Spatial Statistics 29 (2019) 145–159

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Functional SAR models: With application to spatial econometrics

Wilmer Pineda-Ríos [a,b], Ramón Giraldo [b,*], Emilio Porcu [c,d]

[a] Statistics Faculty, Universidad Santo Tomás, Colombia
[b] Department of Statistics, Universidad Nacional de Colombia, Colombia
[c] Chair of Spatial Analytics Methods Center, School of Mathematics and Statistics, Newcastle University, UK
[d] Department of Mathematics, Universidad de Atacama, Chile

ARTICLE INFO

Article history:
Received 18 March 2018
Accepted 11 December 2018
Available online 19 December 2018

Keywords:
Autoregressive errors
Functional linear models
Spatial dependence
Spatial weight matrix

ABSTRACT

Simultaneous autoregressive (SAR) models have been extensively used for the analysis of spatial data in areas as diverse as demography, economy and geography. These are linear models with a scalar response, scalar explanatory variables and autoregressive errors. In this work we extend this modeling approach from scalar to functional covariates. Least squares and maximum likelihood are used as estimation methods of the parameters. A simulation study is considered for evaluating the performance of the proposed methodology. As an illustration, the model is used to establish the relationship between unsatisfied basic needs and curves of gross domestic product obtained in 32 departments of Colombia (districts of the country).

1. Introduction

In the last decades, the advances in equipments of collection and storage of data and in computer technology have enabled researchers in diverse branches of sciences such as, for instance, seismology, meteorology, agronomy, medicine or finance, recording data of characteristics varying over a continuum (time, space, depth, wavelength, etc.). This is the case of data collected by seismographs, rain gauges, thermometers, penetrometers, electroencephalographs, equipments of magnetic resonance and ultra-sound or even just by means of computer programs. In all these cases a number of subjects are observed densely over time, space, or both. Through the application of smoothing techniques,

* Corresponding author.
E-mail address: rgiraldoh@unal.edu.co (R. Giraldo).

https://doi.org/10.1016/j.spasta.2018.12.002

these data become functions (Zhang, 2013). Thus, beyond a series of records are obtained curves, surfaces or images. Functional Data Analysis (FDA) has arisen as a field of statistics which provides tools for modeling this type of information (Ramsay and Silverman, 2005).

The standard statistical techniques for modeling functional data are based on the assumption of independence between the functions, that is, if $X_1(t), \ldots, X_n(t)$ with $t \in T \subset \mathbb{R}$, is a sample of functional data, is assumed that $cov(X_i(t), X_j(t)) = 0$, for all $i \neq j$, $i, j = 1, \ldots, n$. Amongst others, functional linear models (Cardot et al., 1999, 2003; Chiou et al., 2004), functional analysis of variance (Cuevas et al., 2004; Ramsay and Silverman, 2005; Zoglat, 2008; Horváth and Kokoszka, 2012; Zhang, 2013) or multivariate methods for functional data (James and Hastie, 2001; Tarpey and Kinateder, 2003) are usually focused on independent functions. However, in several disciplines of applied sciences there exists an increasing interest in modeling correlated functional data, that is, when $cov(X_i(t), X_j(t)) \neq 0$. This is the case when samples of functions are observed over a discrete set of time points (temporally correlated functional data) or when these functions are observed over different sites or areas of a region (spatially correlated functional data). In these cases the above-mentioned methodologies may not be appropriate, as they do not incorporate dependence among functions into the analysis. For this reason, some statistical methods for modeling correlated variables have been extended to the functional context (Delicado et al., 2010; Ruiz-Medina and Salmerón, 2010).

Specifically in the context of spatially correlated data, let d be a positive integer and $\mathcal{D} \subset \mathbb{R}^d$. *Functional geostatistics* deals with functions observed at a given set of points $s_1, \ldots, s_n$ in $\mathcal{D}$ (Dabo-Niang and Yao, 2007; Giraldo et al., 2011, 2012; Caballero et al., 2013). *Functional areal data* (functional data in lattices) correspond to the case of $\mathcal{D}$ being a fixed and countable set. Usually there is a bijection between $\mathcal{D}$ and a partition of a geographical area and to any $\mathbf{s} \in \mathcal{D}$ is associated a function that summarizes the event in the area represented by $\mathbf{s}$. Some contributions in functional areal data can be found in Delicado and Broner (2008) and Zhang et al. (2016).

The motivation of this work is to provide a theoretical approximation to functional linear models on lattices. Simultaneous autoregressive (SAR) models (Ord, 1975) have been used for the analysis of spatial data on lattices. These are linear models which assume spatial autoregressive errors (Waller and Gotway, 2004; Schabenberger and Gotway, 2005). In this work we propose an extension of this model to the case where the covariate is functional. The main advantage of our model is that the spatial dependence is not restricted to the closer areas as in a CAR model (Schabenberger and Gotway, 2005). When there is a strong structure of spatial dependence, the SAR model can be a better option, because it takes into account all the information available. A preliminary account about the methodology here considered is given in Pineda-Ríos and Giraldo (2016). Some works related to functional autoregressive models are focused on functional responses (Ruiz-Medina, 2011; Aguilera-Morillo et al., 2017) or scalar response with conditional autoregressive (CAR) structure (Zhang et al., 2016; Ahmed, 2017).

The remainder of this paper is organized as follows: Section 2 gives a short review about SAR models. Section 3 defines its extension to the functional context. We show two estimation procedures and a method to find confidence bands for the functional parameter. Section 4 provides a simulation study to evaluate the performance of our estimation methods. Section 5 illustrates the methodology proposed with an application to an econometrical dataset. The article concludes with a brief discussion and suggestions for further research.

2. Simultaneous autoregressive model

In time series analysis, autoregressive models (Box et al., 2015) represent observations at time t as linear combinations of past values (AR model) or past errors (MA model). The extensions of these models to the spatial context are known as the conditionally autoregressive model (CAR) and the simultaneously autoregressive model (SAR), where spatial dependence is defined in terms of neighboring observations or neighboring errors, respectively.

In particular, the SAR model considers the following structure. Let $\{(Y_\mathbf{s}, X_\mathbf{s}); \mathbf{s} \in \mathcal{D} \subset \mathbb{R}^d\}$ be a bivariate stochastic process observed over a discrete fixed subset of $\mathcal{D}$ consisting of $\mathbf{s}_1, \ldots, \mathbf{s}_n$ spatial locations and where, for every $\mathbf{s} \in \mathcal{D}$, $Y_\mathbf{s}$ and $X_\mathbf{s}$ are two real-valued random variables. The SAR model is defined as (Waller and Gotway, 2004)

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\nu}$$

$$\boldsymbol{\nu} = \mathbf{B}\boldsymbol{\nu} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\mathbf{Y} = \left(Y_{\mathbf{s}_1}, \ldots, Y_{\mathbf{s}_n}\right)^T$, $\mathbf{X}$ is a $n \times 2$ design matrix, $\boldsymbol{\beta}$ is a two-dimensional vector of parameters, $\boldsymbol{\nu}$ is a vector of spatially dependent errors, $\mathbf{B} = [b_{ij}]$ is a $n \times n$ matrix of spatial dependence with $b_{ii} = 0$ and $\boldsymbol{\epsilon}$ is a vector of *iid* errors (i.e., $\mathbb{E}(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\epsilon}) = \Sigma$, a diagonal matrix). If all b_{ij} are zero, there is no spatial dependence and the model reduces to the traditional linear regression model with uncorrelated errors.

