Section: Technologies of Geodesy and Cadastre

An analysis of processing bulk data with regard to the efficiency of the creation of the GRID structure used in the generation of the digital terrain mode

Dariusz Gosciewski

University of Warmia and Mazury in Olsztyn , Faculty of Geodesy and Land Management, Institute of Geodesy,
ul. Heweliusza 12, 10-957 Olsztyn, Poland

Abstract

Current survey systems (LiDAR, multibeam echo sounders) allow the automated obtaining of a large amount of surface object measurements in a relatively short time. The large amount of information, survey specificity and the character of the created data sets, as a rule, do not allow them to be used directly for the generation of the digital terrain model (DTM) in spatial information systems (SIS, GIS). The handling of measurement results of this type consists in optimization of set sizes and adaptation of their structure to both the user's requirements and technical processing capabilities. The specificity and universal character of the DTM information layer are an inducement to place particular emphasis on data recording organization and keeping their amount to a minimum. At the same time, it is intended to maintain maximum space description fidelity and its dynamic transformation. Efficient generation of information stored in digital map databases often requires data structure analysis and rationalization of the whole process with regard to processing speed. This paper shows an example of optimizing the efficiency of handling bulk measurement results for DTM creation using the GRID structure. Measurement sets and their processing sequence are also analysed. Methods for selecting the location of measurement points and efficient interpolation algorithms allowing an accurate GRID structure to be generated in a relatively short time are then presented.

Keywords: GIS; digital terrain model; regular GRID structure; interpolation algorithms; large observation sets.

1. Introduction

The currently widely-applied spatial information systems (SIS, GIS) use the digital terrain model (DTM) as one of the information layers [1–3]. It is the fundamental layer and provides the basis for spatial organization of stored object information and helps to locate physical phenomena [4–6]. Different groups of users have reported the need for ever-more accurate terrain surface models. These users have increasing requirements, placing emphasis on data quality, accuracy, reliability and its currentness [7], [8]. Equally great weight is currently attached to data processing dynamics and the possibility of performing real-time analysis is becoming an important element of different handling methods [9]. The specificity and universal character of the DTM information layer is an inducement to put special emphasis on data recording organization and reduce data quantity to a minimum, while also maintaining maximum space description fidelity resulting from survey accuracy. Reducing the data volume stored in this information layer allows to accelerate its use and shorten the processing time to a minimum. Minimizing the size of archived sets ensures quick access to information and dynamic real-time analysis.

The creation of a digital relief model requires prior collection of survey information. There are many methods for terrain information acquisition widely described in the literature [2–6]. Direct surveys are the most accurate and most reliable group among them. By using modern survey techniques, it is possible to obtain high accuracies with maximum automation of the survey process and with minimum human work input [10], [11]. Widely-used integrated mass survey systems, such as LiDAR laser scanning (TLS, ALS, SLS) or multibeam echo sounders ensure a considerable degree of information acquisition automation [12]. Such systems can obtain a very large data volume in a relatively short time. Ever-newer technical solutions enable ever-more accurate determination of the position of points while increasing their density.

When a large data volume is available, the surface model can be described very accurately, but real-time dynamic processing of such a data volume is practically impossible. The acquired data is often disordered and rational handling of a large volume of mass survey results requires the application of a proper processing technology [13], [14]. One of the methods for reducing the data volume needed to create a digital surface representation is the use of a GRID structure. It orders the data and prevents its redundancy, which contributes significantly to speeding up its processing and affects the dynamics of the whole analytical process. During the processing of survey data for terrain surface modeling using a GRID structure, it is worth paying close attention to several important stages in the procedure.

2. Analysis of survey sets

The first stage, preceding data processing proper, involves the analysis and evaluation of survey sets. It allows selecting the correct processing method in further stages of DTM generation. Analysis of the structure and recording logic of sets storing recorded data also helps to determine the optimal method for their processing during GRID structure generation.

Survey resources obtained by integrated mass survey systems are characterized by special features which later determine their processing [15–16]. Because measurements are recorded by using many survey beams (e.g. 500 simultaneously) sent from the survey device in a number of cycles, the volume of acquired data is very large (e.g. 50,000 pts./s). Fig. 1a shows an example distribution of survey points acquired from multi-beam scanning from one scanning head. The recorded points are located in characteristic survey lines (lying perpendicular to the survey direction), which form during each survey cycle of the scanning head. The points are distributed in a scattered manner with non-uniform saturation in different survey areas.

