# **EFFECTS OF DEM GENERATING ALGORITHMS ON WATER RETENTION CALCULATIONS IN POLDERS – A CASE STUDY**

# N. Sz., Suba <sup>a, \*</sup>, S. Nistor <sup>a</sup>, St., Suba <sup>a</sup>

<sup>a</sup> University of Oradea, Faculty of Constructions, Cadastre and Architecture, Department of Cadastre and Architecture, 4 B. St. Delavrancea, 410058, Oradea, Romania

Received: 02.10.2017 / Accepted: 15.10.2017 / Revised: 27.11.2017 / Available online: 15.12.2017

#### DOI: 10.1515/jaes-2017-0015

KEY WORDS: dem algorithm, grid, volumetric calculations, retention volume estimation, data interpolation

## **ABSTRACT:**

Digital Elevation Models have been an important topic in the last decades in order to accurately describe one's surroundings for various purposes. The most wide-spread applications of Digital Elevation Modelling are related to volumetric calculations over large areas, as well as to hydrographic simulations. This paper will analyse the impact of various data interpolation models on the result of volumetric calculations in a longitudinal shaped polder. In order to achieve general conclusions instead of case specific solutions, the raw parametric Digital Elevation Models have been used, rather than the specific parameters suiting the actual scenario. In order to achieve this, data smoothing has also been neglected. The result of the conducted study presents the impact of the digital elevation modelling methods over the volumetric calculations. When applied to the same dataset, the methods yielded different results, partially confirming the predicted usefulness of the algorithms.

# 1. INTRODUCTION

Digital elevation models (DEMs) have been an important topic in geography and surveying sciences for decades due to their geomorphological importance as the reference surface for gravitation-driven material flow, as well as the wide range of uses and applications. (Oksanen, 2006). The most wide-spread applications of Digital Elevation Modelling are related to volumetric calculations in hydrography, mining, but they can also be used for various geographic, surveying and other applications. In order to achieve this, the correct modelling must be carried out, having in mind some decisive factors.

In the process of modelling the surface of a specific area, both the data acquisition and the data interpolation methods must be taken into account. Elevation models have been around for some time. Maps with level curves / isohypses were extensively used in the past for height data or terrain model visualisation. While the accuracy of these maps was largely dependent on the surveying methods used to acquire the necessary data, their use (and again, accuracy) was also limited by the scale at which the map was drawn. In the case of Digital Elevation Models, we can define the resolution of the model rather than the scale as being the influential factor when speaking about the possible accuracy. But is this the only parameter defining it?

Generally speaking, the purpose of the DEM will define the data acquisition method, as well as the data density. The aim is to obtain a dataset, which can de described with the unambiguous formula:

$$z = f(x, y) \tag{1}$$

which means that a certain point defined by X and Y coordinates can have a single Z (height) value. This is a fundamental difference when compared to CAD software modelling.

Data acquisition – to a certain extent – can be done using existing topographic maps with level curves. But, in addition to the positional accuracy problems described above, we must also take into account the temporary accuracy of these maps. The precision of maps is not necessarily expressed in terms of pinpoint positional accuracy of the details contained. Outdated or misleading/incomplete information can be as harmful as mispositioned details (Suba, 2015).

Current raw data sources for creating digital elevation models include:

- photogrammetry
- total station / GNSS measurements
- LIDAR imagery
- SAR imagery
- bathymetric measurements

Regardless of the above methods used, these are raw field data. When they go through different correction processes (coordinate transformations, systematic error correction, random data filtering, data compression etc.), the reference data is obtained. But, technically, the reference data only consists of discretely collected field points in a 3D coordinate system.

Presently, discretely collected height points have to be interpolated to form curved faces, the selection of spatial interpolation methods decide the quality, accuracy and follow-up analysis applications. Interpolation methods are used here to

Corresponding author. e-mail: suba\_norbert@yahoo.com. Tel.: +40 740 59 00 98 (N.Sz. Suba)

calculate the unknown heights of interested points by referring to the elevation information of neighbouring points. There are a great many commercial interpolation software, however, most of them are designed to solve specific problems with limited versatility (Yang, 2004).

Interpolation techniques are based on the principles of spatial autocorrelation, which assumes that closest points are more similar compared to further ones (Patel, 2016). The main goal of developing a GRID is to establish a surface composed of regularly distributed XYZ points in such a way, that the input points would be contained within (Litwin, 2013).

