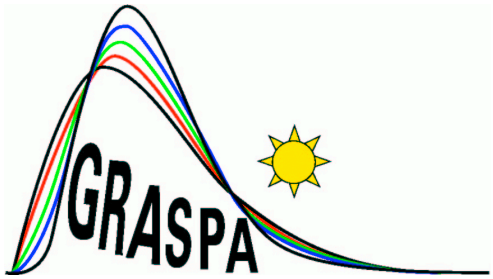
ISSN: 2037-7738
GRASPA Working Papers
[online]

Web Working Papers
by
The Italian Group of Environmental Statistics



Gruppo di Ricerca per le Applicazioni della Statistica
ai Problemi Ambientali

www.graspa.org

Recent developments of functional kriging

Ferdinand Ndongo, Alessandro Fassò,
Emilio Porcu, Jorge Mateu

GRASPA Working paper n. 50, November 2016

Recent developments of functional kriging

Ferdinand Ndongo, Alessandro Fassò, Emilio Porcu, Jorge Mateu

2nd November 2016

Abstract

The continuous improvement of technology have allowed the high frequency recording and storing of data measurements. Considering those data for each observation not any more as vector of points seen as finite set of points but as functions seen as infinite set of points and applying statistical techniques on it, is the core concept of functional data analysis. So several functional versions of a wide range of classical statistical tools are being proposed and developed since the rise of functional data analysis as an important branch of statistics for the analysis of curves. Therefore in spatial statistics different functional versions of kriging predictor have been proposed according to the type of spatial data observed. Here we are mainly interested in the case of geostatistical data observed as functions. This paper provides an overview of how classical geostatiscal tools have been extended in order to deal with spatial prediction of georeferenced data observed or approximated as functions or curves.

Introduction

The topic of functional data analysis (*fda*) had already been introduced by Rao (1958) but has become very popular by the works of J. Ramsay (1982), J. O. Ramsay and Dalzell (1991) and J. O. Ramsay and Silverman (2002). The particularity here is that each record observed as a finite set of measured values is considered as a single entity. This new concept of data brought a lot of research interests in the statistical community and now we can find various statistical books in literature dealing with this topic. For instance Ferraty and Vieu (2006), Shi and Choi (2011), Horváth and Kokoszka (2012), Zhang (2013), or recently Hsing and Eubank (2015). Nowadays many functional versions of classical statistical techniques are proposed in order to extend the statistical thinking to the data observed as functions. An example of such an

extension in control chart is done in Fassò, Toccu, and Magno (2016). For an overview of *fda* the most recent review in our knowledge is Wang, Chiou, and Mueller (2015) dealing not only with functional versions of simple statistical notions such as mean and covariance functions, but also with functional versions of some core techniques such as functional principal component analysis, functional linear regression and functional cluster analysis.

Following this trend, also geostatisticians developed the functional versions of classical geostatistical techniques in order to deal with spatially correlated data observed as functions in several areas. For instance, in climatology see Aguilera-Morillo, Durbán, and Aguilera (2016); in oceanology see Monestiez and Nerini (2008); in optimal design see Rasekhi, Jamshidi, and Rivaz (2014); in air pollution monitoring see Montero and Fernández-Avilés (2015) or Ignaccolo, Mateu, and Giraldo (2014); in studies on agricultural soils see Cortés-D, Camacho-Tamayo, and Giraldo (2016).

This paper is mainly focused in functional versions of classical geostatistical techniques and our intent is to provide an overview of the different geostatistical approaches dealing with spatial prediction of geostatistical data observed (or approximated) as functions.

In the first section, firstly we present the necessary basic knowledge in functional data analysis explaining how to build a smooth function from a vector of points; and then we will introduce the basic concepts of geostatistics from the notion of random field to the spatial prediction technique called kriging. The second section is dedicated to the main topic of this paper: spatial prediction of georeferenced functional data. In this section we will go through the different approaches for spatial prediction by separating it according to their underlying assumptions. In one side the stationarity of the process is assumed and in the other one it is removed. Finally we will conclude with some considerations.

1 Background

In functional geostatistics we put together the core concept of functional data analysis and the statistical methods of geostatistics with the scope to apply geostatistical techniques on functional data. This section is dedicated to the basic knowledge necessary to understand functional geostatistics. We will treat separately functional data analysis and geostatistics in order to give an overview of the core concepts of these two areas. For more details in *fda* see J. O. Ramsay, Hooker, and Graves (2009) and in geostatistics see Cressie (2015). The first subsection is dedicated to functional data analysis where we will present an overview of functional data. In the second subsection

we will briefly introduce the geostatistical technique for spatial prediction which is the same concept extended in its functional form by the functional geostatistics.

