

About regression-kriging: From equations to case studies[☆]

Tomislav Hengl^{a,*}, Gerard B.M. Heuvelink^b, David G. Rossiter^c

^aEuropean Commission, Institute for Environment and Sustainability, TP 280, Via E. Fermi 1, I-21020 Ispra (VA), Italy

^bSoil Science Centre, Wageningen University and Research Centre, P.O. Box 37, 6700 AA Wageningen, The Netherlands

^cInternational Institute for Geo-Information Science and Earth Observation (ITC), P.O. Box 6, 7500 AA Enschede, The Netherlands

Received 7 February 2006; accepted 25 May 2007

Abstract

This paper discusses the characteristics of regression-kriging (RK), its strengths and limitations, and illustrates these with a simple example and three case studies. RK is a spatial interpolation technique that combines a regression of the dependent variable on auxiliary variables (such as land surface parameters, remote sensing imagery and thematic maps) with simple kriging of the regression residuals. It is mathematically equivalent to the interpolation method variously called “Universal Kriging” (UK) and “Kriging with External Drift” (KED), where auxiliary predictors are used directly to solve the kriging weights. The advantage of RK is the ability to extend the method to a broader range of regression techniques and to allow separate interpretation of the two interpolated components. Data processing and interpretation of results are illustrated with three case studies covering the national territory of Croatia. The case studies use land surface parameters derived from combined Shuttle Radar Topography Mission and contour-based digital elevation models and multitemporal-enhanced vegetation indices derived from the MODIS imagery as auxiliary predictors. These are used to improve mapping of two continuous variables (soil organic matter content and mean annual land surface temperature) and one binary variable (presence of yew). In the case of mapping temperature, a physical model is used to estimate values of temperature at unvisited locations and RK is then used to calibrate the model with ground observations. The discussion addresses pragmatic issues: implementation of RK in existing software packages, comparison of RK with alternative interpolation techniques, and practical limitations to using RK. The most serious constraint to wider use of RK is that the analyst must carry out various steps in different software environments, both statistical and GIS.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Spatial prediction; Multiple regression; GSTAT; Environmental predictors; SRTM; MODIS

1. Introduction

In recent years, there has been an increasing interest in hybrid interpolation techniques which combine two conceptually different approaches to modelling and mapping spatial variability: (a) interpolation relying solely on point observations of the target variable; and (b) interpolation based on regression of the target variable on spatially exhaustive auxiliary information. Several studies have shown that

[☆]Supplementary information: A practical guide to regression-kriging in various software packages is available at <http://spatial-analyst.net>

*Corresponding author. Tel.: +39 0332 785535; fax: +39 0332 786394.

E-mail addresses: tomislav.hengl@jrc.it (T. Hengl), gerard.heuvelink@wur.nl (G.B.M. Heuvelink), rossiter@itc.nl (D.G. Rossiter).

hybrid techniques can give better predictions than either single approach (Knotters et al., 1995; Bishop and McBratney, 2001; Bourennane and King, 2003; Lloyd, 2005; Yemefack et al., 2005). These interpolators are currently used in a variety of applications, ranging from modelling spatial variability in tropical rainforest soils (Yemefack et al., 2005), soil mapping (Lopez-Granados et al., 2005; Leopold et al., 2005), mapping of leaf area index (LAI) using Landsat ETM+ data (Berterretche et al., 2005), modelling spatial distribution of human diseases (Pleydell et al., 2004), mapping water table (Desbarats et al., 2002; Finke et al., 2004), mapping abundance of fish in the ocean (Rivoirard, 2002), mapping rainfall over Great Britain (Lloyd, 2005), and mapping rainfall erosivity in the Algarve region of Portugal (Goovaerts, 1999).

One of these hybrid interpolation techniques is known as regression-kriging (RK) (Odeh et al., 1995; Hengl et al., 2004b). It first uses regression on auxiliary information and then uses simple kriging (SK) with known mean (0) to interpolate the residuals from the regression model. This allows the use of arbitrarily-complex regression methods, including generalized linear models. In spite of this and other attractive properties of RK, it is not as widely used in geosciences as might be expected. This paper presents the theory behind RK, explains it with a simple example, and demonstrates its utility with three diverse case studies. It also shows how RK can be implemented in today's GIS and statistical computing environments, identifies barriers to its wider use, and proposes solutions for these.

2. Theory

2.1. Regression-kriging

In the pure *geostatistical* approach, predictions are commonly made by calculating some weighted average of the observations (Webster and Oliver, 2001, p. 38):

$$\hat{z}(s_0) = \sum_{i=1}^n \lambda_i \cdot z(s_i), \quad (1)$$

where $\hat{z}(s_0)$ is the predicted value of the target variable at an unvisited location s_0 given its map coordinates, the sample data $z(s_1), z(s_2), \dots, z(s_n)$, and their coordinates. The weights λ_i are chosen such that the prediction error variance is minimized,

yielding weights that depend on the spatial autocorrelation structure of the variable. This interpolation procedure is popularly known as *ordinary kriging* (OK).

An alternative to kriging is the *regression* approach, which makes predictions by modelling the relationship between the target and auxiliary environmental variables at sample locations, and applying it to unvisited locations using the known value of the auxiliary variables at those locations. Common auxiliary environmental predictors are land surface parameters, remote sensing images, and geological, soil, and land-use maps (McKenzie and Ryan, 1999). A common regression approach is linear multiple regression (Draper and Smith, 1981; Christensen, 1996), where the prediction is again a weighted average, this time of the predictors:

$$\hat{z}(s_0) = \sum_{k=0}^p \hat{\beta}_k \cdot q_k(s_0); \quad q_0(s_0) \equiv 1, \quad (2)$$

where $q_k(s_0)$ are the values of the auxiliary variables at the target location, $\hat{\beta}_k$ are the estimated regression coefficients and p is the number of predictors or auxiliary variables. (To avoid confusion with geographical coordinates, we use the symbol q , instead of the more common x , to denote a predictor.)

RK combines these two approaches: regression is used to fit the explanatory variation and SK with expected value 0 is used to fit the residuals, i.e. unexplained variation (Hengl et al., 2004b):

$$\begin{aligned} \hat{z}(s_0) &= \hat{m}(s_0) + \hat{e}(s_0) \\ &= \sum_{k=0}^p \hat{\beta}_k \cdot q_k(s_0) + \sum_{i=1}^n \lambda_i \cdot e(s_i), \end{aligned} \quad (3)$$

where $\hat{m}(s_0)$ is the fitted drift, $\hat{e}(s_0)$ is the interpolated residual, $\hat{\beta}_k$ are estimated drift model coefficients ($\hat{\beta}_0$ is the estimated intercept), λ_i are kriging weights determined by the spatial dependence structure of the residual and where $e(s_i)$ is the residual at location s_i . The regression coefficients $\hat{\beta}_k$ are estimated from the sample by some fitting method, e.g. ordinary least squares (OLS) or, optimally, using generalized least squares (GLS), to take the spatial correlation between individual observations into account (Cressie, 1993, p. 166):

$$\hat{\beta}_{\text{GLS}} = (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{z}, \quad (4)$$

where $\hat{\beta}_{\text{GLS}}$ is the vector of estimated regression coefficients, \mathbf{C} is the covariance matrix of the

residuals, \mathbf{q} is a matrix of predictors at the sampling locations, and \mathbf{z} is the vector of measured values of the target variable. Once the trend has been estimated the residual can be interpolated with kriging and added to the estimated trend. In matrix notation, this is written as (Christensen, 2001, p. 277):

$$\hat{z}(s_0) = \mathbf{q}_0^T \cdot \hat{\boldsymbol{\beta}}_{\text{GLS}} + \boldsymbol{\lambda}_0^T \cdot (\mathbf{z} - \mathbf{q} \cdot \hat{\boldsymbol{\beta}}_{\text{GLS}}), \quad (5)$$

where $\hat{z}(s_0)$ is the predicted value at location s_0 , \mathbf{q}_0 is the vector of $p + 1$ predictors, and $\boldsymbol{\lambda}_0$ is the vector of n kriging weights used to interpolate the residuals. This prediction model has an error that reflects the position of new locations (extrapolation) in both geographical and feature space:

$$\begin{aligned} \sigma_{\text{RK}}^2(s_0) = & (C_0 + C_1) - \mathbf{c}_0^T \cdot \mathbf{C}^{-1} \cdot \mathbf{c}_0 \\ & + (\mathbf{q}_0 - \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{c}_0)^T \cdot (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \\ & \cdot (\mathbf{q}_0 - \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{c}_0), \end{aligned} \quad (6)$$

where $C_0 + C_1$ is the sill variation and \mathbf{c}_0 is the vector of covariances of residuals at the unvisited location.

