrow{p} 0.$$

That implies

$$\left| \sqrt{\hat{\mathbf{n}}(\phi_x(h_{\mathbf{n}})/\log(\hat{\mathbf{n}}))} S_2 \right| \xrightarrow{p} 0 \quad (39)$$

Thus, from (37)-(39) the proof is over. ■

Proof of Theorem 3.

From 36, we have

$$\begin{aligned}\sqrt{\hat{\mathbf{n}}\phi_x(h_{\mathbf{n}})} (\hat{m}_{\mathbf{n}}(z) - m(t)) &= \sqrt{\hat{\mathbf{n}}\phi_x(h_{\mathbf{n}})} (\hat{m}_{y,\mathbf{n}}(Z) - m_{y,\mathbf{n}}(Z)) - \sqrt{\hat{\mathbf{n}}\phi_x(h_{\mathbf{n}})} (\hat{m}_{X,\mathbf{n}}(Z) (\hat{\beta} - \beta)) \\ &= S_3 - S_4,\end{aligned}\quad (40)$$

Now, by lemma 3, we have

$$S_3 \xrightarrow{D} \mathcal{N}(0, \sigma^2(z)), \quad (41)$$

and by the theorem 1, we have $|S_4| \xrightarrow{p} 0$

Then, from (40)-(41) the proof is over. ■

6. Computational study

In this section, we are interested in the behavior of the estimators proposed on samples of finite size, with particular attention to the influence of spatial correlation and the effect of MAR on the efficiency of the estimators. To do that, we conducted simulations based on observations denoted as $(X_i, Z_i, Y_i, \delta_i)$ with $\mathbf{i} = (i_1, i_2)$, $1 \leq i_1 \leq n_1$, $1 \leq i_2 \leq n_2$, and we generate the SFPLR model with MAR as follows:

$$\delta_i Y_i = \delta_i r(Z_i, X_i) + \delta_i \varepsilon_i = \delta_i X_i^T \beta + \delta_i m(Z_i) + \delta_i \varepsilon_i, \quad (42)$$

where $\beta = (1, 2)^T$, $X_i^j \sim \text{Exponential}(0.5)$, $j = 1, 2$ and take the nonparametric operator m as follows

$$m(z) = \frac{1}{\int_0^1 |1 + z(t)| dt}.$$

The curves $Z_i(t)$ were $Z_i(t) = B_i t \sin(t - A_i) + v_i(t)$ for $t \in [0, 1]$.

For simulated the curves Z_i , we take $v_i(t) \sim \mathcal{N}(0, 2)$, $A = D * \sin\left(\frac{G}{2} + .5\right)$ with $G = \text{GRF}(0, 5, 3)$, $B = \text{GRF}(2.5, 5, 3)$ and $\varepsilon = \text{GRF}(0, .1, 5)$ where the function $\text{GRF}(\mu, \sigma^2, s)$ denotes a stationary Gaussian random field with mean μ and covariance function defined by $C(l) = \sigma^2 \exp\left(-\left(\frac{\|l\|}{s}\right)^2\right)$, $l \in \mathbb{R}^2$ and $s > 0$. and the function D is defined by $D_i = \frac{1}{n_1 \times n_2} \sum_j \exp\left(-\frac{\|i-j\|}{a}\right)$ (i.e. $D_{(i,j)} = \frac{1}{n_1 \times n_2} \sum_{\substack{1 \leq j_1 \leq n_1 \\ 1 \leq j_2 \leq n_2}} \exp\left(-\frac{\|(i_1, j_1) - (j_1, j_2)\|}{a}\right)$). The curves, following the values of a , are displayed in Figure 1. Note that we consider the same curves as in [34] where the function D is used to control the spatial mixing condition. Therefore, these observations are a mixture of independent and dependent data points, as illustrated in Figure 2 (for more details, see [34]). To reduce the level of independence in the data, it is sufficient to decrease the value of a . For our analysis, we have employed a value of $\sigma^2 = 5$ and $s = 5$. Moreover, similar to that described in [29], we adopted the following missing data mechanism:

$$p(x^1, x^2, z) = \mathbb{P}(\delta = 1 | X^1 = x^1, X^2 = x^2, Z = z) = \text{expit}\left(2\alpha \left(\sum_{j=1}^2 |x^j| + \int_0^1 z^2(t) dt\right)\right),$$

where $\text{expit}(u) = e^u / (1 + e^u)$ for all $u \in \mathbb{R}$ and we will take for α the following values: $\alpha = 0.05, 0.5, 5$.

Let's remember that the degree of dependence between the functional variable Z and the variable δ is controlled by the parameter α and to check the value of $p(x)$, we compute $\bar{\delta} = 1 - \frac{1}{n_1 \times n_2} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \delta_{(i_1, i_2)}$.

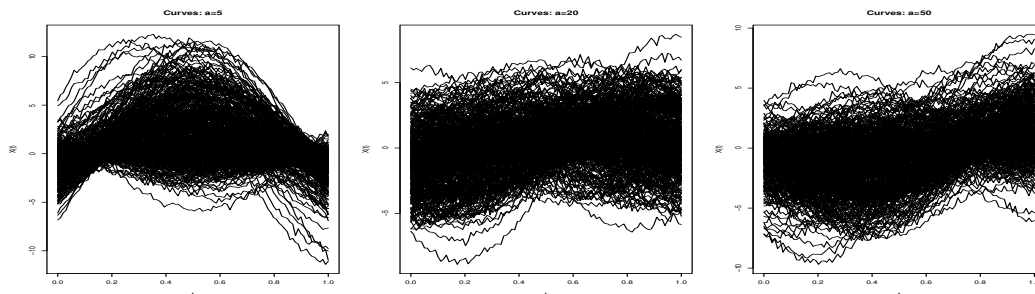


Figure 1. The curves $Z_i, t \in [0, 1]$ for $a = 5, 20, 50$.

