Extending Functional kriging to a multivariate context

Francesca Di Salvo[1], Mariantonietta Ruggieri[2], Antonella Plaia[2]

Abstract

Environmental data usually have a spatio-temporal structure; pollutant concentrations, for example, are recorded along time and space. Generalized Additive Models (GAMs) represent a suitable tool to model spatial and/or temporal trends of this kind of data, that can be treated as functional, although they are collected as discrete observations. Frequently, the attention is focused on the prediction of a single pollutant at an unmonitored site and, at this aim, we extend kriging for functional data to a multivariate context by exploiting the correlation with the other pollutants. In particular, we propose two procedures: the first one (FKED) combines the regression of a variable (pollutant), of primary interest on the other variables, with functional kriging of the regression residuals; the second one (FCK) is based on linear unbiased prediction of spatially correlated multivariate random processes. The performance of the two proposed procedures is assessed by cross validation; data recorded during a year (2011) from the monitoring network of the state of California (USA) are considered.

Keywords: FDA, GAM, FUNCTIONAL KRIGING, KED

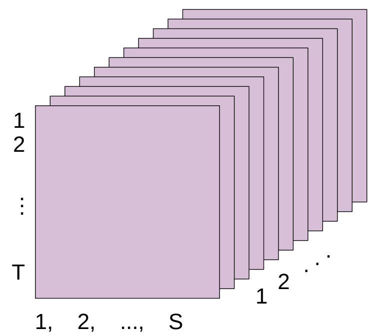
1. Introduction

Environmental data are usually multivariate spatio-temporal data, that can be organized in three way arrays where two dimension domains (both structured) are time and space (Fig. 1).

Let us consider, as motivating example, PM_{10} and the main daily gaseous pollutant concentrations (CO, NO_2, O_3, SO_2) recorded during a year (2011) by the monitoring

[1]Department of Agricultural Sciences, Food and Forestry, University of Palermo, Italy.
[2]Department of Economics, Business and Statistics, University of Palermo, Italy.
Corresponding author: mariantonietta.ruggieri@unipa.it

Figure 1: Three dimensional array for space-time data



network of the State of California: we may recognize time series along one of the dimensions (Fig. 2) and spatial series along another (Fig. 3).

Functional Data Analysis (FDA) [Ramsay and Silverman, 2005] provides a suitable framework when large amount of data are recorded over time and/or space and Generalized Additive Models (GAMs) [Hastie and Tibshirani, 1990] are a useful tool for modelling and describing temporal and/or spatial trends of pollutant concentrations.

Over the last years there has been an increasing interest within the statistical community on FDA and, recently, attention has been focused on Spatial Functional Statistics, considering spatially dependent functional data [Delicado *et al.*, 2010]. In this context, one of the main issues is the spatial prediction. The Functional kriging [Giraldo *et al.*, 2011b, Nerini *et al.*, 2010] extends the ordinary kriging to the functional context, which allows to predict a curve at an unmonitored site by exploiting the curves related to other monitored sites. Giraldo *et al.* [2011b] present a methodology to make spatial predictions at non-data locations when the data values are functions. In particular, they propose both an estimator of the spatial correlation and a functional kriging predictor. Nerini *et al.* [2010] propose to generalize the method of kriging when data are spatially sampled curves and construct a spatial functional linear model including spatial dependencies between curves. Giraldo *et al.* [2010] present an approach

Figure 2: Time series of five pollutants from the monitoring network of California, 2011

