

INFLUENCE OF MEASUREMENT POINTS POSITIONING ON ACCURACY OF INTERPOLATION SURFACES

Dariusz Gosciewski
University of Warmia and Mazury in Olsztyn
Faculty of Geodesy and Land Management
Institute of Geodesy, Poland
e-mail: chillis@uwm.edu.pl

Abstract: Contemporary land information systems allow generating models of numeric surfaces by applying a number of interpolation algorithms. Appropriate selection of measurement points included in calculations is of key importance among the factors influencing the quality and accuracy of interpolation. Appropriate selection of those points allows not only improvement of the accuracy of generated surface models but also shortening the computation process. The paper analyzes the influence of location and density of measurement points on accuracy of the created interpolation surfaces in the aspect of morphological differentiation of the generated models.

1. Introduction

Spatial information systems (SIT, GIS) are frequently used for processing, analyses and maintenance of large volumes of information originating from a variety of sources [8,9,12]. Generally, data originating from the digital terrain model (DTM) form the base for spatial organization in such systems. Source data forming the DTM is obtained most frequently from direct measurements or indirectly from digitalization of analogue materials [8, 10]. Contemporary measurement systems such as aerial laser scanning or multiple beam sonic sounders allow obtaining large volumes of information within a relatively short time [4,6]. With a large volume of measurement data available, it is possible to generate a very accurate terrain model. Direct use of such sets, however, is difficult because of the volume and non-standardized nature of information recording [5,7]. The pace of current development and the particular role of information require, next to quality and reliability of data, the possibility of dynamic data processing and conduct of comprehensive comparative analyses for the same elements during different measurement periods. As a consequence, attempts are made to organize the spatial structure and limit the number of points generating the digital terrain model maintaining at the same time the required accuracy. Presenting the information describing the surface in the form of a regular GRID type nodes structure offers one of the methods of organizing that information [4,5,7]. That allows reducing the number of data, limiting redundancy and significantly accelerating information processing. A particular role in generating the grid of squares is played by appropriate selection of interpolation algorithms' parameters. Thanks to those algorithms the values at individual nodes are interpolated on the basis of measurement points. Correct selection of measurement points participating in the calculations is of key importance for precision of determining the values at the nodes.

2. Test model

A specially developed theoretical test model was used for the purpose of analyzing the results of interpolation carried out using different algorithm at different densities of measurement points. The function of two variables (1) allowing generation of mathematical surface presented in figure 1 was used for development of the model. On the basis of that function, within the specified range of coordinates x, y (1), pseudo-measurement points (further referred to as measurement points) have been generated at random.

$$f(x, y) = 2(\text{SIN}(e^x y) + (y^2 - x^2)) \quad (1)$$

$$x \in \langle -1; 4 \rangle, \quad y \in \langle -2; 5 \rangle$$

They formed a standard surface, which was then rescaled and shifted in a way that allowed obtaining a DTM model of 500m by 700m and the area of 350000m², containing exclusively measurement points with positive values of x, y and z coordinates (fig. 1B). The measurement points generated in that way served calculation of values in nodal points using a variety of interpolation algorithms. The base square of the grid was determined in such a way that the morphology of the DTM model generated on the basis of function (1) should not allow precise interpolation of values in all nodes. The spacing S between nodes was 10m for each tested case (fig. 1C).

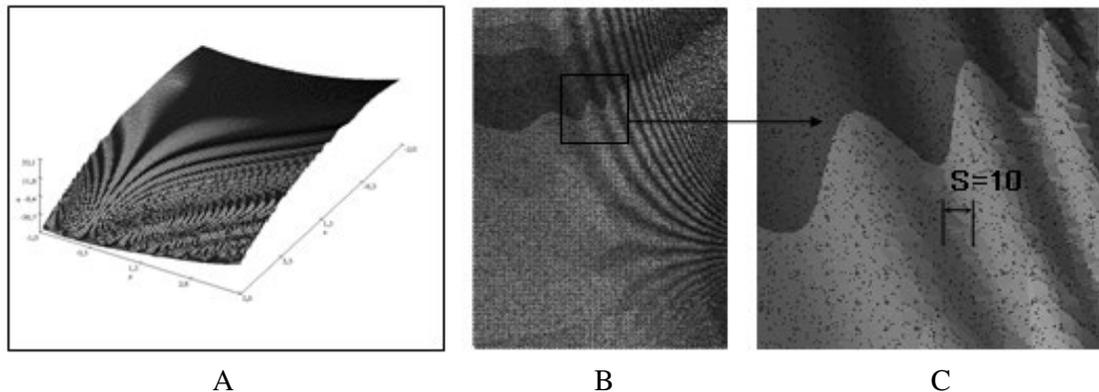


Figure 1: Structure of the test model

Next, on the basis of the same function (1) the theoretical nodular points were generated with the same positioning parameters as the interpolated nodes. On the basis of the test model developed in that way analyses were carried out that allowed comparing the practical value calculated by interpolation with the theoretical value calculated using the function for all nodular points of the GRID structure.

