http://dx.doi.org/10.15576/GLL/2017.3.35

GLL

Geomatics, Landmanagement and Landscape No. 3 • 2017, 35–45

ANALYSIS OF THE DISTRIBUTION AND DENSITY OF MEASUREMENT POINTS IN TERMS OF TERRAIN MODELLING

Robert Gradka

Summary

For the implementation of direct measurements, proper understanding of the existing relationships and spatial variability, and at later stages, for obtaining reliable results of geostatistical analysis, adequate planning network measurement and correct placement of, and/or the evaluation of the number of measurement points in the measurement network are not the only necessary conditions. Another key prerequisite is choosing the right model for creating a DTM, which depends on the shape of the terrain. Correct spatial sampling should provide much information on the spatial distribution of the studied variable in an area, at minimal cost and with minimal effort. Faithful reproduction of the land surface that reflects any of the characteristics of the environment is not possible through DTM, due to a number of restrictions, manifesting themselves in the form and size of the data set; due to time and economic constraints; and also because the full complexity of the terrain's surface cannot be measured or expressed. The present work undertakes to analyse the density and distribution of measuring points on four areas that have specific characteristics in common, yet they remain different in terms of surfaces, height differences, as well as their complexity. After selecting the research areas, these were designed and laid out in a grid with the shape of rectangles that were similar in structure to the GRID model. The data were analysed using geostatistical interpolation by ordinary kriging, in order to conduct a proper analysis of the distribution and density of the measuring points, to calculate the surface properties of a particular point, and in order to attempt to reduce the workload and cost factor.

Keywords

DTM • distribution • density • GRID • ordinary kriging

1. Introduction

In surveyor's work, the following are of utmost importance: the time required for data acquisition, processing and compilation, along with high accuracy of that data. The ratio of speed of data acquisition to the accuracy varies immensely depending on the purpose of the geodetic work, the equipment used, the size of the area, land relief, etc. Sometimes there is a need to make a model of the terrain over a wide area in a short time. Then questions should be asked as to how quickly and accurately it is possible to do so, without compromising on the accuracy. If we either limit or increase the number of points, is this going to produce the same exact results as the full and accurate coverage of all the measurement points? Or perhaps it is their proper placement, rather than their density, that is the key to achieving satisfactory results from the study? Is there a way to optimize this process at all? Is the GRID model applicable to large areas? The author will try to answer these questions in the present paper, while giving examples of several research areas with specific characteristics.

The aim of the study is geostatistical analysis of spatial data, with particular emphasis upon its impact on land surface modelling, while taking into account the distribution and density of measuring points in areas of different surfaces and reliefs. Analysis of terrain modelling, GRID (also relative to each other), as well as location and density of points, was performed on the basis of data obtained from direct measurements. At the time of measurement, grid points were stabilized permanently, in order to enable an overlaying and a comparison between the various models. Data interpolation was done using the ordinary kriging method. Comparisons of GRID models were carried out by way of overlapping the maps in the Surfer software. Due to the large amount of data and the number of resultant maps in this study, only one of the research areas was given as an example, however, the overall findings apply to all four analysed areas.

Due to the variety of research areas, the accuracy of the measurements, and the detailed elaboration of the data obtained, the work can serve as a model for deploying and applying the appropriate density of measurement points in large areas with the view to reducing cost and workload. The work focuses on surveys, where it is possible or even appropriate to apply typical geodetic measurement techniques, while it is not feasible to use techniques such as Unmanned Aerial Vehicle (UAV) or Airborne Laser Scanning (ALS), due to the cost of the equipment itself, the time of data processing, and the accuracy of the data.