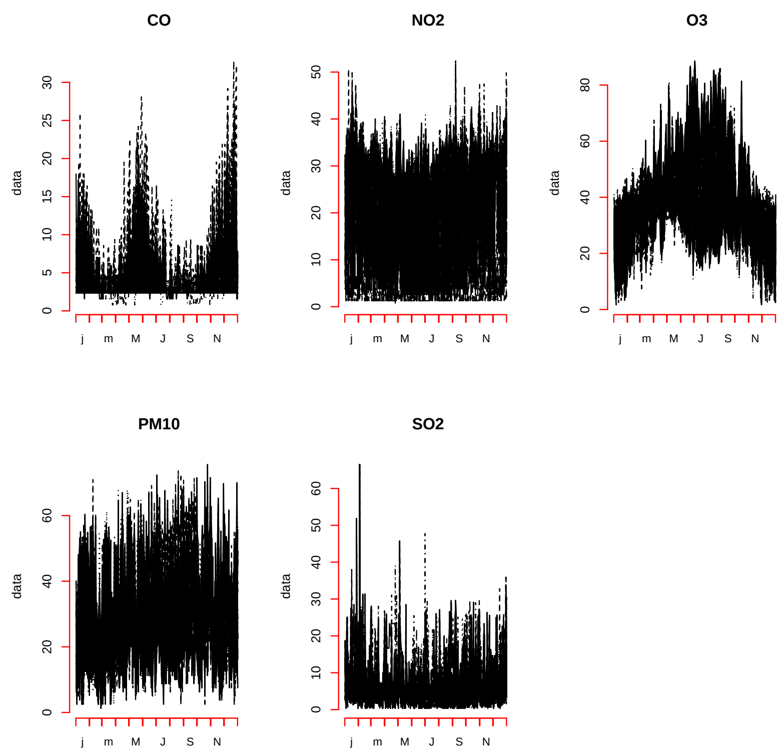
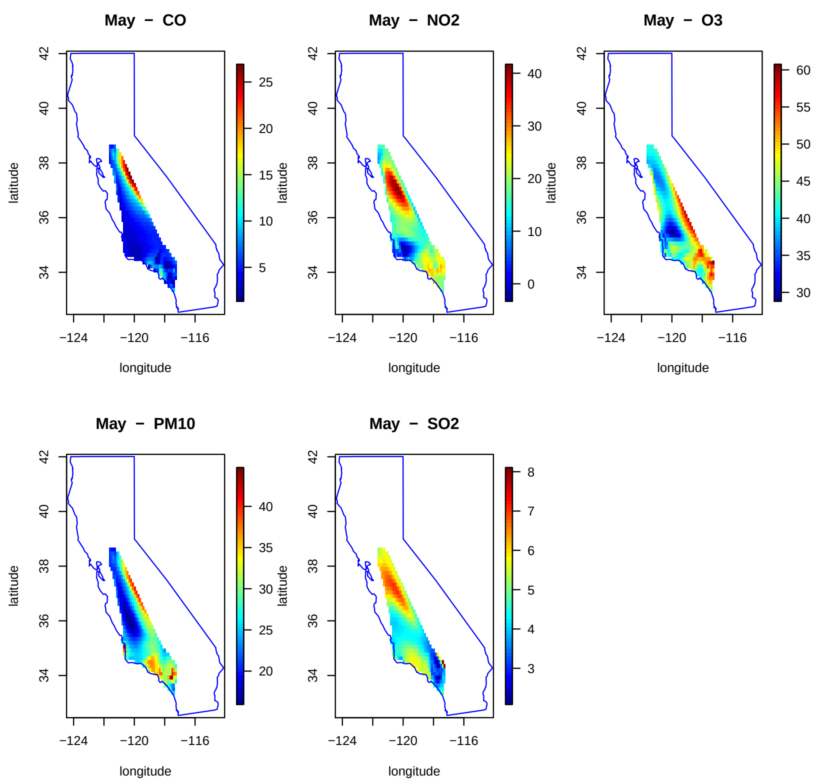


Figure 3: Spatial interpolation of 5 pollutants on May 2011



for spatial prediction, based on the functional linear point-wise model, adapted to the case of spatially correlated curves. Giraldo *et al.* [2011a] extend cokriging analysis and multivariable spatial prediction to the case where the observations at each sampling location consist of samples of random functions, that is they extend two classical multivariable geostatistical methods to the functional context. Giraldo [2014] gives an overview of cokriging analysis and multivariable spatial prediction when the observations at each sampling location consist of samples of random functions, extending classical cokriging multivariable geostatistical methods to the functional context. Suitable methodologies have also been developed in the more realistic cases of absence of stationarity [Caballero *et al.*, 2013], that is for processes with non-constant mean function (non-stationary functional data). In order to take into account exogenous variables, such as meteorological information, Kriging with External Drift (KED), or regression kriging, is extended to the functional data, involving functional modelling for the trend (drift) and spatial interpolation of functional residuals [Ignaccolo *et al.*, 2014].

In this paper we want to consider a recurrent case, when more than a single variable (pollutants, for example) is recorded and a variable has to be predicted in a site where a) no other variables are recorded; b) other variables are recorded. Actually, even if we are interested in predicting a single variable, in an unmonitored site, exploiting its correlation with the other variables can improve the estimation. In particular, in this paper, we want to focus on case a).

