Stochastic Environmental Research and Risk Assessment
https://doi.org/10.1007/s00477-019-01712-z

ORIGINAL PAPER

Prediction of air pollutants PM_{10} by ARBX(1) processes

J. Álvarez-Liébana[1] · M. D. Ruiz-Medina[2]

Abstract
This work adopts a Banach-valued time series framework for component-wise estimation and prediction, from temporal correlated functional data, in presence of exogenous variables. The strong-consistency of the proposed functional estimator and associated plug-in predictor is formulated. The simulation study undertaken illustrates their large-sample size properties. Air pollutants PM_{10} curve forecasting, in the Haute-Normandie region (France), is addressed by implementation of the functional time series approach presented.

Keywords Air pollutants forecasting · Banach spaces · Functional time series · Meteorological variables · Strong consistency

1 Introduction

Several approaches have been adopted in the analysis of pollution data (see, e.g., Pang et al. 2009, for a comparative study). In Zhang et al. (2018), the singular value decomposition is applied to identify spatial air pollution index (API) patterns, in relation to meteorological conditions in China. A novel hybrid model combining Multilayer perceptron model and Principal Component Analysis (PCA) is introduced in He et al. (2015), to improve the air quality prediction accuracy in urban areas. Factor analysis and Box–Jenkins methodology are considered in Gocheva-Ilieva et al. (2014), to examine concentrations of primary air pollutants such as NO, NO_2 , NO_x , PM_{10}, SO_2 and ground level O_3 in the town of Blagoevgrad, Bulgaria. Since PM_{10} are inhalable atmospheric particles, their forecasting has became crucial aimed at adopting efficient public transport policies. In the recent literature, one can find several modelling approaches for PM_{10} forecasting. Among the most common statistical techniques applied, we mention multiple regression, non-linear state space modelling and artificial neural networks (see, e.g., Grivas and Chaloulakou 2006; Paschalidou et al. 2011; Slini et al. 2006; Stadlober et al. 2008; Zolghadri and Cazaurang 2006). Functional Data Analysis (FDA) techniques also play a crucial role in air quality forecasting (see Febrero-Bande et al. 2008; Fernández de Castro et al. 2005; Ignaccolo et al. 2014, among others). Related meteorological variables can also be functional predicted from a functional time series framework (see, e.g., Besse et al. 2000; Ruiz-Medina and Espejo 2012; Ruiz-Medina et al. 2014).

Computational advances have made possible the implementation of flexible models for random elements in function spaces. FDA techniques have emerged in the local analysis of high-dimensional data, which are functional in nature (see the monographs Goia and Vieu 2016; Horváth and Kokoszka 2012; Hsing and Eubank 2015, and the references therein). Parametric functional linear time series techniques are fast and computational low-cost. In contrast with the more flexible nonparametric functional statistical approach (see, e.g., Ferraty and Vieu 2006), where the so-called curse of dimensionality problem arises (see Geenens 2011; Vieu 2018). The semi-parametric framework also provides a partial solution to this problem (see, e.g., Aneiros-Pérez and Vieu 2008; Goia and Vieu 2015, in the semi-functional regression context). Particularly, the approach presented allows a flexible analysis of the local

✉ M. D. Ruiz-Medina
mruiz@ugr.es

J. Álvarez-Liébana
alvarezljavier@uniovi.es

[1] Department of Statistics, O.R. and Didactics of Mathematics, University of Oviedo, Oviedo, Spain

[2] Department of Statistics and O.R., University of Granada, Granada, Spain

Published online: 05 August 2019

variability of the functional values of the random variables studied, as well as the derivation of strongly-consistent functional plug-in predictors, under a state space based framework.

From a theoretical point of view, parametric functional linear time series techniques have been widely studied in the last few decades. In particular, in the autoregressive Hilbertian process framework, the asymptotic properties of componentwise estimators of the autocorrelation operator, and their associated plug-in predictors have been derived in Bosq (2000), Mas (2004, 2007), among others. Recently, in Álvarez-Liébana et al. (2017) and Ruiz-Medina and Álvarez-Liébana (2019a), alternative operator norms for consistency have been investigated. See also Álvarez-Liébana et al. (2016), for the case of Ornstein–Uhlenbeck process in Hilbert and Banach spaces.

The separable Banach space context has also been adopted in linear functional time series modeling, under a state space based approach. This literature has mainly been focused on the spaces of continuous functions $\mathcal{C}([0,1])$ with the supremum norm (see Bueno-Larraz and Klepsch 2018; Dehling and Sharipov 2005; Labbas and Mourid 2002; Parvardeh et al. 2017, among others), and on the Skorokhod space of right-continuous functions on [0, 1], having limit at the left at each $t \in [0,1]$, equipped with the Skorokhod topology, usually denoted as $\mathcal{J}_1$-topology (see, e.g., Blanke and Bosq 2016; El Hajj 2011). The lack of an inner product structure, in the abstract Banach-valued time series framework, is supplied in Ruiz-Medina and Álvarez-Liébana (2019b) by considering suitable embeddings into related Hilbert spaces. Strong consistency of a componentwise estimator of the autocorrelation operator and its associated plug-in predictor is then proved.

A first attempt for the inclusion of exogenous information in the functional time series framework can be found in Damon and Guillas (2002, 2005), where the so-called ARHX(1) processes (Hilbert-valued autoregressive processes of order one with exogenous variables) are introduced. Enhancements were subsequently proposed by Marion and Pumo (2004). First order conditional autoregressive Hilbertian processes were introduced in Guillas (2002). The present paper extends the time series framework in Ruiz-Medina and Álvarez-Liébana (2019b) to the case of first-order Banach-valued autoregressive processes with exogenous variables (ARBX(1) processes). Functional parameter estimation and plug-in prediction can be addressed in our ARBX(1) context, from the multivariate infinite-dimensional formulation of the results in Ruiz-Medina and Álvarez-Liébana (2019b). Specifically, a matrix-operator-based formulation of the ARB(1) process (Banach-valued autoregressive process of order one) state equation is considered. The required Hilbert space embeddings, and sufficient conditions for the strong-consistency of the autocorrelation operator estimator (reflecting temporal correlations between endogenous and exogenous variables), and the associated plug-in functional predictor are then obtained in a direct way. We refer to the reader to Triebel (1983), where several examples of the Banach space context introduced in Ruiz-Medina and Álvarez-Liébana (2019b) can be found.

The outline of the paper is as follows. The ARBX(1) based estimation and prediction methodologies presented are described in Sect. 2. This section also contains the main results of this paper (see Theorem 1). A simulation study is undertaken to illustrate the consistency of ARBX(1) predictors in Sect. 3. PM_{10} short-term forecasting, based on the introduced ARBX(1) framework, is addressed in Sect. 4. Final comments are provided in Sect. 5.

2 ARBX(1) estimation and prediction

In the following, the functional random variables and stochastic processes introduced below are defined on the basic probability space $(\Omega, \mathcal{A}, \mathcal{P})$. Let $\left(B, \|\cdot\|_B\right)$ be a real separable Banach space with associated norm $\|\cdot\|_B$. Consider $X = \{X_n,\ n \in \mathbb{Z}\}$ to be a zero-mean ARB(1) process, with $\mathcal{P}(X_n \in B) = 1, n \in \mathbb{Z}$, satisfying the following state equation (see, e.g., Bosq 2000):

$$X_n = \rho(X_{n-1}) + \varepsilon_n, \quad n \in \mathbb{Z}, \tag{1}$$

where ρ is the autocorrelation operator, which is assumed to be a bounded linear operator on B, that is, $\rho \in \mathcal{L}(B)$, with $\left(\mathcal{L}(B), \|\cdot\|_{\mathcal{L}(B)}\right)$ denoting the Banach space of continuous operators with the supremum norm. Here, $\varepsilon = \{\varepsilon_n,\ n \in \mathbb{Z}\}$ represents the innovation process, which is assumed to be a B-valued strong white noise, and uncorrelated with the random initial condition. In particular, $\sigma_\varepsilon^2 = \mathrm{E}\left[\|\varepsilon_n\|_B^2\right] < \infty,\ n \in \mathbb{Z}$. From Bosq (2000, Theorem 6.1), if there exists $j_0 \geq 1$ such that $\|\rho^j\|_{\mathcal{L}(B)} < 1$, for every $j \geq j_0$, then, Eq. (1) admits a unique stationary solution $X_n = \sum_{j=0}^{\infty} \rho^j(\varepsilon_{n-j})$, with $\sigma_X^2 = \mathrm{E}\left[\|X_n\|_B^2\right] < \infty$, $n \in \mathbb{Z}$.

In this paper, exogenous information is incorporated to Eq. (1) in an additive way. Thus, the state space equation of an ARBX(1) process is given by:

$$X_n = \rho(X_{n-1}) + \sum_{i=1}^{b} a_i(Z_{n,i}) + \varepsilon_n, \quad n \in \mathbb{Z}, \tag{2}$$

where $\{a_i,\ i = 1, \ldots, b\}$ are bounded linear operators on B. The exogenous functional random variables $Z_i = \{Z_{n,i},\ n \in \mathbb{Z}\}, i = 1, \ldots, b$, are assumed to satisfy the following ARB(1) equation, for $i = 1, \ldots, b$,

$$Z_{n,i} = u_i\left(Z_{n-1,i}\right) + \eta_{n,i}, \quad u_i \in \mathcal{L}(B), \quad n \in \mathbb{Z}. \tag{3}$$

For $i = 1, \ldots, b$, $\eta_i = \{\eta_{n,i}\ n \in \mathbb{Z}\}$ is a B-valued strong white noise. Particularly, $\sigma^2_{\eta_i} = \mathrm{E}\left[\left\|\eta_{n,i}\right\|^2_B\right] < \infty, n \in \mathbb{Z}$, $i = 1, \ldots, b$. Here, $\mathcal{P}\left(Z_{n,i} \in B\right) = 1$, $\mathrm{E}\left[Z_{n,i}\right] = 0_B$, $n \in \mathbb{Z}$, for $i = 1, \ldots, b$. The symbol 0_B means the zero element (i.e., null function) in B. Equations (2)–(3) can be rewritten as (see Damon and Guillas 2002),

$$\begin{aligned} &\overline{X}_n = \overline{\rho}\left(\overline{X}_{n-1}\right) + \overline{\varepsilon}_n, \quad \overline{\rho} \in \mathcal{L}(\overline{B}), \\ &\mathcal{P}\left(\overline{X}_n \in \overline{B}\right) = \mathcal{P}\left(\overline{\varepsilon}_n \in \overline{B}\right) = 1, \end{aligned} \tag{4}$$

where $\overline{B} = B^{b+1}$ is also a real separable Banach space, and

$$\begin{aligned} \overline{X}_n &= \begin{Bmatrix} X_n \\ Z_{n+1,1} \\ Z_{n+1,2} \\ \vdots \\ Z_{n+1,b} \end{Bmatrix}, \quad \overline{\varepsilon}_n = \begin{Bmatrix} \varepsilon_n \\ \eta_{n,1} \\ \eta_{n,2} \\ \vdots \\ \eta_{n,b} \end{Bmatrix}, \\ \overline{\rho} &= \begin{Bmatrix} \rho & a_1 & \cdots & \cdots & a_b \\ \mathbf{0}_B & u_1 & \mathbf{0}_B & \cdots & \mathbf{0}_B \\ \mathbf{0}_B & \mathbf{0}_B & u_2 & \mathbf{0}_B & \vdots \\ \vdots & \vdots & & \ddots & \mathbf{0}_B \\ \mathbf{0}_B & \mathbf{0}_B & \mathbf{0}_B & \cdots & u_b \end{Bmatrix}. \end{aligned} \tag{5}$$

