THE ACCURACY OF DETERMINISTIC MODELS OF INTERPOLATION IN THE PROCESS OF GENERATING A DIGITAL TERRAIN MODEL – THE EXAMPLE OF THE VRANA LAKE NATURE PARK

Ante Šiljeg, Sanja Lozić, Silvija Šiljeg

Original scientific paper

The aim of this research is to choose the most appropriate interpolators for the purpose of geomorphologic research, by comparing 7 deterministic methods. For the purpose of developing a model and comparing interpolators, the research used a set of elevation data gathered by aero-photogrammetry and stereo-restitution. The accuracy of interpolation methods was tested by analysing 8 statistic parameters, which were obtained by methods of cross-validation, split sampling and jackknifing. Apart from analysing the parameters, the research also compared the interpolation methods by visual means, through graphic representation of data (two-dimensional and three-dimensional) on the basis of credible graphic representations of sets of data. The research also tested the effect of power, number of neighbours, distance and sectors on the quality of output data. The multiquadratic radial basis function (MQ-RBF) proved to be the best deterministic method of interpolation by all the relevant parameters.

Keywords: digital terrain model; deterministic interpolation methods; methods of comparison of interpolation methods; root mean square error; Vrana Lake Nature Park

Točnost determinističkih metoda interpolacije u generiranju digitalnog modela reljefa – primjer PP Vransko jezero

Izvorni znanstveni članak

Cilj istraživanja je, usporedbom 7 determinističkih metoda interpolacije, odabrati najprikladnije interpolatore za potrebe geomorfoloških istraživanja. Za izradu modela i usporedbu metoda interpolacije korišten je skup visinskih podataka prikupljenih aerofotogrametrijskom izmjerom i stereorestitucijskom obradom. Točnost metoda interpolacije ispitana je analizom 8 statističkih parametara koji su dobiveni metodama unakrsnog vrednovanja, podijeljenih uzoraka i ponovno uzorkovanje (jackknifing). Osim analize parametara, interpolacijske metode su uspoređene i vizualno, kroz grafički prikaz podataka (dvodimenzionalno i trodimenzionalno) na osnovi vjerodostojnih grafičkih prikaza skupa podataka. Ispitan je utjecaj eksponenta udaljenosti, broja susjeda, udaljenosti i sektora na izlazne rezultate. Utvrđeno je da je multikvadratna radijalna osnovna funkcija (MK-ROF) po svim parametrima najbolja deterministička metoda interpolacije.

Ključne riječi: digitalni model reljefa; determinističke metode interpolacije; metode usporedbe metoda interpolacije; PP Vransko jezero; srednja kvadratna pogreška

1 Introduction

Earth's terrain structures can be extremely complex, and many scientists opt for researches by developing and analyzing digital terrain models (DTM) [9, 4, 27, 46, 43, 35, 5, 47]. A model is an object or a concept, which is a simplified version of reality (abstraction) transformed into an understandable shape [26, 48]. It can have many specific applications, such as research, predictions, risk calculations or decision making in environmental managing [34, 17, 11, 23]. DTM has a wide range of usage in scientific disciplines such as geomorphology, hydrology, climatology, landscape ecology, geology, cartography, transport, civil engineering and others. Depending on its application, a model should be developed in such a way that it fulfills the role for which it is made because various scientific disciplines do not require the same level of accuracy.

Despite the rapid advancements in technology, most of today's gathered (measured) data are point samples, meaning that they have a precise value only on specific x and y coordinates. In order to get a continual representation of a surface necessary for research and knowledge about our living space, it is necessary to approximate various data values on those surfaces which have not been sampled by using various methods of interpolation [6, 15, 19, 29].

The final result of an applied interpolation method is a model that approximates or simplifies Earth's surface. Each method offers a different representation, so the main challenge is to generate the most precise surface possible from the available data samples, and to identify the characteristics of errors and variability of approximate values. This is done by testing and comparing various methods of interpolation.

For this research, 8 statistical parameters were used in order to find out the most appropriate methods: minimum value, maximum value, range, value sum, mean value, variance and standard deviation. It is worth to specifically mention the statistical parameter of standard deviation or root mean square error, which is the most world-wide used measure of digital terrain model precision [50, 1].

For the purpose of comparing the accuracy of different interpolation methods, the method of cross-validation was used. Most authors suggest the use of this method [7, 37, 44, 18]. Two more methods were used in addition to cross-validation, split sampling and jackknifing.

In order to develop and compare digital terrain models, 7 deterministic methods were analyzed: inverse distance weighting (IDW), local polynomial (LP), radial basis function – thin-plate spline (RBF - TPS), RBF - spline with tension (SWT), RBF – completely regularized spline (CRS), RBF multiquadratic (MQ) and RBF inverse multiquadratic (IMQ).

For the input data, the research used elevation data obtained by aero-photogrammetry and stereo-restitution. The basic rule of gathering data was to measure representative points of the terrain on which the model could generate an approximate realistic surface.

Depending on whether the terrain features flat surfaces, knolls or hills, the terrain dictates the method of data gathering.

Elevation data is usually organized in three basic structures used in order to represent terrain structures in a digital form: 1) regular network (grid or matrix), 2) TIN (triangulated irregular network), and 3) contour lines.

There are various programs and modules used to interpolate the measured data in order to generate a regular quadratic network necessary for the analysis and visualization [22, 32, 30, 49, 16]. In this research, *Geostatistical Analyst* tool was used to interpolate, analyze and interpret the measured data. It offers two sets of interpolation techniques: 1) deterministic and 2) stochastic.

The aims of this research are: 1) to develop and compare digital terrain models; 2) to determine the most appropriate deterministic methods of interpolation for developing raster models used in geomorphological analyses; 3) to test the effect of power, number of neighbors, distance and sectors on the output results; 4) to find out elevation errors for DTM.

The basic hypotheses of this research are: 1) the multiquadratic radial basic function is the best deterministic method of interpolation; 2) a higher level of terrain dissection negatively affects the accuracy of interpolation algorithms; 3) the visual analysis of graphic representations (two-dimensional and three-dimensional) affects the choice of interpolation method.

2 Methods and techniques of research

The research employed various methods, techniques and procedures, in addition to general scientific methods. The aim was to integrate the various methods in order to obtain a high quality output results. The research used deterministic methods of interpolation, comparison of interpolation methods and methods of spatial resolution selection.

2.1 Methods of comparison and approximation of quality of interpolation methods

The use of interpolation methods enables a result in which the output data is a continuous sample or, in other words, a realistic representation of various objects. Each method results in a different result, so the main goal was to generate the most precise surface on the basis of available data, and to identify the characteristics of errors and variability of approximate values. This was done by comparing different interpolation methods. There are several methods of comparing interpolation methods: cross-validation [3, 21, 38, 44], split sampling [42, 21, 38, 33], jackknifing [37, 12] and residual method [25]. Most authors suggest using the cross-validation method for testing the accuracy of an interpolation method [7, 37, 44, 18]. In this method, all points are used to develop and compare models [37]. The most common form of cross-validation is the so-called *leave one technique*. This technique leaves out one point before the interpolation process in order to approximate its value. After the process the difference between the measured and approximate point is calculated [42, 21]. This process is

then repeated for every sample (measured point). The method is reliable when used on surfaces with a high enough number of representative input points [12]. It is a useful indicator of general characteristic of an interpolation method, but it cannot be used as a measure of reliability of an algorithm [37]. The method of split sampling is used to estimate the stability and accuracy of interpolation algorithm [38]. Its characteristic feature is the separation of data into two parts, training and test data, while keeping the information about the ratio of those two. Test data (random samples) are used as control points, while training data is used to develop a model. The difference between the test data and the elevation data from a developed model is used to calculate the accuracy of an interpolation method.