The prediction of a geophysical quantity based on observations at nearby locations of the same quantity and other related variables, so-called covariables, is often of interest and, in this paper, we explore two alternative ways of including the influence of the covariates in prediction. The classical approach in the geostatistical framework is cokriging and in the functional context the proposed approaches deal with univariate stochastic process, under stationary assumptions (Giraldo [2009], Delicado *et al.* [2010], [Menafoglio *et al.*, 2014]) and non stationary assumptions ([Menafoglio *et al.*, 2013], [Ignaccolo *et al.*, 2014]). In practical and methodological considerations on kriging of functional data the problem of the high dimensionality occurs. In this context, our first proposal, the Functional Kriging with External Drift (FKED), combines the regression of a variable of primary interest on the other variables, with functional kriging of the regression residuals; alternatively, a second procedure, the Functional Cokriging (FCK), is based on linear unbiased prediction of spatially correlated multivariate random processes.

The paper is organized as follows: Sections 2 describes the state of art and 3 introduces the proposed methodology; Section 4 presents the data and the performance of the spatial prediction is assessed; Section 5 reports the conclusions and further developments.

2. GAMs and Functional kriging

2.1. P-spline smoothing

In the geostatistical functional data framework, considering a site $\mathbf{s} \in D \subseteq R^2$, $Y^p_{\mathbf{s}t}$, $p = 1, ..., P$, is a realization of a set of p curves, functions of time $t \in T \subseteq R$:

$$\underbrace{Y^p_{\mathbf{s}t}}_{data} = \underbrace{X^p(\mathbf{s}, t)}_{signal} + \underbrace{\varepsilon^p_{\mathbf{s}t}}_{noise}; \tag{1}$$

the set $X^p(\mathbf{s}, t)$ is a non-stationary functional random field and the set $\varepsilon^p_{\mathbf{s}t}$ is stationary Gaussian process with a zero first moment and isotropic spherical covariance functions. both the proposed procedures fit GAMs to spatio-temporal data via the penalized likelihood approach, assuming separable structures in the data. In a two-step estimation procedure we assume separable spatio-temporal structures, i.e. the spatial correlation structure does not change over time, The following underlying functional form is provided:

$$X^p(\mathbf{s}, t) = Z^p(\mathbf{s}) + \chi^p_{\mathbf{s}}(t). \tag{2}$$

Throughout this paper, the process $Z^p(\mathbf{s})$ has a non constant mean and describes the main spatial effects, that we model through penalized splines in the GAMs framework [Hastie and Tibshirani, 1990]. These models assume that the mean of the response variable depends on an additive predictor through a link function. The space-dependent function $Z^p(\mathbf{s})$ is expanded in terms of basis matrix $\mathbf{B} = (\mathbf{B_1}(\mathbf{s}), ..., \mathbf{B_2}(\mathbf{s}), ..., \mathbf{B_k}(\mathbf{s}))$ and coefficients $\mathbf{u}^p = (u_1^p, u_2^p, \ldots, u_k^p)$:

$$Z^p(\mathbf{s}) = \mathbf{B}(\mathbf{s})\,\mathbf{u}^p. \tag{3}$$

The functions are estimated by minimizing the Penalized Residual Sum of Squares, for each dimension $p = 1, ..., P$:

$$PENSSE_\lambda(\mathbf{y}) = ||\mathbf{y} - \mathbf{B}\mathbf{u}^p||^2 + \mathbf{H}.$$

Depending on the data structure, the model basis $\mathbf{B}$ can be defined as Kronecker product (data in a regular grid) or box product (irregularly spaced data); this last choice is outlined in this paper, as we deal with irregularly spaced data. The box product, or rows-wise Kronecker product, denoted by $\square$ symbol, was defined in [Eilers and Marx, 1996] and proposed by Lee and Durban [2013] in multidimensional smoothing:

$$\mathbf{B} = \mathbf{B}^2 \square \mathbf{B}^1 = (\mathbf{B}^2 \otimes \mathbf{1}'_{k_1}) \odot (\mathbf{1}'_{k_2} \otimes \mathbf{B}^1), \tag{4}$$

where $\mathbf{B}^1 = \left(\mathbf{B}^1{}_1(s_1), ..., \mathbf{B}^1_{k_1}(s_1)\right)$ and $\mathbf{B}^2 = \left(\mathbf{B}^2{}_1(s_2), ..., \mathbf{B}^2_{k_2}(s_2)\right)$, are the $(n \times k_1)$ and $(n \times k_2)$ marginal B-spline bases for the geographical coordinates.

