
DATA INTEGRATION IN SPATIAL MODELLING

Ladislav Vizi¹

¹Ústav geodézie, kartografie a GIS, Fakulta BERG, TU v Košiciach, Pk. Komenského 19,
042 00, Košice, Slovakia
ladislav.vizi@tuke.sk

Abstract. The general purpose of this paper is to introduce the most popular geostatistical methods for integration of the different data sources measure the same characteristic but in different ways. The information available on a natural phenomenon is rarely limited to the values assumed by a single variable over a set of samples points. Direct measurements of the primary attribute, or variable under study, are often supplemented by secondary information originating from other related categorical or continuous attributes, or secondary (auxiliary) variables. Most real studies involve more than one variable. In general, the estimation improves when this additional information is taken into consideration, particularly when primary data are sparse or poorly correlated in space.

The paper deals with the most popular geostatistical methods of data integration: cokriging, collocated cokriging and kriging with an external drift. The paper describes some simple mathematical background of the presented methods, and also some practical examples are given with comments on the obtained results.

Keywords: coregionalization, cross-variogram, primary and secondary variable.

1 Introduction

An important task in the spatial modelling is to investigate various information and their relationships. The greatest challenge in the spatial modelling is "an integration" of different sources of information, describing the same phenomenon in different ways, with unwritten rule: "Do not avoid what is known.". Geostatistics, and its multivariate techniques, provides space and tools to build that consistent models[2]. Except general cokriging, extending kriging to multivariate interpolation, there are another, more sophisticated, methods for spatial data integration. One of way of dealing with multiple variables is to treat one of them as a second-order stationary spatial random variable and treat the remaining variables in a deterministic manner, analogous to the way in which spatial drift (trend) is modelled for the non-stationary case. The trend model is then taken in kriging estimation method as an external drift. Another possibility is to take advantage of jointed spatial variability through direct and cross-variogram models and to use cokriging procedure or its collocated version.

2 Multivariate geostatistical models

The natural extension of the concept of a single-variable regionalisation to some several variables is a *coregionalisation*. It is essentially the joint regionalisation of two or more variables which may, or may not, be spatially inter-correlated [7]. In practical situation, however, we are not interested in the trivial case in which there is no correlation among the variables as each variable can be treated independently using univariate geostatistical

methods. The scattergram of collocated variables Z and V values, or the values of the variables $Z(x)$ and $V(x)$ measured in the same locations x_α ; $\alpha = 1, \dots, n$; provides the first assessment of the correlation between two variables. In general, the relationship between paired values $z(x_\alpha)$ and $v(x_\alpha)$ does not suffice to capture the full spatial relationship between two variables. The cross-spatial relationship between pairs of values separated by some lag distances h , $z(x_\alpha)$, $v(x_\alpha + h)$; $\alpha = 1, \dots, n(h)$; must also be considered [4]. The cross-variogram, defined below, measures this cross spatial dependence between two variables.

2.1 Cross-variogram

For the multivariate case, in which variables are spatially inter-correlated, we can make assumptions of stationarity that are analogous to the univariate case. On the assumption of second-order stationarity we have:

1. The expectation of each variable is constant:

$$E[Z_k(x)] = \mu_{Z_k} \quad \text{for a set of variables } k = 1, \dots, K. \quad (1)$$

2. A cross-covariance can be defined for each pair of variables $Z(x)$ and $V(x)$:

$$\begin{aligned} C_{ZY}(h) &= \text{Cov}[Z(x)V(x+h)] \\ &= E[Z(x) \cdot V(x+h)] - \underbrace{E[Z(x)]}_{\mu_z} \cdot \underbrace{E[V(x+h)]}_{\mu_v}. \end{aligned} \quad (2)$$

3. A cross-variogram can be defined for each pair of variables $Z(x)$ and $V(x)$:

$$2\gamma_{ZV}(h) = E\{[Z(x) - Z(x+h)] \cdot [V(x) - V(x+h)]\}. \quad (3)$$

In case of $V = Z$ is cross-variogram (3) reduced to the classical definition of the variogram for one variable of Z :

$$\begin{aligned} 2\gamma_{ZZ}(h) &= E\{[Z(x) - Z(x+h)] \cdot [Z(x) - Z(x+h)]\} \\ &= E\{[Z(x) - Z(x+h)]^2\}, \end{aligned} \quad (4)$$

what is the average product of difference of the $z(x) - z(x+h)$ multiplied by itself. The cross-variogram is then the average product of the z difference, $z(x) - z(x+h)$, and the v difference for the same location and lag distance h , $v(x) - v(x+h)$. The cross-variogram satisfies the inequality:

$$\gamma_{ZZ}(h) \cdot \gamma_{VV}(h) = |\gamma_{ZV}(h)|^2. \quad (5)$$

Unlike covariance and variogram functions, cross-covariance and cross-variogram can take negative values, which would indicate a negative spatial correlation between $Z(x)$ and $V(x)$. **Figure 1** shows an example of the cross-variography results taken from [15], where is given more detail. The values of the cross-variograms (**G.**, **H.** and **I.**) reflect the signs of the correlation coefficients (**A.**, **B.** and **C.**). Note the values of the direct, or single variograms (**D.**, **E.** and **F.**) that are always positive.