Here, $\mathbf{0}_B$ represents the null operator on B. In Eq. (4), $\mathcal{L}(\overline{B})$ denotes the space of bounded linear operators on $\overline{B}$. The norm in the separable Banach space $\overline{B}$ is given by

$$\begin{aligned} \|\overline{y}\|_{\overline{B}} &= \sup_{n \geq 1} \left|\overline{F}_n(\overline{y})\right| = \sup_{n \geq 1} \sup_{i \in \{1, \ldots, b+1\}} |F_{ni}(y_i)|, \\ \overline{F}_n &= \left(F_{n1}, \ldots, F_{n(b+1)}\right), \end{aligned} \tag{6}$$

for every $\overline{y} = (y_1, \ldots, y_{b+1}) \in \overline{B} = B^{b+1}$. Here, for $i = 1, \ldots, b+1$, $\{F_{ni},\ n \geq 1\} \subset B^{\star}$ is a sequence of bounded linear functionals on B satisfying

$$F_{ni}(x_{ni}) = \|x_{ni}\|_B, \quad \|F_{ni}\| = 1, \quad n \geq 1, \tag{7}$$

with $\{x_{ni},\ n \geq 1\} \subset B$ being a dense sequence in B (see Lemma 2.1 in Kuelbs 1970 for more details). For simplification purposes, we consider a common dense system in B, i.e., $x_n = x_{ni}$, and $F_{ni} = F_n$, for $i = 1, \ldots, b+1$, and $n \geq 1$. We assume the conditions ensuring the existence of a unique stationary solution to Eq. (4). That is, assume that there exists a j_0 such that $\|\overline{\rho}^j\|_{\mathcal{L}(\overline{B})} < 1$, for all $j \geq j_0$. The following componentwise estimator of the autocorrelation operator $\overline{\rho}$ in (4), based on a functional sample of size n, is then formulated:

$$\begin{aligned} \widetilde{\overline{\rho}}_{k_n}(\overline{x}) &= \left(\widetilde{\overline{\Pi}}^{k_n} \overline{D}_n \overline{C}_n^{-1} \widetilde{\overline{\Pi}}^{k_n}\right)(\overline{x}) \\ &= \left(\sum_{j=1}^{k_n} \frac{1}{C_{n,j}} \langle \overline{x}, \overline{\phi}_{n,j} \rangle_{\widetilde{\overline{H}}} \widetilde{\overline{\Pi}}^{k_n} \overline{D}_n(\overline{\phi}_{n,j})\right), \end{aligned} \tag{8}$$

where for $j \geq 1$,

$$\begin{aligned} &\langle \overline{x}, \overline{\phi}_{n,j} \rangle_{\widetilde{\overline{H}}} = \langle x, \phi_{n,j1} \rangle_{\widetilde{H}} + \sum_{i=1}^{b} \left\langle x_i, \phi_{n,j(i+1)} \right\rangle_{\widetilde{H}}, \\ &\forall \overline{x} = (x, x_1, \ldots, x_b) \in \overline{B}, \end{aligned}$$

with $\{\overline{\phi}_{n,j} = (\phi_{n,j1}, \ldots, \phi_{n,j(b+1)}),\ j \geq 1\}$ being the orthonormal eigenvector system associated with

$$\overline{C}_n = \frac{1}{n} \sum_{i=1}^{n} \overline{X}_i \otimes \overline{X}_i,$$

the empirical autocovariance operator of the extended version to $\widetilde{\overline{H}} = \widetilde{H}^{b+1}$, of $\overline{X} = \{\overline{X}_n,\ n \in \mathbb{Z}\}$. Here, the Hilbert space $\widetilde{H}$ is defined in Lemma 2.1 in Kuelbs (1970), as the continuous extension of the separable real-valued Banach space B (see also Lemma 1 in Ruiz-Medina and Álvarez-Liébana 2019b). In particular, its inner product is given by $\langle f, g \rangle_{\widetilde{H}} = \sum_{n=1}^{\infty} t_n F_n(f) F_n(g)$, for $f, g \in \widetilde{H}$, with $\sum_{n=1}^{\infty} t_n = 1$, $t_n > 0, n \geq 1$. Note that $\widetilde{H}$ has weaker topology than B, allowing the continuous inclusion $B \hookrightarrow \widetilde{H}$, and hence, $\overline{B} \hookrightarrow \widetilde{\overline{H}}$, holds. In (8), for each functional sample size n, we have denoted

$$\widetilde{\overline{\Pi}}^{k_n}(\overline{x}) = \sum_{j=1}^{k_n} \langle \overline{x}, \overline{\phi}_{n,j} \rangle_{\widetilde{\overline{H}}} \overline{\phi}_{n,j}, \quad \forall \overline{x} \in \overline{B} = B^{b+1} \subset \widetilde{\overline{H}} = \widetilde{H}^{b+1}. \tag{9}$$

Denote also by $\{C_{n,j},\ j \geq 1\}$, with $C_{n,1} \geq \cdots \geq C_{n,n} \geq 0 = C_{n,n+1} = C_{n,n+2} = \cdots$, the eigenvalues of $\overline{C}_n$ respective associated with the empirical eigenvectors $\{\overline{\phi}_{n,j},\ j \geq 1\}$. The operator $\overline{D}_n = \frac{1}{n-1} \sum_{i=1}^{n-1} \overline{X}_i \otimes \overline{X}_{i+1}$ denotes the empirical cross-covariance operator of the extended version of $\overline{X}_n$ to $\widetilde{\overline{H}}$.

As in Ruiz-Medina and Álvarez-Liébana (2019b), sufficient conditions are now formulated to ensure the strong consistency in the space $\mathcal{L}(\overline{B})$ (i.e., with respect to the supremum norm in $\mathcal{L}(\overline{B})$) of the componentwise functional parameter estimator (8) of $\overline{\rho}$, defined in terms of the countable orthogonal system $\{\overline{\phi}_{n,j},\ j \geq 1\}$ in $\widetilde{\overline{H}}$. Specifically, the following conditions are assumed:

- **Assumption A1** $\|\overline{X}_0\|_{\overline{B}}$ is almost surely bounded. The eigenspaces associated with the eigenvalues of $\overline{C} = E\left[\overline{X}_n \otimes \overline{X}_n\right]$ are one-dimensional.

- **Assumption A2** Let k_n be such that $C_{n,k_n} > 0$ a.s., and both $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. Here, C_{n,k_n} denotes the k_n-th eigenvalue of $\overline{C}_n$.
- **Assumption A3** As $k \to \infty$,

$$\sup_{\overline{x}\in\overline{B},\ \|\overline{x}\|_{\overline{B}}\leq 1}\left\|\overline{\rho}(\overline{x})-\sum_{j=1}^{k}\langle\overline{\rho}(\overline{x}),\overline{\phi}_j\rangle_{\widetilde{\overline{H}}}\widetilde{\overline{\phi}}_j\right\|_{\overline{B}}\to 0,$$

$$\overline{\phi}_j=(\phi_{j1},\ldots,\phi_{j(b+1)}),\ j\geq 1,$$

  where $\overline{C}(\overline{\phi}_j)=C_j\overline{\phi}_j$, $j\geq 1$, in $\widetilde{\overline{H}}$.
- **Assumption A4** Denote by $\mathcal{H}(\overline{X})$ the Reproducing Kernel Hilbert Space (RKHS) generated by $\overline{C}$. The inclusion of $\mathcal{H}(\overline{X})$ into $\widetilde{\overline{H}}^{\star}$ is continuous, i.e., $\mathcal{H}(\overline{X})\hookrightarrow\widetilde{\overline{H}}^{\star}$, is a continuous mapping, where $\widetilde{\overline{H}}^{\star}$ denotes the dual Hilbert space of $\widetilde{\overline{H}}$.
- **Assumption A5** The embedding $i_{\mathcal{H}(\overline{X}),\overline{H}}:\mathcal{H}(\overline{X})\hookrightarrow\overline{H}$ is Hilbert–Schmidt. Here, $\overline{H}=H^{b+1}$, with H being a real separable Hilbert space such that $\widetilde{\overline{H}}^{\star}\hookrightarrow\overline{H}\hookrightarrow\widetilde{\overline{H}}$ conforms a Rigged Hilbert space structure or Gelfand triple.

The following result provides the strong consistency of the componentwise estimator $\widetilde{\overline{\rho}}_{k_n}$ of $\overline{\rho}$, as well as of its associated plug-in predictor $\widetilde{\overline{\rho}}_{k_n}(\overline{X}_n)$ of $\overline{X}_{n+1}$.

Theorem 1 *Let X be the ARBX(1) process introduced in Eqs. (4)–(5). Under Assumptions A1–A5, and the conditions assumed in* Ruiz-Medina and Álvarez-Liébana (2019b, Lemmas 8–9), *for all $\eta>0$,*

$$\mathcal{P}\left(\|\widetilde{\overline{\rho}}_{k_n}-\overline{\rho}\|_{\mathcal{L}(\overline{B})}\geq\eta\right)\leq\mathcal{K}\exp(-n\eta^2/Q_n),$$

where

$$Q_n=\mathcal{O}\left\{\left(C_{k_n}^{-1}k_n\sum_{j=1}^{k_n}a_j\right)^2\right\}$$

as $n\to\infty$. Here,

$$a_1=2\sqrt{2}\,\frac{1}{C_1-C_2},$$
$$a_j=2\sqrt{2}\,\max\left(\frac{1}{C_{j-1}-C_j},\frac{1}{C_j-C_{j+1}}\right),\quad j\geq 2. \tag{10}$$

Therefore, if

$$k_nC_{k_n}^{-1}\sum_{j=1}^{k_n}a_j=o\{\sqrt{n/\ln(n)}\},$$

as $n\to\infty$, then, $\|\overline{\rho}-\widetilde{\overline{\rho}}_{k_n}\|_{\mathcal{L}(\overline{B})}\to_{a.s}0$, and

$$\left\|(\overline{\rho}-\widetilde{\overline{\rho}}_{k_n})(\overline{X}_n)\right\|_{\overline{B}}\to^{a.s.}0,\quad n\to\infty,$$

where $\to_{a.s.}$ means the almost surely convergence.

The proof follows as in Ruiz-Medina and Álvarez-Liébana (2019b, Theorem 1), considering the special case of the separable Banach space $\overline{B}=B^{b+1}$.

3 Simulation study

The goal of the simulation study undertaken is to illustrate the flexibility, and the large-sample-size properties of the ARBX(1) parameter estimator, and associated functional plug-in predictor. The effect of the discretization step size is investigated as well. Note that, under the conditions assumed, the function space scenario selected (fractional Besov and, in particular, Sobolev spaces) is quite flexible (see "Appendix" below).

The following ARBX(1) process has been generated:

$$\overline{X}_n=\overline{\rho}(\overline{X}_{n-1})+\overline{\varepsilon}_n,\quad n\in\mathbb{Z},$$

considering, without loss of generality, $b=3$ (i.e., three exogenous variables) in the general formulation (33) provided in the "Appendix". Hence,

$$\overline{X}_n=\begin{Bmatrix}X_n\\Z_{n+1,1}\\Z_{n+1,2}\\Z_{n+1,3}\end{Bmatrix},\quad \overline{\varepsilon}_n=\begin{Bmatrix}\varepsilon_n\\\eta_{n,1}\\\eta_{n,2}\\\eta_{n,3}\end{Bmatrix},$$

$$\overline{\rho}=\begin{Bmatrix}\rho & a_1 & a_2 & a_3\\ \mathbf{0}_B & u_1 & \mathbf{0}_B & \mathbf{0}_B\\ \mathbf{0}_B & \mathbf{0}_B & u_2 & \mathbf{0}_B\\ \mathbf{0}_B & \mathbf{0}_B & \mathbf{0}_B & u_3\end{Bmatrix}.$$

where X_n, and $Z_{n,i},\eta_{n,i},i=1,2,3$, are also valued in $\mathcal{B}^0_{\infty,\infty}([0,1])$, for each $n\in\mathbb{Z}$ (see again the "Appendix" section for more details). Figure 1 displays the values of $X_n^{\star}$ in $\widetilde{H}^{\star}=\mathcal{B}^{\beta}_{2,2}([0,1])$, obtained by smoothing, with *cubicspline* option of `fit.m MatLab` function, the functional values of X_n at times $n=5000,25{,}000,50{,}000,100{,}000$.