The estimation of the residuals is an iterative process: first the drift model is estimated using OLS, then the covariance function of the residuals is used to obtain the GLS coefficients. Next, these are used to re-compute the residuals, from which an updated covariance function is computed, and so on. Although this is by many geostatisticians recommended as the proper procedure, Kitaniidis (1994) showed that use of the covariance function derived from the OLS residuals (i.e. a single iteration) is often satisfactory, because it is not different enough from the function derived after several iterations to affect the kriging interpolation.

The geostatistical literature uses many different terms for what are essentially the same or at least very similar techniques. All these follow the universal kriging (UK) model that was introduced by Matheron (1969) and that is by many statisticians considered to be the (only) best linear unbiased prediction model of spatial data (Christensen, 2001, Section 6). Originally, UK was intended as a generalized case of kriging where the trend is modelled as a function of coordinates, within the kriging system. Thus, many authors (Deutsch and Journel, 1998; Wackernagel, 1998; Papritz and Stein, 1999) reserve the term *Universal Kriging* for this case. If the drift is defined externally as a linear function of some auxiliary variables,

rather than the coordinates, the term *Kriging with External Drift* (KED) is preferred (Wackernagel, 1998; Chiles and Delfiner, 1999, p. 355). In the case of UK or KED, the predictions are made as with kriging, with the difference that the covariance matrix of residuals is extended with the auxiliary predictors $q_k(s_i)$'s (Webster and Oliver, 2001, p. 183). However, the drift and residuals can also be estimated separately and then summed. This procedure was suggested by Ahmed and de Marsily (1987) and Odeh et al. (1995) later named it *regression-kriging*, while Goovaerts (1997, Section 5.4) uses the term *Kriging with a trend model* to refer to a family of interpolators and refers to RK as *simple kriging with varying local means*. KED and RK differ in the computational steps used, however, the resulting predictions and prediction variances are the same, given the same point set, auxiliary variables, regression functional form, and regression fitting method. The mathematical proof is given in the Appendix.

Although the KED seems, at first glance, to be computationally more straightforward than RK, the variogram parameters for KED must also be estimated from regression residuals, thus requiring a separate regression modelling step. This regression should be GLS because of the likely spatial correlation between residuals. Note that many analysts use instead the OLS residuals, which may not be too different from the GLS residuals (Yemefack et al., 2005). However, they are not optimal if there is any spatial correlation, and indeed they may be quite different in the case of highly correlated, clustered sample points. Also a limitation of KED is the instability of the extended matrix in the case that the covariate does not vary smoothly in space (Goovaerts, 1997, p. 195). RK has the advantage that it explicitly separates trend estimation from residual interpolation, allowing the use of arbitrarily complex forms of regression, rather than the simple linear techniques that can be used with KED. In addition, it allows the separate interpretation of the two interpolated components. For these reasons we advocate the use of the term *regression-kriging over universal kriging*. Hence, RK is a more descriptive synonym of the same generic interpolation method.

2.2. A simple example of regression-kriging

The next section illustrates how RK computations work and compares it to OK using the

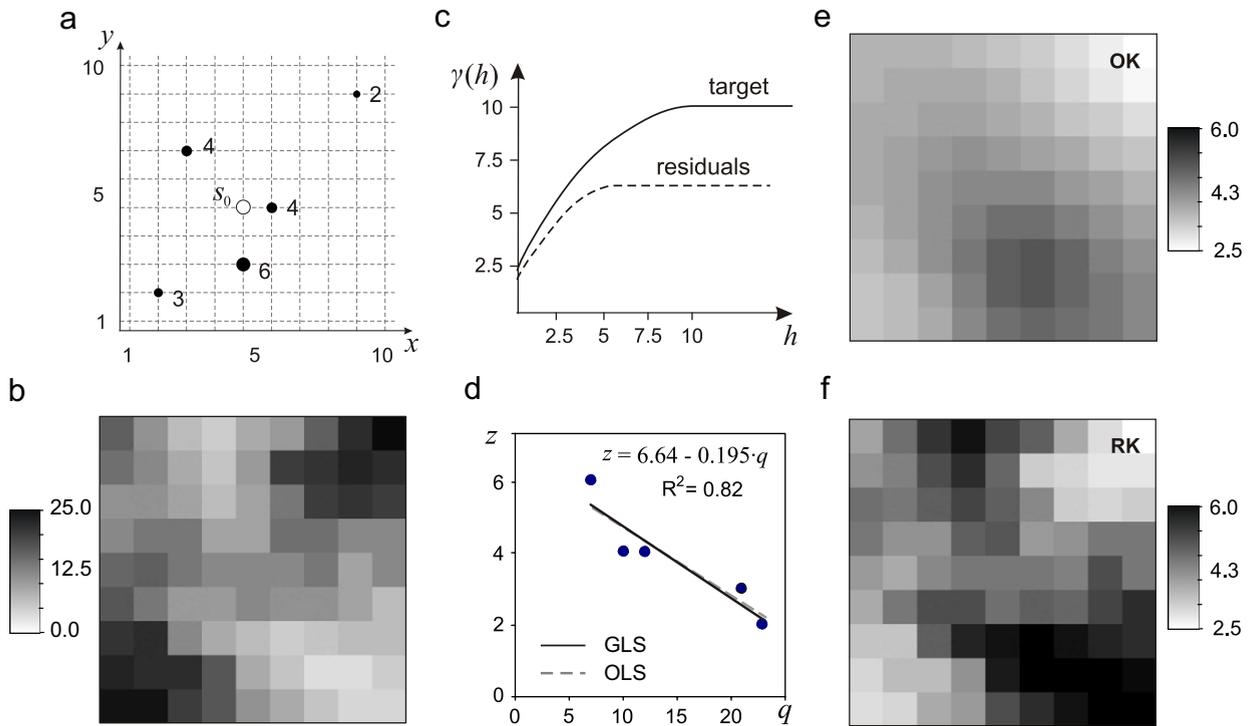


Fig. 1. Comparison of ordinary kriging and regression-kriging using a simple example with 5 points (Burrough and McDonnell, 1998, pp. 139–141): (a) location of points and unvisited site; (b) values of covariate q ; (c) variogram for target and residuals, (d) OLS and GLS estimates of the regression model and results of prediction for a 10×10 grid using ordinary kriging (e) and regression-kriging (f). Note how the RK maps reflects the pattern of the covariate.

textbook example from Burrough and McDonnell (1998, pp. 139–141), in which five measurements are used to predict a value of the target variable (z) at an unvisited location (s_0) (Fig. 1a). We extend this example by adding a hypothetical auxiliary data source: a raster image (10×10 pixels) (Fig. 1b), which has been constructed to show a strong negative correlation with the target variable at the sample points.

The RK predictions are computed as follows:

- (1) Determine a linear model of the variable as predicted by the auxiliary map q . In this case the correlation is high and negative with OLS coefficients $b_0 = 6.64$ and $b_1 = -0.195$ (Fig. 1d).
- (2) Derive the OLS residuals at all sample locations as

$$e^*(s_i) = z(s_i) - [b_0 + b_1 \cdot q(s_i)]. \tag{7}$$

For example, the point at ($x = 9, y = 9$) with $z = 2$ has a prediction of $6.64 - 0.195 \cdot 23 = 1.836$, resulting in an OLS residual of $e^* = -0.164$.