1.1 Functional data analysis

Assuming that a functional observation χ is observed as a finite set of measured values $y_j = 1, \ldots, M$, firstly we have to convert those M observed values into a function with values $\chi(u)$. We can estimate the function using parametric or non-parametric approach. The parametric approach is very limited because the function is supposed to be known while in many applications we do not know the function. That is why in many cases a more flexible non-parametric approach which only assumes smoothness is used. The conversion process involves smoothing since the function is supposed to be observed with error according to the model:

$$y_j = \chi(u_j) + e_j = \mathbf{c}'\phi(u) + e_j = \boldsymbol{\phi}(u_j)'\mathbf{c} + e_j \tag{1}$$

where the residuals e_j are statistically independent and have a normal distribution with mean 0 and constant variance.

The function is built as a linear combination of Q basis functions expressed as:

$$\chi(u) = \sum_{q=1}^{Q} c_q \phi_q(u) = \mathbf{c}'\phi(u) \tag{2}$$

and called a basis function expansion. The parameters $c_1, \ldots, c_Q$ are the coefficients of the expansion and ϕ in the last term of (2) is a vector of length Q containing the basis functions.

Using matrix notation, equation (1) is rewritten as

$$\mathbf{y} = \boldsymbol{\phi}\mathbf{c} + \mathbf{e} \tag{3}$$

where $\mathbf{y}$ is the vector of M observed points, $\boldsymbol{\phi}$ is a $M \times Q$ matrix containing the basis function values $\phi_q(u_j)$, and $\mathbf{e}$ contains the errors.

The estimated $\hat{\mathbf{c}}$ is then computed through smoothing by regression with ordinary least squares approach or with roughness penalty approach.

Using ordinary least squares approach, the estimated coefficients of splines are computed as the minimization of the sum of squared errors:

$$sse(\chi) = \sum_{j}^{M} \left[y_j - \chi(u_j)\right]^2. \tag{4}$$

Replacing (2) in (4) the least-squares estimation becomes

$$sse(\mathbf{c}) = \sum_{j=1}^{M} \left[y_j - \sum_{q=1}^{Q} c_q \phi_q(u_j) \right]^2 = \sum_{j=1}^{M} \left[y_j - \boldsymbol{\phi}(u_j)'\mathbf{c} \right] \tag{5}$$

and the least-squares estimate of the coefficient vector $\mathbf{c}$ is

$$\hat{\mathbf{c}} = (\boldsymbol{\phi}'\boldsymbol{\phi})^{-1}\boldsymbol{\phi}'\mathbf{y}. \tag{6}$$

Using a roughness penalty approach, we have to define a measure of roughness of the fitted curve, and then minimize a fitting criterion that trades off curve roughness against lack of data fit. A common measure of function's roughness is its integrated squared second derivative expressed as

$$PEN_2(\chi) = \int \left[D^2\chi(u) \right]^2 du. \tag{7}$$

The compound fitting criterion is defined adding some multiple of the roughness penalty to the error sum of squares

$$F(\mathbf{c}) = \sum_{j=1}^{M} \left[y_j - \chi(u_j) \right]^2 + \eta \int \left[D^2\chi(u) \right]^2 du \tag{8}$$

where $\chi(u) = \mathbf{c}'\phi(u)$. The smoothing parameter η specifies the emphasis on the second term penalizing curvature relative to goodness of fit quantified on the sum of squared residuals in the first term.

1.2 Geostatistics

Here we have to deal with geostatistical data defined as a type of spatial data with a continuous variation. That means the latent spatial process supposed having generated the data is indexed over a continuous space.

The observed value y_s of an environmental variable at spatial location s, s denoting its spatial coordinates in two dimensions, is regarded as realization of a random variable $\chi(s)$ with assumed distribution, mean, variance, and/or higher-order moments. The set of random variable for $s_i, i = 1, \ldots N$ constitutes a stochastic process

$$\{\chi(s), s \in D \subseteq \mathbb{R}^d\} \tag{9}$$

characterized by correlation in space and assumed properties.

The stochastic process (9) is usually assumed to be second order stationary. That means its mean and variance are constant over the space and its covariance function is only function of h the distance in space. A second order stationary stochastic process can be represent by the model

$$\chi(s) = \mu + \epsilon(s) \tag{10}$$

where μ is the mean of the process and $\epsilon(s)$ is a random error with zero mean and covariance

$$\mathbb{C}(h) = \mathbb{E}[\{\chi(s) - \mu\}\{\chi(s+h) - \mu\}] \tag{11}$$

depends on h the separation between samples in distance and direction and it is a function of h.