The smoother functional values of the exogenous random variables are extended to the space $H=L^2([0,1])$, by projection into the elements of the basis

$$\phi_j(x)=\sqrt{\frac{2}{b-a}}\sin\left(\frac{\pi jx}{b-a}\right),\quad j\geq 1,\quad x\in[a,b],\quad a=0,\ b=1. \tag{11}$$

To ensure the first part of Assumption A1 holds, a truncated multivariate infinite-dimensional Gaussian measure in $\overline{H}=H^4$ is generated:

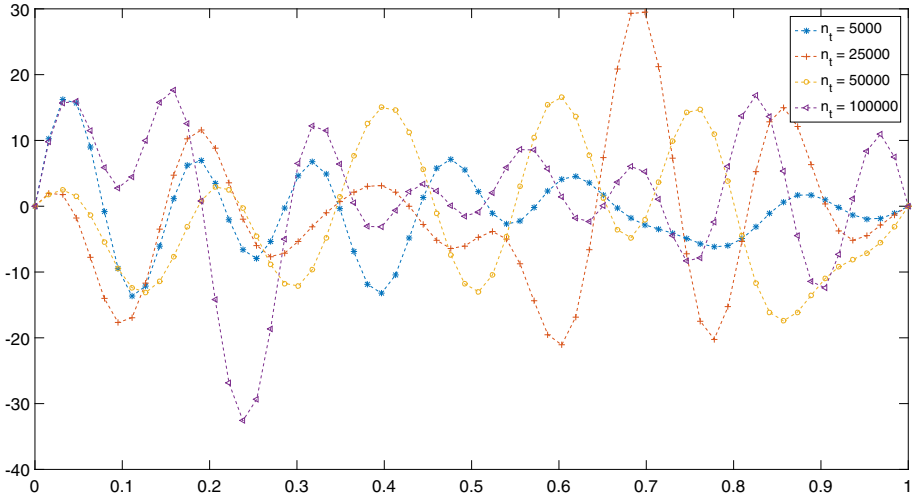

Fig. 1 Functional values of $X_n^{\star}$ at $n = 5000, 25{,}000, 50{,}000, 100{,}000$, with discretization step size $\Delta h = 0.0159$

$$\overline{X}_0 = \begin{Bmatrix} X_0 \\ Z_{1,1} \\ Z_{1,2} \\ Z_{1,3} \end{Bmatrix} \sim \mathcal{N}_T\left(\mathbf{0}, \overline{C}\right), \qquad \overline{\varepsilon}_0 = \begin{Bmatrix} \varepsilon_0 \\ \eta_{0,1} \\ \eta_{0,2} \\ \eta_{0,3} \end{Bmatrix} \sim \mathcal{N}_T\left(\mathbf{0}, \overline{C}_{\boldsymbol{\eta}}\right), \tag{12}$$

where

$$\overline{C} = \begin{Bmatrix} C_{X_0,X_0} & C_{X_0,Z_{1,1}} & C_{X_0,Z_{1,2}} & C_{X_0,Z_{1,3}} \\ C_{Z_{1,1},X_0} & C_{Z_{1,1},Z_{1,1}} & C_{Z_{1,1},Z_{1,2}} & C_{Z_{1,1},Z_{1,3}} \\ C_{Z_{1,2},X_0} & C_{Z_{1,2},Z_{1,1}} & C_{Z_{1,2},Z_{1,2}} & C_{Z_{1,2},Z_{1,3}} \\ C_{Z_{1,3},X_0} & C_{Z_{1,3},Z_{1,1}} & C_{Z_{1,3},Z_{1,2}} & C_{Z_{1,3},Z_{1,3}} \end{Bmatrix}, \tag{13}$$

and

$$\overline{C}_{\boldsymbol{\eta}} = \begin{Bmatrix} C_{\varepsilon_0,\varepsilon_0} & C_{\varepsilon_0,\eta_{0,1}} & C_{\varepsilon_0,\eta_{0,2}} & C_{\varepsilon_0,\eta_{0,3}} \\ C_{\eta_{0,1},\varepsilon_0} & C_{\eta_{0,1},\eta_{0,1}} & C_{\eta_{0,1},\eta_{0,2}} & C_{\eta_{0,1},\eta_{0,3}} \\ C_{\eta_{0,2},\varepsilon_0} & C_{\eta_{0,2},\eta_{0,1}} & C_{\eta_{0,2},\eta_{0,2}} & C_{\eta_{0,2},\eta_{0,3}} \\ C_{\eta_{0,3},\varepsilon_0} & C_{\eta_{0,3},\eta_{0,1}} & C_{\eta_{0,3},\eta_{0,2}} & C_{\eta_{0,3},\eta_{0,3}} \end{Bmatrix}. \tag{14}$$

In Eq. (13), we have denoted

$$C_{X_0,X_0} = (I - \Delta)^{-\gamma}, \quad \gamma > 2\beta > 1, \quad (\beta > 1/2), \tag{15}$$

with $(I - \Delta)^{-\gamma}$ being the Bessel potential of order 2γ, and

$$C_{Z_{1,l},Z_{1,i}} = \mathrm{E}[Z_{1,l} \otimes Z_{1,i}], \quad C_{Z_{1,i},X_0} = \mathrm{E}[Z_{1,i} \otimes X_0], \quad C_{X_0,Z_{1,i}} = \mathrm{E}[X_0 \otimes Z_{1,i}],$$

for $i, l = 1, 2, 3$. Also, in Eq. (14),

$$\begin{aligned} &C_{\varepsilon_0,\varepsilon_0} = \mathrm{E}[\varepsilon_0 \otimes \varepsilon_0], \quad C_{\eta_{0,i},\eta_{0,l}} = \mathrm{E}\left[\eta_{0,i} \otimes \eta_{0,l}\right], \\ &\quad i, l = 1, 2, 3 \\ &C_{\eta_{0,i},\varepsilon_0} = \mathrm{E}\left[\eta_{0,i} \otimes \varepsilon_0\right], \quad C_{\varepsilon_0,\eta_{0,i}} = \mathrm{E}\left[\varepsilon_0 \otimes \eta_{0,i}\right], \\ &\quad i = 1, 2, 3. \end{aligned} \tag{16}$$

The functional entries of (13) are now explicitly defined, ensuring, in particular, the one-dimensionality of their eigenspaces, in order to get the second part of Assumption A1 to hold. Furthermore, for every $f \in H = L^2([0,1])$, and for $i, l = 1, 2, 3$,

$$\begin{aligned} C_{X_0,X_0}(f) &= \sum_{j=1}^{\infty} (1+j)^{-\gamma_1} \langle \phi_j, f \rangle_H \phi_j, \\ C_{X_0,Z_{1,i}}(f) &= \sum_{j=1}^{\infty} (1+j)^{-\frac{\gamma_1+\gamma_{i+1}}{2}} \langle \phi_j, f \rangle_H \phi_j, \\ C_{Z_{1,i},X_0}(f) &= \sum_{j=1}^{\infty} (1+j)^{-\frac{\gamma_1+\gamma_{i+1}}{2}} \langle \phi_j, f \rangle_H \phi_j, \\ C_{Z_{1,l},Z_{1,i}}(f) &= \sum_{j=1}^{\infty} (1+j)^{-\frac{\gamma_{l+1}+\gamma_{i+1}}{2}} \langle \phi_j, f \rangle_H \phi_j. \end{aligned} \tag{17}$$

In the selection of the model parameters, in the ARBX(1) process generation, according to Assumptions A4–A5, $\beta > 1/2$, and $\gamma_i > 2\beta, i = 1, 2, 3, 4$ (see also the "Appendix"). In particular, we have considered $\beta = 3/5$. In those generations, to illustrate the flexibility of the modeling approach presented, the most local singular behaviour corresponds to the functional values of the endogenous variable (see also the air pollutants ARBX(1)-based data analysis implemented in the next section). Specifically, the γ_i, $i = 1, 2, 3, 4$, parameter values, characterizing the local regularity of the functions in the RKHSs of the endogenous and exogenous variables, have been assigned according to the following scheme:

$$\gamma = \{\gamma_i = 2\beta + f(i),\ i = 1, \ldots, b+1\}, \quad f(i) > 0, i = 1, \ldots, b+1. \tag{18}$$

Particularly, the results displayed in Tables 1 and 2 correspond to the following two parametric families:

$$\gamma^1 = \left\{\gamma_i = 2\beta + \frac{i}{10},\ i = 1,2,3,4\right\}, \quad \gamma^2 = \{\gamma_i = 2\beta + \log_{10}(i+1),\ i = 1,2,3,4\}. \tag{19}$$

Thus, as in Sect. 4, a more regular local behaviour is assumed for the functional values of the exogenous variables. Note also that γ^2-family in (19) corresponds to a smoother version of the involved variables. Under the above-referred parametric scenarios, the functional entries C_{X_0,X_0} and $C_{Z_{0,3},Z_{0,3}}$ are represented in Fig. 2.

The ARBX(1) model generated lives in the function space scenario described in the "Appendix". Particularly, see Eq. (33) below.

For simplifications purposes, we have considered, in the generations, $C_{\eta_{0,i},\eta_{0,l}} = C_{\varepsilon_0,\eta_{0,i}} = \mathbf{0}_H, i, l = 1,2,3, i \neq l$. The following identities characterise the diagonal functional entries of $\overline{C}_{\boldsymbol{\eta}}$, in terms of the elements of the basis $\{\phi_j,\ j \geq 1\}$ introduced in Eq. (11): For $i = 1,2,3$, and $j, h \geq 1$,

$$\langle C_{\varepsilon_0,\varepsilon_0}(\phi_j), \phi_h\rangle_H = \begin{cases} C_{X_0X_0}(\phi_j)(\phi_j)(1 - [\rho(\phi_j)(\phi_j)]^2) & j = h, \\ e^{-|j-h|^2/W^2} & j \neq h \end{cases},$$

$$\langle C_{\eta_{0,i},\eta_{0,i}}(\phi_j), \phi_h\rangle_H = \begin{cases} C_{Z_{0,i},Z_{0,i}}(\phi_j)(\phi_j)\left(1 - [u_i(\phi_j)(\phi_j)]^2\right) & j = h \\ e^{-|j-h|^3/W^2} & j \neq h \end{cases},$$

where for $i = 1,2,3$, and $j, h \geq 1$, considering $W = 0.4$,

Table 1 Percentage of simulations from the 200 generations obtained for each sample size, where the error $\overline{B}$-norm is larger than the upper bound (23)

n_t	γ^1	γ^2
$n_1 = 1500$	11.5% ($\frac{23}{200}$)	12% ($\frac{24}{200}$)
$n_2 = 2500$	9.5% ($\frac{19}{200}$)	9% ($\frac{18}{200}$)
$n_3 = 5000$	8% ($\frac{16}{200}$)	8.5% ($\frac{17}{200}$)
$n_4 = 15{,}000$	4.5% ($\frac{9}{200}$)	4.5% ($\frac{9}{200}$)
$n_5 = 25{,}000$	3.5% ($\frac{7}{200}$)	2.5% ($\frac{5}{200}$)
$n_6 = 50{,}000$	2.5% ($\frac{5}{200}$)	1.5% ($\frac{3}{200}$)
$n_7 = 75{,}000$	2% ($\frac{4}{200}$)	1% ($\frac{2}{200}$)
$n_8 = 100{,}000$	1% ($\frac{2}{200}$)	0.5% ($\frac{1}{200}$)
$n_9 = 130{,}000$	0% ($\frac{0}{200}$)	0% ($\frac{0}{200}$)