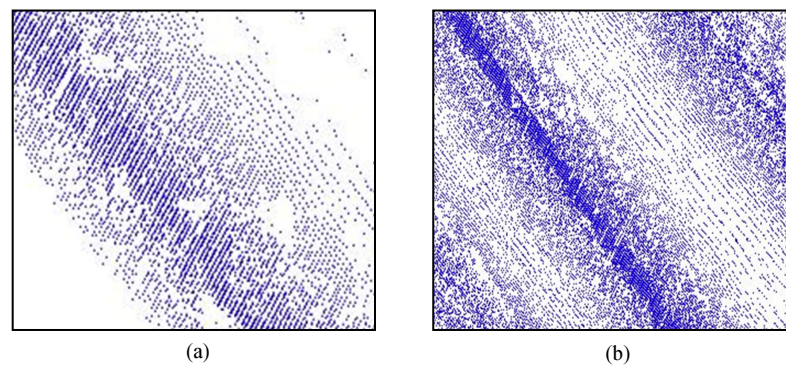(a)   (b)

Fig. 1. Location of recorded survey points

This partly results from the relief of the scanned surface. Breaks and gaps in the survey spectrum are also caused by the obstruction of survey beams by protruding morphological surface elements (trees, buildings) (Fig. 1a). Additionally, to measure the whole area, scanning is performed in several passes and recorded areas often overlap to different degrees (Fig. 1b), causing the duplication of survey zones. Non-uniform density of recorded information is also caused by the fact that the measuring range of the recording beam and the resolution of the recorded points depend on the distance of the scanning head from elements of the scanned surface.

An additional difficulty during survey handling is also sequential recording in consecutive survey cycles on relatively small surface fragments. Data recording is discontinuous and successively supplemented in consecutive parts of the survey file. Consequently, before starting data processing proper, it is advisable to divide the whole set into a number of working subareas containing only points which will be processed simultaneously in one computational cycle.

3. Selection of the digital model representation structure

A DTM is usually created based on points distributed on the terrain surface organized into specific structures. They are most often a regular square grid (GRID) or an irregular triangle network (TIN) [1–3]. The selection of the TIN or GRID representation method for the created model mainly depends on the aim and scope of handling [17]. Assuming that the priority is dynamic real-time data processing, the application of a GRID structure is more advisable. This is connected with the fact that this structure enables a considerable reduction in the volume of data needed for digital surface representation. The advantage of the GRID structure is also easier surface modeling and determination of desired values at any point or section as well as the possibility of using simpler and faster modeling and visualization algorithms. An equally important feature of the GRID is the possibility of using compression algorithms to considerably reduce the set sizes, which allows clearer data recording organization and speeds up processing.

Taking into account the data volume obtainable using integrated mass survey systems, it is possible to generate a GRID structure with a resolution sufficient to create an accurate DTM. Proper selection of node density in the created network remains the fundamental element allowing accurate surface reproduction using a GRID [18], [19]. Fig. 2 shows the node distribution against the background of survey points. A heterogeneous survey set is converted to a homogeneous square grid

by interpolation. The value at each node of the structure is computed by an interpolation algorithm based on survey points found around the node.
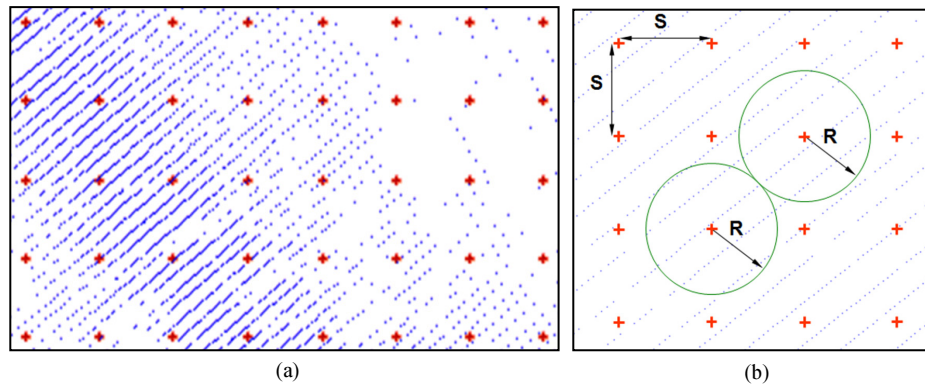

