THE EFFECT OF INTERPOLATION METHODS ON THE QUALITY OF A DIGITAL TERRAIN MODEL FOR GEOMORPHOMETRIC ANALYSES

Ante Šiljeg, Sanja Lozić, Denis Radoš

Despite the modern technological advancements, contemporary methods of data gathering still rely on point samples, which means they feature a precise variable info only on specific x, y - coordinates. In order to get a continual representation of a surface necessary for research, it is necessary to approximate various data values on those surfaces which have not been sampled by using various methods of interpolation. The main aim of this research is to evaluate two different methods of interpolation based on the quality of digital terrain models used for geomorphometric analyses. For the purposes of developing and comparing these methods, a set of elevation data was attained by vectorization of contour lines from HOK. Two methods of interpolation were tested: triangulated irregular network (TIN) and ANUDEM (Topo to Raster). In order to get a more appropriate method, the research employed eight statistical parameters, various graphical representations (both two- and three-dimensional) and geomorphometric parameter of slope inclination. The research shows that ANUDEM is a better method for geomorphometric analyses than TIN is. After the analysis of geomorphometric parameters (slope), the conclusion is that there are no significant differences between the digital terrain models developed on the bases of different sets of elevation data.

Keywords: digital terrain model; HOK; interpolation methods; root mean square error; vectorization

1 Introduction

Since the middle of 20th century, various techniques of digital representation and analysis of terrain have developed, due to the rapid development of computer technology, mathematics and computer graphics [31, 32]. In the concept of terrain representation, one of the most important notions is the digital terrain model [24]. The process of quantitative digital terrain analysis is known as digital geomorphometry [32], digital terrain analysis or DTA [47,16] and digital terrain modeling or DTM [27, 45, 24]. One of the main goals of geomorphometry is to develop and analyze DTM. Simply put, a digital terrain model is a statistical representation of continual terrain surfaces (vector or raster) with a set of known x, y and z coordinates within a custom coordinate system [25].

Scientific literature has not yet established a single term for it, which means that various terms are used depending on the scientific discipline which employs it or a country in which the research was conducted. This research generally uses the term digital terrain model to describe the entire area of Lake Vrana Nature Park attained by vectorization and interpolation of contour lines from HOK. A digital terrain model is a simplified representation of a part of terrain in raster form and features a specific resolution. It was developed by scientifically based method of interpolation of elevation data attained by a specific measurement and analysis method [40]. Despite the modern technological advancements, contemporary methods of data gathering still rely on point samples, which means they feature a precise variable info only on specific x, y - coordinates. In order to get a continual representation of a surface necessary for research, it is necessary to approximate various data values on those surfaces which have not been sampled by using various methods of interpolation. The main aim of this research is to evaluate two different methods of interpolation based on the quality of digital terrain models used for geomorphometric analyses. For the purposes of developing and comparing these methods, a set of elevation data was attained by vectorization of contour lines from HOK. Two methods of interpolation were tested: triangulated irregular network (TIN) and ANUDEM (Topo to Raster). In order to get a more appropriate method, the research employed eight statistical parameters, various graphical representations (both two- and three-dimensional) and geomorphometric parameter of slope inclination. The research shows that ANUDEM is a better method for geomorphometric analyses than TIN is. After the analysis of geomorphometric parameters (slope), the conclusion is that there are no significant differences between the digital terrain models developed on the bases of different sets of elevation data.

Keywords: digital terrain model; HOK; interpolation methods; root mean square error; vectorization
The effect of interpolation methods on the quality of a digital terrain model for geomorphometric analyses

A. Šiljeg et al.

2 Methods and results

2.1 Triangulated irregular network

Triangulated irregular network (TIN) is a method of interpolation employed by linear and non-linear functions for the approximation of unknown values [35, 26]. In TIN method, all point-samples are connected by a series of triangles, mostly based on Delaunay’s triangulation (linear function) [43]. Every triangle is a section of a plane determined by three edges and vertices. For every triangle (ABC) there is a rule stating that every point apart from A, B and C is outside its circumcircle (Fig. 1).