3. Assumptions for the study

Four interpolation algorithms were used for the analyses. The first algorithm – $1/R^2$ – assumed calculation of the value at the node point on the basis of the sum of the value of n measurement points balanced by the reverse square of the distance from the node [2]. The second algorithm – kriging – assumed calculation of the value in the node on the basis of semi-variogram adjusted to the height distribution of measurement points around the node [1]. The third algorithm calculated the value at the nodal point using the function of first level local polynomial [3]. The fourth algorithm for interpolation of the value in the nodal point used multiquadric radial functions [11].

Each of the algorithms used the same base of measurement points randomly distributed on the function surface model for calculations.

Aiming at investigating the influence of measurement points' position around the node on the calculations results, seven cases of their positioning were analyzed (fig.2). Only in the first case presented in figure 2 measurement points were chosen from the entire space around the node without division into sectors.

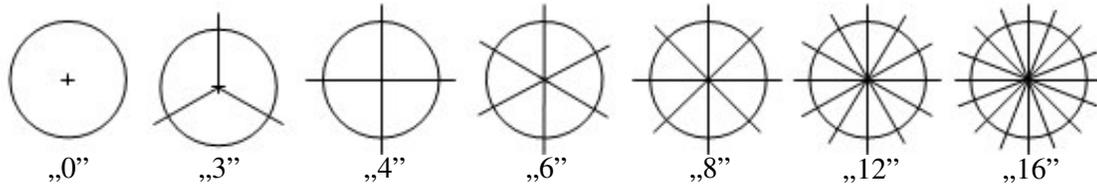


Figure 2: Division of the area around the node into search sectors

In the other cases the search space was divided into equal sectors. The angle of lines creating the first sector was fixed at 0°. The search radius around individual nodes was not limited and was the same in all directions.

Aiming at analyzing the influence of different densities of measurement points on the interpolation results, four surface models were generated consisting of 1000 (fig. 3A), 5000 (fig. 3B), 20000 (fig. 3C) and 50000 (fig. 3D) points respectively. At the top of DTM models presented in figure 3 the densities of measurement points per grid base square with side length of 10m are given.

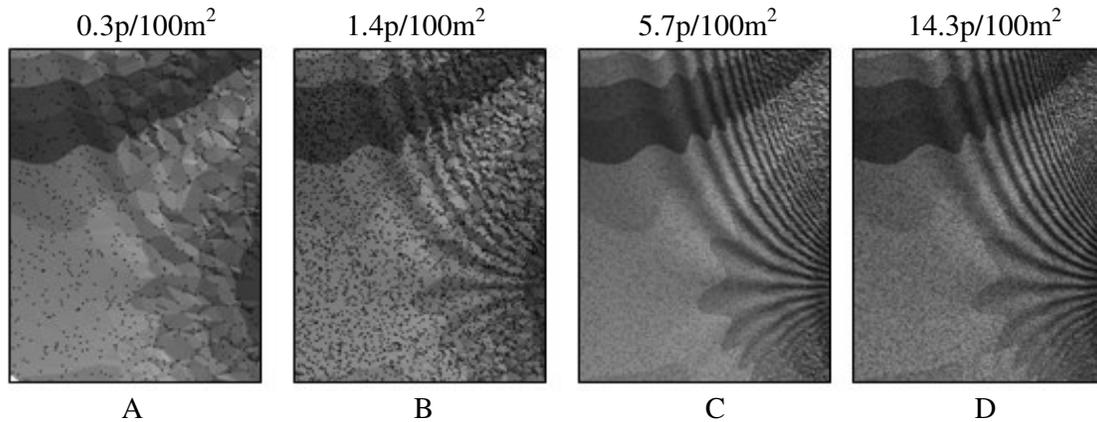


Figure 3: DTM model with different densities of measurement points

The *RMS* (2) coefficient was applied for comparison of the interpolated surface and the standard surface matching accuracy.

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(x_i, y_i) - z_i)^2} \quad (2)$$

where:

- $f(x,y)$ – value of function (1) in the theoretical nodal point with coordinates $x y$,
- z – value calculated using a given interpolation algorithm on the basis of measurement points at the nodal point with coordinates $x y$,
- n – number of nodal points.