In order to calculate our estimators, we use the class of semi-metrics based on the principal component analysis (PCA) metric which is best suited to the treatment of this type of data (discontinuous functional variable). Furthermore, we have chosen the standard kernel function defined as follows: $K(u) = \frac{3}{2} (1 - u^2) \mathbb{1}_{[0,1]}(u)$.

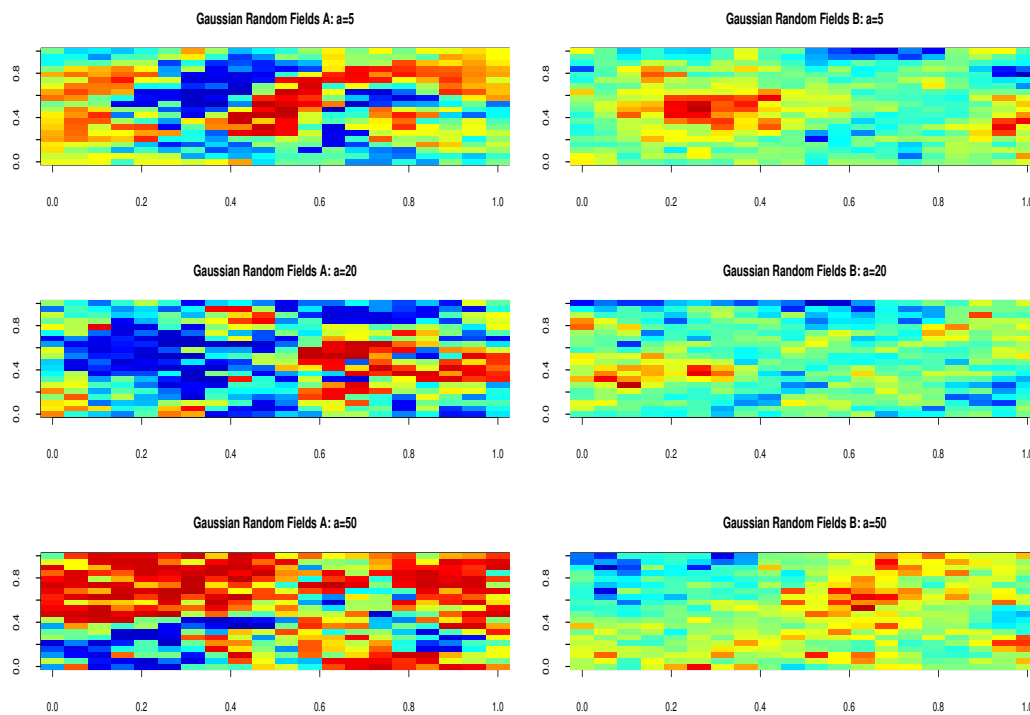


Figure 2. Simulations of the random field were generated for different values of a , specifically $a = 5, 20, 50$.

The objective of this computer study is to conduct a comparison between our semi-partially linear model with MAR estimator (SFPLRM), semi-partially linear model estimator \tilde{r}_n in the complete case (SFPLRC) (see [20]) and the nonparametric functional model estimator with MAR (FNPM).

Recall that the FNPM model is defined as follows:

$$Y = r(Z) + \epsilon,$$

and the estimator of operator regression r (see [32]) is given by

$$\hat{r}_n(z) = \frac{\sum_{j \in \mathcal{I}_n} \delta_j Y_j K(d(z, Z_j)/h_n)}{\sum_{j \in \mathcal{I}_n} \delta_j K(d(z, Z_j)/h_n)}.$$

For that purpose, we conducted a random split of our data, denoted as $(X_i, Z_i, Y_i, \delta_i)_i$, into two subsets: a test sample $(X_i, Z_i, Y_i)_{i \in I'}$ (consisting of data without missing values) and training sample $(X_i, Z_i, Y_i, \delta_i)_{i \in I}$ that they will be used to select the optimal smoothing parameters h_{opt} , via cross-validation procedures (for more details on the selected tuning parameters, see [4]).

To evaluate the precision of the estimators of the three models (SFPLRM, SFPLRC, and FNPM), we use the square and mean square errors, denoted MSE :

$$\begin{aligned} MSE_{FNPM} &= \frac{1}{\#(I')} \sum_{i \in I'} (\hat{r}_n(Z_i) - Y_i)^2, \\ MSE_{SFPLRC} &= \frac{1}{\#(I')} \sum_{i \in I'} (\tilde{r}_n(Z_i, X_i) - Y_i)^2, \\ MSE_{SFPLRM} &= \frac{1}{\#(I')} \sum_{i \in I'} (\hat{r}_n(Z_i, X_i) - Y_i)^2, \end{aligned}$$

where $\#(I')$ represents the size of the testing sample I' . The experiment was replicated $M = 100$ times, which allows us to compute M values for MSE and display their distribution through a boxplot. The results obtained (the prediction values compared with the true values) are presented in Figure 3 for the three models with different values of a . Figure 4 displays boxplots constructed from MSE obtained for the three models.