Topographic analysis of watershed-scale soil and hydrological processes using digital elevation models is commonplace, but most studies have used DEMs of 10 m resolution or coarser. Although availability of higher-resolution DEMs created from light detection and ranging (LIDAR) data is increasing (Gillin, 2015), other studies have pointed out that groundwater flow directions typically follow the general topography of the landscape rather than small-scale surface variations, and thus smoother topography represented in coarser resolutions may better represent near-surface flow pathways and water table positions (Thomas, 2017).

As different interpolation methods applied over the same data sources may result in different results (Arun, 2013), we conducted a study on a specific dataset in order to address the impact on the resulting volume calculations based on coarse, unsmoothed DEM data. This dataset was provided by total station and GNSS equipment data acquisition, as vertical accuracy of all the space borne DEMs is influenced by relief of the terrain as well as type and height of vegetation (Jobin, 2014). Although studies suggest that certain methods are more suitable than others for such applications (Dumitru, 2013), different scenarios can result in different approaches, and with the adopted mathematical procedures, each method may impose particular landscapes after correction and flat area treatment (Fernandez, 2016). Also, there is a chance that the predicted spatial pattern of surface runoff is strongly affected by the presence of interpolation artefacts in the DEM, with completely unrealistic predictions in the case of the worst quality DEMs (Wise, 2007).

### 2. MATERIALS AND METHODS

### 2.1 Interpolation methods

### Inverse Distance Weighting (IDW)

The IDW is a relatively simple weighted average interpolation method used to describe general tendencies defining a certain phenomenon rather than resulting in a rigorous analysis. The method is based on the presumption that the influence of one point on another is decreasing with the distance increasing between them, and it's best suited when the dataset consists of high density points, with approximately equal distance between them. The obtained model is the result of each point in the dataset, with closer distributed points having a bigger impact on the final result. When modelling certain areas that have hills and valleys, the accuracy tends to suffer because of this approach. Also, this method is known to create the so-called" bulls eye" areas, where certain areas are mistakenly created as elevations from the dataset.

#### Kriging

Kriging is a geostatic gridding method, based on linear algorithmic estimations. The method is similar to the IDW method, with the weight factor based on probabilities rather than distances. When applied correctly, Kriging allows users to derivate weight factors that result from optimal and unbiased estimations. The method will try to minimise the variation of errors and to set the prediction mean error to a zero value, in order to avoid sub- or overestimations. In the calculation process a semi-variogram of the data is created, which is used in the weighting process of the sampled date. This option makes the method a rather versatile one, because the default parameters may be accepted to produce an accurate grid of the source data; alternatively, Kriging can be custom-fit to a data set by specifying an appropriate variogram model.

A rather unique characteristic of the method is that it offers an error estimation for each point created by interpolation and thus a certain confidence level in the resulting DEM, making it rather a statistical modelling technique rather than a deterministic one.

# Natural Neighbour

Natural Neighbour is a geometrical estimation method, which uses naturally created neighbouring areas around each point of the dataset. It is considered to be a useful method for not uniformly distributed dataset points, and it can be widely used since it lacks the necessity to define certain parameters for the estimation (radius, number of neighbouring points, weight factor etc).

The method is developed to honour minimum and maximum local values, and thus, it cannot be used to extrapolate data outside the dataset. The gridding method uses a weighted average of the neighbouring observations and generates good contours from data sets containing dense data in some areas and sparse data in other areas.

### Nearest Neighbour

This calculation method relies on assigning the height value of the nearest dataset point to the interpolated grid point, by using Voronoi diagrams to determine the nearest neighbour. The method is considered to be useful either when the points in the dataset are evenly distributed or when there's a need to fill up a few missing areas in regular datasets.

### Minimum Curvature

Minimum Curvature is widely used in earth sciences. The interpolated surface generated by Minimum Curvature is analogous to a thin, linearly elastic plate passing through each of the data values, with a minimum amount of bending. Minimum Curvature generates the smoothest possible surface while attempting to honour your data as closely as possible. Minimum Curvature is not an exact interpolator, however. This means that your data are not always honoured exactly.

# 2.2 DEM generating principles and volumetric calculations

After defining the DEM creation algorithms for the datasets, it is necessary to define a volume calculation method as well, in order to properly compare the results obtained on different DEMs. Additionally, correct level curve determination is also worth comparing.