2. Characteristics of research areas

Based on the considerations pertaining to the selection of research criteria, it was decided that we should focus on the altimetric survey of four areas, of which two are covered with low vegetation; third is an area without vegetation (concrete, asphalt); and fourth is a mixed area, consisting of low vegetation and hardened surfaces (asphalt). All four research areas are located within the city of Kraków, and they were measured using one of the geodetic direct measurement methods – i.e. technical levelling. Common features of testing areas are as follows:

- the areas are located away from high buildings, power grid, etc.,
- in the immediate vicinity and directly on the studied area, there is lack of high vegetation (that would make it impossible to perform regular measurements);
- detailed elements of the basic map (e.g. wells, curbs, sewage grates) are not included.

2.1. Geometry of selected areas

The analysis was preceded by the design and placement of rectangles in each area of the study, which were similar in structure to the DTM GRID model [Flotron 2000, Aguilar 2005, Gotlib 2006, Gościewski 2005, 2007; Hejmanowska 2007, Suchocki 2013]. Creating a grid of points in the form of a simple geometric figure makes it possible to determine the height of the field points, the vertices of those figures, and additional points within the mesh of the grid that describe the relief of measured area. Forms of rectangles and squares are designed for accurate measurement of relief in open and horizontal areas (as in this case).

The basic figure (rectangle) covering the entire measured area was connected to a horizontal and vertical geodetic control network. The rectangle was designed in such a way that its dimensions contained the total number of filling figures (mesh) and that the vertices of these figures could be determined by means of crossings and line measurements, without the use of theodolite. The vertices of the basic figures were marked out in the areas using a precision reflector and an electronic total station with accuracy up to 15cc and a length of 1mm. The grid points, all indirect points on the perimeter and throughout the basic figure, were also marked out at right angles and with accuracy of ± 10mm. For the duration of the measurement, intermediate points were marked in the field with metal pins or nails (in the case of concrete and/or asphalt).

After the initial data processing, that is, after the verification of the data, elimination of measurement errors was carried out, followed by quantitative and qualitative data analysis. Table 1 below presents a comparison between the major elements of the data sets collected in each of the research areas.

Table 1. Summary of key data elements

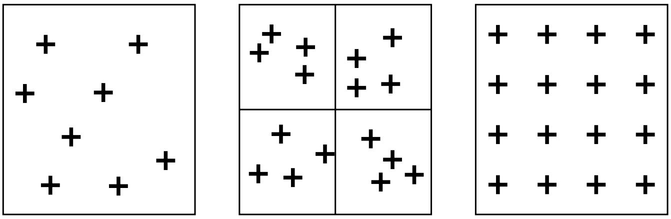
	Area 1	**Area 2**	**Area 3**	**Area 4**
Number of measured points [pcs]	121	121	170	121/234*
Surface size [ha]	0.602	0.598	0.753	0.570
Average number of points per ha	201	202	226	212/411*
Maximum height [m]	222.559	220.718	219.115	202.934
Minimum height [m]	218.036	217.867	218.072	202.396
Average height [m]	219.982	218.905	218.566	202.552
Maximum difference in elevation [m]	4.523	2.851	1.042	0.538

* 113 asphalt track points were measured in addition

Source: author’s study

3. Methodology of elaboration

The key factors influencing the “quality” of the level drawing include the amount and distribution of measurement points in the examined area. In classic test methods, the choice of a two-dimensional measurement network, also called a measurement plan, plays the crucial role (Figure 1). The most popular network types are [Zawadzki 2011]:



Source: author's study

Fig. 1. Classic types of measurement networks (from left to right): simple random, layered random, systematic

