Acta Geodyn. Geomater., Vol. 13, No. 4 (184), 363–368, 2016
DOI: 10.13168/AGG.2016.0018

journal homepage: https://www.irsm.cas.cz/acta

ORIGINAL PAPER

INFLUENCE OF DATASETS DECREASED BY APPLYING REDUCTION AND GENERATION METHODS ON DIGITAL TERRAIN MODELS

Wioleta BLASZCZAK-BAK [1] *, Katarzyna PAJAK [2] and Anna SOBIERAJ [3]

[1] Institute of Geodesy, University of Warmia and Mazury in Olsztyn, Oczapowski St. 1, Olsztyn, Poland
[2] Land Surveying and Geomatics, University of Warmia and Mazury in Olsztyn, Heweliusza St. 12, Olsztyn, Poland
[3] Department of Geodesy, Gdansk University of Technology, Narutowicza St. 11/12, Gdansk, Olsztyn

**Corresponding author's e-mail:* wioleta.blaszczak@uwm.edu.pl

ARTICLE INFO

Article history:
Received 19 January 2016
Accepted 13 June 2016
Available online 20 June 2016

Keywords:
Reduction
Generation
DTM

ABSTRACT

The number of point clouds provided by LiDAR technology can be sometimes seen as a problem in development and further processing for given purposes (e.g. Digital Terrain Model (DTM) generation). Therefore, there is still a need to reduce the obtained big datasets. Reducing can be done, inter alia, by reducing the size of the set or by generating the set. This paper presents two variants of the reduction of point clouds in order to effectively generate the DTM. There is also a comparison of two DTMs generated on the basis of sets reduced by various methods.

1. INTRODUCTION

LiDAR technology allows for the acquisition in a short time a large amount of data - big dataset. This data is the basis for the development of a variety of products (e.g. DTM - Digital Terrain Model, DSM - Digital Surface Model). Obtaining them involves the point cloud processing carried out by different approaches (e.g. Sithole and Vosselman, 2005; Tovari and Pfeifer, 2005; Hebel and Stilla, 2008; Suchocki, 2009, 2013; Reitberger et al., 2009; Vosselman, 2008; Saeedi et al., 2009; Vosselman and Maas, 2010). In these papers methods for development of big datasets, particularly filtration methods related to DTM generation, are presented.

Due to the large number of dataset obtained in measurement there is a problem with their development. Decreasing the number of dataset can be done in the pre-processing stage. It is carried out in a way as to not lose the data necessary for the proper implementation of this objective study. Decreasing the number of big datasets can be conducted by means of:

- generation,
- reduction.

Both methods can be used for decrease of the big dataset. However, the effect of each is totally different. The choice of the method depends on the type of the data and the purpose of the study. In this paper the effects of each method on the same set and with the same purpose were compared. Such comparison allowed to indicate advantages and disadvantages of these method and quality assessment of DTMs obtained from decreased sets.

1.1. GENERATION

Generation is decreasing the size of dataset by creating a GRID. In this method we have new points instead of points with the original coordinates (Gościewski, 2013; Bauer-Marschallinger et al., 2014).

Generation involves creating a grid of regular figures. The grid nodes are interpolated, they have new coordinate values.

These coordinates are calculated on the basis of measurement data located in vicinity of interpolated points. For interpolation the following methods, inter alia, can be used (Stateczny and Lubczonek, 2004): kriging, radial basis function – multiquadric, triangulation with linear interpolation, natural neighbour, minimum curvature, nearest neighbour, inverse distance to a power.

In this study the kriging method was used. The result of interpolation in this method is based on the model of covariance and does not depend only on the distance between the points (sought, and given). Covariance model describes how to change the weight of the measured data depending on the distance between these points. The values of points close to each other vary less than the points located far away from each other. Generation leading to decreasing the number of points in measurement set can be conducted in various softwares, e.g. Surfer v.8, CloudCompare v.2.6.0.

Cite this article as: Blaszczak-Bak W, Pajak K, Sobieraj A.: Influence of datasets decreased by applying reduction and generation methods on Digital Terrain Models. Acta Geodyn. Geomater., 13, No. 4 (184), 361–366, 2016. DOI: 10.13168/AGG.2016.0018

1.2. REDUCTION

Reduction decreases the size of dataset by removing some points according to given algorithm, remaining points are original points from measurement (Błaszczak, 2006; Błaszczak et al., 2011a, 2011b; Chen, 2012).