Fig. 2. Generation of a regular square grid (GRID) based on survey sets

Depending on the morphological terrain diversity, the distances (S) between individual nodes (Fig. 2b) should be selected so that the created grid accurately characterizes the surface. It is also intended to find a sufficient number of survey points for a given interpolation method. The selection of such points within an unlimited search radius causes the need to analyze a large data volume and considerably prolongs the GRID generation time. Assuming the analysis of all survey points, the minimum search radius (R) around the nodes should be adopted, equal to half the diagonal of the square of the grid formed by the nodes (Fig. 2b) and the nearest points searched for within the radius. However, linear distribution of survey points (Fig. 2), characteristic for mass survey systems, does not allow free use of the nearest of them during interpolation. This particularly refers to simplified algorithms. Because accurate and reliable interpolation by such algorithms requires correct point distribution around newly-created nodes, algorithms to find properly-located survey points should be applied.

4. Location of survey points for interpolation

Aiming at the most efficient acquisition of values at GRID nodes based on a large survey data volume, the focus should be on the fastest (and at the same time simplest) interpolation methods. Such methods, using a minimal number of survey points, do not always give sufficiently precise results. The speed of a given interpolation algorithm can be combined with maintaining sufficient accuracy by the proper preparation of survey data for interpolation. For this purpose, the interpolation of the value at a grid point should be preceded by finding survey points with a proper location around the node. Determination of such location also aims at eliminating cases of nodal point value extrapolation, which may particularly occur for data acquired by devices recording a linear survey beam (Fig. 1). To fulfill the above conditions, it is necessary to use algorithms which can locate the points in the survey set which lie nearest to the node, which are at the same time distributed in the vertices of the surrounding figure. The triangle is the simplest figure of this type. Ultimately, all generated nodes of the created GRID structure should be inside triangles formed by local survey points. This assumption maintains high handling accuracy, while considerably shortening the computational time due to the minimal number of survey points.

There are several methods for finding the proper location of survey points. Fig. 3 shows different methods for the location of properly distributed points around a node within a specific search radius (R). In individual methods, the space around a node is divided into three (Fig. 3a), six (Fig. 3b) or four (Fig. 3c) segments, respectively. In all the cases, the first step is finding points lying nearest to the node in each segment. When the neighborhood of the node is divided into three segments (Fig. 3a), a surrounding triangle or one lying near the node can be created on the basis of the found points. This method is the fastest computationally, but does not strictly fulfill the assumption presented above – finding the surrounding figure.
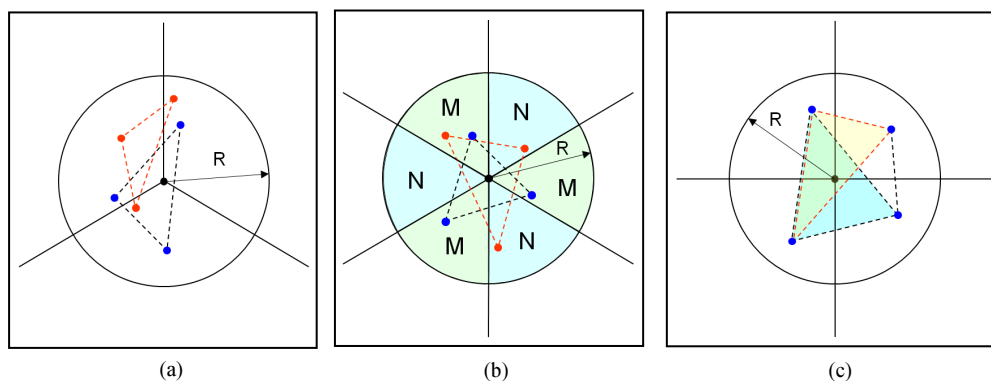

Fig. 3. Methods for location of survey points around a node

The second method, shown in Fig. 3b, allows the selection of points lying only in alternate segments M or N, respectively by dividing the space around a node into six parts. This enables finding even two surrounding triangles and the selection of the one in which the sum of squares of the distances from the node to the vertices is lower. This is also a relatively very fast search method. One should, however, be prepared for situations when a triangle surrounding the node is not found although proper points are present nearby (the points lie simultaneously in segments M and N). Assuming the condition of searching through all variants of point distribution around the node by successive, rotational passing of search segments causes excessive extension of the operation time.