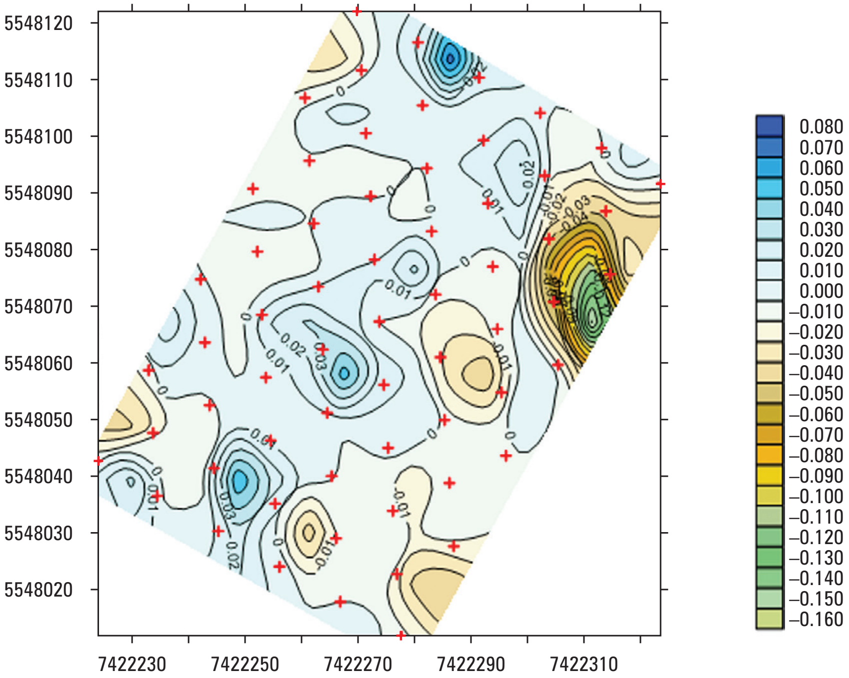
In my work, I first used a simple random method [Hansen 1953] in which the measurement points are chosen independently of each other, in a completely random way, followed by the second, systematic, deterministic method. The second method is based on the premise that the first point is not randomly selected (contrary to the systematic random method), but it is chosen according to certain predetermined rules. Systematic plans for testing the most commonly used networks include rectangular or triangular networks.

The data obtained from one of the basic measuring techniques – that is, the technical levelling – was chosen as the standard method of measuring altitude. According to the author, this produces the most accurate mapping of relief. Density and location analysis relied on point selection in a regular or random way.

In the case of a regular type, the points forming the mesh of the basic figure will be used. This type will be considered for reducing (diluting) the number of points on the basis of "every second" (50%) and "every fourth" (25%) point from the list of all points in a given test area. The random sampling type will also include points forming and filling the basic figure, but these will be randomly selected from about 25% and about 50% of all the measurement points.

In accordance with the above rules, filtered points will create contour maps that will be superimposed and compared to maps from the full set of data collected during the measurement. These are called *result maps* [Watson 1992, Sauer 2006]. The models of the result maps show height differences, expressed in meters, resulting from the overlapping of these maps. Each map was created by means of interpolation of the *ordinary kriging* type [Royle 1981; Oliver 1990; Wackernagel 2003; Li, Leap 2008], and the interpolation lines will pass only around the points of the grid, filling the basic figure. Kriging interpolation is very often used in environmental research, producing better results of spatial interpolation than deterministic methods [Krige 1951; Matheron 1962; Franke 1982; Lam 1983; Mulugeta 1999; Childs 2012; Mund 2013].

On the result maps (Figures 2–5) the following information is shown: differences in altitude [m], designated reference area, measured and selected for analysis of complementary point figures, marked as red crosses. In addition, the points of the basic figure – vertices (max: 4) – are added so that each area overlaps and has the same limited size. The maps below (Figures 2–5) refer to the same research area, showing all dependencies and discrepancies. Each map was made using the Surfer software.