The differential diagram was used to present the areas of deformation in the analyzed surface models. The diagrams presented the absolute differences in values calculated in all nodes between the practical interpolation model – DTM „P” (fig. 4A) and the theoretical function model – DTM „T” (fig. 4B). Figure 4C presents the differential diagram representing the ranges of deformations determined on the basis of the values calculated for each node. Above the diagram the *RMS* coefficient in [m] calculated jointly for the entire differential surface is shown. The legend presented under the diagram shows the assumed ranges of differential values in [m].

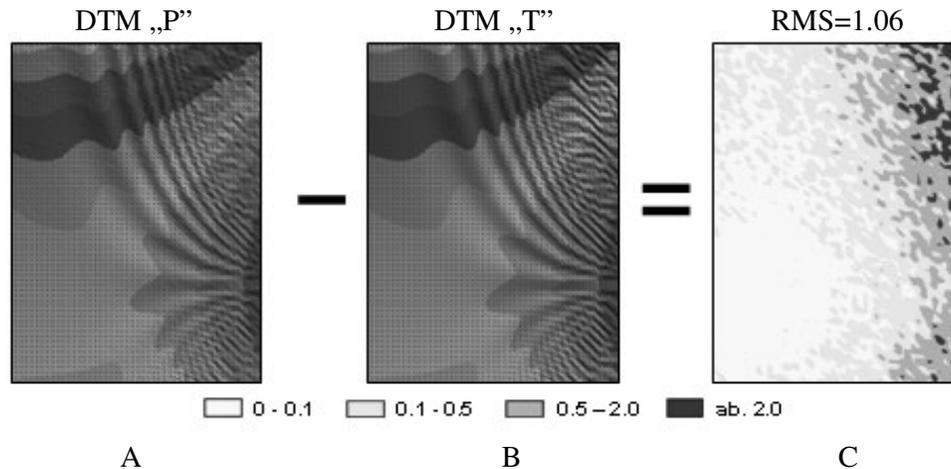


Figure 4: Generation of the differential diagram

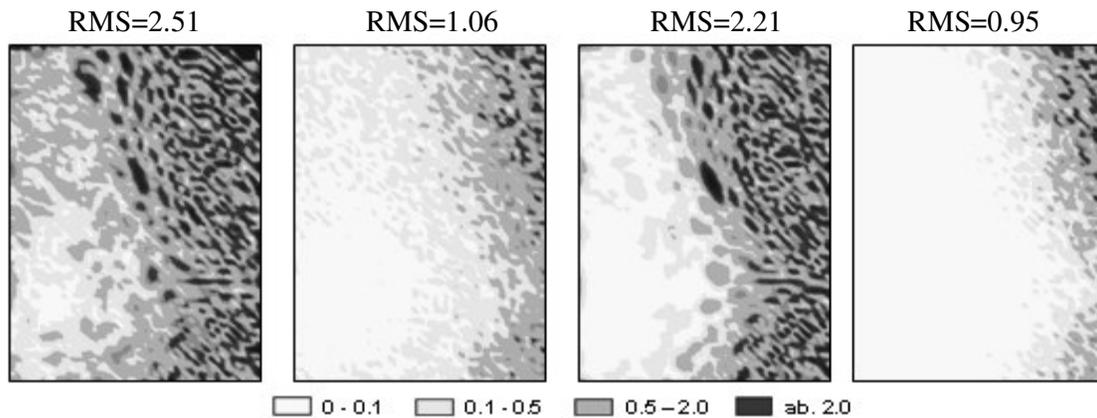
By applying the differential models it is possible to locate deformations in interpolation models as compared to the theoretical model. Application of the *RMS* coefficient allows, on the other hand, determining the accuracy of matching the entire surface by means of a single value. That in turn offers the possibility of comparing different interpolation surfaces by showing on the graphs the dependences between location of measurement points and the value of *RMS* coefficient. That coefficient also offers the possibility of comparing the accuracy of different algorithms.

Further analysis was divided into two steps. In the first one the situation where all nodal points of the test model were taken into consideration was analyzed. In the second step the internal points of the object selected using the rectangle were analyzed.