The third described method (Fig. 3c) allows an analysis of all combinations of point positions measured around the determined point. The method's algorithm divides the area around the node into four segments, in which the nearest survey points are found. A quadrilateral is constructed on the basis of the found points, which always contains two triangles surrounding the given node. Ultimately, one is selected in which the sum of squares of the distances from the node to the vertices is lower.

The effect of selecting points for interpolation by the third method, for four example nodes, is shown in Fig. 4a. Reducing the number of points needed for interpolation enables considerable shortening of the computational time in the next processing stage. Elimination of most survey points allows fast determination of node elevation using simplified interpolation algorithms.

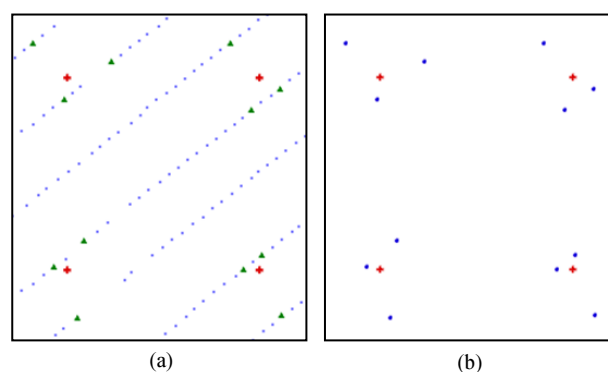(a)   (b)

Fig. 4. Location of survey points for interpolation

The generation of a set storing only points used to conduct interpolation can also be included in data processing (Fig. 4b). Such a set enables analysis of the selection of points for interpolation and a possible change in the search method. This also allows the selection of the interpolation method, depending on the distribution type of points recorded in a particular survey process.

5. Interpolation algorithms

There are many different interpolation methods widely described in the literature [20–22]. They allow value determination at the set node based on the surrounding survey points. Computationally complex algorithms (e.g. Kriging, polynomial regression, radial functions) allow precise determination of the sought values. However, the degree of mathematical complexity of an interpolation method (the number of operations or iterations) closely affects the time of performed computations [23]. Complex algorithms, moreover, need many survey points distributed around the node for computations. Finding proper points in a large volume of data, which were also recorded sequentially in many sets, is a very time-consuming task.

To efficiently generate the nodes of a GRID structure on the basis of a large survey data volume, it is most advisable to select the relatively fastest interpolation methods, which use a minimal number of found points for computations (the least algorithmically complex method). The assumption of using only points forming the surrounding triangle for interpolation considerably reduces the operation complexity degree and speeds up computations. Simple and efficient interpolation algorithms can be used for this task. Operation diagrams for three different interpolation methods are shown in Fig. 5. The surface models they generate are provided under the diagrams. The survey points participating in successive computations (the same for all algorithms) were generated based on the theoretical surface model created by a mathematical function.

The least complex and most efficient computational algorithms are used in the inverse distance power weighting method (Fig. 5a Up). The nodal values are determined from the grand mean, where the weight for each point is its distance from the node raised to the proper power n (here: $n = 2$). This method also allows to determine, with relatively low error, the node elevation for cases where there are not enough points to create a triangle within the adopted search radius or this triangle lies outside the node. The drawback of this method is the excessively high dependence of the computed value at the sought point on the nearest point, which results in model inaccuracies occurring across the entire surface (Fig. 5a Down).