Reduction of big dataset can be conducted by using, e.g. Cyclone v.7.3.3. This software allows for reduction during unification of point clouds derived from different measurement stations. Unification can be executed with reduction – low, medium, high, highest or no reduction at all. Point elimination is determined mainly by spacing between them. When object has a complex structure such approach may lead to loss of significant points. In such situation it is desirable to apply algorithm which adjust reduction to the complexity of the measured object. The use of this algorithm enables to reduce the number of points in the point cloud with varying degrees of reduction in different areas. Thus, points, which are not essential in the generation process (e.g. modeling) will be removed from the measurement dataset. The reduction is not random, and each point is tested a priori to removal because of its usefulness.

The size of the point cloud can be reduced, e.g. by applying the algorithm for decreasing the size of the survey result data set (Błaszczak, 2006; Błaszczak-Bak et al., 2011a).

The reducing algorithm for the LiDAR point cloud, which decreases the size of point cloud, applies known methods of cartographic generalization. Firstly, the search strips in the XOY plane need to be created. Detailed calculations are carried out in these search strips. In the case of LiDAR, due to the data acquisition method and their internal record, search strips most frequently correspond to survey strips. The width of the survey strips results directly from the survey and depends on the angle of laser scanning, altitude and flight speed. The reducing algorithm consists of the following stages:

- *stage 1*: Defining survey strips in the X0Y plane, parallel to the 0Y axis.
- *stage 2*: Selecting the method of cartographic generalization to reduce the size of the survey data set.
- *stage 3*: Application of the chosen generalization method in each strip (in the YOZ plane).

An important stage is the choice of the generalization method. For the present study, the Visvalingam–Whyatt method (V–W) (Visvalingam and Whyatt, 1992) has been selected. In the V–W method a generalized line of triangles from the nearest points is created. The surface of the calculated areas of triangles is compared to the area of the tolerance triangle, the size of which determines how many points will be removed. The size of the tolerance triangle area is established by the user on the basis of statistical characteristics of the survey result data set. It is possible, for example, to assume that its value equals the area of the equilateral triangle with a side corresponding to a minimum distance between the points of the data set (Błaszczak, 2006). If the area of the triangle defined on the basis of the survey results exceeds the area of the tolerance triangle, then the second point of the analysed triangle is maintained, otherwise it is removed.

2. MATERIALS OF THE RESEARCH

The tested LiDAR dataset was obtained from the Idaho Geospatial Data Clearinghouse, University of Idaho Library (Internet access at http://inside.uidaho.edu). LiDAR acquisition was obtained on December 2003 by a subcontractor, EarthData Aviation. The Navajo Chieftain was equipped with an LH System ALS40. The LiDAR system included an inertial measuring unit and a dual frequency airborne GPS receiver. LiDAR data were collected with a 2.0 - 2.2 m nominal post spacing. For the purposes of this study, a subset containing 10 361 points was used, as presented in Figure 1.

The selected fragment was filtered by using adaptive TIN method (Axelsson, 2000) on own software. As a result of the filtration, there are two sets of data: a) the set of points showing the topography (topographic surface dataset - TS*set*) (8683 points), b) a set of points showing the situational details (1376 points). Point clod after filtration is shown in Figure 2.

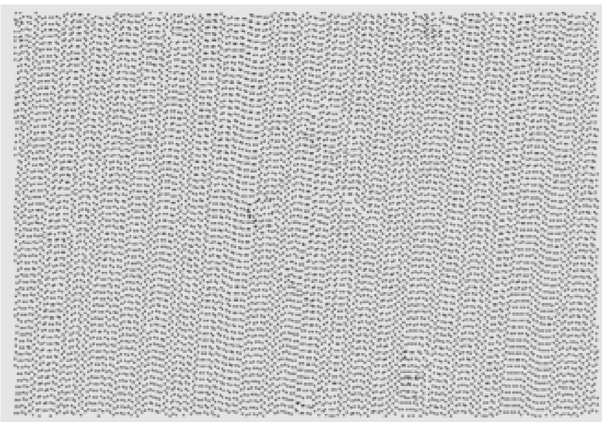

Fig. 1 Original big dataset (source: own study in own software).

