AStA Adv Stat Anal
DOI 10.1007/s10182-017-0305-4

ORIGINAL PAPER

A distance-based model for spatial prediction using radial basis functions

Carlos E. Melo[1] · **Oscar O. Melo**[2] · **Jorge Mateu**[3]

Received: 18 June 2016 / Accepted: 17 July 2017

Abstract In the context of local interpolators, radial basis functions (RBFs) are known to reduce the computational time by using a subset of the data for prediction purposes. In this paper, we propose a new distance-based spatial RBFs method which allows modeling spatial continuous random variables. The trend is incorporated into a RBF according to a detrending procedure with mixed variables, among which we may have categorical variables. In order to evaluate the efficiency of the proposed method, a simulation study is carried out for a variety of practical scenarios for five distinct RBFs, incorporating principal coordinates. Finally, the proposed method is illustrated with an application of prediction of calcium concentration measured at a depth of 0–20 cm in Brazil, selecting the smoothing parameter by cross-validation.

Work partially funded by Grant MTM2016-78917-R from the Spanish Ministry of Science and Education.

Electronic supplementary material The online version of this article (doi:10.1007/s10182-017-0305-4) contains supplementary material, which is available to authorized users.

✉ Oscar O. Melo
oomelom@unal.edu.co

Carlos E. Melo
cmelo@udistrital.edu.co

Jorge Mateu
mateu@mat.uji.es

1 Faculty of Engineering, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia

2 Department of Statistics, Faculty of Sciences, Universidad Nacional de Colombia, Crr. 30 # 45-03, Bogotá, Colombia

3 Department of Mathematics, University Jaume I, Campus Riu Sec, 12071 Castellón de la Plana, Castellón, Spain

Published online: 26 July 2017

Keywords Detrending · Distance-based methods · Radial basis functions · Random function models · Smoothing parameter · Spatial prediction

1 Introduction

Observing and saving complete functions as a result of random experiments are nowadays possible by the development of real-time measurement instruments and data storage resources. In this context, global methods, such as trend surface analysis, use all available data for prediction; while local methods, including inverse distance weighting, usually use only some subset of the data for prediction. One benefit of local methods is that computational time is reduced by using only some subset of the data for prediction purposes. Some methods may make use of all available data, but take into account distances only from the prediction location.

Radial basis functions (RBFs) such as multiquadratic (MQ) or completely regularized splines (CRS) are useful in the construction of digital elevation models (DEM), as shown in Mitášová and Hofierka (1993), in which the spline is incorporated into a geographic information system to study soil erosion. One variation of the multiquadratic function is called the inverse multiquadratic (IMQ), introduced by Hardy and Gopfert (1975). The thin plate spline (TPS) was introduced in geometric design by Duchon (1976). The name TPS refers to a physical analogy involving the bending of a thin sheet of metal. Later, Thiébaux and Pedder (1987) described the TPS as a two-dimensional (surface) cubic spline. Franke (1982) developed a computer program for the solution of the scattered data interpolation problem; the algorithm is based on a weighted sum of locally defined thin plate splines, and yields an interpolation function which is differentiable. Another popular variant to TPS is the Gaussian (GAU) approximation used in Schagen (1979). Back in time, Späth (1969) described a method that allows to avoid inflexion points and contains cubic splines as a special case, using piecewise cubic and exponential (EXP) spline interpolation. Finally, Mitáš and Mitášová (1988), Mitášová and Hofierka (1993), and Mitášová and Mitáš (1993) developed the formulation of the spline with tension (ST), and implemented a segmentation algorithm with a flexible size of overlapping neighborhood.

The link between splines and kriging was called "near" equivalent, because the TPS corresponds to a specific generalized covariance, whereas the kriging estimator or the RBF interpolator only requires the use of a kernel with appropriate positive definiteness properties. In general, this allows adapting the kernel function to a particular data set (Cressie 1989; Myers 1992). Then, for both splines and kriging, the smoothing is determined objectively.

Recent studies using RBFs on irregular domains in two dimensions through the process of conformal transplantation are presented in Heryudono and Driscoll (2010). Zhang (2012) developed a fast algorithm for univariate spline smoothing estimator in a multivariate regression through the use of compactly supported RBFs. Yavuz and Erdogan (2012) performed an analysis of monthly and annual precipitation trends, using ordinary kriging, inverse distance weighting, and completely regularized spline interpolation methods. These studies demonstrate the usefulness of working with RBFs.

In many disciplines related to spatial data analysis (mining, hydrogeology, ecology, earth sciences and environment, among others), we often have to deal with variables of different nature that are associated with the spatial response variable: categorical and binary variables such as type of soil or rock, and continuous variables (the spatial coordinates or additional variables, for example). However, the above-mentioned methods are not entirely appropriate when modeling a mixture of explanatory variables.

Our aim here is to present a unified approach that uses RBFs in such spatial contexts where the explanatory variables are of mixed nature. This paper proposes a new method using distances between individuals such as the Gower (1968) distance, although some other Euclidean distances may also be used. We deal with spatial regression modeling with covariates. Rather than the covariates entering the model directly as predictors, they are used to define distances between observations in the covariate space using a measure proposed by Gower. The eigen-decomposition of the resulting distance matrix is then used to define new covariates (principal coordinates), and those having the greatest correlation with the response are selected for use. In addition, RBFs based on geographic separation of the observations are used to model residual spatial correlation. Therefore, we obtain the distance-based spatial RBFs (DBSRBFs) method, which can be applied in the geostatistical context to predict the trend and to estimate the covariance structure when the principal coordinates as covariates are used.

The DBSRBF is an excellent alternative because it takes full advantage of the information obtained due to the relationship between observations, which can be established through the use of the spectral decomposition using any Euclidean distance. This approach aims at improving the predictions by allowing for inclusion of any number of principal coordinates at the sample locations. Note that universal kriging traditionally considers only information of a matrix of size $n \times p$, where n is the number of individuals and p the number of attributes with $p < n$. However, the DBSRBF proposed method considers a distance matrix between individuals with $n(n-1)/2$ different elements. Then, we build a matrix of principal coordinates using the spectral decomposition from this distance matrix getting information of $(n-1)$ independent columns, obtaining much more than p attributes for the trend. However, to avoid the problem of non-identifiability in the parameter estimation process for the trend in the DBSRBF proposed model, we select the most relevant principal coordinates (see Sect. 2), identifying which principal coordinates are most associated with the response variable. It is important to emphasize that principal coordinates transformation helps to reduce the noise in the geostatistical model because uncorrelated features are typically moved to higher order coordinates.

The plan of the paper is the following. Section 2 develops the methodological proposal introducing the local linear distance-based drift. In this section, some RBFs are described and constructed from the distance-based drift, and an approximation from the spline interpolation to the kriging model for the prediction is shown. Section 3 presents a simulation study based on a variety of practical scenarios that include five different RBFs and incorporate the principal coordinates. Section 4 illustrates the proposed methodology with an application of prediction of calcium concentration measured at a depth of 0–20 cm in Brazil. The paper ends with some conclusions.

2 A distance-based geostatistical model with RBFs

Suppose we are interested in a continuous geo-referenced response variable associated with binary, categorical or other continuous explanatory variables. Let $\boldsymbol{s} \in \mathbb{R}^d$ be a location in a d-dimensional Euclidean space, and suppose that $\boldsymbol{Z}(\boldsymbol{s})$ is a random vector at each spatial location $\boldsymbol{s}$. If $\boldsymbol{s}$ is allowed to vary on a given set $D \subseteq \mathbb{R}^d$, we have a stochastic process $\{Z(\boldsymbol{s}), \boldsymbol{s} \in D\}$, which is object of study within the context of geostatistics (Cressie 1993). Also assume that D is a certain fixed and continuous region, and the spatial index $\boldsymbol{s}$ varies continuously in D, i.e., there is an infinite number of possible locations where the process can be observed.

Assume the stochastic process follows a random function model, and can be written as

$$Z(\boldsymbol{s}_i) = \mu(\boldsymbol{s}_i) + e(\boldsymbol{s}_i), \quad i = 1, \ldots, n \tag{1}$$

where $Z(\boldsymbol{s}_i)$ is the regionalized variable given by the sum of a deterministic function associated with the trend $\mu(\boldsymbol{s}_i)$, and $e(\boldsymbol{s}_i)$ is a stationary stochastic component with zero mean and variogram $2\gamma(\cdot)$. The trend is formed by categorical, continuous and binary variables and it is modeled as

$$\mu(\boldsymbol{s}_i) = \theta_0 + \boldsymbol{v}^t(\boldsymbol{s}_i)\boldsymbol{\theta}, \quad i = 1, \ldots, n \tag{2}$$

where $\boldsymbol{v}(\boldsymbol{s}_i) = (v_1(\boldsymbol{s}_i), \ldots, v_p(\boldsymbol{s}_i))^t$ is a vector containing explanatory variables associated with the spatial location $\boldsymbol{s}_i$, and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^t$ is a vector of unknown parameters.

In matrix form, the model (1) can be expressed as

$$\boldsymbol{Z}_s = \mathbf{1}\theta_0 + \boldsymbol{V}\boldsymbol{\theta} + \boldsymbol{e}_s \tag{3}$$

where $\boldsymbol{Z}_s = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n))^t$, $\mathbf{1}$ is a vector of dimension $n \times 1$ which is associated with the intercept, $\boldsymbol{V} = (\boldsymbol{V}_1, \ldots, \boldsymbol{V}_p)$ is a design matrix of dimension $n \times p$ with p explanatory variables which are given by $\boldsymbol{V}_j = (v_j(\boldsymbol{s}_1), \ldots, v_j(\boldsymbol{s}_n))^t$, $j = 1, \ldots, p$, and $\boldsymbol{e}_s = (e(\boldsymbol{s}_1), \ldots, e(\boldsymbol{s}_n))^t$. By simplicity, we assume that $\boldsymbol{V} = \boldsymbol{H}\boldsymbol{V}^*$ where $\boldsymbol{H} = \boldsymbol{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^t$ is the centering matrix with $\boldsymbol{I}$ the identity matrix of size $n \times n$ and $\mathbf{1}$ a $n \times 1$ vector of ones. $\boldsymbol{V}^*$ is the matrix of original explanatory variables, it may involve continuous, categorical and binary variables, or even a mixture of them. Note that it is not necessary to centralize the covariates in the proposed model; this is performed with the purpose to build the distances, but depending on the distance used we do not always need to centralize the covariates.

Now, our distance-based approach applies to a transformation of the explanatory variables. To this aim, we need to define some similarity (or Euclidean distance) measures, which depend on the explanatory variable characteristics. The distance based is an excellent alternative because it takes full advantage of the information obtained due to the relationship between observations, which can be established through the use of the spectral decomposition using any Euclidean distance.

According to Cuadras (1989) and Cuadras and Arenas (1990), let $\Omega = \{\omega_1, \ldots, \omega_n\}$ be a set consisting of n individuals, and let $\delta_{ii'} = \delta(\omega_i, \omega_{i'}) =$

$\delta(\omega_{i'}, \omega_i) \geq \delta(\omega_i, \omega_i) = 0$ be a distance (or dissimilarity) function defined on Ω. Suppose that the distance matrix with dimension $n \times n$, $\Delta = (\delta_{ii'})$ is Euclidean.

