PREDICTION OF POSITION ERRORS OF POINTS IN FIRST ORDER TRIGONOMETRIC NETWORK

Stefan Miljković [1]
Ognjen Antonijević [2]
Milan Kilibarda [3]

UDK: 519.22/.25:55
DOI: 10.14415/konferencijaGFS2018.070

Summary: *During the establishment of first order trigonometric network in Serbia, the influence of Earth's gravitational field on measurements was not considered. As a consequence, all points have position errors. With additional measurements and calculations, these errors were later determined for X and Y coordinates of each point. This paper uses geostatistical methods (regression kriging) to estimate error prediction model for any location in Serbia.*

Keywords: *Geostatistics, coordinate error, error prediction, trigonometric network, regression kriging*

1. INTRODUCTION

First order trigonometric network is the basis for all cadastral measurements in Republic of Serbia. The network was established through several epochs during the period 1900-1951, using the principles of trigonometry. Angles and baselines were measured using theodolites and invarbasis wires. At the end of each establishment phase, network was adjusted using constrained adjustment and point coordinates were determined on Bessel ellipsoid. After completely establishing this first order trigonometric network, the network was densed (through several phases) by inserting new trigonometric points which comprise lower level networks. This process was further repeated and until finally traverse networks and line networks were established and used for cadastral surveying of the country [12].

During establishment of the first order network, only geometrical quantities on the physical surface of the Earth were measured - angles and baselines. During the 1950s the scientists discovered significant impact of Earth's gravitational field on measurements. Accounting for this impact is necessary in developing and establishing geodetic networks spreading over large areas of the Earth's surface. At the time of these discoveries, the aforementioned primary trigonometric network was already established

---

[1] Stefan Miljković, PhD student., University of Belgrade, Faculty of Civil Engineering, Bul. Kralja Aleksandra 73, Belgrade, Serbia, e – mail: stefan.miljkovic@grf.bg.ac.rs

[2] Ognjen Antonijević, PhD student., University of Belgrade, Faculty of Civil Engineering, Bul. Kralja Aleksandra 73, Belgrade, Serbia, e – mail: antonijevic.ognjen@gmail.com

[3] Milan Kilibarda, PhD, University of Belgrade, Faculty of Civil Engineering, Bul. Kralja Aleksandra 73, Belgrade, Serbia, e – mail: kili@grf.bg.ac.rs

and in everyday practical use. Errors resulting from not taking into account the impact of Earth's gravitational field were present in coordinates of first order network points, and in all subsequent networks derived from this network. All these networks and cadastral data from first and other early country-level surveying campaigns are still in use today [12] [13].

Subsequent analysis and measurements were conducted in order to determine errors resulting from neglecting the influence of Earth's gravitational field. Errors in X (Northing) and Y (Easting) direction were calculated for all points of the first order trigonometric network. This paper presents a geostatistical analysis of these errors and proposes a regression kriging model for prediction at regular grid with resolution of 200 m. This paper is organized as follows: Methods and Materials, Results and Conclusions.

2. METHODS AND MATERIALS

Main data set used for the realisation of this paper are the coordinates of the first order trigonometric network points, in Gauss Kruger projection (Fig. 1a), along with their respective errors on X and Y coordinate axes. These errors present the difference between coordinates in use and the corrected coordinates. Corrected coordinates were calculated after the network has been in use for some time, based on additional analyses and measurements to determine the influence of Earth gravitational field.

Total of 130 points with information about X and Y direction errors are analysed in this paper. Errors in X direction range from -0.717 m to 2.875 m, while in Y direction the range of errors is higher – from -4.724 m to 3.256 m. Root mean square error (RMSE) in X direction is 1.24 m and in Y direction is 1.90 m.

Besides this basic data set, three additional geostatistical predictors were considered for their possible influence on coordinate errors: EU-DEM v1.1 with 25m spatial resolution (Figure 1b), global geoid model (EGM 2008) height undulations grid with 1" resolution (Figure 1c), and distance from the measured network baselines (Figure 1d).

