

Parameters of interpolation methods of creation of digital model of landscape

Róbert Fencík, Margita Vajsáblóvá
Faculty of Civil Engineering, Slovak University of Technology
Bratislava, Slovakia
robert.fencik@stuba.sk, gitka@math.sk

INTRODUCTION

Digital model of landscape is an important part within creation of geo-information systems used in a public administration and in a commercial sphere. It is an important tool in applications, which model an Earth's surface like geomorphology, hydrology, geology, cartography, ecology, etc. Many software products offer different interpolation methods for creation of digital model of landscape. Its accuracy and quality is impacted by selection of an interpolation method and precision input data. A choice of the interpolation method and her parameters is depends on many factors, for example a type of interpolate phenomenon, character of surface (horizontal and elevation segmentation, break lines etc.) or a purpose of a digital model (Fencík, 2005). The most used interpolation methods are the method of inverse distance weighting, the method of natural neighbours, trend, kriging and the spline method (Burrough, 1998, Booth, 2000).

INTERPOLATION METHODS

The paper is aimed to analysis of digital models of landscape created by geostatistic methods with a different type of modelling parameters and factoring of autocorrelation input data depend on they a mutual distance.

The method of inverse distance weighting determines elevation of grid cells using a value weighted average of elevation points lying in a fixed distance from a centre of cells. A weight used for calculation is the reciprocal value of the measuring distance from a local estimation with a power p . A searching radius (fixed or variable) defines points, which been included in an interpolation process.

The method of natural neighbours is on the edge of TIN modelling and an interpolation process of a raster. The value of the cells of the raster model is interpolated from input data, which are in natural neighbourhood of cells. The method in first stage creates a triangular network based on Delaunay criterion and the weight of the neighbouring points is determined by Thiessen-Voronoi technique. Next step is determining the local estimation of value by the linear interpolation from the points at vertexes of a triangle.

We could calculate the points of an interpolation surface by a polynomial function called trend using a multiple regression. The most suitable coefficients for each polynomial degree have been chosen by the least square method.

The spline method belongs to methods of patch modelling. Interpolation or approximation curves use the most frequently polynomial segments, which are connected in the nodes. A general shape of the polynomial segment depends on selection of a spline function, control points and the parameters. We use B-splines for the surface interpolation. The most useable types of splines in computer environment are the regular and the tight spline.

Kriging belongs among geostatistic methods where is supposed an autocorrelation among input data. The paper is aimed to kriging and we pay attention to description her types and parameters in the next section.

KRIGING METHOD

The interpolation surface by kriging is created from three components: drift, regional variable and random noises. The kriging method is the exact interpolation method, which providing very precise results unless is not a noise ratio so large.

Components of the interpolation surface are defined by a variogram, which provides quantification of the correlation by two random variables (Scheer, 2001, Stehlíková, 2002). We talk about the autocorrelation of input data and assume that with decreasing input data distance h is also decreasing of spatial variability of spatial values and simultaneously increasing their autocorrelation. It means, that covariance between whichever pair of spatial variables x , $x+h$ depends on vector h , which separate each from other. We could estimate the variogram $\gamma(h)$ from actual data. The result is the experimental variogram, which we calculate as an average of square differences of experimental data of all pair values elevations $z(x_i)$, $z(x_{i+h})$ separated by vector h :

$$\gamma^*(h) = \frac{1}{2n} \sum_{i=1}^n (z_{(x_i)} - z_{(x_{i+h})})^2 \quad (1)$$

where n is a number of pair spatial values.

Important step within the geostatistic analysis is adjustment of the experimental model by appropriate type of the variogram model. Each model has its own basic structure, which is the function of the distance among data and we express her by the parameters *sill* (upper limit of the variogram model) and *range* (boundary distance between selection locality at which the model reaching of the maximum value or „sill“). The example of a variogram with typical components is in figure 1.