If the vector $v(\boldsymbol{s}_i)$ given in (2) is formed by binary, categorical and continuous variables, then the similarity according to Gower (1968) can be defined for mixed variables as

$$m_{ii'} = \frac{\sum_{h=1}^{p_1} \left(1 - \frac{|v_h(\boldsymbol{s}_i) - v_h(\boldsymbol{s}_{i'})|}{G_h}\right) + c_{1ii'} + \omega_{ii'}}{p_1 + (p_2 - c_{4ii'}) + p_3} \tag{4}$$

where G_h is the range of the h-th continuous variable, p_1 is the number of continuous variables, $c_{1ii'} = c_1(\boldsymbol{s}_i, \boldsymbol{s}_{i'})$ and $c_{4ii'} = c_4(\boldsymbol{s}_i, \boldsymbol{s}_{i'})$ are the number of positive and negative matches, respectively, for p_2 binary variables, and $\omega_{ii'} = \omega(\boldsymbol{s}_i, \boldsymbol{s}_{i'})$ is the number of matches for p_3 multistate variables. In the case that all explanatory variables are binary or categorical, the similarity (4) reduces to

$$m_{ii'} = \frac{c_{1ii'} + c_{4ii'}}{c_{1ii'} + c_{2ii'} + c_{3ii'} + c_{4ii'}} \quad \text{(Sokal–Michener)}$$

where $c_{1ii'}$, $c_{2ii'} = c_2(\boldsymbol{s}_i, \boldsymbol{s}_{i'})$, $c_{3ii'} = c_3(\boldsymbol{s}_i, \boldsymbol{s}_{i'})$, $c_{4ii'}$ are the frequencies of (1,1), (1,0), (0,1) and (0,0), respectively. Note that for l categories of one categorical variable, $l - 1$ dummy variables are created of ones and zeros corresponding to the presence or absence of one category (e.g., with 1 for the presence of A category in that variable and zero for the non-presence of that category, and we did the same procedure for $B, C, \ldots$ categories). Therefore, the Sokal–Michener formula can be applied because we have $l - 1$ binary variables.

Through the transformation

$$\delta_{ii'} = \sqrt{1 - m_{ii'}},$$

it is possible to obtain Euclidean distances. If all explanatory variables in (2) were continuous, the squared distance would be defined as

$$\begin{aligned} \delta^2_{ii'} &= (\boldsymbol{v}(\boldsymbol{s}_i) - \boldsymbol{v}(\boldsymbol{s}_{i'}))^t (\boldsymbol{v}(\boldsymbol{s}_i) - \boldsymbol{v}(\boldsymbol{s}_{i'})) \\ &= \boldsymbol{v}^t(\boldsymbol{s}_i)\boldsymbol{v}(\boldsymbol{s}_i) + \boldsymbol{v}^t(\boldsymbol{s}_{i'})\boldsymbol{v}(\boldsymbol{s}_{i'}) - 2\boldsymbol{v}^t(\boldsymbol{s}_i)\boldsymbol{v}(\boldsymbol{s}_{i'}) \end{aligned} \tag{5}$$

or alternatively, the absolute distance $\delta^2_{ii'} = \sum_{h=1}^{p} |v_h(\boldsymbol{s}_i) - v_h(\boldsymbol{s}_{i'})|$ can be used. Thus, in those cases in which the information is based only on the spatial locations (w_x, w_y), the spatial distances are then directly given by

$$\delta_{ii'} = \sqrt{(w_{x_i} - w_{x_{i'}})^2 + (w_{y_i} - w_{y_{i'}})^2}$$

Expressions for the Gower similarity as given in Eq. (4) will be useful to the extent of having information associated with mixed variables, not only for the sampling sites but also for the non-sampled, which limits its use in unsampled areas.

After selecting a distance, let $\boldsymbol{A}_{n\times n} = (a_{ii'})$ be the matrix with elements $a_{ii'} = -\delta^2_{ii'}/2$ and $\boldsymbol{B} = \boldsymbol{HAH}$. It is known that $\boldsymbol{B}$ is a semi-definite positive matrix of rank

k (Mardia et al. 2002), so the principal coordinates matrix $\boldsymbol{X}$ is obtained from the spectral decomposition as

$$\boldsymbol{B} = \boldsymbol{HAH} = \boldsymbol{U\Lambda U}^t = \boldsymbol{XX}^t$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{B}$, and $\boldsymbol{X} = \boldsymbol{U\Lambda}^{1/2}$ is a $n \times n$ matrix of rank $k \leq n-1$.

Thus, the trend given in (3) becomes

$$\boldsymbol{Z}_s = \beta_0 \mathbf{1} + \boldsymbol{X\beta} + \boldsymbol{\xi}_s \tag{6}$$

where $\boldsymbol{Z}_s$ comes in (3), β_0 and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^t$ are the unknown parameters, $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k)$ is the principal coordinates matrix, and $\boldsymbol{\xi}_s = (\xi(\boldsymbol{s}_1), \ldots, \xi(\boldsymbol{s}_n))^t$ with each $\xi(\boldsymbol{s}_i)$ a stationary stochastic component with zero mean and variogram $2\gamma(\cdot)$. Note that $\mathbf{1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$ are eigenvectors of $\boldsymbol{B}$ with eigenvalues $0, \lambda_1, \ldots, \lambda_k$, respectively, and $\boldsymbol{X}_i^t \boldsymbol{X}_i = \lambda_i$, $\boldsymbol{X}_i^t \boldsymbol{X}_j = 0$ $(i \neq j)$, and $\boldsymbol{X}_i^t \mathbf{1} = 0$, $i, j = 1, \ldots, k$.

To avoid the problem of having a coefficient of determination $R^2 \simeq 1$ when the rank of $\boldsymbol{X}$ is $k = n-1$, it is necessary to consider only the most correlated eigenvectors of $\boldsymbol{B}$, given by $\mathbf{1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$, with the regionalized variable $\boldsymbol{Z}_s$, i.e., the most significantly correlated covariates with $\boldsymbol{Z}_s$. In order to decide whether a predictor variable, i.e., a column of $\boldsymbol{X}$, has to be included or deleted, the variables can be arranged in descending order of absolute correlations with $\boldsymbol{Z}_s$,

$$r^2(\boldsymbol{Z}_s, \boldsymbol{X}_1) > \cdots > r^2(\boldsymbol{Z}_s, \boldsymbol{X}_k) > \cdots > r^2(\boldsymbol{Z}_s, \boldsymbol{X}_{n-1}) \tag{7}$$

where $r^2(\boldsymbol{Z}_s, \boldsymbol{X}_i) = \dfrac{\boldsymbol{Z}_s^t \boldsymbol{X}_i \boldsymbol{X}_i^t \boldsymbol{Z}_s}{\lambda_i \sum_{i'=1}^{n} \left(Z(\boldsymbol{s}_{i'}) - \bar{Z}\right)^2}$, $i = 1, \ldots, n-1$, with $\bar{Z} = \sum_{i'=1}^{n} Z(\boldsymbol{s}_{i'})/n$.

Moreover, a principal coordinate $\boldsymbol{X}_i$ should be deleted if the null hypothesis $\beta_i = 0$ is not rejected. Following Cuadras et al. (1996), we can build a statistical test of the form

$$t_i = \frac{\hat{\beta}_i}{\|\boldsymbol{Z}_s - \hat{\beta}_0 \mathbf{1} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\|^2} \sqrt{\lambda_i (n-k-1)} \tag{8}$$

where $\hat{\beta}_0 = \bar{Z}$, $\widehat{\boldsymbol{\beta}} = \boldsymbol{\Lambda}^{-1} \boldsymbol{X}^t \boldsymbol{Z}_s$ and $\hat{\beta}_i$ is the i-th component of $\widehat{\boldsymbol{\beta}}$. (8) follows a t-student distribution with $(n-k-1)$ degrees of freedom.

Another possibility is through the explained variabilities given by the predictor variables, which are given by the largest eigenvalues $\lambda_1 > \cdots > \lambda_k > \lambda_{k+1} > \cdots > \lambda_{n-1}$, and then choosing the k first principal coordinates. However, as noted by Cuadras (1993), if a dimension with a small eigenvalue is correlated with the response variable, this dimension will be correlated with "noise" rather than with the main variability of the data.

Another good alternative for the selection of principal coordinates is performed in a similar way to the selection of the number of variables in multivariate regression, using the statistic called C_p − Mallows. That is, we draw a graph representing the points $(i, 1-c(i))$ $i = 1, \ldots, k, k+1, \ldots, n-1$, and then the point with a significant

fall in the lack of predictability, given by $1 - c(i)$, is determined. The predictability $c(i)$ is given by

$$c(0) = 0, \qquad c(i) = \frac{\sum_{j=1}^{i} r_j^2 \lambda_j}{\sum_{j=1}^{n-1} r_j^2 \lambda_j}, \quad i = 1, \ldots, n-1 \tag{9}$$

where $r_j^2 = r^2(\boldsymbol{Z}_s, \boldsymbol{X}_j)$, and λ_j is the j-th eigenvalue associated with $\boldsymbol{X}_j$, $j = 1, 2, \ldots, n-1$ [see Cuadras et al. (1996) for more details].

Finally, following any of the methods presented above, the $\boldsymbol{X}_{k+1}, \ldots, \boldsymbol{X}_{n-1}$ principal coordinates must be removed.

2.1 Relation with the classical geostatistical model

Model (6) depends on the chosen distance δ_{ij}. Therefore, when the predictor variables are continuous and the Euclidean distance is used, the distance-based geostatistical model is compatible with the classical geostatistical model. This equivalence also holds for qualitative and mixture (continuous and categorical) explanatory variables when an appropriate dissimilarity measure is used.

If all the explanatory variables in (3), $\boldsymbol{V}^*$, are continuous, the squared distance is given by (5). Thus, the distance matrix $\boldsymbol{D} = (\delta_{ij})$ is obtained, and

$$\boldsymbol{A} = -\frac{1}{2}\left[\text{diag}(\boldsymbol{V}^*\boldsymbol{V}^{*t})\mathbf{1}^t + \mathbf{1}\text{diag}(\boldsymbol{V}^*\boldsymbol{V}^{*t})^t - 2\boldsymbol{V}^*\boldsymbol{V}^{*t}\right]$$

where $\text{diag}(\boldsymbol{V}^*\boldsymbol{V}^{*t})$ is a vector which contains the diagonal terms of the matrix $\boldsymbol{V}^*\boldsymbol{V}^{*t}$. Therefore,

$$\boldsymbol{B} = \boldsymbol{H}\boldsymbol{A}\boldsymbol{H} = \boldsymbol{H}\boldsymbol{V}^*\boldsymbol{V}^{*t}\boldsymbol{H} = \boldsymbol{V}\boldsymbol{V}^t = \boldsymbol{X}\boldsymbol{X}^t$$

because $\boldsymbol{H}[\text{diag}(\boldsymbol{V}^*\boldsymbol{V}^{*t})\mathbf{1}^t]\boldsymbol{H} = \mathbf{0}$ and $\boldsymbol{H}[\mathbf{1}(\text{diag}(\boldsymbol{V}\boldsymbol{V}^t))^t]\boldsymbol{H} = \mathbf{0}$. Then, the geostatistical model based on distances introduced in (6) is a centered geostatistical model (3), i.e., it produces the same predictions at p dimensions than the model given in (3).