The penalty matrix:

$$\mathbf{H} = \lambda_1 \mathbf{I}_{k_2} \otimes \mathbf{D}_1{}'\mathbf{D}_1 + \lambda_2 \mathbf{D}_2{}'\mathbf{D}_2 \otimes \mathbf{I}_{k_1}, \tag{5}$$

allows for anisotropic smoothing structures, λ_1 and λ_2 being the smoothing parameters; $\mathbf{I}_{k_1}$ and $\mathbf{I}_{k_2}$ are the identity matrices of order k_1 and k_2, respectively; $\mathbf{D}_1$ and $\mathbf{D}_2$ are second-order difference matrices of order k_1 and k_2, respectively. The values of λ, can be readily estimated by means of criterions as AIC, BIC or Generalized Cross Validation (GCV), by using the **mgcv** library [Wood, 2016] in the statistical platform R.

The temporal dynamic is estimated from the residuals of the model 3 through a P-spline smoothing model, with a basis matrix $\Phi(t)$ spanning the space of the time and a vector of parameters θ^p estimated by penalized least square:

$$\chi^p_{\mathbf{s}}(t) = \Phi(t)\theta^p. \tag{6}$$

2.2. Functional Kriging

Functional Kriging is the prediction of spatially referred curves in an unvisited site, based on curves at nearby locations weighted by the strength of their correlation with the location of interest $\mathbf{s}_0$, in such a way that curves from those locations closer to the prediction point will have greater influence. For a single variable of interest, some contributions, discussed in [Delicado *et al.*, 2010], extend the classical geostatistical techniques to the functional context, providing a definition of functional variogram.

Among them, the best linear unbiased predictor, the ordinary kriging for function-valued spatial data, proposed by [Giraldo *et al.*, 2011b], is the approach here adopted for the univariate component $\chi^p_{\mathbf{s}}(t)$ of our functional model. In other words, we consider a second-order stationary and isotropic functional random process, that is, the mean function is constant in the domain $D \subseteq R^2$: $E\left[\chi^p_{\mathbf{s}}(t)\right] = \mu(t)$. The second order properties of the process are described by a covariance function, depending only on the distance h between two sampling points $\mathbf{s}_i, \mathbf{s}_j$ and on time t: $C(h;t) = Cov(\chi^p_{\mathbf{s}_i}(t),\ \chi^p_{\mathbf{s}_j}(t))$, and by the functional variogram:

$$\gamma(h;t) = \frac{1}{2}Var(\chi^p_{\mathbf{s}_i}(t) - \chi^p_{\mathbf{s}_j}(t)),\quad h = \left\|\mathbf{s}_i - \mathbf{s}_j\right\|.$$

It also implies that the variance is constant.

The predictor $\hat{\chi}^p_{\mathbf{s}_0}(t)$ in an unvisited site $\mathbf{s}_0$ is a linear combination of the available curves $\chi^p_{\mathbf{s}}(t)$ with the optimal weight determined on the trace-variogram, the mean function obtained by integrating the variogram function over the time:

$$\gamma(h) = \frac{1}{2}E\left[\int_T \left(\chi^p_{\mathbf{s}_i}(t) - \chi^p_{\mathbf{s}_j}(t)\right)^2 dt\right]. \tag{7}$$

In order to have the best linear unbiased predictor (BLUP), the weights are estimated by minimizing:

$$min_\alpha = \int_T E\left\|\hat{\chi^p}_{\mathbf{s}_0}(t) - \chi^p_{\mathbf{s}_0}(t)\right\|^2 dt. \tag{8}$$

Since the curves are estimated by means of a linear combination of B-Spline and coefficients, the kriging prediction is carried out, at an unvisited location, by kriging on the coefficients of the spline:

$$min_\alpha = \int_T E \left\| \sum\nolimits_{i=1}^{n} \alpha_i(t)\chi^p_{\mathbf{s}_i}(t) - \chi^p_{\mathbf{s}_0}(t) \right\|^2 dt, \tag{9}$$

subject to the constraints on the weights: $\sum_{i=1}^{n} \alpha_i = 1$.

As result, each curve is weighted by a scalar parameter:

$$\hat{\chi^p}_{\mathbf{s}_0}(t) = \sum\nolimits_{i=1}^{n} \alpha_i(t)\chi^p_{\mathbf{s}_i}(t). \tag{10}$$

The R package **geofd** [Giraldo *et al.*, 2015] implements the ordinary kriging prediction for functional data.

3. Two methodological proposals

The two proposed procedures aim to include the multivariate information in the two components of (2), the first one, FCK, taking into account the cross dependence in $\chi^p_{\mathbf{s}}(t)$, while the second, FKED, involving regression models for $Z^p(\mathbf{s})$.

