Archives of Photogrammetry, Cartography and Remote Sensing, Vol. 22, 2011, pp. 51-61
ISSN 2083-2214

REDUCTION OF DTM OBTAINED FROM LIDAR DATA FOR FLOOD MODELING

Krzysztof Bakuła [1]

[1] Department of Photogrammetry Remote Sensing and Spatial Information Systems
Faculty of Geodesy and Cartography, Warsaw University of Technology
Pl. Politechniki 1, 00-661 Warszawa, Poland
k.bakula@gik.pw.edu.pl

KEY WORDS: DTM, generalization, LiDAR, algorithm, data structures

ABSTRACT: Recent years the cataclysm of flood has occurred in many regions around the world. For this reason, so much attention is focused on prediction of this cataclysm by creating flood risk maps and hydrodynamic – numerical simulation of flood water which are based on Digital Terrain Model (DTM). The modern techniques for automatic data acquisition provide very abundant amount of points. Actually, Light Detection and Ranging (LiDAR) is the most effective data source for DTM creation with density of one to few points per square meter and good height accuracy of less than 15 cm. This high redundancy of data is essential problem for algorithms used in programs for flood modeling. Many software generating such models are restricted with respect to the maximum number of points in DTM. Hundreds of thousands of points are too large number for complex calculations which describe fluid model of the flood water. In order to obtain reliable and accurate results, it is necessary to have DTM with an appropriate accuracy. The flood disaster also occurs in large areas what usually is associated with large data sets. However, it is possible to provide suitable DTM for flood modeling by its generalization without losing its accuracy, which could still ensure sufficient precision for hydrodynamic – numerical calculations. In this paper six reduction algorithms were tested to obtain DTM with small number of points and with accuracy comparable to the original model created from LiDAR data. The main criteria for this comparison was the relation between accuracy and reduction coefficient of final result. Methods used in this research were based on different DTM structures. GRID, TIN and hierarchical structures were compared in various approaches to obtain the most reduced and the most accurate terrain model of two study areas. As the result of the experiment the best methods for data reduction were chosen. Over 90% reduction rate and less than 20 cm root mean standard error were achieved in practice for different types of terrain with respect to input DTM. It was noted that hybrid and quad-tree grid based models can be even more efficient than a typical uniform GRID or TIN one.

1. INTRODUCTION

In recent years, lots of flood disasters have occurred in many regions around the world. In 2010 risk assessment of this disaster has became once again a major topic of highest interest in Central Europe, USA, China, Australia and many other countries where floods took place. For this reason, so much attention is focused on prediction of this global cataclysm by creating flood risk maps and hydrodynamic – numerical simulation of flood water. The most influential input for such product like Computational Fluid Dynamic (CFD), beside land cover, is the topography provided by Digital Terrain Model (DTM, Digital Elevation Model – DEM). This model must realistically describe terrain surface and be very accurate to secure concerned area from the risks and consequences of flooding.

The modern techniques for automatic data acquisition provide very abundant amount of points. Actually, Light Detection and Ranging (LiDAR) is the most effective data source for environment application. It can be used in 3D urban modeling, hydrological and glacier modeling, landform or soil classification, river bank or costal and forest management (Liu and Zhang, 2008). Especially LiDAR is the most appropriate data source for DTM creation in flood modeling, because of its high density of one to few points per square meter and good height accuracy of less than 15 – 20 cm (Cobby et al., 2001; Mandlburger and Briese, 2007). LiDAR data is also useful for other various river management application and it has started to replace other technologies (Brügelmann and Bollweg, 2004). Besides DTM creation, water level and river bed morphology measurements or wave amplitudes determination can be mentioned during discussion about influence of LiDAR data for flood modeling. Moreover, the fundamental for LiDAR data is the vegetation classification what has principal impact on roughness which is very important parameter for hydrodynamic simulations (Casas et al., 2010).