In order to transpose the triangulation problem into a plane, it is important to note that a terrain cannot have multiple points on the same x and y coordinates and the same elevation [12]. Delaunay plane triangulation starts from the premise that the points in xy-plane are projected onto a paraboloid \( z=x^2+y^2 \) parallel to z-axis, a condition that meets the triangulation rule of the least possible number of sharp angles (optimal solution) [12]. The unknown values of points are determined mostly via one of the fastest and simplest algorithms, the so-called improving algorithm. The basis of this algorithm is the addition of \( n \) unknown points into the existing network (Delaunay network 1 – DN1), which follows the rule of removing any triangles with circumcircles containing the new point (P.) (Fig. 1). In this way a new network of triangles is constructed (Delaunay network 2 – DN2) according to previously set rules, and this action repeats (Fig. 1). TIN is a method which had a wide range of applications, although most of authors point out numerous deficiencies, mostly in cases when the distance between the points is large [4, 26]. With the development of aero and terrestrial laser data gathering technology (elevation density), TIN is showing up as an optimal method [2].

2.2 ANUDEM (Australian National University DEM) or Topo to Raster

ANUDEM is an interpolation method specifically designed for generating hydrologically accurate digital terrain model [7, 6]. This interpolation method was developed at Australian National University, but is widely used in many countries and research papers [23, 50]. With its erosive effects, water affects the terrain in a significant manner, so it is important to accurately describe relations in a complicated drainage system within a digital model. The method is based on the ANUDEM program developed by Hutchinson [19] and is in fact a modified version of thin-plate spline method, which takes into consideration significant changes in terrain, such as ridges, ravines, creeks etc. [42]. This method employs Gauss-Seidel iterative calculating method [7, 30], which is optimized to be effective as a local interpolation method (for example, inverse distance) but without the loss of area continuity characteristic of global interpolation methods [7]. Dissection is defined as first and second degree of partial derivation \( f \) of the interpolation method described by the following equations [30]:

\[
J_1(f) = \int (f_x^2 + f_y^2)dx\,dy, \\
J_2(f) = \int (f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2)dx\,dy.
\]

Minimizing \( J_2 \) results in an unrealistically flat surface, while minimizing \( J_1 \) creates a more protrusive local maximum and minimum on samples that feature better resolution. Hutchinson [19] found a better empirical compromise between \( J_1 \) and \( J_2 \), by defining the algorithm for dissection according to formula [30]:

\[
J(f) = 0.5 \times h^2 J_1(f) + J_2(f)
\]

where \( h \) stands for spatial resolution.

Hutchinson (1996) implemented a spatially variable statistical estimation for error by introducing a quadratic network to display elevation. This modification results in a level of flattening which adjusts to local terrain slopes. The purpose of the method is to identify and remove any "false" local erosion bases (sinks) in the output model, since the program assumes that all the unknown sinks are in fact errors, considering their presence in the nature is very rare [13]. The algorithm does not remove those sinks that have a vertical difference higher than specified by the user via vertical tolerance (the usual value is half the equidistance). The algorithm enables the input of data pertaining to stream lines, which means every stream can be adjusted to the valley’s bottom. In addition, the algorithm prefers the use of contour lines as primary input data in order to generate a correct model, since it is based

1150

Technical Gazette 22, 5(2015), 1149-1156
on the identification of the most significant curvature on each contour line.

2.3 Generating DTM from data attained by vectorization of contour lines

Contour lines obtained from HOK were interpolated by experts for stereo-restitution by using methods and techniques which were available during the creation of maps. Regardless of the fact that the newer data, attained and processed by modern photogrammetry, is geodetically more accurate in terms of location and elevation (the maps were first scanned and then geo-referenced), the analysis of the terrain in Vrana Lake Nature Park can be made by vectorising and interpolation contour lines from HOK. The reasons are the following: 1) lower spatial resolution enables a more realistic display of sudden terrain changes, as opposed to 25-meter resolution suggested by DGU (2003) (for data acquired by aero-photogrammetry) which flattens the surface and usually ignores some important details in the terrain; 2) for the purpose of determining elevation values, various aero-photogrammetry methods were used, which employed a detailed precise levelling (up to a millimetre level of accuracy); 3) contour lines are manually interpolated and 4) the developed model clearly shows a realistic approximation.