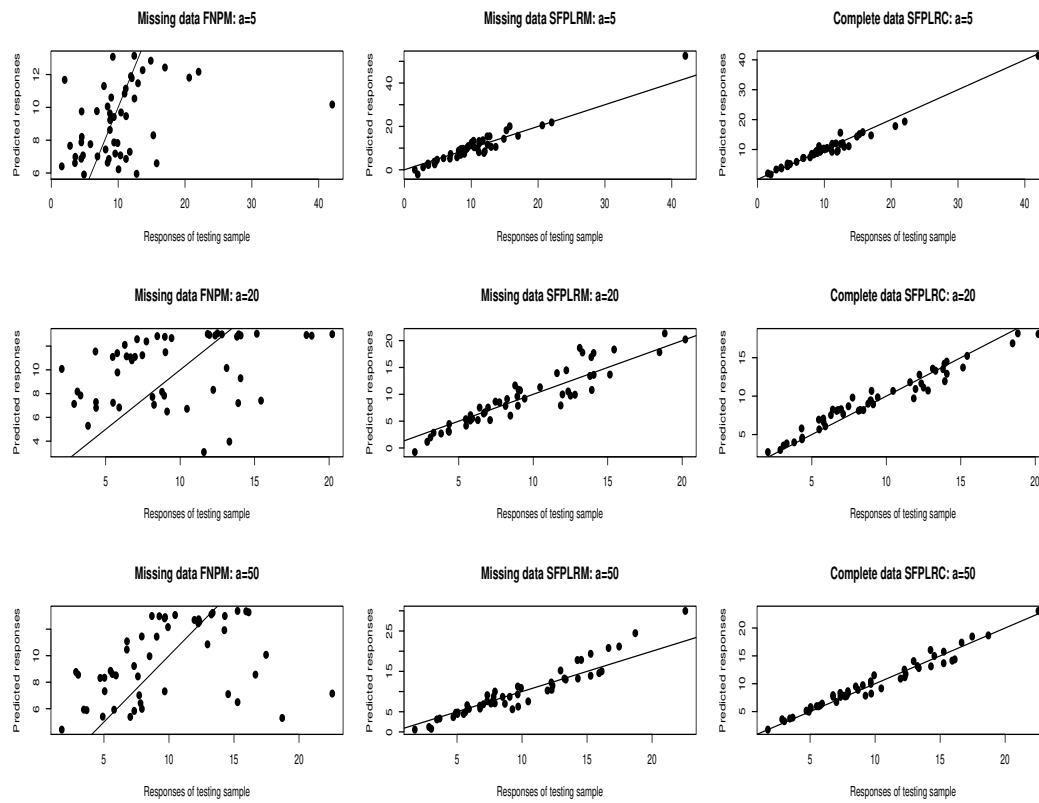


Figure 3. Predictions of the 2 models for $a = 5, 20, 50$.

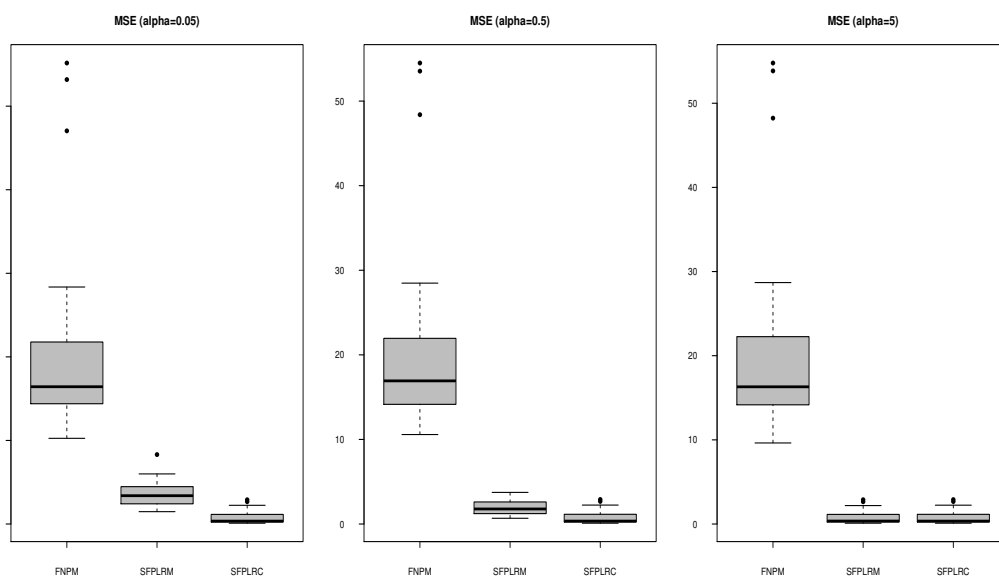


Figure 4. Boxplot MSE of the 2 models for $\alpha = 0.05, 0.5, 5$.

By Figure 4, we observe that the SFPLRM estimator shows better prediction effects than the FNPM in comparison to the SFPLRC estimator.

In Table 1 below, we present the mean squared error (MSE) results for three methods, FNPM, SFPLM, and SFPLR, under various combinations of n_1 , n_2 , and significance level α .