The stochastic process is said to be intrinsic stationary if the expected differences are zero

$$\mathbb{E}[\chi(s) - \chi(s+h)] = 0 \tag{12}$$

and the covariance of the residuals is replaced by the variance of the differences to measure spatial variations

$$\mathbb{V}[\chi(s) - \chi(s+h)] = \mathbb{E}\{[\chi(s) - \chi(s+h)]^2\} = 2\gamma(h). \tag{13}$$

$\gamma(h)$ is the semivariance at lag h, and as a function of h it is the variogram.

For second-order stationary processes the covariance function and variogram are equivalent

$$\gamma(h) = C(0) - C(h),$$

where $C(0) = \sigma^2$ is the variance of the process.

Geostatistics is based on the theory of random spatial process and provides unbiased spatial prediction of environmental variables at unsampled site taking into account spatial correlation. Using a suitable model to describe the spatial covariance, expressed as a variogram, spatial prediction is done by kriging which can be seen as a weighted averaging technique.

The kriging equations are obtained by imposing on the predictor the classical conditions of unbiasedness and minimum variance. Kriging is a best linear unbiased predictor ($BLUP$).

Suppose that the functional data χ_{s_i}, $i = 1, \ldots, N$ are realizations of the stochastic process (9), the kriging predicts χ at any new point s_0, by

$$\hat{\chi}_{s_0} = \sum_{i=1}^{N} \lambda_i \chi_{s_i} \tag{14}$$

where χ_{s_i} are the data observed at the N spatial locations in the neighbourhood of s_0 and λ_i are the kriging weights obtained by imposing the classical conditions of unbiasedness and minimum variance

$$\sum_{i=1}^{N} \lambda_i = 1 \tag{15}$$

$$\mathbb{E}(\hat{\chi}_{s_0} - \chi_{s_0}) = 0. \tag{16}$$

The prediction variance is given by

$$\mathbb{V}(\hat{\chi}_{s_0}) = \mathbb{E}\left[(\hat{\chi}_{s_0} - \chi_{s_0})^2\right]. \tag{17}$$

2 Spatial prediction in functional geostatistics

This section is dedicated to how geostatisticians deals with georeferenced spatially correlated data observed as functions in order to make spatial prediction of a function in spatial locations where the function has not been observed.

Following Cressie (1993) for the definition of spatial processes and Ferraty and Vieu (2006) for that of functional random variables, Delicado, Giraldo, Comas, and Mateu (2010) define a spatial functional process as

$$\{\boldsymbol{\chi}_s,\ s \in D \subseteq \mathbb{R}^d\} \tag{18}$$

where s is a generic spatial location in the d-dimensional Euclidean space, the set D is a fixed subset of $\mathbb{R}^d$ with positive volume and N points $s_1, \ldots, s_N$ in D are chosen to observe the random functions $\chi_{s_i}, i = 1, \ldots, N$. Typically χ_s is a real function from $U = [a, b] \subseteq \mathbb{R}$ to $\mathbb{R}$. For every fixed $u \in U$, $\{\boldsymbol{\chi}_s(u), s \in D \subseteq \mathbb{R}^d\}$ is a scalar-valued random process defined in D with values in $\mathbb{R}$. For fixed u and s, $\chi_s(u)$ is a random variable.

The functional random process (18) is second order stationary and isotropic if its mean and variance are constant, and its covariance function depends only on the distance between sampling points. Formally, that is:

$$\mathbb{E}[\chi_s(u)] = m(u),\ \text{for all } u \in U,\ s \in D \tag{19}$$

$$\mathbb{V}[\chi_s(u)] = \sigma^2(u),\ \text{for all } u \in U,\ s \in D \tag{20}$$

$$\mathbb{C}[\chi_{s_i}(u) - \chi_{s_{i'}}(u')] = \mathbb{C}(h; u, u'),\ s_i, s_{i'} \in D,\ u, u' \in U,\ h = \|s_i - s_{i'}\|, \tag{21}$$

if $u = u'$, $\mathbb{C}[\chi_{s_i}(u) - \chi_{s_j}(u)] = \mathbb{C}(h; u)$

$$\frac{1}{2}\mathbb{V}[\chi_{s_i}(u) - \chi_{s_{i'}}(u')] = \gamma(h; u, u'), s_i, s_{i'} \in D, u, u' \in U, h = \|s_i - s_{i'}\|,$$

if $u = u'$,

$$\frac{1}{2}\mathbb{V}[\chi_{s_i}(u) - \chi_{s_{i'}}(u)] = \gamma(h; u) \tag{22}$$

which is called variogram.