Raw data filtering is important before the DEM generating process. Usually, we can obtain data which is not specific to an extended surrounding, and it can negatively influence the DEM calculation, by introducing non-characteristic entities. Such data can consist of height values determined on the top of vegetation or crops rather than on ground level (when using remote sensing data or airborne photogrammetry) or certain measured entities with total station / GNSS equipment that are useful for the overall study of the polder, but may induce certain errors in DEM generation (such as building heights in the area, bottom of manholes, and generally any non-characteristic element that will barely affect the water flow or volume calculation, but it can affect the DEM calculation method). Since different algorithms tend to evaluate these entities in different ways, it is recommended that a preliminary filtering is made.

For each measured dataset (in nature representing polders with dams), after filtering and DEM generation, the minimum and maximum height value must be determined for volume calculations. Generally, the lowest height value is the one close to the natural or artificial drainage of the polder, while the maximum value is given by the polder mean dam height (any values above this will result in a water overflow, so further volume calculations are pointless).

In order to study both of these aspects, we have opted for a simple volume calculation method rather than the automatic volume calculation capabilities offered by certain software. After generating the grids with the methods presented above (resolution of 10 meters), and after obtaining the characteristic level curves (minimum and maximum for each scenario, as well as 0.5 meter curves between them), the following formula allows us to estimate the volume between two successive level curves:

$$v_{i-i+1} = \frac{(S_i + S_{i+1})}{2} \cdot E$$
 (2)

where:

 $V_{i-i+1}$  represents the volume between two level curves

 $S_i, S_{i+1}$  the area of the level curves

*E* equidistance / distance between the curves

The total water retention volume is calculated as the sum of the individual volumes between the neighbouring level curves:

$$V = \sum_{i}^{i+1} v \tag{3}$$

### 2.3 Case study – longitudinal retention basin

In the case study, field measurements have been made at the Velju Mare polder, situated in the vicinity of the village of Inand, Bihor County, Romania. Measurements have been made in order to reach the polder dam mean height value on the field for correct volume calculations.

On-the-field measurements were conducted in such a way that the dataset would have a rather coarse, but even density, with sample points being taken on an approximately 50 m x 50 m grid, with additional points being measured at certain interest points (characteristic areas of slope changes etc.). Preliminary DEM generation and level curve calculations resulted in a hydrographic basin that is over 1.5 kilometres long, but at most 190 meters wide, so it could provide a good scenario to study the DEM creation models in this type of terrain configuration. Our aim is to generate a GRID with different interpolation methods, which can be used to manage and visualize the resulting DEM.



Figure 1. Extrapolated DEM with field survey data

In the follow-up, the 10 meter resolution grid was generated with Golden Software's Surfer software for each of the methods described above, containing the field survey data. Level curves were generated and their areas have been calculated using Geotop Romania's MapSys software package. The visual representation of the resulting DEMs can be observed in Figures 2 to 6 (with the red line representing the level curve of 111 meters above sea level, the mean value for the polder dam), while the minimum height above sea level was 106 meters (with little exceptions, see table 9):



Figure 2. DEM generated with IDW interpolation



Figure 3. DEM generated with Ordinary Kriging interpolation



Figure 4. DEM generated with Natural Neighbour interpolation



Figure 5. DEM generated with Nearest Neighbour interpolation



Figure 6. DEM generated with Minimum Curvature interpolation

As it has been mentioned before, no smoothing operations have been done on the resulting DEMs. Visual inspection of the resulted elevation models indicate that in our case, the general principles regarding the modelling algorithms have been confirmed.

The unsmoothed IDW generated the predicted bull's eye regions, creating hills and valleys that are not specific to the terrain configuration, so further data processing or volume calculations have not been made upon this model. Kriging and Natural Neighbour interpolation methods seems to have worked well. The Nearest Neighbour method (as expected) didn't manage to follow the natural landscape, but has generated additional retention areas and an additional surface beyond the minimum expected height, whereas the Minimum Curvature method, although similar to the results of Kriging, was unable to match the height values of the dataset points in certain areas.

Tables 7 to 10 summarize both the area and the volume calculations for each DEM generation algorithm that was used. The *volume w/prev* column contains data about the water retention volume calculated with the previous level curve area, rather than containing the cumulated volume until that level curve (total volumes are calculated at the end of each table).