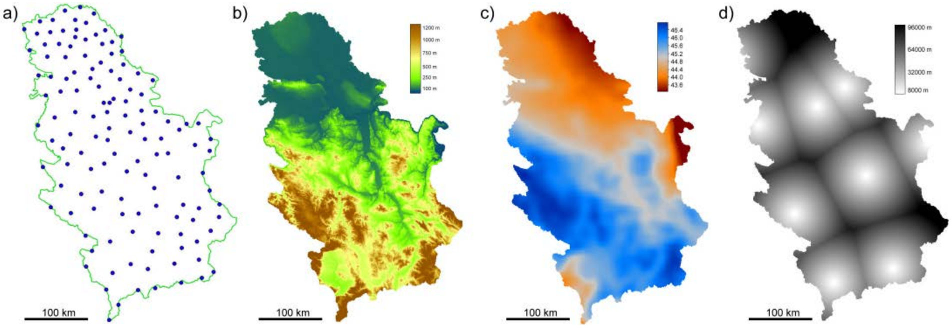


Figure 1. a) Basic data set b) DEM predictor, c) EGM 2008 1' predictor, and d) Distance grid from network baselines *predictor*

EU-DEM is a digital surface model (DSM) of EEA member and cooperating countries representing the first surface as illuminated by the sensors. It is a hybrid product based on SRTM and ASTER GDEM data fused by a weighted averaging approach. EU-DEM v1.1 is available in GeoTiff 32 bits format. It is a contiguous dataset divided into 100 x 100 km tiles, resulting in a total of 1992 tiles of 4000 x 4000 pixel at 25 m resolution with vertical accuracy: +/- 7 meters RMSE [6].

Earth Gravitational Model EGM2008 has been publicly released by the U.S. National Geospatial-Intelligence Agency (NGA). This gravitational model is complete to spherical harmonic degree and order 2159, and contains additional coefficients extending to degree 2190 and order 2159. In this paper, a grid of pre-computed geoid undulations at 1 x 1-minute resolution is used [7].

Angles are basic measurements in trigonometric networks. However it is necessary to measure a certain number of baselines between the points to define the network scale. These baselines are measured with high accuracy and as the network grows farther from them, the error accumulates. Therefore, the more distant new points are from the network baselines, the coordinates accuracy is lower. To include this distance as a predictor, distance grid from network baselines was calculated.

After generating three predictor grids, they were resampled to 200 x 200 m resolution, and pixels outside of country border were removed. Final predictors over territory of Serbia are shown in Figure 2.

Geostatistics is a branch of statistics focusing on spatial or spatiotemporal datasets. It incorporates the spatial (and in some cases temporal) coordinates of the data within the analyses, and aims at providing accurate estimates of phenomena at unsampled locations together with a quantification of the related uncertainty.

Kriging is the most important interpolation function in geostatistics - practically most of geostatistics is based and uses this method. It is often described with the abbreviation “BLUE” - best linear unbiased estimation. Main idea of kriging is to estimate function value in an unvisited location by calculating weighted average of values in nearby known points:

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

where $\hat{z}(s_0)$ is the predicted value of the target variable at unmeasured location s_0 with its given coordinates, and $z(s_i)$ are the values of target variable at known locations with their coordinates and respective weights (λ_i). The weights are chosen with the aim of minimizing the prediction error variance, resulting in weights depending on the spatial autocorrelation structure of the variable, it is based on the variogram model. This interpolation procedure is called ordinary kriging [2][1].

Regression models the relationship between the target variable and predictors (auxiliary environmental variables) at sample locations. This model is then used to make predictions at new locations, using the known value of predictors at those locations. This is an alternative to kriging. Linear multiple regression is commonly used in this approach [3][1], where predicted value is a weighted average of predictors:

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

where $q_k(s_0)$ are the values of the predictors at the target location, $\hat{\beta}_k$ are estimated regression coefficients, and p is the number of predictors (auxiliary variables).