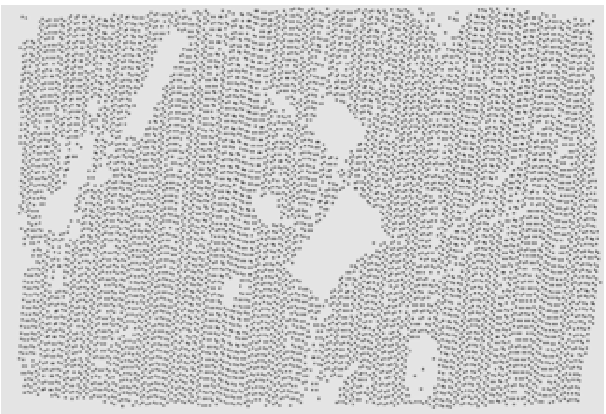

Fig. 2 Big dataset after filtration based on adaptive TIN Metod (source: own study in own software).

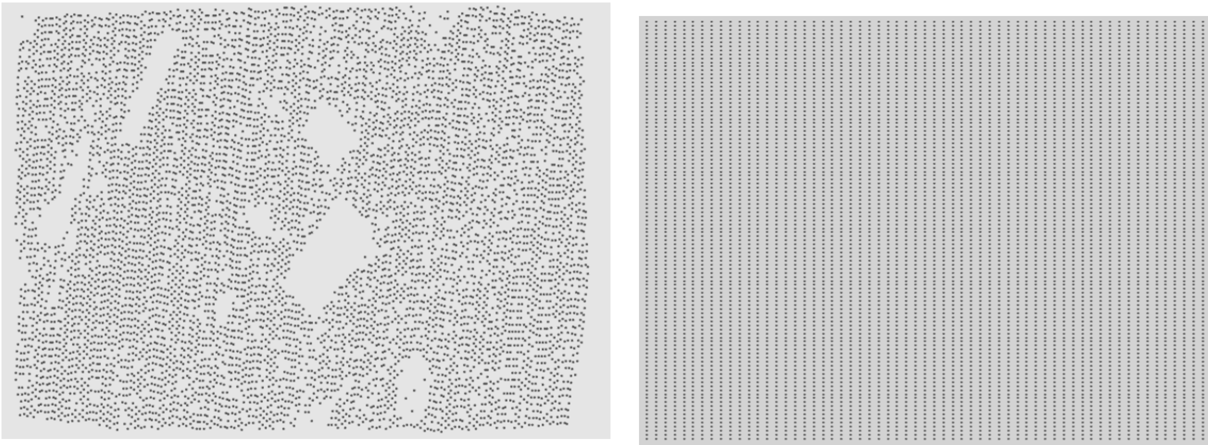

Fig. 3 Decreased TS*set* a) by means of reduction – R*set* (source: own study in own software), b) by conducting GRID generation - G*set* (source: own study in Surfer v.8).

Table 1 Comparison of the reduction method and generation method.

Total number of points in original LiDAR dataset 10361 Number of terrain points in TS*set* 8683 z_{mean} = 415.984 m Range (R) = 14.400 m m_0 = 3.699 m			
Reduction (R*set*)		Generation (G*set*)	
Number of terrain points	7458	Number of terrain points	6344
Z_{mean}	415.869 m	Z_{mean}	415.916 m
m_0	3.729 m	m_0	3.751 m
R	14.090 m	R	14.351 m

3. NUMERICAL TESTS

3.1. *DECREASING THE NUMBER OF TSset*

Authors assumed, that DTM generated on the basis of TS*set* is the best representation of terrain. However, TS*set* is a big dataset, therefore it was decided to decrease the set and generated DTM again using two methods, reduction and generation. TS*set* after applying reduction algorithm presented in (Błaszczak-Bak et al., 2011) is shown in Figure 3a. A set decreased by GRID generation is presented in Figure 3b.

In a reduction variant following parameters were adopted: width of search strip is 5 m, surface tolerance in V-W method is 0.08 m. In a generation variant it was assumed, that the grid size is 5 m, and in the interpolation points kriging method was used. Any further reduction of the grid size resulted in an increasing the number of points.