4. Analysis of DTM models accuracy for all nodal points

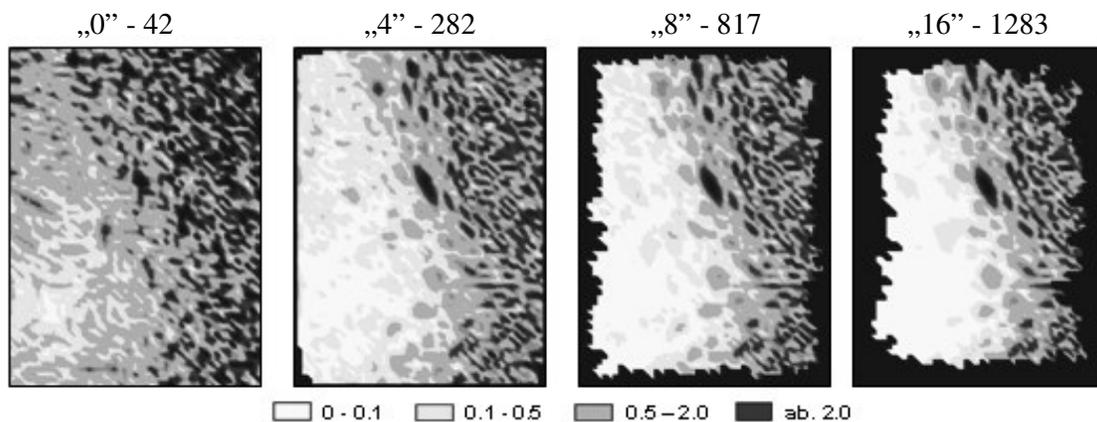
Using the differential diagrams, the influence of measurement points' density on DTM model accuracy was analyzed. The diagrams presented in figure 5 show location of deformations at different measurement points' densities for two interpolation algorithms. In all cases 16 measurement points selected from the entire space around the node without division into sectors were used. Interpolation using $1/R^2$ algorithm at the density of $0.3p/100m^2$ (fig. 5A) caused relatively large deformations across the entire surface ($RMS=2.51m$). Increasing the measurement points' density to $14.3p/100m^2$ (fig. 5B) significantly improved the matching of the model ($RMS=1.06$). A similar situation occurred when kriging was used for interpolation. At measurement points' density of $0.3p/100m^2$ (fig. 5C) interpolation offered clearly worse results than in case of the density of $14.3p/100m^2$ (fig. 5D). Comparing the diagrams and *RMS* coefficient for both algorithms a clear advantage of the algorithm using kriging over the $1/R^2$ algorithm can be shown in both lower and higher density of measurement points. As in the discussed cases selection of points occurred without division into sectors, the influence of

edge nodes on accuracy of the interpolation models was not observed. Similar dependences were obtained for the remaining algorithms.



A B C D
Figure 5: Differential diagrams for different measurement points densities

The diagrams shown in figure 6 allow analyzing the influence of the number of search sectors on DTM model accuracy. The further examples present deformations after interpolation with the algorithm using the radial functions. The density of measurement points in all cases was $0.3p/100m^2$. In all cases 16 measurement points were used located in the following way: without sectors (fig. 6A), in 4 sectors (fig. 6B), in 8 sectors (fig. 6C) and in 16 sectors (fig. 6D).



A B C D
Figure 6: Differential diagrams for different numbers of search sectors

Above the diagrams, next to the number of sectors, the number of nodes rejected as a consequence of inaccuracies in calculations caused by insufficient number of measurement points in individual sectors is given. In consecutive diagrams an increasing share of rejected edge nodes with increasing number of sectors can be seen. On the other hand, in the area situated inside the diagrams, the accuracy increases with the increase in the number of sectors. A similar situation occurred during interpolation using the other algorithms.

The graphs presented in figure 7 allow analyzing the dependence of *RMS* coefficient on the number of sectors for different algorithms at different densities of measurement points. On the vertical axes of graphs the values of *RMS* coefficient are represented while on the horizontal

axes of category the number of sectors is presented. The density of measurement points is represented by series lines. Consecutive graphs were generated for different interpolation algorithms: $1/R^2$ (fig. 7A), local polynomial (fig. 7B), kriging (fig. 7C) and radial functions (fig. 7D). Graphs 7A and 7B present the situation where there is one measurement point per sector while in graphs 7C and 7D the number of measurement points per sector depends on the number of sectors and in all cases equals 16.