In order to predict a function in an unsampled site, two cases are distinguished. The first one assumes the stochastic process (18) to be second order stationary and isotropic, while the second one supposes this assumption to be violated. The two following subsections present how spatial prediction of a function in an unsampled site is performed, firstly in the case of a stationary and isotropic stochastic process, and secondly when it is not.

2.1 Second order stationary and isotropic case

The different spatial predictor found in literature dealing with spatial prediction of functions when the functional spatial process is assumed to be second order stationary and isotropic are: the curve kriging predictor (CKP), the ordinary kriging for functional data ($OKFD$), the pointwise kriging predictor for functional data ($PKFD$) and the functional kriging for total model($FKTM$).

2.1.1 *Curve Kriging Predictor* (*CKP*)

Predict a function in an unsampled location is not a recent topic. Goulard and Voltz (1993) face the problem to predict continuous changes of curves in space considering a set of curves $\{\chi_{s_i},\ i = 1, \ldots, N\}$. The authors assume some restrictive assumptions: (i) the curves are only are known only up to a finite set of small number of points $\{u_j\}$, $j = 1, \ldots, M$; (ii) a parametric model with small number of parameters is known and fitted for the reconstruction of the curve in space. For predicting directly the curves, they propose the *Curve Kriging Predictor* (CKP) as the best linear unbiased predictor for χ_{s_0} by:

$$\hat{\chi}_{s_0} = \sum_{i=1}^{N} \lambda_i \chi_{s_i}(u),\ u \in U = [a, b],\ \lambda_1, \ldots, \lambda_N \in \mathbb{R} \tag{23}$$

where the coefficients λ_i are such that:

$$\mathbb{E}(\hat{\chi}_{s_0} - \chi_{s_0}) = 0 \tag{24}$$

and

$$\mathbb{E}\left\{\int_a^b [\hat{\chi}_{s_0}(u) - \chi_{s_0}(u)]^2 du\right\} = \int_a^b \mathbb{V}[\hat{\chi}_{s_0}(u) - \chi_{s_0}(u)] du \tag{25}$$

is minimized. Thus the optimization problem to be solved is

$$\min_{\lambda_1,\ldots,\lambda_N} \int_a^b \mathbb{V}[\hat{\chi}_{s_0}(u) - \chi_{s_0}(u)]du, \ s.t. \ \sum_{i=1}^{N} \lambda_i = 1 \tag{26}$$

where $\sum_{i=1}^{N} \lambda_i = 1$ is the unbiasedness constraint.

The authors propose to use a set parametric models $\chi(.;\theta)$ and then estimate to obtain $\chi(.;\hat{\theta}_{s_i})$ as an approximation of the function χ_{s_i}. Then the *CKP* can be written as:

$$\hat{\chi}_{s_0} = \sum_{i=1}^{N} \lambda_i \ \chi(.;\hat{\theta}_{s_i}). \tag{27}$$

In order to overcome the restrictive assumptions of Goulard and Voltz (1993), some spatial predictors have been proposed such as ordinary kriging for function-value spatial data (*OKFD*), pointwise kriging predictor for functional data (*PKFD*) and functional kriging total model (*FKTM*). Many of them are extensions to the functional case of classical geostatistical predictors.

2.1.2 *Ordinary kriging for function-value spatial data (OKFD)*

Giraldo, Delicado, and Mateu (2011b) extend the *CKP* with a non-parametric approach through B-spline smoothing and choosing the smoothing parameter by functional cross-validation. The family of linear predictors of (23) is defined and the resulting predicted curve is then a linear combination of observed curves. The λ's in predictor give the weight of the curves around the unsampled location such that curves from locations closer to the prediction points (unsampled locations) will have heavier weights than the other ones far way. The kriging predictor is given by the solution of the optimization problem (26). In functional cross-validation: each functional data location is removed from the data set and a function is predicted at this location using a functional kriging predictor based on the remaining smoothed functions; then the sum of square errors is computed as

$$SSE_{FCV} = \sum_{i=1}^{N} SSE_{FCV}(i) = \sum_{i=1}^{N} \sum_{j=1}^{M} [\hat{\chi}_{s_i}^{(i)}(u_j) - \chi_{s_i}^{(i)}(u_j)]^2 \tag{28}$$

where $\hat{\chi}_{s_i}^{(i)}(u_j)$ is the functional kriging prediction on s_i evaluated at u_j, $j = 1, \ldots, M$, by leaving the site s_i temporarily out of the sample. The involved smoothing parameters are chosen by minimization of SSE_{FCV}.