The second relatively fast method is the application of weighting by the areas of opposite triangles (Fig. 5b Up). The nodal values are determined from the grand mean, where the weight for each point is the triangle area lying on the opposite

side of the node (e.g. for point P1 – triangle T1). This allows more advantageous selection of computational coefficients and better balancing of the interpolated value. This method enables the generation of a surface characterized by only local model inaccuracies, solely for when the areas of one or two triangles are extremely small (Fig. 5b Down).
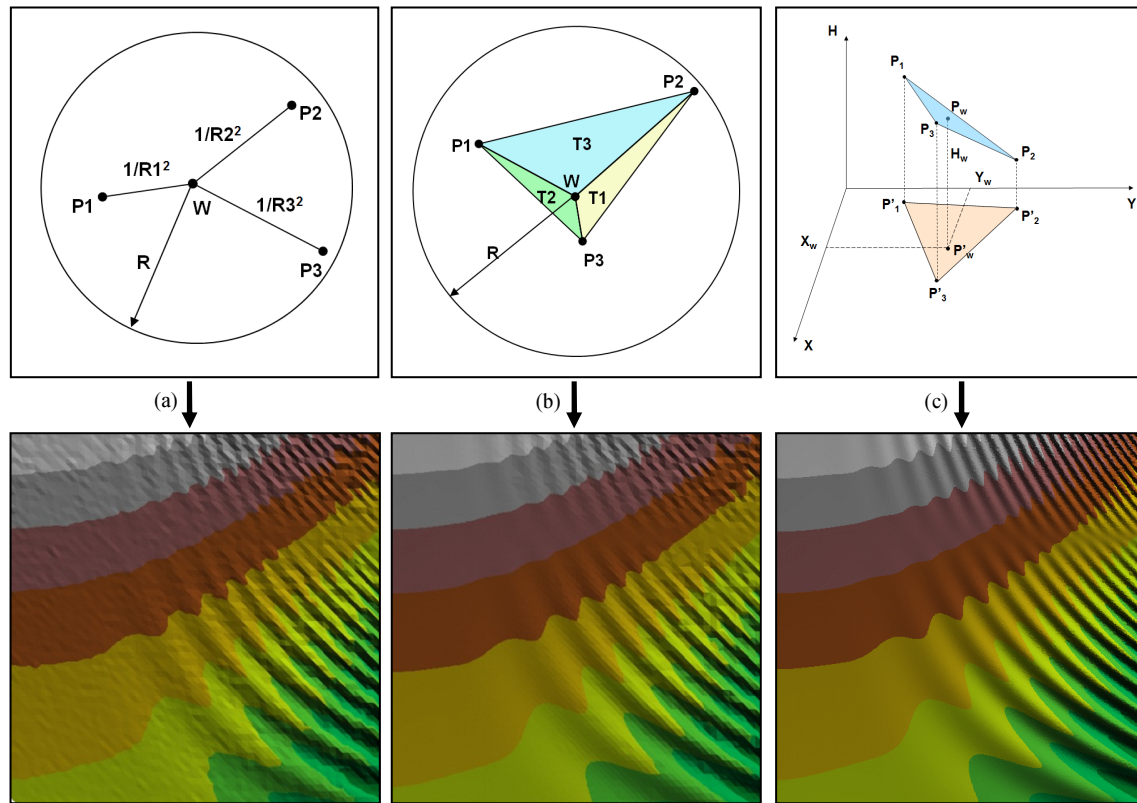

Fig. 5. Interpolation algorithm diagrams and interpolation effects

The third method allows the computation of the elevation of the interpolated point as the solution of a two equation system: of the plane passing through the three survey points forming the triangle with the node inside and of the straight line passing through the determined node (Fig. 5c Up). This is also a fast interpolation method enabling precise determination of the sought point. The interpolation surface determined by this algorithm is the best quality compared to the other methods (Fig. 5c Down). However, owing to more extended computational operations, this method is slightly slower than the others.

When conducting interpolation, combinations of different methods can be applied, making their use dependent on survey data analysis and changing them for different point locations. The averaging of obtained results can also be applied when the number of found surrounding triangles is higher. This causes increased accuracy of the determined nodal value, but prolongs the computational time.