The redundancy of high resolution DTMs is essential problem for algorithms used in programs for the flood modeling. Many software generating such models are restricted with respect to the maximum number of points in DTM. Hundreds of thousands of points are too large number (usually less than 500 000 points) for complex calculations which describe fluid model of the flood water. In order to obtain reliable and accurate results, it is necessary to have DTM with approximated accuracy of 20 cm. The flood disaster also occurs in large areas what usually is associated with large data sets. However, it is possible to provide suitable DTM for flood modeling by its generalization, which could still ensure sufficient accuracy for hydrodynamic – numerical calculations.

Numerous works have been reported on DTM generalization, but the purpose is often very various. DTMs can be generalized for terrain analysis (Zhou and Chen, 2011) with retaining its main geographical characteristics (Ai and Li, 2010), visualization purposes (Zakšek and Podobnikar, 2005; Martín et al., 2009), flood and its risk areas modeling (Kraus, 2003; Haile and Rientjes, 2005) by using hydraulic models (Mandlburger and Briese, 2007; Mandlburger et al., 2008) or just to improve data processing efficiency in terms of both storage and processing time (Liu and Zhang, 2008). No matter why DTM is generalized it has to be done in intelligent approach to reduce data redundancy and keep the accuracy which original model presents.

There are many categorizations of methodologies which can be used in DTM generalization. Zhou and Chen (2011) propose five different groups of algorithms: 3D line generalization, filtering, point-additive, point-subtractive and feature-point methods. First method is widely used in cartography so it will not be a subject of this paper as others, which will be shortly described while some of them have been also used during this research. Generally speaking, filtering methods relate filtering techniques from the image processing. The most common methods like grid width unit downgrading (Haile and Rientjes, 2005) or low-pass filtering can be mentioned here (Casas et al., 2010). Unfortunately, these approaches cannot distinguish characteristic and uncharacteristic features (Zakšek and Podobnikar, 2005) so their results might be not precise enough for

DTM reduction in flood modeling for large areas. On the other hand they are the most common methods which almost every software is equipped with. It is a reason why their results will be shown in this paper. Another example of filtering methods is selective filtering which preserves characteristic points. This group of methods uses a moving window which evaluates the importance of each central point by its neighborhood analysis. This approach will be used in the research described in this paper, because it removes significant amount of point in flat areas and retains points where height difference is observed. The point-additive method is an iterative method which starts with the minimum initial approximation. Each iteration adds points which have the maximum variation to the surface defined by TIN in previous iterations until the threshold is reached. This approach can be implemented by using hierarchical or hybrid structure (Mandlburger and Briese, 2007). The point-subtractive method (decimation) starts triangulation of all points and iteratively drop them until threshold is reached. The last group of methods in this categorization is feature-point method which selects characteristic, important points (e.g. peaks, pits, points of valleys and saddles) (Zakšek and Podobnikar, 2005; Zhou and Chen, 2011) and uses them for TIN or hybrid DTM generation. Basically, each of mentioned methods used for DTM generalization has advantages and disadvantages which depend on its application. What is important, the method significantly reducing number of points in flatter areas and retaining important points from river surroundings for flood modeling is required. In addition, there are many methods of DTM generalization that do not reduce DTM itself, but original LiDAR data (Liu and Zhang, 2008) which DTM is generated from. It is wide range of knowledge about data obtained from aerial scanning systems but it is not analyzed in this paper.

As discussed before, generalization of DTMs should meet all demands of its later application. For hydrologist analysis special requirements like accuracy and relatively small data set must be respected. Mandlburger and Briese (2007) mention additionally that the product like Computational Fluid Dynamic (CFD) hydraulic model is associated with DTM in unstructured geometries i.e. a computation grid based on irregularly distributed points. Two appropriate data structures, therefore, can be proposed: hierarchical and irregular to reduce number of points and maintain required accuracy. Hierarchical division is based-on quadtree-like data structure – if grid cell does not meet maximum height tolerance, it is divided into four parts in each pass. The most common irregular structure is triangular irregular network (TIN) (Lee, 1991; Ai and Li, 2010; Zhou and Chen, 2011), but there is more sufficient structure applied in DTM reduction. The hybrid structure is very appropriate in this case. It is based on a regular grid and intermeshed points defining breaklines and hot spots, what has been used in works of many authors involved in this area (Kraus, 2003; Mandlburger and Briese, 2007; Mandlburger et al., 2008; Martín et al., 2009).