2.4 Accuracy of HOK

Since the WW I, cartographic development was mostly connected with military institutions outside Croatia. All topographic maps (except map 1:5000) were made in the Military Geographical Institute in Belgrade, which still has the original versions, while only a number of copies remained in the Republic of Croatia [17]. During the 1960s, topographic cards with scale 1:5000 started to be developed in Croatia as well, first by aero-photogrammetry and precise levelling, and later exclusively by aero-photogrammetry. The original papers of such maps were made on drawing foils. It is important to note that these maps were made with high scientific and expert standards, and that their quality in terms of technicality is impeccable [3]. A study was made, named “Study about the substitution of the originals and renovation of the content of topographic maps”[3], which aimed at gaining at least a partial insight into the accuracy of 1:5000 maps. The study tested the accuracy of coordinate network as well as geodetic basis, on the case of several papers of the 1:5000 map. Root mean square error of the coordinate network was approximately ±0,7m, while the root mean square error of the geodetic basis was between ±0,58 m and ±1,50 m. The accuracy of the map is directly connected to the accuracy of photogrammetric restitution. The accuracy of aero-triangulation is 4 µm ±±0,12 m. The accuracy of restitution of line and point elements on analytical stere-instruments in 10 µm ±±0,30 m. The maximum error is approximately ±1 m [3].

2.5 Vectorization of contour lines

The values of elevation data necessary for the development of DTM of Vrana Lake Nature Park were obtained by vectorising the contour lines from Croatian basic maps in 1:5000 scale. Contour lines on digital raster maps represent lines which connect the areas of the same elevation with the equidistance of 5 meters. Before the process of vectorization, contour lines in raster model can essentially have an infinite number of elevation data. However, during the vectorization process, a specific number points is defined within each line that represent elevation data, and this number depends on the chosen method. Vectorization is process of transforming raster model of data into vector model, and can be manual, semiautomatic or automatic. Manual vectorization is physically most demanding, since it includes three activities: preparation of the original, navigation of the pointer across previously defined objects and, finally, adding attributes. Automatic vectorization means that the entire process is automatic and based on various algorithms which detect and evaluate raster content [5]. Considering that the automatic vectorization from HOK resources is virtually impossible due to their low resolution and numerous objects on the maps (roads, names, dry-stone walls, water streams etc.), this research employed manual and semi-automatic vectorization. In semiautomatic vectorization, a contour line is automatically vectorised up until the first obstruction, such as crossing another contour line, which means that the operator has to manually take the cursor over the crossing and onto the line which is being vectorised [11].

Semiautomatic vectorization was made by the program WinTopo. Nineteen pages of HOK maps were used. The process of vectorization included several steps2: 1) adding the pages to the program (individually), 2) selecting contour lines in raster model (sampling colours)

---

1 HOK (hrvatska osnovna karta) eng. Croatian base map

2 Vectorization process on smaller scale maps (e.g. 1:25 000) is more complicated due to more contents (intersecting layers) [33, 34].
The aim in the process of vectorization of the line elements is to vectorise the middle line or axis. After the process of scanning, the line element is defined with numerous pixels, so it is necessary to reduce the thickness of it, but following two criteria: the line has to have the thickness of one pixel and the new representation has to follow the original line [5]. This procedure is called Median Axis Transformation or skeletalization. Statistically speaking, for the greater area of Vrana Lake Median Axis Transformation or skeletalization is to vectorise the middle line or axis. After the automatic vectorization there can be numerous errors that have to be manually removed (many of the contour lines were wrongly connected, so they needed to be separated and reconnected appropriately), 9) converting the format into a GIS-compatible one, 10) topological processing – using the program Topology, several topological rules were applied (the lines should not intercept themselves or any other line; they need to be connected where there is continuity), so that the output data should be as high quality as possible, and 11) adding attributes (elevation). In order to avoid any potential errors that can usually appear when manually dealing with elevation attributes, a regular 500×500 meter grid was constructed. This grid was used to evaluate the data, so that each of the 228 squares was manually checked. Fig. 2 shows the examples of manual and semiautomatic vectorization.