Solving Eq. (2.1) with respect to $\boldsymbol{\epsilon}$, we have

$$(\mathbf{I} - \mathbf{B})(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = \boldsymbol{\epsilon}, \tag{2.2}$$

where $\mathbf{I}$ is the identity matrix of size n. The model in (2.2) was introduced by Whittle (1954) and called "simultaneous" to describe the n autoregressions that occur simultaneously at each data location. The matrix of spatial dependence parameters $\mathbf{B}$ plays an important role. Usually the parameterization $\mathbf{B} = \rho\mathbf{W}$ is adopted, where $\mathbf{W}$ is a known spatial proximity matrix and ρ is a correlation coefficient (Waller and Gotway, 2004; Schabenberger and Gotway, 2005). Under such a parameterization, the SAR model can be written as (Waller and Gotway, 2004)

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon} \\ &= \mathbf{X}\boldsymbol{\beta} - \rho\mathbf{W}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Y}) + \boldsymbol{\epsilon}, \end{aligned} \tag{2.3}$$

showing that, in a SAR model, both response and covariate are spatially *lagged* (Waller and Gotway, 2004). For a well-defined model, invertibility of $(\mathbf{I} - \rho\mathbf{W})$ is necessary. This restriction imposes conditions on $\mathbf{W}$ and ρ. Haining (1993) showed that if λ_{max} and λ_{min} are the largest and smallest eigenvalues of $\mathbf{W}$, and if $\lambda_{\text{min}} < 0$ and $\lambda_{\text{max}} > 0$, then

$$\frac{1}{\lambda_{\text{min}}} \le \rho < \frac{1}{\lambda_{\text{max}}}.$$

Often, the rows of $\mathbf{W}$ are standardized in the sense that each element is divided by the corresponding sum by row. Then, $\lambda_{\text{max}} = 1$ and $\lambda_{\text{min}} \le -1$, so that $\rho < 1$. If ρ is known and $\Sigma = \sigma^2\mathbf{I}$, then generalized least squares (GLS) can be used to estimate $\boldsymbol{\beta}$ and σ^2. If ρ is unknown, iterative GLS is used to estimate all the parameters.

3. Functional SAR model

We now provide a SAR model with scalar response and a functional covariate. Let $\{(Y_{\mathbf{s}}, X_{\mathbf{s}}(t)); \mathbf{s} \in \mathcal{D} \subset \mathbb{R}^d, t \in \mathcal{T}\}$ be a bivariate stochastic process observed over a discrete fixed subset of $\mathcal{D}$ consisting of $\mathbf{s}_1, \ldots, \mathbf{s}_n$ spatial locations and where for every $\mathbf{s} \in \mathcal{D}$, $Y_{\mathbf{s}}$ is a real-valued random variable and $X_{\mathbf{s}}(t)$, $t \in \mathcal{T}$ is a real-valued random function (assumed to be in $L^2(\mathcal{T})$ with respect to the Lebesgue measure for every fixed s, i.e., corresponds to a square integrable stochastic process on the real interval $\mathcal{T}$). The functional SAR (FSAR) model is given by

$$\begin{cases} \mathbf{Y} = \boldsymbol{\beta}_0 + \displaystyle\int_{\mathcal{T}} \mathbf{X}(t)\beta(t)\mathrm{d}t + \boldsymbol{\nu} \\ \boldsymbol{\nu} = \rho\mathbf{W}\boldsymbol{\nu} + \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} \sim NMV(\mathbf{0}, \sigma^2\mathbf{I}), \end{cases} \tag{3.1}$$

where $\mathbf{Y}$ is defined as in Eq. (2.1), $\mathbf{X}(t) = (X_{s_1}(t), \ldots, X_{s_n}(t))^T$, $\boldsymbol{\beta}_0$ is a $n \times 1$ constant vector, $\beta(\cdot) \in L^2(\mathcal{T})$ is a functional parameter, $\mathbf{W}$ is a symmetric proximity matrix and NMV notation corresponds to a multivariate normal distribution. Notice that integration in (3.1) is intended as componentwise.

Without loss of generality, we assume $\boldsymbol{\beta}_0 = \mathbf{0}$. Thus the model (3.1) can be written as

$$\begin{aligned}\mathbf{Y} &= \int_{\mathcal{T}} \mathbf{X}(t)\beta(t)\mathrm{d}t + \rho\mathbf{W}\boldsymbol{\nu} + \boldsymbol{\epsilon} \\ &= \int_{\mathcal{T}} \mathbf{X}(t)\beta(t)\mathrm{d}t + \rho\mathbf{W}\mathbf{Y} - \rho\mathbf{W}\int_{\mathcal{T}} \mathbf{X}(t)\beta(t)\mathrm{d}t + \boldsymbol{\epsilon}.\end{aligned} \tag{3.2}$$

Solving with respect to $\boldsymbol{\epsilon}$, we obtain

$$\boldsymbol{\epsilon} = (I - \rho\mathbf{W})\left(\mathbf{Y} - \int_{\mathcal{T}} \mathbf{X}(t)\beta(t)\mathrm{d}t\right). \tag{3.3}$$

In Eq. (3.1), the functional parameter of interest is $\beta(t)$. Specifically, we estimate $\beta(\cdot)$ through least squares (Section 3.1) as well as through maximum likelihood techniques (Section 3.2). Both approaches are based on Eq. (3.3). Before going into estimation details, some preliminary results are needed. Our presentation throughout is simplified to avoid technical obfuscation that might distract the reader from the main message. We assume that both parameter and covariate can be expanded in terms of some basis functions. For an orthonormal basis $\{\varphi_j\}_{j\in\mathbb{N}}$ of $L^2(\mathcal{T})$, we have

$$\begin{cases} X_{\mathbf{s}_i}(t) = \sum_{j=1}^{\infty} a_{ij}\varphi_j(t) \\ \beta(t) = \sum_{j=1}^{\infty} b_j\varphi_j(t), \quad t \in \mathcal{T}. \end{cases} \tag{3.4}$$

In (3.4), it is assumed that $\{a_{ij}\}_{j\geq 1}$ are random variables related the ith area (represented by the ith location $\mathbf{s}_i$). Recall that the usual inner product on $L^2(\mathcal{T})$ is $\langle f, g\rangle := \langle f, g\rangle_{L^2(\mathcal{T})}$ defined as

$$\langle f, g\rangle = \int_{\mathcal{T}} f(t)g(t)\mathrm{d}t, \quad f, g \in L^2(\mathcal{T})$$

with induced norm $\|f\| = \sqrt{\langle f, f\rangle}$. Using Parseval's identity, we have

$$\sum_{j=1}^{\infty} \left|\langle \beta, \varphi_j\rangle\right|^2 = \|\beta\|^2 < \infty.$$

On the other hand, using dominated convergence, and since $\{\varphi_j\}_{j\geq 1}$ is an orthonormal basis, we have

$$\begin{aligned}\sum_{j=1}^{\infty} \left|\langle \beta, \varphi_j\rangle\right|^2 &= \sum_{j=1}^{\infty}\left|\left\langle \sum_{i=1}^{\infty} b_i\varphi_i, \varphi_j\right\rangle\right|^2 \\ &= \sum_{j=1}^{\infty}\left|\sum_{i=1}^{\infty} b_i\langle\varphi_i, \varphi_j\rangle\right|^2 \\ &= \sum_{j=1}^{\infty}\left|\sum_{i=1}^{\infty} b_i\delta_{ij}\right|^2 \\ &= \sum_{j=1}^{\infty} b_j^2,\end{aligned}$$

where δ_{ij} is the Kronecker delta. Hence,

$$\sum_{j=1}^{\infty} b_j^2 = \|\beta\|^2.$$

Now, if $\mathbb{E}(a_{ij}) = 0$ and $\mathbb{E}\left(a_{ij}^2\right) = \sigma_j^2$, for all $i = 1, \ldots, n$, i.e., if we assume spatial homoscedasticity, we have

$$\sum_{j=1}^{\infty} \sigma_j^2 = \int_{\mathcal{T}} \mathbb{E}\left(X_i^2(t)\right) \mathrm{d}t < \infty.$$