6. Data processing after interpolation

The last processing stage is ordering the data of the GRID structure after interpolation and its archiving. Due to the overlapping of successive survey areas and the sequential character of information processing, nodal values at the contact points of working areas can be computed more than once. The values at duplicated nodal points should be averaged and then, to avoid redundancy, repeated points are eliminated. The next step is the sorting of GRID structure nodes performed, depending on the needs, along columns or lines, enabling analysis by profiles. The maximum number of computational records stored in single files is determined during the final archiving of sets, because of limitations as to the volume of simultaneously processed data in a given GIS system. Finally, depending on the requirements set by the hardware and software platform, the proper form of recording and organization of computational records is generated. The universal form of data exchange in this case are ASCII sequential files, but because of a high number of records in computational sets, the preferred form of information exchange are binary files with random access. The structure of information recording in such a file is dictated by the requirements of the specific land information system. Fig. 6 shows different examples of the digital terrain model developed in the discussed stages.

The DTM generated using a GRID structure, as is the case for a TIN structure, can be presented as a set of individual nodes (Fig. 6a) or as a continuous surface representation (Fig. 6b). The created model can be used for complex dynamic analysis because it is built of a relatively low number of points and contains their ordered structure.
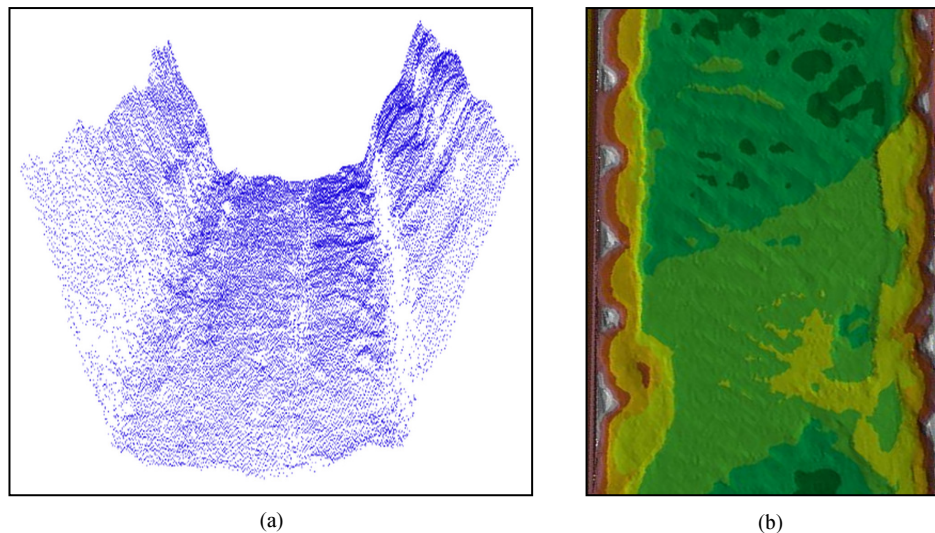
(a)   (b)

Fig. 6. Examples of DTM development

7. Summary

Bearing in mind the constantly growing needs for land data collection, it is necessary to develop methods for fast, up-to-date and accurate data processing. The constant need for increasing the accuracy of handling for ever larger observation sets, with simultaneous time limitations, is partly satisfied by the continuous development of computer hardware. However, despite a constant increase in the information flow rate, its dynamic processing still has limitations. Rational use of mass information during DTM construction should be preceded by efficient data preparation. It is therefore necessary to optimize data processing methods, whose reduction while maintaining proper handling accuracy gives measurable benefits. By selecting appropriate processing algorithms, it is possible to reduce the volume of information needed for DTM generation while maintaining maximum model accuracy. The application of proper procedures and stages during data processing also allows rational use of the available computer hardware and software. The implementation of the principles specified above allows efficient optimization of resultant sets, which can serve for dynamic DTM creation. This enables a considerable shortening of the times needed for handling the results, which is particularly important in the analysis of real-time processes.

References

[1] Goodchild, M. F.; Longley, P. A. 2005. The future of GIS and spatial analysis, in P.A. Longley; M. F.Goodchild; D. J. Maguire; D. W. Rhind (Eds.). *Geographical Information Systems: Principles, Techniques, Management and Applications*. Hoboken, NJ.: Wiley, 567–580. http://dx.doi.org/10.1111/j.1467-9671.2005.00233.x

[2] Carlisle, B. H. 2002. *Digital Elevation Model Quality and Uncertainty in DEM-based Spatial Modelling*. Ph.D. Thesis, University of Greenwich, London, UK.

[3] Li, Z.; Zhu, Q.; Gold, C. 2005. *Digital Terrain Modeling: Principles and Methodology*. CRC Press: Boca Raton, FL.