The generated contour lines and elevation points were used to develop a DTM by using two methods of interpolation: TIN and ANUDEM. In the case of TIN method, layers of contour lines and points were used. During the TIN interpolation, break points within the contour lines are used and connected to triangles according to Delaunay’s triangulation (Fig. 3).

For the ANUDEM-generated model, five layers were used: contour lines, elevation points, shape lines, break lines, lake borders and border of the Nature Park. This method is both technically and scientifically more complex and difficult since it requires the input and evaluation of a number of parameters: the area of pixel, cell margins (the distance outside the given area of interpolation has to be equal or greater than 0), the highest or lowest z-value inside the given interpolated area, choice of drainage method (depressions or sinks), choice of the primary input data (a contour line or elevation point), the number of iterations (40 optimally), dissection factor, vertical standard errors and tolerances (which depend on the primary input data). Fig. 4 shows a clear difference between the processes of interpolation (ANUDEM is more appropriate for geomorphological representations). The differences are most evident in flattened areas (such as those that feature a slope inclination lower than 5°), which make up for a significant part of the researched Nature Park. For the purpose of demonstration, an area on the island Lastovo is presented (Fig. 4). Areas with inclination lower than 2° (dark green) are shown as triangular areas (TIN) (since slopes with inclination lower than 2° are areas of accumulation); evidently, those areas cannot actually look like they are represented in the picture. Because of that, TIN interpolation is not recommended if the areas included in the vectorization process are those of karst terrain such as poljes, ravines, dry valleys etc. While using contours lines and elevation points generated from smaller scale maps (e.g. 1:25 000), TIN interpolation method shows more disadvantages due to higher equidistance, especially on flat areas [33, 34].

ANUDEM method was developed for the purpose of creating a hydrologically correct and highly optimal digital model. The problem in the ANUDEM method is the partial flattening of extremely inclined slopes (e.g. this method is not recommended for vectorising cliffs).
2.7 Comparison of DTM s developed from different sets of elevation data

The geomorphometric parameter of slope inclination (Geomorphological definition of slope inclination, based on the dominant morphological processes active in relation to inclination as well as terrain shapes, is accepted by the IGU (International Geographical Union, 1968)) was analysed in order to find out any differences between digital terrain models made from elevation data gathered by photo-photogrammetry and those made by vectorization of contour lines from HOK. However, the condition is that the spatial resolution remains the same. It is the result of elevation point density, applied interpolation methods, gathering techniques and methods and elevation data processing. For example, the spatial resolution of a model made by vectorization and interpolation of contour lines (according to the cartographic rule method) measures 5 meters, although it can be lesser than that [14], while the resolution of a model developed by interpolation of data gathered by aero-photogrammetry and stereo-restitution is 25 meters [9] or 15 meters (sample density analysis method) [40].

Because of that, the models were compared only on the examples of ravines Mernjača and Mednjača (area of 230,33 ha) which feature a more dense sampling (elevation point) so the spatial resolution could be as precise as 5 m [40].

The compared methods were Topo to Raster and ordinary cokriging (OCK-e) (the best method for data gathered by aero-photogrammetry) [40] (Fig. 5). Both models feature spatial resolution of 5 meters (pixel area is 25 m²) which enables the comparison of grid elements (pixels) and the values within them. The greatest differences in areas were noted in slopes with inclination from 12.01° to 32°, which is logical since that category is the most present in the observed area (Tab. 1). The least differences were in the category from 2.01° to 5° (Tab. 1). Both models are basically appropriate for general geomorphological analyses (meso-level). However, in the case of specific analyses (micro-level), such as cliffs or micro-denudation shapes, better results will come from aero-photogrammetry and ordinary cokriging interpolation method. Those differences are notable in three-dimensional representations.