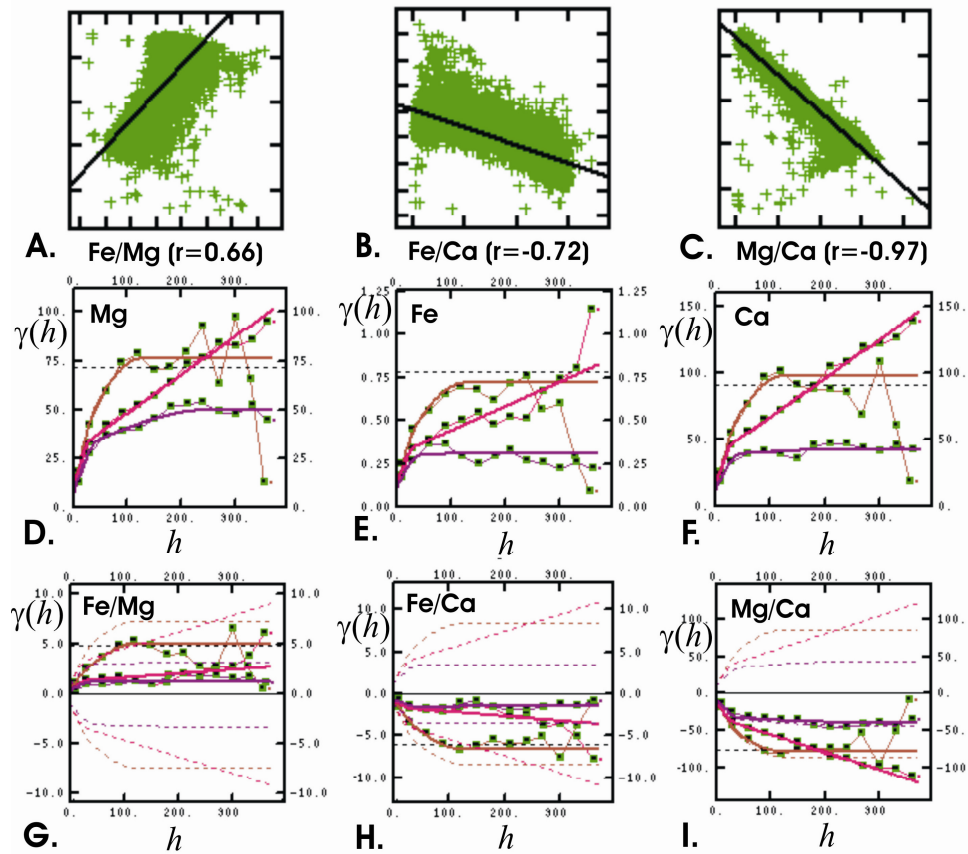


Figure 1. Experimental directional variograms and cross-variogram for Fe, Mg and Ca variables. The values of the cross-variograms (**G.**, **H.** and **I.**) reflect the signs of the correlation coefficients (**A.**, **B.** and **C.**). Row (**D.**, **E.** and **F.**) represents single variograms for three variables under study.

On the assumption of second-order stationarity the relationship between the cross-covariance and cross-variogram is:

$$2\gamma_{ZV}(h) = 2C_{ZV}(0) - C_{ZV}(h) - C_{ZV}(-h). \quad (6)$$

Whilst a cross-variogram is symmetrical in terms of the pairs of variables and in the direction of calculation:

$$\gamma_{ZV}(h) = \gamma_{VZ}(h) \quad \text{and} \quad \gamma_{ZV}(h) = \gamma_{ZV}(-h), \quad (7)$$

this is not necessarily so for the cross-covariance:

$$C_{ZV}(h) = C_{VZ}(-h) \quad \text{but} \quad C_{ZV}(-h) \neq C_{ZV}(h) \quad \text{and} \quad C_{ZV}(h) \neq C_{VZ}(h), \quad (8)$$

i.e., changing the order in which the variables are entered into calculation or changing the direction of the separation vector h may change the value of the cross-covariance. That means, cross-variogram, which covers only the even term of covariance, should not be used where asymmetry is thought to be significant, for example in time-series application [18].