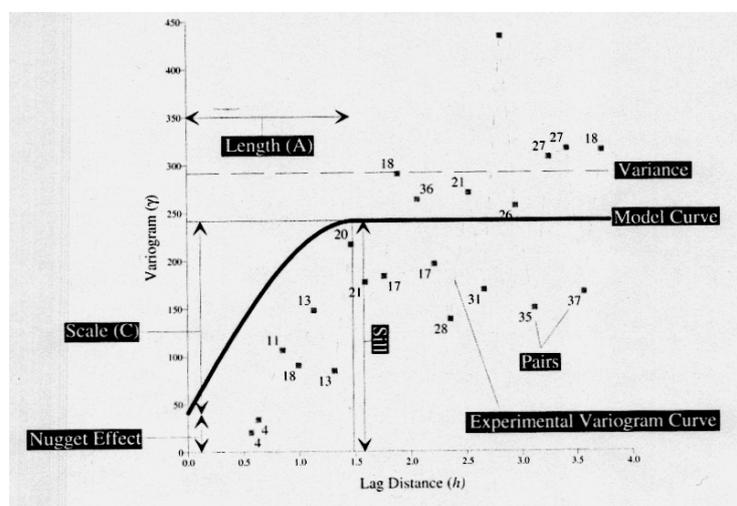


Figure 1: Example of a variogram model (Golden Software, 1999)

If $Z(x)$ is a standard random function with the variogram $\gamma(h)$, then during selection of the correspondent model for his adjustment we have to consider:

- variation and tendency of average incremental values $Z_{(x)}-Z_{(x+h)}$ within of vector h (average increments we mark as *drift*),
- properties of the variogram in a proximity origin (parabolical, linear, nugget-effect – discontinuity of the variogram with the origin),
- presence or absence of *sill* value at during increasing $\gamma_{(h)}$ ($\gamma_{(h)} = \text{constant}$ if $h > a$, where a is range value).

We can sort variogram models into next groups:

- *models with sill*:
 - with linear behaviour in the proximity origin (spherical model, exponential model),
 - with parabolic behaviour in the origin neighbourhood (Gauss model),
- *models without sill*:
 - linear model,
 - logarithmic model,
- *models with nugget-effect*.

We decided to use software Surfer 8 for modelling interpolation surfaces, which includes an extensive variogram modelling subsystem. Users create a concrete type of the variogram with choices of Length (A) (range), Scale(C) and Nugget effect (sum of them is a sill). Except variograms the very important choices in gridding methods are searching of input data and breaklines.

Kriging interpolation method constructs an internal matrix based on the number of points including in interpolating a grid node value, which may be defined by several options of *Search*. The *No Search* option tells Surfer to use all data points, we can use it when the absolute maximum number of data points is 750. The *number of sectors to search* option divides a search area into smaller sections to which we can apply the following search rules:

- the Maximum number of data to use ALL sectors,
- the Maximum number of data to use EACH sectors,
- the Minimum number of data to use all sectors.

The *Search Ellipse* defines neighbourhood of points to consider when interpolating each grid node. Data points outside the search ellipse are not considered during grid node interpolation. Search ellipses are specified by defining radius and angle between the positive X axis and the ellipse axis. In most cases it works well to set the search ellipse ratio and direction to coincide with the anisotropy ratio and directions. Anisotropy is used if data are representation of processes with preferred orientations.

Breaklines are components of discontinuities in the surface. A breakline is a tree-dimensional blanking file that defines a line with X, Y and Z values at each vertex. Surfer uses linear interpolation to determine the values between breakline vertices when gridding.

DEM OF MODRA-HARMÓNIA TERRITORY

An experimental territory was a part of the cadastral district MODRA with dimensions 1000 x 1000 meters. It's situated approximately 20 kilometres on the north from Bratislava. A relief of the experimental territory is irregular and broken. The larger part of the experimental territory is under forest cover.

As input data for creation of DEM we used data processed with the digital photogrammetry. A photograph scale of aerial photographs was 1:7 500 – 1:8 000. A file of input data contains 4 792 points, their position is showed in figure 2.