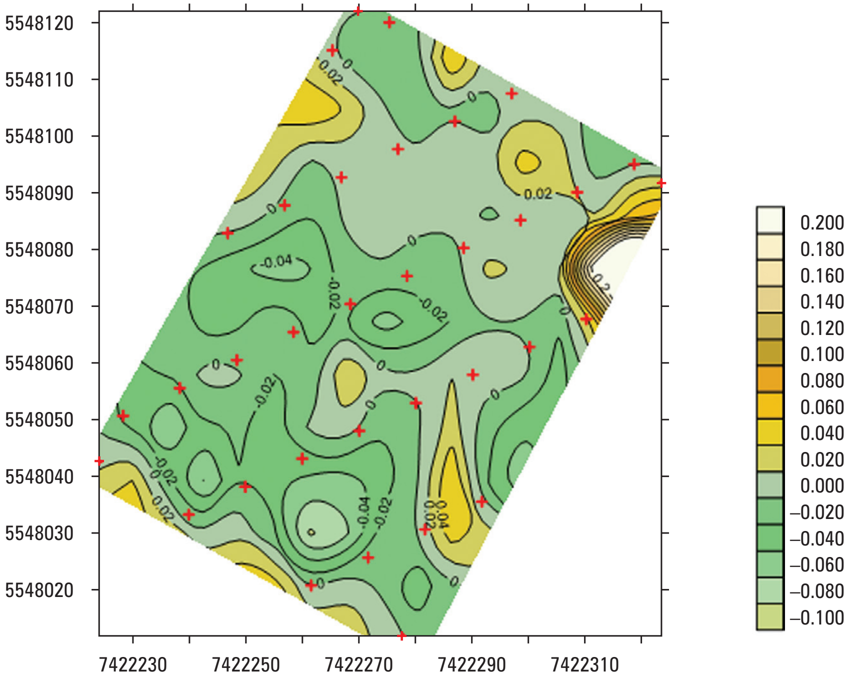


Source: author's study

Fig. 2. The result map – "every second" point

The above result map (Figure 2) shows the regular sampling type and the points selected on the "every second" point basis. It can be read that the values of the difference in actual heights range between –160 mm and +80 mm, therefore the range of differences is 240 mm. For the most part of the studied object, these differences do not exceed ± 20 mm. There are cases of "contour mesh" occurring here: two smaller ones (up to +50 mm) in the central and southern sections, and two larger ones, one of which is located on the northern edge of the area (+80 mm). What is characteristic for this area is a small difference in height of the terrain (slightly over 0.5 m), which was observed as a hill in the north-east corner of the studied area. The effect (lack of points on the hill, which had been omitted during the analysis) is a visible lowering of the area in this part, with the value of as much as –160 mm.

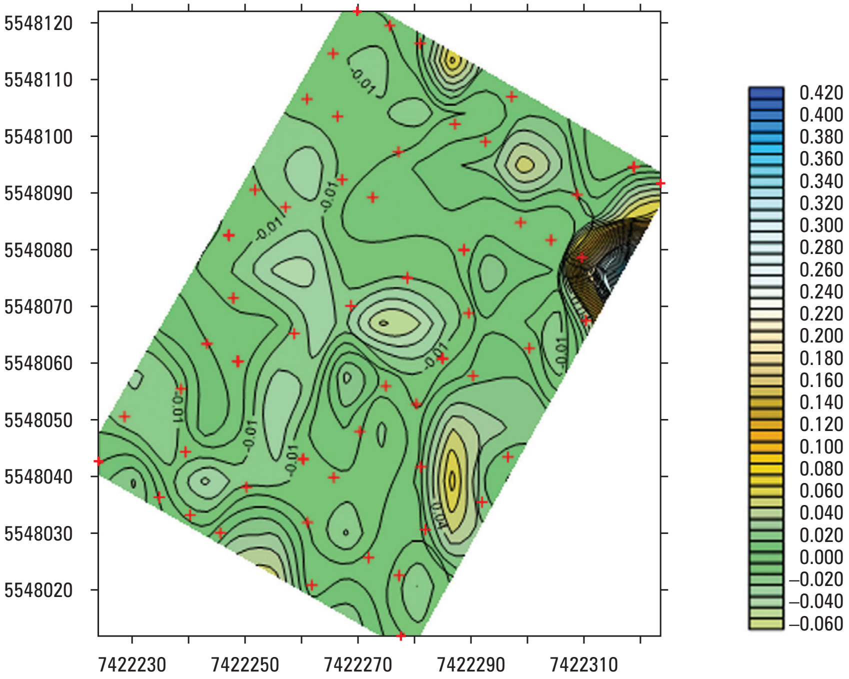
On the result map (Figure 3), the distribution of altitude differences is almost uniform across the entire area. Differences are in the range between –40 mm and +40 mm, except for the part of the area where the hill is located (north-east corner of the studied area), and where the differences are more than +400 mm (colour scale limited to +200 mm). The absence of a characteristic point (accepted for analysis) on an existing hill causes such divergences. In the northern part of the studied area, the differences are up to +40 mm, and they are predominantly between 0 mm and + 20 mm. The middle part remains in the range between –20 mm and 0 mm, while in the south, the differences are larger, and they reach –40 mm. There is a single case of a "contour eye" in the south-east section, which has a value of +40 mm.