The sample sizes $n = 1500, 2500, 5000, 15{,}000, 25{,}000, 50{,}000, 75{,}000, 100{,}000, 130{,}000$ have been tested, under the truncation rule $k_n = [\ln(n)]^-$. As indicated, the results displayed correspond to the two parametric families γ^1 and γ^2 in Eqs. (18)–(19), characterizing the RKHSs of the endogenous and exogenous functional variables

$$\rho(\phi_j)(\phi_h) = \langle \rho(\phi_j), \phi_h\rangle_H = \begin{cases} (1+j)^{-1.5} & j = h, \\ e^{-|j-h|/W} & j \neq h \end{cases},$$

$$u_i(\phi_j)(\phi_h) = \langle u_i(\phi_j), \phi_h\rangle_H = \begin{cases} (1+j)^{-(3+0.5i)} & j = h, \\ e^{-|j-h|^2/W} & j \neq h. \end{cases} \tag{20}$$

Furthermore, for $i = 1,2,3$, and $j, h \geq 1$,

$$\langle a_i(\phi_j), \phi_h\rangle_H = \begin{cases} (1+j)^{-(4+0.5i)} & j = h, \\ e^{-|j-h|^3/W} & j \neq h \end{cases}. \tag{21}$$

In the computation of the $\overline{B}$ norm, and the involved Besov and Sobolev norms, in terms of wavelets (see "Appendix", in particular, Eqs. (32)–(35)), the coarser space and highest resolution level have been fixed at $J = 2$ and $K = 6$, respectively. Specifically, in the model generated, we have considered $r = 1$, and then, $2^J \geq 2^{(\lceil r \rceil + 1)}$ is required, according to Angelini et al. (2003). Thus, $J = 2$ is fixed, to define the coarser space in the multiresolution analysis performed by wavelets (see also "Appendix"). As deeply discussed in pp. 152–156 in Antoniadis and Sapatinas (2003), the choice of the optimal K can just be reduced to being such that $K < \log_2(\sqrt{L})$, where 2^L denotes the number of grid points considered in the evaluation of the elements of the wavelet basis selected. Our choice has been Daubechies wavelets of order $N = 10$ (see Fig. 3). Then, $L = 13$ and $K = 6$ respectively define the number of grid points, and the highest resolution level studied in the next section (see also Daubechies 1992). Note that, in our case, K has been selected by a cross-validation procedure, applied within the range of allowed values above-referred.

3.1 Assumptions made and numerical results

As given in Sect. 2, the componentwise estimator (8) of $\overline{\rho}$ is strongly consistent in $\mathcal{L}(\overline{B})$, under the formulated Assumptions A1–A5, and the conditions in Ruiz-Medina and Álvarez-Liébana (2019b, Lemmas 8–9). See also Theorem 1. As commented before, the model generated in Sect. 3 ensures Assumptions A1 and A4–A5 hold. Regarding Assumption A2, in the following, for a given functional sample size n, we consider $k_n = [\ln(n)]^-$, where $[\cdot]^-$ denotes the integer part function. This truncation parameter ensures that, for all the sample sizes studied

$$n = 1500, 2500, 5000, 15{,}000, 25{,}000, 50{,}000, 75{,}000, 100{,}000, 130{,}000,$$

the empirical eigenvalue C_{n,k_n} is positive (i.e., Assumption A2 is satisfied). Figure 4 below shows that Assumption A3 also holds, under the ARBX(1) process scenario generated in the previous section.

Table 2 Percentage of simulations from the 200 generations of each sample size, where the error $\overline{B}$-norm is larger than the upper bound (23)

	n					
	γ^1			γ^2		
	5000	15,000	30,000	5000	15,000	30,000
Δh_1	12% ($\frac{24}{200}$)	10% ($\frac{20}{200}$)	7.5% ($\frac{15}{200}$)	13.5% ($\frac{27}{200}$)	10.5% ($\frac{21}{200}$)	6.5% ($\frac{13}{200}$)
Δh_2	9% ($\frac{18}{200}$)	5.5% ($\frac{11}{200}$)	4% ($\frac{8}{200}$)	9.5% ($\frac{19}{200}$)	6% ($\frac{12}{200}$)	4% ($\frac{8}{200}$)
Δh_3	7% ($\frac{14}{200}$)	4% ($\frac{8}{200}$)	3.5% ($\frac{7}{200}$)	7.5% ($\frac{15}{200}$)	4% ($\frac{8}{200}$)	3% ($\frac{6}{200}$)
Δh_4	5.5% ($\frac{11}{200}$)	3.5% ($\frac{7}{200}$)	2% ($\frac{4}{200}$)	4.5% ($\frac{9}{200}$)	2.5% ($\frac{5}{200}$)	2% ($\frac{4}{200}$)
Δh_5	2.5% ($\frac{5}{200}$)	1.5% ($\frac{3}{200}$)	1% ($\frac{2}{200}$)	1.5% ($\frac{3}{200}$)	1% ($\frac{2}{200}$)	1% ($\frac{2}{200}$)
Δh_6	1.5% ($\frac{3}{200}$)	0.5% ($\frac{1}{200}$)	0.5% ($\frac{1}{200}$)	1% ($\frac{2}{200}$)	0.5% ($\frac{1}{200}$)	0% ($\frac{0}{200}$)
Δh_7	1% ($\frac{2}{200}$)	0.5% ($\frac{1}{200}$)	0% ($\frac{0}{200}$)	0.5% ($\frac{1}{200}$)	0% ($\frac{0}{200}$)	0% ($\frac{0}{200}$)

The sample sizes $n = 5000, 15{,}000, 30{,}000$, and the truncation rule $k_n = [\ln(n)]^-$ have been studied. As indicated, the results displayed correspond to the two parametric families γ^1 and γ^2 in Eqs. (18)–(19), characterizing the involved RKHSs. The discretization steps $\{\Delta h_j = 3^{-(2+j)},\ j = 1, \ldots, 7\}$ are considered

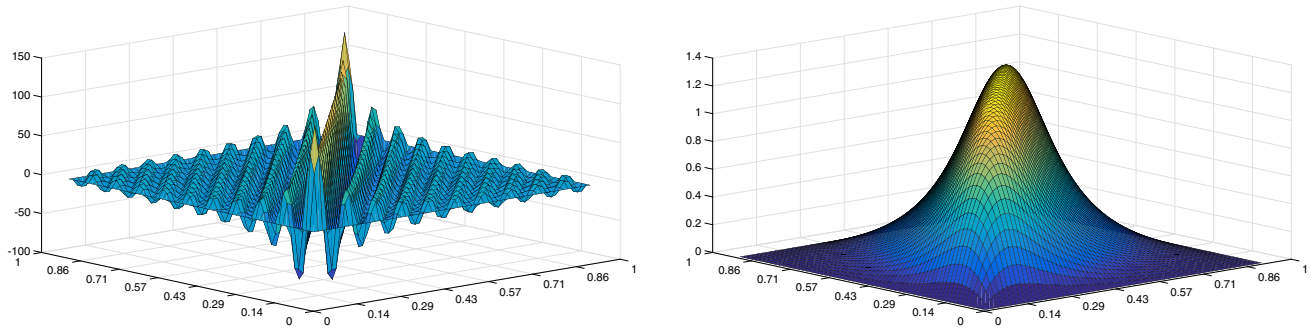

Fig. 2 Covariance kernels defining operators C_{X_0,X_0} (left-hand-side) and $C_{Z_{1,3},Z_{1,3}}$ (right-hand-side), respectively given in terms of $\gamma_1 = 2\beta + 1/10$ and $\gamma_4 = 2\beta + 4/10$, with $\beta = 3/5$, and discretization step size $\Delta h = 0.0159$

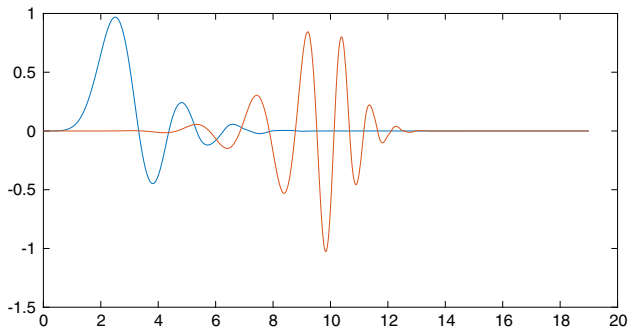

Fig. 3 Father (blue) and mother (red) Daubechies wavelets of order $N = 10$

In addition, according to Theorem 1 [see also Ruiz-Medina and Álvarez-Liébana (2019b, Theorem 1)], the strong-consistency of the componentwise estimator (8) in $\mathcal{L}(\overline{B})$ follows when

$$k_n C_{k_n}^{-1} \sum_{j=1}^{k_n} a_j = o\left(\sqrt{n/\ln(n)}\right), \quad n \to \infty, \tag{22}$$

where as before,

$$a_1 = \frac{2\sqrt{2}}{C_1 - C_2};\ a_j = 2\sqrt{2}\max\left(\frac{1}{C_{j-1} - C_j}, \frac{1}{C_j - C_{j+1}}\right), \quad j \geq 2,$$

and $\{C_j\}_{j\geq 1}$ denotes the system of eigenvalues of the extended matrix autocovariance operator $\overline{C}$ in (13). We can observe in Fig. 5 that condition (22) is satisfied.

Recall that, to ensure strong consistency in $\overline{B}-$norm, the upper bound in Theorem 1 [see also Ruiz-Medina and Álvarez-Liébana (2019b, Theorem 1 and Corollary 1)], for the functional prediction error $\left\|\overline{\rho}(\overline{X}_n) - \widehat{\overline{X}}_n\right\|_{\overline{B}}$ is given by

$$\mathcal{M}\xi_n = \mathcal{M}\exp\left(\frac{-n}{C_{k_n}^{-2}k_n^2\left(\sum_{j=1}^{k_n} a_j\right)^2}\right), \tag{23}$$

where $\mathcal{M} = \sup_{\omega\in\Omega\setminus\Omega_0}\|\overline{X}_0(\omega)\|_{\overline{B}}$, $\mathcal{P}(\Omega_0) = 0$, and for each functional sample size n, as before, C_{k_n} is the k_n-th eigenvalue of the extended autocovariance operator $\overline{C}$ in (13), with $k_n = [\ln(n)]^-$. Table 1 reflects the percentage of values of the computed $\overline{B}$-norm of the functional error greater than the upper bound (23).

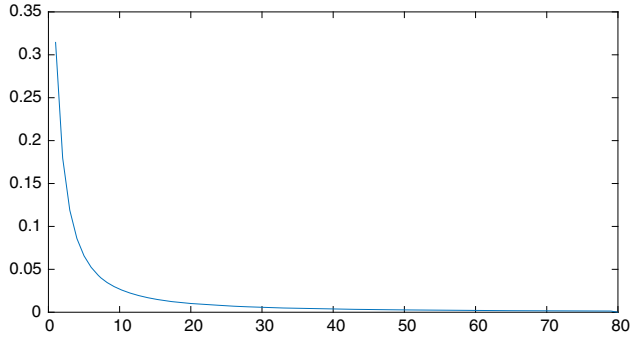

Fig. 4 Values of

$$\left\| \overline{\rho}(\cdot) - \sum_{l=1}^{K} \langle \overline{\rho}(\cdot), \overline{\phi}_l \rangle_{[\mathcal{B}_{2,2}^{-\beta}([0,1])]^{b+1}} \overline{\phi}_l \right\|_{\mathcal{L}\left([\mathcal{B}_{\infty,\infty}([0,1])]^{b+1}\right)}$$

, for K-values reflected in the horizontal axis

As commented, the numerical results in Table 1 focus on the illustration of the strong consistency of the derived ARBX(1) plug-in predictor. While the effect of the discretization step size is analyzed in Table 2. Particularly, the following discretization step scheme has been considered:

$$\left\{ \Delta h_j = \frac{1}{3^{2+j}},\ j = 1, \ldots, 7 \right\},$$

and the particular values

$$\Delta h_1 = 3.70(10^{-2}), \quad \Delta h_2 = 1.23(10^{-2}),$$
$$\Delta h_3 = 4.12(10^{-3}), \quad \Delta h_4 = 1.37(10^{-3}),$$
$$\Delta h_5 = 4.57(10^{-4}), \quad \Delta h_6 = 1.52(10^{-4}),$$
$$\Delta h_7 = 5.08(10^{-5}),$$

have been studied (see Fig. 6 below), corresponding to $\left\{3^{2+j}+1,\ j=1,\ldots,7\right\}$ grid points.