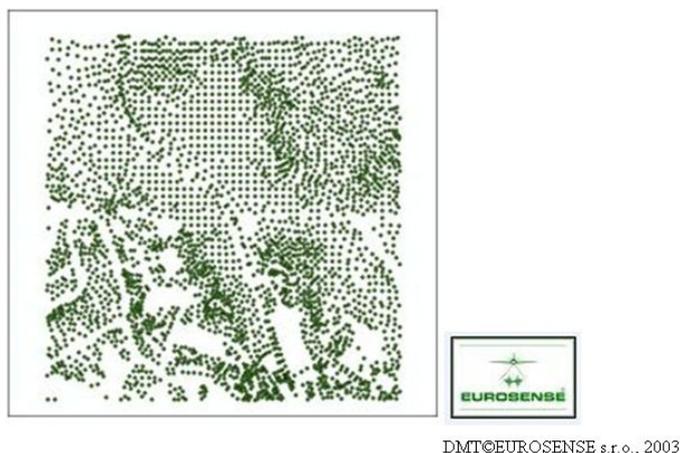


Figure 2: Position of points on the experimental territory

We used for the creation of the DEM kriging method using SURFER 8 (Golden software, 1999). We created the raster DEM with a grid 1m. We considered the breaklines within the creation of the DEM. We defined a few geometric types of the breaklines, we used: a road, a quarry and a lake.

The software SURFER 8 offers two basic types of the kriging methods: *point kriging* and *block kriging*. We can choose a drift parameter and a variogram model for each method. We used for the creation of the DEM following types of the methods:

- Point kriging with linear variogram model without a drift (fig. 3),
- Block kriging with:
 - linear variogram model without the drift (fig. 4),
 - linear variogram model with the linear drift (fig. 5),
 - Gaussian variogram model (fig. 6),
 - quadratic variogram model,
 - spherical variogram model (fig. 7),
 - cubic variogram model.

A linear variogram model is characterized by a slope. We used the slope value 1 in our non-anisotropic data. *Point Kriging* estimates the values of the points at the grid nodes. *Block Kriging* estimates the average value of the rectangular blocks centered on the grid nodes. A visual effect of *Point Kriging* and *Block Kriging* with the linear variogram model is not very different. Using non-linear variogram model in kriging method resulted in a rougher surface.

