v.kriging: Preliminary case studies

Eva Stopková

The purpose of these two case studies is to outline a brief introduction how to use the module *v.kriging*. However, no exact rules leading to optimal result of the interpolation cannot be specified, as modelling of any phenomenon depends on statistical characteristics of particular dataset. To get relevant interpolated (2D/3D) raster model, it is necessary to try different anisotropic ratios, testing various functions for theoretical variogram modelling and careful analysis of the results.

3D kriging: a brief introduction

Input layer should contain 3D coordinates (xyz) and values to be interpolated (in attribute table). Basically, definition of these parameters is required:

```
v.kriging phase=initial in=input_layer icol=column report=file.txt
```

```
v.kriging in=input_layer phase=middle icol=column file=png \
hz_fun=exponential vert_fun=exponential hz_range=double vert_range=double -u
```

```
v.kriging in=input_layer phase=final icol=column file=png out=raster \
final_fun=exponential final_range=double crossval=crossval_file.txt
```

In the middle phase, there is possible also to modify nugget effect (default: 0.0) and sill (default: calculated from variogram values, more details in (*Stopková, 2014*)).

Case study: Slovakia 3D precipitation

3D interpolation is based on the input points of annual precipitation dataset (Mitasova and Hofierka, 2004). As the algorithm still needs to be optimized for large datasets, the points in smaller region (*Table 1*) were extracted.

N = 5 468 000 m
W = 4 361 000 m   E = 4 465 500 m
S = 5 374 500 m
top: 2 250 m   bottom: 200 m

Table 1: Smaller region extent (resolution hz: 500 m, vert: 100 m)

In the initial phase, **experimental variograms** (horizontal and vertical) were computed:

```
v.kriging phase=initial in=precip3d@PERMANENT ic=precip report=precip3d.txt
```

In the middle phase, there were empirically estimated function types and coefficients of **theoretical variograms**:

```
v.kriging in=precip3d@PERMANENT phase=middle hz_fun=exponential
vert_fun=gaussian ic=precip file=png hz_range=20000. vert_range=2200. -u
```

Horizontal and vertical variograms (experimental and theoretical) are available in *Figure 1*. These variograms provide the base for computation of univariate anisotropic theoretical variogram in the final phase (*Figure 1a*) that results into the 3D raster.

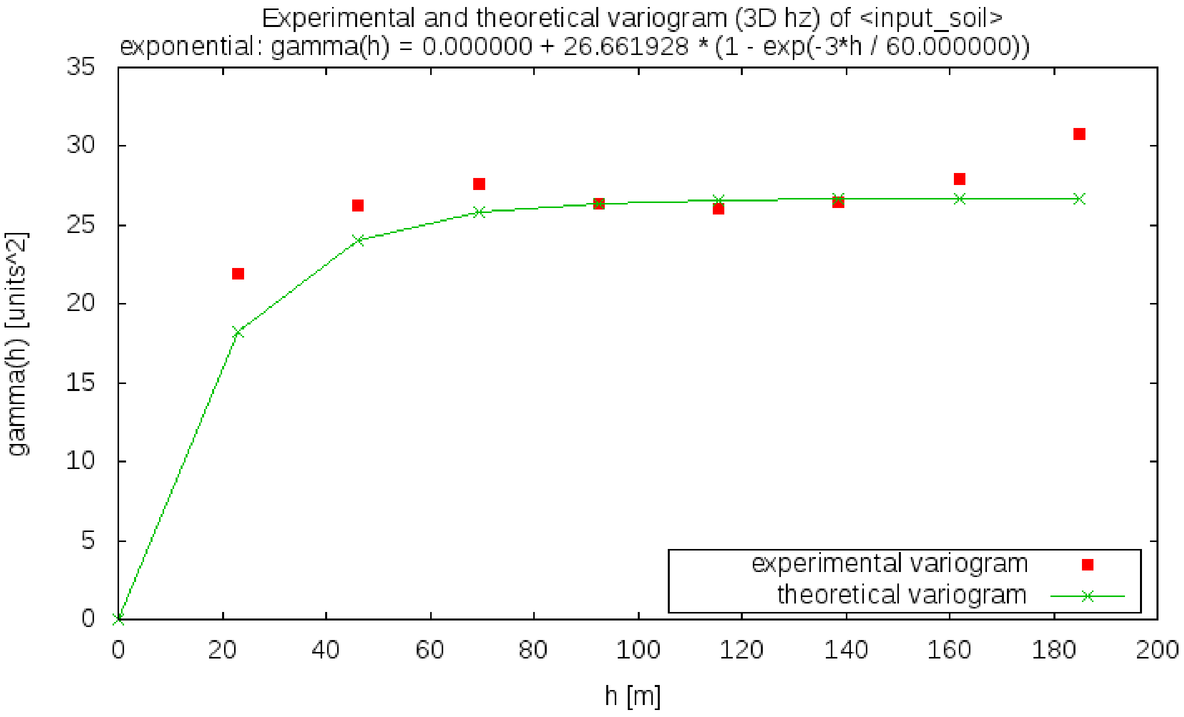
```
v.kriging in=precip3d@PERMANENT phase=final hz_fun=exponential
vert_fun=gaussian ic=precip  file=png hz_range=100000. vert_range=1600. -u
```