The aim of this paper, therefore, is to derive a high quality DTM with much reduced points number for study areas. The main purpose is to decrease data set in flat parts of analyzed terrain and retain important points representing rough areas and particularly being part of river surrounding (river bank, river bed, flood embankments etc.) what has a major impact on flood modeling. The following parts of this paper present methodology, description and results of the experiment. The evaluation of all methods and proposal of the most respective parameters selection providing high quality products for flood modeling will be discussed and summarizing conclusions will be drawn in final part of this paper.

2. METHODOLOGY

2.1 Data

The research was based on high spatial resolution (1 m GRID width) DTMs that present different type of terrain as shown in Fig. 1. Each of them covers 9 square kilometers (3 kilometers by 3 kilometers). The first test area in this study is an urban terrain of the city of Wrocław in Lower Silesia, Poland, which is situated on the Oder River. Maximum height difference for this case study is 10 meters. The town was flooded during the cataclysm in 1997 and 2010. DTM of this area was created on basis of data from OPTECH ALTM 2050 system obtained in 2006 with density 3-4 points per square meter. The primary purpose of that LiDAR data collection was the creation of color orthophotomap, DSM and DTM for the city of Wrocław.

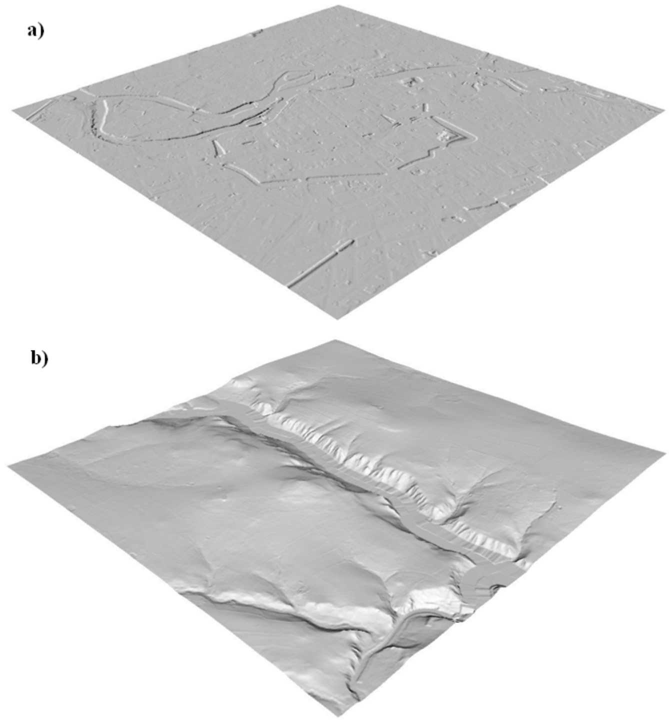

Fig. 1. DTM of two study areas: urban terrain in Wrocław (a) and hilly river valley in Lubań District (b)

Second test area is part of Lubań District which is located in the Kwisa river valley, Lower Silesia, Poland. This river has caused numerous flooding because of intensive rainfalls in 2010. This study site features very hilly terrain with maximum height difference equals

100 meters. DTM of this area was generated from data obtained by OPTECH ALTM 3100 in 2009 with density of 4-5 points per square meter. The purpose of data collection was the creation of national computer system which can protect the society, the economy and the environment against extraordinary threats.

2.2 Software used

General programming tools and available Inpho software were used to reduce DTMs for this study. The LiDAR data were classified into terrain and non-terrain points by using data filter algorithms in DTM Master software. The segmentation of LiDAR data is a crucial prerequisite for the modeling of surface objects (Cobby et al., 2001) so it was decided to use well-known and professional software as Inpho product. After data segmentation, GRID DTMs were created from ASCII files and then they were generalized in various software using the algorithms described in Subsection 2.3. Subsequently, models in GRID structures were interpolated from reduced DTMs to compare with input model in statistic and visual analysis in ArcGIS 10 by ESRI. To assess reduction methods for each created differential model (i.e. model of height differences spatial distribution) root mean standard error and reduction coefficient were calculated.