In Table 2, the sample sizes $n = 5000, 15{,}000, 30{,}000$ are studied. This table displays the percentage of simulations where the error B-norm is larger than the upper bound in (23), for each discretization step size, functional sample size and parametric families γ^1 and γ^2.

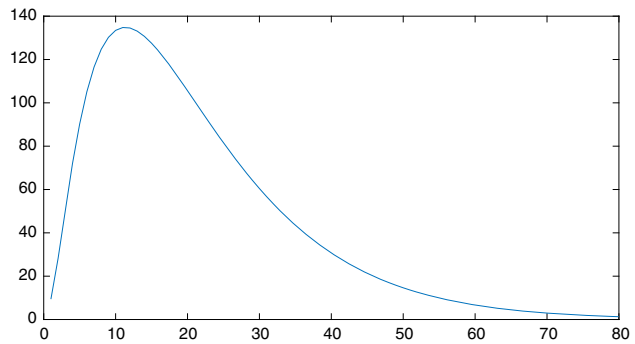

Fig. 5 Values of $\frac{k_n C_{k_n}^{-1} \sum_{j=1}^{k_n} a_j}{\sqrt{n/\ln(n)}}$, considering $k_n = [\ln(n)]^-$, as reflected in the horizontal axis

4 Real-data application: short-term forecasting of air pollutants

In this section, the performance of the ARBX(1) based prediction approach presented is illustrated in a real-data example. Specifically, the short-term forecasting of daily average concentrations of atmospheric aerosol particles with diameters less than 10 μm, also known as PM_{10} (coarse particles), is achieved from a functional perspective. The importance of the accurate forecasting of this kind of particles relies on being inhalable atmospheric pollution particles, which impact the public health. Following the suggestions by the World Health Organization, the European Union developed in 2008 (in particular, directive 2008/50/EU) a complete legislative package, establishing health based standards for the levels of PM_{10}: daily mean concentration of PM_{10} should not be greater than 50 μg m^{-3} more than 35 days per year, neither the annual average of concentration of PM_{10} shall not be greater than 40 μg m^{-3}. However, this limit has been exceed during the last years in heavily industrialized areas, deriving in severe people's health problems. Therefore, PM_{10} forecasting is crucial to adopting efficient public transport policies. The dataset is analyzed in Sect. 4.1, while Sect. 4.2 describes the previous processing procedure required, for the implementation of our functional prediction methodology in Sect. 4.3.

4.1 Description of our dataset

The dataset considered is comprised of daily average concentrations of PM_{10}, coming from hourly measurements, from January 1, 2007 to March 31, 2011, collected by the air quality Normand (French) authority, known as Air Normand (see http://www.atmonormandie.fr). Specifically, we have been able to access the records, collected at 6 of the 13 pollution monitoring stations, in the fixed network located throughout Haute-Normandie region, which is one of the most heavily industrialized areas in France (see locations in the maps displayed in Fig. 7).

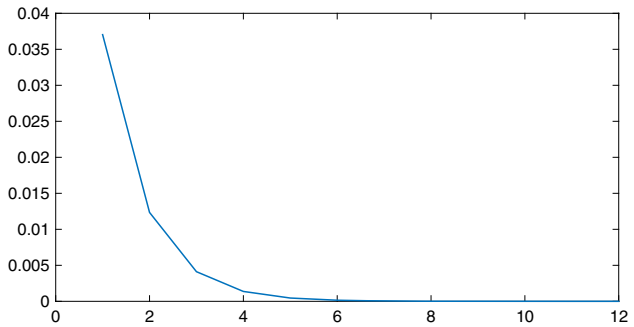

Fig. 6 Zero-convergence-rate of $\Delta h_j = \frac{1}{3^{2+j}}, j = 1, \ldots, 12$

In the following, these monitoring stations will be denoted as $\{S_c,\ c = 1, \ldots, 6\}$, which have been chosen aimed at reflecting a wide variety of scenarios, such that roadsides, urban, industrial and rural regions. As reflected in Fig. 8, the monitoring station S_6 (located in a rural area) has collected the smallest PM_{10} concentrations, while stations S_2 and S_5 (located at roadsides) collected the highest pollution levels. These stations also display the highest variability, which seems logical, since pollution levels in roadside are strongly dependent on the road traffic.

The ARBX(1) modelling is adopted since pollution particles are mainly procured from natural sources (i.e., influenced by meteorological variables), or due to human activity. In our study we incorporate the exogenous information coming from meteorological variables. Specifically, we consider the following four exogenous variables ($b = 4$ in our ARBX(1) model): Daily average temperature (°C), daily average atmospheric pressure (hPa), daily average wind speed (ms^{-1}), and daily maximum gradient of temperature (°C), computed all of them from hourly measurements. Note that daily maximum gradient of temperature denotes the daily maximum of the hourly differences between the temperature at 2 and 100 m. Measurements of these meteorological variables were collected at three meteorological stations belonging to the French national meteorological service, such that each air pollution station is associated with the closest meteorological station. Thus, pollution stations S_1, S_2 and S_3 are associated to a common meteorological station. Also, a second meteorological station covers the pollution stations S_4 and S_5. Finally, a third meteorological station is associated with the pollution station S_6. From the records collected at the three meteorological stations, covering pollution stations by spatial proximity, the sample behaviour of the four exogenous variables affecting PM_{10} concentrations at the six pollution stations, is displayed in the boxplots in Fig. 9.

4.2 Data preprocessing

To obtain our functional dataset the following steps are implemented:

Step 1: Missing-data imputation In the caption of Fig. 8, the percentage of missing data is provided. Note that the R package used (see, e.g., Moritz and Bartz-Beielstein 2017) performs the imputation, based on an average of the previous and posterior non missing values. Since pollution data is strongly linked with routines and consumption patterns in business days, the past and the next five values are considered. At each station, $S_c, c = 1, \ldots, 6$, 1551 records are then available corresponding to daily measurements from January 1, 2007 to March 31, 2011. Daily observations of the endogenous variable at pollution station S_1 are displayed after imputation in Fig. 10 below.

Step 2: Balanced data (31-days per month). A common support is required for all the observed curves, evaluated in the same function space. *Cubicspline* option in `fit.m` `MatLab` function is applied, to obtaining 31 measurements per month. Thus, the final grid contains $31 \times 51 = 1581$ points.

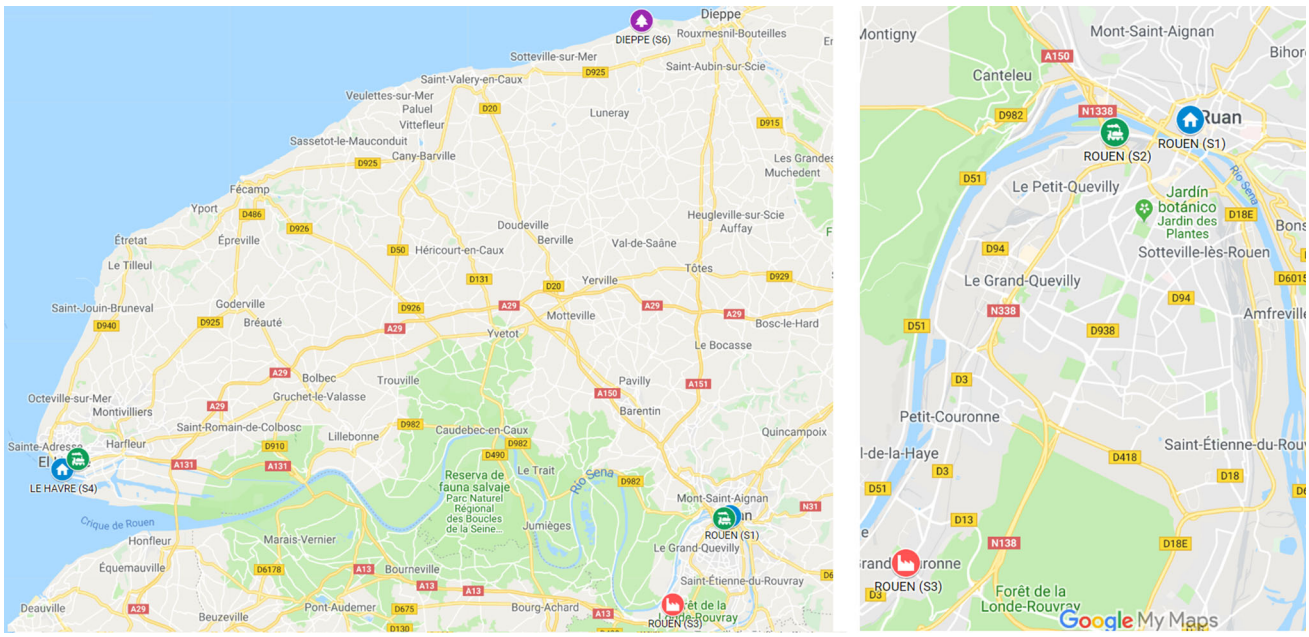

Fig. 7 At the left-hand-side: map (extracted from Google Maps) displaying the location of the six pollution monitoring stations analyzed. At the right-hand-side: enlarged map displaying the stations near Rouen

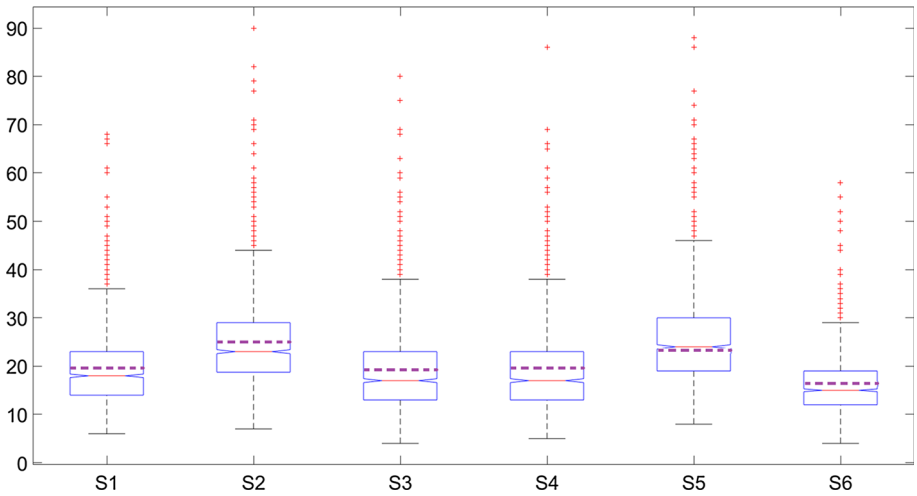

Fig. 8 Boxplots on the observed sample behaviour of PM_{10} concentrations (μg m^{-3}). Purple dotted lines reflect the average concentrations, while red solid lines splitting the box reflect the medians. Percentage of missing observations: S_1 Rouen (Urban) 1%; S_2 Rouen (Roadside) 0.4%; S_3 Rouen (Industrial) 1.4%; S_4 Le Havre (Urban) 3 %; S_5 Le Havre (Roadside) 1.2 %; S_6 Dieppe (Rural) 1.7 %