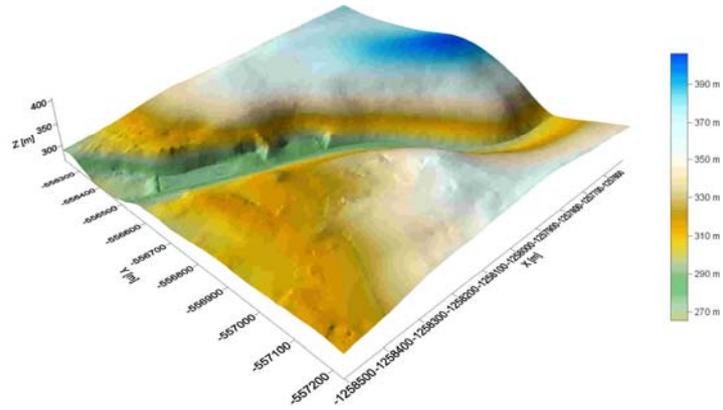


Figure 3: Model created by kriging with a point linear variogram

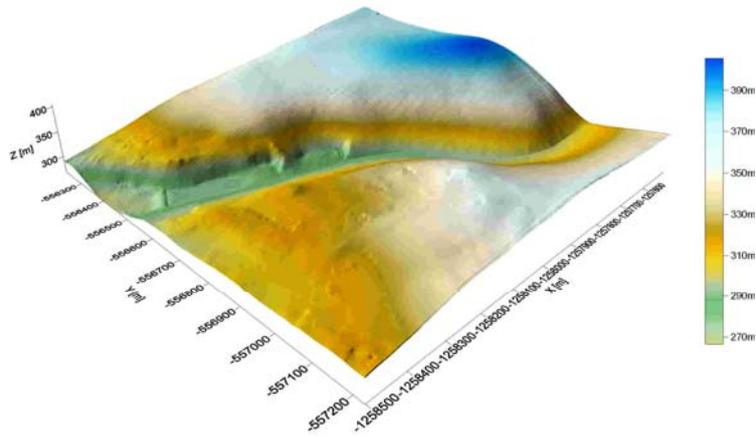


Figure 4: Model created by kriging with a block linear variogram without drift

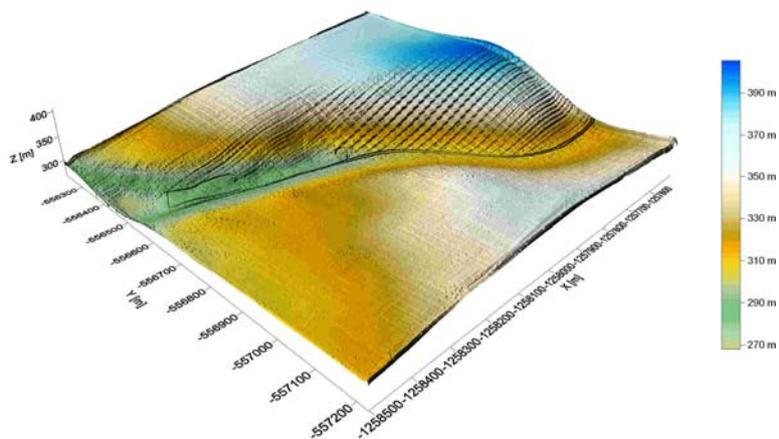


Figure 5: Model created by kriging with a Gaussian variogram

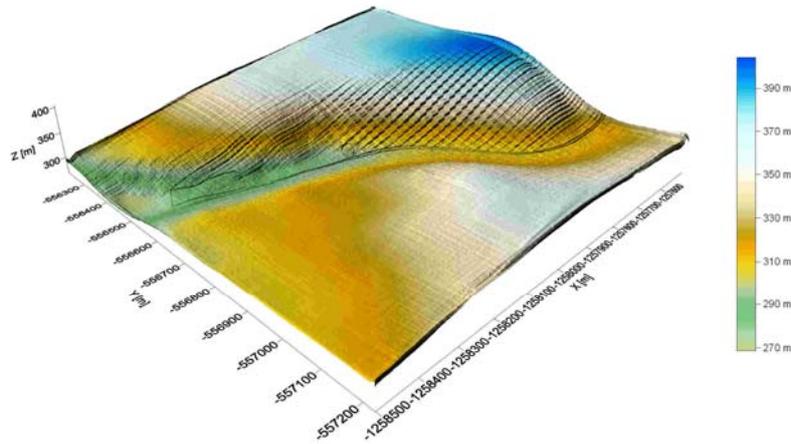


Figure 6: Model created by kriging with a spherical variogram

ANALYSIS OF DEM ACCURACY

We could evaluate the DEM accuracy from a few aspects, for example: the accuracy in control points, an error autocorrelation, the relative accuracy in relation to a neighbourhood etc. We considered the accuracy of each model on the dataset of selected control points of the experimental territory. The file of control points (210) we got from a direct observation in the field by the electronic tacheometry (fig.7). The results of the tested DEM and their accuracy characteristics are showed in Table 1.

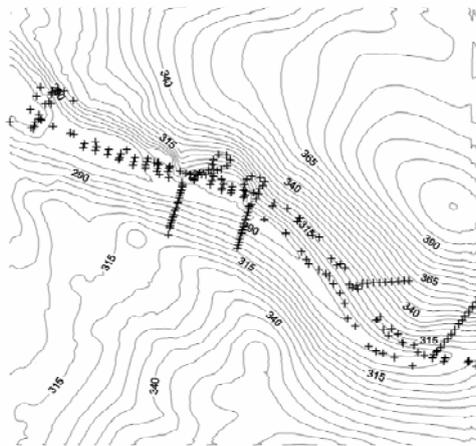


Figure 7: Control points in the experimental territory

We evaluated the results (tab. 1) according to a few statistical indexes and the elevation accuracy criterions of the irregular relief. The models created using the kriging method with linear variogram complied with the criterions. A large *Range* of the deviations have been caused by data capture in woodland during a vegetation season.