<table>
<thead>
<tr>
<th>Class/°</th>
<th>TOPO / E</th>
<th>OCK – e / E</th>
<th>TOPO / ha</th>
<th>OCK – e / ha</th>
<th>Difference / ha</th>
<th>TOPO / %</th>
<th>OCK – e / %</th>
<th>Difference / %</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2</td>
<td>1111</td>
<td>1052</td>
<td>2,778</td>
<td>2,630</td>
<td>−0,148</td>
<td>1,206</td>
<td>1,142</td>
<td>−0,064</td>
</tr>
<tr>
<td>2,01−5</td>
<td>5286</td>
<td>5301</td>
<td>13,215</td>
<td>13,253</td>
<td>0,037</td>
<td>5,737</td>
<td>5,754</td>
<td>0,016</td>
</tr>
<tr>
<td>5,01−12</td>
<td>18338</td>
<td>19358</td>
<td>45,845</td>
<td>48,395</td>
<td>2,550</td>
<td>19,904</td>
<td>21,011</td>
<td>1,107</td>
</tr>
<tr>
<td>12,01−32</td>
<td>64757</td>
<td>63115</td>
<td>161,893</td>
<td>157,788</td>
<td>−4,105</td>
<td>70,287</td>
<td>68,505</td>
<td>−1,782</td>
</tr>
<tr>
<td>&gt;32,01</td>
<td>2641</td>
<td>3307</td>
<td>6,603</td>
<td>8,268</td>
<td>1,665</td>
<td>2,867</td>
<td>3,589</td>
<td>0,723</td>
</tr>
</tbody>
</table>

Figure 5 The differences between models made from different sets of elevation data
The effect of interpolation methods on the quality of a digital terrain model for geomorphometric analyses

A. Šiljeg et al.

1154

Technical Gazette 22, 5(2015), 1149-1156

\[ RMSE = \left( \frac{1}{N} \sum_{i=1}^{N} \left( z(x_i) - z(x_i) \right)^2 \right)^{\frac{1}{2}}, \]

where: \( x_i \) – approximate value on the location \( i \); \( z(x_i) \) – measured (real) value on the location \( i \); \( N \) – number of measured points.

Root mean square error expresses the level at which the interpolated values differ from the measured ones. It is based on the assumption that the accidental errors average 0 and that they are normally distributed [8]. Most of the papers dealing with that subject found out that the root mean square error is not 0 [41, 36, 35, 36]. In ANUDEM interpolation method, 54,338% of the points have an error smaller than 1 meter (Tab. 2). Root mean square error is 1,859 meters, while the maximum error is −12,078 (Tab. 3). Standard deviation in TIN method is 1,876 meters, whereas maximum one is −17,546 meters (Tab. 3). The percentage of errors within 1 meter in TIN method is 53,107%. The highest percentage of errors >5,01 meters was present in the vertically dissected part of Vrana Lake Natural Park or Mernjača and Mednjača ravines, for both methods (approximately 2%) (Tab. 2; Fig. 6).