Again, since $\{\varphi_j\}_{j\geq 1}$ is an orthonormal basis of $L^2(\mathcal{T})$ and using dominated convergence, we conclude that

$$\begin{aligned}
\int_{\mathcal{T}} X_i(t)\beta(t)\mathrm{d}t &= \int_{\mathcal{T}} \left(\sum_{k=1}^{\infty} a_{ik}\varphi_k(t)\right)\left(\sum_{j=1}^{\infty} b_j\varphi_j(t)\right)\mathrm{d}t \\
&= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} a_{ik}b_j\left(\int_{\mathcal{T}} \varphi_k(t)\varphi_j(t)\mathrm{d}t\right) \\
&= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} a_{ik}b_j\delta_{jk} \\
&= \sum_{j=1}^{\infty} a_{ij}b_j. \qquad (3.5)
\end{aligned}$$

The idea of the estimation of $\beta(\cdot)$ is to consider estimates of the form

$$\beta^*(t) = \sum_{j=1}^{k} b_j\varphi_j(t),$$

where k is a relatively small number which is used as a smoothing parameter (Horváth and Kokoszka, 2012). This value can be estimated by using cross-validation (Ramsay and Silverman, 2005). Truncating the series (3.5) at k, we have

$$\int_{\mathcal{T}} X_i(t)\beta(t)\mathrm{d}t = \sum_{j=1}^{k} a_{ij}b_j. \qquad (3.6)$$

Using (3.6) and defining $\mathbf{A}_k = (a_{ij})_{n\times k}$ and $\mathbf{b}_k = (b_1, \ldots, b_k)^T$ the k-truncated errors vector in Eq. (3.3) is given by

$$\boldsymbol{\epsilon}_k = (\mathbf{I} - \rho_k\mathbf{W})(\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k). \qquad (3.7)$$

Note that when $\boldsymbol{\beta}_0 = b_0\mathbf{1}$, $b_0 \neq 0$ in (3.1), Eq. (3.7) holds defining $\mathbf{A}_k = (\mathbf{1}, a_{ij})_{n\times(k+1)}$ and $\mathbf{b}_k = (b_0, b_1, \ldots, b_k)^T$.

3.1. Least squares estimation

We propose to estimate $\beta(t)$ and ρ in model (3.1) based on the errors in (3.7). We can assume, as in the classical regression model (Fox, 1984), $\mathbf{X}(t)$ fixed or stochastic. In the first case, the assumption is $\boldsymbol{\epsilon} \sim NMV(\mathbf{0}, \sigma^2\mathbf{I})$. On the second one, it is assumed that $\mathbf{X}(t)$ and $\boldsymbol{\epsilon}$ are independent and thus $\boldsymbol{\epsilon}\big|\mathbf{X}(t) \sim NMV(\mathbf{0}, \sigma^2\mathbf{I})$. Specifically, we minimize respect to $\mathbf{b}_k$ and ρ_k the sum of squares

$$\begin{aligned}
\boldsymbol{\epsilon}_k^T\boldsymbol{\epsilon}_k &= (\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k)^T (I - \rho_k\mathbf{W})^2 (\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k) \\
&= (\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k)^T \mathbf{Z}_k (\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k) \\
&= \mathbf{Y}^T\mathbf{Z}_k\mathbf{Y} - \mathbf{Y}^T\mathbf{Z}_k\mathbf{A}_k\mathbf{b}_k - \mathbf{b}_k^T\mathbf{A}_k^T\mathbf{Z}_k\mathbf{Y} + \mathbf{b}_k^T\mathbf{A}_k^T\mathbf{Z}_k\mathbf{A}_k\mathbf{b}_k \\
&= \mathbf{Y}^T\mathbf{Z}_k\mathbf{Y} - 2\mathbf{b}_k^T\mathbf{A}_k^T\mathbf{Z}_k\mathbf{Y} + \mathbf{b}_k^T\mathbf{A}_k^T\mathbf{Z}_k\mathbf{A}_k\mathbf{b}_k,
\end{aligned}$$

with $\mathbf{Z}_k = (I - \rho_k \mathbf{W})^2$. Taking derivative on the right hand side in the last identity with respect to $\mathbf{b}_k$ we get

$$\frac{d\boldsymbol{\epsilon}_k^T\boldsymbol{\epsilon}_k}{d\mathbf{b}_k} = -2\mathbf{A}_k^T\mathbf{Z}_k\mathbf{Y} + 2\mathbf{A}_k^T\mathbf{Z}_k\mathbf{A}_k\mathbf{b}_k.$$

Now, taking derivative with respect to ρ_k we get

$$\begin{aligned}\frac{d\boldsymbol{\epsilon}_k^T\boldsymbol{\epsilon}_k}{d\rho_k} &= \mathbf{Y}^T\left(2\rho_k\mathbf{W}^2 - 2\mathbf{W}\right)\mathbf{Y} - 2\mathbf{b}_k^T\mathbf{A}_k^T\left(2\rho_k\mathbf{W}^2 - 2\mathbf{W}\right)\mathbf{Y} + \mathbf{b}_k^T\mathbf{A}_k^T\left(2\rho_k\mathbf{W}^2 - 2\mathbf{W}\right)\mathbf{A}_k\mathbf{b}_k \\ &= 2\rho_k\left(\mathbf{Y}^T\mathbf{W}^2\mathbf{Y} - 2\mathbf{b}_k^T\mathbf{A}_k^T\mathbf{W}^2\mathbf{Y} + \mathbf{b}_k^T\mathbf{A}_k^T\mathbf{W}^2\mathbf{A}_k\mathbf{b}_k\right) - 2\left(\mathbf{Y}^T\mathbf{W}\mathbf{Y} - 2\mathbf{b}_k^T\mathbf{A}_k^T\mathbf{W}\mathbf{Y} + \mathbf{b}_k^T\mathbf{A}_k^T\mathbf{W}\mathbf{A}_k\mathbf{b}_k\right) \\ &= 2\rho_k\left(\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k\right)^T\mathbf{W}^2\left(\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k\right) - 2\left(\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k\right)^T\mathbf{W}\left(\mathbf{Y} - \mathbf{A}_k\mathbf{b}_k\right).\end{aligned}$$

Hence, we obtain the following system of equations

$$\begin{cases}\mathbf{A}_k^T\left(\mathbf{I} - \hat{\rho}_k\mathbf{W}\right)^2\mathbf{A}_k\hat{\mathbf{b}}_k = \mathbf{A}_k^T\left(\mathbf{I} - \hat{\rho}_k\mathbf{W}\right)^2\mathbf{Y} \\ \rho_k\left(\mathbf{Y} - \mathbf{A}_k\hat{\mathbf{b}}_k\right)^T\mathbf{W}^2\left(\mathbf{Y} - \mathbf{A}_k\hat{\mathbf{b}}_k\right) = \left(\mathbf{Y} - \mathbf{A}_k\hat{\mathbf{b}}_k\right)^T\mathbf{W}\left(\mathbf{Y} - \mathbf{A}_k\hat{\mathbf{b}}_k\right).\end{cases} \tag{3.8}$$

The system (3.8) cannot be solved analytically. Following Schabenberger and Gotway (2005), we use an iterative procedure to find $\hat{\rho}_k$ and $\hat{\mathbf{b}}_k$:

1. Estimate $\mathbf{b}_k$, taking $\rho_k = 0$;
2. Estimate ρ_k using the last estimator for $\mathbf{b}_k$;
3. Estimate $\mathbf{b}_k$ using the last estimator for ρ_k;
4. Repeat steps 2 and 3 until convergence.