Regression kriging combines these two approaches, using regression to fit the explanatory variation and simple kriging (with expected value 0) to fit the residuals – this is the fitting of unexplained variation [1]:

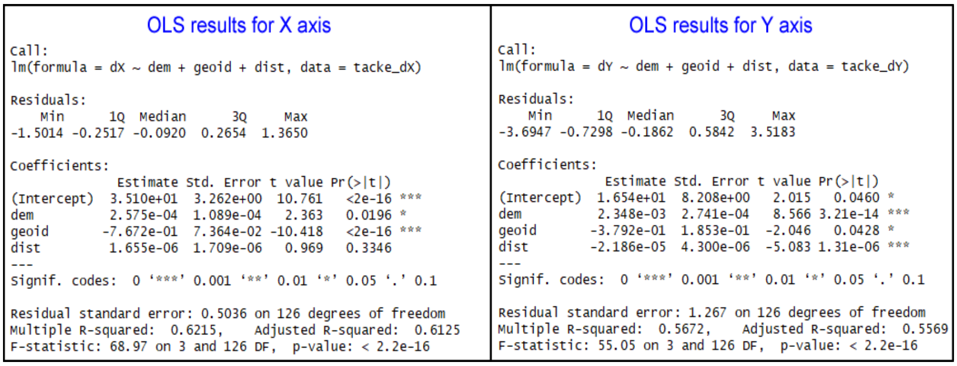
$$\hat{z}(s_0)=\hat{m}(s_0)+\hat{e}(s_0)=\sum_{k=0}^{p}\hat{\beta}_k\cdot q_k(s_0)+\sum_{i=1}^{n}\lambda_i\cdot e(s_i) \tag{3}$$

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, 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.

3. RESULTS

Firstly, linear regression of point coordinates errors and predictors was performed to assess the importance of predictors in modeling point errors. Results of simple linear regression for both input variables (X and Y coordinate errors) can be seen in Figure 2. Here global measures of quality - RMSE and coefficient of determination (adjusted R squared) can be seen, together with significance of each predictor in the model.

From the linear regession model, it is evident that height undulations (variable geoid from Figure 2.) is most significant predictor for X coordinate error, while DEM has little lower significance and distance from the baselines is unsignificant regarding to ANOVA t-test. Regression model for Y axis coordinate errors show high significance of DEM and baseline distance predictors, while height undulations have lower impact on prediction results. Because of these results, it was decided to include all 3 predictors in regression krigning model.



OLS results for X axis

```
Call:
lm(formula = dX ~ dem + geoid + dist, data = tacke_dX)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.5014 -0.2517 -0.0920  0.2654  1.3650 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  3.510e+01  3.262e+00  10.761   <2e-16 ***
dem          2.575e-04  1.089e-04   2.363   0.0196 *  
geoid       -7.672e-01  7.364e-02 -10.418   <2e-16 ***
dist         1.655e-06  1.709e-06   0.969   0.3346    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1

Residual standard error: 0.5036 on 126 degrees of freedom
Multiple R-squared:  0.6215,	Adjusted R-squared:  0.6125 
F-statistic: 68.97 on 3 and 126 DF,  p-value: < 2.2e-16
```

OLS results for Y axis

```
Call:
lm(formula = dY ~ dem + geoid + dist, data = tacke_dY)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.6947 -0.7298 -0.1862  0.5842  3.5183 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1.654e+01  8.208e+00   2.015   0.0460 *  
dem          2.348e-03  2.741e-04   8.566 3.21e-14 ***
geoid       -3.792e-01  1.853e-01  -2.046   0.0428 *  
dist        -2.186e-05  4.300e-06  -5.083 1.31e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1

Residual standard error: 1.267 on 126 degrees of freedom
Multiple R-squared:  0.5672,	Adjusted R-squared:  0.5569 
F-statistic: 55.05 on 3 and 126 DF,  p-value: < 2.2e-16
```

Figure 2. Results of OLS (Ordinary Least Squares) linear regression

After performing linear regession on the data, variograms were calculated for linear regression residuals. Results of variogram analysis and selected variogram models can be seen in Figure 3. Exponential model was chosen for both input datasets.