However, it would not be necessary to consider an Euclidean p-dimensional distance as given in Eq. (5). Let E_l ($k \leq l \leq n-1$) be the space spanned by the columns of $\boldsymbol{X}$, where $\boldsymbol{X}$ is a metric scaling solutions obtained from a distance applied to the same data. Then, by taking $k > p$, i.e., the most suitable columns of $\boldsymbol{X}$, the distance-based geostatistical model improves the classical geostatistical model when $(\boldsymbol{Z}_s - \widehat{\theta}_0\mathbf{1}) \in E_l$. Note that this is always true for $l = n-1$ with $l > p$.

To illustrate the comparison between the distance-based geostatistical model and the classical geostatistical model, let us focus on the R^2. The R_k^2 can be written as

$$R_k^2 = \sum_{j=1}^{k} r^2(\boldsymbol{Z}_s, \boldsymbol{X}_j)$$

Note that $R_k^2 < R_{k+1}^2$ because the principal coordinates $\boldsymbol{X}_j$'s ($j = 1, \ldots, k$) are uncorrelated with the principal coordinate $\boldsymbol{X}_{k+1}$. Moreover, R_k^2 is maximized for a given k provided that the first k ordered principal coordinates are selected.

On the other hand, we can write

$$R^2(\boldsymbol{Z}_s, \boldsymbol{V}^*) \geq \sum_{j=1}^{p} r^2\left(\boldsymbol{Z}_s, \boldsymbol{V}_j^*\right)$$

where $R^2(\boldsymbol{Z}_s, \boldsymbol{V}^*)$ is the coefficient of determination between $\boldsymbol{Z}_s$ and $\boldsymbol{V}^*$ in the classical geostatistical model, and $r^2(\boldsymbol{Z}_s, \boldsymbol{V}_j^*)$ is the correlation coefficient between $\boldsymbol{Z}_s$ and the j-th explanatory variable ($j = 1, \ldots, p$), $\boldsymbol{V}_j^*$.

Without loss of generality, suppose $p = 1$ and note that applying the distance-based geostatistical model, the R_l^2 can be written as

$$R_l^2 = \sum_{j=1}^{l} r^2(\boldsymbol{Z}_s, \boldsymbol{X}_j), \quad 1 \leq k \leq l \leq n-1$$

because the principal coordinates $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{n-1})$ are uncorrelated. Also, we can represent as $R^2(\boldsymbol{Z}_s, \boldsymbol{V}_1^*)$, the coefficient of determination between $\boldsymbol{Z}_s$ and the explanatory variable $\boldsymbol{V}_1^*$ using the classical geostatistical model. Therefore, if we take $l = k > 1$ (e.g., $k = 2$) then we find that

$$R_l^2 = \sum_{j=1}^{l} r^2(\boldsymbol{Z}_s, \boldsymbol{X}_j) \geq R^2\left(\boldsymbol{Z}_s, \boldsymbol{V}_1^*\right)$$

So, the distance-based geostatistical model improves the classical geostatistical model. However, this inequality is not too relevant if $R^2(\boldsymbol{Z}_s, \boldsymbol{V}_1^*)$ is close to 1, and so, the classical geostatistical model is sufficiently good, and thus the distance-based geostatistical model improvement is not necessary. This result holds when the explanatory variables are qualitative or mixed. This approach aims at improving the predictions by allowing for inclusion of any number of principal coordinates at the sample locations. Note that universal kriging traditionally considers only information of a matrix of size $n \times p$, $\boldsymbol{V}$. Therefore, when we want to improve predictions regardless of the effect that some covariates have on the spatial response variable, the methodology proposed in this paper can be used.

2.2 RBFs built from the distance-based drift

In spatial interpolation, there exist methods that do not require information from a spatial dependence model such as the variogram or covariogram, these are called deterministic and are the ones of interest in this section. Model (1), using a distance-based format, can be expressed in general form by

$$Z(\boldsymbol{s}_i) = g(\boldsymbol{s}_i) + \varepsilon(\boldsymbol{s}_i), \quad i = 1, \ldots, n \tag{10}$$

where $\varepsilon(\boldsymbol{s}_i)$ is the unobserved noise term which may or may not be included depending on the application, and it is assumed to be normal mutually independent with mean 0 and variance τ^2. In addition, $g(\boldsymbol{s}_i)$ is a real-valued function given by

$$\begin{aligned} g(\boldsymbol{s}_i) &= \sum_{l=0}^{k} \nu_l x_l(\boldsymbol{s}_i) + \kappa(\boldsymbol{s}_i) = \sum_{l=0}^{k} \nu_l x_l(\boldsymbol{s}_i) + \sum_{i'=1}^{n} \omega_{i'} \phi(\boldsymbol{s}_i - \boldsymbol{s}_{i'}) \\ &= \boldsymbol{x}^t(\boldsymbol{s}_i)\boldsymbol{\nu}_s + \boldsymbol{\phi}^t(\boldsymbol{s}_i)\boldsymbol{\omega}_s \end{aligned} \tag{11}$$

where $\kappa(\boldsymbol{s}_i) = \sum_{i'=1}^{n} \omega_{i'} \phi(\boldsymbol{s}_i - \boldsymbol{s}_{i'})$, $\boldsymbol{x}(\boldsymbol{s}_i) = (x_0(\boldsymbol{s}_i) = 1, x_1(\boldsymbol{s}_i), \ldots, x_k(\boldsymbol{s}_i))^t$ with each $x_l(\boldsymbol{s}_i)$ a linearly independent real-valued function in the location coordinates of $\boldsymbol{s}_i$, $\boldsymbol{\nu}_s = (\nu_0, \nu_1, \ldots, \nu_k)^t$ with each ν_l a trend model coefficient, $\boldsymbol{\phi}(\boldsymbol{s}_i) = (\phi(\boldsymbol{s}_i - \boldsymbol{s}_1), \ldots, \phi(\boldsymbol{s}_i - \boldsymbol{s}_n))^t$ with $\phi(\boldsymbol{s}_i - \boldsymbol{s}_{i'})$ a basis function, i.e., a scalar function of the Euclidean distance between $\boldsymbol{s}_i$ and $\boldsymbol{s}_{i'}$, and $\boldsymbol{\omega}_s = (\omega_1, \ldots, \omega_n)^t$, with ω_i an unknown weight.

Writing (11) in matrix form,

$$\boldsymbol{g}_s = \boldsymbol{X}_s \boldsymbol{\nu}_s + \boldsymbol{\Phi}_s \boldsymbol{\omega}_s \tag{12}$$

where $\boldsymbol{g}_s = (g(\boldsymbol{s}_1), \ldots, g(\boldsymbol{s}_n))^t$, $\boldsymbol{X}_s = (\mathbf{1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_k)$ is a $n \times (k+1)$ matrix with elements $\mathbf{1}$ and $\boldsymbol{X}_l = (x_l(\boldsymbol{s}_1), \ldots, x_l(\boldsymbol{s}_n))^t$, $l = 1, \ldots, k$, and $\boldsymbol{\Phi}_s$ is a $n \times n$ matrix with elements $\phi(\boldsymbol{s}_i - \boldsymbol{s}_{i'})$, $i, i' = 1, \ldots, n$.

The parameters $\boldsymbol{\nu}_s$ and $\boldsymbol{\omega}_s$ can be estimated by penalized least squares, minimizing the following expression

$$\sum_{i=1}^{n} [Z(\boldsymbol{s}_i) - g(\boldsymbol{s}_i))]^2 + \rho \int_{\mathbb{R}^2} J_m(\kappa(\boldsymbol{s}))\mathrm{d}\boldsymbol{s} \tag{13}$$

where $J_m(\kappa(\boldsymbol{s}))$ is a measure of roughness of the spline function κ (defined in terms of m-th degree derivatives of κ) when $\rho > 0$, which acts as a smoothing parameter. If $\rho = 0$, we can use RBFs such as EXP, GAU, MQ and IMQ.

Writing $m = 2$ and replacing (12) in (13), we obtain

$$\begin{aligned} L(\boldsymbol{\omega}_s, \boldsymbol{\nu}_s) &= (\boldsymbol{Z}_s - \boldsymbol{X}_s\boldsymbol{\nu}_s - \boldsymbol{\Phi}_s\boldsymbol{\omega}_s)^t (\boldsymbol{Z}_s - \boldsymbol{X}_s\boldsymbol{\nu}_s - \boldsymbol{\Phi}_s\boldsymbol{\omega}_s) + \rho \int_{\mathbb{R}^2} \left[\kappa''(\boldsymbol{s})\right]^2 \mathrm{d}\boldsymbol{s} \\ &= (\boldsymbol{Z}_s - \boldsymbol{X}_s\boldsymbol{\nu}_s - \boldsymbol{\Phi}_s\boldsymbol{\omega}_s)^t (\boldsymbol{Z}_s - \boldsymbol{X}_s\boldsymbol{\nu}_s - \boldsymbol{\Phi}_s\boldsymbol{\omega}_s) + \rho\boldsymbol{\omega}_s^t\boldsymbol{\Phi}_s\boldsymbol{\omega}_s \end{aligned}$$

where $\int_{\mathbb{R}^2} \left[\kappa''(\boldsymbol{s})\right]^2 \mathrm{d}\boldsymbol{s} = \int_{\mathbb{R}^2} \kappa''(\boldsymbol{s}) \left[\kappa''(\boldsymbol{s})\right]^t \mathrm{d}\boldsymbol{s} = \boldsymbol{\omega}_s^t\boldsymbol{\Phi}_s\boldsymbol{\omega}_s$ with $\kappa''(\boldsymbol{s}) = \sum_{i'=1}^{n} \omega_{i'} \phi''(\boldsymbol{s} - \boldsymbol{s}_{i'}) = [\boldsymbol{\phi}''(\boldsymbol{s})]^t\boldsymbol{\omega}_s$, $\boldsymbol{\Phi}_s = \int_{\mathbb{R}^2} \boldsymbol{\phi}''(\boldsymbol{s})[\boldsymbol{\phi}''(\boldsymbol{s})]^t \mathrm{d}\boldsymbol{s}$ and $\boldsymbol{\phi}''(\boldsymbol{s}) = (\phi''(\boldsymbol{s} - \boldsymbol{s}_1), \ldots, \phi''(\boldsymbol{s} - \boldsymbol{s}_n))^t$.

After differentiating with respect to the vectors $\boldsymbol{\omega}_s$ and $\boldsymbol{\nu}_s$, we find that $(\boldsymbol{\omega}_s, \boldsymbol{\nu}_s)$ are the solution of the following system of linear equations

$$\begin{pmatrix} \boldsymbol{\Phi}_s + \rho \boldsymbol{I} & \boldsymbol{X}_s \\ \boldsymbol{X}_s^t & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\omega}_s \\ \boldsymbol{\nu}_s \end{pmatrix} = \begin{pmatrix} \boldsymbol{Z}_s \\ \boldsymbol{0} \end{pmatrix}$$

where $\boldsymbol{I}$ is the $n \times n$ identity matrix and ρ can be interpreted as a white noise added to the variances at the data locations. If there were no trend, $\boldsymbol{X}_s$ becomes a vector of ones, and $\boldsymbol{\nu}_s$ a bias parameter.

We note that positive definiteness plays a central role with respect to the uniqueness of the coefficients in the interpolator $\hat{g}(\boldsymbol{s}_i)$. In order to generalize the interpolators, it is necessary to consider more general forms of positive definiteness.