3.1. Functional Kriging with External Drift (FKED)

The following procedure includes the influence of other covariates, combining a regression of a variable of primary interest on the other variables with functional kriging of the regression residuals (FKED). In this subsection we extend the general procedure known as kriging with external drift to a broader range of regression techniques. In our procedure we identify one of the dimensions of the multivariate process as the primary variable of interest and consider the other as secondary variables. We aim at predicting the primary variable at an unvisited location getting information from curves of primary and secondary variables at a possibly different set of distinct locations. The residuals of the estimated regression are the input of the procedure of functional ordinary kriging predictor. From a practical point of view, this hybrid techniques, based on regression and kriging, may play an interesting role in dealing with missing values through predictive models, incorporating available information from several different variables.

Due to the high flexibility in model specification, GAMs provide a proper framework for including covariates in the spatial predictor rather straightforwardly: predictions are drawn by estimating the relationship between the p^{th} variable of interest, denoted by $Z_t^{p^*}(\mathbf{s})$ and the set of the other $P-1$ auxiliary variables $Z_t^{\{P-1\}}(\mathbf{s})$ at sample locations, and applying the model to unvisited locations:

$$Z^{p^*}(\mathbf{s}) = f(\mathbf{s}) + g(\mathbf{Z}_t^{\{P-1\}}(\mathbf{s})). \tag{11}$$

Both for computational reasons an for interpretability, the model assumes an additive structures: for each covariate a penalized regression spline of order m smooths the data and quite simple expressions can be derived for the estimator of the functional data:

$$Z^{p^*}(\mathbf{s}) = f(\mathbf{s}) + \sum_{p=1}^{P-1} g(\mathbf{Z}_t^{p}(\mathbf{s})), \tag{12}$$

and for the penalty matrix:

$$\mathbf{H}_\lambda = \lambda_1 \mathbf{I}_{k_2} \otimes \mathbf{D}_1{}'\mathbf{D}_1 + \lambda_2 \mathbf{D}_2{}'\mathbf{D}_2 \otimes \mathbf{I}_{k_1} + \sum_{p=1}^{P-1} \lambda_p \mathbf{D}_p{}'\mathbf{D}_p. \tag{13}$$

For the reconstructed functional datum in the location $\mathbf{s}_0$, the standard error of prediction is also known [Giraldo *et al.*, 2011b]. In the subsequent step we focus on the observed residuals of model (12) in order to estimate the ordinary kriging predictor $\hat{\chi}^p_{\mathbf{s}_0}(t)$ and the functional datum is obtained as:

$$\hat{X}^p(\mathbf{s}_0, t) = \hat{Z}^p(\mathbf{s}_0) + \hat{\chi}^p_{\mathbf{s}_0}(t).$$

Depending on the strength of the auxiliary information in the maps of covariates and on the spatial correlation among curves, the model might turn to pure kriging (no influence from covariates) or pure regression (pure nugget variogram).

3.2. Functional CoKriging (FCK)

An alternative procedure includes the information of other covariates in the functional prediction of spatially correlated multivariate random processes, accounting for the cross-dependence between the different p dimensions. We denote it as Functional

Cokriging (FCK). The aim is to predict a curve at a location of interest weighting all the p dimensional curves from those locations closer to the prediction point. For the initial model (2) we adopt the definition of the component $Z^p(\mathbf{s})$ as a smoothing function of coordinates and we focus on the component $\chi^p_{\mathbf{s}}(t)$ for which we derive a linear predictor with weights determined by the strength of the correlations among the curves in the same site and in different sites. The most natural way to generalize the functional prediction is to generalize the trace variogram, defining a similar measure of cross-dependence between curves. Referring to [Cressie, 1993], let generalize the cross-variance between the curves, referred to two dimensions p and p' in two sites $\mathbf{s}_i$ and $\mathbf{s}_j$, in the functional context:

$$\gamma^{p,p'}(h;t) = \frac{1}{2} Var(\chi^p_{\mathbf{s}_i}(t) - \chi^{p'}_{\mathbf{s}_j}(t)), \tag{14}$$

for $h = \left\|\mathbf{s}_i - \mathbf{s}_j\right\|$, p and p' in $1, \ldots, P$.