The results[1] were compared with:

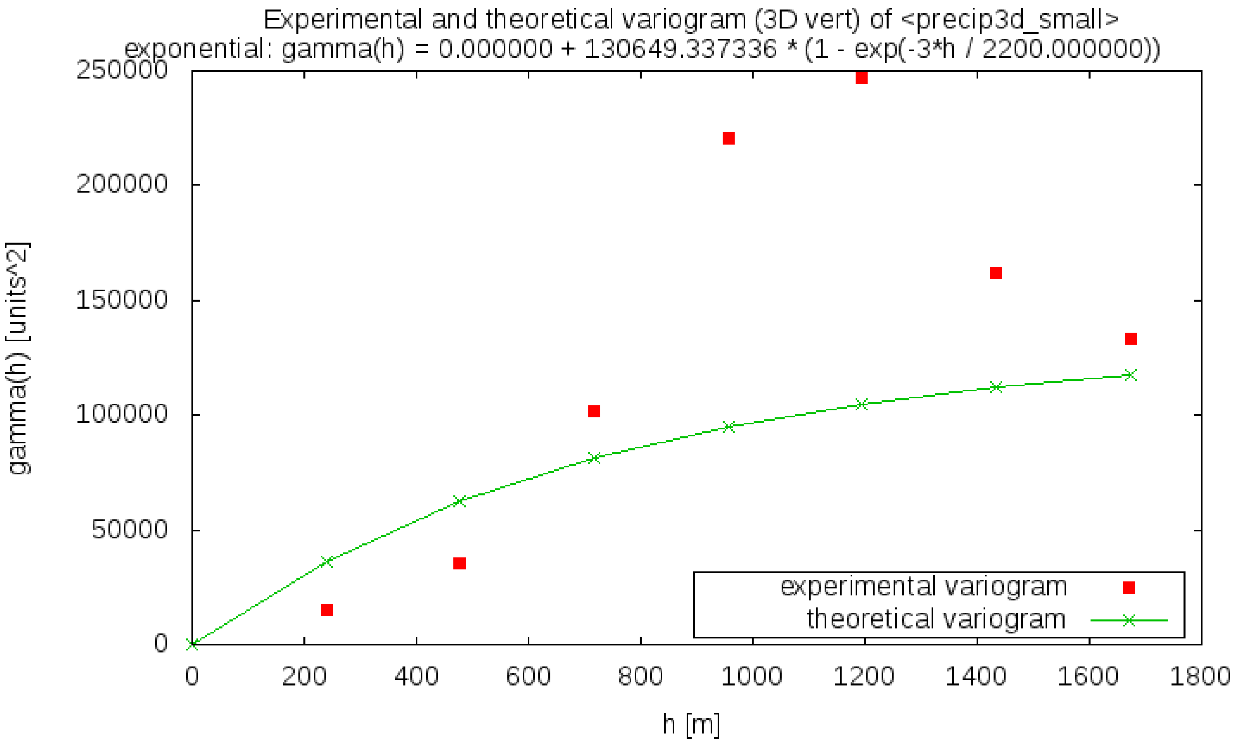
- the values interpolated using *v.vol.rst*,
- the values interpolated using two-dimensional mode of *v.kriging*.

---

[1] Only cross-sections of 3D rasters were compared because of the distribution of the points. They are positioned in three-dimensional space (thus interpolated value can be determined by function of three spatial coordinates), but they are located just on the terrain. Therefore interpolation above and below the terrain would become imprecise in deeper/higher areas of the dataset.