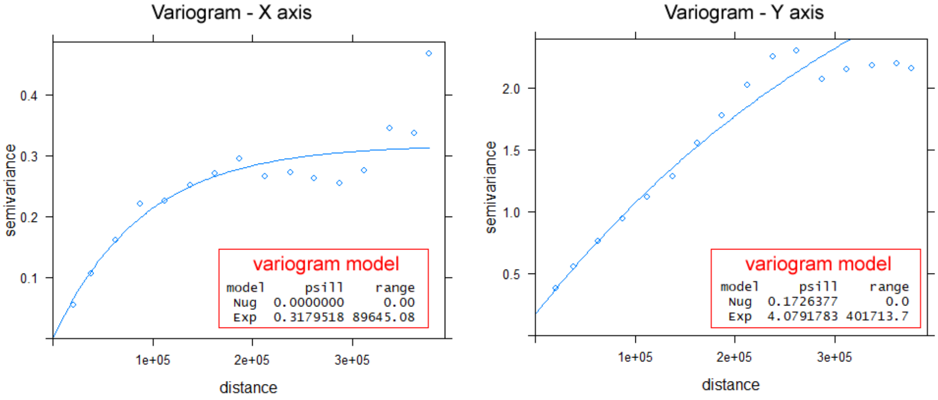


Figure 3. Variograms models for point error regression residuals in X and Y direction

Regression kriging was performed, separate prediction models were calculated for X and Y axis errors (Figure 4). Validation of regression kriging results was performed with cross validation technique. Leave-one-out procedure was used.

Global statistical measures of regression kriging quality are shown in Table 1. Plot of cross validation residuals can be seen in Figure 5., and histogram of residuals is shown in Figure 6. The results of cross validation show that residuals for X axis range from -1.225 m to 0.599 m for X coordinate errors and from -1.312 m to 2.621 m for Y coordinate errors.

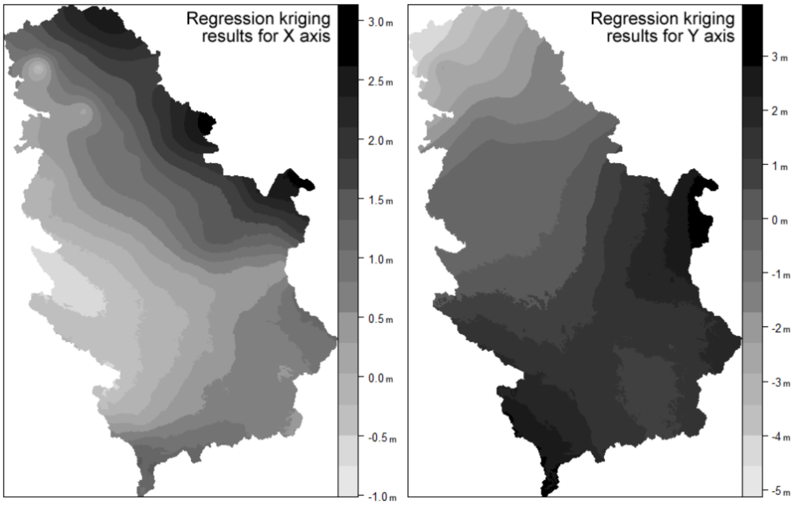


Figure 4. Regression kriging error prediction in X and Y directions

Table 1. Global statistical measures of regression kriging quality

	Regression kriging for X axis	*Regression kriging for Y axis*
RMSE	0.2343	0.3506
R^2	0.9154	0.9658