2.1.3 *Pointwise linear Kriging predictor for Functional Data (PKFD)*

Giraldo, Delicado, and Mateu (2010) propose the pointwise linear predictor for functional data (*PKFD*) where both curves and weight are functions and are expanded in terms of a set of basis functions. Estimating the coefficients of these basis functions for each functional parameter and for each functional observation is performed through the *PKFD* for $\chi_{s_0}(u)$, $u \in U$, define as

$$\hat{\chi}_{s_0}(u) = \sum_{i=1}^{N} \lambda_i(u)\chi_{s_i}(u), \ i = 1, \ldots, N. \tag{29}$$

Therefore, in order to find the best linear unbiased predictor, the N functional parameters λ_i are given by the solution of the following optimization problem

$$\min_{\lambda_{1(.)},\ldots,\lambda_n(.)} \int_a^b \mathbb{V}[\hat{\chi}_{s_0}(u) - \chi_{s_0}(u)]du, \ s.t. \ \sum_{i=1}^{N} \lambda_i(u) = 1, \text{ for all } u \in [a, b] \tag{30}$$

Here the coefficients $\lambda_i(.)$ are functions that have to be determined to yield the best linear unbiased predictor. The optimization problem is solved by fitting non-parametrically the observed functions $\chi_{s_i}(u)$, as well as the parameter functions $\lambda_i(.)$. It is assumed that these functions can be expressed in terms of Q basis functions $\phi_1(u), \ldots, \phi_Q(u)$ as

$$\boldsymbol{\chi}_{s_i}(u) = \sum_{q=1}^{Q} c_{iq}\phi_q(u) = \mathbf{c}_i'\boldsymbol{\phi}(u) \tag{31}$$

$$\lambda_i(u) = \sum_{q=1}^{Q} b_{iq}\phi_q(u) = \mathbf{b}_i'\boldsymbol{\phi}(u) \tag{32}$$

with $i = 1, \ldots, N$. The choice of Q, the number of basis functions in the expansions, can be solved by functional cross-validation (28) or using standard methods in non-parametric regression for the choice of the smoothing parameter.

2.1.4 *Functional kriging total model (FKTM)*

Giraldo, Delicado, and Mateu (2011a) propose the functional kriging for total model (*FKTM*) which is an adaptation of functional linear model for

functional responses (total model) introduced by *J. O. Ramsay (2006)* where λ_i is be defined in $[a,b]\times[a,b]$. Then, the predictor of χ_{s_0} is

$$\hat{\chi}_{s_0}(v)=\sum_{i=1}^{N}\int_a^b \lambda_i(u,\nu)\ \chi_{s_i}(u)dv \tag{33}$$

with $u\in[a,b]$, $\lambda_i\in[a,b]\times[a,b]\longmapsto\mathbb{R}$, $i=1,\ldots,N$. The functional parameter $\lambda_i(u,\nu)$ determines the impact of the ith observed function at time u on an unobserved function at time v.

Monestiez and Nerini (2008) used a truncated expansions in a functional basis, followed by cokriging over the basis coefficients to define the predictor (33). Orthogonality is the main methodology difference between the two approaches. In *Monestiez and Nerini (2008)* the orthonormal basis functions defining $\chi_{s_i}(u)$ are expressed in terms of a linear combination of Q known basis function as in (31).

The expansion of $\lambda_i(u,\nu)$ into the *ϕ-basis* is expressed as:

$$\lambda_i(u,\nu)=\sum_{p=1}^{Q}\sum_{q=1}^{Q} b_{pq}^i\phi_p(s)\phi_q(u)=\phi'(s)B_i\phi(u).$$

resulting in the predictor

$$\hat{\chi}_{s_0}(v)=\sum_{i=1}^{N}\sum_{p,q=1}^{Q} b_{pq}^i c_p(s_i)\phi_q(u)$$

and the cokriging estimator of c_0 at location s_0 is defined as:

$$\hat{c}_0=\sum_{i=1}^{N} B_i' c_i.$$

Condition of unbiasedness $\mathbb{E}(\hat{c}_0-c_0)=0$ is satisfied by choosing weights that fulfil the constraints

$$\sum_{i=1}^{N} b_{pq}^i=\begin{cases} 1 \text{ if } c_p=c_q \\ 0 \text{ otherwise}\end{cases}$$

2.2 Non stationary case

The case when the assumption of stationarity of the functional spatial process is violated is treated considering the non-stationarity due to the non-constant spatial mean variation. This problem is solved using an extension to the functional form of residual kriging, universal kriging and kriging with external drift.