Conceptually similar method is jackknifing [24], however, this this method's output results depend on the technique of choosing test data (random samples). The idea of this technique is to test interpolation algorithms, specifically their consistency. A high quality algorithm is the one that shows balanced results in relation to reduced elevation data. Most commonly, about 5 %, 25 % or 50 % of input data is chosen as test data.

Unlike the cross-validation method, residual method relies on the difference between the given Z values and interpolated values, and is calculated according to the formula $Z_{res} = Z_{dat} - Z_{grd}$, where Z_{dat} is the value of a residual, Z_{dat} is the value of input data and Z_{grd} is the value of interpolated point [25]. The characteristic of this method is that the value of standard deviation is the same as in cross-validation method, while other parameters gain different values.

The most appropriate method of interpolation was chosen on the basis of 8 established parameters (variables): minimum value, maximum value, range, value sum, mean value, variance and standard deviation. The success of each interpolation method was evaluated by these parameters. All parameters are important when evaluating an interpolation method. However, one parameter usually stands out as the most frequently used one, that of standard deviation or root mean square error. It is the most world-wide used measure of digital terrain model accuracy [50, 1]. The main strength of this parameter is in its simple concept and calculation [45]. Mathematically, it is expressed with the following formula:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\{z(x_i) - z(x_j)\right\}^2}, \qquad (1)$$

where: x_i – approximate value on location i; $z(x_i)$ – measured (real) value on location i; N – the 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 premise that accidental errors equal 0 and that they are normally distributed [8]. Most researches on this subject have confirmed that the root mean square error is not 0 [42, 20, 36, 39, 40]. In addition to the analysis of parameters, the methods were compared visually, through a graphic representation of data (two-

and three-dimensional) based on realistic graphic sets of data. The research also used the method of calculation and comparison of profiles [31, 25].

2.2 Source of data

Point, line and area elevation data obtained by aero-photogrammetry are found in the Topographic database of the National Geodetic Office. They are an important part of each national data infrastructure. The rules of development and gathering of the data were established through a study and an idea project of Official Topographic Cartographic System (STOKIS). Elevation data is used for various purposes, but mostly for the creation of contour lines on topographic maps as well as a basis for rectification of orthophotos. In stereo-measuring, two types of photogrammetry devices are used: analytical instruments and digital photogrammetry stations. Operators gather information on the types of objects, in accordance with criteria defined by CROTIS (Croatia Topographic Information System). The instruments feature a level of measuring precision of 5μm. The density of the gathered data depends on the type of terrain, slope inclination and vertical dissection. The average distance between the points in breakpoints and shape lines is approximately 25 m, while the average distance in the raster of elevation points is approximately 90 m. For the development of raster DTM, 14 layers were used: embankment, notch, narrow road, pathway, coastline, canal, narrow canal, still water, creek, individual demarcation point, raster of elevation points, break-line and shape line [10]. Each layer has a numeric code and sub-code which differentiate it from other layers.

The development of DTM from the available data included several steps: 1) converting the data from one format into another – they were converted using *Conversation Tool (To Geodatabase) – Feature Class to Feature Class*, from DGN format into *Geodatabase* format. The tool also enabled the search of layers via SQL option by their codes and sub-codes, as well as the selection of a specific cartographic projection; 2) converting the lines, or breakpoints into individual points. For example, a specific break-line, regardless of its length, featured one elevation value that referred to its initial break-point. However, that line is made up of several breakpoints, where each point has an x and y coordinate and elevation. Considering that most of interpolation methods use points as input data, it was necessary to perform the conversion process. It was performed by *ET GeoWizard* extension, specifically the *Polyline M (Z) to Point* tool; 3) topological correction of the data – numerous lines and points did not meet the requirements of topological rules.

3 Results
3.1 Data interpolation

Interpolation of elevation data was processed in two phases. In the first phase the software optimized the parameters for deterministic methods of interpolation (Tab. 1).In the second phase the parameters were manually decided with for the purpose of comparing

manually and automatically chosen parameters. For the statistic comparison of methods of interpolation, the method of cross-validation was used, as well as split sampling and jackknifing. Descriptive statistics was first calculated on the basis of 83657 elevation points (of broader area of the Vrana Lake Nature Park)[1], 70806 of which refer to points gathered by photogrammetry and stereo-restitution, and 12851 points obtained by bathymetry [41].

Table 1 Parameters of interpolation methods

IM	Power	Model	Distance	*NN*	*NS*
IDW	2	circular	4912,4	15	1
LP	1	circular	127,8	Varies	1
CRS	2,368	circular	4912,4	15	1
SWT	0,426	circular	4912,4	15	1
M	0	circular	4912,4	15	1
IM	4,997	circular	4912,4	15	1
TPS	1e20	circular	4912,4	15	1

IM – interpolation method, *NN* − Number of neighbours, *NS* − Number of sectors

Bathymetrically measured points were used to avoid extrapolation in the coastal part of the lake (Fig 1). This problem (in case the lake was not re-measured) can be solved by adding a raster of points to the lake (e.g. with 10 m of distance) with the same values (e.g. 0,4 m) (Fig 2). The identical type of problem appears when making a digital model of an island. Therefore, if the sea surface is not marked by the same elevation values, the errors in the coastal areas will be extremely high, which will negatively affect the output results.

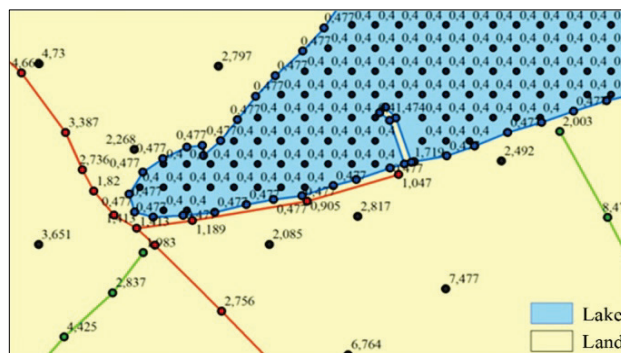**Figure 1** The problem of extrapolation on the coastal areas of a lake

Figure 2 An example of a possible solution to the coastal extrapolation problem

---

[1]The broader area is 202,5 km² (30 pages). The reasons for using elevation points outside the Vrana Lake Nature Park are: to point out the variability of statistic results in regard to the number of input data (elevation points) and to avoid extrapolation (unreliability of the model) in the bordering parts of the Vrana Lake Nature Park.