In the site $\mathbf{s}_0$, where the set of the other $P-1$ covariates $\chi^p_t(\mathbf{s})$ is available, the prediction is:

$$\hat{\chi}^{p^*}_{\mathbf{s}_0}(t) = \sum_{i=1}^{n} \sum_{p=1}^{P} \alpha_{ij}(t) \chi^p_{\mathbf{s}_i}(t). \tag{15}$$

The vector α being the solution that minimizes under the uniform-unbiasedness assumptions:

$$min_\alpha = \int_T E \left\| \hat{\chi}^{p^*}_{\mathbf{s}_0}(t) - \chi^{p^*}_{\mathbf{s}_0}(t) \right\|^2 dt \tag{16}$$

$$min_\alpha = \int_T E \left\| \sum_{i=1}^{n} \sum_{p=1}^{P} \alpha_{ij}(t) \chi^p_{\mathbf{s}_i}(t) - \chi^{p^*}_{\mathbf{s}_0}(t) \right\|^2 dt \tag{17}$$

subject to the constraints:

$$\sum_{i=1}^{n} \alpha_{ij} = 1 \ , \ for \ p = p^* \tag{18}$$

$$\sum_{j=1}^{p} \alpha_{ij} = 0 \ , \ for \ p \neq p^*. \tag{19}$$

By analogy with Functional kriging, the proposal goes through the definition of the trace - covariogram:

$$\Gamma^{pp'}(h) = \frac{1}{2} E \left[\int_T (\chi^p_{\mathbf{s}_i}(t) - \chi^{p'}_{\mathbf{s}_j}(t))^2 dt \right], \ \ h = \left\|\mathbf{s}_i - \mathbf{s}_j\right\|, \tag{20}$$

and the implementation of the trace covariogram in a optimization procedure.

To implement our proposal, all computations are coded in R (R Development Core 2016). The conversion to functional data is realized by using the **fda** package [Ramsay *et al.*, 2014] and **mgcv** package [Wood, 2016], while the **geofd** package [Giraldo *et al.*, 2015] is also used to implement the proposed kriging procedure. The R code is available on request.

4. Dealing with real data

In order to show the behavior of the two proposed procedures, a spatio-temporal multivariate data set related to air quality is here considered.

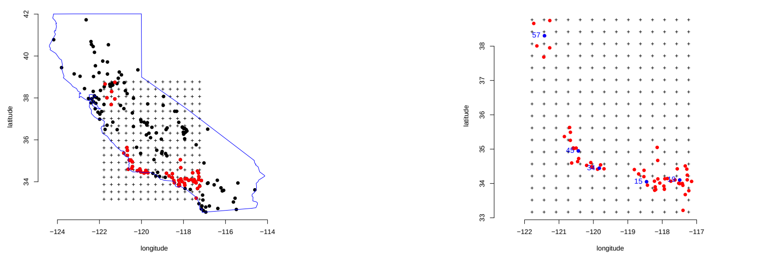
In particular, our case study considers PM_{10} and the main daily gaseous pollutant concentrations (CO, NO_2, O_3, SO_2), recorded during 2011 and aggregated by month, at 59 monitoring stations dislocated along the State of California (raw data are available at: http://www.epa.gov).

The sites in our map make up a regular space-time grid with respect to the five pollutants, in the sense that there is the same configuration of spatial points at each time. Data on the regular space-time grid consists of 295 time series, arranged in a $12 \times 59 \times 5$ array; five of the monitoring sites are excluded from the analysis and used for assessing the performance of the proposed procedures. A map of the monitored area, with the observed sites, is reported in Fig. 4; the five sites chosen as validation set are highlighted in blue (Fig. 4, right).

The concentrations of the pollutants are opportunely standardized and scaled in [0, 100], through the linear interpolation introduced by Ott and F. [1976] and used by US EPA (Environmental Protection Agency); as shown in [Ruggieri and Plaia, 2012], the standardization by segmented linear function with respect to the standardization by threshold value, allows accounting for different effects of each pollutant on human health, as well as for short and long-term effects.

The 2-step GAM procedure estimates the curves using P-spline: functions of the coordinates only are estimated in the preliminary step (eq. 3), in order to take into account the main spatial variations; then, the underlying temporal variability of the residuals of the previous model is modelled in order to obtain estimations of the 59×5 functions of time (eq. 6). The parameters (number of knots and smoothing parameters) are selected by mean of Generalized Cross Validation.

Then we performed the two proposed procedures, in order to assess the spatial prediction capability: combining a regression of a variable of primary interest on the other variables, with functional kriging of the regression residuals (FKED) (eq.10 and eq.11); or including the information of other covariates in prediction of spatially correlated multivariate random processes (FCK) (eq.3 and eq.15). Cross-validation is applied to compare their performances.