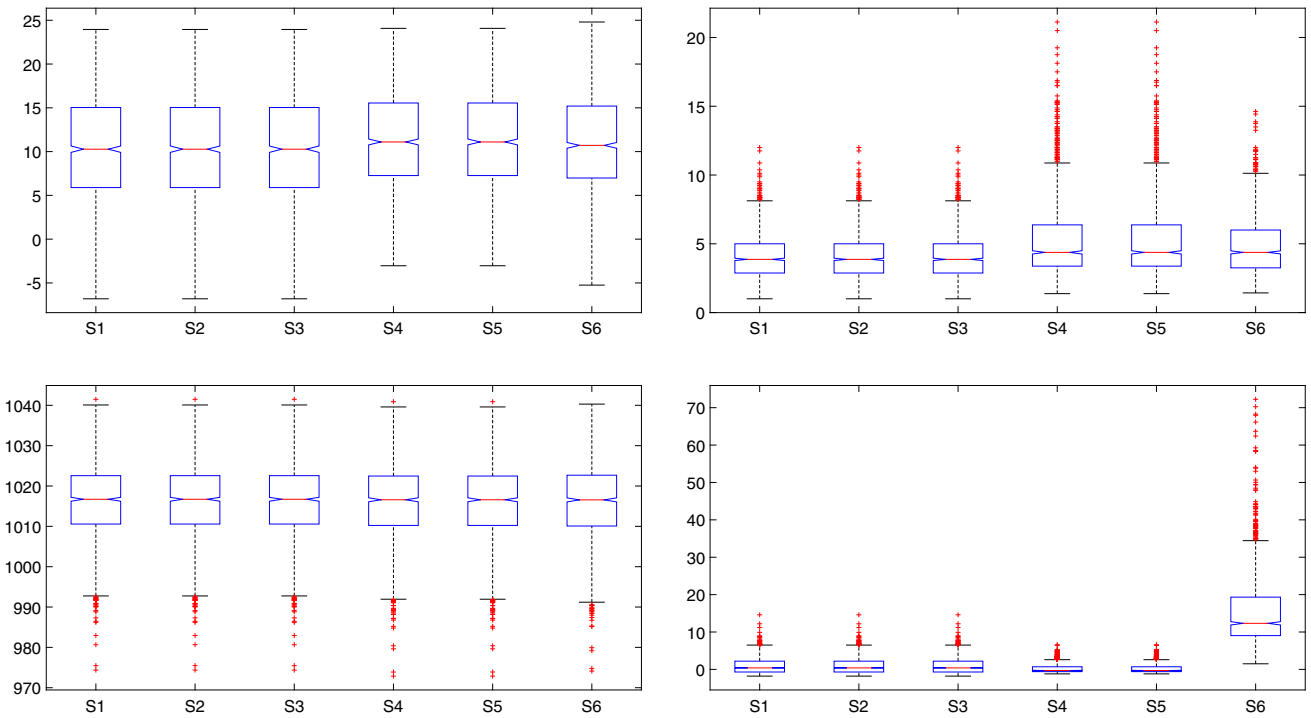

Fig. 9 Boxplots displaying the sample behaviour of exogenous variables at each meteorological station (mean temperature on the top left, mean wind speed on the top right, mean air pressure at the bottom left, and maximum temperature gradient at the bottom right), as reflected in this figure, respectively associated with the pollution stations (S_1, S_2, S_3), (S_4, S_5) and (S_6)

Step 3: Splitting our dataset At each pollution station $S_c, c = 1, \ldots, 6$, we will construct our plug-in predictor $\widehat{X}_{51}^c = \widetilde{\overline{\rho}}_{k_n}(\overline{X}_{50}^c)$, from a functional sample $\{\overline{X}_1, \ldots, \overline{X}_{50}\}$ of size 50. Thus, functional prediction is achieved for the last month, March, 2011, at each pollution station.

Step 4: Detrending and deseasonalizing A polynomial $a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \cdots$ is fitted for detrending all the curve data, including the last month. Thus, we have checked the trends fitting polynomials of increasing degree, stopping when the fitting trends between two successive degrees display a similar behaviour. As shown in Fig. 11 below, a quadratic polynomial trend $a_0 + a_1 t + a_2 t^2$ has finally been fitted. After detrending, annual seasonality is also removed

Step 5: Modelling by an ARBX(1) process Summarizing, from the previous steps, our functional sample is constituted by 50 detrended, and annually deseasonalized observed curves for the endogenous and exogenous variables (on a grid of 1581 points). Plug-in functional prediction is achieved for the 51-th month, from the observed curves at the previous 50 months by fitting ARBX(1)

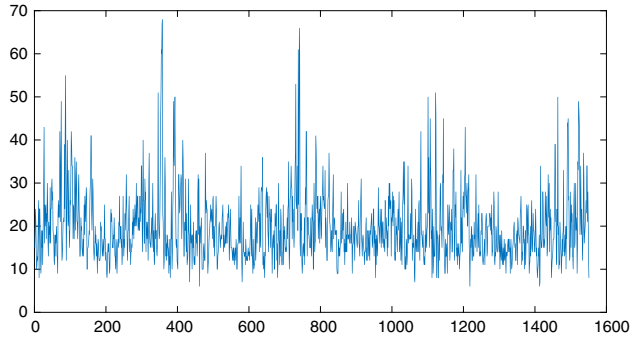

Fig. 10 Daily observations (from 01/01/2007 to 31/03/2011, 1551 grid points) of PM_{10} at station S_1 after imputation

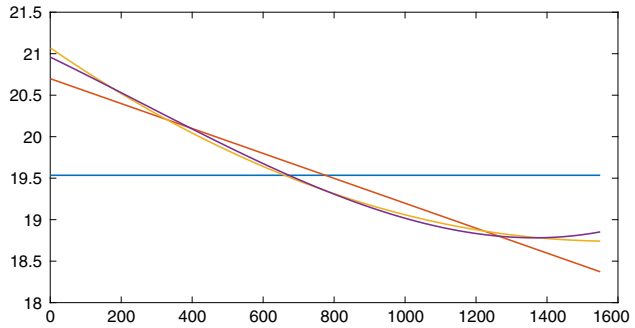

Fig. 11 Fitting trends as polynomial of degrees 0 (null trend; blue line), 1 (linear trend; red line), 2 (quadratic trend; yellow line) and 3 (cubic trend; purple line)

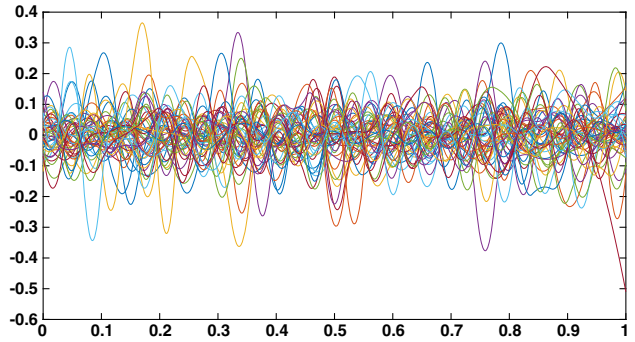

Fig. 12 Observed PM_{10} curves at station S_1 for the period January 2007–February 2011

model in Eqs. (24)–(25) below. Figure 12 displays our PM_{10} functional dataset (after interpolation and smoothing) at station S_1, corresponding to the period January 2007–February 2011.

Specifically, the following ARBX(1) model is fitted:

$$\begin{aligned} &\overline{X}_n^c = \overline{\rho}^c\left(\overline{X}_{n-1}^c\right) + \overline{\varepsilon}_n^c, \quad n = 1, \ldots, T = 50, \\ &c = 1, \ldots, 6, \end{aligned} \tag{24}$$

where $b = 4$, and hence,

$$\overline{X}_n^c = \begin{Bmatrix} X_n^c \\ Z_{n+1,1}^c \\ Z_{n+1,2}^c \\ Z_{n+1,3}^c \\ Z_{n+1,4}^c \end{Bmatrix}, \quad \overline{\varepsilon}_n^c = \begin{Bmatrix} \varepsilon_n^c \\ \eta_{n,1}^c \\ \eta_{n,2}^c \\ \eta_{n,3}^c \\ \eta_{n,4}^c \end{Bmatrix},$$

$$\overline{\rho}^c = \begin{Bmatrix} \rho^c & a_1^c & a_2^c & a_3^c & a_4^c \\ \mathbf{0}_B & u_1^c & \mathbf{0}_B & \mathbf{0}_B & \mathbf{0}_B \\ \mathbf{0}_B & \mathbf{0}_B & u_2^c & \mathbf{0}_B & \mathbf{0}_B \\ \mathbf{0}_B & \mathbf{0}_B & \mathbf{0}_B & u_3^c & \mathbf{0}_B \\ \mathbf{0}_B & \mathbf{0}_B & \mathbf{0}_B & \mathbf{0}_B & u_4^c \end{Bmatrix}, \tag{25}$$

$c = 1, \ldots, 6.$

Equivalently, for $c = 1, \ldots, 6$,

$$X_n^c = \rho^c\left(X_{n-1}^c\right) + \sum_{i=1}^{4} a_i^c\left(Z_{n,i}^c\right) + \varepsilon_n^c, \quad n = 1, \ldots, T = 50, \tag{26}$$

with, for $i = 1, \ldots, 4$, $a_i^c \in \mathcal{L}(B)$, $\rho^c \in \mathcal{L}(B)$, and

$$Z_{n,i}^c = u_i^c\left(Z_{n-1,i}^c\right) + \eta_{n,i}^c, \quad u_i^c \in \mathcal{L}(B), \quad n = 1, \ldots, T = 50. \tag{27}$$

It can be observed in Fig. 12 that PM_{10} curves are continuous. Hence, $\beta > 1/2$. In particular, we look for the minimal local regularity order, in our fitting of the previous introduced ARBX(1) model, with $b = 4$. Thus, the parameter value $\beta = 3/5$ is considered, and $\gamma = \gamma_i = 2\beta + \epsilon$, with $\epsilon = 0.01$, for $i = 1, 2, 3, 4$, being fitted, to define the corresponding RKHSs structures.

$$\begin{aligned} &\overline{B} = \left[\mathcal{B}_{\infty,\infty}^0([0,1])\right]^5, \\ &\widetilde{\overline{H}} = \left[\mathcal{B}_{2,2}^{-\beta}([0,1])\right]^5, \ \overline{H} = [L^2([0,1])]^5, \ \mathcal{H}(\overline{X}) = [\mathcal{B}_{2,2}^{\gamma}([0,1])]^5. \end{aligned} \tag{28}$$

4.3 The performance of the ARBX(1) plug-in predictor

In the implementation of the leave-one-out cross validation procedure, at each pollution station $S_c, c = 1, \ldots, 6$, the following functional sample is considered:

$$\overline{Y}^{h,c} = \left\{\overline{Y}_i^{h,c}, \quad i = 1, \ldots, 48\right\} = \left\{\overline{X}_i^c, \quad i = 1, \ldots, 49\right\} \setminus \left\{\overline{X}_h^c\right\},$$

by leaving aside the functional data $\overline{X}_h^c$, $h = 1, \ldots, 49$, as well as the functional data $\overline{X}_{50}^c, c = 1, \ldots, 6$, in the computation of the componentwise estimator (8) of the autocorrelation operator $\overline{\rho}$. Thus, at each iteration $h \in \{1, \ldots, 49\}$ of the implemented leave-one-out cross validation procedure, the ARBX(1) plug-in predictor

$[\widehat{\overline{X}}^c_{51}]_{h,k_n}$ of $\overline{X}^c_{51}$ is computed, from a functional sample of size 49, considering the truncation order k_n, as follows

$$[\widehat{\overline{X}}^c_{51}]_{h,k_n} = \sum_{j,l=1}^{k_n} \frac{1}{\widetilde{C}^{c,h}_{n,j}} \langle \overline{X}^{c,h}_{50}, \widetilde{\phi}^{c,h}_{n,j} \rangle_{\widetilde{\overline{H}}} \times \left(\frac{1}{47} \sum_{i=1}^{47} \langle \overline{Y}^{c,h}_{i+1}, \widetilde{\phi}^{c,h}_{n,j} \rangle_{\widetilde{\overline{H}}} \langle \overline{Y}^{c,h}_{i}, \widetilde{\phi}^{c,h}_{n,l} \rangle_{\widetilde{\overline{H}}} \widetilde{\phi}^{c,h}_{n,l} \right), \tag{29}$$

for $c = 1, \ldots, 6$. Here, the empirical eigenvalues

$$\left\{ \widetilde{C}^{c,h}_{n,j},\ j = 1, \ldots, 48,\ h = 1, \ldots, 49,\ c = 1, \ldots, 6 \right\}$$

are calculated by the formula

$$\widetilde{C}^{c,h}_{n,j} = \frac{1}{48} \sum_{i=1}^{48} \langle \overline{Y}^{c,h}_{i}, \widetilde{\phi}^{c,h}_{n,j} \rangle^2_{\widetilde{\overline{H}}} = \frac{1}{48} \sum_{\substack{i=1 \\ i \neq h}}^{49} \langle \overline{X}^{c}_{i}, \widetilde{\phi}^{c,h}_{n,j} \rangle^2_{\widetilde{\overline{H}}}.$$

In the above equations, for each $c = 1, \ldots, 6$, and $h = 1, \ldots, 49$, $\left\{ \widetilde{\phi}^{c,h}_{n,j},\ j \geq 1 \right\}$ denotes the system of eigenvectors of the extended empirical autocovariance operator, based on a functional sample of size 48. The truncation parameter values $k_{n,1} = [\log_2(\sqrt{n})]^-$ and $k_{n,2} = \left[\ln\left(n^{5/2}\right)\right]^-$ are tested, in the computation of (29). Similarly to the simulation study undertaken, all the required conditions for the strong-consistency are checked. In particular, Assumptions A4–A5 directly follow from the function space scenario (28) adopted. Figure 13 displays the convergence to zero required in Assumption A3 at pollution stations S_1 and S_6.