2.2 Permissible models of cross-variograms

Modelling of a coregionalization calls for inferring $K(K + 1)/2$ direct and cross-variogram models for K inter-correlated random functions $Z_k(x)$. The difficulty does not lie in the number of models to infer, but in the fact that these models cannot be built independently from one another. The cross-variogram can be modelled in the same way as that variogram, and the same restricted set of mathematical functions, or basic structures, is available. But to describe the coregionalization there is another condition. Any linear combination of the K variables is itself a regionalized variable, and its variance must be positive or zero. It may not be negative because there is nothing like the negative variance. Only models that ensure the validity of positive variance calculations can be used for a coregionalisation modelling purpose and in kriging. Such models are called permissible models [17].

Denote $\{Z_k(x); k = 1, \dots, K\}$ a set of K inter-correlated random functions. The variance of any finite linear combination, LC , of the random functions $Z_k(x_\alpha)$ can be expressed as a linear combination of cross-covariance values:

$$\begin{aligned} \text{VAR}[LC] &= \text{VAR} \left[\sum_{k=1}^K \sum_{\alpha=1}^n \omega_{\alpha k} Z_k(x_\alpha) \right] \\ &= \sum_{\alpha=1}^n \sum_{\beta=1}^n \omega_{\alpha}^T \omega_{\beta} \mathbf{C}_{\alpha\beta} \geq 0, \end{aligned} \quad (9)$$

where ω_{α} is $K \times 1$ vector of weights $\omega_{\alpha k}$, $[\omega_{\alpha 1}, \dots, \omega_{\alpha K}]^T$, and $\mathbf{C}_{\alpha\beta}$ is the $K \times K$ matrix of stationary cross-covariances between any two random functions $Z_k(x_\alpha)$ and $Z_l(x_{\beta=\alpha+h})$, $C_{kl}(h)$.

From the relation between covariance and variogram $C(h) = C(0) - \gamma(h)$, the variance (9) can be rewritten in terms of matrix of variogram models $\gamma(h)$:

$$\text{VAR}[LC] = \sum_{\alpha=1}^n \sum_{\beta=1}^n \omega_{\alpha}^T \omega_{\beta} \mathbf{C}(0) - \sum_{\alpha=1}^n \sum_{\beta=1}^n \omega_{\alpha}^T \omega_{\beta} \gamma(h) \geq 0, \quad (10)$$

where $\mathbf{C}(0)$ is the variance-covariance matrix. For the variogram models there are the unbounded ones (models without sill) that have no covariance counterpart, corresponding to the intrinsic but non-stacionary random function [9]. For such models, the variance of LC is defined on the condition that the vectors of ω_{α} sum to null vector, which allows the filtering of the terms $\mathbf{C}(0)$ from (10):

$$\text{VAR}[LC] = - \sum_{\alpha=1}^n \sum_{\beta=1}^n \omega_{\alpha}^T \omega_{\beta} \gamma(h) \geq 0 \quad \text{with} \quad \sum_{\alpha=1}^n \omega_{\alpha} = 0. \quad (11)$$

Expression (11) shows that to ensure non-negativity of the variance of LC , the matrix of variogram models $\gamma(h)$ must be conditionally negative semi-definite under condition that the sum of vectors ω_{α} is the null vector [9].

2.3 Linear model of coregionalization

In the linear model of coregionalization we assume that each variable $Z_k(x)$ is a linear sum of $(S + 1)$ orthogonal, i.e. independent, random variables $Y_r^s(x)$, each with mean 0 and basic variogram function, or structure, $\Gamma_s(h)$. Subscript s is simply an index, not a power:

$$Z_k(x) = \sum_{s=0}^S \sum_{r=1}^R a_{kr}^s Y_r^s(x) + \mu_{Z_k} \quad (12)$$

In this expression we have:

$$E[Z_k(x)] = \mu_{Z_k} \quad \text{and} \quad E[Y_r^s(x)] = 0 \quad \text{for } \forall r \text{ and } s \quad (13)$$

and:

$$\frac{1}{2} E\left\{ \left[Y_r^s(x) - Y_r^s(x+h) \right] \cdot \left[Y_{r'}^{s'}(x) - Y_{r'}^{s'}(x+h) \right] \right\} = \begin{cases} \Gamma_s(h) & \text{if } r = r' \text{ and } s = s', \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

These conditions express the mutual (two by two) independence of the random functions $Y_r^s(x)$ [9].

The the variogram for any pair of variables k and l is defined as the set of $K \times K$ direct and cross-variograms such as:

$$\gamma_{kl}(h) = \sum_{s=0}^S \underbrace{\sum_{r=1}^R a_{kr}^s a_{lr}^s}_{b_{kl}^s} \Gamma_s(h), \quad (15)$$

where the b_{kl}^s is sill of the basic variogram model $\Gamma_s(h)$. These b_{kl}^s represent the variances and covariances, i.e. nugget and sill variances, for the independent components if they are bounded. For unbounded variograms the b_{kl}^s are nugget variances and gradients [18]. By construction, the coefficients b_{kl}^s and b_{lk}^s are identical for all s , and for each s the matrix of coefficients:

$$\mathbf{b}_s = [b_{kl}^s] = \begin{bmatrix} b_{kk}^s & b_{kl}^s \\ b_{lk}^s & b_{ll}^s \end{bmatrix}, \quad (16)$$

must be positive definite. Since the matrix is symmetric, it is sufficient that $b_{kk}^s \geq 0$ and $b_{ll}^s \geq 0$ and that its determinant is positive or zero. For K coregionalized variable the full matrix of coefficients will be order K , and its determinant and all its principal minors must be positive or zero [11].