Source: author's study

Fig. 3. The result map – "every fourth" point

The result map (Figure 4) presents a very uniform distribution of height differences, but there are divergences between models, ranging between –20 mm and +300 mm. Throughout the test area, the difference values range between –20 mm and +10 mm, apart from the fragment of an object located in the north-eastern part, whose value amounts to +300 mm. This is due to the existence of a small elevation on which the measurement point is not included in the analysis of the random selection of points, and which is consequently manifested in the form of a "contour eye" of such substantial value.



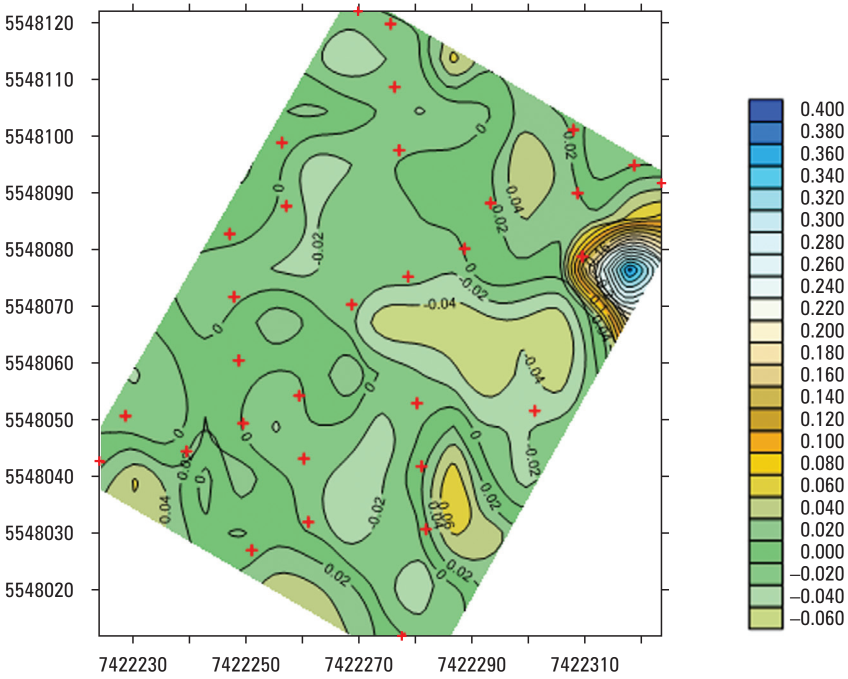
Source: author's study

Fig. 4. The result map – points selected at random (50%)

From the result map (Figure 5), it can be seen that the altitude difference values in almost the entire test area remain within the actual range between –40 mm and +40 mm. In the northern part of the object, discrepancies between the models amounted to between 0 and 20 mm. The central part presents a difference between –40 mm and 0 mm, while in the southern portion, it is between 0 and +40 mm. Three cases of "contour eyes" were determined on the map. Their values (+40 mm, +60 mm, +400 mm) are associated with the absence or small amount of measured points located near each other. In the north-east corner, there is a "contour eye", with a value of up to +400 mm, which is caused by the hill existing in the area.

The table below (Table 2) shows the results of the analysis of the accuracy and dependence obtained from field measurements. This table shows that both the measurement and the accuracy of the terrain (RMSE) are high.