3.2. Maximum likelihood estimation

Now, we aim at finding the maximum likelihood estimator for $\beta(\cdot)$ and ρ. Using (3.3) and the assumption on $\boldsymbol{\epsilon}$, we have

$$\mathbb{E}\left((\mathbf{I} - \rho\mathbf{W})\left(\mathbf{Y} - \int_{\mathcal{T}}\mathbf{X}(t)\beta(t)\mathrm{d}t\right)\Big|\mathbf{X}\right) = \mathbf{0}.$$

If $(\mathbf{I} - \rho\mathbf{W})$ is an invertible matrix, using properties of expected value, we get

$$\mathbb{E}(\mathbf{Y}|\mathbf{X}) = \int_{\mathcal{T}}\mathbf{X}(t)\beta(t)\mathrm{d}t.$$

We thus obtain

$$\begin{aligned}\mathrm{cov}(\mathbf{Y}|\mathbf{X}) &= \mathrm{cov}(\boldsymbol{\nu}) \\ &= \mathrm{cov}\left((\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}\right) \\ &= (\mathbf{I} - \rho\mathbf{W})^{-1}\mathrm{cov}(\boldsymbol{\epsilon})(\mathbf{I} - \rho\mathbf{W})^{-1} \\ &= \sigma^2\left((\mathbf{I} - \rho\mathbf{W})^2\right)^{-1}.\end{aligned}$$

Using the last identity, in concert with the assumption on $\boldsymbol{\epsilon}$ in (3.1) we get

$$\mathbf{Y}|\mathbf{X} \sim NMV\left(\int_{\mathcal{T}}\mathbf{X}(t)\beta(t)\mathrm{d}t, \sigma^2\left((\mathbf{I} - \rho\mathbf{W})^2\right)^{-1}\right).$$

Consequently, the likelihood function (L), conditioning to $\mathbf{X}$, is given by

$$L(\beta(t), \sigma^2, \rho) = (2\pi)^{-n/2}(\sigma^2)^{-n/2}\left|(\mathbf{I} - \rho\mathbf{W})^2\right|^{1/2} \exp\left\{-\frac{1}{2\sigma^2}\left(\mathbf{Y} - \int_{\mathcal{T}}\mathbf{X}(t)\beta(t)\mathrm{d}t\right)^T(\mathbf{I} - \rho\mathbf{W})^2\left(\mathbf{Y} - \int_{\mathcal{T}}\mathbf{X}(t)\beta(t)\mathrm{d}t\right)\right\}.$$

Taking natural logarithm, we obtain (modulo some constant) that $l = ln(L)$ admits the expression

$$\begin{aligned} l(\beta(t), \sigma^2, \rho) = &-\frac{n}{2}\ln(\sigma^2) + \ln\Big(|(\mathbf{I}-\rho\mathbf{W})|\Big) \\ &- \frac{1}{2\sigma^2}\left(\mathbf{Y} - \int_{\mathcal{T}} \mathbf{X}(t)\beta(t)\mathrm{d}t\right)^T (\mathbf{I}-\rho\mathbf{W})^2 \left(\mathbf{Y} - \int_{\mathcal{T}} \mathbf{X}(t)\beta(t)\mathrm{d}t\right). \end{aligned}$$

Here we need again to truncate the model at some $k = k_n$. Using the errors vector in (3.7), the function to maximize is approximated to

$$l_k(\mathbf{b}_k, \sigma^2, \rho) = -\frac{n}{2}\ln(\sigma^2) + \ln(|(\mathbf{I}-\rho_k\mathbf{W})|) - \frac{1}{2\sigma^2}(\mathbf{Y}-\mathbf{A}_k\mathbf{b}_k)^T(\mathbf{I}-\rho_k\mathbf{W})^2(\mathbf{Y}-\mathbf{A}_k\mathbf{b}_k).$$

Then,

$$\begin{aligned} \frac{\partial l_k}{\partial \mathbf{b}_k} &= -\frac{1}{2\sigma^2}\frac{d\boldsymbol{\epsilon}_k^T\boldsymbol{\epsilon}_k}{d\mathbf{b}_k}, \\ \frac{\partial l_k}{\partial \rho_k} &= -\mathrm{Tr}\left((\mathbf{I}-\rho_k\mathbf{W})^{-1}\mathbf{W}\right) - \frac{d\boldsymbol{\epsilon}_k^T\boldsymbol{\epsilon}_k}{d\rho_k}, \\ \frac{\partial l_k}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(\mathbf{Y}-\mathbf{A}_k\mathbf{b}_k)^T(\mathbf{I}-\rho_k\mathbf{W})^2(\mathbf{Y}-\mathbf{A}_k\mathbf{b}_k), \end{aligned}$$

where ∂ denotes partial derivative. Hence, the log-likelihood function should be maximized by numerical methods. Note that maximizing the log-likelihood is equivalent to minimize the sum of squares errors. Using any of the proposed estimation methods, we write

$$\hat{\beta}_k(t) = \Phi_k(t)^T\hat{\mathbf{b}}_k, \tag{3.9}$$

where $\Phi_k(t) = (\varphi_1(t), \dots, \varphi_k(t))^T$ is the k-dimensional vector of the first k elements of the initial basis of $L^2(\mathcal{T})$.

3.3. Confidence bands for $\beta(t)$

Now, let us assume that ρ is known and $k = k_n$ is fixed, in order to establish a confidence band for $\beta(t)$. We write

$$\beta_k(t) = \Phi_k(t)^T\mathbf{b}_k.$$

Set $\Sigma_k = \mathbf{A}_k^T(\mathbf{I}-\rho\mathbf{W})^2\mathbf{A}_k$. Notice that Σ_k is a symmetric matrix with real entries and consequently can be diagonalized as

$$\Sigma_k = \mathbf{PJP}^{-1},$$

where $\mathbf{J}$ is the diagonal matrix of eigenvalues of Σ_k. Using the assumption that ρ is known, we have

$$\hat{\mathbf{b}}_k = \Sigma_k^{-1}\mathbf{A}_k^T(\mathbf{I}-\rho\mathbf{W})^2\mathbf{Y}.$$

Then, computing the expected value of $\hat{\mathbf{b}}_k$ conditionally on $\mathbf{U}_k$, we get

$$\begin{aligned} \mathbb{E}\left(\hat{\mathbf{b}}_k\right) &= \Sigma_k^{-1}\mathbf{A}_k^T(\mathbf{I}-\rho\mathbf{W})^2\,\mathbb{E}(\mathbf{Y}) \\ &= \Sigma_k^{-1}\mathbf{A}_k^T(\mathbf{I}-\rho\mathbf{W})^2\mathbf{A}_k\mathbf{b}_k \\ &= \Sigma_k^{-1}\Sigma_k\mathbf{b}_k \\ &= \mathbf{b}_k. \end{aligned}$$

Hence, $\hat{\mathbf{b}}_k$ is unbiased estimator for $\mathbf{b}_k$. On the other hand,

$$\begin{aligned} \mathrm{cov}\left(\hat{\mathbf{b}}_k\right) &= \Sigma_k^{-1}\mathbf{A}_k^T(\mathbf{I}-\rho\mathbf{W})^2\mathrm{cov}(\mathbf{Y})(\mathbf{I}-\rho\mathbf{W})^2\mathbf{A}_k\Sigma_k^{-1} \\ &= \sigma^2\Sigma_k^{-1}\mathbf{A}_k^T(\mathbf{I}-\rho\mathbf{W})^2(\mathbf{I}-\rho\mathbf{W})^{-2}(\mathbf{I}-\rho\mathbf{W})^2\mathbf{A}_k\Sigma_k^{-1} \end{aligned}$$

$$
\begin{aligned}
&= \sigma^2 \Sigma_k^{-1} \mathbf{A}_k^T (\mathbf{I} - \rho \mathbf{W})^2 \mathbf{A}_k \Sigma_k^{-1} \\
&= \sigma^2 \Sigma_k^{-1} \Sigma_k \Sigma_k^{-1} \\
&= \sigma^2 \Sigma_k^{-1}.
\end{aligned}
$$

Using the results in Schabenberger and Gotway (2005), the generalized least squares can be used to estimate σ^2. Thus,

$$
\hat{\sigma}^2 = \frac{\left(\mathbf{Y} - \mathbf{A}_k \hat{\mathbf{b}}_k\right)^T (\mathbf{I} - \rho \mathbf{W})^2 \left(\mathbf{Y} - \mathbf{A}_k \hat{\mathbf{b}}_k\right)}{n - r}, \tag{3.10}
$$

where $r = \operatorname{tr}\left(\mathbf{A}_k \Sigma_k^{-1} \mathbf{A}_k^T (\mathbf{I} - \rho \mathbf{W})^2\right)$, with tr being the trace operator. Hence, a confidence band of $100(1 - \alpha)\%$ is

$$
\hat{\beta}_k(t) \pm z_{1-\frac{\alpha}{2}} \hat{\sigma} \sqrt{\Phi_k(t)^T \Sigma_k^{-1} \Phi_k(t)}. \tag{3.11}
$$

If ρ is unknown, we can use a feasible least squares estimator based on $\hat{\rho}_k$ to estimate Σ_k in (3.11).