The intrinsic coregionalization model is but a particular linear model (16) in which all the $K(K + 1)/2$ coefficients b_{kl}^s of any variogram function $\Gamma_s(h)$ are proportional to each other with coefficients ϕ_{kl} :

$$\gamma_{kl}(h) = \phi_{kl} \underbrace{\sum_{s=0}^S b_s \Gamma_s(h)}_{\Gamma_0(h)} \quad \text{for } \forall k, l \quad \text{with} \quad \sum_{s=0}^S b_s = 0, \quad (17)$$

where $\Gamma_0(h)$ is the same standardized linear model of regionalization. The intrinsic coregionalization model is much more restrictive than the linear model of coregionalization. The $K(K + 1)/2$ direct and cross-variogram models must include all $(S + 1)$ structures in the same proportion b_s . Any linear model of coregionalization that consists of a single basic structure of variogram is an intrinsic model of coregionalization.

3 Cokriging

The term kriging is traditionally reserved for linear regression using data on the same attribute as that being estimated. The term cokriging is reserved for linear regression that also uses data defined on different attributes [5]. Consider the simplest case where we have two spatially correlated variables, $Z(x)$ and $V(x)$ with auto-variograms $\gamma_Z(h)$ and $\gamma_V(h)$ and cross-variogram $\gamma_{ZV}(h)$. Variable $Z(x)$ has been sampled at a set of locations x_α^Z and variable $V(x)$ has been sampled at a set locations x_α^V . It is assumed that there is some overlap between x_α^Z and x_α^V (the variables share some sample locations – *partial heterotopic*). In case of a single secondary variable $V(x)$, the ordinary cokriging estimator of $Z(x_0)$ is written:

$$Z_{CK}^*(x_0) = \sum_{\alpha_z=1}^{n_z} \omega_\alpha^Z Z(x_\alpha) + \sum_{\alpha_v=1}^{n_v} \omega_\alpha^V V(x_\alpha). \quad (18)$$

The cokriging equations system is found by minimising the estimation variance subject to the two non-bias condition $\sum_{\alpha_z=1}^{n_z} \omega_\alpha^Z = 1$ and $\sum_{\alpha_v=1}^{n_v} \omega_\alpha^V = 0$ by incorporated two

Lagrange multipliers λ_Z and λ_V :

$$\begin{cases} \sum_{\beta_z=1}^{n_z} \omega_\beta^Z \gamma_{\alpha\beta}^Z + \sum_{\beta_v=1}^{n_v} \omega_\beta^V \gamma_{\alpha\beta}^{ZV} + \lambda_Z = \gamma_{\alpha 0}^Z, \\ \sum_{\beta_z=1}^{n_z} \omega_\beta^Z \gamma_{\alpha\beta}^{ZV} + \sum_{\beta_v=1}^{n_v} \omega_\beta^V \gamma_{\alpha\beta}^V + \lambda_V = \gamma_{\alpha 0}^{ZV}, \\ \sum_{\alpha_z=1}^{n_z} \omega_\alpha^Z = 1, \\ \sum_{\alpha_v=1}^{n_v} \omega_\alpha^V = 0. \end{cases} \quad (19)$$

The associated kriging variance is:

$$\sigma_{CK}^2 = \sum_{\alpha_z=1}^{n_z} \omega_a^Z \gamma_{\alpha 0}^Z + \sum_{\alpha_v=1}^{n_v} \omega_a^V \gamma_{\alpha 0}^{ZV} + \lambda_Z - \gamma_0^Z. \quad (20)$$

In *isotopic* case (the variables share all sample locations), the advantage of cokriging is that the cokriging estimator of a sum of variables is equal to the sum of the cokriging estimators. For example, the thickness of a geological layer $T(x)$ is defined as the difference between the depths of the bottom limit $Z_B(x)$ and top one $Z_T(x)$ of that layer. The difference of cokriging estimators of $Z_B(x)$ and $Z_T(x)$ ensures the positive estimation of $T(x)$. **Figure 2** shows an example of thickness of the layer estimation as a product of difference of cokriged estimations of the top and bottom elevations of the layer. More detail can be found in [16].