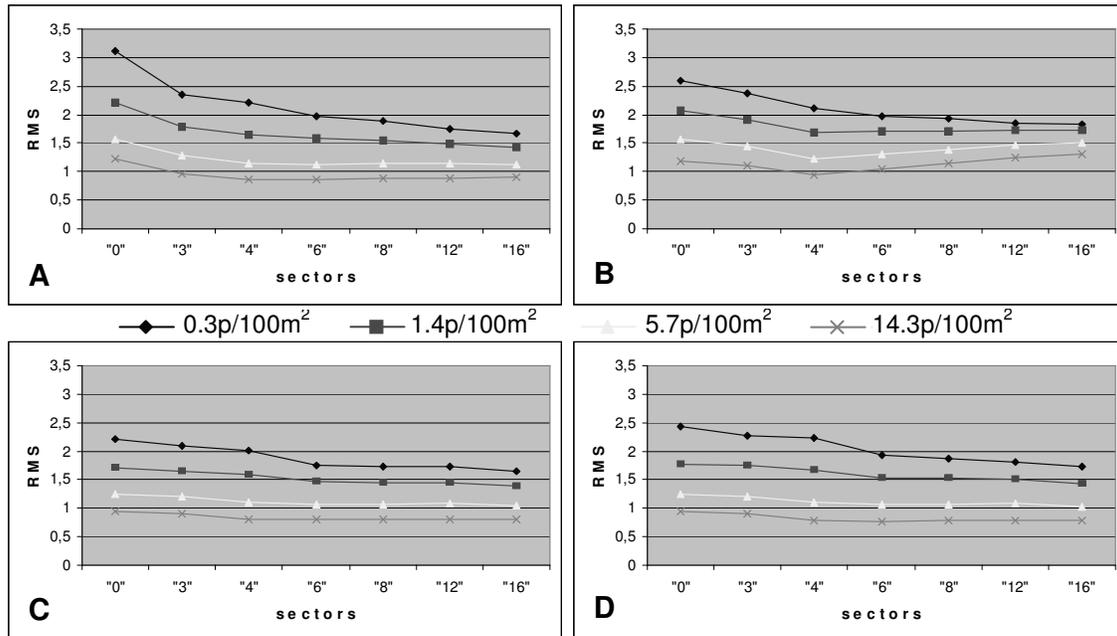


Figure 7: Graphs of RMS coefficient dependence on the number of sectors at different densities for all measurement points

Analyzing the graphs a clear improvement of accuracy of models with the increase of measurement points' density can be observed in all of them (the *RMS* coefficient value decreases). In graphs 7A, C and D the influence of the number of sectors on *RMS* coefficient value decreases with the increase in density of measurement points meaning that the dependence on model quality on the number of sectors decreases. In graph 7A (algorithm $1/R^2$) a clear influence of the number of sectors on the DTM model quality is visible. The *RMS* coefficient value decreases successively with the increase in search sectors number in case of both $0.3p/100m^2$ and $1.4p/100m^2$ densities. Stabilization of accuracy at the level of 4 sectors occurs only in case of a higher density of measurement points. In case of interpolation using local polynomial (fig. 7B) the best accuracy of the model was achieved also in case of 4 sectors. In case of higher densities of measurement points the value of *RMS* coefficient increases with the increase of the number of sectors. In case of graphs 7C and 7D and densities of $0.3p/100m^2$ and $1.4p/100m^2$ model quality stabilizes at the level of 6 sectors while for densities of $5.7p/100m^2$ and $14.3p/100m^2$ at the level of 4 sectors. Further increase in the number of sectors does not improve the interpolation model quality and the *RMS* coefficient value stays at a similar level. In both graphs similar trends can be observed in case of algorithms using kriging (fig. 7C) and the radial functions (fig. 7D).

5. Analysis of DTM models accuracy for selected nodal points

During the second stage of the study, aiming at elimination of the influence of edge points, the area selected using a rectangle 250m X 290m, giving the area of 72500m² was analyzed (fig. 8A,B). Additionally, the area for analysis was selected in a way aiming at minimizing the influence of mismatch between the size of the base square of nodes' grid and morphology of the DTM model.