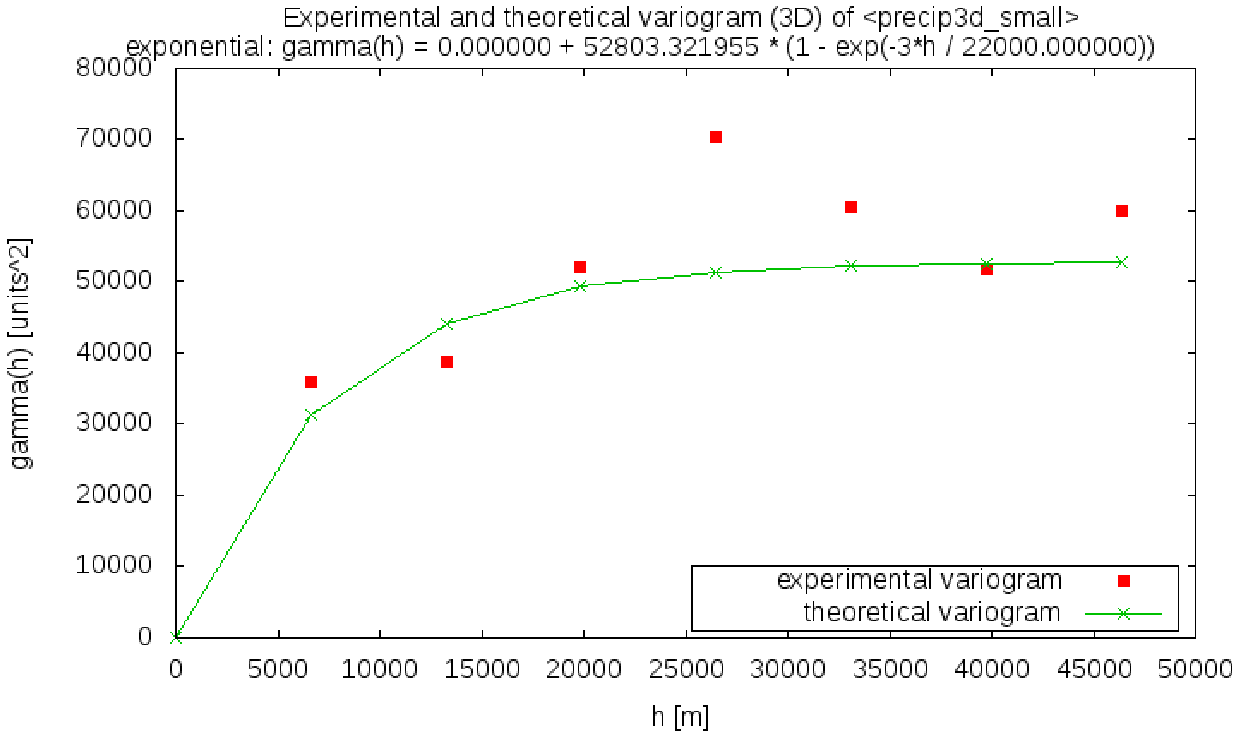


a: horizontal direction

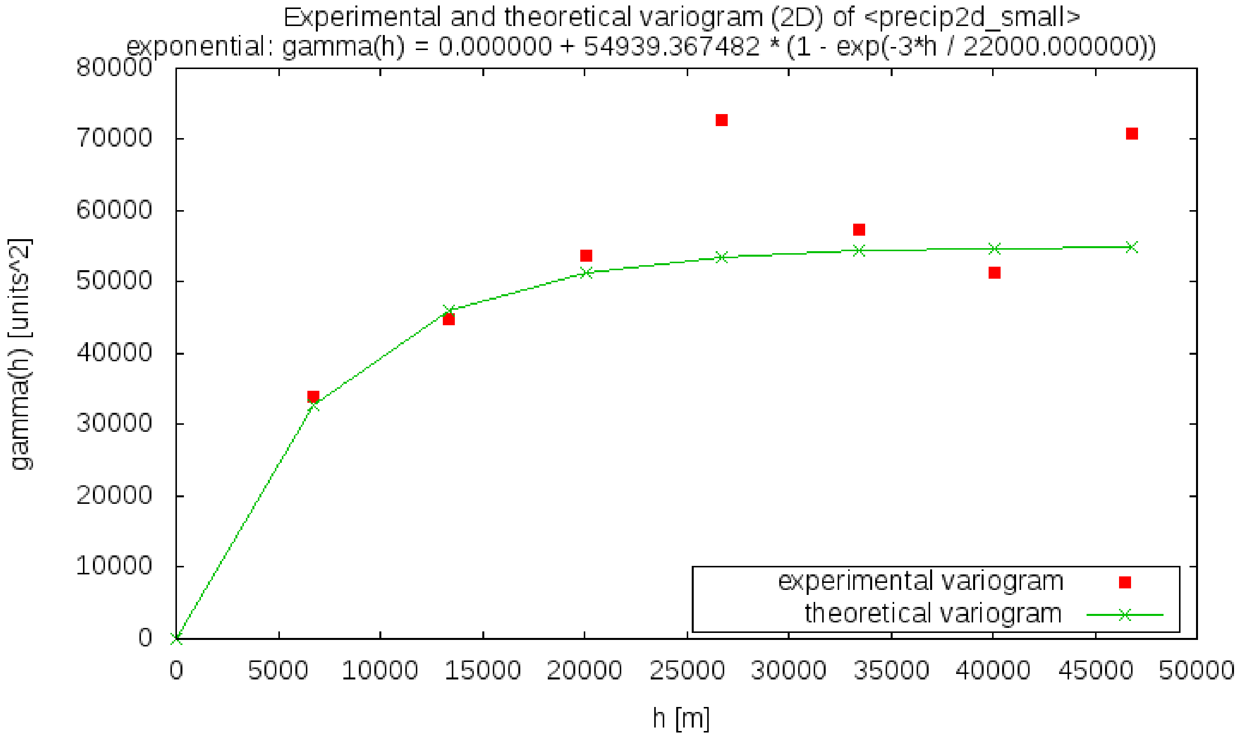


b: vertical direction

Figure 1: Experimental and theoretical variogram in horizontal and vertical direction



a: Anisotropic variogram of three-dimensional data



b: Variogram of two-dimensional data

Figure 2: Experimental and theoretical variogram of the reduced dataset

Comparison of the results RST interpolation was performed using modified settings (tension and smoothing parameters) according to (*Neteler and Mitasova, 2004*, page 173):

```
v.vol.rst -c input="precip3dPERMANENT" wcolumn="precip" tension=100.
smooth=0. \
cvdev="cxvalidation_rst_final" segmax=50 npmin=200 npmax=700 wscale=1.0
zscale=50
```

Interpolation in two-dimensional space was performed in 2D mode of *v.kriging* using following commands:

```
v.kriging phase=initial input=precip2d_small icolumn=dbl_4
report=precip2d_exp1.txt --o -2 v.kriging phase=final input=precip2d_small

icolumn=dbl_4 file=png final_function=exponential final_range=22000.
output=precip_exp1_rev crossval=precip2d_xval_exp1.txt --o -2
```

Statistical characteristics of the cross-validation results in *Table 2* vary probably because of the difference in interpolation algorithm (RST vs. kriging). The differences in two-dimensional and three-dimensional kriging interpolation (that are supposed to be identical on the cross-section with the terrain) could be based on the dataset configuration. As mentioned above, although the dataset is positioned in three-dimensional space, input values are not truly vertically stratified. This could influence the vertical variogram modelling, as proven by ongoing module testing on synthetic datasets.

	v.vol.rst	3D exponential	2D exponential
Minimum [mm]	-908.303	-33.144	-33.262
Maximum [mm]	257.69	20.490	20.144
Mean [mm]	0.554	0.218	0.295
Variance [mm^2]	28891.6	39.326	43.793
Std. deviation [mm]	169.975	6.271	6.618