Table 1. Mean squared error (MSE) for FNPM, SFPLRM and SFPLRC.

n_1	n_2	FNPM $\alpha = 0.05$	SFPLRM $\alpha = 0.05$	FNPM $\alpha = 0.5$	SFPLRM $\alpha = 0.5$	FNPM $\alpha = 5$	SFPLRM $\alpha = 5$	SFPLRC Complete
10	10	22.29	6.562	21.64	3.782	21.13	2.611	2.504
10	20	26.29	5.869	24.88	3.329	25.06	1.232	1.228
10	50	22.62	5.132	21.89	2.929	21.71	1.105	1.098
20	10	22.98	4.865	21.62	3.079	21.37	1.282	1.253
20	20	20.00	4.539	19.22	2.845	18.82	0.991	0.987
20	50	21.63	4.518	21.18	2.428	21.03	0.763	0.755
50	10	20.14	4.578	20.19	2.973	19.51	1.642	1.624
50	20	22.27	3.817	21.83	2.285	21.93	0.631	0.630
50	50	20.64	3.892	20.37	2.246	20.29	0.491	0.488

In Table 1, which presents the mean squared error (MSE) for FNPM, SFPLRM, and SFPLRC under various conditions defined by combinations of n_1 , n_2 and α , it becomes evident that the SFPLRM estimator consistently shows superior predictive accuracy when compared to the FNPM estimator across a range of settings. This is notably evident as the MSE values for the SFPLRM estimator consistently remain lower across the majority of combinations of n_1 , n_2 and α . Furthermore, a notable observation is that higher values of α (e.g., $\alpha = 5$) tend to correspond with lower MSE values, indicating enhanced predictive performance for the SFPLRM estimator. It's also worth mentioning that as the number of samples ($n_1 \times n_2$) increases, the results suggest that the missing at random (MAR) mechanism has little to no discernible effect on the prediction MSE. This observation implies that the SFPLRM estimator maintains consistent performance regardless of the presence of missing data, mainly when dealing with larger sample sizes.

Now, we aim to examine the bias in the estimation of β and its associated Square Error (SE), which is defined as:

$$SE = |\hat{\beta}_n - \beta|^2,$$

The results are presented in Table 2

Table 2. Square Error (SE) for $\hat{\beta}$.

n_1	n_2	SFPLRM $\hat{\beta}_n^1$ $\alpha = 0.05$	SFPLRM $\hat{\beta}_n^2$ $\alpha = 0.05$	SFPLRM $\hat{\beta}_n^1$ $\alpha = 0.5$	SFPLRM $\hat{\beta}_n^2$ $\alpha = 0.5$	SFPLRM $\hat{\beta}_n^1$ $\alpha = 5$	SFPLRM $\hat{\beta}_n^2$ $\alpha = 5$	SFPLRC $\hat{\beta}_n^1$ C	SFPLRC $\hat{\beta}_n^2$ Complete
10	10	0.504	1.148	0.477	0.807	0.574	0.475	0.463	0.456
10	20	0.479	1.309	0.471	1.025	0.280	0.491	0.331	0.475
10	50	0.485	1.415	0.472	1.144	0.337	0.492	0.287	0.479
20	10	0.485	1.488	0.464	1.086	0.348	0.510	0.320	0.471
20	20	0.488	1.597	0.477	1.163	0.311	0.488	0.275	0.478
20	50	0.469	1.432	0.463	1.029	0.317	0.488	0.258	0.482
50	10	0.491	1.534	0.475	1.048	0.274	0.516	0.322	0.498
50	20	0.489	1.426	0.479	1.096	0.237	0.481	0.271	0.472
50	50	0.477	1.463	0.468	1.106	0.290	0.487	0.246	0.480

The results from Table 2 suggest that the SFPLRM estimator displays good accuracy in estimating the parameters $\hat{\beta}_n$, with relatively small Square Error values. This implies that the estimator provides

reasonably accurate estimates of the true parameters, even in cases with missing data and varying degrees of dependence. The theoretical conclusions of Theorems 3.1 and 3.2 are consistent with such a numerical outcome.

7. Real data application

The objective of this part is to compare, on a set of real data consisting of particle pollution indices, the effectiveness of the SFPLR model by our estimators when the data are MAR. The source of these data is the AriaWeb information system, managed by CSI Piemonte and Regione Piemonte, and our analysis is obtained from 34 monitoring sites using gravimetric instruments recorded during the winter season from October 2005 to March 2006 (daily measurements including $T = 182$ days). This involves analyzing the levels of pollution, which allowed us to detect higher levels of pollution in the plains closer to urban centers while lower concentrations of this index are observed near the Alps (for more detailed information about the data, we recommend referring to the publication by [36]).

To select the appropriate covariates, a preliminary regression analysis was carried out and the following covariates were selected

- $X^1 = HMIX$: maximum daily mixing height (in meters),
- $X^2 = EMI(s)$: daily primary aerosol emission rates (in g/s).
- $X^3 = PREC(s)$: total daily precipitation (in millimeters),
- $Z = TEMP$: the average daily temperature (in Kelvin $^{\circ}K$).

To implement the theoretical conclusions of the previous section on real data, we will analyze the effectiveness of our estimators built with MAR data in the context of spatial functional prediction which highlights the importance of considering spatial locations. Specifically, we assume that the observations are linked via an SFPLR model (42), where the response variable is $Y = PM_{10}(s)$ (in $\mu g/m^3$) (for each $s = 1, \dots, 182$) represents pollution levels, the functional predictor $Z_i(t)$ represents the daily mean temperature curve recorded at the i th station, with its precise location determined by the coordinates $\mathbf{i} = (UTMX; UTM Y)$, $Z = TEMP(t); t = 1, \dots, 182$ and the parametric part are: $X = (X^1, X^2, X^3)$, (for 182 days). $\delta = 1$ if $PM_{10}(s)$ is observed and $\delta = 0$ otherwise. Note that our data have some missing values (473 NaN of $PM_{10}(s)$ for each $s = 1, \dots, 182$ and each station, about 7.64% missing data).