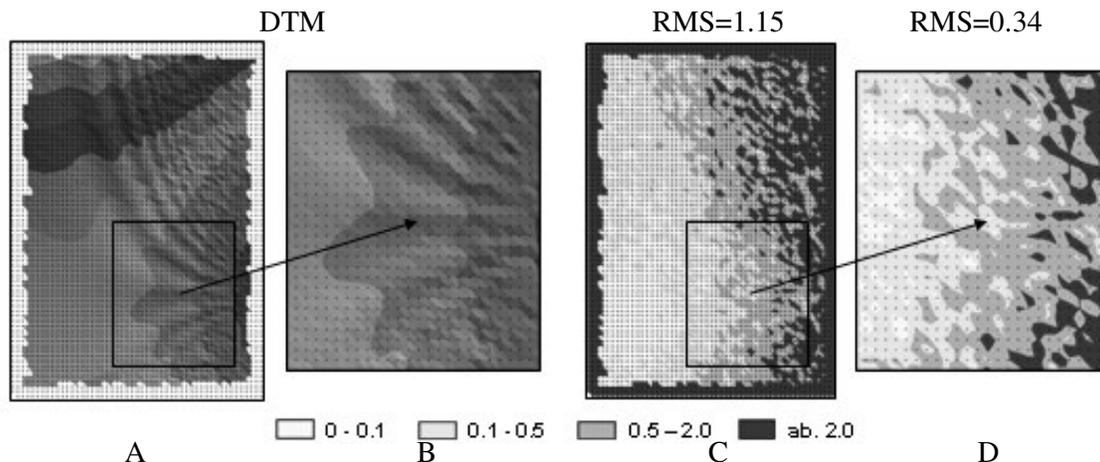


Figure 8: Selection of measurement points in the DTM model and in the differential diagram

Figure 8C presents the differential diagram showing model deformation after interpolation using the $1/R^2$ algorithms in case of 6 search sectors. Above the diagram the *RMS* coefficient is shown for the entire interpolation area after elimination of edge points. Figure 8D shows the differential diagram generated for selected 672 (24 x 28) nodal points calculated in the same way as above. The *RMS* coefficient presented above the diagram indicates a much more accurate match of the area selected using the rectangle to the standard surface.

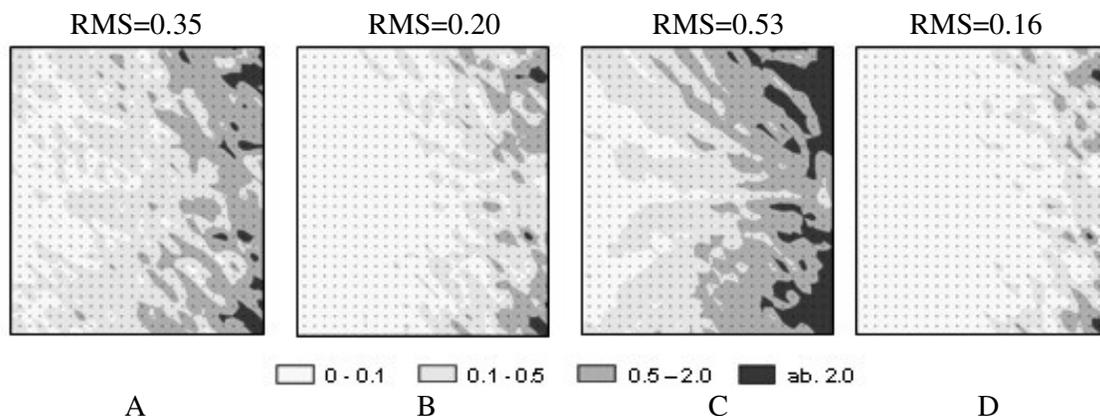


Figure 9: Comparison of interpolation algorithms accuracy for four search sectors at the same density of measurement points

The coefficient value is lower because the selected area had the base square size better matching the DTM morphology and the influence of errors developing at edge points was eliminated. A similar situation also occurs in cases presented in figure 9. Further differential diagrams present results of analyses for different interpolation algorithms: $1/R^2$ (fig. 9A), kriging (fig. 9B), local polynomial (fig. 9C) and radial functions (fig. 9D). In all cases 16

measurement points, 4 search sectors and density of 14.3p/100m² were used for interpolation of each node. The worst matching of the DTM model was achieved in case of local polynomial while the best was offered by kriging and radial functions algorithms.

Similar to the preceding case, the dependence of *RMS* coefficient on the number of sectors was analyzed for the selected areas. For that purpose graphs for different interpolation algorithms: 1/R² (fig. 10A), local polynomial (fig. 10B), kriging (fig. 10C) and radial functions (fig. 10D) were generated. Graphs 10A and 10B present the situation where there is one measurement point per each sector while in graphs 10C and 10D the number of points per sector depends on the number of sectors but in each case totals 16. Thanks to selection of areas the interpolation results were improved for all algorithms. As a consequence of a significant improvement in DTM models quality the scale of the vertical axis of values was changed.