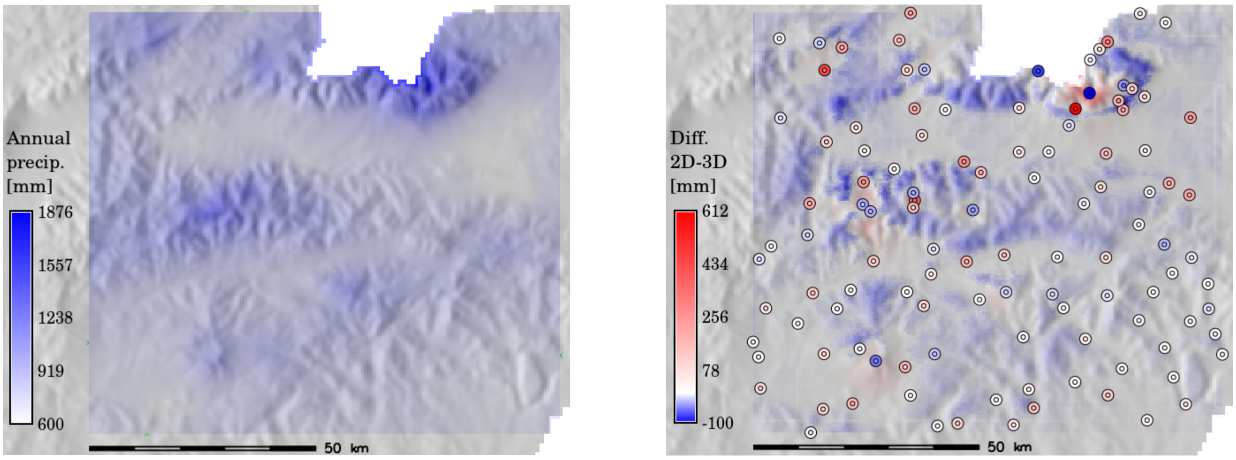
Table 2: Statistical characteristics of the cross-validation results

Table 3 summarizes the differences between the cross-sections. These might be considered as too high, but they were probably caused by vertical modelling of the planar data, despite of positions of the points in three-dimensional space.

	v.vol.rst	Exponential 3D		Exponential 2D		Differences (2D-3D)
		Value	Diff.	Value	Diff.	
# of cells	37639	37639		37639		37639
Minimum [mm]	600.922	599.579	-454.706	592.657	-660.768	-100.421
Maximum [mm]	1675.090	1875.990	261.531	2057.74	68.559	612.461
Mean [mm]	883.909	862.077	-29.88	869.01	-24.073	-4.711
Variance mm^2	11102	21942	1784.56	27780.9	1813.92	689.308
Std. deviation [mm]	105.366	148.128	42.244	166.676	42.590	26.255

Table 3: Statistical characteristics of the results and of the differences

Figure 2b shows that the differences have increased mainly in the areas with steeper slopes. In flatlands, even the differences seem to be stable. This could support the idea of “not really 3D data”. The figure presents the cross-validation results as well; the behaviour of 2D (bigger circles) and 3D (smaller circles) residuals seems to be almost identical.



a: The cross-section from interpolated 3D raster

b: The differences with 2D kriging result, together with the cross-validation results

Figure 3: The result of 3D interpolation

2D kriging: a brief introduction

Input layer should contain 2D coordinates (xy) and values to be interpolated (in attribute table). The commands can in general look like this:

```
v.kriging phase=initial in=input_layer icol=name report=file.txt file=png -2
```

```
v.kriging in=input_layer phase=final icol=name file=png out=name \
final_function=linear crossval=crossval_file.txt -2
```

Case study: elev_lid792_randpts

The case study is based on 500 random points that were extracted from input points of Digital Elevation Model (DEM) *elev_lid792_randpts* from the North Carolina dataset (Neteler and Mitasova, 2008). In the initial phase, temporary **experimental variogram** was computed:

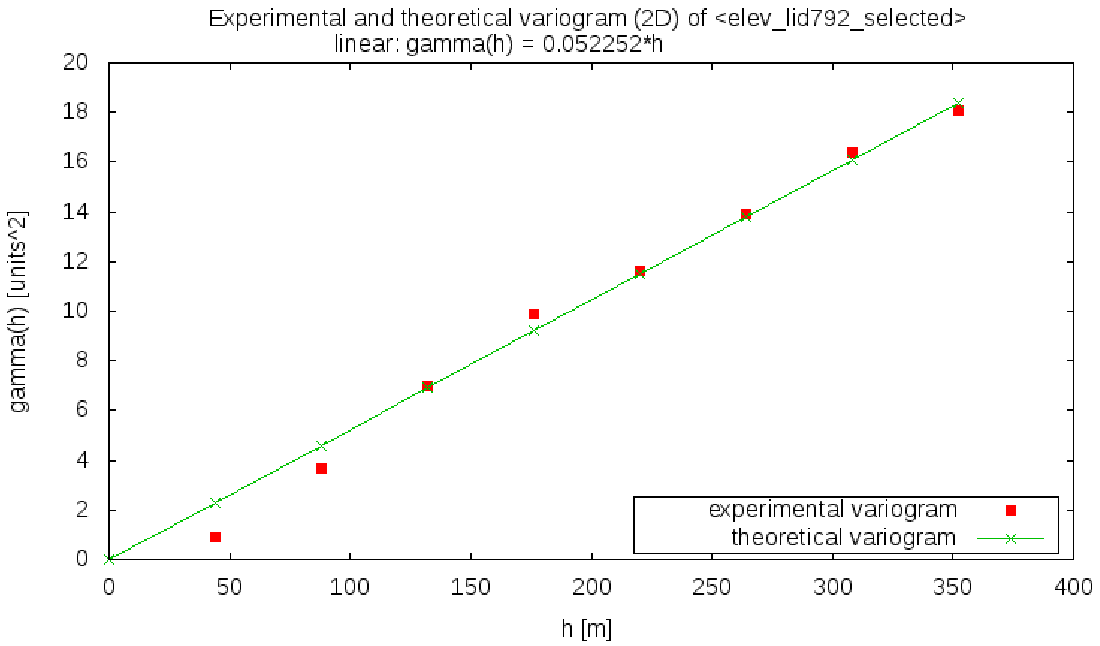
```
v.kriging phase=initial in=elev_lid792_selected ic=value azimuth=45.  td=45.
report=lid792_500_linear.txt -2 --o
```

Then, in final phase (middle phase is skipped in 2D kriging) the **theoretical variogram** was computed and interpolation of unknown values was performed:

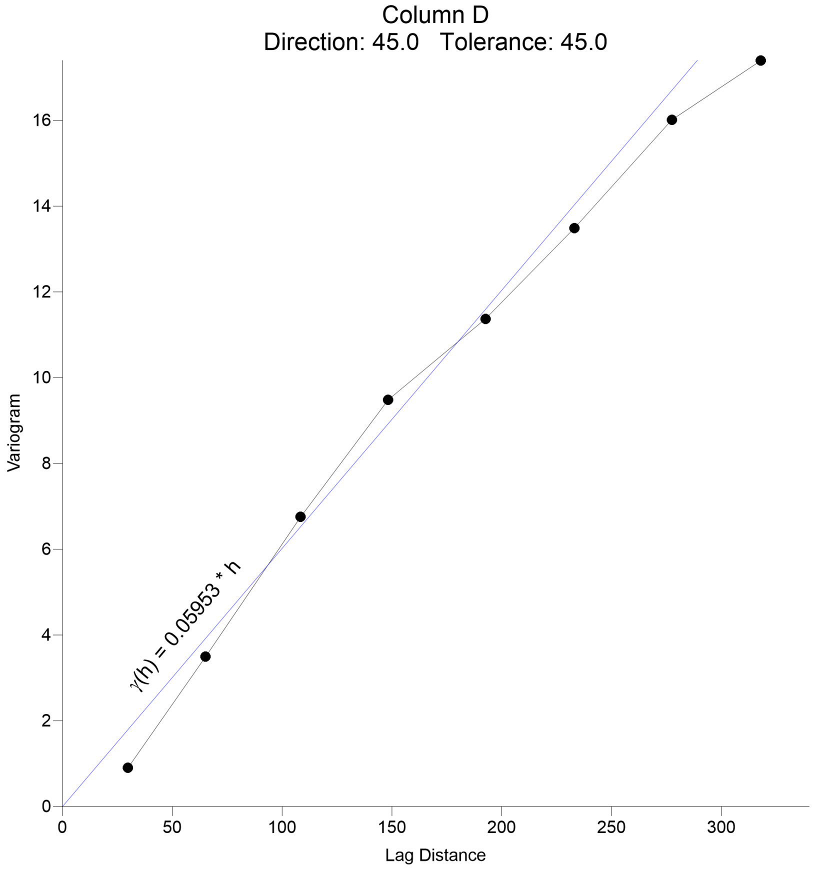
```
v.kriging in=elev_lid792_selected phase=final final_function=linear ic=value
file=png out=lid792_500_linear crossval=lid792_500_xval_linear.txt -2 --o
```

Variogram modelling was compared with the result of variogram analysis in *Surfer* (*Golden Software, Inc.*), see *Figure 4*. The difference in coefficient of the linear function might be caused by using slightly different lag size or by different approach to the computation (spatial index in *v.kriging*, optimization algorithms in *Surfer* (*Golden Software, Inc.*) etc.).