- (3) Model the covariance structure of the OLS residuals. In this example the number of points is far too small to estimate the autocorrelation function, so we follow the original text in using a hypothetical variogram of the target variable (spherical model, nugget $C_0 = 2.5$, sill $C_1 = 7.5$ and range $R = 10$) and residuals (spherical model, $C_0 = 2, C_1 = 4.5, R = 5$). The residual model is derived from the target variable model of the text by assuming that the residual variogram will have approximately the same form and nugget but a somewhat smaller sill and range (Fig. 1c), which is often found in practice (Hengl et al., 2004b).
- (4) Estimate the GLS coefficients using Eq. (4). In this case we get just slightly different coefficients $b_0 = 6.68$ and $b_1 = -0.199$. The GLS coefficients will not differ much from the OLS coefficients as long there is no significant clustering of the sampling locations (Fig. 1d) as in this case.
- (5) Derive the GLS residuals at all sample locations as

$$e^{**}(s_i) = z(s_i) - [b_0 + b_1 \cdot q(s_i)]. \tag{8}$$

Note that the b now refer to the GLS coefficients.

- (6) Model the covariance structure of the GLS residuals as a variogram. In practice this will hardly differ from the covariance structure of the OLS residuals.
- (7) Interpolate the GLS residuals using SK with known expected mean of the residuals (by definition 0) and the modelled variogram. In this case at the unvisited point location (5, 5) the interpolated residual is -0.081 .
- (8) Add the GLS surface to the interpolated GLS residuals at each prediction point. At the unvisited point location (5, 5) the auxiliary variable has a value 12, so that the prediction is then:

$$\begin{aligned}\hat{z}(5, 5) &= b_0 + b_1 \cdot q_i + \sum_{i=1}^n \lambda_i(s_0) \cdot e(s_i) \\ &= 6.68 - 0.199 \cdot 12 - 0.081 = 4.21,\end{aligned}\quad (9)$$

which is, in this specific case, a slightly different result than that derived by OK with the hypothetical variogram of the target variable ($\hat{z} = 4.30$).

The results of OK (Fig. 1e) and RK (Fig. 1f) over the entire spatial field are quite different in this case, because of the strong relation between the covariate and the samples. In the case of RK, most of variation in the target variable (82%) has been accounted for by the trend. Depending on the strength of the correlation, the RK might turn to pure kriging (no correlation) or pure regression (high correlation, pure nugget variogram). In that sense, pure kriging and pure regression should be considered as only special cases of a generic spatial prediction technique (Gotway and Stroup, 1997; Christensen, 2001; Pebesma, 2004).

3. Case studies

We will now demonstrate how can RK be used to map different types of environmental variables. We present three mapping exercises at the national level, using the same set of predictors and the generic mapping framework based on RK explained in detail in Hengl et al. (2004b). The key characteristics of this framework are:

- Factor analysis of covariates (raster maps) is used prior to interpolation to reduce the multi-

collinearity and to be able to compare results of fit for different predictors (Hengl et al., 2004b).

- The logistic transformation of the target variable is used to account for skewed distribution and prevent predictions outside the physical range (Hengl et al., 2004b, pp. 90–91).
- Stepwise regression is used to select regression predictors, to avoid spurious detail in the prediction maps.

Sixteen environmental predictors were prepared for the national territory of Croatia from two sources of data: land surface and multitemporal satellite radiometric images. These are the two most commonly used types of environmental predictors of soil and vegetation (Dobos et al., 2000; Hengl et al., 2002). The topography was parameterized with elevation data from two sources: (a) Shuttle Radar Topography Mission (SRTM) and (b) a DEM interpolated from the contour lines digitized from the 1:25K topo-maps. An SRTM 90m resolution DEM was downloaded from the International Agriculture Research Consortium for Spatial Information server (<http://srtm.csi.cgiar.org>).

Prior to derivation of land surface parameters, the SRTM elevation must be estimated from the raw SRTM data (which is a surface, not elevation, model) by filtering out the canopy (Rabus et al., 2003). Another problem with SRTM data is that it may contain substantial noise and artefacts, so that the nominal vertical accuracy (RMSE) of 15 m is often over-optimistic. Therefore, the average elevation was computed as a weighted average between the DEM derived from the contours from the 1:25K topo maps and the SRTM DEM. This is a good compromise because the SRTM DEM will show more detail in the plain areas, while the DEM derived from the topo maps is more accurate in the areas of dense contours. The land surface parameters slope gradient in % (SLOPE), wetness index (CTI) and direct incoming solar radiation (SOLAR) expressed in kWh/m² were computed from the DEM. The height of canopy (CANH) was computed as the difference between the two surface models (Kellendorfer et al., 2004) (Fig. 2). SLOPE and CANH were derived in the Integrated Land and Water Information System (ILWIS) GIS (Hengl et al., 2003) and CTI and SOLAR were derived in the SAGA GIS package (<http://saga-gis.org>).

Vegetation was represented by a set of enhanced vegetation index (EVI) images provided as band 2

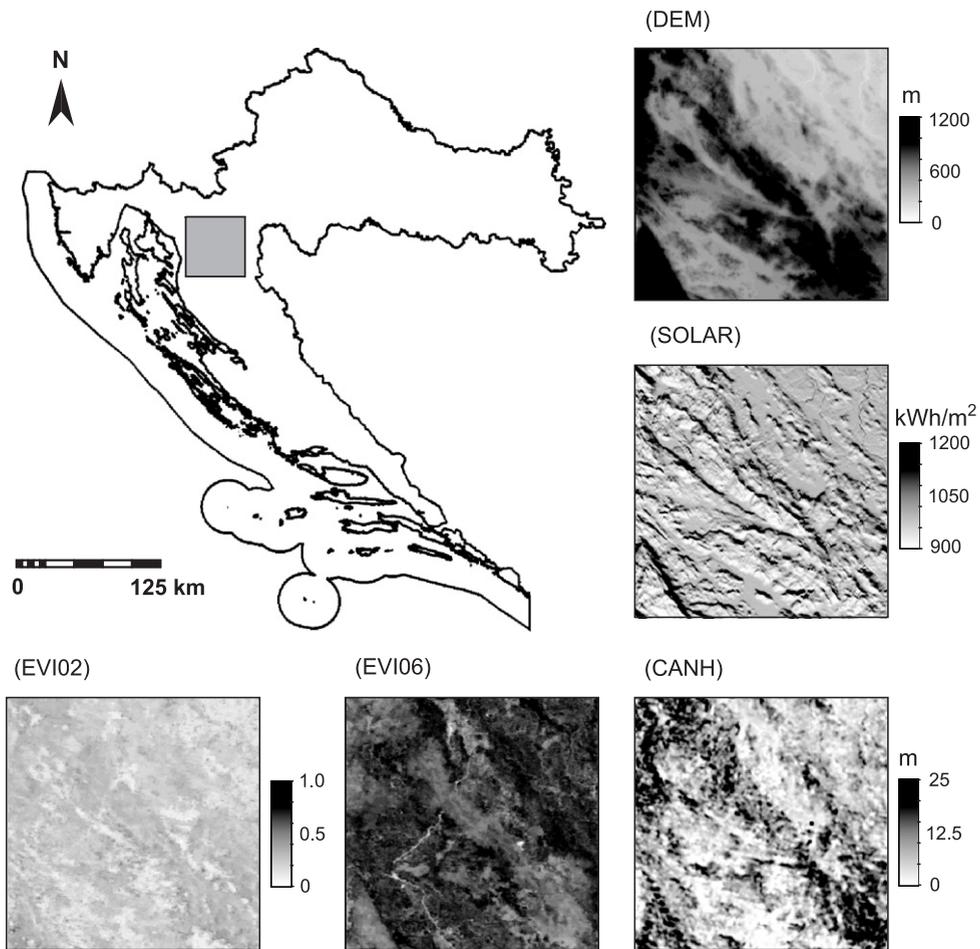


Fig. 2. Window showing some auxiliary maps used as environmental predictors: digital elevation model (DEM), solar irradiation (SOLAR), canopy height (CANH), and enhanced vegetation index for February (EVI02) and July (EVI06) of 2004.

of the 11-band Moderate Resolution Imaging Spectroradiometer (MODIS) imagery (<http://modis.gsfc.nasa.gov>), freely available from the NASA servers. Sixteen-day MODIS composites at a spatial resolution of 250 m for the year 2004 (a total of 22) were first filtered to remove artefacts caused by clouds and snow, using the statistical procedure described in Hengl et al. (2004a, pp. 100–101). These were then combined as pair averages to obtain 11 images, i.e. approximately one image per month (Fig. 2). All raster maps were brought to the same grid resolution of 200 m (full image 2353×2370 pixels), which corresponds to a working scale of about 1:200K (Hengl, 2006).