Based on the results (presented in Table 2), which *inter alia* present the size of the average square error of the studied model (RMSE), it was not possible to distinguish the magnitude of this error depending, for example, on the inclination angle of the terrain. In flat and hilly areas, the magnitudes of these errors do not differ significantly enough for such distinction to be made. An important element of the study was to determine the magnitude of the systematic error of the model. For this purpose, the average values of all differences for the entire area under control were additionally calculated. The errors of field measurement could have contributed to the systematic error.



Source: author's study

Fig. 5. The result map – points selected at random (25%)

Table 2. Results of the analysis of accuracy and dependence obtained from field measurements

	Area 1	Area 2	Area 3	Area 4
Systematic error (average)[m]	–0.006	0.011	–0.001	–0.003
RMSE [m]	±0.012	±0.020	±0.005	±0.008
Standard deviation [σ]	±0.010	±0.017	±0.005	±0.008
Correlation coefficient [r]	0.9995	0.9993	0.9948	0.9937
Determination factor R2	0.9999	0.9993	0.9996	0.9985

Source: author's study

Very high correlation and determination coefficients (Table 2), obtained in the correlation analysis of geodetic measurements, confirm the accuracy of the terrain model obtained by the technical levelling method.

4. Conclusions

Regular sampling (Figures 2–3) can be performed in the form of profiles, or in a grid of rectangles or squares. The advantage of this sampling is the ability to fully automate the acquisition of altitude values. The disadvantage is its limitation to areas with only slight height changes. The number of points obtained is inadequate to the area, because in flat areas it is too large, but in undulating areas, it is too small. In some cases – for instance in areas with small and medium altitude differences – regular measurement would have to be complemented by some characteristic points, in order to avoid substantial errors. In areas with large altitude differences, the height differences are relatively high, and the measurement of the land features will not be sufficient to produce an altitude model that would be exact and precise.

When analysing the distribution and density of points for random sampling (Figures 4–5), large differences in elevation were observed in areas of large area density and small size. Divergences increase as the number of points decreases, and their concentration in one place improves the accuracy of the model only in that particular section of space, around the points. On the remaining part, this causes deterioration in accuracy of the model, manifested in broad and large values of height differences – the "contour eyes". In the case of random sampling of data, it is impossible to achieve high accuracy when operating on a small number of points.

If the characteristic points of the terrain are not included in the analysis, it is not possible to clearly and precisely define the differences in heights. In areas with lower altitude differences, these discrepancies are smaller, but as the result maps show, some of the most distinctive field points, those that deviate from the object the most, should be measured (except for random points). The interpolation model of the measured points obtained in this way, whether regular or random, will be more accurate. Dense testing of the examined area is more accurate, but it is burdened with very high costs, while testing too sparsely causes inaccurate analyses due to lack of spatial correlation information. The density of data sampling should always depend on the changes in terrain.

Excessive concentration of points, or linear structure of measurement points, both have negative impact on the results of interpolation. The best way would be to distribute the points evenly, and for geostatistical methods it is not necessary to use systematic measurement networks or to determine the exact distance between points. This distance may sufficiently be given as the approximate distance between the measuring points (as it has been done in the present work). Differences between distances were ± 5cm. By using geostatistical methods, the distribution of measurement points should not be too chaotic; they should be confined to regular areas, bounded by a rectangle or an ellipse. In the case of an irregular area, we may expect major irregularities in the models as well as significant height differences, following from the occurrence of the "boundary effects" that may be the result of insufficient number of points, measurement pairs in certain areas, or in certain directions. It is assumed that even in the case of a small area, the number of measurement points in a two-dimensional space should not be less than several dozen, but this obviously depends on the purpose and accuracy of the given survey, on the size and shape of the area, and on the spatial continuity of the phenomenon being tested. The lesser the continuity, the more measuring points are needed.