Definition 1 Let $\boldsymbol{X}_0, \ldots, \boldsymbol{X}_k$ be linearly independent real-valued variables defined on $\mathbb{R}^d$, and $\boldsymbol{\Phi}_s$ a real symmetric matrix. Then, $\boldsymbol{\Phi}_s$ is positive definite with respect to $\boldsymbol{X}_0, \ldots, \boldsymbol{X}_k$ if and only if for all sets of points $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ in $\mathbb{R}^d$ we have that $\sum_{i=1}^{n} \sum_{i'=1}^{n} q_i q_{i'} \phi(\boldsymbol{s}_i - \boldsymbol{s}_{i'}) \geq 0$ for all q_i $(i = 1, \ldots, n)$, where q_i is a scalar (not all zero), and such that $\sum_{i'=1}^{n} x_l(\boldsymbol{s}_{i'}) q_{i'} = 0$ for $l = 1, \ldots, k$.

By Definition 1, $\boldsymbol{\Phi}_s$ is a positive definite matrix then $\boldsymbol{\Phi}_s + \rho \boldsymbol{I}$ is a non-singular matrix. Therefore, the weights $\boldsymbol{\omega}_s$ and the parameter vector $\boldsymbol{\nu}_s$ are estimated, respectively, by

$$\widehat{\boldsymbol{\omega}}_s = (\boldsymbol{\Phi}_s + \rho \boldsymbol{I})^{-1} \left\{ \boldsymbol{I} - \boldsymbol{X}_s \left[\boldsymbol{X}_s^t (\boldsymbol{\Phi}_s + \rho \boldsymbol{I})^{-1} \boldsymbol{X}_s \right]^{-1} \boldsymbol{X}_s^t (\boldsymbol{\Phi}_s + \rho \boldsymbol{I})^{-1} \right\} \boldsymbol{Z}_s$$

$$\widehat{\boldsymbol{\nu}}_s = \left[\boldsymbol{X}_s^t (\boldsymbol{\Phi}_s + \rho \boldsymbol{I})^{-1} \boldsymbol{X}_s \right]^{-1} \boldsymbol{X}_s^t (\boldsymbol{\Phi}_s + \rho \boldsymbol{I})^{-1} \boldsymbol{Z}_s$$

Several RBFs that are considered in this paper using the distance-based approach are presented in Table 1. The optimal smoothing parameter η, which is a parameter of free choice, is found by minimizing the root-mean-square prediction error (RMSPE) using cross-validation. Further descriptions of RBFs and their relationships to splines and kriging can be found in Bishop (1995, p. 164), Chilés and Delfiner (1999, p. 272), and Cressie (1993, p. 180).

2.3 Distance-based spatial prediction using RBFs

Having estimated the parameters $\boldsymbol{\nu}_s$ and $\boldsymbol{\omega}_s$, we are now ready to perform spatial prediction at a new location, $\boldsymbol{s}_0$, where a set of mixed explanatory variables are observed. Therefore, assume that on the set of mixed explanatory variables, a new individual $(n + 1)$ is observed, with known $\boldsymbol{v}(\boldsymbol{s}_0) = (v_1(\boldsymbol{s}_0), \ldots, v_p(\boldsymbol{s}_0))^t$. Then, the distances between the new individual and each of the individuals involved in model (1) are to be calculated, i.e., $\delta_{0i} = \delta\left(v(\boldsymbol{s}_0), v(\boldsymbol{s}_i)\right)$, $i = 1, \ldots, n$. From these distances, a prediction can be given using a result proposed by Cuadras and Arenas (1990) and Gower (1971), which relates the vector $\boldsymbol{\delta}_0 = \left(\delta_{01}^2, \ldots, \delta_{0n}^2\right)^t$ of squared distances with the vector $\boldsymbol{x}\left(\boldsymbol{s}_0\right) = \left(x_1\left(\boldsymbol{s}_0\right), \ldots, x_k\left(\boldsymbol{s}_0\right)\right)^t$ of principal coordinates associated with the new individual. This relation is given by

Table 1 Functional forms of some RBFs

RBF	Functional form
EXP	$\phi(\delta) = e^{-\eta\delta}, \quad \eta \neq 0$
GAU	$\phi(\delta) = e^{-\eta\delta^2}, \quad \eta \neq 0$
MQ	$\phi(\delta) = \sqrt{\eta^2 + \delta^2}, \quad \eta \neq 0$
IMQ	$\phi(\delta) = 1/\sqrt{\eta^2 + \delta^2}, \quad \eta \neq 0$
TPS	$\phi(\delta) = \begin{cases} (\eta \cdot \delta)^2 \log(\eta \cdot \delta) & \text{if } \delta \neq 0, \eta > 0 \\ 0 & \text{if } \delta = 0 \end{cases}$
CRS	$\phi(\delta) = \begin{cases} \ln\left(\eta\frac{\delta}{2}\right)^2 + E_1\left(\eta\frac{\delta}{2}\right)^2 + C_E & \text{if } \delta \neq 0 \\ 0 & \text{if } \delta = 0 \end{cases}$ where $\eta > 0$, ln is natural logarithm, $E_1(x)$ is the exponential integral function, and C_E is the Euler constant
ST	$\phi(\delta) = \begin{cases} \ln(\eta \cdot \delta/2) + K_0(\eta \cdot \delta) + C_E & \text{if } \delta \neq 0 \\ 0 & \text{if } \delta = 0 \end{cases}$ where $K_0(x)$ is the modified Bessel function and C_E is the Euler constant

$$\delta_{0i}^2 = (\boldsymbol{x}(\boldsymbol{s}_0) - \boldsymbol{x}(\boldsymbol{s}_i))^t (\boldsymbol{x}(\boldsymbol{s}_0) - \boldsymbol{x}(\boldsymbol{s}_i))$$

with $i = 1, \ldots, n$. Then, we have that

$$\boldsymbol{x}(\boldsymbol{s}_0) = \frac{1}{2}\boldsymbol{\Lambda}^{-1}\boldsymbol{X}^t(\boldsymbol{b} - \boldsymbol{\delta}_0)$$

where $\boldsymbol{b} = (b_{11}, \ldots, b_{nn})^t$ is a vector formed by the diagonal elements of $\boldsymbol{B}$, and $b_{ii} = \boldsymbol{x}(\boldsymbol{s}_i)^t\boldsymbol{x}(\boldsymbol{s}_i)$ with $i = 1, \ldots, n$.

The aim is to predict the value of $Z(\boldsymbol{s}_0)$ based on a set of observations $\boldsymbol{Z}_s^*$. For this, the RBF predictor is given by

$$\widehat{Z}(\boldsymbol{s}_0) = \widehat{g}(\boldsymbol{s}_0) = \sum_{i=1}^{n_h} \varphi_i Z(\boldsymbol{s}_i) = \boldsymbol{\varphi}^t \boldsymbol{Z}_s^* \tag{14}$$

subject to

$$\sum_{r=1}^{n_h} \varphi_r x_l(\boldsymbol{s}_r) = \boldsymbol{\varphi}^t \boldsymbol{X}_l = x_l(\boldsymbol{s}_0), \quad l = 0, \ldots, k$$

where n_h is the size of the neighborhood, $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_{n_h})^t$ and $\boldsymbol{Z}_s^* = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_{n_h}))^t$.

The expected error is equal to zero, i.e.,

$$\mathrm{E}\left(\hat{Z}(\boldsymbol{s}_0) - Z(\boldsymbol{s}_0)\right) = 0$$

and the mean square error of prediction by kriging, σ_K^2, using the approximation with a distance-based RBF is given by

$$\begin{aligned}\sigma_K^2(\boldsymbol{s}_0) &= \mathrm{E}\left\{\left[\hat{Z}(\boldsymbol{s}_0) - Z(\boldsymbol{s}_0)\right]^2\right\} \\ &\cong -\sum_{r=1}^{n_h}\sum_{r'=1}^{n_h}\varphi_r\varphi_{r'}\phi(\boldsymbol{s}_r - \boldsymbol{s}_{r'}) + 2\sum_{r=1}^{n_h}\varphi_r\phi(\boldsymbol{s}_r - \boldsymbol{s}_0) \\ &\cong -\boldsymbol{\varphi}^t\boldsymbol{\Phi}_{s_0}\boldsymbol{\varphi} + 2\boldsymbol{\varphi}^t\boldsymbol{\phi}_0\end{aligned}$$

where $\boldsymbol{\Phi}_{s_0}$ is a $n_h \times n_h$ matrix, $\boldsymbol{\phi}_0 = (\phi(\boldsymbol{s}_1 - \boldsymbol{s}_0), \ldots, \phi(\boldsymbol{s}_{n_h} - \boldsymbol{s}_0))^t$ corresponds to the radial function vector evaluated between the neighbors and the point where we aim to predict, i.e., $\phi(\boldsymbol{s}_r - \boldsymbol{s}_0)$, $r = 1, \ldots, n_h$.

The weights are determined minimizing the following penalized expression

$$\begin{aligned}l(\boldsymbol{\varphi}, \boldsymbol{\alpha}) = &\sum_{r=1}^{n_h}\sum_{r'=1}^{n_h}\varphi_r\varphi_{r'}\phi(\boldsymbol{s}_r - \boldsymbol{s}_{r'}) - 2\sum_{r=1}^{n_h}\varphi_r\phi(\boldsymbol{s}_r - \boldsymbol{s}_0) + \rho\int_{\mathbb{R}^2} J_m(\kappa(\boldsymbol{s}))\mathrm{d}\boldsymbol{s} \\ &+ 2\sum_{l=0}^{k}\alpha_l\left(\sum_{r=1}^{n_h}\varphi_r x_l(\boldsymbol{s}_r) - x_l(\boldsymbol{s}_0)\right)\end{aligned}$$

or equivalently, the previous expression becomes

$$l(\boldsymbol{\varphi}, \boldsymbol{\alpha}) = \boldsymbol{\varphi}^t\left(\boldsymbol{\Phi}_{s_0} + \rho\boldsymbol{I}_0\right)\boldsymbol{\varphi} - 2\boldsymbol{\varphi}^t\boldsymbol{\phi}_0 + 2\boldsymbol{\alpha}^t\left(\boldsymbol{X}_{s_0}^t\boldsymbol{\varphi} - \boldsymbol{x}_s(\boldsymbol{s}_0)\right)$$

where $\boldsymbol{I}_0$ is the $n_h \times n_h$ identity matrix, $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_k)^t$ is the vector of $k+1$ Lagrange multipliers associated with the unbiasedness constraint, $\boldsymbol{X}_{s_0}$ is similar to (12) using only the n_h observations, and $\boldsymbol{x}_s(s_0) = (x_0(s_0), \ldots, x_k(s_0))^t$.