Table 3 below and Fig. 14 display the mean leave-one out cross validation functional errors

$$E^{k_{n,m}}_c = \frac{1}{49} \sum_{h=1}^{49} \left\| \overline{X}^c_{51} - \left[\widehat{\overline{X}}^c_{51} \right]_{h,k_{n,m}} \right\|_{\overline{B}}, \quad c = 1, \ldots, 6, \quad m = 1, 2, \tag{30}$$

at the six pollution stations studied, for the two truncation orders analyzed. In the calculation of (30), the Besov and Sobolev norms involved in our function space scenario (28) are computed by projection into Daubechies wavelets of order $N = 10$ (see Daubechies 1992), with six resolution levels (see also Fig. 3).

In Table 3, the worst performance is observed at pollution stations S_2 and S_5, corresponding to roadside stations. As commented before, the traffic flow is one of the main factors inducing the higher-variability displayed by PM_{10} concentrations in these stations. A slightly better performance can be observed with truncation order $k_{n,1}$, but, indeed, a significant improvement cannot be concluded. When larger values of parameter β, and hence, of parameter γ, defining our function space scenario, are considered, a stronger smoothing of our original data set is achieved, in terms of Sobolev and Besov norms. Thus, a better performance is obtained, with a loss of accuracy, in the approximation of the local behaviour of PM_{10} concentrations.

5 Final comments

It is well-known that FDA techniques provide a flexible framework for the local analysis of high-dimensional data which are continuous in nature. One of the main subjects in FDA is the suitable choice of the function space, where the observed data take their values. In particular, the norm of the selected space should provide an accurate measure of the local variability of the observed endogenous and exogenous variables, that could be crucial in the posterior representation of the possible interactions with the phenomena of interest and its evolution. That is the case of the real-data example analyzed in Sect. 4

This paper adopts an abstract Banach space framework, assuming an autoregressive dynamics in time, for all the

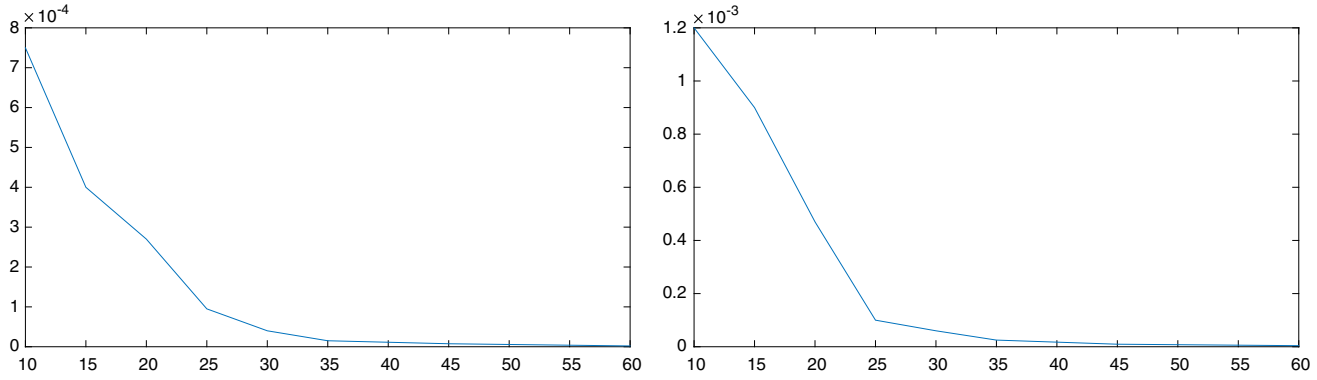

Fig. 13 Evaluation at stations S_1 (left-hand-side) and S_6 (right-hand-side) of the empirical norm $L_k = \sup_{\overline{x} \in \overline{B},\ \|\overline{x}\|_{\overline{B}} \leq 1} \left\| \overline{\rho}(\overline{x}) - \sum_{j=1}^{k} \langle \overline{\rho}(\overline{x}), \overline{\phi}_{n,j} \rangle_{\overline{H}} \overline{\phi}_{n,j} \right\|_{\overline{B}}$, for values $k = 10, 15, 20, 25, 30, 35, 45, 60$, displayed in the horizontal axis

Table 3 Mean leave-one out cross validation functional errors (30) at pollution stations $\{S_c,\ c = 1, \ldots, 6\}$

$k_{n,1}$		$k_{n,2}$	
Station S_c	$E_c^{k_{n,1}}$	Station S_c	$E_c^{k_{n,2}}$
S_1	0.00150	S_1	0.00193
S_2	0.00772	S_2	0.00940
S_3	0.00186	S_3	0.00202
S_4	0.00167	S_4	0.00182
S_5	0.00990	S_5	0.01020
S_6	0.00091	S_6	0.00127

The truncation parameters $k_{n,1} = [\log_2(\sqrt{n})]^-$ and $k_{n,2} = \left[\ln\left(n^{5/2}\right)\right]^-$ have been tested

functional random variables involved in the model. Specifically, an ARBX(1) model is considered. The endogenous and exogenous information affecting the functional response at a given time is incorporated through a suitable linear model, involving a matrix autocorrelation operator. This operator model reflects possible interactions between all endogenous and exogenous functional random variables at any time.

Particularly, the scale of fractional Besov spaces provides a suitable functional framework, where the presented approach can be implemented, modelling local regularity/singularity in an accurate way. Note that the norms in these spaces can be characterized in terms of the wavelet transform. Specifically, wavelet bases provide countable dense systems in Besov spaces, that can be used in the definition of the inner product and associated norms in weighted fractional Sobolev spaces, constructed from the space of square integrable functions on an interval (see Triebel 1983; Ruiz-Medina and Álvarez-Liébana 2019b). Thus, suitable embeddings can be established for applying the construction in Lemma 2.1 in Kuelbs (1970). As special cases of well-known Banach spaces within our framework, we refer to $\mathcal{C}([0,1])$ the space of continuous functions on [0, 1], with the supremum norm, and $\mathcal{D}([0,1])$ the Skorokhod space of right-continuous functions on the interval [0, 1], having a left limit at all $t \in [0,1]$. Note that these spaces have been widely used in the FDA literature in a Banach-valued time series context (see Bosq 2000).

The simulation study and real-data application illustrate the fact that our approach is sufficiently flexible to describing the local behaviour of both, regular and singular functional data. Note that, in the singular case, we can choose a suitable norm that measures the local fluctuations in a precise way. This information is relevant, for example, in the analysis of PM_{10} concentrations, as was illustrated in Sect. 4.

Fig. 14 Map displaying mean leave-one out cross validation functional errors (30) at pollution stations $\{S_c, c = 1, \ldots, 6\}$. The truncation parameters $k_{n,1} = [\log_2(\sqrt{n})]^-$ (left-hand-side) and $k_{n,2} = \left[\ln\left(n^{5/2}\right)\right]^-$ (right-hand-side) have been tested

An individual statistical analysis has been performed in Sect. 4 at each pollution station. The incorporation of spatial interactions in the analysis could be addressed in a multivariate infinite-dimensional spatial framework, and constitutes the subject of a subsequent paper. Specifically, most of the authors consider, for example, a joint analysis of wind speed and wind direction as exogenous variables. In that case, a spatial functional correlation analysis should be considered. There are several available methods in the current spatial functional statistical literature (see Delicado et al. 2010) to address this issue. Particularly, we refer to the frameworks of multivariate functional random field based prediction (see, e.g., Bohorquez et al. 2017); and, spatial functional kriging-based techniques (see, e.g., Bohorquez et al. 2017; Delicado et al. 2010; Giraldo et al. 2010; Ignaccolo et al. 2014; Nerini et al. 2010, among others).

The state space based approach is relatively new in the spatial functional framework (see, e.g., Ruiz-Medina 2011, where spatial autoregressive series in Hilbert spaces are introduced). The invertibility of the linear state space equation leads to important technical improvements, in the derivation of the asymptotic properties of componentwise functional parameter estimators, and associated plug-in predictors. Particularly, to ensure strong-consistency, the truncation rule depends on the sample size, the separation, and the rate of convergence to zero of the eigenvalues of the autocovariance operator.

Acknowledgements This work was supported in part by projects MTM2015–71839–P and PGC2018-099549-B-I00 (co-funded by Feder funds), of the DGI, MINECO, Spain.

Appendix

It is well-known that Besov spaces, $\left\{\left(\mathcal{B}^r_{p,q}, \|\cdot\|^r_{p,q}\right), r \in \mathbb{R}, 1 \leq p, q \leq \infty\right\}$, and their norms can be characterized in terms of the wavelet transform (see, e.g., Triebel 1983). Specifically, for every $f \in \mathcal{B}^r_{p,q}$,

$$\|f\|^r_{p,q} \equiv \|\varphi_J * f\|_p + \left[\sum_{j=J}^{\infty}\left(2^{jr}\|\psi_j * f\|_p\right)^q\right]^{1/q} < \infty, \quad (31)$$

where φ and ψ denote the father and mother wavelets, whose translations and dilations provide a multiresolution analysis of a suitable space of square-integrable functions. Particularly, consider the space $L^2([0,1])$, and its orthogonal decomposition from an $(\lceil r \rceil + 1)$-regular Multiresolution Analysis, induced by an orthogonal basis of wavelets, for certain $r > 0$. Then, father and mother wavelets belong to $\mathcal{C}^{(\lceil r \rceil+1)}([0,1])$. For every $f \in L^2([0,1])$,

$$f(t) = \sum_{k=0}^{2^J-1} \alpha^f_{J,k}\varphi_{J,k}(t) + \sum_{j=J}^{K}\sum_{k=0}^{2^j-1} \beta^f_{j,k}\psi_{j,k}(t), \quad t \in [0,1], \quad (32)$$

where J is such that $2^J \geq 2^{(\lceil r \rceil+1)}$, and for $k = 0, \ldots, 2^{j-1}, j = J, \ldots, K$,

$$\alpha^f_{J,k} = \int_{\mathbb{R}} f(x)\overline{\varphi_{J,k}(x)}dx, \quad \beta^f_{j,k} = \int_{\mathbb{R}} f(x)\overline{\psi_{j,k}(x)}dx$$

(see Daubechies 1992). Here, K is the truncation parameter defining the last (or highest) resolution level considered in the finite-dimensional wavelet approximation (32).