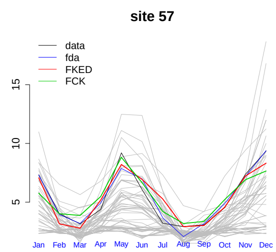
(a) The air monitoring network in California

(b) The cross-validation procedure

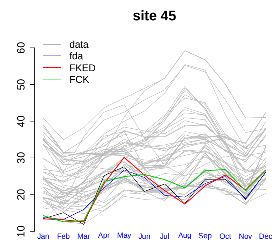
Figure 4: Maps

Plots of the predicted curves for the validation set (blue points in the maps reported in Fig. 4) are presented in Fig. 5 for some pollutants and sites. In each site of the validation set, each pollutant in turn is considered not observed. Its prediction, FKED (red line) and FCK (green line), is compared to the observed time series (black line) and to the functional estimation (blue line), the last obtained including the site in the estimation procedure; in the figures, the smoothed curves in all the observed sites (gray lines) are also represented in the background. The predictions appear overall consistent, being very close to the smoothed and observed data for both the approaches. They both catch the main variations in time and this suggests that the results are improved when the spatial kriging exploits common dynamics in different pollutants.

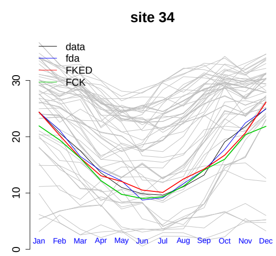
Results from the leave-one-out cross-validation (not distinguishing between test set and validation set) for testing the two algorithms may be also evaluated compared the



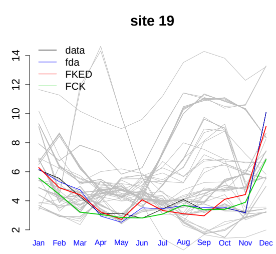
(a) CO



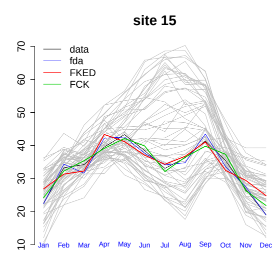
(b) PM_{10}



(c) NO_2



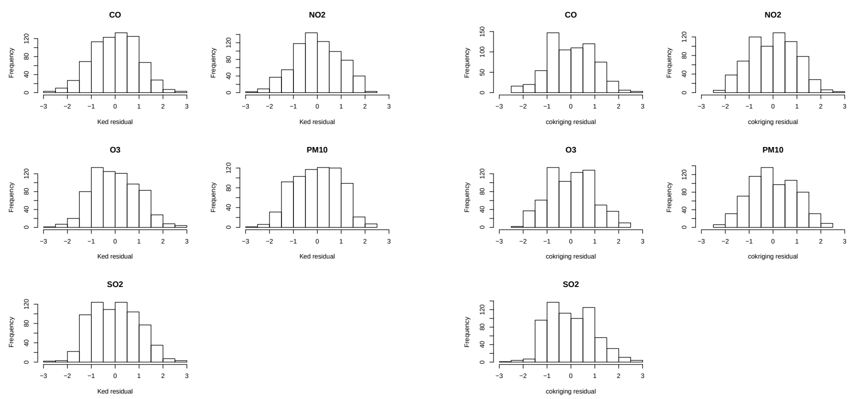
(d) SO_2



(e) O_3

Figure 5: FKED and FCK prediction

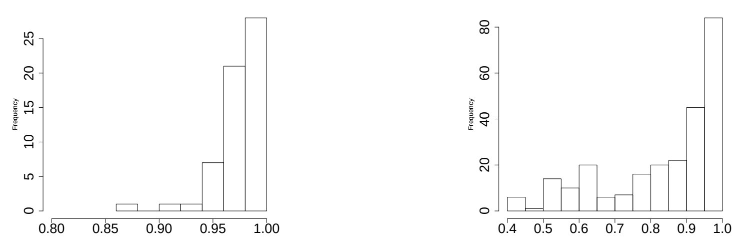
histograms of the standardized residuals (Fig. 6), i.e. the predicted values minus the fda values, divided by the kriging variance; they confirm unbiased predictors for both approaches.



(a) FKED (b) FCK

Figure 6: Standardized residuals from FKED and FCK

The correlation, as well as the root mean square deviation (RMSD), between the estimated functional data and predictors are also presented in Figg. 7 and 8, respectively, for the FKED and FCK approaches. In both cases, the FKED approach performs slightly better, as it emerges from the direct comparison of the two distributions.