After differentiating with respect to $\boldsymbol{\varphi}$ and $\boldsymbol{\alpha}$, equating the result to zero and performing some algebraic procedures, the following matrix system is found

$$\begin{pmatrix}\boldsymbol{\Phi}_{s_0} + \rho\boldsymbol{I}_0 & \boldsymbol{X}_{s_0} \\ \boldsymbol{X}_{s_0}^t & \boldsymbol{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{\varphi} \\ \boldsymbol{\alpha}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\phi}_0 \\ \boldsymbol{x}_s(\boldsymbol{s}_0)\end{pmatrix} \tag{15}$$

By Definition 1, $\boldsymbol{\Phi}_{s_0}$ is positive definite, then $\boldsymbol{\Phi}_{s_0} + \rho\boldsymbol{I}_0$ is invertible, and the system can solve with the coefficients $\boldsymbol{\varphi}$ and $\boldsymbol{\alpha}$ given by

$$\begin{aligned}\widehat{\boldsymbol{\varphi}}^t = &\left\{\boldsymbol{\phi}_0 + \boldsymbol{X}_{s_0}\left[\boldsymbol{X}_{s_0}^t(\boldsymbol{\Phi}_{s_0} + \rho\boldsymbol{I}_0)^{-1}\boldsymbol{X}_{s_0}\right]^{-1}\left[\boldsymbol{x}_s(\boldsymbol{s}_0) - \boldsymbol{X}_{s_0}^t(\boldsymbol{\Phi}_{s_0} + \rho\boldsymbol{I}_0)^{-1}\boldsymbol{\phi}_0\right]\right\}^t \\ &(\boldsymbol{\Phi}_{s_0} + \rho\boldsymbol{I}_0)^{-1} \\ \widehat{\boldsymbol{\alpha}} = &-\left[\boldsymbol{X}_{s_0}^t(\boldsymbol{\Phi}_{s_0} + \rho\boldsymbol{I}_0)^{-1}\boldsymbol{X}_{s_0}\right]^{-1}\left[\boldsymbol{x}_s(\boldsymbol{s}_0) - \boldsymbol{X}_{s_0}^t(\boldsymbol{\Phi}_{s_0} + \rho\boldsymbol{I}_0)^{-1}\boldsymbol{\phi}_0\right]\end{aligned} \tag{16}$$

Once we know $\boldsymbol{\alpha}$ and $\boldsymbol{\varphi}$, an approximate estimation for the mean square error of prediction can be written as

$$\widehat{\sigma}_K^2(\boldsymbol{s}_0) \cong \sum_{r=1}^{n_h} \widehat{\varphi}_r \phi(\boldsymbol{s}_r - \boldsymbol{s}_0) + \rho \sum_{r=1}^{n_h} \widehat{\varphi}_r^2 + \sum_{l=0}^{k} \widehat{\alpha}_l x_l(\boldsymbol{s}_0)$$

If we want to evaluate or compare adjustments between DBSRBFs, it is recommended to use cross-validation (leave-one-out). This criterium is now presented.

2.4 Cross-validation

We consider the root-mean-square prediction error (RMSPE) statistic to evaluate the DBSRBFs interpolation methods. Cross-validation (leave-one-out) is a technique used to evaluate the accuracy in the interpolation method. The method consists of removing one observation from the n sample points (usually related to a neighborhood), and then, with the remaining $n-1$ values and a selected RBF, we predict the variable value of study at the point location that was removed. This procedure is done sequentially with each sample point, and thus, a set of n prediction errors are obtained. If the DBSRBF model describes well the spatial structure, then the difference between the observed and predicted values should be small. This procedure is justified because the RBF interpolation methods are accurate, i.e., the predicted and observed values match at the sampled points. In this way, the cross-validation procedure gives an idea of how good the forecasts are, and provides information about which model gives more accurate predictions. The expression for the RMSPE is given by

$$\text{RMSPE} = \sqrt{\frac{\sum_{i=1}^{n} \left(\hat{Z}_{[i]}(\boldsymbol{s}_i) - Z(\boldsymbol{s}_i) \right)^2}{n}} \tag{17}$$

where $\hat{Z}_{[i]}(\boldsymbol{s}_i)$ is the predicted value obtained from cross-validation, and $Z(\boldsymbol{s}_i)$ is the sampled value at location $\boldsymbol{s}_i$.

A variation is given by dividing the sample into two subsamples; the first subsample is used for variogram modeling, and the other subsample is used to validate the kriging method. After that, validation measures can be constructed from the observed and predicted values (Bivand et al. 2008). If all goes well, the $RMSPE$ should be as small as possible (close to zero).

Finally, our procedure can be summarized as follows:

1. Obtain the principal coordinates from the spectral decomposition of the matrix of similarities (or distances) calculated from the explanatory variables.
2. Select the most correlated or significant principal coordinates with the regionalized variable, $\boldsymbol{Z}_s$. In this step, we recommend using the criterium given in (7) to make a first selection in order to leave out principal coordinates which are poorly correlated with the regionalized variable, and then, using the criteria (8) or (9) to choose the most significant principal coordinates using the DB regression.

Table 2 Scenarios considered in the simulated spatial experiments

Factor	Generic form
Noise level	$z_j(\boldsymbol{s}_i) = \beta_0 + \beta_1 V_i + \beta_2 D_{i2} + \beta_3 D_{i3} + f(w_{x_i}) + f(w_{y_i}) + f(w_{x_i}) f(w_{y_i}) + \sigma_j e(\boldsymbol{s}_i)$, with $\sigma_j = 0.02 + 0.04(j-1)^2$
Design density	$z_j(\boldsymbol{s}_i) = \beta_0 + \beta_1 V_i + \beta_2 D_{i2} + \beta_3 D_{i3} + f(X_{ji}) + f(Y_{ji}) + f(X_{ji}) f(Y_{ji}) + \sigma e(\boldsymbol{s}_i)$, where $\sigma = 0.1$, $X_{ji} = F_j^{-1}(X_i)$, $Y_{ji} = F_j^{-1}(Y_i)$
Spatial variation	$z_j(\boldsymbol{s}_i) = \beta_0 + \beta_1 V_i + \beta_2 D_{i2} + \beta_3 D_{i3} + f_j(w_{x_i}) + f_j(w_{y_i}) + f_j(w_{x_i}) f_j(w_{y_i}) + \sigma e(\boldsymbol{s}_i)$, where $\sigma = 0.2$, $f_j(l_i) = \sqrt{l_i(1-l_i)} \sin\left(\frac{2\pi\left[1+2^{(9-4j)/5}\right]}{l_i + 2^{(9-4j)/5}}\right)$
Variance function	$z_j(\boldsymbol{s}_i) = \beta_0 + \beta_1 V_i + \beta_2 D_{i2} + \beta_3 D_{i3} + f(w_{x_i}) + f(w_{y_i}) + f(w_{x_i}) f(w_{y_i}) + \sqrt{\upsilon_j(w_{x_i}) + \upsilon_j(w_{y_i}) + \upsilon_j(w_{x_i})\upsilon_j(w_{y_i})} e(\boldsymbol{s}_i)$ with $\upsilon_j(l_i) = \left[0.15\{1 + 0.4(2j-7)(l_i - 0.5)\}\right]^2$
Assumptions and other choices	
$V_i \sim Bi(n, 0.4)$, $i = 1, \ldots, 100$, $n = 50, 100, 150$; $X_i, Y_i \overset{iid}{\sim} Uniform\,(0,1)$; F_j is the $Beta\left(\frac{j+4}{5}, \frac{11-j}{5}\right)$ c.d.f., $j = 1, 3$; $f(l_i) = 1.5 f_1\left(\frac{l_i - 0.5}{0.15}\right) - f_1\left(\frac{l_i - 0.8}{0.04}\right)$ where $l_i = w_{x_i}$ or w_{y_i} and $f_1(u) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-u^2}{2}\right)$	

3. Optimize the parameters η and ρ of the distance-based spatial interpolator using RBFs by means of cross-validation (leave-one-out) from expression (17), and using (16) in different neighborhoods of a predetermined size. The size of the neighborhood, n_h, can be also chosen within the optimization process.
4. Make predictions at sampled and unsampled points to generate the prediction map using the DBSRBFs method, i.e., by mean of $\hat{Z}(\boldsymbol{s}_0) = \hat{\boldsymbol{\varphi}}^t \boldsymbol{Z}_s^*$.

3 Simulated experiments

This section describes a simulation study conducted to evaluate the efficiency of the proposed method under different conditions associated with smoothing parameters and RBFs using the distance-based method. The scenarios are designed considering those proposed in Wand (2000) with some modifications. In particular, we aim to study the effects of: (i) noise level, (ii) design density, (iii) degree of spatial variation, and (iv) noise variance function. These configurations and scenarios are presented in Table 2.

This study considers simulations in two dimensions (w_x, w_y), plus a Binomial random variable with $V_1 \sim Bi(n, 0.4)$, sample sizes $n = 50, 100, 150$, neighborhood sizes $n_h = 8, 16$, smoothing parameters $\eta = 0.01, 0.1$, and $j = 1, 3$ in the variance factor. Assume we have a nominal variable associated with three particular regions in the unit square, as shown in Fig. 1. Since there are three regions, only two dummy variables (D_2 and D_3) are considered to avoid problems of singularity. In addition, $e(s_i)$ follows a stationary isotropic Gaussian process with mean zero and covariance

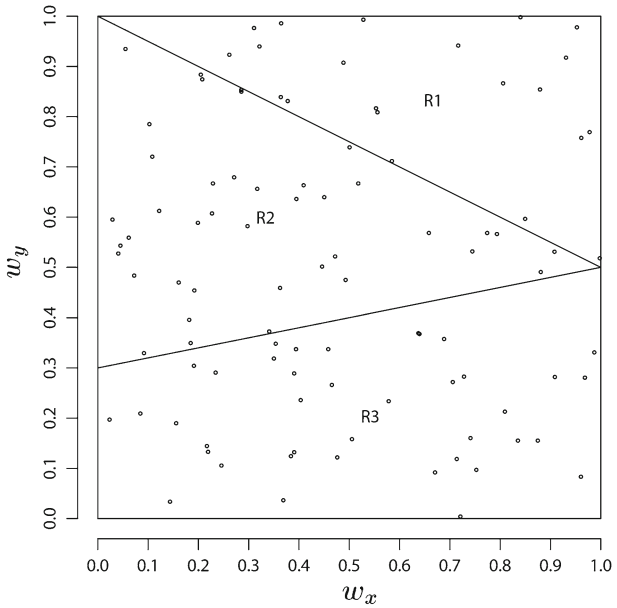

Fig. 1 Location of sampling points and associated regions defining the nominal variable

function generated from a Matérn variogram model with the following parameters: nugget ($\tau^2 = 0.1$), range ($\phi = 0.5$), partial sill ($\sigma^2 = 1$), kappa ($\kappa = 0$). For trend parameters, we assume the following values $\beta_0 = 10$, $\beta_1 = -4$, $\beta_2 = 2$ and $\beta_3 = -4$, with w_{x_i} and w_{y_i} associated with the spatial coordinates, where i is the i-th simulated observation. In Table 3, the simulated scenarios are presented. We use the rbf, rbf.cv and rbf.tcv functions in the geospt library of the R package (Melo et al. 2012). The proposed method is tested with five RBFs (MQ, TPS, CRS, ST, and EXP). A total of 120 scenarios were simulated, and for each of them the process was repeated 100 times.

For each simulated data set, we assessed the quality of the fit with the RMSPE, which was obtained by cross-validation (leave-one-out) method. The results are shown in Tables 4 and 5. We initially considered using a positive parameter ρ, but the values of RMSPE showed no significant differences with those obtained when $\rho = 0$, in particular when the multiquadratic, exponential and inverse multiquadratic functions were used.

Tables 4 and 5 show the average values of RMSPE from 100 simulations per case and for the 120 cases described in Table 4. The DBSRBF method worked well for large neighborhoods, indicating a gain (a decrease) of the RMSPE average values when $n_h = 32$ with respect to $n_h = 8$. Just on the contrary, for $n_h = 8$, when $j = 3$ the RMSPE average values were larger than those obtained when $j = 1$. However, for $n_h = 32$, the RMSPE average values are almost similar when $j = 1$ and $j = 3$. Taking into account parameter η, there was similar values in the RMSPE average when $\eta = 0.01$ compared to $\eta = 0.1$.