From sixteen original images we then derived the same number of factors using principal component analysis in ILWIS. Visual inspection of the PCs showed that the components 9–11 repeated previous

features with much less contrast, while components 12–16 just reflect noise in the input images. Much of the noise in the components comes from the MODIS images, which indicates that even more sophisticated methods to filter such images are needed. Thus the first eight PCs were used as orthogonal variables for the stepwise regression.

In all case studies we used the GSTAT package both to automatically fit the variograms of residuals and to produce final predictions (Pebesma, 2004). For fitting of the variograms, we used the exponential model and weighted least squares method to emphasize shorter distances and lags with higher number of point pairs (N_j/h_j^2). In addition to the map showing final predictions, GSTAT also produces the UK variance map, which is the estimate of the uncertainty of the prediction model, i.e. precision of prediction. The GSTAT command files and

detailed explanation of the procedures is available at <http://spatial-analyst.net>.

3.1. Example 1: mapping soil organic matter

In the first example we demonstrate interpolation of a continuous variable that is assumed to be correlated with relief and vegetation, but the physical model is unknown. We used a set of 2087 measurements of the soil organic matter content (Fig. 3a) from the Croatian Soil Database (Martinović and Vranković, 1997; Antonić et al., 2003). Values for OM ranged from 0% to 64% with an average value of 7.6 and s.d. 7.5%. The data were transformed using logistic transformation to reduce right skewness by using the physical limits $OM_{min} =$

0% and $OM_{max} = 100\%$. The few locations with $OM = 0\%$ were replaced with the estimated laboratory measurement error of 0.1%, otherwise it would not be possible to convert these values to logits. After logistic transformation the average value was $OM^{++} = -2.594$ with s.d. of 0.923.

Stepwise regression selected PC1,2,4,5,6 and 8 as significant predictors. The best predictors were PC2 and PC6, which reflect CTI, mean EVI and elevations—as we get to higher elevations the OM tends to increase because of the cooler temperatures and consequently reduced soil microbiological activity (Hengl et al., 2002). The predictors explained 40% of the variation in OM (Fig. 3b). The variogram of residuals was fitted with exponential model ($C_0 = 0.443$, $C_1 = 0.071$, and $R = 31178$ m),

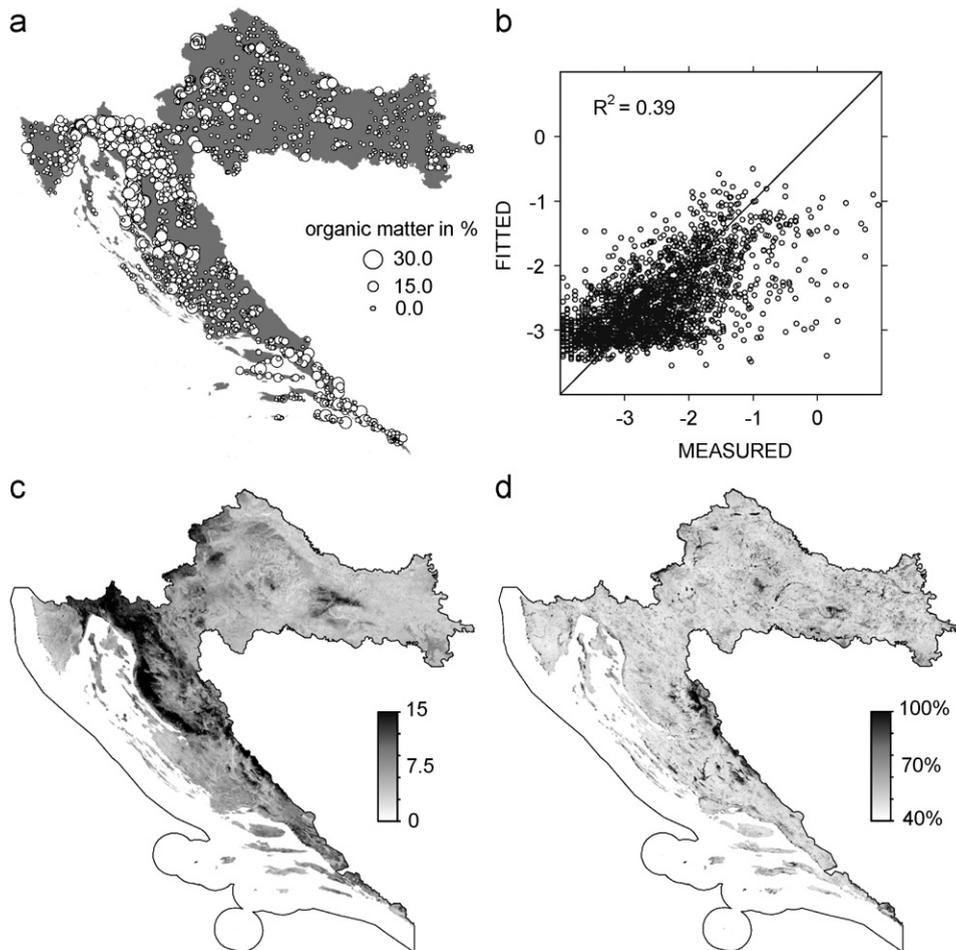


Fig. 3. Interpolation of organic matter in topsoil using RK: (a) original sampled values, (b) regression plot, (c) predicted values of whole area of interest, and (d) prediction error (RK variance).

with a relatively high nugget to sill ratio. Although the variogram model indicates poor predictive possibilities, almost half of variation has been explained by the regression model. Such variograms a typical effect of removing the feature-space structure: the remaining nugget is smaller or the same as in the original variogram, but the total sill is reduced. In this case, the residuals showed about two times smaller sill and about three times shorter range of spatial correlation than the target variable ($C_0 = 0.553$, $C_1 = 0.365$, and $R = 84129$ m). Consequently, the RK prediction map (Fig. 3c) closely follows the map of elevation, with few hot-spots in regions where the residuals were high. Finally, the prediction model explained 64% of total variation. The remaining areas of high prediction error (Eq. (6)) can be seen in Fig. 3d. The map of the prediction error can be now used to locate

additional samples and consider using either larger support size or more detailed predictors.

3.2. Example 2: mapping presence/absence of yew

The second example is an interpolation of a binary variable: presence or absence of a plant species in each grid cell, in this case yew (*Taxus baccata* L.). The result is the probability of occurrence (logistic-regression model), based on 364 sample plots arranged in a regular grid, in which yew was either present or absent (Fig. 4a). This grid was acquired from the on-line *Flora Croatica* database (Nikolić and Topić, 2005). In this case the predictors accounted for only 22% of total variation (Fig. 4b). The stepwise regression procedure selected six PCs (2, 4, 5, 6, 8, 10), the best predictor in fact being CANH. This is probably because yew is a