2.3 Methods

In this paper six different generalization algorithms were used to obtain DTM with small number of points and with accuracy comparable to the original model created from LiDAR data. Each approach reduced input data points in a different way which can be seen in Figure 2. The main criteria for this comparison was the relation between accuracy (described by: root mean standard error of generalized DTMs with respect to input one) and reduction rate of final result.

First method for data reduction was grid width unit downgrading. The original DTMs were resampled, thus generalized DTMs of smaller resolution were created. This method shows errors caused by discontinuous representation of the terrain surface. The bigger grid size is selected, the more approximate shape of terrain is noticed. Haile and Rientjes (2005) proved in their investigation that DTM resolution has significant impact on flood simulation results. It was observed that inundation extent, flow velocity, depth and patterns across the model domain were affected by significant errors of resampled DTM. In spite of fact that this approach is widespread, simple and fast, it does not give a sufficient approximation result. Nonetheless, in this paper, the described method is used as a reference for further five approaches to show better their efficiency in comparison with grid resampling.

Next two algorithms were used with the available ESRI software and they can be qualified as feature-point methods. Both of them are applicable especially in the generalization of terrain models in large scale studies in cartography, but in this research, their efficiency for dense DTM was investigated. Chen and Guevara (1987) described a method to select 'very important points' (VIP) from GRID DTMs to generate TIN models. This approach from ArcToolbox (ArcGIS 10) assesses the significance using 3x3 window to calculate how well analyzed point is approximated by its eight grid neighbors. In result, the extracted TIN model should contain as much information about the terrain surface as it is possible. Second method implemented in ArcGIS was Kernel MinMax algorithm choosing minimum and maximum height in each window which can be modified by its size.

Another two approaches are based on hybrid models (regular grid plus additional points describing breaklines, structure lines, spot heights etc). Author's hybrid DTM creation method determined these additional points by using ΔZ values between eight neighboring cells in four directions around analyzed point. In case when mentioned ΔZ values had been less than declared in reduction process, the point was removed.

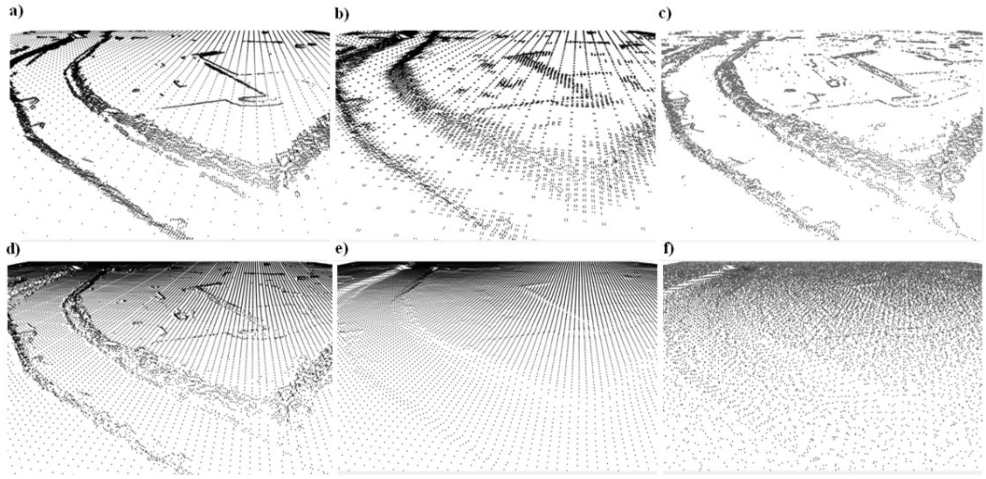

Fig. 2. Points distribution of generalized DTM for a fragment of Wrocław case study with use of: author’s hybrid algorithm (95% reduction, RMSE – 0.14m) (a), author’s hierarchic approach (92%, 0.14m) (b), VIP (95%, 0.24m) (c), DTM Master hybrid (92%, 0.09m) (d), Grid downgrading (94%, 0.20m) (e), Kernel MinMax (92%, 0.21m) (f)

In the second method two parameters can be set – XY and Z distance – for data thin out in DTM Master 5.2.1 (Inpho software). A point was deleted, if the distance to the nearest point was larger than XY distance value or if the difference between the height of that point and the height interpolated from its neighbors was larger than this value.