3.2. *COPMARISON OF OBTAINED DECREASED SETS*

To compare obtained results authors decide to use following parameters:

(a) mean height calculated from heights of R*set* and G*set*: z_{mean}– mean height

(b) mean error:

$$m_0 = \sqrt{\frac{\sum \left(z_{mean} - z_i \right)^2}{k-1}} \qquad (1)$$

where: z_i (i=1,2…, k) – heights of the point, k – size of the R*set* or G*set*,

(c) range: R = z_{max} – z_{min}, where z_{max} – is a maximum height and z_{min} – minimal height.

The characteristics of the obtained sets (after reduction R*set* and the after generation G*set*) are shown in Table 1.

The original set after filtering TS*set* consists of 8683 points. Applying the reduction method caused a decreasing the number of the set by 1225 points (14 %), while generating GRID – decrease by 2339 (27 %). Error m_0 increased in both cases: about 3.0 cm for a R*set* and about 5.2 cm for the G*set*. Z_{mean} in the R*set* decreased by 11.5 cm, and in G*set* by 6.8 cm. Large differences are observed for parameter R (range). R is smaller about 31cm for R*set* and only 4.9 cm for G*set*.

On the basis of obtained sets three DTMs were generated: DTM TS*set* (Fig. 4), DTM R*set* (Fig. 5a), DTM G*set* (Fig. 5b). For all DTMs grid 1m x 1m was adopted, so they can be compared.

Analyzing the construction of obtained models it can be seen that the model based on the G*set* is more smoothed. More details can be seen in DTM generated from the R*set*.

4. COMPARISON OF OBTAINED DTMS

On the basis of R*set* and G*set* two models were generated: DTM R*set* and DTM G*set*, respectively. Those DTMs were compared with the DTM based on a TS*set*. Thus, differences between the obtained surfaces were calculated for two variants: 1) the DTM TS*set* – DTM R*set*; 2) the DTM TS*set* – DTM G*set*. For comparison the following parameters were used:

(a) ΔZ_{min} - minimum height difference between DTMs,

(b) ΔZ_{max} - the maximum height difference between DTMs,

(c) ΔZ_{mean} - the average height difference between DTMs,

(d) RMSE – root mean square error,

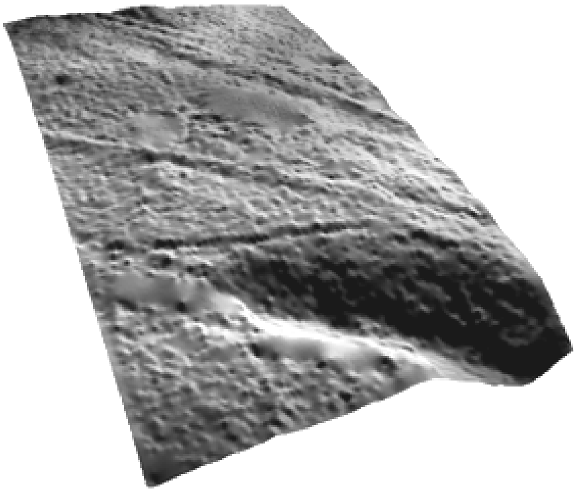

Fig. 4 DTM TS*set* (source: own study in Surfer v.8).

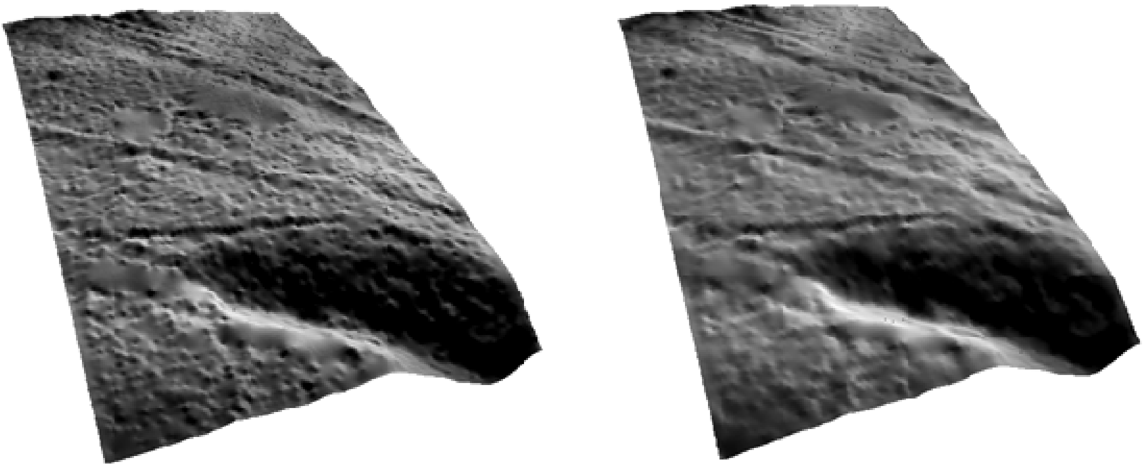

Fig. 5 DTMs a) DTM R*set*, b) DTM G*set* (source: own study in Surfer v.8).