Table 1: Comparison of Z coordinate values of the DEM with control points

INTERPOLATION METHODS		Range [m]	Root Mean Square [m]	Variance [m]	Standard deviation [m]	Mean [m]
POINT KRIGING without drift		15.64	4.44	13.13	3.6	-2.58
BLOCK KRIGING	Linear without drift	16.22	4.88	12.31	3.51	-3.40
	Linear with drift	16.22	4.88	12.42	3.52	-3.39
	Gaussian	24.49	10.30	31.09	5.57	-8.66
	Quadratic	24.49	11.06	29.38	5.42	-9.65
	Spherical	24.49	10.99	29.34	5.42	-9.57
	Cubic	24.49	11.03	29.40	5.42	-9.62

We used *Cross Validation* during the creation of the DEM. *Cross Validation* can be considered an objective method of assessing the quality of a gridding method or compare the relative quality two or more candidate gridding methods. The cross validation process can be realized for all dataset or on points, which are selected by following options:

- Number of random points to validate,
- Select validation points within limits of direction X, Y, Z,
- Exclude data validation within distance of the validate point.

We used 4 000 points for the cross validation process. We evaluated the basic accuracy characteristics of the DEM for determination of suitability mentioned interpolation methods. We determined for each method following accuracy characteristics: *Range*, *Root Mean Square*, *Variance*, *Standard deviation* and *Mean*. The results of our analysis using *Cross Validation* are in Table 2.

Table 2: Cross validation of grid values

INTERPOLATION METHODS		Range [m]	Root Mean Square [m]	Variance [m]	Standard deviation [m]	Mean [m]
POINT KRIGING without drift		7.12	0.47	0.22	0.47	0.01
BLOCK KRIGING	Linear without drift	6.51	0.43	0.19	0.43	-0.03
	Linear with drift	4.56	0.41	0.17	0.41	-0.02
	Gaussian	12.00	2.46	3.29	1.81	1.66
	Quadratic	33.77	3.08	9.47	3.08	-0.11
	Spherical	5.29	0.43	0.18	0.43	-0.01
	Cubic	37.66	3.07	9.41	3.06	-0.18

Evaluation of the models in Table 2 confirmed suitability of the linear variogram, potentially the spherical variogram for our data. The value of RMS indicates a reliability of Z coordinates in the input dataset. The autocorrelation of input data represented by the others non-linear models showed as unfounded.

CONCLUSION

During studying of spatial variability we work usually with the omnidirectional variogram (requirement is the isotropic model) where we do not consider orientation (direction) of vector h , but only his size. The result is combination all orientations (directions) in one variogram, which could be understandable as “average” of different directional variograms. The important factor from the point of view creation of digital model of landscape is the autocorrelative prediction of input data. The aim of our paper is deeper research of properties of the spatial model created by geostatistic methods, especially by parts of kriging. We want to point out on variety of digital models, which were created by different types of variograms. We verified accuracy of the raster elevation calculated with geostatistic methods by comparison with the file of control points from a direct observation. This analysis shows suitability of selected parameters of the model for specified features of the modelling territory. The results of our research shows, that the most suitable type of the variogram model in kriging interpolation method is linear type.

ACKNOWLEDGEMENT

The authors thanks the Grant Agency of the Slovak Republic, VEGA (Grant No. 1/1034/04) for supporting this work.

BIBLIOGRAPHY

- Booth, B., 2000: Using of ArcGIS 3D Analyst. User guide. ESRI. CA, USA.
- Burrough, P.A., McDonell, R.A., 1998: Principles of geographical information systems. Oxford University Press.
- Fencík, R., Vajsáblová, M., Vaníková, E., 2005: Comparison of interpolating methods of creation of DEM. In: 16th Cartographic conference, University of Defence press. Brno, Czech Republic. pp. 77-87.
- Golden Software, 1999: Surfer Win-32 Software ‘Help menus’.Colorado, USA.
- Scheer, L., 2001: Using of Geostatistics in Investigation of Condition of Forest Geosystems. Technical University press. Zvolen, Slovakia, pp. 1-35.
- Stehlíková, B., 2002: Spatial statistics. Agricultural University press. Nitra, Slovakia, pp.73-85.