2.2.1 *Residual kriging (RK)*

Reyes, Giraldo, and Mateu (2010) the authors propose a methodology to extend kriging predictors for functional data to the case where the mean function is not constant (19) through the region of interest. They consider an approach based on the classical residual kriging based in three steps: first, they perform a functional regression model in order to de-trend the mean, secondly they apply kriging methods for functional data to the regression residuals for doing prediction of a residual curve on a non-data location, finally the prediction curve is obtained as the sum of the trend and the residual prediction.

The authors first estimate a functional regression model (*FRM*) with functional response and scalar covariates using the model:

$$\chi_{s_i}(u) = \alpha(u) + \beta_1(u)x_i + \beta_2(u)y_i + \epsilon_i(u) \tag{34}$$

where $\chi_{s_i}(u)$, $i = 1, \ldots, N$ are the functions at visited locations, (x_i, y_i) are geographical coordinates (scalar covariates), $\alpha(u), \beta_1(u), \beta_2(u)$ are the functional parameters of interest and $\epsilon(u)$ is a white noise for each $u \in U = [a, b]$.

Once estimated the regression model, in a second step we obtain the residuals

$$\begin{aligned} \epsilon_{s_i}(u) &= \chi_{s_i}(u) - \hat{\mu}_{s_i}(u) \\ &= \chi_{s_i}(u) - (\hat{\alpha}(u) + \hat{\beta}_1(u)x_i + \hat{\beta}_2(u)y_i) \end{aligned} \tag{35}$$

They then compute the predicted residual curves as residual versions of *OKFD*, *PKFD* and *FKTM* given by:

$$\hat{\epsilon}_{s_0}(u) = \sum_{i=1}^{N} \lambda_i \epsilon_{s_i}(u) \tag{36}$$

$$\hat{\epsilon}_{s_0}(u) = \sum_{i=1}^{N} \lambda_i(u) \epsilon_{s_i}(u) \tag{37}$$

$$\hat{\epsilon}_{s_0}(\nu) = \sum_{i=1}^{N} \int_a^b \lambda_i(u, \nu) \epsilon_{s_i}(u) dv \tag{38}$$

where $\epsilon_i(u)$, $i = 1, \ldots, N$ are the residual curves obtained from the estimated *FRM*.

And finally the predicted curve at the location s_0 is computed with the so called residual kriging predictor for functional data expressed as:

$$\hat{\chi}_{s_0}(u) = \hat{\mu}_{s_0}(u) + \hat{\epsilon}_0(u) \tag{39}$$

2.2.2 *Functional Universal Kriging Predictor (FUKP)*

Caballero, Giraldo, and Mateu (2013) propose a solution to the problem of spatial prediction of non-stationary functional data by extending the classical universal kriging predictor for univariate data to the context of functional data. They specify that the non stationarity is related to the non constant mean (19) and propose the *Functional Universal Kriging Predictor (FUKP)*. The functional process is defined as:

$$\begin{aligned} \chi_s(u) &= \mu(s,u) + \epsilon_s(u) \\ &= \sum_{p=1}^{P} \beta_p(u) f_p(s) + \epsilon_s(u) \end{aligned} \tag{40}$$

for all $s \in D$. In addition $\mathbb{E}[\epsilon_s(u)] = 0$ and $\mathbb{V}[\epsilon_s(u)] = \sigma^2(u)$.

Consequently, the functional vector $\chi_s(u)$ follows a spatial functional regression model of the form:

$$\chi_s(u) = X(u)\beta(u) + \epsilon_s(u)$$

They propose the functional universal kriging predictor ($FUKP$) of $\chi_{s_0}(u)$ as $\hat{\chi}_{s_0}(u) = \boldsymbol{\lambda}^T \boldsymbol{\chi}_s(u)$ where the components of the weighting vector k are real numbers and are chosen in such a way that $FUKP$ is an unbiased predictor of minimum variance. The $\boldsymbol{\lambda}'s$ are obtained by minimizing the variance of the prediction error.