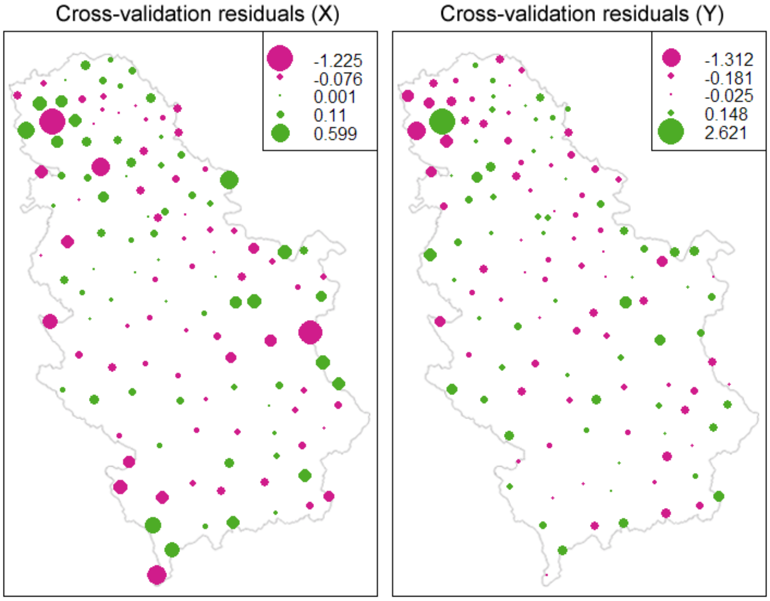


Figure 5. Regression kriging cross validation residuals in X direction(left) and in Y direction(right)

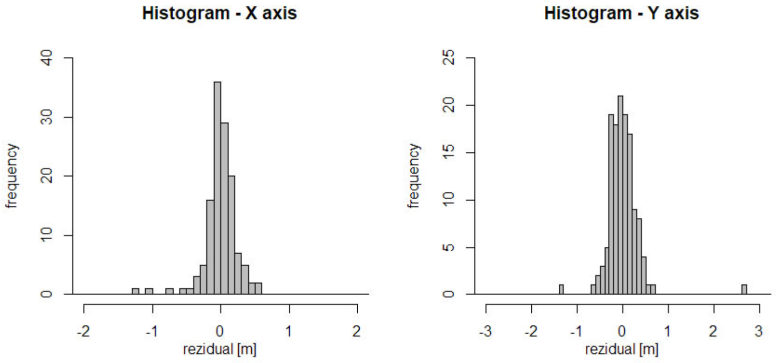


Figure 6. Histogram of cross-validation residuals

Finally, to determine the goodness of model for particular regions in Serbia, data from cross-validation was aggregated for each administrative region in Serbia. Regions used were downloaded from GADM [11] – level 1 administrative units were used. RMSE of cross-validation residuals were calculated for each region, using the points which fall in that region. This is a way to show how well the regression kriging model fits different parts of the country. Results are shown in Figure 7.

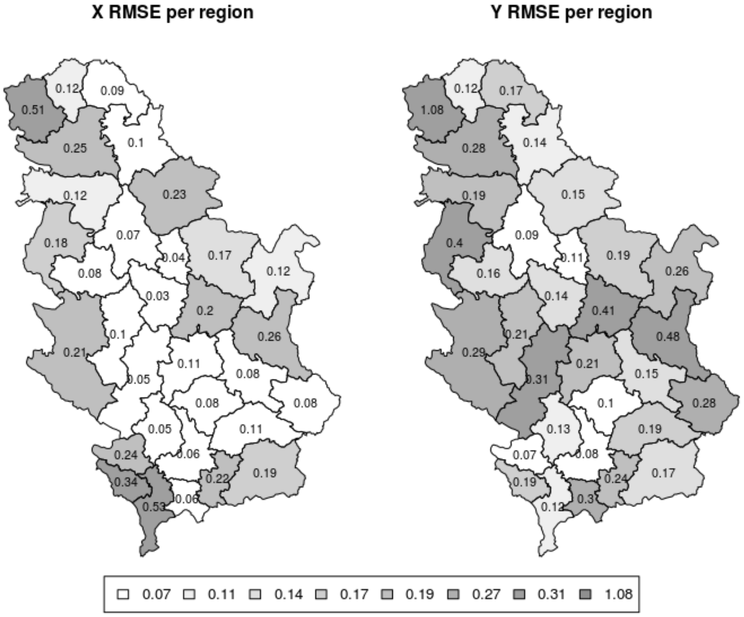


Figure 7. RMSE based on cross validation for regions in Serbia

4. CONCLUSIONS

The global statistical measures of model quality seen in Figure 2 (for linear regression model) and in Table 1 (for regression kriging model), show improvement in both RMSE and R^2 that can be seen for both X and Y point coordinate error models. RMSE for X direction errors improved from 0.50 m to 0.23 m, and for Y direction from 1.27 m to 0.35 m. R^2 improved to 0.92 (from 0.61) and to 0.97 (from 0.56), for X and Y coordinate error models respectively. Variogram model for Y direction errors has a nugget of 0.1726 m, which represents a minimal error that model makes with each prediction. This is not desirable, but it is probably a consequence of errors in basic dataset and couldn't be avoided.