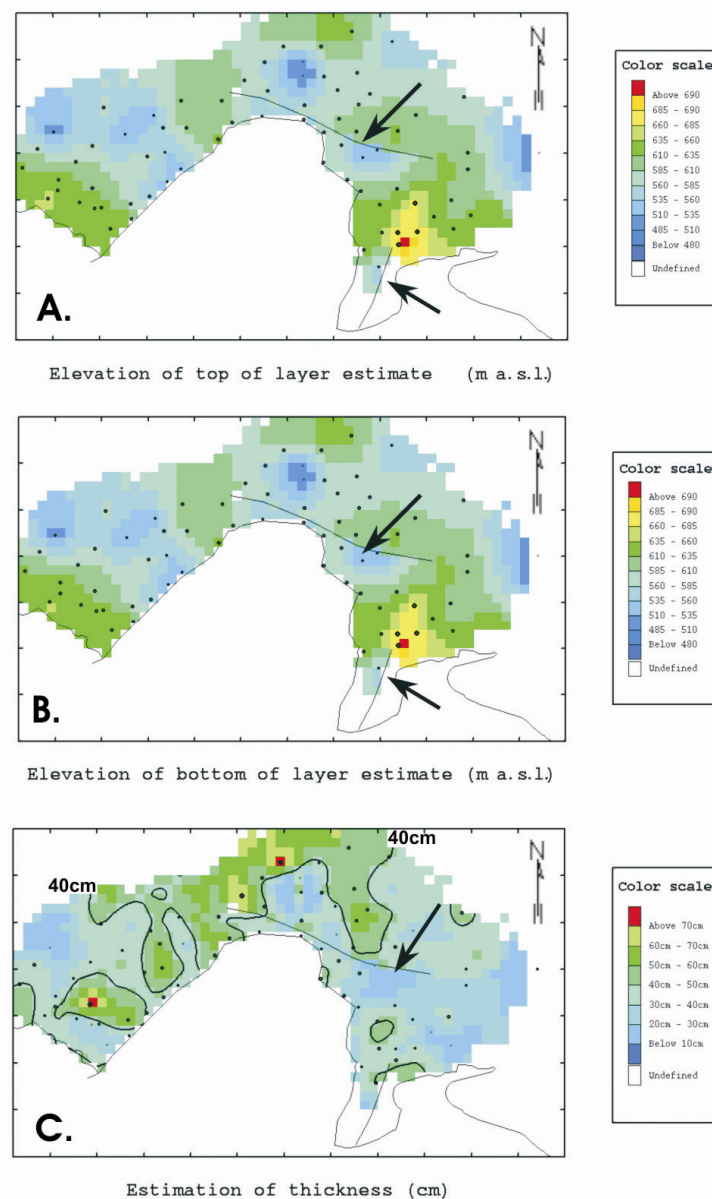


Figure 2. Estimation of thickness of the layer **C.** as a product of difference of cokriged estimations of the top **A.** and bottom **B.** elevation of the layer. The arrows show the main faults positions.

The cokriging estimator is theoretically better than the kriging one because its error variance is always smaller or equal than the error variance of kriging which ignores all secondary variables, or auxiliary information [9]. The cokriging estimator (18) can be extended to any K number secondary variables. The variogram matrix requires $K(K + 1)/2$ variogram functions when K different variables are considered in a cokriging. The inference becomes extremely demanding in terms of data and subsequent joint modelling is particularly tedious. This is the main reason why cokriging has not been extensively used in practice for a high number of variables. Another reason that cokriging is not used extensively in practice is the screen effect of the better correlated data (usually the z samples) over the data less correlated with the z unknown (the v samples). Unless the primary variable, that which is being estimated, is undersampled with respect to the secondary data, the weights given to the secondary data tend to be small and the reduction in estimation variance brought by cokriging is not worth the additional inference and modelling effort [5].

4 Collocated cokriging

When secondary variable is much more densely sampled than the primary variable, or secondary variable is available at each estimated position, the left side of kriging system with variogram values may be unstable because the correlation between close secondary data is much greater than the correlation between distant primary data. Moreover, secondary data that are very close or collocated with the position of an unknown primary value $z(x_0)$ tend to screen the influence of secondary data that are further away [9].

A reduced form of cokriging consist of retaining only collocated secondary data $v(x_\alpha = x_0)$, or relocated data $v'(x_\alpha \rightarrow x_0)$ to the nearest node x_0 being estimated. That means the collocated cokriging method uses $(n_z + 1)$ values. The collocated cokriging estimator is written:

$$Z_{CCK}^*(x_0) = \sum_{\alpha_z=1}^{n_z} \omega_\alpha^Z Z(x_\alpha) + \omega_0^V V(x_0), \quad (21)$$

with the constraint that all weights must sum to one:

$$\sum_{\alpha_z=1}^{n_z} \omega_\alpha^Z + \omega_0^V = 1, \quad (22)$$

The corresponding cokriging system requires knowledge of only the variogram $\gamma_Z(h)$ a cross-variogram $\gamma_{ZV}(h)$. The weights are obtained by solving the following system of $(n_z + 2)$ linear equation:

$$\begin{cases} \sum_{\beta_z=1}^{n_z} \omega_\beta^Z \gamma_{\alpha\beta}^Z + \omega_0^V \gamma_{\alpha\beta}^{ZV} + \lambda = \gamma_{\alpha 0}^Z, \\ \sum_{\beta_z=1}^{n_z} \omega_\beta^Z \gamma_{\alpha\beta}^{ZV} + \omega_0^V \gamma_0^V + \lambda = \gamma_0^{ZV}, \\ \sum_{\alpha_z=1}^{n_z} \omega_\alpha^Z + \omega_0^V = 1. \end{cases} \quad (23)$$

4.1 Multi-located cokriging

The idea of the multicollocated cokriging is to enhance the cokriging process by adding for each target point the value of the secondary variable value at this location. The system resembles the traditional cokriging techniques where one additional sample, which coincides with a target point, and for which only secondary variable value is provided. The technique requires a bivariate structure model of primary and secondary variables defined at primary data locations. In general case where there are n primary data values, accompanied by n secondary data values, in the cokriging neighbourhood, the estimator become:

$$Z_{CCK}^*(x_0) = \sum_{\alpha_z=1}^{n_z} \{ \omega_{\alpha}^Z Z(x_{\alpha}) + \omega_{\alpha}^V V(x_{\alpha}) \} + \omega_0^V V(x_0), \quad (24)$$

with $(2n_z + 1)$ values for corresponding kriging system [1].

4.2 Markov – Bayes approximation of collocated cokriging

The aim of collocated cokriging with Markov-Bayes approximation is to take the full advantage of the densely sampled secondary variable, especially during the structural analysis and variography where the cross-variogram and variogram of the primary variable are derived by simply scaling the secondary variogram [9]. The scaling factors are obtained by comparing the experimental variances of primary and secondary datasets and using their correlation coefficient.

Figure 3 shows an example three different results of Late Sarmatian sediment thickness estimations [m]. Figure **A.** shows the result obtained by kriging based only on drill-hole data as primary variable. Figure **B.** shows result base on collocated cokriging with seismic map as secondary variable. Due to low correlation between primary and secondary variable, the result resemble the kriging solution without using the seismic information. The kriged sediment thickness map takes more of seismic background as the correlation increase. The result of collocated cokriging based on Markov-Bayes approximation is shown in Figure **C.** Accounting for secondary variable through approximation of variogram function of the primary variable and cross-variogram yields more detailed map to the one based on cross-variography of collocated primary and secondary variable.

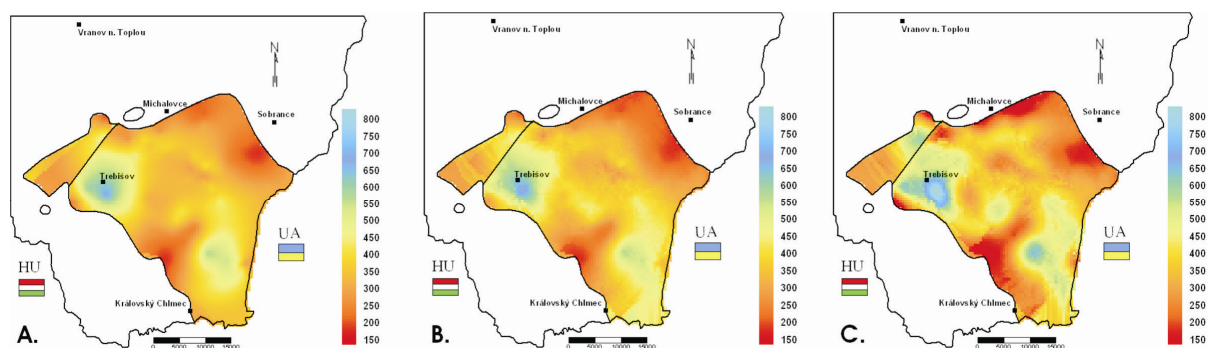


Figure 3. Three different results of the sediment thickness estimation [m]. Figure **A.** shows estimation based only on primary variable (drill-holes). Figure **B.** shows the result obtained with collocated cokriging using seismic information as secondary variable. Figure **C.** shows the results of Markov-Bayes approximation of collocated cokriging.