3.4. Asymptotic inference

In this section we assume that k in (3.6) is fixed. We also consider two scenarios for the spatial dependence parameter (ρ known and unknown, respectively).

3.4.1. ρ known

Under this assumption, the maximum likelihood estimator of $\mathbf{b}_k$ is equal to the generalized least squares estimator. Thus,

$$
\begin{aligned}
\hat{\mathbf{b}}_k &= \Sigma_k^{-1} \mathbf{A}_k^T (\mathbf{I} - \rho \mathbf{W})^2 \mathbf{Y} \\
&= \Sigma_k^{-1} \mathbf{A}_k^T (\mathbf{I} - \rho \mathbf{W})^2 \mathbf{A}_k \mathbf{b}_k + \Sigma_k^{-1} \mathbf{A}_k^T (\mathbf{I} - \rho \mathbf{W})^2 (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\epsilon}_k \\
&= \mathbf{b}_k + \Sigma_k^{-1} \mathbf{A}_k^T (\mathbf{I} - \rho \mathbf{W}) \boldsymbol{\epsilon}_k \\
&= \mathbf{b}_k + \left(\frac{1}{n} \Sigma_k\right)^{-1} \left(\frac{1}{n} \mathbf{A}_k^T (\mathbf{I} - \rho \mathbf{W}) \boldsymbol{\epsilon}_k\right).
\end{aligned}
$$

A condition that ensures the consistency of $\hat{\mathbf{b}}_k$ for $\mathbf{b}_k$ is

(A1) $\lim_{n \to \infty} \frac{1}{n} \Sigma_k = \mathbf{Q}_k$,

where $\mathbf{Q}_k$ is a nonsingular matrix. Now, using the Lindeberg–Feller central limit theorem (Resnick, 2013) and the above condition we have (Schabenberger and Gotway, 2005)

$$
\sqrt{n}\left(\hat{\mathbf{b}}_k - \mathbf{b}_k\right) = \left(\frac{1}{n} \Sigma_k\right)^{-1} \left(\frac{1}{\sqrt{n}} \mathbf{A}_k^T (\mathbf{I} - \rho \mathbf{W}) \boldsymbol{\epsilon}_k\right) \xrightarrow{d} NMV\left(\mathbf{0}, \mathbf{Q}_k^{-1}\right). \tag{3.12}
$$

3.4.2. ρ unknown

In Section 3.1 we obtained estimators of ρ_k and $\mathbf{b}_k$. Based on these ones $\mathbf{b}_k$ can be re-estimated as

$$
\tilde{\mathbf{b}}_k = \Sigma_k^{-1} \mathbf{A}_k^T \left(\mathbf{I} - \hat{\rho}_k \mathbf{W}\right)^2 \mathbf{Y}. \tag{3.13}
$$

Now taking into account the conditions

(A2) $\frac{1}{n} \mathbf{A}_k^T \left((\mathbf{I} - \hat{\rho}_k \mathbf{W})^2 - (\mathbf{I} - \rho \mathbf{W})^2\right) \mathbf{A}_k \xrightarrow{p} 0$;

(A3) $\frac{1}{n} \mathbf{A}_k^T \left((\mathbf{I} - \hat{\rho}_k \mathbf{W})^2 - (\mathbf{I} - \rho \mathbf{W})^2\right) \boldsymbol{\epsilon}_k \xrightarrow{p} 0$,

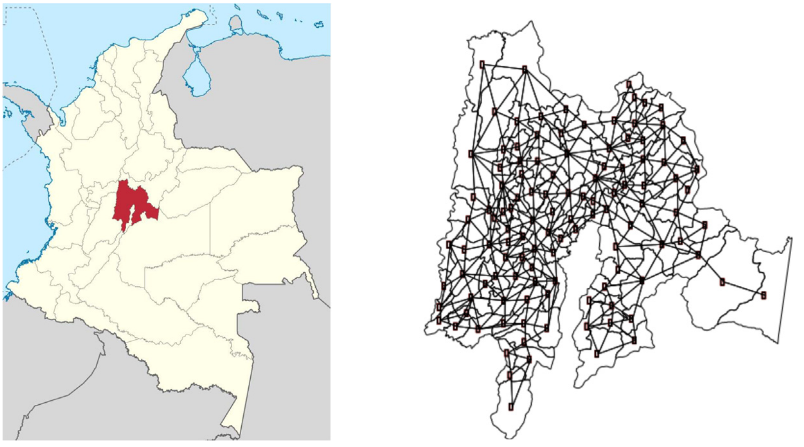

Fig. 1. Spatial layout of Colombia (left panel). The Department (county) highlighted in left panel corresponds to Cundinamarca. Centroids of municipalities of this Department (right panel).

Remark 1. When ρ is unknown, we can use the above conditions to give a partial result about the convergence of the estimator $\hat{\beta}(t)$ for $\beta(t)$, assuming k fixed. More exactly, if the assumptions (A1)–(A3) are satisfied, then

$$\sqrt{n}\left(\tilde{\mathbf{b}}_k - \mathbf{b}_k\right) \xrightarrow{d} NMV\left(\mathbf{0}, \mathbf{Q}_k^{-1}\right). \tag{3.14}$$

In the case where k converges towards infinity, Ahmed (2017) has established certain assumptions related to the spatial dependence structure that can be considered for the model presented here.

4. Simulation study

We consider several simulation scenarios for the model in Eq. (3.1) to assess the performance of the algorithms proposed. To carry out the simulations, we use the spatial layout of 117 municipalities of the Department of Cundinamarca, located in the central region of Colombia (Fig. 1). Under each scenario, the simulations are generated adding to a sinusoidal trend 117 realizations of *iid* Gaussian processes discretized on a fine grid of 101 points $\{t_1, \dots, t_{101}\} \in [0, 100]$. Usually to carry out simulations of covariates and parameters in a functional regression model are chosen trigonometric, exponential and polynomial functions (Ramsay and Silverman, 2005; Amato et al., 2006; Ramsay et al., 2009; Giraldo et al., 2018), because posteriorly its behavior can be easily fitted by means of basis functions.

Specifically, we follow the next steps:

1. For each municipality in Fig. 1, we simulate a discrete set of values of a functional covariate through $X_i(t_j) = \cos(t_j) + \xi_i(t_j)$, $i = 1, \dots, 117$; $j = 0, 1, \dots, 100$, with $\xi_i(t_j)$ a realization of a white-noise. These values are then smoothed by using a B-spline basis for obtaining the set of curves $X_i(t)$, $i = 1, \dots, 117$ (Fig. 2). We use the same number of basis functions considered in Section 5. $X_i(t_j)$ is defined as in Amato et al. (2006) and Giraldo et al. (2018).
2. A spatial weight matrix $\mathbf{W}_{117\times117}$ (row-standardized) is built based on the first order contiguity relations of the 117 municipalities in right panel of Fig. 1. In this matrix $w_{ij} > 0$ if the municipalities i and j are neighbors (if their centroids are connected), and 0 otherwise. We also set $w_{ii} = 0$.
3. To evaluate the effect of the spatial correlation in the procedure proposed, we consider several values for ρ parameter. In particular we take $\rho = 0.1, 0.3, 0.5, 0.7$ and 0.9.
4. The functional regression parameter is defined as $\beta(t) = e^{-\frac{t}{10}}\left(\left(\frac{t}{10}\right)^2 + 3\left(\frac{t}{10}\right) - 4\right)$ (see Fig. 3).