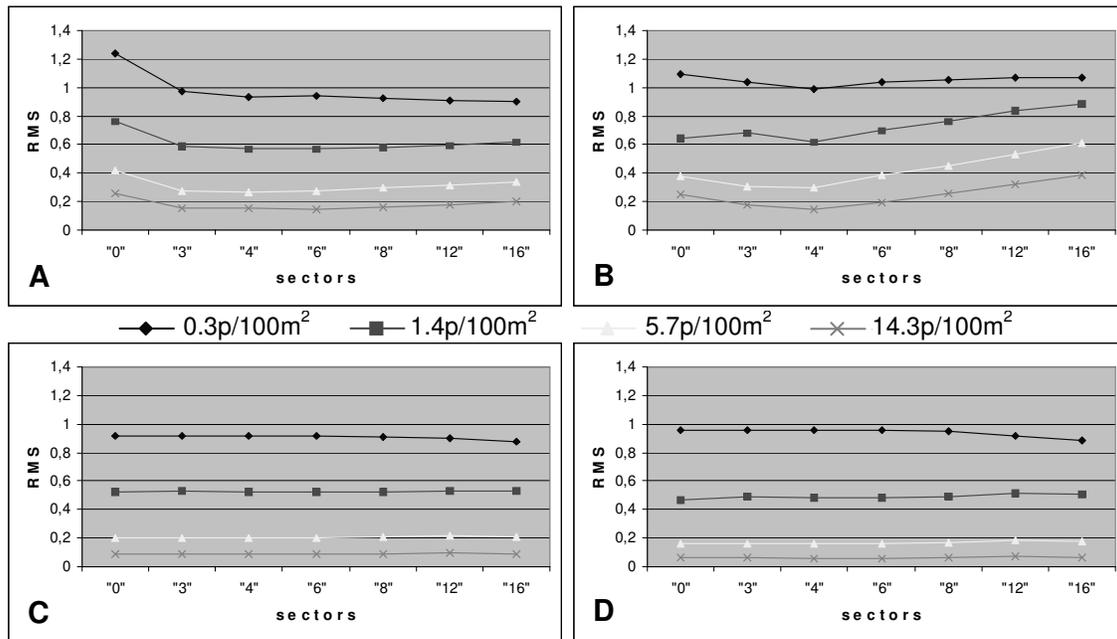


Figure 10: Graphs of RMS coefficient dependence on the number of sectors in case of different densities for selected measurement points

Also in this case a model quality improvement can be observed in all graphs with the increase in measurement points' number. Interpolation using the 1/R² algorithm offered the worst results during measurement points search without sectors. In case of that algorithm use of 4 search sectors offered optimum results (fig. 10A). A similar situation occurred for the local polynomial based algorithm (fig. 10B). In this case increasing the number of sectors above 4 deteriorated model quality for each case of measurement points' density. Very similar accuracies of interpolated models were obtained while using algorithms based on kriging (fig. 10C) and radial functions (fig. 10D). In both cases the accuracy of interpolation did not depend on the number of sectors. After increasing the density to 14.3p/100m² the accuracies of calculations using those algorithms were the best among all tested cases.

6. Comparison of interpolation algorithms accuracy

For the purpose of accuracy comparison of interpolation algorithms DTM models generated for all (fig. 11A) and selected (fig. 11B) nodal points were analyzed. The interpolation was carried out for 16 measurement points positioned in 4 search sectors for each nodal point. The

vertical axis shows the *RMS* coefficient value, the density was placed along the horizontal axis while series consisted of consecutive interpolation algorithms.

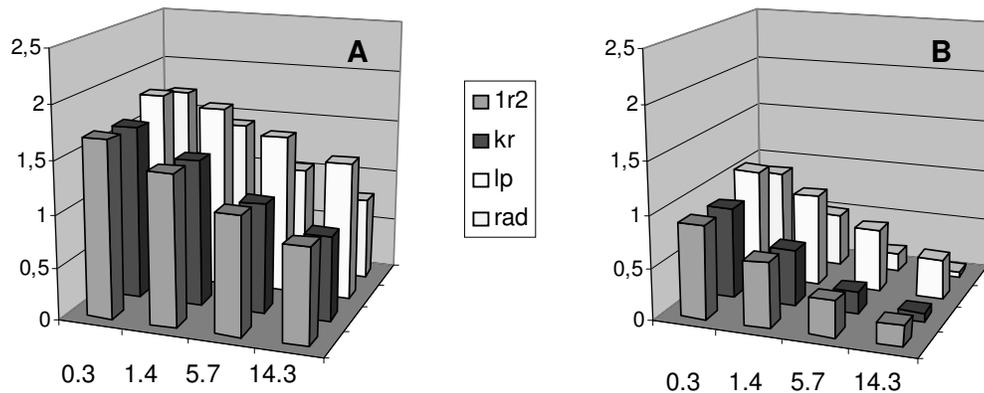


Figure 11: Comparison of interpolation algorithms accuracy

The graphs allow noticing an increase in DTM models accuracy with increasing the density of measurement points. A clear improvement in accuracy for all algorithms occurred during interpolation made for selected points (free from edge errors and with adjusted size of the base square) (fig. 11B). In both cases the higher quality of models using algorithms based on kriging and radial functions is visible. Interpolation with those two algorithms offers compatible results.