Figure 5 provides the curves of the functional variable Z_i and Figure 4 represents the spatial position of the 34 monitoring stations in the Piemonte region (northern Italy)

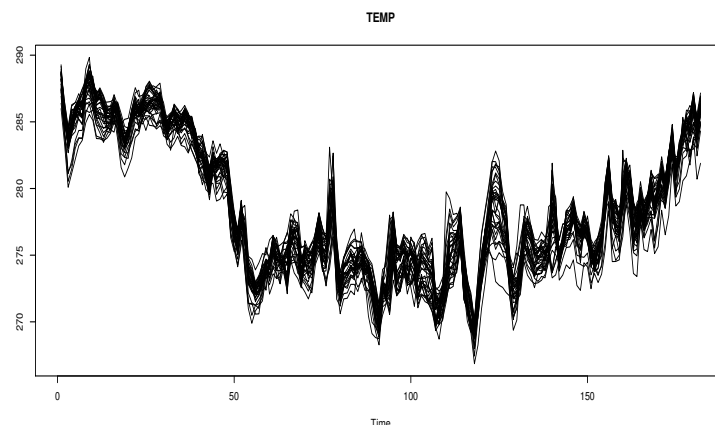


Figure 5. Temperature curves Z .

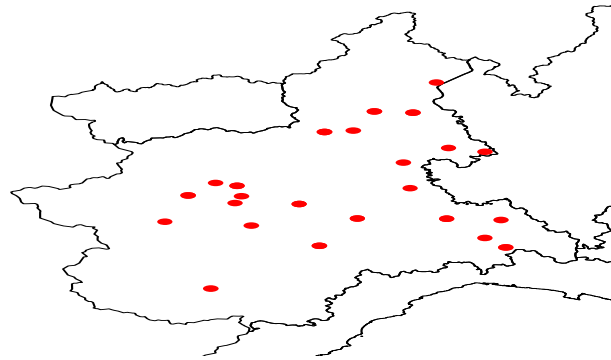


Figure 6. Locations of the stations in Piemonte (northern Italy).

However, implementing this spatial modeling approach requires preliminary data preparation to validate the stationarity assumption that handles spatial heterogeneity resulting from variations in the effects of space on sampled units. To answer this, we will use an "detrending step" introduced by [37], which is designed for the multivariate case of the three variables (response, functional, and vectorial explanatory). This algorithm is defined by the following regression.

$$\tilde{X}_i = m_1(\mathbf{i}) + X_i \quad , \quad \tilde{Z}_i = m_2(\mathbf{i}) + Z_i \quad \text{and} \quad \tilde{Y}_i = m_3(\mathbf{i}) + Y_i.$$

Thus, instead to the initial observations $(X_i, Z_i, Y_i, \delta_i)_i$, we compute the SFPLRM estimator from the statistics $(\tilde{X}_i, \tilde{Z}_i, \tilde{Y}_i, \delta_i)_i$ (see Figure 7). The latter are obtained by

$$\hat{X}_i = \tilde{X}_i - \hat{m}_1(\mathbf{i}) \quad , \quad \hat{Z}_i = \tilde{Z}_i - \hat{m}_2(\mathbf{i}) \quad \text{and} \quad \hat{Y}_i = \tilde{Y}_i - \hat{m}_3(\mathbf{i}),$$

and $\hat{m}_1(\cdot)$, \hat{m}_2 and \hat{m}_3 are the kernel estimators of the regression functions $m_1(\cdot)$, $m_2(\cdot)$ and $m_3(\cdot)$ which are expressed by

$$\hat{m}_1(\mathbf{i}_0) = \frac{\sum_{i \in \mathcal{I}_n} \delta_i X_i H_1(\|\mathbf{i}_0 - \mathbf{i}\|/h_n^1)}{\sum_{i \in \mathcal{I}_n} \delta_i H_1(\|\mathbf{i}_0 - \mathbf{i}\|/h_n^1)} \quad , \quad \hat{m}_2(\mathbf{i}_0) = \frac{\sum_{i \in \mathcal{I}_n} \delta_i Z_i H_2(\|\mathbf{i}_0 - \mathbf{i}\|/h_n^2)}{\sum_{i \in \mathcal{I}_n} \delta_i H_2(\|\mathbf{i}_0 - \mathbf{i}\|/h_n^2)}$$

$$\text{and } \hat{m}_3(\mathbf{i}_0) = \frac{\sum_{i \in \mathcal{I}_n} \delta_i Y_i H_3(\|\mathbf{i}_0 - \mathbf{i}\|/h_n^3)}{\sum_{i \in \mathcal{I}_n} \delta_i H_3(\|\mathbf{i}_0 - \mathbf{i}\|/h_n^3)},$$

where the functions H_j , $j = 1, 2, 3$ represent kernel functions, while h_n^j , $j = 1, 2, 3$ are the bandwidth parameters associated with the actual regression. To illustrate the practical implications of this detrending step on our dataset, we will examine its impact by performing a comparative analysis of the performance of the SFPLRM regression in the two cases, one with detrending and the other without this one.