(e) VAR – variance,

(f) coefficient of determination, which is the measure of model adjustment (the closer to 1, the better the match of the model to another model):

$$D^2 = \frac{\sum_{i=1}^{k}\left(z_{DTM\,Rset} - z_{mean}\right)^2}{\sum_{i=1}^{k}\left(z_{DTM\,Tset} - z_{mean}\right)^2} \quad \text{and}$$

$$D^2 = \frac{\sum_{i=1}^{k}\left(z_{DTM\,Gset} - z_{mean}\right)^2}{\sum_{i=1}^{k}\left(z_{DTM\,Tset} - z_{mean}\right)^2} \tag{2}$$

The obtained results are shown in Table 2.

Analyzing the differences between DTMs, in particular the minimum and maximum heights, it can be concluded that the DTM generated from the R*set* in relation to the DTM TS*set* differs more than DTM G*set*. Values ΔZ_{min} and ΔZ_{max} for the second variant are smaller by 18.52 cm and 5.94 cm, respectively, in comparison to first variant. However, a more meaningful measure of the DTMs variation in relations to DTM TS*set* is the variance and standard deviation. The variation in the first variant is twice lower than for the second variant. The standard

Table 2 Comparison of the DTMs generated on the basis of R*set* and G*set* with DTM generated from TS*set.*

DTM differences	ΔZ_{min} [cm]	ΔZ_{max} [cm]	ΔZ_{mean} [cm]	RMSE [cm]	VAR [cm]	D^2
DTM TS*set* – DTM R*set*	-37.99	26.16	-0.06	0.0040	2.49	0.970
DTM TS*set* – DTM G*set*	-19.47	20.22	0.01	0.0061	5.84	0.980