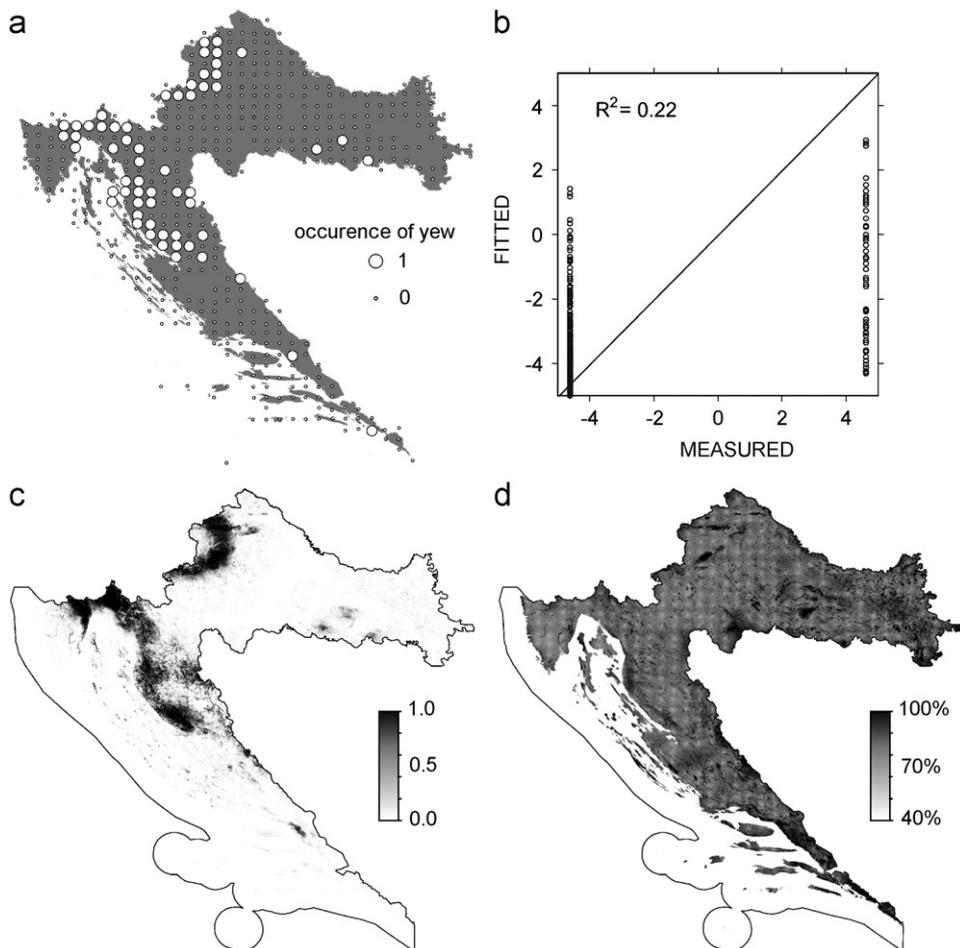


Fig. 4. Interpolation of occurrence of yew (*Taxus baccata* L.): (a) 364 observations on a regular grid, (b) logistic regression plot, (c) predicted values of whole area of interest, and (d) prediction error (RK variance).

species that favours shady sites, e.g. areas of deep boreal forest. The variogram of residuals was fitted, as in previous case study, with an exponential model ($C_0 = 4.126$, $C_1 = 5.336$ and $R = 13797$ m). Again, the nugget variation was rather significant. In addition, the prediction error is rather high, indicating that only 46% of variation has been explained by the model (Fig. 4d). Note also from Fig. 4d that the prediction error mainly reflects spatial location of points. The problem with this data set is obviously in the sampling design and sampling density. In this case, the short-range variation is completely under-sampled, the shortest spacing between the points is about 13 km and such sampling density probably does not comply with the targeted 200 m grid resolution.

3.3. Example 3: mapping land surface temperature

The last example demonstrates that RK can also be used for calibration of predictions made by running a physically based simulation model. In this case the target variable (z) is first modelled using a deterministic model which does not need any measurements but only input parameters. The observations can then be used to calibrate the model by kriging the residuals:

$$\hat{z}(s_0) = b_0 + b_1 \cdot \hat{z}_{mod}(s_0) + \sum_{i=1}^n \lambda_i \cdot [z(s_i) - (b_0 + b_1 \cdot \hat{z}_{mod}(s_0))], \quad (10)$$

where \hat{z}_{mod} is the soil property predicted using the physical model, and b_0 and b_1 are calibration coefficients, which are also derived using the GLS procedure. The example we use here is mapping of average annual land surface temperature (TEMP), which is modelled as a function of elevation, short-wave radiation ratio and leaf area index (Wilson and Gallant, 2000, p. 98):

$$\text{TEMP} = T_b - \frac{\Delta T \cdot (\text{DEM} - \text{DEM}_b)}{1000} + C \cdot \left(S - \frac{1}{S}\right) \cdot \left(1 - \frac{\text{LAI}}{\text{LAI}_{max}}\right), \quad (11)$$

where DEM is the elevation at the prediction location, DEM_b is the elevation of the reference climatic station (Zagreb), T_b is the temperature at the reference station, ΔT is the temperature gradient (here 5.06 °C per 1000 m elevation), C is an empirical constant (here 1 °C), S is the short-wave radiation ratio, LAI is the leaf area index at the grid cell and LAI_{max} is the

maximum leaf area index. TEMP was derived using ILWIS scripts (Hengl et al., 2003), following Eq. (11).

The map of TEMP from the modelled temperature was calibrated using actual measurements from year 2004, here annual temperatures measured at 127 climatic stations (Fig. 5a) provided by the Croatian meteorological and hydrological service. These ranged from 4.7 to 16.5 °C, average 11.6 °C. The physical limits for the logit transformation were set at long-term minimum and maximum annual temperatures ($z_{min} = -5$ and $z_{max} = 30$). The model used in Eq. (11) proved to be good estimator of the temperatures, achieving $R^2 = 0.564$ (Fig. 5b). The regression coefficients of actual vs. predicted were $b_0 = -0.193$ and $b_1 = 1.073$ (they would be 0 and 1 for a perfect, unbiased model). The calibration residuals can then be used to improve local predictions. These were fitted with an exponential model ($C_0 = 1.0885$, $C_1 = 14.4617$, and $R = 1,031,960$ m). The final predictions are shown in Fig. 5c. Note that the prediction model accounted for almost 95% of the total variation in the data, so that the prediction errors are rather small (Fig. 5d). Still, there is a nugget variation of about ± 1 °C that should not be ignored.

4. Discussion and conclusions

Clearly, RK is powerful spatial prediction technique that can be used to interpolate sampled environmental variables (both continuous and categorical) from large point sets. The barriers to widespread routine use of RK in environmental modelling and mapping are as follows. First, the statistical analysis in the case of RK is more sophisticated than for simple mechanistic or kriging techniques. Second, RK is computationally demanding and often cannot be run on standard PCs. The third problem is that many users are confused by the quantity of interpolation options, so that they are never sure which one is the most appropriate. In addition, there is a lack of user-friendly GIS environments to run RK. This is because, for many years GIS technologies and geostatistical techniques have been developing independently. We now address these and other issues regarding the practical use of RK.

4.1. Competitors to RK

The competitors to RK include completely different methods that may fit certain situations