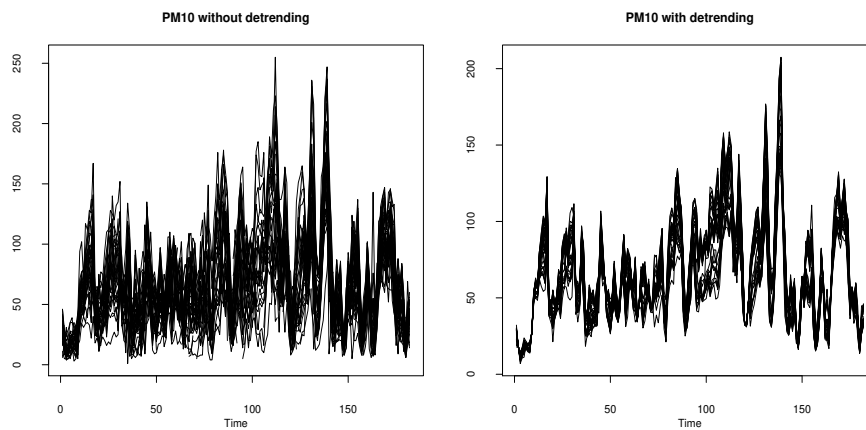


Figure 7. No detrending observations $PM_{10}(s)$ (Y) versus detrending observations $PM_{10}(s)$ (\hat{Y}).

To conduct this analysis, we employ the same methodology as employed in the simulation example for the selection of the estimator's parameters. Specifically, we use the quadratic kernel on the interval $(0, 1)$ in combination with the PCA metric and the cross-validation (CV) criterion to determine the smoothing parameter h_n . For the real regressions $m_1(\cdot)$, $m_2(\cdot)$, and $m_3(\cdot)$, we use the `npreg` routine in the R-package `np`, with $K = H_1 = H_2 = H_3$.

To assess the feasibility of this approach, we randomly split the data sample multiple times (precisely 100 times). The data is divided into two subsets: a learning sample consisting of 24 observations and a test sample containing 10 observations. We then evaluate the significance of the proposed detrending procedure by examining the Mean Squared Error (*MSE*), as used in the simulation example. This analysis allows us to determine the impact of detrending on the performance of the estimators in practice.

The results in Figure 8 reveal a significant enhancement in model accuracy with the inclusion of the detrending step. The Mean Squared Error (*MSE*) values are substantially reduced when detrending is applied, indicating a clear advantage of this data preprocessing technique. As illustrated in Figure 9, it becomes apparent that the detrending step consistently outperforms the non-detrending, clearly showcasing its superior capability to accurately capture the underlying functional relationships. These results underscore the pivotal role of detrending in enhancing model performance and underscore the inherent advantages of the SFPLRM estimator in the realm of non-parametric spatial data analysis.