In order to determine the features of data gathered by aero-photogrammetry and stereo-restitution, statistic indicators for 15 543 points were calculated which refer to the land part of the Vrana Lake Nature Park.
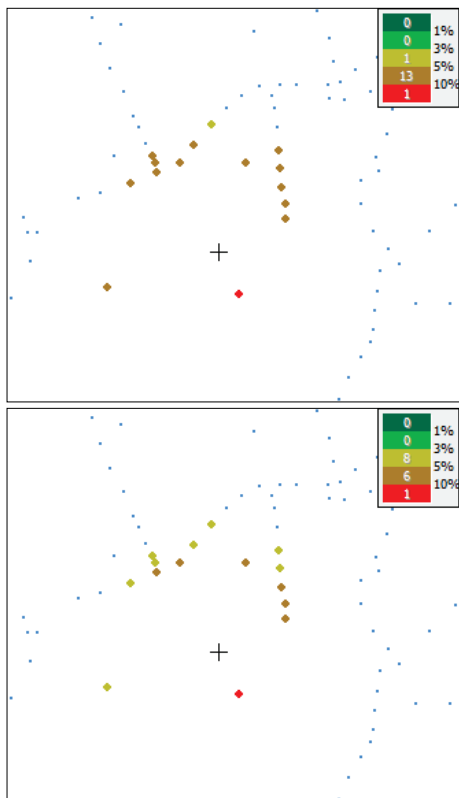

Figure 3 The effect of the power on the arrangement of weight coefficients (IDW)

Figure 4 An example of adding weight coefficients to 15 neighbouring points (IDW)

In the case of deterministic methods, the output results and other parameters are affected by the power, number of neighbours, distance and sector type (Fig. 3). If the power equals 0 in the inverse distance method, then all the neighbours have equal weigh factor for the point that is being approximated. By increasing the power, the effect of farther points is reduced (Fig. 3). For example, if the power is 1, then the value of the approximated point (a cross) is 17,918 m; if the exponent is 2, then the value for the same point increases by 1,582 m (19,5); if the exponent is 3, then the value is 21,376 m. The reason for that is that the points closest to the approximated one have the highest weight coefficient. The closest point (the

marked one) has the highest coefficient of 0,416, while the lowest ponder is 0,019 (Fig 4).

In the case of radial basis function each method has a different power, which affects the arrangement of weight coefficients in relation to the number of neighbours. For every deterministic method the rule is that the power depends on the number of neighbours and the type of sector.

For every deterministic method of interpolation the distance equalled 4912,4 m, except for local polynomial (127,8 m). A circular isotropic model was used. The general assumption is that the surrounding points have an equal effect on the central point that is being approximated (with no defined direction).

The number of neighbours determines the number of points that affect the output results for the point that is being approximated. The number of neighbours was 15 (except for local polynomial method). For the local polynomial method, the number of neighbours that affect the central point constantly changes. It primarily depends on the distance and density of points surrounding the approximated one. Therefore, the number of points involved in the calculation is defined by the distance. Fig 5, left, shows 18 neighbours that affect the central point, while the right one shows 85.
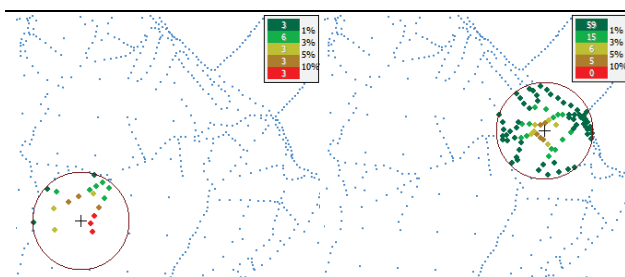

Figure 5 The effect of distance on the number of neighbouring points in the local polynomial method
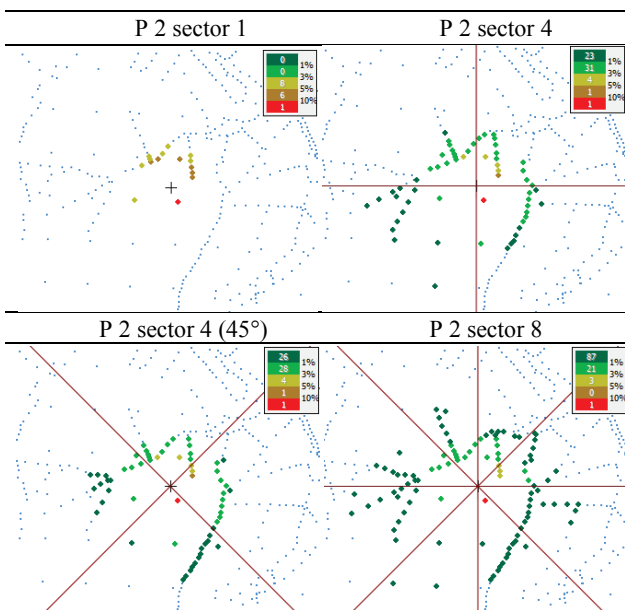

Figure 6 The role of sectors in the distribution and effect of neighbouring points on the approximated point

There are four types of sectors in *Geostatistical Analyst* extension: undivided sector, sector divided into four parts, sector divided into four parts with 45°

inclination and sector divided into eight parts. Therefore, the number and distribution of neighbouring points depend on the chosen sector (Fig 6). By changing the number of sectors, the number of neighbours multiplies by the number of divisions. For example, in the method IDW, using power 2 and sector 1 (15 neighbours), the approximated point measures 19,501. In case of sector 4 (60 neighbours, 15 in each sector), the approximated point is 19,847; in case of sector 4 (45°) with 60 neighbours the value is 20,116; and in case of sector 8 (120 neighbours, 15 in each sector), the value is 19,628. The minimum difference between the approximated values is between sectors 1 and 8 (0,127). Since most of the interpolation

process is based on rule that everything is related to everything else, but near things are more related than distant things, it can be concluded that it is unnecessary to change the number of neighbours in order to change the approximated output results. In the case of LP method, sectors do not affect the output results for the pints being approximated. In RBF method, the software optimized sector 1 and assigned 15 neighbours that affect the approximated point. By changing sectors 4 or 8, the number of neighbours also changes. Unlike in IDW, in RBF method the number of neighbours changes from 120 to 64 if the sector type changes to 8, since the assigned number of neighbours automatically lowers to 8.

Table 2 Results of the cross-validation method

IM	Number of measured points	Minimum value / m	Maximum value / m	Range / m	Value sum / m	Mean value / m	Variance / m²	Standard deviation / m
IDW	83 657	−24,68	35,80	60,48	−2419,16	−0,028	3,842	1,960
LP	83 657	−23,97	20,70	44,68	−1306,87	−0,015	3,287	1,813
CRS	83 657	−67,28	29,66	96,95	−2024,93	−0,024	2,782	1,668
SWT	83 657	−112,41	43,28	155,69	−1622,81	−0,019	2,217	1,489
MQ	83 657	−25,02	26,50	51,52	−306,38	−0,003	1,020	1,010
IMQ	83 657	−248,47	75,47	323,95	−2613,79	−0,031	4,796	2,190
TPS	83 657	−10 888,37	11 071,25	21 959,62	4516,15	0,053	3294,530	57,398

By comparing 7 deterministic methods of interpolation, it is possible to note some significant differences in the output results for each statistic parameter. Tab. 2 and Tab. 3 show how much the input data, i.e. the number of elevation points and terrain dissection, can affect the output results of a specific interpolation method. Statistic parameters were calculated for 83 657 and 15 542 elevation points in order to prove the effect of the number of points, vertical dissection and to avoid extrapolation. The range of minimum values (in m) goes from 83 657 points is from −10 888,37 (TPS) to −22,43 (OKK), while the range of maximum values (in m) goes from 12,52 (SKK) to 11 071,25 (TPS). The errors in TPS method are connected with bathymetric data gathering (increased distance of 200 m between profiles and high density of sampling) [41]. Because of that, this method was tested only on those elevation data that was gathered by photogrammetry. However, the method still showed some problems in the north-eastern part of the

Park. The method seems to be problematic at interpolating surfaces in those areas that feature high elevation data variability at shorter distances and at longer distance as well.