The last method applied in DTM reduction was non-uniform quad-tree data structure which is adaptive hierarchical grid system. To produce quad-tree grid model, ΔZ value was verified for all cells in 16x16 window. If the height difference in that part of area was higher that ΔZ declared, initial window was subdivided into four smaller windows. Otherwise, average height was defined for four corner cells of analyzed window. The process was terminated on 2x2 windows and its advantage is giving a different cell size depending on terrain what is effective way to reduce large data sets.

To compare the methods described above, reduced models were interpolated into grid DTM once again to have possibility for spatial distribution of height differences presentation. For each approach the reduction coefficient was calculated as percentage of point removed from input data while, on the basis of differential models generated, root mean standard error was estimated.

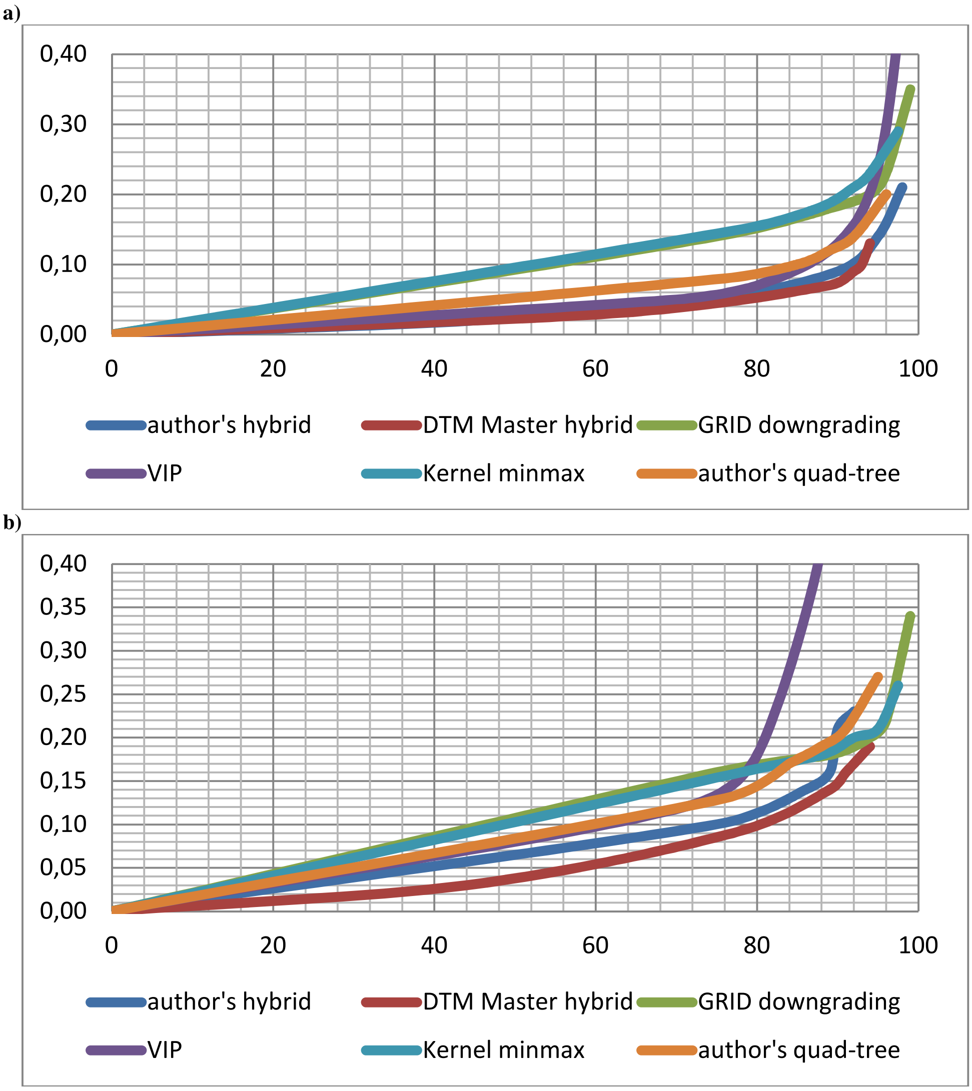

Fig. 3. Comparison of RMSE [m] versus reduction rate [%] between 6 algorithms used in research for urban terrain of Wrocław case study (a) and hilly valley of Kwisa river in Lubań District (b)