5 Kriging with an external drift

Kriging with an external drift considers a non-parametric trend shape that could come from a secondary variable such as seismic. Kriging with an external drift is but a variant of kriging with a trend model for an intrinsic but non-stationary random function. The trend model is limited to the two terms $d(x) = a_0 + a_1 f_1(x)$, with the term $f_1(x)$ set equal to a secondary, or external, variable instead of as a function of the spatial coordinates. The smooth variability of the secondary variable $V(x)$ is related to that primary variable $Z(x)$ being estimated [17]. The trend model is then:

$$E[Z(x)] = \mu_Z(x) = a_0 + a_1 v(x), \quad (25)$$

where $v(x)$ is assumed to reflect the spatial trends of the z variability up to a linear rescaling of units (corresponding to the two parameters a_0 and a_1). The estimator is:

$$Z_{KED}^*(x_0) = \sum_{\alpha=1}^n \omega_{\alpha} Z(x_{\alpha}). \quad (26)$$

The solution of the following kriging system with $(n + 2)$ linear equation are the kriging weights ω_{α} :

$$\begin{cases} \sum_{\beta=1}^n \omega_{\beta} \gamma_{\alpha\beta} + \lambda_0 + \lambda_1 d(x_{\alpha}) = \gamma_{\alpha 0}, \\ \sum_{\alpha=1}^n \omega_{\alpha} = 1, \\ \sum_{\alpha=1}^n \omega_{\alpha} d(x_{\alpha}) = d(x_0). \end{cases} \quad (27)$$

The kriging with an external drift is a particular case of the kriging with a trend system where the trend function is a linear and the trend component $f_1(x)$ at any location x is identified with value $v(x)$. The trend model (26) states that the local average of the primary variable $z(x)$ is linearly related to the secondary datum $v(x)$. It is critical to validate that assumption. For instance, it makes sense to assume that the seismic travel time to a reflecting horizon is linearly related to the depth of that horizon. Seismic data can then be used as an external drift for mapping from a few boreholes data [8, 12]. Another example relates to using elevation data to model a trend in meteorological data [10, 17], hydrogeology [3], soil science [14], biology [13], etc.

The relation between primary and secondary data must be linear, otherwise an appropriate transformation of the secondary variable could make that relation linear form. The value of v variable must be available at all data location x_{α} and at all location x_0 being estimated.

Figure 4 A. shows the estimation of Pannonian sediment thickness [m] based on only drill-hole data as primary variable by IRF-k kriging method [6]. To improve the spatial modelling, the secondary exhaustive information, represented by seismic TWT map [ms] (**B.**), was incorporated in estimation by kriging with external drift (**C.**). Maps **A.** and **C.** show the same long-range features of spatial variability, but kriging with external drift yields more

local details. Such short short-range variation results from the local re-evaluation of the linear regression of primary and secondary variable. The larger variation of the estimates obtained by kriging with external drift relates to the larger slopes of the trend model. The combination of a steep trend model and low seismic values (look at south-west area within geological boundaries) yields negative estimates of kriging with external drift of the sediment thickness, what is totally unacceptable! On the other side, the black pixels show the estimates of thickness higher than the maximum one observed in the experimental data. This extrapolation of high value is correct and shows a depocentrum of the sediments, where is no drillhole located.

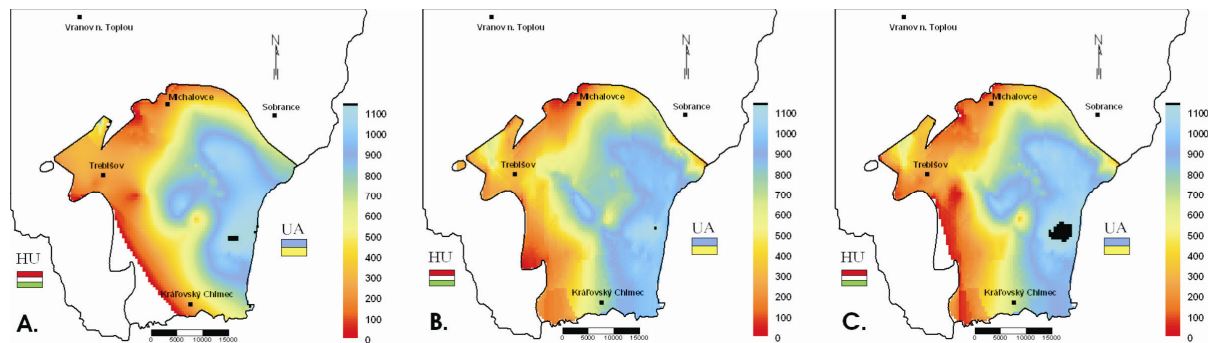


Figure 4. Map of Pannonian sediment thickness obtained using IRF-k kriging (A.) and accounting for the seismic map (B.) as kriging with external drift (C.)

6 Summary

Cokriging is much more demanding than kriging of one variable because of $K(K+1)/2$ single and cross-variograms must be inferred and jointly modelled, and a large cokriging system must be solved. In case of low correlation between primary and secondary data in combination with isotopic sampling, (or at least isotopic sampling within search neighbourhood) and intrinsic model of regionalisation, the weights attached to the secondary variable will be systematically zero, and therefore the cokriging result is similar to the one of kriging.