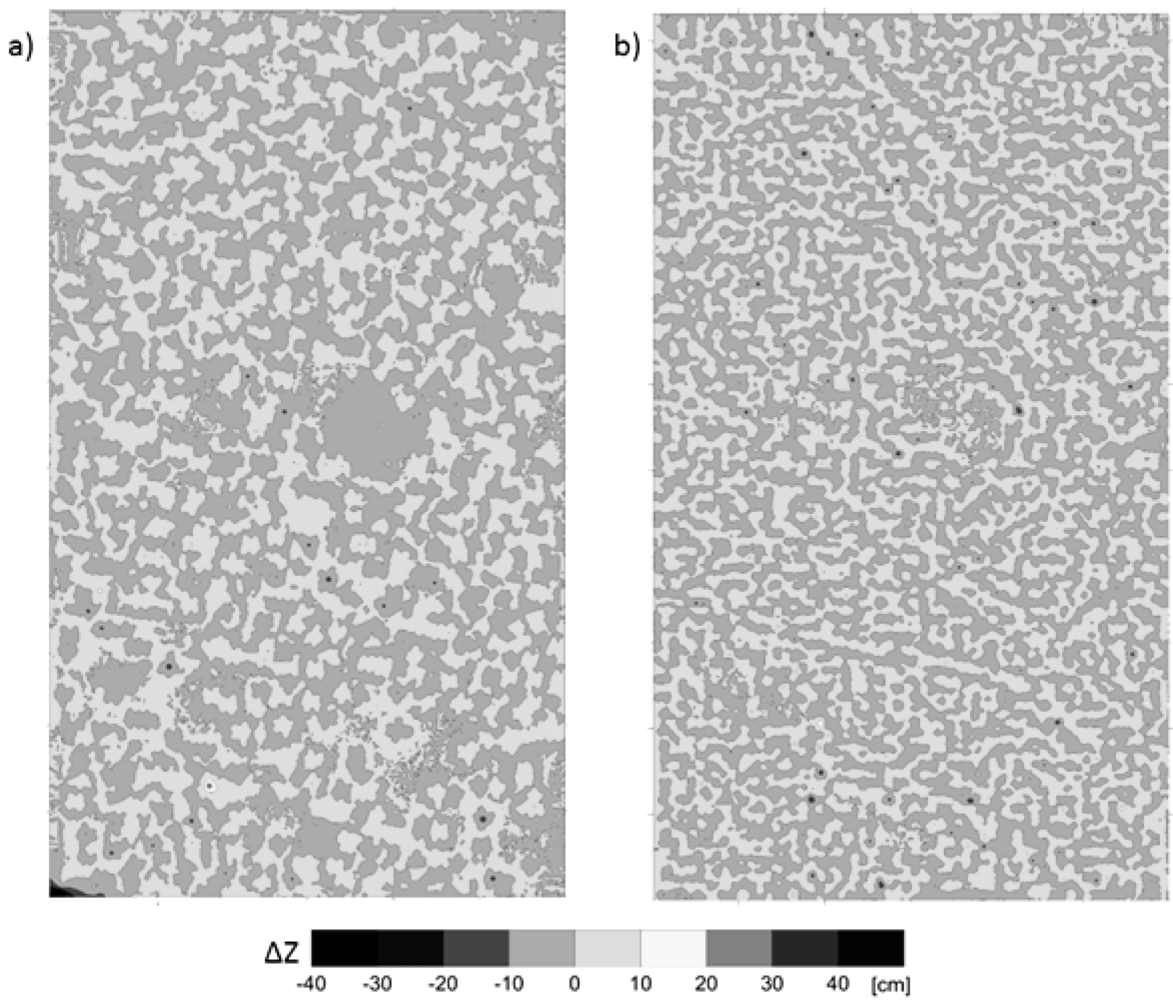

Fig. 6 DTMs differences a) DTM TSset – DTM Rset, b) DTM TSset – DTM Gset (source: own study in Surfer v.8).

deviation is 1.5 times smaller for the first variant. The coefficient of determination is higher by 1 % for reduction method. On this basis, it can be assumed that the DTM generated on the basis of the set decreased by reduction will be more fitted to the DTM generated from the original set (TS*set*).

Influence of reduction and generation on DTM generation was also tested by presenting spatial distribution of differences between DTMs (Fig. 6).

As it can be seen, in both variants DTMs created on the basis of decreased sets are generally characterized by the variability of the differences in height in the range of <-10.00cm, 10.00cm>. For the first variant those small differences create more compact areas, it is especially seen in the middle of Figure 6a. It resulted from the fact, that reduction leaves points which are characteristic for relief within selected area. Thus, reduction allows to preserved the relief a little better in comparison to generation method. For both variants point areas with the altitude differences of approx. 20.00 cm can also be observed, however much more point areas are visible in the second variant. It can be explained by local peaks which are eliminated during interpolation conducted within generation. The visible in Figure 6a in the left-bottom corner difference up to 40.00 cm can indicate that reduction sometimes removes significant points, particularly in close localisation of border of analyzed area.

5. CONCLUSIONS

The paper presents two variants of the decreasing the number of point clouds in order to effectively generate the DTM. Both variants can be used to reduce the size of a set. The reduced sets, obtained by using selected methods (for reduction - algorithm (Błaszczak et al., 2011a), for generating - kriging method) are different in number of points in the set and its spatial distribution. Also characteristic of these sets as well as DMTs generated based on them are different.

General conclusions from the presented analyses are as follows:

1. The reduction method always causes a set decreasing, a reduction degree is decided by the user by introducing appropriate criteria for reductions in the algorithm.

2. The method of GRID generation can cause a decrease or increase of the number of points in the point cloud. The user must select the size of the grid to generate a set with decreased number. User does not have a significant influence on the result as in the reduction method.
3. DTMs generated on the basis of a decreased sets (R*set* and G*set*) can be used for further studies. They do not depart from the DTM generated from a TS*set*. However, if there is need to have more detailed DTM, authors recommend to use of DTM R*set*.
4. The difference of DTM surfaces DTM TS*set* and DTM G*set* shows that DTM G*set* is smoother and the accuracy throughout the DTM is the same. Analyzing the difference of DTM surfaces DTM TS*set* and DTM R*set* it can be assumed, that the accuracy seems to be variable in different areas of the DTM R*set*. Although, in this study, the accuracy of the DTM R*set* is characterized by a value of RMSE and VAR, accuracy of this variability will be further analyze by authors.