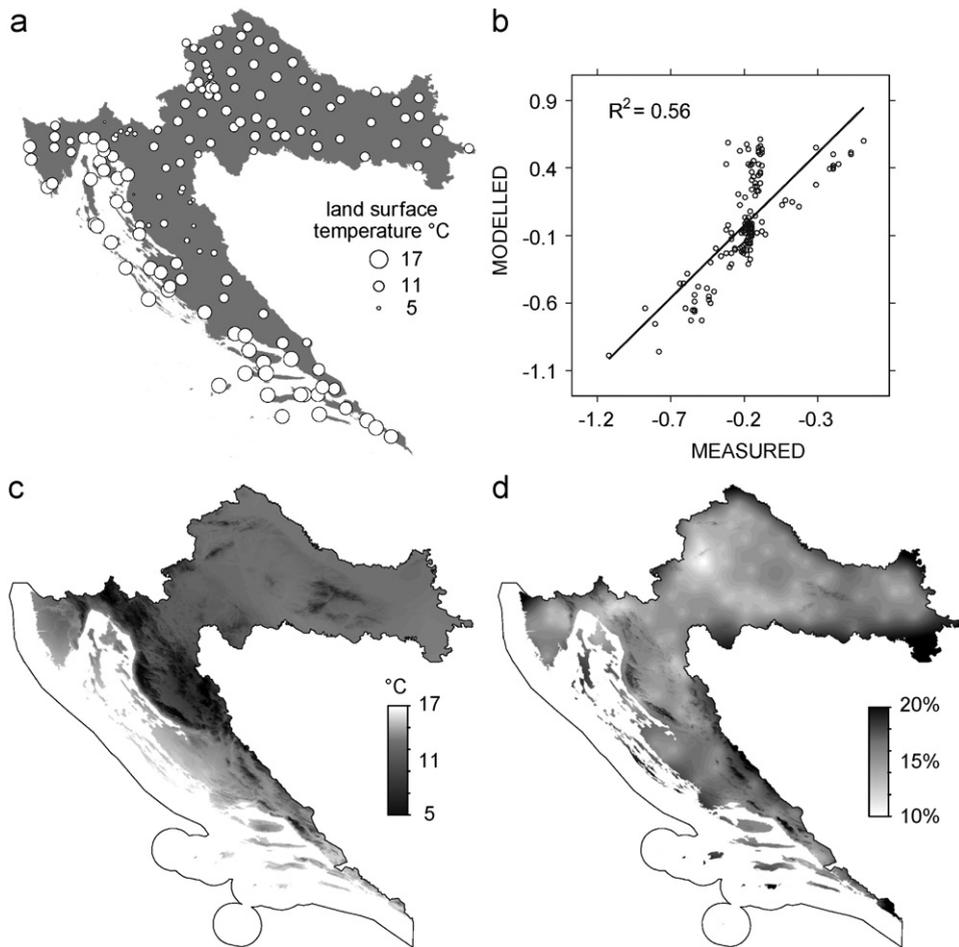


Fig. 5. Interpolation of land surface temperature: (a) original sampled values, (b) calibration plot, (c) predicted values of whole area of interest, and (d) prediction error (RK variance).

better. If the auxiliary data are of different origin and reliability, the Bayesian maximum entropy approach might be a better alternative (D'Or, 2003). There are also machine-learning techniques that combine neural network algorithms and robust interpolators. Henderson et al. (2005) used decision trees to predict various soil parameters from large quantity of soil profile data and with the help of land surface and remote sensing attributes. This technique is flexible, optimizes local fits and can be used within a GIS. However, it is statistically suboptimal because it ignores spatial location of points during the derivation of classification trees. The same authors further reported (Henderson et al., 2005, pp. 394–396) that, although there is still some spatial correlation in the residuals, it is not clear how to employ it.

RK must also be compared with alternative kriging techniques, such as OK and cokriging (CK). The advantage of OK is that it is less complicated in its use and included in most software packages. However, when auxiliary information is available in the form of maps of covariates that can explain part of the variation in the target variable, then RK outperforms OK because it exploits the extra information. Colocated CK does make use of the auxiliary information, but is developed for situations in which the auxiliary information is not spatially exhaustive (Knotters et al., 1995). CK also requires simultaneous modelling of both direct and cross variograms, which can be time-consuming for large number of covariates. In the case where the covariates are available as maps, RK will generally be preferred over CK, although CK may in some circumstances give superior results (Goovaerts, 1999).

4.2. Software implementation

There is still a large gap between what is possible for some (researchers) and what is available to many (users). No GIS package includes all of generalized linear models, variogram fitting, iterative estimation of residual variograms, and kriging, let alone their seamless integration. We have compared different aspects of geostatistical packages listed by the AI-GEOSTATS group (<http://ai-geostats.org>) and several well-known GIS packages (Table 1, see also supplementary materials at <http://spatial-analyst.net>). Although the UK (using coordinates) is available in most geostatistical packages, KED with multiple auxiliary maps can be run in only a limited number of packages. In fact, only Isatis (<http://geovariances.com>), SAGA (<http://saga-gis.org>), and GSTAT (<http://gstat.org>) as stand-alone application or integrated into R (R Development Core Team, 2004; Pebesma, 2004), GRASS (<http://geog.uni-hannover.de/grass>) or Idrisi (<http://clarklabs.org>) offer a possibility to interpolate a variable using auxiliary maps. We have tested RK in all these packages to discover that RK in Isatis is limited to a use of a single (three in script mode) auxiliary maps (Bleines et al., 2005). In Idrisi GLS regression coefficients cannot be estimated and the system is rather unstable. In GSTAT, both RK predictions and simulations (predictors as base maps) at both point and block support can be run by defining short scripts, which can help automatize interpolation of large amounts of data. However, GSTAT implements the algorithm with extended matrix (KED), which means that both the values of predictors and of target variable are used to estimate the values at each new location, which for large data sets can be time-consuming or can lead to computational problems (Leopold et al., 2005).

Setting UK in GSTAT to a smaller window search can lead to termination of the program due to the singular matrix problems. In fact, local UK with a global variogram model is not valid because the regression model will differ locally, hence the algorithm should also estimate the variogram model for residuals for each local neighbourhood. The singular matrix problem will happen especially when indicator variables are used as predictors or if the two predictor maps are highly correlated. Our experience with GSTAT was that interpolation of more than 1000 points over 1M of pixels can last up

Table 1
Comparison of computing capabilities of some popular statistical and GIS packages (versions in year 2005)

Aspect	S-PLUS		GSTAT/ R		SURFER		ISATIS		GEOEas		GSLIB		GRASS		PC Raster		ILWIS		IDRISI		ArcGIS		SAGA		
	II	A, B	IV	B	III	B, E	I	B	IV	B	IV	B, C	IV	B, C	IV	B, C	II	II	I	I	IV	B, C	IV	B, C	
Commercial price category	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Main application	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
User-friendly environment to non-expert	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Quality of support and description of algorithms	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Standard GIS capabilities	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Standard descriptive statistical analysis	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Image processing tools (orthorectification, filtering, land surface analysis)	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Comprehensive regression analysis (regression trees, GLM)	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Interactive (automated) variogram modelling	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Regression-kriging with auxiliary maps	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆
Dynamic modelling (simulations, spatial iterations, propagation, animations)	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆	☆

☆—Full capability, ★—possible but with many limitations, —not possible in this package. Commercial price category: I—> 1000 EUR; II—500–1000 EUR; III—<500 EUR; IV—open source or freeware. Main application: A—statistical analysis and data mining; B—interpolation of point data; C—processing of auxiliary maps; E—preparation and visualization of final maps.

to several hours on a standard PC. To run simulations in GSTAT with the same settings will take even more time. This clearly proves that, although KED procedure is mathematically elegant, such problems show that it might be more effective for real-life applications to fit the trend and residuals separately (RK) instead of through use of an extended matrix (KED). Another limitation of GSTAT is that it is a stand-alone application and the algorithms cannot be adjusted easily.

To allow extension of GSTAT functionalities and integration of its functionalities with other statistical functions, the developer of GSTAT, with a support of colleagues, developed an R package called *spatial* (<http://r-spatial.sourceforge.net>). R itself provides rich facilities for regression modelling, including GLS. The only problem here is that each step must be run by the analyst, who must really be an R expert. The open-source packages open the door to analyses of unlimited sophistication. However, they were not designed with graphical user interface, wizards, or interaction as is typical for commercial GIS, so are not easily used by non-experts. There is thus opportunity both for commercial GIS to incorporate RK ideas, or for open-source software to become more user-friendly.