3. RESULTS

Experiment was conducted on two different type of case studies. Few generalized DTMs from input were made for each method described in Subsection 2.3. The results of the experiment are presented in Table 1, while in Figure 3 they are shown in comparison of

RMSE versus reduction rate. Not many terrain models were generalized with low reduction rate because it would be useless for flood modeling (high reduction needed). The results show that there is no significant decrease in accuracy for the reduction rate lower than 80%. For higher values of that number sudden rapid increase of RMSE can be observed, which depends on the method used and the type of terrain analyzed. Moreover, only some of presented approaches allow to achieve a reduction rate more than 90% with less than 20 cm RMSE.

Tab.1. The result of DTM reduction with using six methods for two case studies in research

Approach		**Wrocław case study**					**Lubań case study**				
author's hybrid	Δx,y [m]	10.0					5.0				
	Δz [m]	0.1	0.2	0.5	1.0	2.0	1.0	1.5	1.7	1.8	2.0
	reduction [%]	45	74	90	95	98	75	86	89	90	92
	RMSE [m]	0.02	0.05	0.09	0.14	0.21	0.10	0.14	0.16	0.21	0.23
author's quad-tree	Δz [m]	0.3	0.5	0.7	1.0	2.0	1.0	1.5	2.0	3.0	5.0
	reduction [%]	76	85	89	92	96	76	84	88	91	95
	RMSE [m]	0.08	0.10	0.12	0.14	0.20	0.13	0.21	0.23	0.29	0.31
DTM Master hybrid	XY distance [m]	10.0					10.0				
	Z distance	0.05	0.2	0.4	0.5	1.0	0.05	0.2	0.4	0.5	1.0
	reduction [%]	62	89	92	92	93	44	77	89	91	94
	RMSE [m]	0.03	0.07	0.09	0.10	0.13	0.03	0.09	0.14	0.16	0.19
Grid unit down-grading	unit width [m]	2.0	3.0	4.0	5.0	10.0	2.0	3.0	4.0	5.0	10.0
	reduction [%]	75	89	94	96	99	75	89	94	96	99
	RMSE [m]	0.14	0.18	0.20	0.23	0.35	0.16	0.18	0.20	0.22	0.34
VIP	Δx,y [m]	10.0					10.0				
	significant ratio	30.0	20.0	10.0	5.0	2.0	30.0	20.0	10.0	5.0	2.0
	reduction [%]	70	80	90	95	98	71	81	91	96	99
	RMSE [m]	0.05	0.07	0.13	0.24	0.47	0.12	0.20	0.58	1.26	2.53
Kernel minmax	window size [pixels]	3.0	4.0	5.0	6.0	9.0	3.0	4.0	5.0	6.0	9.0
	reduction	78	88	92	94	97	78	88	92	95	97
	RMSE [m]	0.15	0.18	0.21	0.23	0.29	0.16	0.18	0.20	0.21	0.26

4. DISCUSSION

Each methods used in this research has its advantages and disadvantages. The study does not exclude any of tested methods and can lead to improve them in more compound algorithms. DTM resolution downgrading has impact on flood simulation what was proved by Haile and Rientjes (2005). Grid width unit downgrading and Kernel MinMax algorithm similarly reduced input DTMs giving comparable RMSE in result. The problem is that these methods do not retain characteristics of the terrain topography and the location of points forming digital terrain model is so regularly determined (Figures 2e and 2f) that the higher is the complexity of the analyzed area, the more likely is that important features for flood modeling like embankments, sloped river banks, etc. can be skipped in reduction. Figure 4d shows spatial distribution of height difference in Kwisa valley study for grid downgrading method where the highest values of RMSE can be noticed on sloped terrain. Additionally, the same size of RMSE in similar reduction rate was observed in both case studies for these approaches. This fact leads to the conclusion that the influence of the terrain slope is less than in other methods used.