Table 3 Simulated scenarios

Model parameters			n	RBF				
η	j	n_h		MQ	TPS	CRS	ST	EXP
0.01	1	8	50	1	25	49	73	97
			100	2	26	50	74	98
			150	3	27	51	75	99
		32	50	4	28	52	76	100
			100	5	29	53	77	101
			150	6	30	54	78	102
	3	8	50	7	31	55	79	103
			100	8	32	56	80	104
			150	9	33	57	81	105
		32	50	10	34	58	82	106
			100	11	35	59	83	107
			150	12	36	60	84	108
0.1	1	8	50	13	37	61	85	109
			100	14	38	62	86	110
			150	15	39	63	87	111
		32	50	16	40	64	88	112
			100	17	41	65	89	113
			150	18	42	66	90	114
	3	8	50	19	43	67	91	115
			100	20	44	68	92	116
			150	21	45	69	93	117
		32	50	22	46	70	94	118
			100	23	47	71	95	119
			150	24	48	72	96	120

The natural numbers in the last five columns (from 1 to 120) represent the number of scenario

Analyzing generic forms, the lowest RMSPE average values corresponded to the noise level case; while spatial variance, design density and variance factor have RMSPE average values almost similar. In terms of the DBSRBF method, we found: (i) for the *noise level* case, the DBSRBF methods that produced the lowest RMSPE average were CRS and MQ, while TPS and ST showed the highest; (ii) for the *design density* case, almost all DBSRBF methods have the same RMSPE average, there is not a clear winner; (iii) in terms of *spatial variance*, the lowest RMSPE average value was observed with CRS when $n_h = 32$, but in general, all DBSRBF methods produce close values of RMSPE average; and (iv) the *variance factor* case with better results in terms of RMSPE averages was the CRS when $n_h = 32$, and again, all DBSRBF methods show similar values. Additionally, note that the RMSE averages decrease as sample size, n, increases.

Since the larger RMSPE average values were in those cases with neighborhood size $n_h = 8$, these cases where shown in separate boxplots to the cases with $n_h = 32$ (see Figs. 2, 3). According to these Figures for $n_h = 32$, the DBSRBF method that

Table 4 RMSPE averages under the DBSRBFs method for the scenarios presented in Table 3 (noise level and design density cases)

Parameter			n	Noise level					Design density				
η	j	n_h		MQ	TPS	CRS	ST	EXP	MQ	TPS	CRS	ST	EXP
0.01	1	8	50	0.07	0.16	0.14	0.15	0.07	0.17	0.17	0.18	0.17	0.17
			100	0.02	0.02	0.02	0.02	0.02	0.04	0.04	0.04	0.04	0.04
			150	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.02	0.02
		32	50	0.03	0.02	0.01	0.02	0.03	0.04	0.02	0.01	0.02	0.04
			100	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.02
			150	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.02
	3	8	50	0.13	0.15	0.14	0.15	0.13	0.18	0.26	0.26	0.26	0.19
			100	0.03	0.05	0.04	0.05	0.04	0.07	0.16	0.16	0.16	0.17
			150	0.01	0.02	0.01	0.02	0.01	0.03	0.04	0.04	0.04	0.04
		32	50	0.03	0.02	0.01	0.02	0.03	0.03	0.02	0.01	0.02	0.03
			100	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.02
			150	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.02
0.1	1	8	50	0.06	0.10	0.26	0.14	0.07	0.18	0.17	0.23	0.17	0.17
			100	0.02	0.02	0.05	0.02	0.02	0.04	0.04	0.04	0.04	0.04
			150	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.02	0.02
		32	50	0.03	0.02	0.00	0.02	0.03	0.04	0.03	0.00	0.02	0.04
			100	0.02	0.01	0.00	0.01	0.02	0.02	0.01	0.00	0.01	0.02
			150	0.01	0.01	0.00	0.01	0.02	0.01	0.01	0.00	0.01	0.02
	3	8	50	0.13	0.14	0.15	0.15	0.13	0.18	0.25	0.29	0.25	0.19
			100	0.04	0.05	0.07	0.05	0.04	0.11	0.14	0.20	0.16	0.17
			150	0.01	0.02	0.01	0.02	0.01	0.03	0.04	0.06	0.04	0.04
		32	50	0.03	0.03	0.00	0.02	0.03	0.03	0.03	0.00	0.02	0.03
			100	0.02	0.01	0.00	0.01	0.02	0.02	0.01	0.00	0.01	0.02
			150	0.01	0.01	0.00	0.01	0.02	0.02	0.01	0.00	0.01	0.02

showed the lowest variability was the CRS, while the others showed close behaviors. When $n_h = 8$, the DBSRBF method that showed the highest variability was the TPS, while the others methods showed similar variability. In general, all methods have less variability when $n_h = 32$ than when $n_h = 8$.

Analyzing Fig. 2, we note the following: (i) in terms of *noise level*, we found less variability when $j = 1$, the RMSPE average values were less for $j = 1$ than for $j = 3$; (ii) in case of the *design density*, the DBSRBF method that showed the highest variability was the TPS, while the others showed similar behaviors; (iii) under the *spatial variation* scenarios, when $j = 1$ and $n = 50$ a general large variability was generated under any particular case; and (iv) for the *variance factor*, no matter the j and η, the variability was higher when $n = 50$, and decreased when $n = 100$ and $n = 150$.

Table 5 RMSPE averages under the DBSRBFs method for the scenarios presented in Table 3 (spatial variation and variance factor cases)

Parameter			n	Spatial variation					Variance function				
η	j	n_h		MQ	TPS	CRS	ST	EXP	MQ	TPS	CRS	ST	EXP
0.01	1	8	50	0.16	0.15	0.16	0.15	0.16	0.13	0.17	0.16	0.16	0.13
			100	0.04	0.04	0.04	0.04	0.04	0.04	0.06	0.05	0.06	0.05
			150	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
		32	50	0.04	0.02	0.01	0.02	0.04	0.04	0.02	0.01	0.02	0.04
			100	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.02
			150	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.02
	3	8	50	0.16	0.19	0.19	0.19	0.17	0.13	0.15	0.14	0.15	0.13
			100	0.06	0.16	0.16	0.15	0.17	0.03	0.05	0.04	0.05	0.04
			150	0.03	0.04	0.03	0.04	0.04	0.02	0.02	0.02	0.02	0.02
		32	50	0.03	0.02	0.01	0.02	0.03	0.04	0.02	0.01	0.02	0.04
			100	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.02
			150	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.02
0.1	1	8	50	0.15	0.15	0.18	0.15	0.16	0.14	0.16	0.18	0.16	0.13
			100	0.04	0.04	0.04	0.04	0.04	0.05	0.06	0.08	0.06	0.05
			150	0.02	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02
		32	50	0.03	0.03	0.00	0.02	0.04	0.03	0.03	0.00	0.02	0.04
			100	0.02	0.01	0.00	0.01	0.02	0.02	0.01	0.00	0.01	0.02
			150	0.01	0.01	0.00	0.01	0.02	0.01	0.01	0.00	0.01	0.02
	3	8	50	0.17	0.17	0.22	0.18	0.17	0.14	0.14	0.15	0.14	0.13
			100	0.10	0.14	0.19	0.15	0.17	0.04	0.04	0.07	0.05	0.04
			150	0.03	0.03	0.05	0.03	0.04	0.02	0.02	0.02	0.02	0.02
		32	50	0.03	0.03	0.00	0.02	0.03	0.03	0.03	0.00	0.02	0.04
			100	0.02	0.01	0.00	0.01	0.02	0.02	0.01	0.00	0.01	0.02
			150	0.02	0.01	0.00	0.01	0.02	0.01	0.01	0.00	0.01	0.02

According to Fig. 3, we note that: (i) when considering the *noise level*, there was a larger variability when $j = 3$, slightly increasing the RMSPE average values, especially when using the MQ, TPS, ST and EXP RBFs; (ii) in terms of the *design density*, the CRS RBF showed the lowest variability: for MQ, TPS, ST and EXP RBFs, the variabilities were similar; (iii) and (iv) when the *spatial variation* and the *variance factor* were taken into account, the largest variability was for $n = 50$, slightly larger for MQ, TPS, ST and EXP RBFs than for CRS basis function: the RMSPE average values slightly increased for $j = 3$ in comparation with $j = 1$.

In general terms, DBSRBF methods were robust enough against different sample sizes because the variabilities were similar and homogeneous under all the scenarios analyzed.

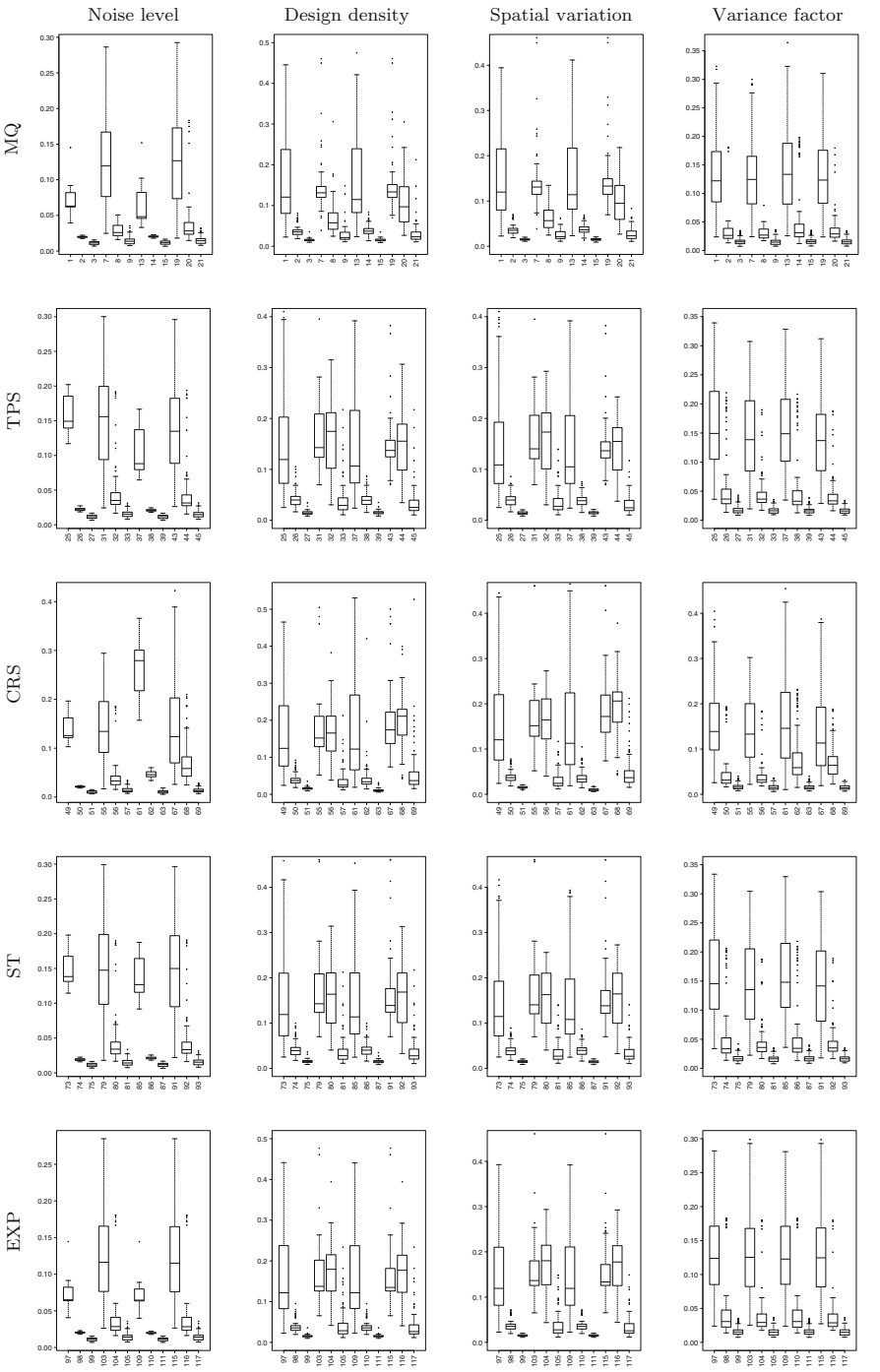

Fig. 2 RMSPE values for the spatial simulation scenarios using DBSRBFs method when the number of neighbors is $n_h = 8$