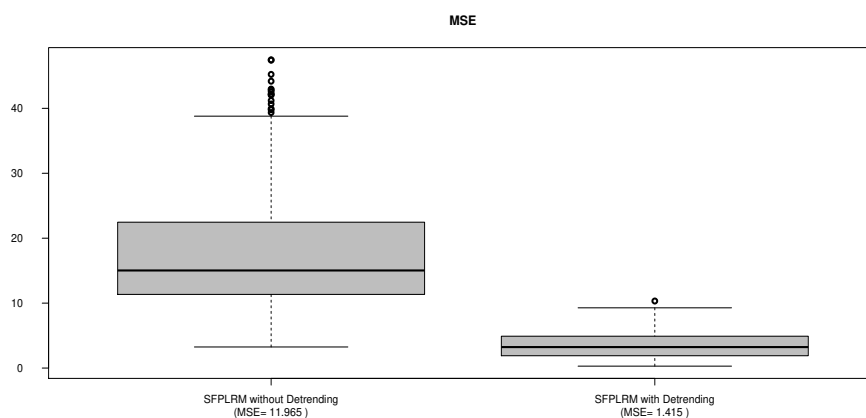


Figure 8. Boxplot of *MSE* Values: Detrending vs. No Detrending

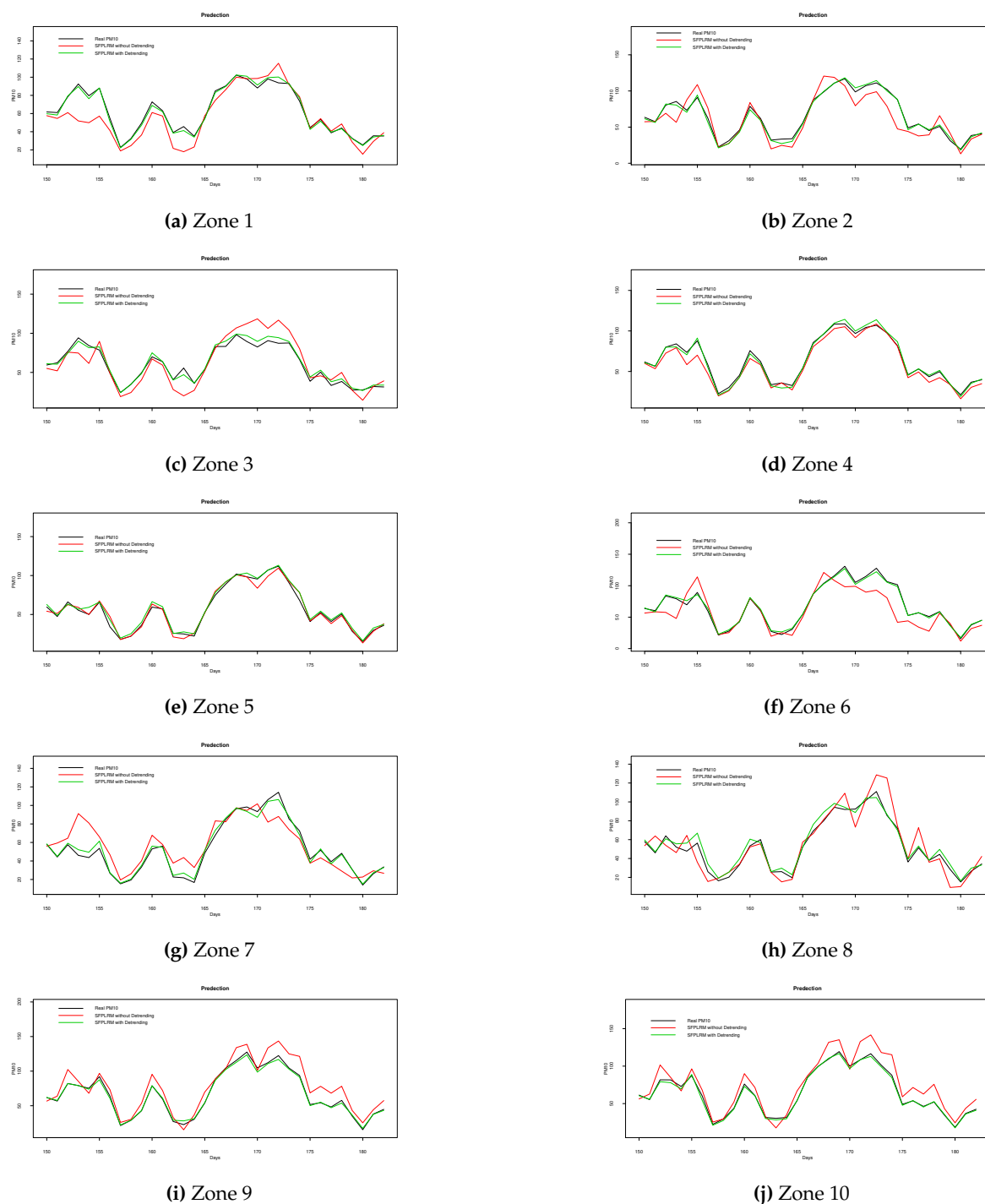


Figure 9. Prediction of the testing sample of the PM_{10} for $s = 150, \dots, 182$ in 10 stations.

8. conclusion

This paper addresses the issue of a semi-functional partial linear regression model for spatial data, assuming that the missing responses occur randomly. The construction of the estimator encompasses both the linear and nonparametric components of the model. One crucial aspect of this study involves the demonstration of the asymptotic normality of the best estimators. This is achieved by imposing certain mild constraints and establishing the probability convergence of the nonparametric component. Furthermore, the utilization of both simulated and real data in the analysis serves to underscore the possible viability and adaptability of the model under investigation, along with the estimators derived

from it, in the context of predictive tasks. This is achieved through a comparative analysis using a non-parametric estimator. It is essential to highlight that significant emphasis was placed on the absence of a missing mechanism that is missing at random within the domain of functional data statistics.

In the subsequent analysis, it would be intriguing to examine the expansion of our framework into alternative avenues. One such direction is expanding the framework from missing at random (MAR) data to censored data, necessitating the utilization of intricate mathematical techniques.

Author Contributions: The authors contributed approximately equally to this work. All authors have read and agreed to the final version of the manuscript. Formal analysis, Tawfik Benchikh; Validation, Omar Fetitah; Writing – review & editing, Ibrahim M. Almanjahie and Mohammad Kadi Attouch.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors thank and extend their appreciation to the funder of this work. This work was supported by the Deanship of Scientific Research at King Khalid University through the Large Research Groups Program under grant number R.G.P. 2/406/44.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mateu, J.; Romano, E. Advances in spatial functional statistics. *Stoch Environ Res Risk Assess*, **2017**, *31*, 1-6.
2. Ramsay, J.; Silverman, B. *Functional Data Analysis*, Second Edition, Springer-Verlag, New York, 2005.
3. Bosq, D.; Blanke, D. *Inference and prediction in large dimension*. Wiley series in probability and statistics. Wiley, Chichester, 2007.
4. Ferraty, F.; Vieu, P. *Nonparametric functional data analysis. Theory and Practice*. Springer Series in Statistics. New York, 2006.
5. Aneiros-Pérez, G.; Horová, I.; Hušková, M.; Vieu P. Editorial for the Special Issue on Functional Data Analysis and Related Fields. *J. Multivariate Anal.*, **2022**, *189*.
6. Ling, N.; Vieu, P. Nonparametric modelling for functional data: selected survey and tracks for future. *Statistics*, **2018**, *52* (4), 934-949.
7. Greven, S.; Scheipl, F. A general framework for functional regression modelling. *Statistical Modelling*, **2017**, *17* (1-2), 1-35.
8. Aneiros-Pérez, G.; Vieu, P. Semi-functional partial linear regression, *Stat. Probab. Lett.*, **2006**, *76*, (11), 1102-1110.
9. Aneiros-Pérez G.; Vieu, P. Nonparametric time series prediction. A semi-functional partial linear modeling. *Journal of Multivariate Analysis*, **2008**, *99*, 834-857.
10. Aneiros-Pérez G.; Vieu, P. Automatic estimation procedure in partial linear model with functional data. *Stat.Pap.*, **2011**, *52* (4), 751-771.
11. H. Lian, Functional partial linear model. *J. Nonparametr. Stat.*, **2011**, *23* (1), 115-128.
12. Aneiros Pérez, G.; Raña, R.; Vieu, P.; Vilar, J. Bootstrap in semi-functional partial linear regression under dependence. *Test*, **2018**, *27* (3), 659-679.
13. Zhao, F-r.; Zhang, B-X. Testing Linearity in Functional Partially Linear Models. *Acta Mathematicae Applicatae Sinica, English Series*, **2022**, DOI <https://doi.org/10.1007/s10255-023-1040-0>.
14. Feng, S.; Xue, L. Partially functional linear varying coefficient model, *Statistics*, **2016**, *50* (4), 717-732.
15. Boente, G.; Vahnovan, A. Robust estimators in semi-functional partial linear regression models. *Journal of Multivariate Analysis*, **2017**, *154* (C), 59-84.
16. Ling, N.; Aneiros-Pérez, G.; Vieu, P. knn estimation in functional partial linear modeling. *Statist. Papers*, **2020**, *61* (1), 423-444.
17. Shang, H. Bayesian bandwidth estimation for a semi-functional partial linear regression model with unknown error density. *Comput. Stat.*, **2014**, *29* (3-4), 829-848.
18. Ling, N.; Vieu, P. On semiparametric regression in functional data analysis. *WIREs Computational Statistics*, **2020**, *12* (6), 20-30.