The algorithm of data thin out, used in DTM Master software, reduces input model in the most effective way, as shown in Figure 3. Generalized DTMs represent the overall terrain surface morphology well but also here, higher height difference can be noticed in sloped part of river surrounding (Figure 4d). What is crucial to efficiency of this algorithm, distribution of points determining hybrid DTM is still dense in flat parts of terrain and important points are preserved. . This can be seen in Figure 2, which shows a comparison of the point distribution after DTM reduction achieved with the six proposed methods.

The worst results were achieved for VIP algorithm in both case studies because of the lack of limit for length of triangles sides (Figure 3 and 4f). Moreover, this approach can rely heavily on algorithm which might not identify some significant points on slopes (Zhou and Chen, 2011) and can cause some systematic height differences . This method is therefore, more appropriate for DTM generalization without losing its accuracy in large scale studies in cartography, where creation of small-sized skeleton model is the most important purpose in the reduction. The compound approach with use of hybrid structure can be proposed as basis on VIP algorithm for additional points detection, what will be investigated in future research work.

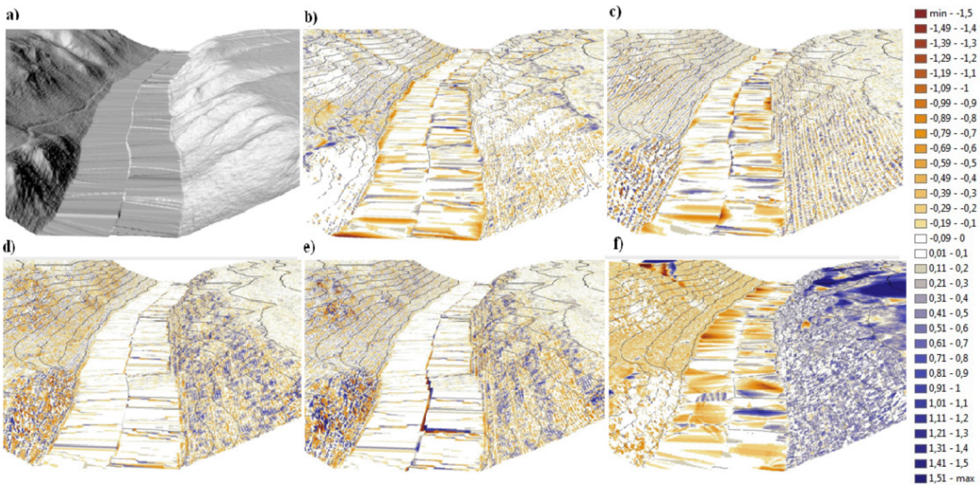

Fig. 4. DTM of Kwisa valley (a) and height differences spatial distribution of generalized DTM in respect to original one with use of: author's hybrid method (86% reduction, RMSE = 0.14m) (b), author's quad-tree approach (88%, 0.23m) (c), DTM Master hybrid (94% reduction, 0.19m) (d), grid downgrading algorithm (94%, 0.20m) (e), VIP approach (91%, 1.35m) (f)

As mentioned before, a the method which significantly reduces number of points in flat areas and retains important points from river surroundings is required for flood modeling. The two of author's methods have met this condition as shown in Figures 2a and 2b. In contrast to other approaches, they slightly generalized the sloped terrain as indicated by the lower reduction rate for the second test study (Lubań). In this way, river surrounding was retained by having low height difference on sloped river banks (Figures 4b and 4c), what is necessary for the application of DTMs in presented kind of modeling. On the other hand, river bed and flatter parts of analyzed area are affected by higher height difference between generalized DTM and input one. This problem should be solved by improved algorithms because of its possible impact on the river flow modeling.

Further work related to the topic of the paper will focus on creating more compound algorithms for DTM generalization, testing them on different, flood hazard case studies. Especially, with a use of reduced terrain models in software producing flood risk maps and hydrodynamic simulation of flood water, which could give an answer about impact of the accuracy of the product obtained during reduction on inundation extent, velocity and depth of river flow.