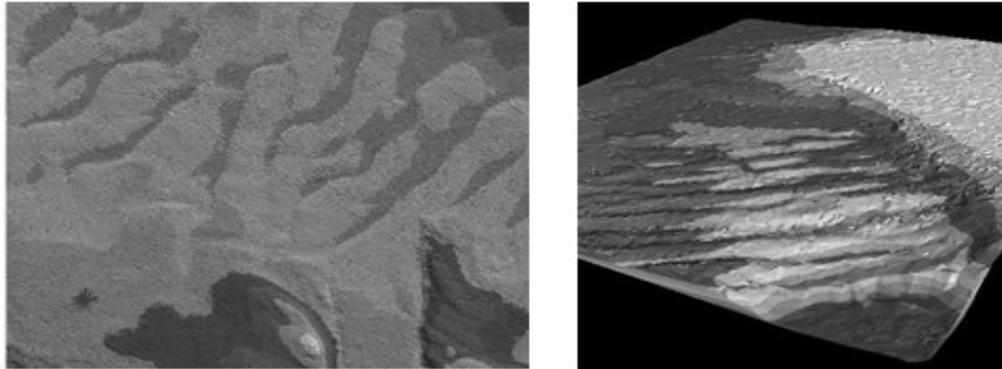


Figure 12: Samples of DTM models

Two DTM models were shown in figure 12 as a practical example. Both models were generated on the basis of measurement points obtained from multiple beam sonic sounder. Interpolation was carried out using the radial functions based algorithm. 16 measurement points positioned in 4 search sectors were used for calculation of each nodal point. Edge nodes were eliminated from the models.

7. Conclusions

Application of differential diagrams for analysis allows locating deformations of interpolation models. Such diagrams allow presentation of the scope of deformations within the determined accuracy ranges. Application of the RMS coefficient, on the other hand, allows determining the accuracy of matching of the entire surface using a single value. That in turn allows

comparative analyses concerning the influence of the distribution of measurement points on accuracy of generated DTM models. During generation of digital terrain models using GRID structure a significant role is played by correct selection of measurement points used for calculations. Such selection is possible using sectors. The analyses carried out indicate that selection of the number of search sectors should depend on the density of measurement points. Determining the optimum number of search sectors was important for the majority of analyzed cases. In case of low densities use of six sectors is the optimum solution; in case of larger densities four sectors are optimal. The algorithm based on radial functions was the most accurate among the analyzed algorithms. During studies aiming at determining interpolation accuracy attention should be paid to the influence of edge points errors. Adjustment of the base square to the morphology of the studied surface has also a major influence on the results. Interpolation accuracy using the optimum number of search sectors was confirmed in practice during interpolation on a real digital terrain model.

References:

- [1] Cressie, N.A.: Statistics for Spatial Data, John Wiley and Sons, Inc., New York, 1991.
- [2] Davis, J.C.: Statistics and Data Analysis in Geology, John Wiley and Sons, New York, 1986.
- [3] Draper, N., H. Smith: Applied Regression Analysis, second edition, Wiley-Interscience, 1981.
- [4] Fisher, P.F.: Spatial Data Sources and Data Problems, Geographical Information Systems, Longman Scientific & Technical, New York, 1991.
- [5] Gościewski, D.: Effect of the distribution of measuring points on GRID network generation, International Scientific-Technical Conference EXPLO-SHIP, Scientific Fascicles WSM. Szczecin, 2004.
- [6] Gościewski, D.: Influence of measuring point location on selection of interpolation algorithms, The 6th International Conference: Environmental Engineering, Gediminas Technical University Press, Vilnius, 2005.
- [7] Lue, Y., K. Novak: Recursive Grid-Dynamic Window Matching for Automatic DEM Generation, ACSM-ASPRS Fali Conception Technical Papers, 1991.
- [8] Marble, D.F.: Geographic Information Systems: Añ Overview, Introductory Readings in Geographic Information Systems,. Bristol, Pennsylvania, 1990.
- [9] Obermeyer, N. J., J. K. Pinto: Managing Geographic Information Systems, Guilford Press, Los Angeles, 1994.
- [10] Parker, D.: Innovations in GIS, University of Newcastle, Newcastle, 1996.
- [11] Powell, M.J.D.: The Theory of Radial Basis Function Approximation in 1990, University of Cambridge Numerical Analysis Reports, DAMTP, 1990.
- [12] Walker, T. C., R. K. Miller: Geographic Information Systems: An Assessment of Technology, Applications, and Products, Madison, Georgia, 1990.