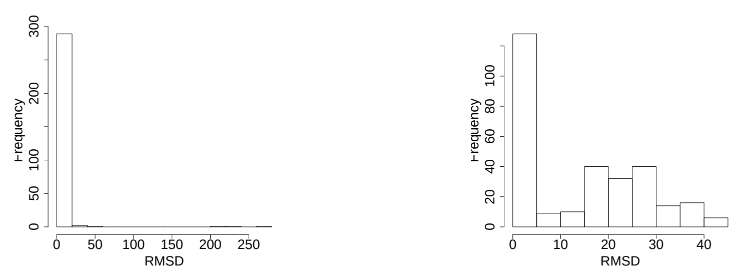


(a) FKED (b) FCK

Figure 7: Distribution of the correlations between estimated functional data and predictors in the validation set

5. Conclusions and further developments

In this paper an integration of Multivariate Spatial FDA with kriging for functional data is proposed, exploiting correlations among variables in order to predict one of them. In particular, we want to consider a recurrent case, when more than a single variable (pollutants, for example) is recorded and a variable has to be predicted in a site where a) no other variables are recorded; b) other variables are recorded. Actually, even if we are interested in predicting a single variable in an unmonitored site, exploiting its correlation with the other variables can improve the estimation. In this paper, we want to focus on case a). The spatial prediction capability of the proposed procedures has been assessed considering a three way array ($time \times space \times variables$) containing the concentrations of 5 main pollutants recorded in 59 monitoring sites in California (USA) over a year. We focus on predicting each pollutant in an unmonitored site. The performance of the proposed procedures has been evaluated first graphically, comparing observed and predicted data at five validation sites. A more detailed performance evaluation has been carried out considering some performance indexes. In particular, the correlation coefficient ρ (the higher the better) and the root mean square deviation RMSD (the lower the better) have been computed by comparing recorded and estimated data considering a leave-one-out procedure. As specified, here we deal only with the case a), getting good performances. An extension of the proposed procedures will be considered in a future work to explore their potentiality when the case b) has to be treated.



(a) FKED (b) FCK

Figure 8: Distribution of the RMSD between estimated functional data and predictors in the validation set

References

Caballero, W., Giraldo, R., and Mateu, J. (2013). *A universal kriging approach for spatial functional data*. Stoch Environ Res Risk Assess.

Cressie, N. (1993). *Statistics for spatial data*. Wiley.

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

Eilers, P. and Marx, B. (1996). Flexible smothing with b-splines and penalties. *Statistical Science*, **11**, 89–121.

Giraldo, R. (2009). *Geostatistical Analysis of Functional Data*. Ph.D. thesis, Universitat Politecnica da Catalunya, Barcellona.

Giraldo, R. (2014). Cokriging based on curves: Prediction and estimation of the prediction variance. InterStat. `http://interstat.statjournals.net/YEAR/2014/abstracts/1405002.php`.

Giraldo, R., Delicado, P., and 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., and Mateu, J. (2011a). Geostatistics with infinite dimensional data: a generalization of cokriging and multivariable spatial prediction. *Matematica: ICM-ESPOL*, **9**(1), 1390–3802.

Giraldo, R., Delicado, P., and Mateu, J. (2011b). Ordinary kriging for function valued spatial data. *Environ Ecol Stat*, **18**(3), 411–426.

Giraldo, R., Delicado, P., and Mateu, J. (2015). geofd: Spatial prediction for function value data. *R package version 1.0*.

Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models*. Chapman & Hall/CRC. Boca Raton.

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

Lee, D. and Durban, M. (2013). P-spline anova-type interaction models for spatio-temporal smoothing. *Statistical Modelling*, **11**(1), 49–69.

Menafoglio, A., Secchi, P., and Dalla Rosa, M. A. (2013). Universal kriging predictor for spatially dependent functional data of a hilbert space. *Electronic Journal of Statistics*, **7**, 2209–2240.

Menafoglio, A., Guadagnini, A., and Secchi, P. A. (2014). Kriging approach based on aitchison geometry for the characterization of particle-size curves in heterogeneous aquifers. *Stochastic Environmental Research and Risk Assessment*, **28**, 1835–1851.

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

Ott, W. R. and F., H. J. W. (1976). A quantitative evaluation of the pollutant standards index. *J Air Pollut Control Assoc*, **26**, 1051–1054.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer-Verlag, second edition.

Ramsay, J. O., Wickham, H., Graves, S., and Hooker, G. (2014). fda: Functional data analysis. *R package version 2.4.4*.

Ruggieri, M. and Plaia, A. (2012). An aggregate aqi: comparing different standardizations and introducing a variability index. *Science of the Total Environment*, **420**, 263–272.

Wood, S. N. (2016). mgcv: mixed gam computation vehicle with gcv/aic/reml smoothness estimation. *R package version 1.8-15*.