The accuracy of the resulting DTM is not only influenced by the degree of compaction of the measuring points and their accuracy, but also by the way that the DTM itself is modelled. Reliable mapping of natural terrain by using available numerical methods is one of the major research problems associated with DTM modelling. Because of the structure and the ease of data archiving, regular GRID models are most commonly used in spatial information systems and geodetic practices. In addition, these models are a great tool to illustrate the relief in large areas, and are easily subjected to statistical and spatial analyses.

References

Aguilar F.J., Agüera F., Aguilar Manuel A., Carvajal F. 2005. Effects of Terrain Morphology, Sampling Density, and Interpolation Methods on Grid DEM Accuracy. Photogram. Engin. Remote Sens., 71(7), 805–816.

Childs C. (ed.). 2012. Interpolating surfaces in ArcGIS Spatial Analyst, ESRI Education Services.

Flotron A., Koelbl O. 2000. Precision terrain models for civil engineering. OEEPE Offical Publication, 38, 37–134.

Franke R. 1982. Scattered Data Interpolation: Test of Some Methods. Mathemat. Comput., 33(157), 181–200.

Gościewski D. (ed.). 2005. Influence of measurement points location on selection of interpolation algorithm, The 6th International Conference Environmental Engineering, Gediminas Technical University Press, Vilnius-Lithunia.

Gościewski D. 2007. Analiza dokładności interpolacyjnych modeli powierzchni typu GRID. Materiały 20. Jesiennej Szkoły Geodezji, Wrocław.

Gotlib D., Olszewski R. 2006. Co z trzecim wymiarem? O modelowaniu rzeźby terenu w referencyjnych bazach danych, Geodeta, 4, 31–34.

Hansen M.H., Hurwitz W.N., Madow G. 1953. Sample Survey Methods and Theory, vol. 1: Methods and Applications, John Wiley & Sons Inc., New York.

Hejmanowska B. 2007. Analiza NTM w postaci GRID i TIN na przykładzie danych z OKI. Arch. Fotogram. Kartogr. Teledet., 17a, 281–289.

Krige Danie G. 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. J. Chem., Metal. Mining Soc. South Africa, 52 (6), 119–139.

Lam N.S. 1983. Spatial interpolation methods: A review, Am. Cartogr., 10(2), 129–149.

Li J., Heap A.D. 2008. A review of Spatial Interpolation Methods for Environmental Scientists, no. 23.

Matheron G. 1962. Traité de géostatistique appliquée. Technip, Paris.

Mulugeta G. 1999. The elusive nature of expertise in spatial interpolation. Cart. Geog. Inf. Sys., 25(1), 33–41.

Mund J.-P. 2013. Geospatial statistics and spatial data interpolation methods, GIS'EM 2013 at Eberswalde, Eberswalde.

Oliver M.A. 1990. Kriging: A Method of Interpolation for Geographical Information Systems. Int. J. Geogr. Inform. Syst., 4, 313–332.

Royle A.G., Clausen F.L., Frederiksen P. 1981. Practical Universal Kriging and Automatic Contouring. Geoprocessing, 1, 377–394.

Sauer T. 2006. Numerical Analysis. Pearson Education Inc., Boston.

Suchocki Cz., Damięcka-Suchocka M., Błoch P., Stec M. 2013. An evaluation of digital terrain model accuracy using direct survey data. Acta Sci. Pol., Geod. Descr. Terr., 12(3), 17–26.

Wackernagel H. 2003. Multivariate geostatistics. An introduction with applications. Third completely revised edition, Springer-Verlag, Heidelberg.

Watson D.F. 1992. Contouring: a guide to the analysis and display of spatial data. Oxford, Pergamon.

Zawadzki J. 2011. Metody geostatystyczne dla kierunków przyrodniczych i technicznych. Oficyna Wyd. Politechniki Warszawskiej, Warszawa.

Dr inż. Robert Gradka
Politechnika Wrocławska
Wydział Geoinżynierii, Górnictwa i Geologii
50-421 Wrocław, ul. Na Grobli 15
e-mail: robert.gradka@pwr.edu.pl