The method of thin plate spline is often used in the development of DTM [14, 1, 12]. The results of this method are often uncertain [28]. The identical problem of "spikes", visible in Fig. 9, appeared in the research by Aguilar et al. [1]. The method's results mostly depend on the method of gathering data, sample density and vertical dissection of the area that is being modelled.

The range (in m) is calculated on the basis of minimum and maximum value (the range between the two). Sum value is an indicator of uniformity between positive and negative values. If the sum is positive, then it indicates that there is more positive values and vice versa. The parameter of mean value is important because it indicates the features of distribution of frequencies.

Table 3 Results of the cross-validation method

IM	Number of measured points	Minimum value / m	Maximum value/ m	Range / m	Value sum / m	Mean value / m	Variance / m²	Standard deviation / m
IDW	15 542	−22,71	35,80	58,51	−3123,48	−0,200	6,980	2,642
LP	15 542	−18,49	20,57	39,07	1820,97	0,117	7,258	2,694
CRS	15 542	−67,28	26,10	93,38	−2449,13	−0,157	5,171	2,274
SWT	15 542	−112,41	43,28	155,69	−1813,76	−0,116	4,653	2,157
MQ	15 542	−16,61	17,83	34,44	−476,41	−0,030	1,850	1,360
IMQ	15 542	−248,47	75,47	323,95	−3090,12	−0,198	11,635	3,411
TPS	15 542	−10 888,37	11 071,25	21 959,62	6774,27	0,435	17 450,146	132,099

According to all indicators, the best method used on 15 542 points (aero-photogrammetrically gathered data within the Park) is multiquadratic function. In this method, the standard deviation measured 1,214. Other methods worth mentioning are completely regularized spline and spline with tension. Although the standard

deviation parameter is important, it is evident that even the best interpolation method experiences problems that generalize through this parameter, since the standard deviation indicates the deviation of minimum and maximum values from the mean value. For example, in spline with tension method, the standard deviation was

2,157, which can be a good statistical indicator. However, there is a significant error in the model at the same time (−112,41 m).

3.2 Split sampling method and jackknifing

In order to show the measure of certainty (stability and accuracy) of the best interpolation algorithm, two more methods of interpolation were used: 1) split sampling and 2) jackknifing. In the case of split sampling method, the test points were sampled by the extension *Geostatistical Analyst* and the tool *Create subsets*. In the jackknifing method, the sampling was done with the extension *Hawths Tool* and the *Sampling tool* (*create random selection)*. The process of separating training and test points in the second method is more complicated and is done semi-automatic. The methods differ in the way the test points are sampled. It is necessary to mention that the output results are similar in both methods, but they are never identical. For example, taking 25 % of test samples

for 5 times, the tools will always sample 25 % of different points. For both methods, 5, 25 and 50 % of measured points were used as test samples (Tabs. 4 and 5). Therefore, 83 657 points were divided into training and test points. Training points were used to develop models, while the test ones enabled the evaluation of stability and accuracy of interpolation algorithm. With 5 % less points (4182) the most appropriate algorithm of multiquadratic method showed excellent results. The value of standard deviation in the case of multiquadratic method was 0,993 (lowered by 0,017 in comparison with cross-validation, Tab. 3). The value of standard deviation with 25 % reduced number of points (20 914) was increased by only 0,099 (and was equal to 1,109). Another argument for its precision is that with even 50 % less points (41 828) the standard deviation in multiquadratic algorithm was increased by only 0,365. In the jackknifing method, the values of standard deviation are very similar to those of split sampling method with some differences.

Table 4 Results of the split sampling method.

IM	% of test points	Number of measured points	Minimum value / m	Maximum value / m	Value sum / m	Mean value / m	Variance / m²	Standard deviation / m
MQ	5	83 657	−11,984	12,600	−6,164	−0,001	0,996	0,993
MQ	25	83 657	−16,818	16,914	−60,410	−0,003	1,053	1,109
MQ	50	83 657	−32,494	24,794	638,226	0,015	1,173	1,375

Table 5 Results of the jackknifing method

IM	% of test points	Number of measured points	Minimum value / m	Maximum value/ m	Value sum / m	Mean value / m	Variance / m²	Standard deviation / m
MQ	5	83 657	−11,422	15,085	82,746	0,019	1,098	1,048
MQ	25	83 657	−25,114	19,138	106,562	0,005	1,208	1,099
MQ	50	83 657	−26,846	24,766	62,293	0,001	1,216	1,478

3.3 Comparison of interpolation methods through spatial representations

For the purpose of determining visual comparisons between deterministic methods of interpolation, two-dimensional and three-dimensional graphic representations of the most dissected parts of the Vrana Lake Nature Park were developed. This area was chosen because vertical dissection is one of the most important indicators of interpolation algorithm's quality. The visual component is one of the key aspects when choosing an interpolation method since it enables the comparison between the digital model and its corresponding real terrain. The visual component is often neglected in scientific works, and its importance was pointed out by Mitas and Mitasova [28]. Two-dimensional and three-dimensional representations, unlike statistic indicators, can clearly show problematic aspects of certain interpolation methods in the way they generate continuous surfaces, since they offer visual insight into illogical calculations that show up during the interpolation between the measured points. In two-dimensional representations the area is divided into 12 classes with equidistance of 20 m (with the exception of first and second class). A small segment of the lake is shown as well in order to demonstrate the differences in generating lake surface among various methods of interpolation.

Visually, methods differ in the way they flatten contour lines that depict the terrain and areas in the given elevation classes (Fig. 7).

Fig. 8 shows 6 graphs generated by the cross-validation method, and the graphs show the relation between the measured (*x*-axis) and predicted (*y*-axis) values expressed in m^{-2}. The graphs were generated on the basis of 83 675 elevation points. The difference between the predicted and the measured value is called an error. The blue lines refer to the so-called line of best solution, while the interrupted grey line shows 1:1 ratio between the measured and predicted values. Therefore, the farther the red dots are from the blue and grey line, the higher the prediction error is and vice versa. The distribution of errors is, generally observing, different for each interpolation method. The lowest digression from the line of optimal solution is connected with the multiquadratic RBF method, while the highest digression appears in thin plate spline method which features an absolute error of significant 11 071,25 m. CRS and IMQ methods feature a relatively even distribution of points in relation to the line of optimal solution. Exception appears within the elevation range from 50 to 75 m, which features errors of up to −248,47 in case of IMQ and 112,41 in case of SWT. Other significant differences between the measured and predicted values are visible in the case of inverse distance method (especially in the 50 ÷

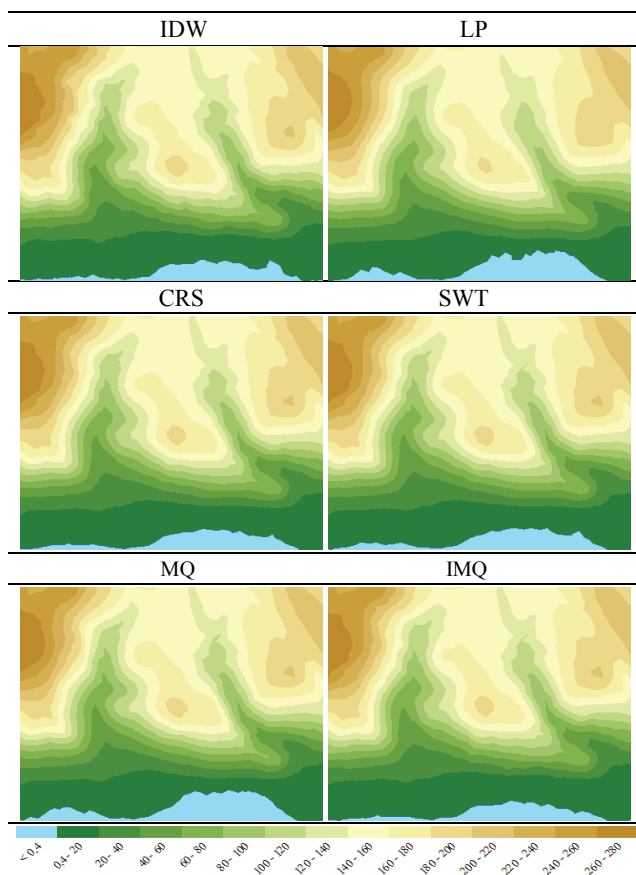120 m elevation range). For these methods, the maximum error was never higher than 39 m.