Detailed conclusions of this study are:

1. Application of reduction method caused a decrease in the number of set by 1225 (14 %), while GRID generation - by 2339 (27 %).
2. DTMs generated from R*set* and G*set* are similar, however DTM R*set* is more detailed, while DTM G*set* is more smoothed.
3. RMSE is smaller for the R*set* and is equal 0.004 cm.
4. The difference in height between the DTM R*set* and DTM G*set* are in range <-10.00 cm, 10.00 cm>.

The results show that the decreased data sets can be used to build the DTM. Such sets allow for a more efficient and faster DTM generation. It has particular significance in the case of very big data sets.

REFERENCES

Axelsson, P.: 2000, DEM generation from laser scanner data using adaptive TIN models. International Archives of Photogrammetry and Remote Sensing, XXXIII/4B, Amsterdam.

Bauer-Marschallinger, B., Sabel, D. and Wagner, W.: 2014, Optimisation of global grids for high-resolution remote sensing data. Computers & Geosciences, 72, 84–93. DOI: 10.1016/j.cageo.2014.07.005

Błaszczak, W.: 2006, Optimization of large measurement results sets for building data base of spacial information system. Doctors thesis, University of Warmia and Mazury in Olsztyn.

Błaszczak-Bąk, W., Janowski, A., Kamiński, W. and Rapiński, J.: 2011a, Optimization algorithm and filtration using the adaptive TIN model at the stage of initial processing of the ALS point cloud. Canadian Journal of Remote Sensing, 37(6), 583–589. DOI: 10.5589/m12-001

Błaszczak-Bąk, W., Janowski, A., Kamiński, W. and Rapiński, J.: 2011b, ALS data filtration with fuzzy logic. Journal of Indian Society of Remote Sensing, 39, 591–597. DOI: 10.1007/s12524-011-0130-2

Chen, Y.: 2012, High performance computing for massive LiDAR data processing with optimized GPU parallel programming. The University of Texas at Dallas, 1519515, 77 pp.

Gościewski, D.: 2013, Selection of interpolation parameters depending on the location of measurement points. Giscience & Remote Sensing, 50(5), 515–526. DOI: 10.1080/15481603.2013.827369

Hebel, M. and Stilla, U.: 2008, The International Archives of the Photogrammetry and Remote Sensing and Spatial Information Sciences, XXXVII(B3a), 105–110.

Reitberger, J., Schnörr, C., Krzystek, P. and Stilla, U.: 2009, 3D segmentation of single trees exploiting full waveform LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing, 64(6), 561–574. DOI: 10.1016/j.isprsjprs.2009.04.002

Saeedi, S., Samadzadegan, F. and El-Sheimy, N.: 2009, Object extraction from Lidar data using an artificial swarm bee colony clustering algorithm. IAPRS, 37, 3/W4, 133–138.

Sithole, G. and Vosselman, G.: 2005, Filtering of airborne laser scanner data based on segmented point clouds. Geo-Information Science, 66–71.

Stateczny, A. and Łubczonek, J.: 2004, collective work, ed. A. Stateczny. Methods of comparative navigation . Gdańsk Scientific Society, Gdynia, 120–159, (in Polish).

Suchocki, C.: 2009, Application of terrestrial laser scanning in cliff shores monitoring. Rocznik Ochrona Srodowiska, 11, 715–725.

Suchocki, C., Damięcka-Suchocka, M., Błoch, P. and Stec, M.: 2013, Accuracy assessment of digital terrain model constructed from the direct data. Acta Sci. Pol., Geodesia et DescriptioTerrarum, 12 (3), 17–26, (in Polish).

Tovari, D. and Pfeifer, N.: 2005, Segmentation based robust interpolation – a new approach to laser data filtering. IAPRSSIS, 363W19, 79–84.

Visvalingam, M. and Whyatt, J.D.: 1992, Line generalization by repeated elimination of point. Cartographic Information Systems Research Group. University of Hull.

Vosselman, G.: 2008, Analysis of planimetric accuracy of airborne laser scanning surveys. The International Archives of the Photogrammetry and Remote Sensing and Spatial Information Sciences, XXXVII(B3a), 99–104.

Vosselman, G. and Maas, H.: 2010, Airborne and terrestrial laser scanning, Whittles Publishing.