### Table 2

<table>
<thead>
<tr>
<th>Classes</th>
<th>TOPO</th>
<th>%</th>
<th>TIN</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1</td>
<td>8437</td>
<td>54,338</td>
<td>8246</td>
<td>53,107</td>
</tr>
<tr>
<td>1,01 ÷ 2,5</td>
<td>4992</td>
<td>32,150</td>
<td>5104</td>
<td>32,872</td>
</tr>
<tr>
<td>2,501 ÷ 5</td>
<td>1705</td>
<td>10,981</td>
<td>1806</td>
<td>11,631</td>
</tr>
<tr>
<td>&gt; 5,01</td>
<td>393</td>
<td>2,531</td>
<td>371</td>
<td>2,389</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>IM</th>
<th>Number of points</th>
<th>Min. value/m</th>
<th>Max. value/m</th>
<th>Range/m</th>
<th>Value sum/m</th>
<th>Mean value/m</th>
<th>Variance/m²</th>
<th>Standard deviation/m</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANUDEM</td>
<td>15 527</td>
<td>−12,078</td>
<td>11,351</td>
<td>23,429</td>
<td>2227,853</td>
<td>0,143</td>
<td>1,363</td>
<td>1,859</td>
</tr>
<tr>
<td>TIN</td>
<td>15 527</td>
<td>−17,546</td>
<td>12,169</td>
<td>29,715</td>
<td>418,396</td>
<td>0,027</td>
<td>1,370</td>
<td>1,876</td>
</tr>
</tbody>
</table>

### Figure 6

Pictogram of elevation errors by classes

3 Discussion

The results of this research showed that the output results of digital models depend on the data gathering methods, sample density, interpolation methods, terrain features (primarily vertical dissection) and pixel size. The main aim of this paper was to determine the most appropriate method of interpolation and spatial resolution for purposes of various geomorphological researches. From the five most common methods of gathering elevation data and comparison of digital interpolation models, two sets of elevation data were used. They were obtained by different methods, techniques and procedures of data gathering: aero-photogrammetry, stereo-restitution and vectorization of contour lines from HOK.

Two methods of interpolation of contour lines were compared: TIN and ANUDEM (Topo to raster). The observed methods showed significant differences in generating surfaces, especially on flat terrain. The differences are most obvious when deriving geomorphometric parameters from those models. Topo to raster method showed better results and is more appropriate for developing models and analyses used by geomorphometric researches. The model that was developed by Topo to raster method was compared (by analysing geomorphometric parameter of slope inclination) to the best model developed by aero-
photogrammetrically gathered data. The research showed that there are no significant differences between the observed methods in the context of general geomorphological analyses (at a meso-level). However, when performing specific analyses (micro-level), such as on cliffs or micro-denudation formations, the results were better in the case of aero-photogrammetrically gathered data and interpolated by ordinary cokriging method. Those differences were notable on the developed three-dimensional representations.

Most of the authors who have researched a similar subject offer information only about root mean square error of the developed models. However, there are other errors in those models, such as minimum and maximum errors, that need to be taken into consideration. This research classified elevation errors into four classes in order to determine the characteristics of errors. The effect of vertical dissection of the output data was also determined.

Elevation errors of the developed models were obtained on the basis of 15 527 points gathered by aero-photogrammetry. It is important to note that points obtained by photogrammetry also feature an aspect of error which is a result of stereo-restitution. It is recommended that the future researches employ overlaying the developed models with the elevation points gathered by a precise base GPS. The drawback is that it is quite difficult to determine the exact number of samples in such case.

4 Conclusion

After the analysis of statistical parameters and elevation errors it can be concluded that there are no significant differences between the two methods, especially in terms of standard deviation (the difference being only 1.7 cm). Such results were not expected, since the interpolation functions were different. The reason behind such result is most probably a large number of samples that were used for checking the accuracy of the developed models. The differences between the models are most evident when they are used to derive the morphometric parameter of slope inclination.

5 References


Authors’ addresses

**Ante Šiljeg, Ph.D.**
Sveučilište u Zadru,
Odjel za geografiju
Trg kneza Višeslava 9, 2300 Zadar, Croatia
E-mail: asiljeg@unizd.hr

**Sanja Lozić, Ph.D. Assist. Prof.**
Sveučilište u Zadru
Odjel za geografiju
Trg kneza Višeslava 9, 2300 Zadar, Croatia
E-mail: slozic@unizd.hr; sanja.lozic@zg.t-com.hr
Cell: 098/540-832

**Denis Radoš, mag. geogr.**
Sveučilište u Zadru
Odjel za geografiju
Trg kneza Višeslava 9, 2300 Zadar, Croatia
E-mail: drados@unizd.hr