Figure 7 Representation of elevation classes on the more vertically dissected part of the Vrana Lake Nature Park

The principles of interpolation functions are clearly visible on the examples of three-dimensional representation in relation to the input data and the given parameters of deterministic method of interpolation (power, distance, number of neighbours, sector type). Fig. 9 shows some problems (dents) connected with the inverse distance method when using the exponent of 2. On higher exponents, the dents are even more expressed. This example shows that the inverse distance method offers limited usage in geomorphometric analyses, especially in the analysis of slope curvature. Local polynomial method, unlike the inverse distance, significantly flattens the curvatures and shows some problems on the lake area due to the high density of gathered points within the profile and the significant elevation difference. Radial basis function indicates differences between kernel functions [13, 21] and points out the effect of power on the output results (Tab. 1).

The best results came out of multiquadratic method, while spline with tension and completely regularized spline showed satisfactory results. The most problems were connected with the thin plate spline method. In this method, some of the Nature Park areas were represented realistically, while other parts featured extreme errors. These errors were connected to the bathymetric data gathering process (larger gap between profiles and high density of sampling). Because of that, this method was tested solely on the basis of photogrammetrically gathered data. However, the method still showed problems,

specifically in the north-eastern part of the Nature Park where the interpolation turned out bad between points with high elevation variability. Another method that did not result in quality output was inverse multiquadratic method. Although its model is similar to the inverse distance method, it is a conceptually different method. The reason for its bad results was the power of 5.
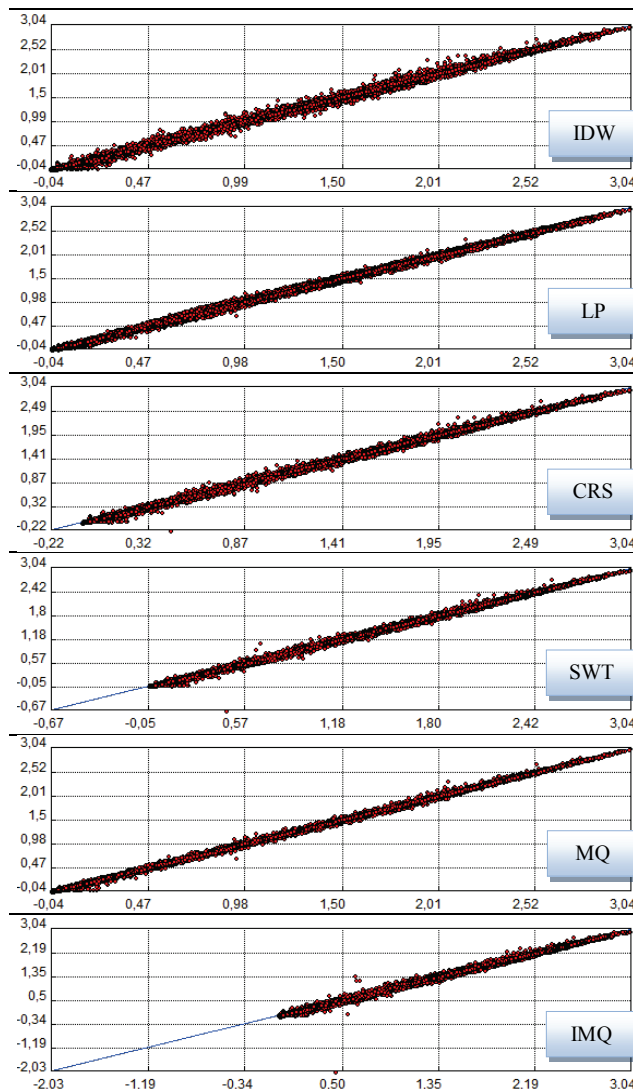

Figure 8 Regression lines of measured and predicted values for the broader area of the Vrana Lake NP

Fig. 10 and Tab. 6 show the percentage of elevation errors obtained by cross-validation method. In order to define the characteristics of errors for each method of interpolation, errors were grouped into 4 classes: <1, 1,01-2,5, 2,501-5, >5. The percentage of errors by classes was made via query function in GIS (by attribute and by location) for 15 542 points gathered by photogrammetry inside the Vrana Lake Nature Park. After the analysis, some significant differences between interpolation methods can be noticed. Among the deterministic methods, the highest percentage of errors (8,05 %) was in local polynomial method (16). By comparing it to the visual representation, it is evident that the method flattens out those surfaces that are on a more vertically dissected terrain (Mernjača and Mednjača gullies) (Fig. 10).
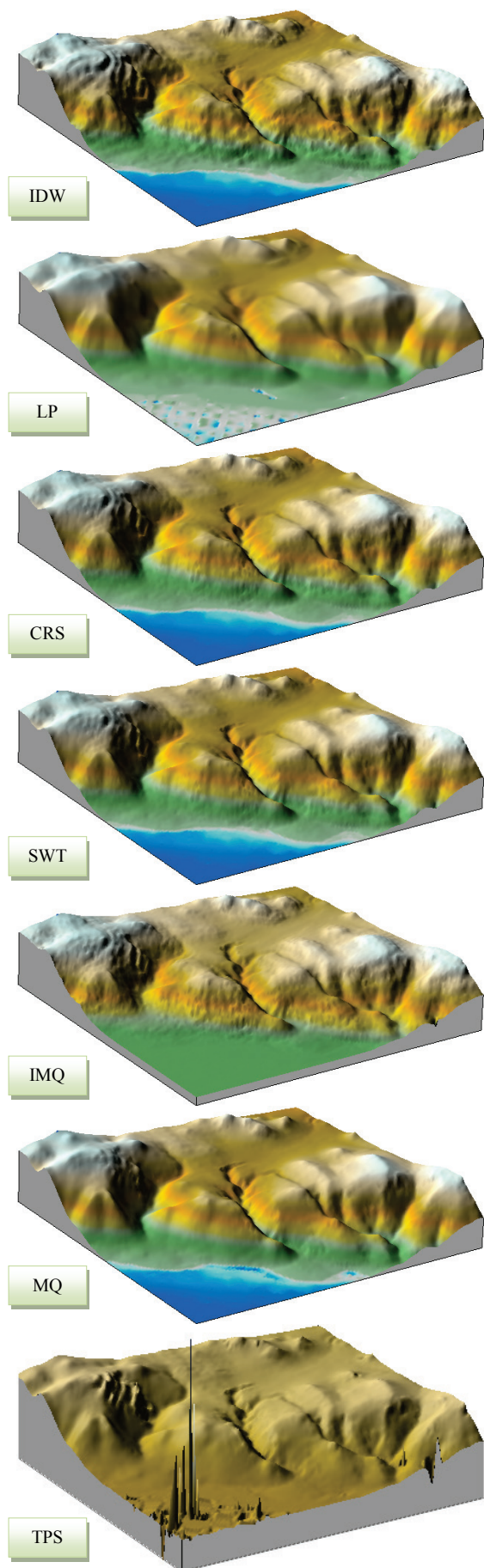
Figure 9 Three-dimensional representation of the vertically dissected part of the Vrana Lake NP