5. CONCLUSSION

LiDAR is the most appropriate data source for DTM creation and other various application in flood modeling. Its high density, however, is essential problem for software which produce flood risk maps and hydrodynamic simulation of flood water. It is desired to reduce data amount without losing information about terrain characteristic. As the result of experiment, over 90% reduction rate and less than 20 cm root mean standard error were achieved for different types of terrain. DTM reduction achieves better results for flatter area than for more hilly study area. It was noted that hybrid and quad-tree grid based models can be even more efficient than a typical uniform GRID or TIN. To what extent DTM can be generalized (to what high reduction rate terrain model can be generalized to meet an appropriate height accuracy) depends mainly on the terrain type and input data density. It is also not recommended to set reduction parameters fixed in chosen method for any type of terrain, because such operation should base on individual approach to each of the analyzed area. It is suggested that along with the process of generalization, where reduction coefficient is calculated, root mean standard error of height for the whole area of interest should be also computed in order to determine the correctness of reduction parameters selection.

6. REFERENCES

Ai T., Li J., 2010. A DEM generalization by minor valley branch detection and grid filling. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, pp. 198-207.

Brügelmann R., Bollweg A.E., 2004. Laser altimetry for river management. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXV, B2, Istanbul, Turkey, pp. 234 – 239

Casas A., Lane S. N., Yu D., Benito G., 2010. A method for parameterising roughness and topographic sub-grid scale effects in hydraulic modelling from LiDAR data. In: *Hydrology and Earth System Sciences (HESS)*, Vol. 7, pp. 2261-2299.

Chen Z. T., Guevara J. A., 1987. Systematic selection of very important point (VIP) from digital terrain model for constructing triangular irregular networks. In: *Proceedings of International Symposium on Computer-Assisted Cartography (AUTO-CARTO 8)*, Baltimore, U.S.A., pp. 50-56.

Cobby D.M., Mason D.C., Davenport I.J., 2001. Image processing of airborne scanning laser altimetry data for improved river flood modelling. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, 56, pp. 121-138.

Haile A.T., Rientjes T.H.M., 2005. Effects of LiDAR resolution in flood modeling: A model sentitivity study for the city of Tegucigalpa, Honduras. ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005", Enschede, the Netherlands, September 12-14, 2005

Kraus K., 2003. LaserScanDTMs for Modeling Flood Risk Areas. Photogrammetric Week 2003. Stuttgart, Germany. http://www.ifp.uni-stuttgart.de/publications/phowo03/kraus.pdf

Lee J., 1991. Comparison of existing methods for building triangular irregular network models of terrain from grid digital elevation models. In: International Journal of Geographical Information, Vol. 5, No. 3, pp. 267-285.

Liu X., Zhang Z., 2008. LiDAR data reduction for efficient and high quality DEM generation. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, Vol. XXXVII, Part B3b, pp. 173-178.

Mandlburger, G., C. Briese, 2007. Using Airborne Laser Scanning for Improved Hydraulic Models. In: International Congress on Modeling and Simulation, Christchurch, New Zealand.
http://www.mssanz.org.au/MODSIM07/papers/12_s27/UsingAirbone_s27_Mandlburger_.pdf

Mandlburger G., Hauer C., Höfle B., Habersack H., Pfeifer N., 2008. Optimisation of LiDAR derived terrain models for river flow modelling. In: Hydrology and Earth System Sciences (HESS), Vol. 5, pp. 3605-3638.

Martín M. T., Rodríguez J., Irigoyen J., Martínez-Llario J. C., Arias P., 2009. Semi-automatic Process for Hybrid DTM Generalization based on Structural Elements Multi-analysis. In: *The Cartographic Journal*, Vol. 46, pp. 146-154(9)

Zakšek K., Podobnikar T., 2005. An Effective DEM generalization with basic GIS operations. In: 8th ICA workshop on Generalisation and Multiple Representation, Coruńa, Spain. http://ica.ign.fr/Acoruna/Papers/Zaksek_Podobnikar.pdf

Zhou Q., Chen Y, 2011. Generalization of DEM for terrain analysis using a compound method. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, pp.38-45.