The results were compared with the values interpolated using the module *v.surf.rst* and kriging tool of *Surfer* (*Golden Software, Inc.*). *Figure 5* and *Figure 6* show the comparisons of interpolated DEM with the results of another interpolation tools. Statistical characteristics of the results are summarized in *Table 4*.



a: by *v.kriging*



b: by *Surfer* (*Golden Software, Inc.*

Figure 4: Comparison of horizontal variogram (experimental and theoretical)

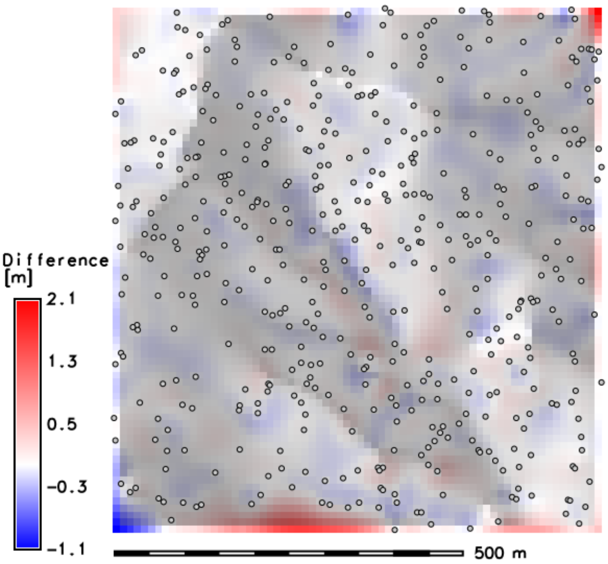


Figure 5: The difference: *v.kriging* and *v.surf.rst*

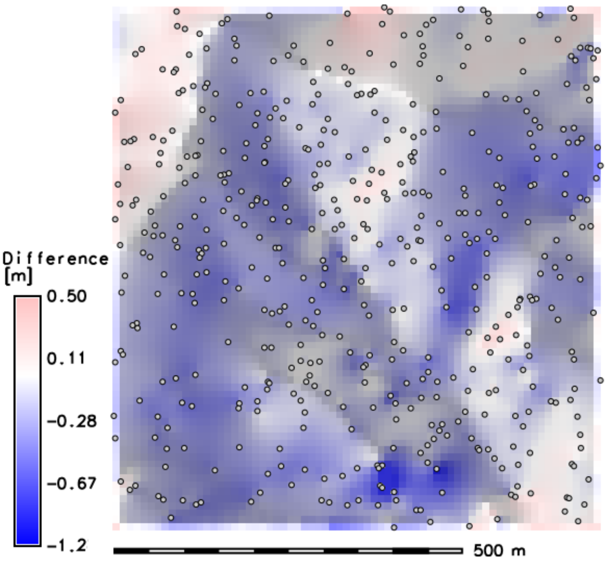


Figure 6: The difference: *v.kriging* and *Surfer*

Results	v.kriging	*v.surf.rst*		Surfer (*Golden Software, Inc.*)	
		Values	Differences	Values	Differences
Minimum [m]	105.114	105.061	-1.065	105.090	-1.181
Maximum [m]	131.510	131.570	2.072	131.510	0.522
Mean [m]	120.763	120.781	0.018	120.584	-0.178
Variance [m^2]	43.7367	43.2701	0.027244	44.2389	0.027244
Standard deviation [m]	6.613	6.578	0.165	6.651	0.213
95% quantile [m]	130.115	130.109	0.225	130.088	0.213
75% quantile [m]	126.580	126.587	0.046	126.434	-0.047
50% quantile [m]	121.315	121.325	0.000	121.080	-0.190
25% quantile [m]	115.749	115.786	-0.047	115.489	-0.328
5% quantile [m]	109.004	109.115	-0.139	108.800	-0.487

Table 4: Comparison of statistical characteristics of interpolated rasters

Cross-validation results of all the methods mentioned above and their statistical characteristics have been compared as well. Cross validation using *v.surf.rst* was performed with these settings:

```
v.surf.rst -c input="elev_lid792_selected" layer="1" zcolumn="value"\
cvdev="lid792_500_rst_xval" tension=40 segmax=30 npmin=120 \
dmin=5.000000 dmax=25.000000 zscale=1.0
```