2.2.3 Functional kriging with external drift (*FKED*)

Ignaccolo et al. (2014) propose the functional predictors ($OKFD$, $PKFD$ and $FKTM$) based on functional residuals with the inclusion of exogenous variables (both scalar and functional) in a functional drift into the model. They propose the functional kriging with external drift ($FKED$) which is an extension of the so-called kriging with external drift - or regression kriging- to the case of functional data.

It is considered to have a spatial functional process (18) that can be non-stationary and whose elements are supposed to follow the model

$$\chi_s = \mu_s + \epsilon_s \tag{41}$$

the term μ_s is interpreted as a drift describing a spatial trend while ϵ_s represents a residual random field that is zero-mean, second-order stationary and isotropic, so that:

$$\mathbb{E}(\chi_s) = \mu_s, s \in D$$

$$\mathbb{E}(\epsilon_s) = 0, s \in D$$

$$\mathbb{C}(\epsilon_{s_i}, \epsilon_{s_{i'}}) = \mathbb{C}(u)$$

At a generic $s_i, i = 1, \dots N$, and at point u model (41) can be rewritten as a functional concurrent linear model introduced in J. O. Ramsay (2006)

$$\chi_{s_i}(u) = \mu_{s_i}(u) + \epsilon_{s_i}(u) \tag{42}$$

with the drift

$$\mu_{s_i}(u) = \alpha(u) + \sum_p a_p(u) X_{p,i} + \sum_q \beta_q(u) f_{q,i}(u) \tag{43}$$

where $\alpha(u)$ is a functional intercept, $X_{p,i}$ is the p_{th} scalar covariate and $f_{q,i}(u)$ is the q_{th} functional covariate at site s_i, $a_p(u)$ and $\beta_q(u)$ are the covariate coefficients. In the drift, both scalare and functional covariates are included, and the coefficients $a_p(u)$ and $\beta_q(u)$ are also of a functional nature, allowing to estimate nonlinear effects of a covariate.

In order to predict a curve at an unmonitored site s_0, taking into account exogenous variables in the drift, they propose a three step procedure: (i) at the first step the functional regression model with functional response and scalar and functional covariates is fitted by generalized cross validation (GCV) in order to estimate the drift coefficient and obtain functional residuals.

$$\begin{aligned} \epsilon_{s_i}(u) &= \chi_{s_i}(u) - \mu_{s_i}(u) \\ &= \chi_{s_i}(u) - [\alpha(u) + \sum_p a_p(u) X_{p,i} + \sum_q \beta_q(u) f_{q,i}(u)] \end{aligned} \tag{44}$$

(ii)At the second step the residual curve prediction at the unmonitored site s_0 can be obtained by ordinary kriging for functional data Giraldo et al. (2011b) according to which

$$\hat{\epsilon}_{s_0}(u) = \sum_{i=1}^{N} \lambda_i \epsilon_{s_i}(u) \tag{45}$$

where the kriging coefficients $\lambda_i \in \mathbb{R}$ are constant, so that the predicted curve is a linear combination of data residual curves. The weights λ_i are determined as the solution of a linear system written to solve the optimization problem

$$\min_{\lambda_1,\dots,\lambda_n} \int_a^b \mathbb{V}[\hat{\epsilon}_{s_0}(u) - \epsilon_{s_0}(u)] du, s.t. \sum_{i=1}^{N} \lambda_i(u) = 1, \tag{46}$$

The authors also present the continuous time varying kriging for functional data Giraldo et al. (2010) always based on functional residuals 45 where the curves and functional parameters are represented in terms of basis functions and also functional kriging total model

$$\hat{\epsilon}_{s_0}(v) = \sum_{i=1}^{N} \int_a^b \lambda_i(u,v)\epsilon_{s_i}(u)dv \tag{47}$$

(iii) at the third step compute the predicted curve at unmonitored site by adding - as in the classical regression kriging- , the two terms, that is :

$$\hat{\chi}_{s_0}(u) = \hat{u}_{s_0}(u) + \hat{\epsilon}_{s_0}(u)$$

where

$$\hat{\mu}_{s_0}(u) = \hat{\alpha}(u) + \sum_p \hat{a}_p(u)X_{p,0} + \sum_q \hat{\beta}_q(u)f_{q,0}(u)]$$

depends on the covariate values $X_{p,0}$ and $f_{q,0}$ at site s_0.

3 Conclusion

Through this paper we provide an overview of the state of the art of a geostatistical technique such as kriging applied on geostatistical data observed as functions. From a first attempt to predict a curve at an unsampled site using a parametric approach, we move into non-parametric approaches treating first the case keeping the assumptions of stationarity and after the non stationarity case. In both cases, it was highlighted the fact that the functional kriging was an extension of the existing classical kriging to the functional data. It came out that the non stationary case is treated considering only the mean non stationary side. The paper being limited to the classical framework, it can be extended in the future taking into account the approach of linear operators in Hilbert space and the non stationarity in (co)variance.