4.3. Limitations of RK

Finally, there are some limitations to routine use of RK. If any of these problems occur, RK can give even worse results than even non-statistical, empirical interpolators such as Thiessen polygons or moving averages. The following difficulties might also be considered as challenges for the geostatisticians:

- (1) *Data quality*: RK relies completely on the quality of data. If the data comes from different sources and have been sampled using biased or unrepresentative design, the predictions might be even worse than with simple mechanistic prediction techniques (Example 2, Fig. 4). Even a single bad data point can make any regression arbitrarily bad, which affects the RK prediction over the whole area.
- (2) *Under-sampling*: For regression modelling, the multivariate feature space must be well-represented in all dimensions. For variogram modelling, an adequate number of point-pairs must be available at various spacings. Webster and Oliver (2001, p. 85) recommend at least 50 and preferably 300 points for variogram estimation. Neter et al. (1996) recommends at least 10 observations per predictor for multiple regression. We strongly recommend using RK only for data sets with more than 50 total observations and at least 10 observations per predictor to prevent over-fitting.
- (3) *Reliable estimation of the covariance/correlation structure*: The major dissatisfaction of using KED or RK is that both the regression model parameters and covariance function parameters need to be estimated simultaneously. However, in order to estimate coefficients we need to know covariance function of residuals, which can only be estimated after the coefficients (the chicken-egg problem). Here, we have assumed that a single iteration is a satisfactory solution, although someone might also look for other iterative solutions (Kitanidis, 1994).
- (4) *Extrapolation outside the sampled feature space*: If the points do not represent feature space or represent only the central part of it, this will often lead to poor estimation of the model and poor spatial prediction (Example 1, Fig. 4d). This is especially important for linear modelling where the prediction variance exponentially increases as we get closer to the edges of the feature space. For this reason it is important that the points be well spread at the edges of the feature space and that they be symmetrically spread around the center of the feature space (Hengl et al., 2004c). An assessment of the extrapolation in feature space can also be used to allocate additional point samples that can be used to improve the existing prediction models. This also justifies use of multiple predictors to fit the target variable, instead of using only the most significant predictor or first principal component, which is, for example, advocated by the Isatis development team (Bleines et al., 2005).
- (5) *Predictors with uneven relation to the target variable*: Auxiliary maps should have a constant physical relationship with the target variable in all parts of the study area, otherwise artefacts will be produced. An example is a single NDVI as a predictor of topsoil organic matter. If an agricultural field has just been harvested (low NDVI), the prediction map will (incorrectly) show very low organic matter content within the crop field.

(6) *Intermediate-scale modelling*: RK has not been adapted to fit data locally, with arbitrary neighbourhoods for the regression as can be done with kriging with moving window (Walter et al., 2001). Many practitioners would like to adjust the neighbourhood to fit their concepts of the scale of processes that are not truly global (across the whole study area) but not fully local either.

4.4. Next steps

What the programmers might consider for future is the refinement of (local) RK in a moving window. This will allow not only better data fitting, but will also allow users to visualize variation in regression (maps of R^2 and regression coefficients) and variogram models (maps of variogram parameters). Note that the RK with moving window would need to be fully automated, which might not be an easy task considering the computational complexity. Also, unlike the OK with moving window (Walter et al., 2001), RK has much higher requirements considering the minimum number of observations (at least 10 per predictor, at least 50 to model variogram). In general, our impression is that much of the procedures (regression and variogram modelling) in RK can be automated and amount of data modelling definitions expanded (local or global modelling, transformations, selection of predictors, type of GLMs etc.), as long as the point data set is large and of high quality. Ideally, the user should be able to easily test various combinations of input parameters and then (in real-time) select the one that produces most satisfactory predictions.

In conclusion, RK is a flexible method for modelling and mapping which offers conceptual advantages over alternative methods. We hope it becomes a routine part of the geostatistical toolbox. A task for the programmers in the near future will be to incorporate statistical procedures, such as step-wise regression, neural networks, automated variogram modelling, simulated annealing, unsupervised fuzzy classification and similar, within GIS user environments. Because many statistical techniques can be automated, integration of GIS and statistical algorithms should open the possibility to easily and quickly interpolate dozens of variables by using dozens of predictors. Nevertheless, analysts should have the final control to adjust the system as needed. To do this, they should have full insight into algorithms used.

Appendix A. Proof of equivalence of RK and KED

Start from kriging with external drift (or universal kriging) where the predictions are made as in OK using $\hat{z}_{\text{KED}}(s_0) = \lambda_{\text{KED}}^T \cdot \mathbf{z}$. The KED kriging weights (λ_{KED}^T) are obtained by solving the system (Wackernagel, 1998, p. 179):

$$\begin{bmatrix} \mathbf{C} & \mathbf{q} \\ \mathbf{q}^T & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \lambda_{\text{KED}} \\ \boldsymbol{\phi} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{q}_0 \end{bmatrix}, \quad (\text{A.1})$$

where $\boldsymbol{\phi}$ is a vector of Lagrange multipliers. Writing this out yields:

$$\begin{aligned} \mathbf{C} \cdot \lambda_{\text{KED}} + \mathbf{q} \cdot \boldsymbol{\phi} &= \mathbf{c}_0, \\ \mathbf{q}^T \cdot \lambda_{\text{KED}} &= \mathbf{q}_0. \end{aligned} \quad (\text{A.2})$$

From this follows:

$$\mathbf{q}^T \cdot \lambda_{\text{KED}} = \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{c}_0 - \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q} \cdot \boldsymbol{\phi} \quad (\text{A.3})$$

and hence:

$$\begin{aligned} \boldsymbol{\phi} &= (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{c}_0 \\ &\quad - (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}_0, \end{aligned} \quad (\text{A.4})$$

where the identity $\mathbf{q}^T \cdot \lambda_{\text{KED}} = \mathbf{q}_0$ has been used. Substituting $\boldsymbol{\phi}$ back into Eq. (A.2) shows that the KED weights equal (Papritz and Stein, 1999, p. 94):

$$\begin{aligned} \lambda_{\text{KED}} &= \mathbf{C}^{-1} \cdot \mathbf{c}_0 - \mathbf{C}^{-1} \cdot \mathbf{q} \\ &\quad \cdot [(\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{c}_0 \\ &\quad - (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}_0] \\ &= \mathbf{C}^{-1} \cdot [\mathbf{c}_0 + \mathbf{q} \cdot (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \\ &\quad \cdot (\mathbf{q}_0 - \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{c}_0)]. \end{aligned} \quad (\text{A.5})$$

Let us now turn to RK. Recall from Eq. (4) that the GLS estimate for the vector of regression coefficients is given by

$$\hat{\boldsymbol{\beta}}_{\text{GLS}} = (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{z} \quad (\text{A.6})$$

and weights for residuals by

$$\lambda_0^T = \mathbf{c}_0^T \cdot \mathbf{C}^{-1}. \quad (\text{A.7})$$

Substituting these in RK formula (Eq. (5)) gives

$$\begin{aligned} \hat{z}_{\text{RK}}(s_0) &= \mathbf{q}_0^T \cdot \hat{\boldsymbol{\beta}}_{\text{GLS}} + \lambda_0^T \cdot (\mathbf{z} - \mathbf{q} \cdot \hat{\boldsymbol{\beta}}_{\text{GLS}}) \\ &= [\mathbf{q}_0^T \cdot (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}^T \cdot \mathbf{C}^{-1} + \mathbf{c}_0^T \cdot \mathbf{C}^{-1} \\ &\quad - \mathbf{c}_0^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q} \cdot (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}^T \cdot \mathbf{C}^{-1}] \cdot \mathbf{z} \end{aligned}$$

$$\begin{aligned}
&= \mathbf{C}^{-1} \cdot [\mathbf{c}_0^T + \mathbf{q}_0^T \cdot (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}^T \\
&\quad - \mathbf{c}_0^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q} \cdot (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \cdot \mathbf{q}^T] \cdot \mathbf{z} \\
&= \mathbf{C}^{-1} \cdot [\mathbf{c}_0 + \mathbf{q} \cdot (\mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{q})^{-1} \\
&\quad \cdot (\mathbf{q}_0 - \mathbf{q}^T \cdot \mathbf{C}^{-1} \cdot \mathbf{c}_0)] \cdot \mathbf{z}. \quad (\text{A.8})
\end{aligned}$$

The left part of the equation is equal to Eq. (A.5), which proves that KED will give the same predictions as RK if same inputs are used. A detailed comparison of RK and KED is also available as supplementary material.