In the presence of densely sampled secondary information, collocated cokriging is a valuable alternative to cokriging because avoid instability caused by highly redundant secondary data, it is faster, since calls for a smaller cokriging system, and it doesn't call for the secondary variogram function for lag $h > 0$. In case of Markov-Bayes approximation, collocated cokriging does not require modelling of cross-variogram function and direct variogram function for primary data, because it is inferred from variogram function of secondary variable.

Collocated cokriging and kriging with an external drift are designed to incorporate exhaustively sampled secondary information. In case of kriging with external drift the secondary datum provides information only about the primary trend at given location. Because the secondary information influences strongly the kriging with external drift, especially when the slope of the local trend model is large, final interpretation of results should be validated by the physics of the studied phenomenon. The combination the large slopes of a trend model with low values of a strictly positive variable yields some negative estimates, especially in an extrapolation areas, what is unacceptable.

Reference

- 1 Chiles, J-P. & Delfiner, P.: *Geostatistics: Modelling Spatial Uncertainty*. John Wiley & Sons, Inc.. 1999. New York. ISBN 0-471-08315-1.
- 2 Daly, C. & Verly, G., W.: Geostatistics and data integration. In *Geostatistics for the Next Century* (pg. 94-107). Dimitrakopoulos, R. (ed.). Kluwers Academic Publisher. 1994.
- 3 Desbarats, A., J. et al.: On the kriging of the water table elevations using collateral information from a digital elevation model. *Journal of Hydrology*. **255**. 25 – 38. Elsevier. 2002. ISSN 022-1694.
- 4 Deutsch, C., V.: *Geostatistical Reservoir Modeling*. Oxford University Press, Inc. 2002. New York. ISBN 0-19-513806-6.
- 5 Deutsch, C., V. & Journel, A., G.: *GSLIB Geostatistical Software Library*. 2nd edition. Oxford University Press. 1998. Inc. New York. ISBN 0-19-510015-8.
- 6 Dowd, P., A.: *MINE5260 Non-Stationarity*. MSc. in Mineral Resources and Environmental Geostatistics. University of Leeds, Leeds. 2004. U.K.
- 7 Dowd, P., A.: *MINE5270 Multivariate Geostatistics*. MSc. in Mineral Resources and Environmental Geostatistics. University of Leeds, Leeds. 2004. U.K.
- 8 Galli, A. & Meunier, G.: Study of a gaz reservoir using the external drift method. In *Geostatistical case studies* (pg. 105-119). Armstrong, M. & Matheron, G. (eds.). Dordrecht Reidel Publishing Company, 1987.
- 9 Goovaerts, P.: *Geostatistics for Natural Resources Evaluation*. Oxford University Press, Inc. 1997. London. ISBN 0-19-511538-4.
- 10 Hlásny, T.: Modelling selected climate parameters in the ISATIS environment. *GIS... Ostrava 2007*. TU-VŠB, Ostrava. 2007. Ostrava. ISSN 1213-239X.
- 11 Isaaks, E., H. & Srivastava, R., M.: *An Introduction to Applied Geostatistics*. Oxford University Press, Inc. 1989. New York. ISBN-0-19-505012-6.
- 12 Moinard, L.: Application of kriging to the mapping of a reef from wireline logs and seismic data: a case history. In *Geostatistical case studies* (pg. 93-103). Armstrong, M. & Matheron, G. (eds.). Dordrecht Reidel Publishing Company. 1987.
- 13 Régneire, J. & Sharov, A.: Simulating temperature dependent ecological processes at the subcontinental scale: male gypsy moth flight phenology as an example. *Int. Journal of Biometeorology*. **42**. 146 – 152. 1999.
- 14 Snepvangers, J. C., et al.: Soil water content interpolation using spatio-temporal kriging with external drift. *Geoderma*. **112**. 253 – 271. Elsevier. 2003. ISSN 0016-7061.
- 15 Vizi, L.: Nelineárny geoštatistický odhad vyťažiteľných zásob (Jelšava - obzor 220). In *GIS... Ostrava 2007*. 17 p. VŠB TU Ostrava. 2007. - ISBN 1213-2454.
- 16 Vizi, L., et al.: GIS for geological survey data and geostatistical model of the Kišovce-Švábovce Mn deposit. In *GIS... Ostrava 2003*. 13 p. VŠB TU Ostrava. 2003. ISSN 1213-239X.
- 17 Wackernagel, H.: *Multivariate Geostatistics*. 3rd edition. Springer-Verlag, Berlin. 2003. Germany. ISBN 3540441425.
- 18 Webster, R. & Oliver, M.: *Geostatistics for Environmental Scientists*. John Wiley & Sons, Inc. 2001. New York. ISBN 0-471-96553-7.