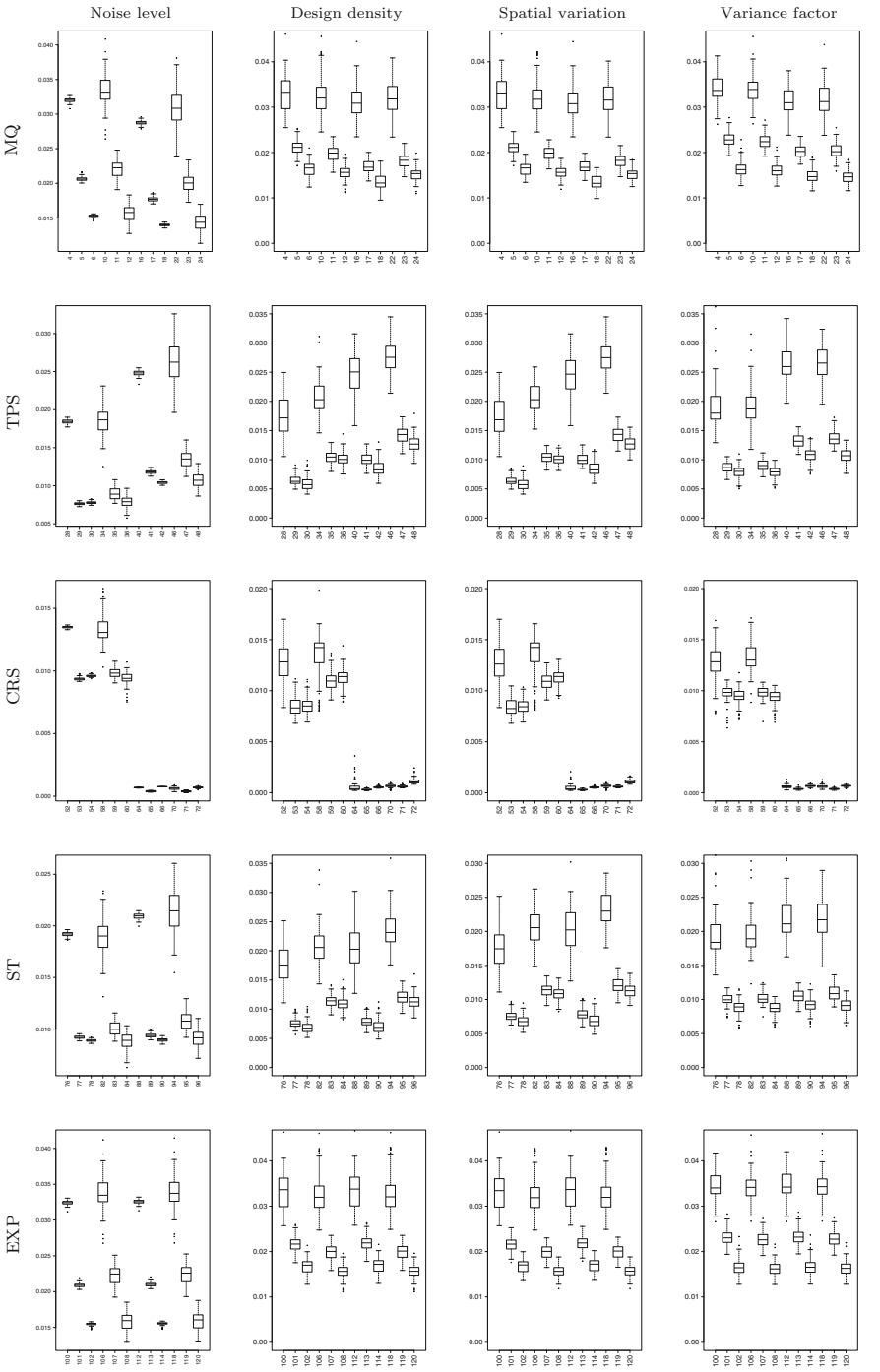

Fig. 3 RMSPE values for the spatial simulation scenarios using DBSRBFs method when the number of neighbors is $n_h = 32$

4 Application

We analyzed a data set in which soil samples were collected with a Dutch type drill on an incomplete regular lattice at a spacing of approximately 50 m, with geographic coordinates (northings and eastings) 900 m apart in both directions. The soil samples were taken from 0 to 20 cm depth layer at each of the 178 locations (Fig. 4a). Magnesium and calcium content were measured in $\text{mmol}_c/\text{dm}^3$, although we focus here only on the calcium content (ca20). The study region was divided into three subregions which have experienced different soil management regimes. This characterization is ideal for the application of our proposed method. We also had additional information on the altitude and subregion code of each sample, which is associated with three periods of fertilization in different places or areas. The data were taken from Capeche et al. (1997), and the main objective of the study was adequate land use planning that allows rational and sustainable management, avoiding the erosion process. This is important in order to allocate subsidies in the experimental fields to perform searches that will be extrapolated to similar soils and climatic zones.

Table 6 contains two times the log-likelihood ($2 \log L$) for both the proposed method (DBSRBF) and the classical one. The explanatory variables taken in the classical method were soil type, altitude and linear spatial trend (spatial coordinates). In the DBSRBF case, the principal coordinates were built using the above explanatory variables. For the selection of the principal coordinates employing the DB regression, we use the criteria given in (7) to make an initial selection in order to remove the principal coordinates poorly correlated with the regionalized variable $\boldsymbol{Z}_s$, and then we use the criterion (8) to choose the most significant principal coordinates. At this point, we do not find differences between the three methods for the selection of the principal components presented in Section 2. The results presented Table 6 show a steady increase in $2 \log L$ when increasing the number of principal coordinates. There is a significant gain as $2 \log L$ goes from -1272.03 to -1178.80 when the number of principal coordinates goes from 0 to 18. The value of $2 \log L$ for the classical case is -1259.17, which is lower than the values obtained by the DB method. Based on this, and to obtain maps of calcium content, we thus used seventeen principal coordinates in the detrending because the other principal coordinates were not significant at 5% level using criterion (8), and finally, a cross-validation (leave-one-out) method was performed selecting the DBSRBF method that showed the lowest RMSPE.

Additionally, we did a comparison of our proposal (DBSRBF method) with the case of a covariate entering directly in the model and adding a RBF to take care of the remaining spatial pattern (classic method). The cross-validation results are shown in Table 7, where it is shown that the best radial functions are the IMQ, EXP and CRS showing the smallest RMSPE values. Further, the proposed method shows, under this application, some advantages in reducing the error independently of which DBSRBF is used because the RMSPE is smaller compared to the case when we are not using the distance-based method. As a practical detail, the optimization functions used for ρ and η were "bobyqa" and "optim" from "minqa" and "stats" packages, respectively, from R Development Core Team (2016) Software, depending on which method provided stable results. This information is also shown in Table 7. The corresponding prediction maps are shown in Figs. 4b–h and 5a–g.

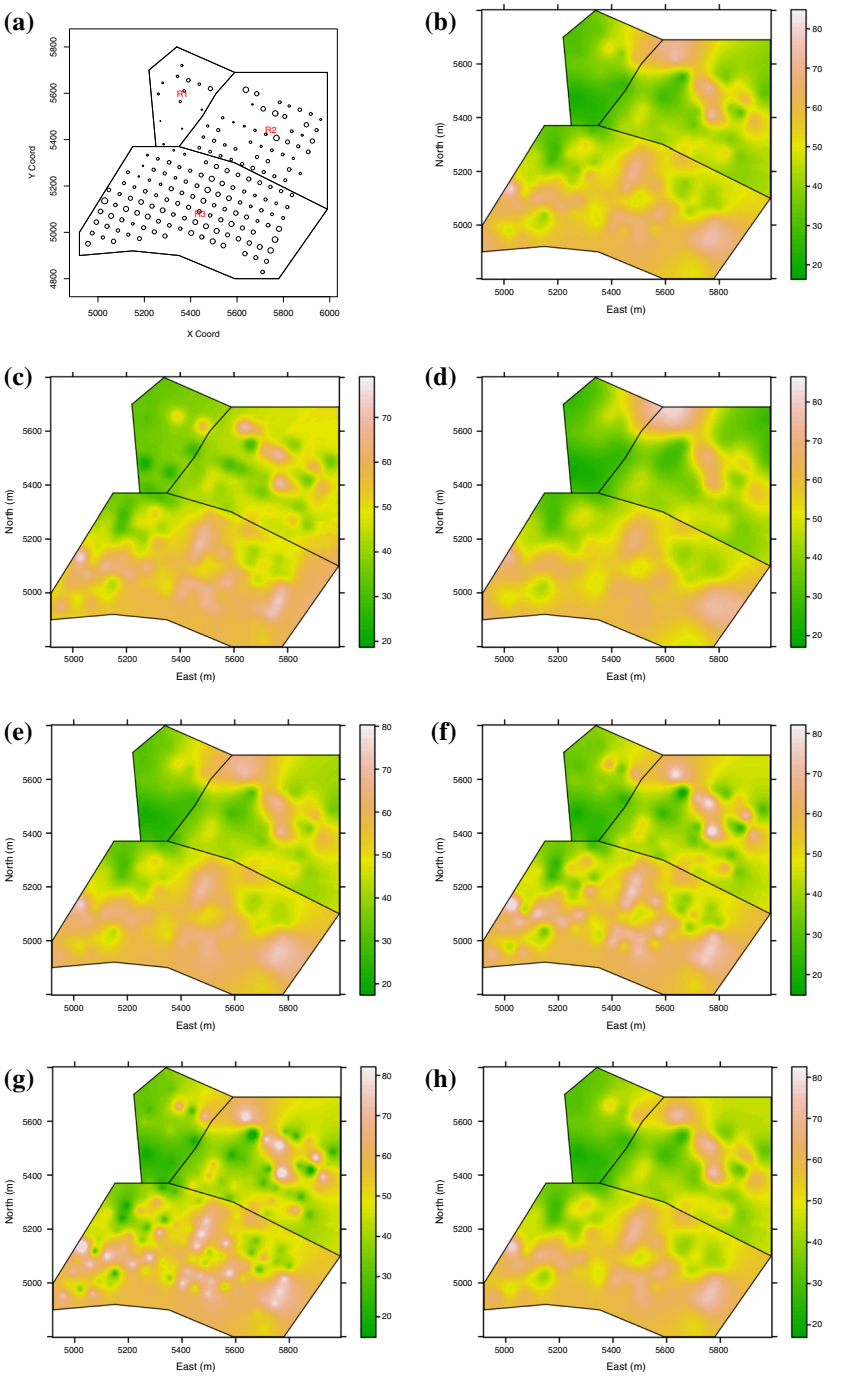

Fig. 4 Sampling locations and prediction maps under classic method for soil calcium content including subregion (soil type). **a** *Circle plot* of calcium content with lines delimiting subregions, **b** MQ, **c** GAU, **d** TPS, **e** EXP, **f** CRS, **g** ST, **h** IMQ