Fable 7. Surfaces and volumes using Krigin	ıg	
--	----	--

Curve height (m)	Area (sq. m.)	Volume
Curve neight (III)		w/prev.(cu. m.)
111	197462	90765.5
110.5	165600	76250
110	139400	63693.25
109.5	115373	49909.5
109	84265	36624.75
108.5	62234	26405.5
108	43388	16851.25
107.5	24017	8484.5
107	9921	3315.5
106.5	3341	880
106	179	0
Total vo	lume	373189.75

Table 8. Surfaces and volumes using Natural Neighbour

Curve height (m)	Area (sq. m.)	Volume w/prev. (cu. m.)
111	193034	88649
110.5	161562	74762
110	137486	61118.5
109.5	106988	46724.5
109	79910	34526
108.5	58194	24345
108	39186	15105.5
107.5	21236	7238.75
107	7719	2253.75
106.5	1296	326.25
106	9	0
Total vo	lume	355049.25

Curve height (m)	Area (sq. m.)	Volume w/prev. (cu. m.)
111	332298	148113.5
110.5	260156	105962.75
110	163695	73132.25
109.5	128834	56921.25
109	98851	44610.75
108.5	79592	31306.5
108	45634	18802
107.5	29574	11295.25
107	15607	5341
106.5	5757	1761.25
106	1288	447.5
105.5	502	0
Total ve	olume	497694

Table 9. Surfaces and volumes using Nearest Neighbour

Table 10. Surfaces and volumes using Minimum Curvature

Curve height (m)	Area (sq. m.)	Volume w/prev. (cu. m.)
111	198777	91636.75
110.5	167770	76458.75
110	138065	62952.5
109.5	113745	50150.25
109	86856	38434.75
108.5	66883	28047.25
108	45306	17955
107.5	26514	9720.75
107	12369	4313
106.5	4883	1323.25
106	410	0
Total vo	lume	380992.25





Figure 11. Volume comparison between the used models

As the above tables and figure clearly shows, the data processing have confirmed what visual inspection suggested - that Kriging and Minimum Curvature have generated similar models, Natural Neighbour resulted in a narrower approximation of the hydrographic basin, while Nearest Neighbour clearly overshot the estimation, resulting 110,000 cubic meters' bigger volume estimation than the next greatest total volume (more than 25% of the value).



Figure 12. Surface area of all the models at different height values

Figure 12 presents the data contained in Tables 7 to 10, regarding the surface area differences obtained at various height values with each of the interpolation methods. As the graph suggests, there has been a significant difference between Nearest Neighbour and the rest of the methods around the height value of 108.5 meters above sea level, then the results follow the tendency of the other three methods until the height value of 110 meters above sea level, where another jump occurred in that specific level curve's area determination.



Figure 13. Cumulated retention volumes at different height values



Figure 14. Volumes at different height values, calculated only with previous level curve

Upon studying the graphical representation of the cumulated water retention volumes at different height values (Figure 13), as well the volumes at different height values, calculated between adjacent level curves (Figure 14), it can be noted that the sudden jump in the area determination can also be noticed in the retention volumes, as expected.

#### 3. RESULTS

The results of the conducted study gave an insight on the impact of the digital elevation modelling methods over the volumetric calculations in polders. When applied to the same dataset, the methods yielded different results, partially confirming the predicted usefulness of the algorithm.

Unsmoothed IDW, as expected, couldn't cope with the sudden changes in elevation, and created the well-known bull's eye areas, which couldn't be accepted for further studies in this scenario.

The Natural Neighbour method had sub-estimated the tendency of the terrain change, and resulted in a narrower hydrographic basin representation. Nearest Neighbour, on the other hand, relied on the values of the nearest points, creating a very coarse DEM with major differences regarding the physical reality, and thus resulting in hugely over-estimated water retention values.

Ordinary Kriging and Minimum Curvature methods proved to be quite efficient in this scenario. Both the visual representation of the resulted DEM and the similarity in the surface area and volume calculations suggest that the algorithms used in these methods managed to represent correctly the studied area. As Minimum Curvature does not honour the measured data exactly because of the aggressive smoothing method implemented in the algorithm, Ordinary Kriging is considered to best suited the studied scenario.