References

- Ahmed, S., de Marsily, G., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. *Water Resources Research* 23 (9), 1717–1737.
- Antonić, O., Pernar, N., Jelaska, S., 2003. Spatial distribution of main forest soil groups in Croatia as a function of basic pedogenetic factors. *Ecological Modelling* 170 (2–3), 363–371.
- Berterretche, M., Hudak, A.T., Cohen, W.B., Maiersperger, T.K., Gower, S.T., Dungan, J., 2005. Comparison of regression and geostatistical methods for mapping leaf area index (LAI) with landsat ETM+ data over a boreal forest. *Remote Sensing of Environment* 96 (3), 49–61.
- Bishop, T., McBratney, A., 2001. A comparison of prediction methods for the creation of field-extent soil property maps. *Geoderma* 103 (1–2), 149–160.
- Bleines, C., Perseval, S., Rambert, F., Renard, D., Touffait, Y., 2005. ISATIS. Isatis Software Manual. Geovariances & Ecole Des Mines De, Paris, 710pp.
- Bourennane, H., King, D., 2003. Using multiple external drifts to estimate a soil variable. *Geoderma* 114 (1–2), 1–18.
- Burrough, P., McDonnell, R., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford, 333pp.
- Chiles, J., Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty. Wiley, New York, 695pp.
- Christensen, R., 1996. Plane Answers to Complex Questions: The Theory of Linear Models, second ed. Springer, New York, 452pp.
- Christensen, R., 2001. Linear Models for Multivariate Time Series and Spatial Data, second ed. Springer, New York, 398pp.
- Cressie, N., 1993. Statistics for Spatial Data, revised ed. Wiley, New York, 900pp.
- Desbarats, A.J., Logan, C.E., Hinton, M.J., Sharpe, D.R., 2002. On the kriging of water table elevations using collateral information from a digital elevation model. *Journal of Hydrology* 255 (1–4), 25–38.
- Deutsch, C., Journel, A., 1998. GSLIB: Geostatistical Software and User's Guide, second ed. Oxford University Press, New York, 369pp.
- Dobos, E., Micheli, E., Baumgardner, M.F., Biehl, L., Helt, T., 2000. Use of combined digital elevation model and satellite radiometric data for regional soil mapping. *Geoderma* 97 (3–4), 367–391.
- D'Or, D., 2003. Spatial Prediction of Soil Properties, the Bayesian Maximum Entropy Approach. Ph.D., Université Catholique de Louvain, 212pp.
- Draper, N., Smith, H., 1981. Applied Regression Analysis, second ed. Wiley, New York, 709pp.
- Finke, P.A., Brus, D.J., Bierkens, M.F.P., Hoogland, T., Knotters, M., de Vries, F., 2004. Mapping groundwater dynamics using multiple sources of exhaustive high resolution data. *Geoderma* 123 (1–2), 23–39.
- Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, 483pp.
- Goovaerts, P., 1999. Using elevation to aid the geostatistical mapping of rainfall erosivity. *Catena* 34 (3–4), 227–242.
- Gotway, C., Stroup, W., 1997. A generalized linear model approach to spatial data analysis and prediction. *Journal of Agricultural, Biological, and Environmental Statistics* 2 (2), 157–198.
- Henderson, B., Bui, E., Moran, C., Simon, D., 2005. Australia-wide predictions of soil properties using decision trees. *Geoderma* 124 (3–4), 383–398.
- Hengl, T., 2006. Finding the right pixel size. *Computers & Geosciences* 32 (9), 1283–1298.
- Hengl, T., Rossiter, D.G., Husnjak, S., 2002. Mapping soil properties from an existing national soil data set using freely available ancillary data. In: Proceedings of the 17th World Congress of Soil Science, Paper no. 1140. IUSS, Bangkok, Thailand, p. 1481.
- Hengl, T., Gruber, S., Shrestha, D., 2003. Digital Terrain Analysis in ILWIS. Lecture Notes. International Institute for Geo-Information Science & Earth Observation (ITC), Enschede, 56pp.
- Hengl, T., Gruber, S., Shrestha, D.P., 2004a. Reduction of errors in digital terrain parameters used in soil–landscape modelling. *International Journal of Applied Earth Observation and Geoinformation (JAG)* 5 (2), 97–112.
- Hengl, T., Heuvelink, G., Stein, A., 2004b. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma* 122 (1–2), 75–93.
- Hengl, T., Rossiter, D., Stein, A., 2004c. Soil sampling strategies for spatial prediction by correlation with auxiliary maps. *Australian Journal of Soil Research* 41 (8), 1403–1422.
- Kellndorfer, J., Walker, W., Pierce, L., Dobson, C., Fites, J.A., Hunsaker, C., Vona, J., Clutter, M., 2004. Vegetation height estimation from shuttle radar topography mission and national elevation datasets. *Remote Sensing of Environment* 93 (3), 339–358.
- Kitanidis, P., 1994. Generalized covariance functions in estimation. *Mathematical Geology* 25, 525–540.
- Knotters, M., Brus, D., Voshaar, J., 1995. A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma* 67 (3–4), 227–246.
- Leopold, U., Heuvelink, G.B., Tiktak, A., Finke, P.A., Schoumans, O., 2005. Accounting for change of support in spatial accuracy assessment of modelled soil mineral phosphorous concentration. *Geoderma* 130 (3–4), 368–386.
- Lloyd, C.D., 2005. Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *Journal of Hydrology* 308 (1–4), 128–150.
- Lopez-Granados, F., Jurado-Exposito, M., Pena-Barragan, J., Garcia-Torres, L., 2005. Using geostatistical and remote

- sensing approaches for mapping soil properties. *European Journal of Agronomy* 23 (33), 279–289.
- Martinović, J., Vranković, A. (Eds.), 1997. *Croatian Soil Database (in Croatian)*, vols. I–III, Ministry of Environmental Protection and Physical Planning, Zagreb.
- Matheron, G., 1969. *Le krigeage universel (Universal kriging)*. vol. 1. Cahiers du Centre de Morphologie Mathematique, Ecole des Mines de Paris, Fontainebleau, 83pp.
- McKenzie, N., Ryan, P., 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma* 89 (1–2), 67–94.
- Neter, J., Kutner, M., Nachtsheim, C., Wasserman, W. (Eds.), 1996. *Applied Linear Statistical Models*, fourth ed. McGraw-Hill/Irwin, Chicago, New York, 1408pp.
- Nikolić, T., Topić, J. (Eds.), 2005. *Red Book of Vascular Flora of Croatia*. Ministry of Culture, State Institute for Nature Protection, Republic of Croatia, Zagreb, 693pp.
- Odeh, I., McBratney, A., Chittleborough, D., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. *Geoderma* 67 (3–4), 215–226.
- Papritz, A., Stein, A., 1999. Spatial prediction by linear kriging. In: Stein, A., van der Meer, F., Gorte, B. (Eds.), *Spatial Statistics for Remote Sensing*. Kluwer Academic Publishers, Dordrecht, pp. 83–113.
- Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30 (7), 683–691.
- Pleydell, D.R.J., Raoul, F., Tourneux, F., Danson, F.M., Graham, A.J., Craig, P.S., Giraudoux, P., 2004. Modelling the spatial distribution of *Echinococcus multilocularis* infection in foxes. *Acta Tropica* 91 (3), 253–265.
- R Development Core Team, 2004. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-00-3.
- Rabus, B., Eineder, M., Roth, A., Bamler, R., 2003. The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing* 57 (4), 241–262.
- Rivoirard, J., 2002. On the structural link between variables in kriging with external drift. *Mathematical Geology* 34 (7), 797–808.
- Wackernagel, H., 1998. *Multivariate Geostatistics: An Introduction with Applications*, second ed. Springer, Berlin, 291pp.
- Walter, C., McBratney, A.B., Donuaoui, A., Minasny, B., 2001. Spatial prediction of topsoil salinity in the Chelif valley, Algeria, using local ordinary kriging with local variograms versus whole-area variogram. *Australian Journal of Soil Research* 39, 259–272.
- Webster, R., Oliver, M., 2001. *Geostatistics for Environmental Scientists Statistics in Practice*. Wiley, Chichester, 271pp.
- Wilson, P.J., Gallant, C.J. (Eds.), 2000. *Terrain Analysis: Principles and Applications*. Wiley, New York, 479pp.
- Yemefack, M., Rossiter, D.G., Njomgang, R., 2005. Multi-scale characterization of soil variability within an agricultural landscape mosaic system in southern Cameroon. *Geoderma* 125 (1–2), 117–143.