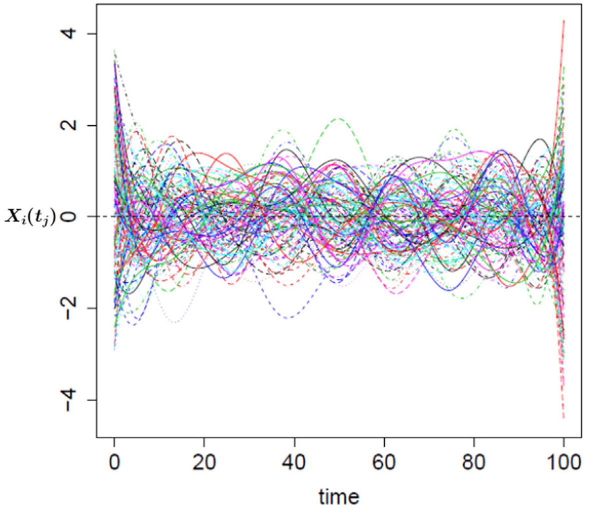

Fig. 2. 117 simulations of $X(t_j) = \cos(t_j) + \xi(t_j)$, $j = 0, 1, \ldots, 100$, with $\xi(t_j)$ a realization of a white-noise. The discrete values obtained are expanding by using a B-spline basis.

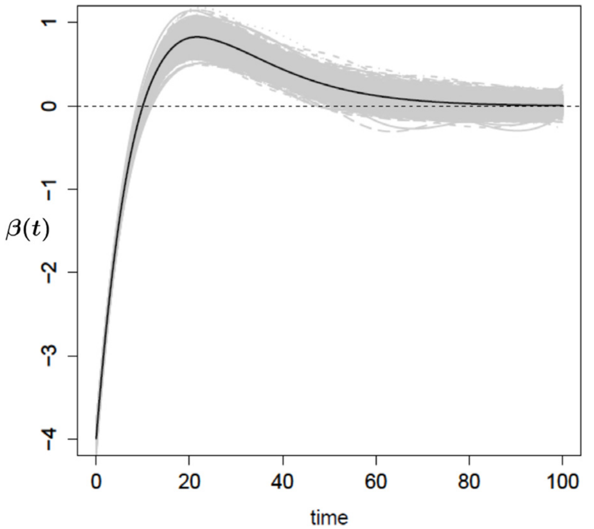

Fig. 3. Functional parameter $\beta(t)$ (black line) and 1000 estimations of this one (gray lines) obtained with $\rho = 0.5$ and $\sigma^2 = 1$.

5. We generate a vector $\boldsymbol{\epsilon}$ of Gaussian errors with zero mean and covariance matrix $\mathbf{I}$, i.e., $\sigma^2 = 1$. After, we simulate a value of vector $\mathbf{Y}$ according to the following equation

$$\mathbf{Y} = \mathbf{A}_k\mathbf{b}_k + (\mathbf{I} - \rho\mathbf{W})^{-1}\epsilon,$$

with $\mathbf{A}_k$ and $\mathbf{b}_k$ defined in Section 3.1.
6. The parameters $\beta(t)$, σ^2 and ρ are estimated, using the iterative procedure detailed in Section 3.2.

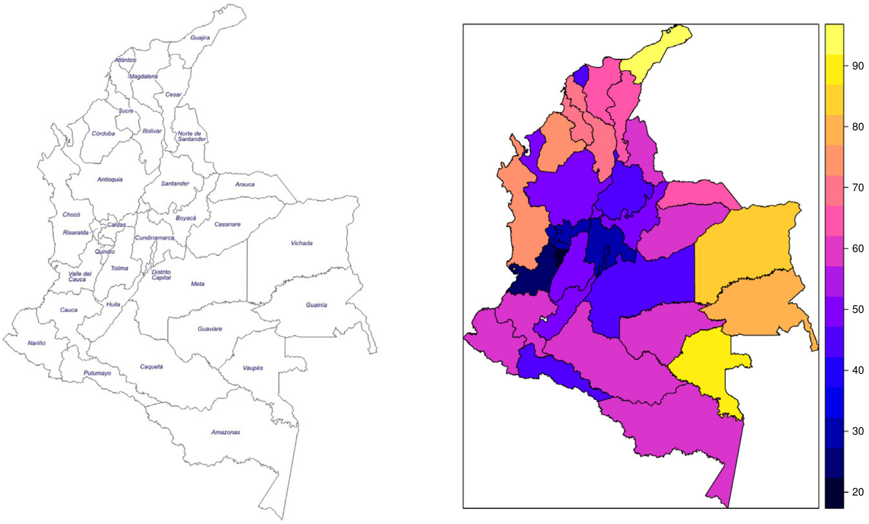

Fig. 4. Departments of Colombia (left panel) and choropleth map of index of Unsatisfied Basic Need in Colombia (right panel).

Table 1
Summary statistics of estimations for the spatial dependence parameter ρ, the mean of integrated squared error (*MISE*) and the variance σ^2.

ρ	$\hat{\rho}$	*MISE*	$\hat{\sigma}^2$
0.1	0.09	0.11	1.07
0.3	0.31	0.11	0.80
0.5	0.50	0.11	0.93
0.7	0.70	0.10	0.81
0.9	0.90	0.10	0.80

7. The steps 5 and 6 are repeated 1000 times. In each case, $\hat{\beta}_i(t)$ is calculated. In Fig. 3 are shown the estimations when $\rho = 0.5$. Also, we calculate $\hat{\sigma}_i^2$, $\hat{\rho}_i$ and the integrated squared error (ISE) of $\hat{\beta}_i(t)$, using the equation
$$\|\hat{\beta}_i - \beta\|^2 = \int_{\mathcal{T}} \left(\hat{\beta}_i(t) - \beta(t)\right)^2 \mathrm{d}t.$$
8. Finally, we calculate the mean of integrated squared error (MISE), $\hat{\sigma}^2 = \frac{1}{1000}\sum_{i=1}^{1000} \hat{\sigma}_i^2$ and $\hat{\rho} = \frac{1}{1000}\sum_{i=1}^{1000} \hat{\rho}_i$.

The results of the procedure described above are shown in Table 1.

The results in Table 1 indicate that the algorithm implemented allows to accurately estimate the spatial dependence parameter ρ. Also show that with a k relatively small are obtained estimations only slightly biased of σ^2 and $\beta(t)$. In this way, we can conclude from a practical point of view that the methodology proposed has a good performance.

5. Relationship between Unsatisfied Basic Need and Gross Domestic Product in Colombia

We show an application of the methodology introduced in Section 3 to an econometric dataset. For each of 32 departments of Colombia (Fig. 4, left), we consider data of Unsatisfied Basic Needs (UBN) index in 2015 and Agriculture's share of the Gross Domestic Product (AGDP) (recorded annually

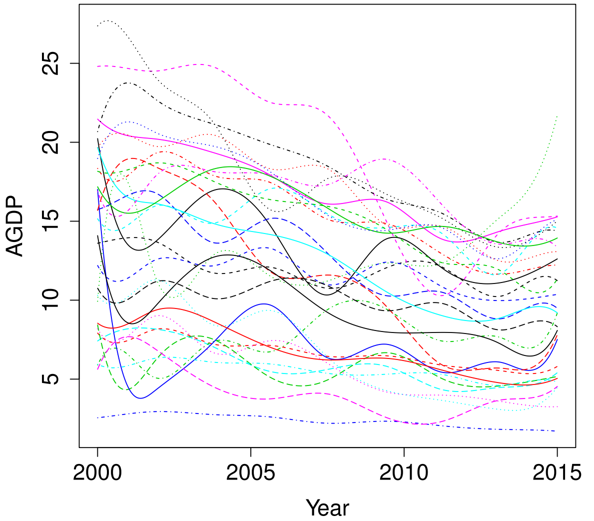

Fig. 5. Curves of AGDP (Agriculture's share of the Gross Domestic Product) for departments of Colombia, obtained after expanding the data by using a 11 B-splines basis.

between 2000 and 2015). These data were obtained from the Colombian Administrative Department of Statistics (DANE, by its acronym in Spanish). Our motivation is to establish the relationship between these variables as a contribution to explain the regional variation of UBN in Colombia. Taking into account that this is an areal dataset (data of UNB and AGDP for each department) we consider appropriate to use the FSAR model here proposed.