Cross validation points are shown in *Figure 7* (bigger circles represent kriging results, smaller circles represent the results of RST). Statistical characteristics of the cross validation are summarized in *Table 5.*

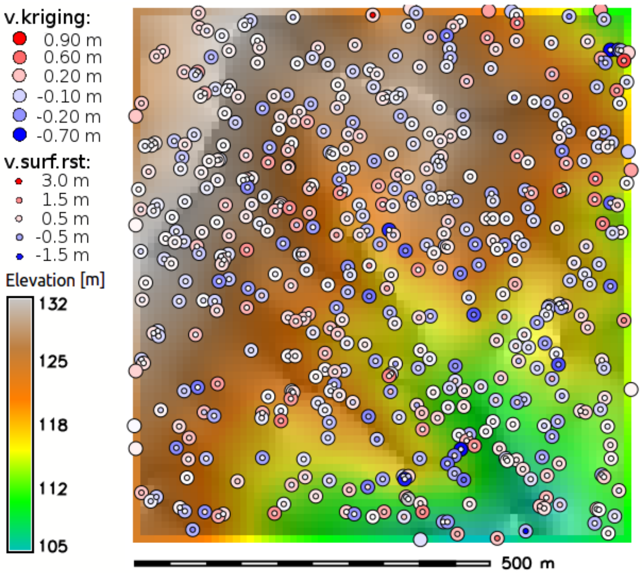


Figure 7: Cross validation by *v.kriging*

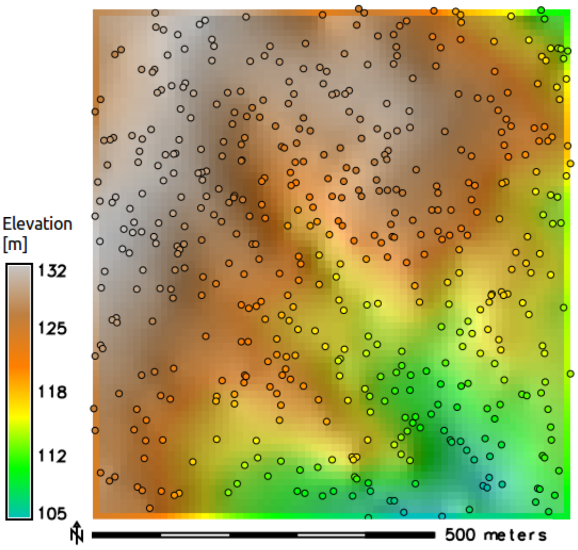


Figure 8: The DEM by using *v.kriging*

Cross validation	v.kriging	*v.surf.rst*	Surfer (*Golden Software, Inc.*)
Minimum [m]	-0.682	-1.593	-2.344
Maximum [m]	0.879	3.362	3.011
Mean [m]	0.005	0.004	0.003
Variance [m^2]	0.033089	0.143742	0.133809
Standard deviation [m]	0.182	0.379	0.366
95% quantile [m]	0.315	0.557	0.584
50% quantile [m]	0.007	-0.013	-0.020
25% quantile [m]	-0.102	-0.144	-0.126
5% quantile [m]	-0.290	-0.499	-0.445

Table 5: Statistical characteristics of the cross validation results

Conclusion

There is no general rule how to define variables for particular input dataset. The user should try several options before finding the best variogram fit. More detailed verification utilizing various datasets will be presented in separate paper.

References

Neteler, M. and Mitasova, H. (2004). *Open Source GIS: A GRASS GIS Approach.* 2nd Ed. 401 pp, Springer, New York. Online Supplement: http://www.grassbook.org

Stopkova, E. (2014). *Development and application of 3D analytical functions in spatial analyses* (Unpublished doctoral dissertation). The Department of Theoretical Geodesy, Faculty of Civil Engineering of Slovak University of Technology in Bratislava, Slovakia.

Datasets

Neteler, M. and Mitasova, H. (2008). Sample data North Carolina [nc_spm_08_grass7]. Available at `https://grass.osgeo.org/download/sample-data/`

Mitasova, H. and Hofierka, J. (2004). Slovakia Precipitation data [slovakia3d_grass]. Available at `https://grass.osgeo.org/download/sample-data/`