References

Aguilera-Morillo, M. C., Durbán, M., & Aguilera, A. M. (2016). Prediction of functional data with spatial dependence: a penalized approach. *Stochastic Environmental Research and Risk Assessment*, 1–16.

Caballero, W., Giraldo, R., & Mateu, J. (2013). A universal kriging approach for spatial functional data. *Stochastic environmental research and risk assessment*, *27*(7), 1553–1563.

Cortés-D, D. L., Camacho-Tamayo, J. H., & Giraldo, R. (2016). Spatial prediction of soil penetration resistance using functional geostatistics. *Scientia Agricola*, *73*(5), 455–461.

Cressie, N. (1993). Statistics for spatial data: Wiley series in probability and statistics. *Wiley-Interscience, New York*, *15*, 105–209.

Cressie, N. (2015). *Statistics for spatial data*. John Wiley & Sons.

Delicado, P., Giraldo, R., Comas, C., & Mateu, J. (2010). Statistics for spatial functional data: some recent contributions. *Environmetrics*, *21*(3-4), 224–239.

Fassò, A., Toccu, M., & Magno, M. (2016). Functional control charts and health monitoring of steam sterilizers. *Quality and Reliability Engineering International*.

Ferraty, F., & Vieu, P. (2006). *Nonparametric functional data analysis: theory and practice*. Springer Science & Business Media.

Giraldo, R., Delicado, P., & Mateu, J. (2010). Continuous time-varying kriging for spatial prediction of functional data: an environmental application. *Journal of Agricultural, Biological, and Environmental Statistics*, *15*(1), 66–82.

Giraldo, R., Delicado, P., & Mateu, J. (2011a). Geostatistics with infinite dimensional data: a generalization of cokriging and multivariable spatial prediction. *Matemática: ICM–ESPOL*, *9*, 16–21.

Giraldo, R., Delicado, P., & Mateu, J. (2011b). Ordinary kriging for function-valued spatial data. *Environmental and Ecological Statistics*, *18*(3), 411–426.

Goulard, M., & Voltz, M. (1993). Geostatistical interpolation of curves: a case study in soil science. In *Geostatistics tróia'92* (pp. 805–816). Springer.

Horváth, L., & Kokoszka, P. (2012). *Inference for functional data with applications* (Vol. 200). Springer Science & Business Media.

Hsing, T., & Eubank, R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*. John Wiley & Sons.

Ignaccolo, R., Mateu, J., & Giraldo, R. (2014). Kriging with external drift

for functional data for air quality monitoring. *Stochastic environmental research and risk assessment*, *28*(5), 1171–1186.

Monestiez, P., & Nerini, D. (2008). A cokriging method for spatial functional data with applications in oceanology. In *Functional and operatorial statistics* (pp. 237–242). Springer.

Montero, J.-M., & Fernández-Avilés, G. (2015). Functional kriging prediction of pollution series: The geostatistical alternative for spatially-fixed data. *Estudios Econ. Apl*, *33*, 145–174.

Ramsay, J. (1982). When the data are functions. *Psychometrika*, *47*(4), 379–396.

Ramsay, J. O. (2006). *Functional data analysis*. Wiley Online Library.

Ramsay, J. O., & Dalzell, C. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 539–572.

Ramsay, J. O., Hooker, G., & Graves, S. (2009). *Functional data analysis with r and matlab*. Springer Science & Business Media.

Ramsay, J. O., & Silverman, B. W. (2002). *Applied functional data analysis: methods and case studies* (Vol. 77). Citeseer.

Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics*, *14*(1), 1–17.

Rasekhi, M., Jamshidi, B., & Rivaz, F. (2014). Optimal location design for prediction of spatial correlated environmental functional data. *Journal of Modern Applied Statistical Methods*, *13*(2), 26.

Reyes, A., Giraldo, R., & Mateu, J. (2010). Residual kriging for functional data. application to the spatial prediction of salinity curves.

Shi, J. Q., & Choi, T. (2011). *Gaussian process regression analysis for functional data*. CRC Press.

Wang, J.-L., Chiou, J.-M., & Mueller, H.-G. (2015). Review of functional data analysis. *arXiv preprint arXiv:1507.05135*.

Zhang, J.-T. (2013). *Analysis of variance for functional data*. CRC Press.