Initially, the discrete data of AGDP were expanded in terms of 11 B-spline basis functions (Fig. 5). The number of basis was selected by using cross-validation (Ramsay and Silverman, 2005). We start the analysis describing the behavior of the AGDP curves (Figs. 5–7) and UBN data (Fig. 4) and posteriorly we estimate a FSAR model (Eq. (3.1)) taking UBN as an scalar response and AGDP as a functional covariate.

According to Fig. 5, we observe that AGDP tends, in general terms, to decrease between 2000 and 2012. This is directly related to the monetary efforts that the departmental governments have concentrated on public administration and defense, i.e., on policies associated with the armed conflict that Colombia has experienced during the last 50 years. We also note that AGDP has slightly increased after 2012. This is due to the fact that the Colombian government has modified some policies in order to implement sustainable rural development and these have been geared towards making agriculture attractive (particularly for the younger generations).

In order to evaluate the presence of atypical functional data, we use the methodology given in Hyndman and Shang (2010). These authors extend the bagplot proposed by Rousseeuw et al. (1999) to the functional context. Initially a robust functional principal component analysis (Bali et al., 2011) is carried out. Then, by using the two first principal components, is plotted a bagplot to identify the atypical scores. Finally is obtained a functional bagplot which displays the median curve (the curve with the greatest depth), and the inner and outer regions, allowing to visualize the behavior of the atypical curves. The methodology identify scale outliers (curves that lie outside the range of the majority of the data) and shape outliers (curves masked among the rest of the curves) (Hyndman and Shang, 2010).

Fig. 6 indicates that there are some outliers (points out the bags in left panel). These correspond to AGDP curves of Arauca, Chocó and Meta departments (see Fig. 4 to identify their location). Also we note in right panel of Fig. 6 that these curves are distanced from the others (the curves of AGDP corresponding to Arauca and Chocó are scale outliers whereas Meta is a shape outlier). This is due

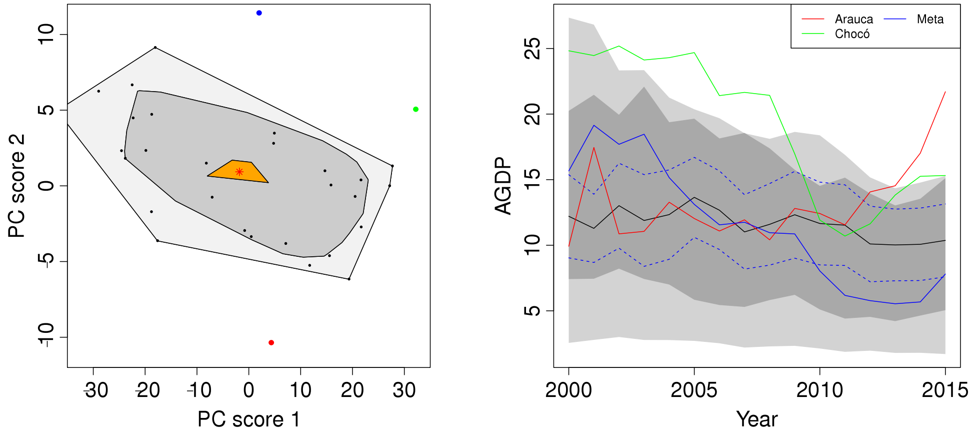

Fig. 6. Bagplot (left) and functional bagplot (right) of AGDP (Agriculture's share of the Gross Domestic Product) curves. The dark and light gray regions show the bag and fence regions, respectively. The red asterisk is the Tukey depth median. Black curve in right figure corresponds to the median. PC score 1 and PC score 2 are the first two functional principal components obtained with the AGDP curves . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

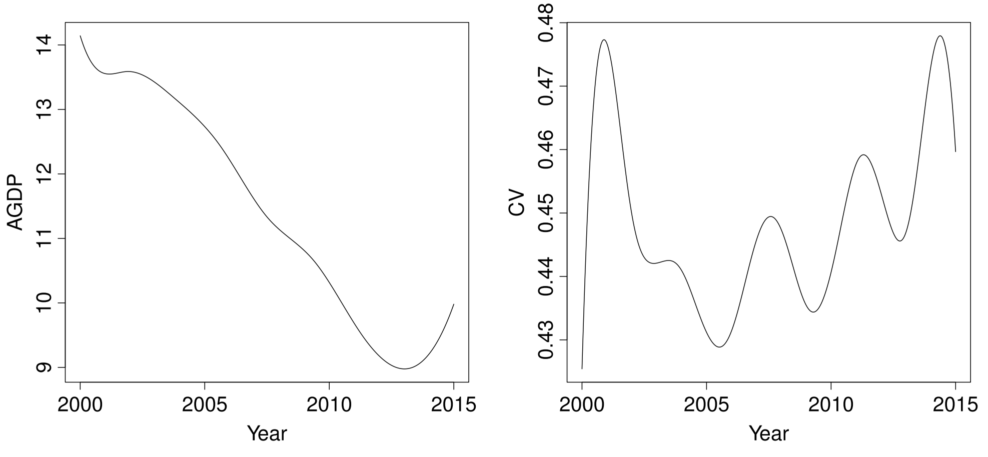

Fig. 7. Estimated functional mean (left panel) and functional coefficient of variation (right panel) of AGDP (Agriculture's share of the Gross Domestic Product).

to different reasons, which include among others geographic and environmental conditions, the presence of illegal crops and the influence of armed groups in the productive activities.

Fig. 7 shows the functional mean (left) and functional coefficient of variation (right) of AGDP. The mean plot confirms that there is a decreasing trend until 2012. On the other hand we can see that the variation coefficient varies approximately between 42% and 47% indicating that there is a relatively high variability in AGDP which remains homogeneous during the study period. This is due to the fact that in the poorest departments investment in agriculture (in percentage terms) is much lower.

With regard to UBN, the choropleth map (Fig. 4, right) indicates that the departments with the most adverse conditions (yellow colors) are Guajira, Vaupés, Vichada and Guainía (Fig. 4, left). In these ones even basic health care is below standard. The other side of the coin corresponds to departments in the central region of the country, where the rural investment has been growing in recent years.

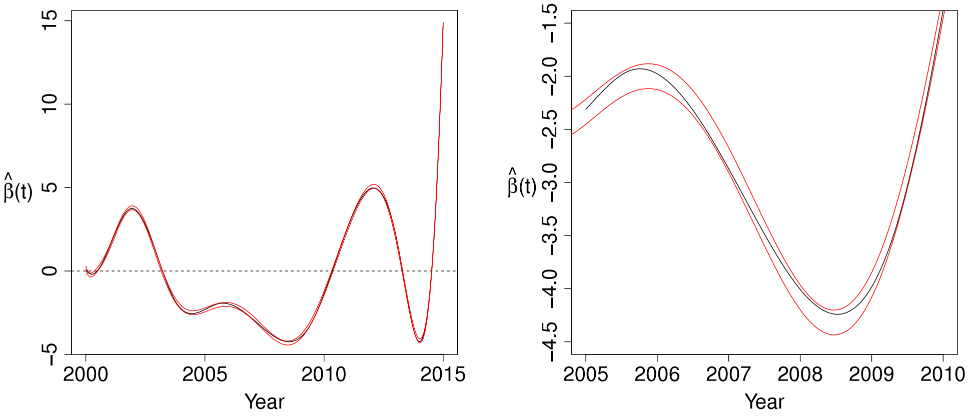

Fig. 8. Left Panel: Estimation of β(t) (black line) and 95% confidence bands (red lines). Right Panel: Enlargement of confidence bands from 2005 to 2010. Confidence bands are calculated by using $\hat{\rho}_k$ in the estimation of Σ_k . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We adopted a FSAR model to establish the relationship between UBN and AGDP. Our interest is the estimation of the functional parameter $\beta(t)$. This can be used as an indicator of the impact of public investments in agriculture (implemented by the departments of Colombia during 16 years) on the UNB index of the year 2015. The model is estimated by using the methodology exposed in Section 3. Fig. 8 shows the estimation of $\beta(t)$ and its 95% confidence bands.