Unlike the local polynomial method, the method that had the lowest percentage of errors above 5 m (1,35 %) and the highest percentage of errors below 1 m (7,18 %) was multiquadratic method. One of the most used methods world-wide [23], inverse distance method, displayed problems with interpolation algorithm connected with areas with significant vertical dissection (Tab. 6). The percentage of errors below 1 m is 20 % lower than that of multiquadratic method, while 966 points (out of 15 542) feature an error above 5 m. Spline with tension showed satisfactory results since the number points with and error below 1 m was 10 246 (out of 15 542).

4   Discussion

The process of modelling terrain is becoming faster due to the development in methods, techniques and procedures. However, it also requires inter-disciplinary knowledge which is necessary for understanding and interpretation of modelling process. There are numerous factors that affect the output results of a DTM. They are of significant importance because models have a number of specific applications, such as research, predictions, risk calculations, decision making in managing environment etc. It is the application of a model that enables the feedback on the quality of modelling process. The purpose of terrain modelling is to ensure that the researcher is aware of what affects the output results in what manner and in any time of research, in order to generate the most appropriate model and to easily interpret the results. The researcher also has to be aware of any deficiencies of the model so that the future generations could solve those problems.

The results of this research showed that the output results of digital modelling and terrain analysis depend on data gathering methods, sample density, interpolation methods, terrain features, pixel size and applied algorithms. The main aim was to find out the most appropriate method of interpolation for the purpose of geomorphological research. Seven deterministic methods of interpolation were compared. It is important to note that there is no such thing as the single best method of interpolation, regardless of the fact that some authors prefer deterministic and others geo-statistic methods, since they are all determined by spatial and temporal component. That means that the result of comparison of various methods is the product of current time and depends on the environment and time that we live in and research, as well as the technology we use and the data we gather and process. Sometimes we are not aware that there are methods and technologies that are not available to us at the time of research.

This paper confirms the results of other world researches in which scientists argue that the multiquadratic method is not only the best radial basis function, but also one of the best deterministic methods of interpolation.

For the purpose of developing the model from data gathered by aero-photogrammetry, 7 methods of interpolation were compared in two phases. In the first phase, 83 657 points were compared by means of cross-validation method. By analysing statistic parameters, the

best results were yielded by the multiquadratic method. In order to test the reliability of interpolation algorithms, the methods were compared by split sampling and jackknifing. The worst results among the deterministic methods were obtained by the thin plate spline method.

The method is problematic at interpolating surfaces of terrain that feature high variability of elevation data on relatively short distances, but seems to be problematic at longer distances as well.
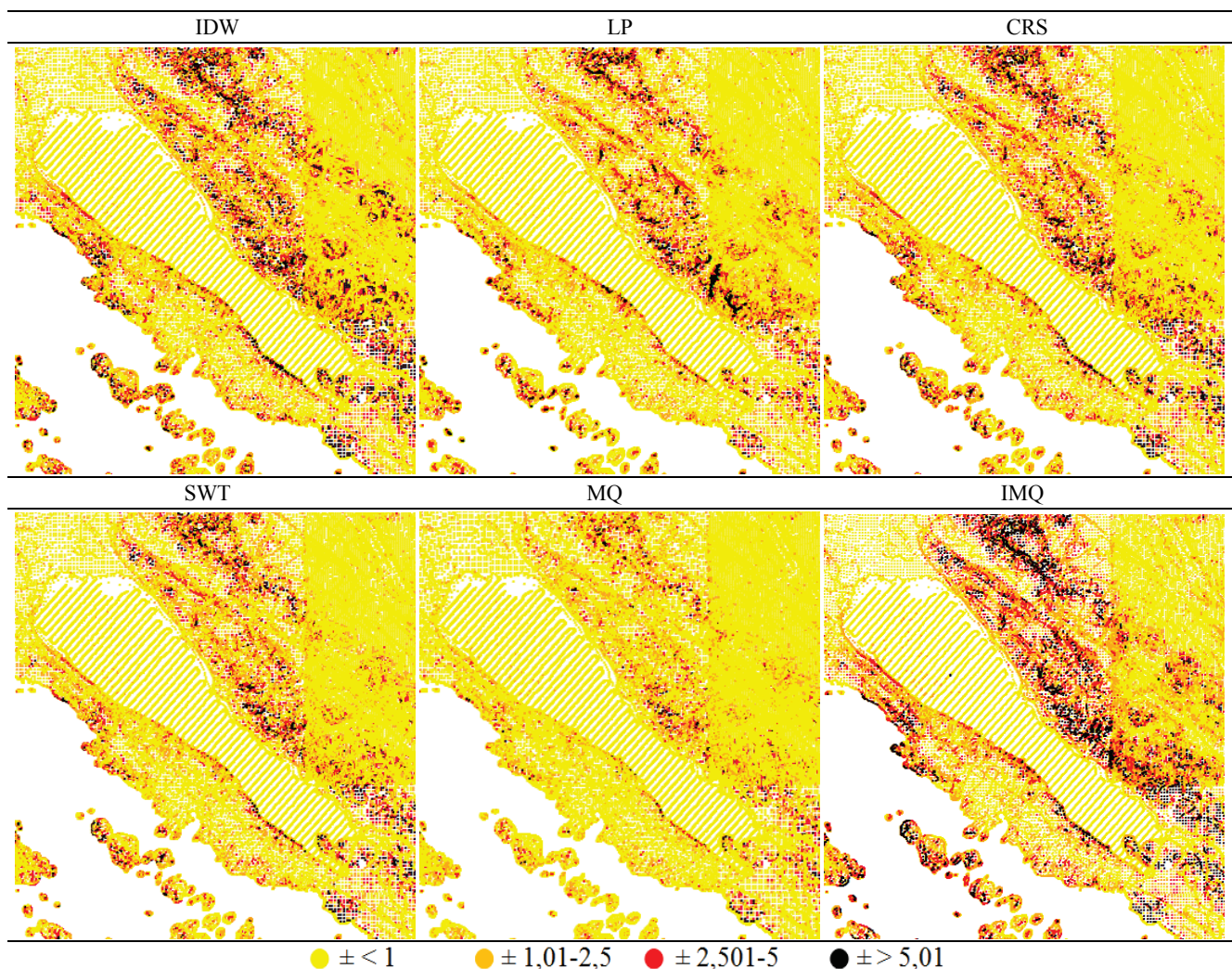
IDW	LP	CRS
SWT	MQ	IMQ

● ± < 1   ● ± 1,01-2,5   ● ± 2,501-5   ● ± > 5,01

Figure10 Pictogram of elevation errors by classes

Table 6 The number and percentage of errors by classes and interpolation methods

15 542	Error/ m				Error/ %			
IM	< 1	1,01-2,5	2,501-5	> 5,01	< 1	1,01-2,5	2,501-5	> 5,01
IDW	8770	3974	1832	966	56,428	25,569	11,787	6,215
LP	8224	3876	2191	1251	52,915	24,939	14,097	8,049
SWT	10 246	3535	1286	475	65,925	22,745	8,274	3,056
CRS	9305	3916	1638	683	59,870	25,196	10,539	4,395
MQ	11 841	2819	672	210	76,187	18,138	4,324	1,351

In the second phase, 15542 elevation points were compared within the Vrana Lake Nature Park. The best method, according to all the relevant parameters, was multiquadratic method. Its standard deviation was 1,124, while the maximum error was 14,86 m. After the analysis of statistic parameters and graphic comparisons (two-dimensional and three-dimensional) the optimal results for the purpose of geomorphological research were achieved by the following methods: completely regularized spline and spline with tension.