[4] Carlisle, B. H. 2005. Modelling the spatial distribution of DEM error, *Transactions in GIS* 9(4): 521–540.

[5] O'Sullivan, D.; Unwin, D. J. 2003. *Geographic Information Analysis*. Hoboken, NJ: Wiley.

[6] Harmon, J. E.; Anderson, S. J. 2003. *The Design and Implementation of Geographic Information Systems*. Hoboken, NJ: Wiley.

[7] Openshaw, S.; Alvanides, S. 2005. Applying geocomputation to the analysis of spatial distributions, in P.A. Longley; M. F. Goodchild; D. J. Maguire; D. W. Rhind (Eds.). *Geographical Information Systems: Principles, Techniques, Management and Applications*. Hoboken, N.J.: Wiley, 267–282.

[8] Wechsler, S. P. 2003. Perceptions of digital elevation model uncertainty by DEM users, *URISA Journal* 15(2): 57–64.

[9] Ehlschlaeger, C. R. 2002. Representing multiple spatial statistics in generalized elevation uncertainty models: Moving beyond the variogram, *International Journal of Geographical Information Science* 16(3): 259–285. http://dx.doi.org/10.1080/13658810110099116

[10] Aruga, K.; Sessions, J.; Akay, A. E. 2005. Application of an airborne laser scanner to forest road design with accurate earthwork volumes, *Journal of Forest Research* 10(2): 113–123. http://dx.doi.org/10.1007/s10310-004-0116-9

[11] Florinsky, I. V. 2002. Errors of signal processing in digital terrain modelling, *International Journal of Geographical Information Science* 16(5): 475–501. http://dx.doi.org/10.1080/13658810210129139

[12] Axelsson, P. 2000. DEM generation from laser scanner data using adaptive TIN models, in K. J. Beek (Ed.). *Molenaar M International Archives of Photogrammetry and Remote Sensing*. GITC: Amsterdam, Netherlands, 110–117.

[13] Shi, W.; Fisher, P. F.; Goodchild, M. F. 2004. Recent developments in modeling uncertainties in geospatial data and analysis, *Photogrammetric Engineering and Remote Sensing* 70(8): 919–920.

[14] Gosciewski, D. 2013. The effect of the distribution of measurement points around the node on the accuracy of interpolation of the digital terrain model, *Springer, Journal of Geographical Systems* 15(4): 513–535. http://dx.doi.org/10.1007/s10109-012-0176-x

[15] Pebesma, E. J. 2004. Multivariable geostatistics in S: the gstat package, *Computers and Geosciences* 30(7): 683–691. http://dx.doi.org/10.1016/j.cageo.2004.03.012

[16] Fowler, R. 2001. Topographic Lidar, in D. F. Maune (Ed.). *Digital Elevation Model Technologies and Applications: The DEM Users Manual*. American Society for Photogrammetry and Remote Sensing: Bethesda, MD, 207–236.

[17] Pitas, I. 2000. *Digital Image Processing Algorithms and Applications*. John Wiley and Sons, NY.

[18] Gosciewski, D. 2013. Selection of interpolation parameters depending on the location of measurement points, *Taylor & Francis, GIScience & Remote Sensing* 50(5): 515–526.

[19] Fisher, P. F.; Tate, N. J. 2006. Causes and consequences of error in digital elevation models, *Progress in Physical Geography* 30(4): 467–489. http://dx.doi.org/10.1191/0309133306pp492ra

[20] Raaflaub, L. D.; Collins, M. J. 2006. The effect of error in gridded digital elevation models on the estimation of topographic parameters, *Environmental Modelling and Software* 21(5): 710–732. http://dx.doi.org/10.1016/j.envsoft.2005.02.003

[21] Wilson, J.; Gallant, J. 2000. *Digital terrain analysis. Terrain Analysis*. Principles and Applications, John Wiley and Sons, New York 2000.

[22] O'Sullivan, D.; Unwin, D. J. 2003. *Geographic Information Analysis*. Hoboken, NJ: Wiley.

[23] Schabenberger, O.; Gotway, C. A. 2005. *Statistical Methods for Spatial Data Analysis*. Chapman and Hall, CRC: Boca Raton, FL.