Table 6 Comparison between the DBSRBF method and the classical method using maximum likelihood

	Number Principal coordinates trend DBSRBF						Classic
	1	2	4	8	17	18	
$2 \log L$	−1272.03	−1260.00	−1253.61	−1236.01	−1193.20	−1178.80	−1259.17

Table 7 Comparison between classic and DB of several spatial RBF methods for the calcium content using ordinary cross-validation

METHOD	RBF	Optimization					
		optim			bobyqa		
		η	RMSPE	η	ρ	iter	RMSPE
CLASSIC	EXP	0.00	7.69	0.00	0.00	207	7.68
	GAU	0.01	11.32	0.01	0.00	29	11.32
	MQ	4.52	7.73	4.52	0.00	55	7.73
	IMQ	38.85	7.69	38.85	0.00	74	7.69
	TPS	0.00	7.78	0.00	0.10	520	7.71
	CRS	0.08	7.67	0.14	1.07	1318	7.66
	ST	0.07	7.69	0.24	1.39	611	7.66
DBSRBF	EXP	0.03	7.32	0.03	0.00	110	7.12
	GAU	0.20	8.01	0.20	0.00	39	7.08
	MQ	0.00	7.38	0.00	0.00	19	7.38
	IMQ	27.36	7.30	25.36	0.00	44	7.13
	TPS	0.81	7.62	0.20	1.24	23	7.61
	CRS	0.10	7.33	0.40	2.41	48	7.20
	ST	0.08	7.34	1.00	2.40	35	7.20

In Figs. 4b–h and 5a–g, we observe that in the classical method, the MQ, IMQ, TPS and EXP functions present a similar behavior on the calcium content in the analyzed region. Another group of functions that show a similar result are the one formed by the CRS and ST (these demand a lot computational time), and the Gaussian shows a slightly different behavior to the other RBF. For the DBSRBF method, we find that the MQ and TPS functions generate a similar prediction surface, different from the result from the others RBFs, whose results are similar.

5 Conclusions

Modeling uncertainty for regionalized variables as a function of mixed variables (continuous, binary or categorical) can be relevant in many disciplines of the geosciences and environmental areas. The proposed method shows, both under simulations and applications, some advantages in reducing the error depending from which DBSRBF

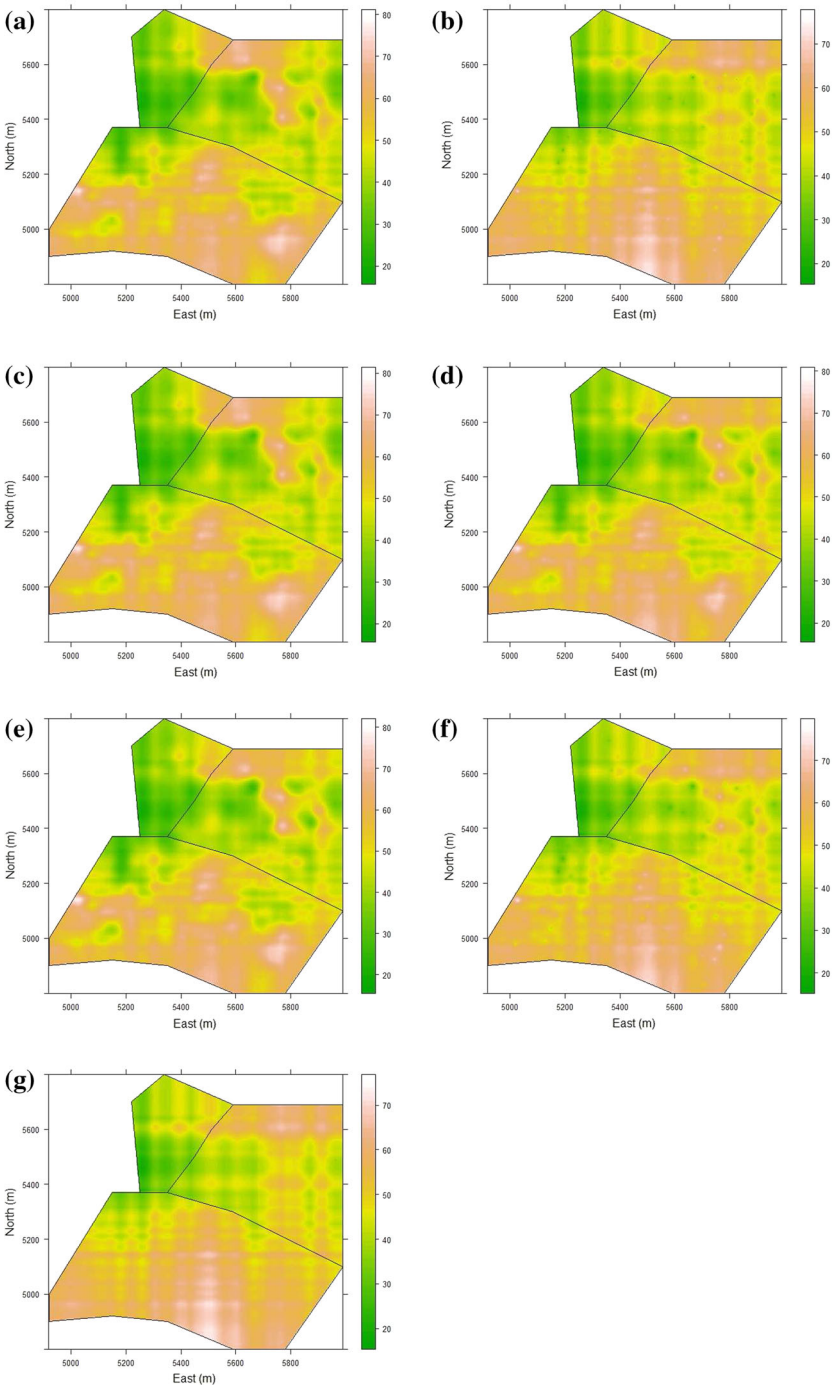

Fig. 5 Prediction maps under the DBSRBFs method for soil calcium content including subregion (soil type). **a** MQ, **b** GAU, **c** TPS, **d** EXP, **e** CRS, **f** ST, **g** IMQ

is employed. The kriging structural analysis minimized by DBSRBF splines is demonstrated by simulations. It is suggested that the cross-validation calculated for splines may be a more reliable measure of overall prediction error. In general, by increasing the number of principal coordinates, the DBRBF method produces good estimates of the regionalized variable. Then, this method is very flexible and gives good results in practice.

In a variety of studies, detection of variability among areas is quite a difficult task, and so the proposed DBSRBF method is expected to be useful as it does not despise the existing information. Although the correlation between an explanatory variable and the response variable can be low, the key point in the proposed method is the correlation between the principal coordinates and the response variable. Therefore, the DBSRBF method can produce better estimates of the regionalized variable if the number of principal coordinates is increased. This is possible by including more significative principal coordinates than explanatory variables in the trend of DBSRBF proposed model.

Acknowledgements Work partially funded and supported by: Grant MTM2016-78917-R from the Spanish Ministry of Science and Education; Core Spatial Data Research (Faculty of Engineering, Francisco José de Caldas District University) (Grant COL0013969); and Applied Statistics in Experimental Research, Industry and Biotechnology (Universidad Nacional de Colombia) (Grant COL0004469).

References

Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford Press, Oxford (1995)

Bivand, R., Pebesma, E., Rubio, V.: Applied Spatial Data Analysis with R. Springer, New York (2008)

Capeche, C. L., Macedo, J. R., Manzatto, H. R. H., Silva, E. F.: Caracterização pedológica da fazenda Angra - PESAGRO/RIO, Technical report, Estação experimental de Campos (RJ). In Informação, globalização, uso do solo, Rio de Janeiro (1997)

Chilès, J.P., Delfiner, P.: Geostatistics: Modeling Spatial Uncertainty. Wiley, New York (1999)

Cressie, N.: Geostatistics. The American Statistician **43**, 197–202 (1989)

Cressie, N.: Statistics for Spatial Data, Revised edn. Wiley, New York (1993)

Cuadras, C.M.: Distance analysis in discrimination and classification using both continuous and categorical variables. In: Dodge, Y. (ed.) Recent Developments in Statistical Data Analysis and Inference, pp. 459–474. Elsevier, North-Holland (1989)

Cuadras, C.M.: Interpreting an inequality in multiple regression. The American Statistician **47**(4), 256–258 (1993)

Cuadras, C.M., Arenas, C.: A distance based regression model for prediction with mixed data. Commun. Stat. A Theory Methods **19**, 2261–2279 (1990)

Cuadras, C.M., Arenas, C., Fortiana, J.: Some computational aspects of a distance-based model for prediction. Commun. Stat. Simul. Comput. **25**(3), 593–609 (1996)

Duchon, J.: Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. Rairo Anal. Numer. **10**, 5–12 (1976)

Franke, R.: Smooth interpolation of scattered data by local thin plate splines. Comput. Math. Appl. **8**, 273–281 (1982)

Gower, J.: Adding a point to vector diagrams in multivariate analysis. Biometrika **55**, 582–585 (1968)

Gower, J.: A general coefficient of similarity and some of its properties. Biometrics **27**, 857–871 (1971)

Hardy, R., Gopfert, W.: Least squares prediction of gravity anomalies, geodial undulations, and detections of the vertical with multiquadric harmonic functions. Geophys. Res. Lett. **2**, 423–426 (1975)

Heryudono, A., Driscoll, T.: Radial basis function interpolation on irregular domain through conformal transplantation. J. Sci. Comput. **44**(3), 286–300 (2010)

Mardia, K.V., Kent, J.T., Bibby, J.M.: Multivariate Analysis. Academic Press Inc, London (2002)

Melo, C.E., Santacruz, A., Melo, O.O.: geospt: an R package for spatial statistics. R package version 1.0-2. http://geospt.r-forge.r-project.org/ (2012)
Mitáš, L., Mitášová, H.: General variational approach to the interpolation problem. Comput. Math. Appl. **16**, 983–992 (1988)
Mitášová, H.: Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry analysis. Math. Geol. **25**, 657–669 (1993)
Mitášová, H., Mitáš, L.: Interpolation by regularized spline with tension: I. Theory and implementation. Math. Geol. **25**, 641–655 (1993)
Myers, D.: Kriging, cokriging, radial basic functions and the role of positive definiteness. Comput. Math. Appl. **24**, 139–148 (1992)
R Development Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/ (2016)
Schagen, I.P.: Interpolation in two dimensions: a new technique. J. Inst. Math. Appl. **23**, 53–59 (1979)
Späth, H.: Exponential spline interpolation. Computing **4**, 225–233 (1969)
Thiébaux, H., Pedder, M.: Spatial Objective Analysis: With Applications in Atmospheric Science. Academic Press, London (1987)
Wand, M.: A comparison of regression spline smoothing procedures. Comput. Stat. **15**, 443–462 (2000)
Yavuz, H., Erdogan, S.: Spatial analysis of monthly and annual precipitation trends in Turkey. Water Resour. Manag. **26**(3), 609–621 (2012)
Zhang, G.: Smoothing splines using compactly supported, positive definite, radial basis functions. Comput. Stat. **27**(3), 573–584 (2012)