Significant differences between interpolation methods were detected after the analysis of errors. Among the deterministic methods, the least errors above 5 m (1,35

%) and most errors below 1 m (76,18 %) appeared in multiquadratic method. By using one of the most used method world-wide, inverse distance method [23], it was detected that the interpolation algorithm performs problematic in those areas where the terrain becomes vertically more dissected.

All of the performed analyses and the conclusions that came out of them can be useful for the future researches of data gathering methods and interpolation methods. Spatial interpolation methods are very useful in various scientific disciplines, so they should be a more prominent subject of research in Croatian science. There are 42 known methods of interpolation in the world [23],

and this research tested 7 of them, integrated into ArcGIS program. It would be useful to analyse integration into other software solutions in the future researches (such as GS+, stats, gstats, geoR and others). In doing so, it would be possible to compare the results of specific interpolation methods between two or several types of software, and to point out the exact differences.

5 Conclusion

Since every method of interpolation gives a different output representation, the main challenge is to generate the most precise surface based on the available data, and to detect the characteristic of errors and variability of approximated values. This would be done by testing and comparing various methods of interpolation. Today, however, scientists can obtain various data gathered by laser (aero or terrestrial) technology. Models developed on the basis of such technology can feature very precise horizontal and vertical accuracy (in millimeters, which means that they could generate near-realistic terrain). The question that arises from such premise is whether the interpolation methods will play any significant role for terrain modeling in the future, when such technologies can gather up to 1 000 000 points per second? In other words, the question is which role will today's methods of interpolation play when comparing a model developed on the basis of data gathered by aero-photogrammetry and stereo-restitution to a model developed on the basis of data gathered by terrestrial lasers of extreme precision.

One of the subjects in future research could be the evaluation of output results (statistic, graphic etc.) of various models developed on the basis of data gathered by high precision aero-photogrammetry (models would be inter-layered and the differences detected). This approach enables precise detection of the accuracy of used interpolation functions. The greatest problem concerning laser data gathering is the process of thinning the amount of data, since it surpasses the capabilities of today's standard hardware and software solutions [2]. At the same time, interpolation models are becoming so precise that there will be a need for defining the specific standards of scaling and researching terrain from micro to macro level. The scale will have to be defined by the level of research and will be connected with the density of the gathered data on a specific surface.

6 References

[1] Aguilar, F. J.; Agüera, F.; Aguilar, M. A.; Carvajal, F. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy // Photogrammetric Engineering and Remote Sensing. 71, 7(2005), pp. 805-816. DOI: 10.14358/PERS.71.7.805

[2] Basa, L.; Juraj, I. The cloud of points and AutoCAD Civil 3D 2011. // Ekscentar. 14(2011), pp. 34-39. URL: http://hrcak.srce.hr/file/106551 (20. 1. 2012.)

[3] Bishop, C. M. Neural Networks for Pattern Recognition. Oxford Press, Oxford, 1995.

[4] Bishop, M. P.; Shroder, J. F. Remote sensing and geomorphometric assessment of topographic complexity and erosion dynamics in the Nanga Parbat massif. // Tectonics of the Nanga Parbat Syntaxis and the Western

Himalaya (Eds. Khan, M. A. et al.). Geological Society London, London, 2000, pp. 181-199.

[5] Chaplot, V.; Darboux, F.; Bourennane, H.; Leguédois, S.; Silvera, N.; Phachomphon, K. Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landform types and data density. // Geomorphology. 77, 1-2(2006), pp. 126-141. DOI: 10.1016/j.geomorph.2005.12.010

[6] Collins, F. C.; Bolstad, P. V. A comparison of spatial interpolation techniques in temperature estimation // In Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling. National Center for Geographic Information Analysis (NCGIA): Santa Fe, NM, Santa Barbara, CA, 1996.

[7] Cressie, N. Statistics for Spatial Data. Revised ed. John Wiley and Sons, New York, 1993.

[8] Desmet, P. J. J. Effects of interpolation errors on the analysis of DEMs. // Earth Surface Processes and Landforms. 22(1997), pp. 563-580. DOI: 10.1002/(SICI)1096-9837(199706)22:6<563::AID-ESP713>3.0.CO;2-3

[9] Dikau, R.; Brabb, E. E.; Mark, R. K.; Pike, R. J. Morphometric landform analysis of New Mexico. // Zeitschrift für Geomorphologie, N. F. Suppl.-Bd. 101(1995), pp. 109-126.

[10] DGU. Product specification: Digital elevation model. CRONO GIP, Zagreb, 2003.

[11] Dobesch, H.; Dumolard, P.; Dyras, I. Spatial interpolation for climate data: The use of GIS in climatology and meteorology. ISTE Ltd., London, 2007. DOI: 10.1002/9780470612262

[12] Erdogan, S. A comparison of interpolation methods for producing digital elevation models at the field scale. // Earth Surface Processes and Landforms. 34, 3(2009), pp. 366-376. DOI: 10.1002/esp.1731

[13] Fasshauer, G. E. Solving partial differential equations with radial basis functions: Multilevel methods and moothing. // Adv. Comp. Math. 11, 2-3(1999), pp. 139-159.

[14] Gousie, M. B.; Franklin, W.R. Constructing a DEM from grid-based data by computing intermediate contours. // GIS 2003: Proceedings of the Eleventh ACM International Symposium on Advances in Geographic Information Systems (Eds.: E. Hoel and P. Rigaux), New Orleans, 2003, pp. 71-77. URL:http://www.ecse.rpiscrews.us/Homepages /wrf/Research/p/p06-gousie.pdf (20.09.2012.).

[15] Hartkamp, A. D.; De Beurs, K.; Stein, A.; White, J. W. Interpolation Techniques for Climate Variables. NRG-GIS Series 99-0, CIMMYT, Mexico, D. F., 1999.

[16] Hengel, T. A Practical Guide to Geostatistical Mapping of Environmental Variables. European Communities, JRC Scientific and Technical Report, Luxembourg, 2007.

[17] Hijmans, R. J.; Cameron, S. E.; Parra, J. L.; Jones, P. G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. // International Journal of Climatology. 25, 15(2005), pp. 1965-1978. DOI: 10.1002/joc.1276

[18] Hofierka, J.; Cebecauer, T.; Šúri, M. Optimisation of Interpolation Parameters Using a Cross-validation. // Digital Terrain Modelling (Eds. Peckham, R. J., Jordan, G.), Development and Applications in a Policy Support Environment, Series: Lecture Notes in Geoinformation and Cartography. Springer, 2007, pp. 67-82.