#### 4. DISCUSSIONS

As it has already been mentioned, in the process of modelling a specific area, both the data acquisition and the data interpolation methods must be taken into account. Regarding the dataset, coarse or uncharacteristic data may negatively influence the results, while high resolution data may prolong the processing time, without noticeable differences regarding the purpose of the end product. In the process of data interpolation, numerous methods can be used, and although exactly modelling the physical reality is impossible, some methods offer better results than others in certain scenarios.

When evaluating the precision and correctness of the generated Digital Elevation Model, not only statistical error reporting must be taken into account. Visual inspection of the physical reality is necessary (both from bird's eye view and terrain level view, if possible) in order to establish whether the generated DEM corresponds to the reality or not. Still, as the study suggests, running multiple algorithms on a certain dataset and directly comparing the results may also improve the confidence level of certain methods.

#### 5. REFERENCES

Atun, P.V., 2013. The Egyptian Journal of Remote Sensing and Space Science. The Egyptian Journal of Remote Sensing and Space Science, Volume 16, Issue 2, pp. 133-139, https://doi.org/10.1016/j.ejrs.2013.09.001 Dumitru, P.D., Plopeanu, M., Badea, D., 2013. Comparative study regarding the methods of interpolation. Recent Advances in Geodesy and Geomatics Engineering., pp. 45-52, ISBN: 978-960-474-335-3

Fernandez, A., Adamowski, J., Petroselli, A., 2016. Analysis of the behavior of three digital elevation model correction methods on critical natural scenarios. Journal of Hydrology: Regional Studies, Volume 8, pp. 304-315, https://doi.org/10.1016/j.ejrh.2016.09.009

Gillin, C.P., Bailey, S.W., McGuire, K.J., Prisleyt, S.P., 2015. Evaluation of Lidar-derived DEMs through Terrain Analysis and Field Comparison. Photogrammetric Engineering & Remote Sensing, Volume 81, Issue 5, pp. 387-396, https://doi.org/10.14358/PERS.81.5.387

Jobin, T., Sabu, J., Thrivikramjic, K.P., Arunkumard, K.S., 2014. Sensitivity of digital elevation models: The scenario from two tropical mountain river basins of the Western Ghats, India. Geoscience Frontiers, Volume 5, Issue 6, pp. 893-909, https://doi.org/10.1016/j.gsf.2013.12.008

Litwin, U., Pijanowski, J. M., Szeptalin, A., Zygmunt, M., 2013. Application of Surfer software in densification of digital terrain model (DTM) grid with the use of scattered points. Geomatics, Landmanagement and Landscape No. 1, pp. 51–61, http://dx.doi.org/10.15576/gLL/2013.1.51

Oksanen, J., 2006. Digital Elevation Model Error in Terrain Analysis. Academic Dissertation in Geography, Publications of the Finnish Geodetic Institute, Helsinki University Press, ISBN-13: 978-952-10-3350-6

Patel, A., Katiyar, S.K., Prasad, V., 2016. Performances evaluation of different open source DEM using Differential Global Positioning System (DGPS). The Egyptian Journal of Remote Sensing and Space Science, Volume 19, Issue 1, pp. 7-16, https://doi.org/10.1016/j.ejrs.2015.12.004

Suba, N.Sz., Suba, Şt., 2015. Mapping data – Quality, quantity or both? Journal of Applied Engineering Sciences, vol. 5(18), issue 1/2015, pp. 101-108, https://doi.org/10.1515/jaes-2015-0013

Thomas, I.A., Jordan, P., Shine, O., Fenton, O., Mellander, P.-E., Dunlop, P., Murphy, P.N.C., 2017. Defining optimal DEM resolutions and point densities for modelling hydrologically sensitive areas in agricultural catchments dominated by microtopography. International Journal of Applied Earth Observation and Geoinformation, Volume 54, pp. 38-52, https://doi.org/10.1016/j.jag.2016.08.012

Wise, S.M., 2007. Effect of differing DEM creation methods on the results from a hydrological model. Computers & Geosciences, Volume 33, Issue 10, pp. 1351-1365, https://doi.org/10.1016/j.cageo.2007.05.003

Yang. C-s., Kao, Sz-p., Lee, F-b., Hung, P-s., 2004. Twelve different interpolation methods: A case study of Surfer 8.0. Proceedings of the XXth ISPRS Congress, Istanbul, pp. 778-785, ISSN: 1682-1750