As commented before, the following function spaces have been considered:

$$\begin{aligned}
\overline{B} &= \left[\mathcal{B}^0_{\infty,\infty}([0,1])\right]^{b+1};\\
\widetilde{\overline{H}} &= \left[H_2^{-\beta}([0,1])\right]^{b+1} = \left[\mathcal{B}^{-\beta}_{2,2}([0,1])\right]^{b+1}\\
\overline{H} &= \left[L^2([0,1])\right]^{b+1};\\
\mathcal{H}(\overline{X}) &= \prod_{i=1}^{b+1} H_2^{\gamma_i}([0,1]) = \prod_{i=1}^{b+1} \mathcal{B}^{\gamma_i}_{2,2}([0,1])\\
\overline{B}^{\star} &= \left[\mathcal{B}^0_{1,1}([0,1])\right]^{b+1};\\
\widetilde{\overline{H}}^{\star} &= \left[H_2^{\beta}([0,1])\right]^{b+1} = \left[\mathcal{B}^{\beta}_{2,2}([0,1])\right]^{b+1}\\
\overline{H}^{\star} &= \left[L^2([0,1])\right]^{b+1};\\
[\mathcal{H}(\overline{X})]^{\star} &= \prod_{i=1}^{b+1} H_2^{-\gamma_i}([0,1]) = \prod_{i=1}^{b+1} \mathcal{B}^{-\gamma_i}_{2,2}([0,1]),
\end{aligned} \quad (33)$$

where the parameters $\{\gamma_i\}_{i=1,\ldots,b+1}$ reflect the second-order local regularity of the functional random components of $\overline{X} = \{\overline{X}_n,\ n \in \mathbb{Z}\}$ in Eq. (5). From embedding theorems between Besov spaces, the following continuous inclusions hold (see Triebel 1983):

$$\mathcal{H}(\overline{X}) \hookrightarrow \widetilde{\overline{H}}^{*} \hookrightarrow \overline{B}^{*} \hookrightarrow \overline{H} \hookrightarrow \overline{B} \hookrightarrow \widetilde{\overline{H}} \hookrightarrow [\mathcal{H}(\overline{X})]^{\star}, \quad (34)$$

for $\gamma_i > 2\beta > 1, i = 1, \ldots, b+1$. Thus, Assumptions A4–A5 are satisfied. The $\overline{B}$ and $\overline{B}^{\star}$ norms are then computed from the following identities: For every $\overline{f} = (f; f_1, \ldots, f_b)$, $\overline{g} = (g; g_1, \ldots, g_b) \in \overline{B} \subset \widetilde{\overline{H}}$,

$$\begin{aligned}\|\overline{f}\|_{\overline{B}} &= \sup_{j\geq J}\sup_{k=0,\ldots,2^j-1}\sup\left(\left|\alpha^f_{J,k}\right|,\left|\beta^f_{j,k}\right|,\sup_{i=1,\ldots,b}\left|\alpha^{f_i}_{J,k}\right|,\right.\\ &\quad\left.\sup_{i=1,\ldots,b}\left|\beta^{f_i}_{j,k}\right|\right)\\ \|\overline{g}\|_{\overline{B}^*} &= \left[\sum_{k=0}^{2^J-1}\left|\alpha^g_{J,k}\right|+\sum_{j=J}^{K}\sum_{k=0}^{2^j-1}\left|\beta^g_{j,k}\right|\right]\\ &\quad+\left[\sum_{k=0}^{2^J-1}\sum_{i=1}^{b}\left|\alpha^{g_i}_{J,k}\right|+\sum_{j=J}^{K}\sum_{k=0}^{2^j-1}\sum_{i=1}^{b}\left|\beta^{g_i}_{j,k}\right|\right],\end{aligned} \tag{35}$$

where for $f \in B$, and $g \in B^\star$,

$$\begin{aligned}\|f\|_B &= \sup\left\{\left|\alpha^f_{J,k}\right|, k=0,\ldots,2^J-1;\left|\beta^f_{j,k}\right|,\right.\\ &\quad\left. k=0,\ldots,2^j-1, j=J,\ldots,K\right\},\\ \|g\|_{B^*} &= \sum_{k=0}^{2^J-1}\left|\alpha^g_{J,k}\right|+\sum_{j=J}^{K}\sum_{k=0}^{2^j-1}\left|\beta^g_{j,k}\right|.\end{aligned}$$

References

Álvarez-Liébana J, Bosq D, Ruiz-Medina MD (2016) Consistency of the plug-in functional predictor of the Ornstein–Uhlenbeck in Hilbert and Banach spaces. Stat Probab Lett 117:12–22

Álvarez-Liébana J, Bosq D, Ruiz-Medina MD (2017) Asymptotic properties of a componentwise ARH(1) plug-in predictor. J Multivar Anal 155:12–34

Aneiros-Pérez G, Vieu P (2008) Nonparametric time series prediction: a semi-functional partial linear modeling. J Multivar Anal 99:834–857

Angelini C, De Candittis D, Leblanc F (2003) Wavelet regression estimation in nonparametric mixed effect models. J Multivar Anal 85:267–291

Antoniadis A, Sapatinas T (2003) Wavelet methods for continuous-time prediction using Hilbert-valued autoregressive processes. J Multivar Anal 87:133–158

Besse PC, Cardot H, Stephenson DB (2000) Autoregressive forecasting of some functional climatic variations. Scand J Stat 27:673–687

Blanke D, Bosq D (2016) Detecting and estimating intensity of jumps for discretely observed ARMAD(1,1) processes. J Multivar Anal 146:119–137

Bohorquez M, Giraldo R, Mateu J (2017) Multivariate functional random fields: prediction and optimal sampling. Stoch Environ Res Risk Assess 31:53–70

Bosq D (2000) Linear processes in function spaces. Springer, New York

Bueno-Larraz B, Klepsch J (2018) Variable selection for the prediction of $C[0,1]$-valued AR processes using RKHS. arXiv: 1710.06660

Damon J, Guillas S (2002) The inclusion of exogenous variables in functional autoregressive ozone forecasting. Environmetrics 13:759–774

Damon J, Guillas S (2005) Estimation and simulation of autoregressie Hilbertian processes with exogenous variables. Stat Inference Stoch Process 8:185–204

Daubechies I (1992) Ten lectures on wavelets, vol 61. CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia

Dehling H, Sharipov OS (2005) Estimation of mean and covariance operator for Banach space valued autoregressive processes with dependent innovations. Stat Inference Stoch Process 8:137–149

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

El Hajj L (2011) Limit theorems for $\mathcal{D}([0,1])$-valued autoregressive processes. C R Acad Sci Paris Sér I Math 349:821–825

Febrero-Bande M, Galeano P, González-Manteiga W (2008) Outlier detection in functional data by depth measures with application to identify abnormal NO_x levels. Environmetrics 19:331–345

Fernández de Castro BM, González-Manteiga W, Guillas S (2005) Functional samples and bootstrap for predicting sulfur dioxide levels. Technometrics 47:212–222

Ferraty F, Vieu P (2006) Nonparametric functional data analysis: theory and practice. Springer, New York

Geenens G (2011) Curse of dimensionality and related issues in nonparametric functional regression. Stat Surv 5:30–43

Giraldo R, Delicado P, Mateu J (2010) Geostatistics for functional data: an ordinary kriging approach. Environ Ecol Stat 18:411–426

Gocheva-Ilieva S, Ivanov A, Voynikova D, Boyadzhiev D (2014) Time series analysis and forecasting for air pollution in small urban area: an SARIMA and factor analysis approach. Stoch Environ Res Risk Assess 28:1045–1060

Goia A, Vieu P (2015) A partitioned single functional index model. Comput Stat 30:673–692

Goia A, Vieu P (2016) An introduction to recent advances in high/infinite dimensional statistics. J Multivar Anal 146:1–6

Grivas G, Chaloulakou A (2006) Artificial neural network models for prediction of PM_{10} hourly concentrations in the greater area of Athens, Greece. Atmos Environ 40:1216–1229

Guillas S (2002) Doubly stochastic Hilbertian processes. J Appl Probab 39:566–580

He H-D, Lu W-Z, Xue Y (2015) Prediction of particulate matters at urban intersection by using multilayer perceptron model based on principal components. Stoch Environ Res Risk Assess 29:2107–2114

Horváth L, Kokoszka P (2012) Inference for functional data with applications. Springer, New York

Hsing T, Eubank R (2015) Theoretical foundations of functional data analysis, with an introduction to linear operators. Wiley, New York

Ignaccolo R, Mateu J, Giraldo R (2014) Kriging with external drift for functional data for air quality monitoring. Stoch Environ Res Risk Assess 28:1171–1186

Kuelbs J (1970) Gaussian measures on a Banach space. J Funct Anal 5:354–367

Labbas A, Mourid T (2002) Estimation et prévision d'un processus autorégressif Banach. C R Acad Sci Paris Sér I 335:767–772

Marion JM, Pumo B (2004) Comparaison des modéles ARH(1) et ARHD(1) sur des données physiologiques. Ann ISUP 48:29–38

Mas A (2004) Consistance du prédicteur dans le modèle ARH(1): le cas compact. Ann ISUP 48:39–48

Mas A (2007) Weak-convergence in the functional autoregressive model. J Multivar Anal 98:1231–1261

Moritz S, Bartz-Beielstein T (2017) imputeTS: time series missing value imputation in R. R J 9:207–218

Nerini D, Monestiez P, Mantea C (2010) Cokriging for spatial functional data. J Multivar Anal 101:409–418

Pang W, Christakos G, Wang J-F (2009) Comparative spatiotemporal analysis of fine particulate matter pollution. Environmetrics 21:305–317

Parvardeh A, Jouzdani NM, Mahmoodi S, Soltani AR (2017) First order autoregressive periodically correlated model in Banach spaces: existence and central limit theorem. J Math Anal Appl 449:756–768

Paschalidou AK, Karakitsios S, Kleanthous S, Kassomenos PA (2011) Forecasting hourly PM_{10} concentration in Cyprus through artificial neural networks and multiple regression models: implications to local environmental management. Environ Sci Pollut Res 18:316–327

Ruiz-Medina MD (2011) Spatial autorregresive and moving average Hilbertian processes. J Multivar Anal 102:292–305

Ruiz-Medina MD, Álvarez-Liébana J (2019) A note on strong-consistency of componentwise ARH(1) predictors. Stat Probab Lett 145:224–248

Ruiz-Medina MD, Álvarez-Liébana J (2019) Strongly consistent autoregressive predictors in abstract Banach spaces. J Multivar Anal. https://doi.org/10.1016/j.jmva.2018.08.001

Ruiz-Medina MD, Espejo RM (2012) Spatial autoregressive functional plug-in prediction of ocean surface temperature. Stoch Environ Res Risk Assess 26:335–344

Ruiz-Medina MD, Espejo RM, Ugarte MD, Militino AF (2014) Functional time series analysis of spatio-temporal epidemiological data. Stoch Environ Res Risk Assess 28:943–954

Slini T, Kaprara A, Karatzas K, Mousiopoulos N (2006) PM_{10} forecasting for Thessaloniki, Greece. Environ Model Softw 21:559–565

Stadlober E, Hormann S, Pfeiler B (2008) Quality and performance of a PM_{10} daily forecasting model. Atmos Environ 42:1098–1109

Triebel T (1983) Theory of function spaces II. Birkhauser, Basel

Vieu P (2018) On dimension reduction models for functional data. Stat Prob Lett 136:134–138

Zhang L, Liu Y, Zhao F (2018) Singular value decomposition analysis of spatial relationships between monthly weather and air pollution index in China. Stoch Environ Res Risk Assess 32:733–748

Zolghadri A, Cazaurang F (2006) Adaptive nonlinear state-space modelling for the prediction of daily mean PM_{10} concentrations. Environ Model Softw 21:885–894