[19] Hu, K.; Li, B.; Lu, Y.; Zhang, F. Comparison of various spatial interpolation methods for non-stationary regional soil mercury content. // Environmental Science. 25, 3(2004), pp. 132-137.

[20] Hutchinson, M. F. Interpolation of Rainfall Data with Thin Plate Smoothing Splines - Part I: Two Dimensional Smoothing of Data with Short Range Correlation. // Journal

of Geographic Information and Decision Analysis. 2, 2(1998), pp. 139-151.

[21] Johnston, K.; Hoef, J. M. V.; Krivoruchko, K.; Lucas, N. Using ArcGIS™ Geostatistical Analyst. ESRI, Redlands, USA, 2001.

[22] Legendre, P.; Legendre, L. Numerical Ecology. Elsevier, Amsterdam, 1998.

[23] Li, J.; Heap, A. D. A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia, Record 2008/23, Canberra, 2008.

[24] Lloyd, C. D.; Atkinson, P. M. Deriving DSMs from LiDAR data with kriging, // International Journal of Remote Sensing. 23, 12(2002), pp. 2519-2524. DOI: 10.1080/01431160110097998

[25] Medved, I.; Pribičević, B.; Medak, D.; Kuzmanić, I. Comparison of Interpolation Methods of Bathymetry Data Used for Monitoring of Lake Volume Change. // Geodetski list. 68, 2(2010), pp. 71-86.

[26] Meyer, W. Concepts of Mathematical Modelling. McGraw-Hill Book Company, New York, 1995.

[27] Millaresis, G. C.; Argialas, D. P. Extraction and delineation of alluvial fansfrom digital elevation models and Landsat Thematic Mapper images. // Photogrammetric Engineering and Remote Sensing. 66, 9(2000), pp. 1093-1101.

[28] Mitas, L.; Mitasova, H. Spatial Interpolation. // Geographical Information Systems: Principles, Techniques, Management and Applications (Second edition) (Eds.: Longley, P.; Goodchild, M. F.; Maguire, D. J.; Rhind, D. W.), Wiley, Chichester, 1999, pp. 481-492.

[29] Naoum, S.; Tsanis, I. K. Ranking spatial interpolation techniques using a GIS-based DSS. // Global Nest: the International Journal. 6, 1(2004), pp. 1-20.

[30] Pebesma, E. J. Multivariable geostatistics in S: the Gstat package. // Computer & Geosciences. 30, 7(2004), pp. 683-691. DOI: 10.1016/j.cageo.2004.03.012

[31] Pribičević, B.; Medak, D.; Kordić, B.Application of Two-Frequencies Bathymetry in Determination of Travertine Layer. // Geodetski list. 1(2007), pp. 1-18.

[32] Robertson, G. P. GS⁺: Geostatistics for the Environmental Sciences. Gamma design Software, Plainwell, Michigan USA, 2000.

[33] Saito, H.; Seki, K.; Šimunek, J. An alternative deterministic method for the spatial interpolation of water retention parameters. // Hydrol. Earth Syst. Sci. 13(2009), pp. 453-465. DOI: 10.5194/hess-13-453-2009

[34] Schloeder, C. A.; Zimmerman, N. E.; Jacobs, M. J. Comparison of methods for interpolating soil properties using limited data. // Soil Science Society of America Journal. 65, 2(2001), pp. 470-479. DOI: 10.2136/sssaj2001.652470x

[35] Shary, P. A; Sharaya, L. S.; Mitusov, A. V. Fundamental quantitative methods of land surface analysis. // Geoderma. 107, 2(2002), pp. 1-32. DOI: 10.1016/S0016-7061(01)00136-7

[36] Sheikhhasan, H. A Comparison of Interpolation Techniques for Spatial Data Prediction. Master's Thesis in Computer Science. Faculty of Science, Universiteit van Amsterdam, The Netherlands, 2006.

[37] Smith, S. L.; Holland, D. A.; Longley, P. A. Interpreting Interpolation: The Pattern of Interpolation Errors in Digital Surface Models Derived from Laser Scanning Data. CASA Working Paper 66. Centre for Advanced Spatial Analysis, University College London, London, 2003.

[38] Smith, S. L.; Holland, D. A.; Longley, P. A. Quantifying interpolation errors in urban airborne laser scanning models. // Geographical Analysis. 37, 2(2005), pp. 200-224. DOI: 10.1111/j.1538-4632.2005.00636.x

[39] Soenario, I.; Plieger, M.; Sluiter, R. Optimization of Rainfall Interpolation. Royal Netherlands Meteorological Institute (KNMI), De Bilt, Nederlands, 2010.

[40] Svobodava, J. Selection of Appropriate Interpolation Methods for Creation DEMs of Various Types of Relief by Copplex Approach to Assessment of DEMs. // GIS Ostrava, (2011), http://gis.vsb.cz/GIS_Ostrava/GIS_Ova_2011/ sbornik/papers/Svobodova.pdf (02.10.2012.)

[41] Šiljeg, A. Digital Terrain Model in the Analysis of Geomorphometrical Parameters - The Example of Nature Park Lake Vrana), Doctoral Thesis, Faculty of Science, University of Zagreb, 2013. 187 p.

[42] Tomczak, M. Spatial Interpolation and its Uncertainty Using Automated Anisotropic Inverse Distance Weighting (IDW) - Cross-Validation/Jackknife Approach. // Journal of Geographic Information and Decision Analysis. 2, 2(1998), pp. 18-30.

[43] Tucker, G. E.; Catani, F.; Rinaldo, A.; Bras, R. L. Statistical analysis of drainage density from digital terrain data. // Geomorphology. 36, 3-4(2001), pp. 187-202. DOI: 10.1016/s0169-555x(00)00056-8

[44] Webster, R.; Oliver, M. A. Geostatistics for Environmental Scientists. 2ⁿᵈ Edition, John Wiley & Sons Ltd, Chichester, 2007. DOI: 10.1002/9780470517277

[45] Weng, Q. An evaluation of spatial interpolation accuracy of elevation data. // Progress in Spatial Data Handling (Eds. Riedl, A.; Kainz, W.; Elmes, G. A.), Springer-Verlag, Berlin, 2006. pp. 805-824.

[46] Wilson, J. P., Gallant, J. C. Digital terrain analysis. // Terrain analysis: principles and applications (Eds. Wilson, J. P.; Gallant, J. C.), John Wiley and Sons: New York, 2000. pp. 1-27.

[47] Wilson, J. Digital terrain modeling. // Geomorphology. 137, 1(2011), pp. 269-297.

[48] Wu, J.; Marceau, D. Modeling complex ecological systems: an introduction. // Ecological Modelling. 153(2002), pp. 1-6. DOI: 10.1016/S0304-3800(01)00498-7

[49] Yang, C. S.; Kao, S. P.; Lee, F. B.; Hung, P. S. Twelve Different Interpolation Methods: a Case Study of Surfer 8.0. // Proceedings of the XXᵗʰ ISPRS Congress, Istanbul, 2004, pp. 778-783.

[50] Yang, X., Hodler, T. Visual and statistical comparisons of surface modeling techniques for point-based environmental data. // Cartography and Geographic Information Science. 27, 2(2000), pp. 165-175. DOI: 10.1559/152304000783547911

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

Silvija Šiljeg, prof.
Sveučilište u Zadru, Odjel za geografiju
Trg kneza Višeslava 9, 2300 Zadar, Croatia
E-mail: stoplek@unizd.hr