19. Li, Y.; Ying, C. Semi-functional partial linear spatial autoregressive model. *Communication in Statistics-Theory and Methods*, **2021**, *50*, 5941-5954.
20. Benallou, M.; Attouch, M. K.; Benchikh, T.; Fetitah, O. Asymptotic results of semi-functional partial linear regression estimate under functional spatial dependency. *Communications in Statistics-Theory and Methods*, **2021**, *51*, 1-21.
21. Graham, J. W. Missing data analysis and design. NY: Springer, New York. 2020.
22. Little, R. J. A.; Rubin, D. B. Statistical Analysis with Missing Data. 3rd Edition; Wiley Series in Probability and Statistics. 2020.
23. Ferraty, F.; Sued, F.; Vieu, P. Mean estimation with data missing at random for functional covariables. *Statistics*, **2013**, *47* (4), 688-706.
24. Cheng, P.E. Nonparametric estimation of mean functionals with data missing at random. *J. Amer. Statist. Assoc.* **1994**, *89*, 81-87.
25. Ling, N.; Liang, L.; Vieu, P. Nonparametric regression estimation for functional stationary ergodic data with missing at random. *J. Stat. Plan. Inference*, **2015**, *162*, 75-87.
26. Rachdi, M.; Laksaci, A.; Kaid, Z.; Benchiha, A.; Fahimah, A. k-Nearest neighbors local linear regression for functional and missing data at random. *Statistica Neerlandica, Netherlands Society for Statistics and Operations Research*, **2021**, *75* (1), 42-65.
27. Wang, Q.H.; Linton, O.; Wolfgang H., Semiparametric regression analysis with missing response at random. *J. Amer. Statist. Assoc.*, **2004**, *99*, 334-345.
28. Wang, Q.H., Sun, Z. Estimation in partially linear models with missing responses at random. *J. Multivariate Anal.*, **2007**, *98*, 1470-1493.
29. Ling, N.; Kan, R.; Vieu, P.; Meng, S. Semi-functional partially linear regression model with responses missing at random. *Metrika*, **2019**, *82* (1), 39-70.
30. Haworth, J.; Cheng, T. Non-parametric regression for space-time forecasting under missing data. *Computers, Environment and Urban Systems*, **2012**, *36* (6), 538-550.
31. Puranik, A.; Binub, V.S.; Seena, B. Estimation of missing values in aggregate level spatial data. *Clinical Epidemiology and Global Health journal*, **2021**, *9*, 304-309.
32. Alshahrani, F.; Almanjahi, I. M.; Benchikh, T.; Fetitah, O.; Attouch; M. K. Asymptotic normality of nonparametric kernel regression estimation for missing at random functional spatial data. *Journal of Mathematics, Hindawi*, **2023**, <https://doi.org/10.1155/2023/8874880>.
33. Carbon, M.; Tran, L. T.; Wu, B. Kernel density estimation for random fields density estimation for random fields. *Stat Probab Lett*, **1997**, *36* (2), 115-125.
34. Dabo-Niang, S.; Rachdi, M.; Yao, A.F. Kernel regression estimation for spatial functional random variables. *Far East Journal of Theoretical Statistics*, **2011**, *37* (2), 77-113.
35. Lin, Z.; Lu, C. Limit theory for mixing dependent random variables. Dordrecht: Kluwer. 1996.
36. Cameletti, M.; Ignaccolo, R.; Bande, S. Comparing spatio-temporal models for particulate matter in Piemonte, *Environmetrise*, **2011**, *22* (8), 985-996.
37. Hallin, M.; Lu, Z.; Yu, K. Local Linear Spatial Quantile Regression. *Bernoulli*, **2009**, *15* (3), 659-686.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.