It is expected that the greater the investment in agriculture, the lower the UBN index. Obviously this effect is not immediate but is reflected years later. According to the estimation in Fig. 8, we can conclude that AGDP has shown a positive effect on the UBN index of the year 2015 in two periods of time (2003–2011 and 2013–2014, respectively). This relationship is particularly evident in departments such as Valle del Cauca (southwest region), Quindío, Risaralda, Caldas, Cundinamarca and Santander (central region), Meta (Eastern plains), Putumayo (Amazon area) and Atlántico in the north coast of the country (see blue colors in the choropleth map in right panel of Fig. 4).

6. Concluding remarks

We have defined the functional spatial autoregressive model for areal data and established a methodology to carry out some inferential procedures. However there are many additional points which deserve study. According to the simulation results, we can conclude that the estimations, based on our iterative procedure, have good performance. Nevertheless, this computational study should be expanded. All the simulations were obtained assuming a fixed k. It would be important to check the behavior of bias and variability (MISE) of $\hat{\sigma}^2$, $\hat{\beta}(t)$, $\hat{\rho}$ when k increases. Other modeling approaches considering a functional SAR model with several functional and scalars covariates could be considered. In this case, the estimation of the parameters and the treatment of collinearity are challenges to be addressed. Another interesting point is the comparison with other functional models for spatial data (for example with a functional CAR model). In this scenario simulations under different spatial correlation structures could be carried out. Our premise is that the functional SAR model here described has a better performance than other similar models when the spatial dependence increases because it allows to catch all the spatial dependence underlying. The definition of a functional Gaussian Markov random field is, at the best of our knowledge, an open research field that can be studied.

From a practical point of view, we found that the model based on AGDP data has allowed to recognize some periods of time where the investment in agriculture has had a positive effect on the reduction of UBN index. This modeling approach is innovative in the field of spatial econometrics. Application to other types of data could be explored.

References

Aguilera-Morillo, M.C., Durbán, M., Aguilera, A.M., 2017. Prediction of functional data with spatial dependence: a penalized approach. Stoch. Environ. Res. Risk Assess. 31 (1), 7–22.
Ahmed, M., 2017. Contribution à la statistique spatiale et l'analyse de données fonctionnelles (Ph.D. thesis), Department of Mathematics, University of Lille.
Amato, U., Antoniadis, A., De Feis, I., 2006. Dimension reduction in functional regression with applications. Comput. Statist. Data Anal. 50 (9), 2422–2446.
Bali, J.L., Boente, G., Tyler, D.E., Wang, J.-L., et al., 2011. Robust functional principal components: a projection-pursuit approach. Ann. Statist. 39 (6), 2852–2882.
Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M., 2015. Time series analysis: forecasting and control. John Wiley & Sons.
Caballero, W., Giraldo, R., Mateu, J., 2013. A universal kriging approach for spatial functional data. Stoch. Environ. Res. Risk Assess. 27 (7), 1553–1563.
Cardot, H., Ferraty, F., Mas, A., Sarda, P., 2003. Testing hypotheses in the functional linear model. Scand. J. Stat. 30 (1), 241–255.
Cardot, H., Ferraty, F., Sarda, P., 1999. Functional linear model. Statist. Probab. Lett. 45 (1), 11–22.
Chiou, J., Müller, H., Wang, J., 2004. Functional response models. Statist. Sinica 14, 675–693.
Cuevas, A., Febrero-Bande, M., Fraiman, R., 2004. An ANOVA test for functional data. Comput. Statist. Data Anal. 47 (1), 111–122.
Dabo-Niang, S., Yao, F., 2007. Kernel regression estimation for continuous spatial processes. Math. Methods Statist. 16 (4), 298–317.
Delicado, P., Broner, S., 2008. Distance-based LISA Maps for Multivariate Lattice Data. In: Technical Report, Universitat Politècnica de Catalunya, pp. 1–20.
Delicado, P., Giraldo, R., Comas, C., Mateu, J., 2010. Statistics for spatial functional data: some recent contributions. Environmetrics 21 (3–4), 224–239.
Fox, J., 1984. Linear Statistical Models and Related Methods with Applications to Social Research. John Wiley & Sons.
Giraldo, R., Caballero, W., Camacho-Tamayo, J., 2018. Mantel test for spatial functional data. Adv. Statist. Anal. 102 (1), 21–39.
Giraldo, R., Delicado, P., Mateu, J., 2011. Ordinary kriging for function-valued spatial data. Environ. Ecol. Stat. 18 (3), 411–426.
Giraldo, R., Mateu, J., Delicado, P., 2012. geofd: an R package for function-valued geostatistical prediction. Rev. Colomb. Estad. 35 (3), 385–407.
Haining, R., 1993. Spatial Data Analysis in the Social and Environmental Sciences. Cambridge University Press.
Horváth, L., Kokoszka, P., 2012. Inference for Functional Data with Applications. Springer.
Hyndman, R.J., Shang, H.L., 2010. Rainbow plots, bagplots, and boxplots for functional data. J. Comput. Graph. Statist. 19 (1), 29–45.
James, G., Hastie, T., 2001. Functional linear discriminant analysis for irregularly sampled curves. J. R. Stat. Soc. Ser. B Stat. Methodol. 63 (3), 533–550.
Ord, K., 1975. Estimation methods for models of spatial interaction. J. Amer. Statist. Assoc. 70 (349), 120–126.
Pineda-Ríos, W., Giraldo, R., 2016. Functional SAR Model, ArXiv e-prints, 1609.03680, http://adsabs.harvard.edu/abs/2016arXiv160903680P.
Ramsay, J., Hooker, G., Graves, S., 2009. Functional Data Analysis with R and MATLAB. Springer Science & Business Media.
Ramsay, J., Silverman, B., 2005. Functional Data Analysis. Springer.
Resnick, S., 2013. A Probability Path. Springer Science & Bussines Media.
Rousseeuw, P.J., Ruts, I., Tukey, J.W., 1999. The bagplot: a bivariate boxplot. Amer. Statist. 53 (4), 382–387.
Ruiz-Medina, M.D., 2011. Spatial autoregressive and moving average hilbertian processes. J. Multivariate Anal. 102 (2), 292–305.
Ruiz-Medina, M., Salmerón, R., 2010. Functional maximum likelihood estimation of ARH(p) models. Stoch. Environ. Res. Risk Assess. 24 (1), 131–146.
Schabenberger, O., Gotway, C.A., 2005. Statistical Methods for Spatial Data Analysis. Chapman & Hall.
Tarpey, T., Kinateder, K., 2003. Clustering functional data. J. Classification 20 (1), 93–114.
Waller, L., Gotway, C., 2004. Applied Spatial Statistics for Public Health Data. John Wiley & Sons.
Whittle, P., 1954. On stationary processes in the plane. Biometrika 41 (3), 434–449.
Zhang, J., 2013. Analysis of Variance for Functional Data. CRC Press.
Zhang, L., Baladandayuthapani, V., Zhu, H., Baggerly, K.A., Majewski, T., Czerniak, B.A., Morris, J.S., 2016. Functional CAR models for large spatially correlated functional datasets. J. Amer. Statist. Assoc. 111 (514), 772–786.
Zoglat, A., 2008. Functional analysis of variance. Appl. Math. Sci. 2 (23), 1115–1129.