Results of cross validation (Figure 5.) show that some points have high residuals, located near the north-west, south and east borders of the country for X coordinate errors, and mostly in the north-west part of the country for Y coordinate errors. These points are also clearly observed in the histograms in Figure 6. It is evident that quality of presented prediction model results varies for different regions of the country, as seen in Figure 7. Lowest quality predictions are in the north-west part of the country. Based on this map, users can see the fitness of model to a particular area of interest.

The reasons for variation in regression model prediction accuracy cannot be determined reliably. It is most likely influenced by two factors: inhomogeneity of network errors or/and lack of adequate predictors. Presented model could be further improved by including the information about the deflection of vertical. This information was not available to the authors at the time of writing this paper.

REFERENCES

[1] T. Hengl, Gerard B.M. Heuvelink, D. G. Rossiter: About regression-kriging: From equations to case studies. Computers & Geosciences, Volume 33, Issue 10, **2007**, Pages 1301-1315

[2] Webster, R., Oliver, M.: Geostatistics for Environmental Scientists Statistics in Practice. Wiley, Chichester, **2001.**, 271 pp.

[3] Draper, N., Smith, H. : Applied Regression Analysis, second ed. Wiley, New York, **1981.** , 709 pp.

[4] Rossiter D G: Tutorial - An example of statistical data analysis using the R environment for statistical computing, **2014**, http://www.css.cornell.edu/faculty/dgr2/teach/R/R_corregr.pdf

[5] Tomislav Malvić: Primjena geostatistike u analizi geoloških podataka, INA, SD Istraživanje i proizvodnja nafte i plina, Grupa za tiskarske i grafičke poslove, Zagreb, **2008.**

[6] T. Hengl: A Practical Guide to Geostatistical Mapping, **2009**, ISBN 978-90-9024981-0, http://spatial-analyst.net/book/system/files/Hengl_009_GEOSTATe2c1w.pdf

[7] QuickR- Help: http://www.statmethods.net/interface/help.html, visited 15.01.2018.

[8] https://land.copernicus.eu/pan-european/satellite-derived-products/eu-dem/eu-dem-v1.1?tab=metadata, downloaded 20.02.2018.

[9] http://www.trimble.com/globalTRL.asp?Nav=Collection-22914, downloaded 20.02.2018.
[10] Pebesma, Edzer J.: The Role of External Variables and GIS Databases in Geostatistical Analysis. **2016.**, Transactions in GIS. 10 (4): 615–632.
[11] http://biogeo.ucdavis.edu/data/gadm2.8/shp/SRB_adm_shp.zip , downloaded 20.02.2018.
[12] S. Delčev, J. Gučević, V. Ogrizović, M. Kuhar: First-order trigonometric network in the former Yugoslavia. Acta Geodaetica et Geophysica. **2015.** Volume 50. Issue 2. pp 219–241.
[13] Oleg Odalović: Fizička Geodezija. Građevinski fakultet Univerziteta u Beogradu. ISBN 978-86-7518-112-5. **2010.** pp 500

ПРЕДИКЦИЈА ПОЛОЖАЈНЕ ГРЕШКЕ ТАЧАКА ТРИГОНОМЕТРИЈСКЕ МРЕЖЕ ПРВОГ РЕДА

Резиме: *Приликом успостављања тригонометријске мреже првог реда у Србији, није узет у обзир утицај Земљиног гравитационог поља на мерења. Као последица, све тачке мреже имају грешку у положају, која је каснијим мерењима и рачунањима одређена посебно за X и Y координате сваке тачке. Коришћењем геостатистике (регресионог кригинга), овај рад се бави одређивањем предикционог модела грешке за било коју тачку на територији Србије.*

Кључне речи: *Геостатистика, деформација мреже, предикција грешке, тригонометријска мрежа, регресиони кригинг*