Comparative study regarding the methods of interpolation

PAUL DANIEL DUMITRU, MARIN PLOPEANU, DRAGOS BADEA
Department of Geodesy and Photogrammetry
Faculty of Geodesy
Technical University for Civil Engineering Bucharest
122-124, Bvd. Lacul Tei, Sector 2, Bucharest
ROMANIA
paul.dumitru@gmail.com, marin.plopeanu@geodezie.utcb.ro, badea_dragos@yahoo.com
http://geodezie.utcb.ro

Abstract: - An important problem between specialists is the interpolation of the surfaces determined by classics or modern techniques. Choosing the proper technique is practically a rigorous study regarding all the interpolation methods and using that method according to the needs of the project. The determination of the Digital terrain model, the digital surface model from GNSS, LiDAR or classic measurements, the determination of the surface of the quasi – geoid surface with the geometric method from GNSS and leveling observations are one of the projects that use the interpolation techniques. In this paper are shortly presented the main methods for interpolating the surfaces as: polynomial interpolation, Delaunay triangulation, nearest neighbor, natural neighbor, Kriging, inverse distance weighting (IDW) and spline functions. The study ends with a comparative study regarding the characteristics of each method presented.

Key-Words: -interpolation, surface, Delaunay triangulation, Kriging, Inverse distance weighting (IDW)

1 Introduction

Surface modeling is the process of a natural or artificial surface determination by using one or more mathematical equations.

An universal algorithm for surface modeling is not available for all applications, each method of generating surfaces have a number of advantages and disadvantages that must be taken into account at its discretion. Modeling the 3-dimensional surface in space involves finding a function z = f (x, y) that represents the entire surface of the values z = f (x, y) associated with the point P (x, y) arranged irregularly. In addition, this function can predict the values z = f (x, y) and for other positions regularly arranged. Such a feature is known as interpolation function. There are two types of interpolation functions, exact and approximate. In a known point of a certain value, to determine the exact interpolation of the same value is the problem proposed. In fact, a method is exact only when it is known beforehand the expression of the function z = f (x, y), if any. There are some accurate methods that can be used with a smoothing factor and in this case it can go from exact methods to approximate methods.

2 Interpolating methods

For interpolating scatter or regular distributed data there are approximate or exactly methods. They can be used for multiple purposes and depends of the type of the works. Below are shortly presented the main interpolating methods and based on this analysis could be chosen the proper one. Some of the methods could be used for interpolating data in a regular grid and other in irregular grid. They could be used to generate surfaces for 3D analyses or for determine other values for unknown points. Knowing the 2D position that has attached a value for the third coordinate the surfaces could be model for any type of project.

Some of the methods are implemented in powerful applications that can be used to generate surfaces as Digital Terrain Model, Digital Terrain Surface, (quasi)geoid conversion surface, etc..:Polynomial interpolation, Delaunay triangulation with linear or spline interpolation, nearest neighbor, natural neighbor, inverse distance weighted, Kriging method, radial basis functions, polyharmonic functions, modified Shepard, local polynomial, minimum curvature, etc..

ISBN: 978-960-474-335-3

2.1. Polynomial interpolation

For the polynomial interpolation method it is necessary to determine a polynomial that has the property to go through some data points by using different methods. This method is used to determine the general trend of the values of a polynomial function z = f (x, y) for a certain area. Polynomials can vary for the degrees number, representing different geometric surfaces: a plane, a bilinear surface (etc.) quadrant area, a cubic surface, another appropriately defined area. In addition to the variables (x, y), the maximum power of the polynomial equation of these variables may represent other parameters. The general equation (1) of a spatial surface defined by a polynomial of order n + m is:

$$f(x,y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{20}x^2 + a_{02}y^2 + a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3 + \ldots + a_{mn}x^my^n \quad (1)$$

where a_{mn} - coefficients of the variables (x, y)
m - the degree of x the function f (x, y)
n - the degree of the variable y of the function f (x, y)

The relation (1) is a polynomial of order m + n which describes the tendency of a surface determined by a set of m + n +1 points.

2.2. Multivariate interpolation

The multivariate interpolation or spatial interpolation is represented by a function that consists of more than one variable. The interpolation function is known in the set-point given by coordinates (x_i, y_i, H_i), and the problem is to find the values of interpolation functions for an arbitrary point (x, y, H).

Multivariate interpolation is performed using several methods depending on the arrangement of points in the data set. Thus, one can speak of interpolation in a regular grid (with a predefined spacing and not necessarily uniform) or the interpolation in an irregular network for a random arrangement of the data set.

Regular network is a spatial representation by rectangles. Each cell of the network can be represented by its index (i, j) in the case of two-dimensional spatial representation (Fig.1) and the (i, j, k) in the case of three-dimensional representations, each node of the network of cells having the coordinates (i · dx, dy · j) in two-dimensional space and (i · dx, dy · j, k · dz) in three-dimensional space, dx, dy and dz spacing representing the network or network step in any direction.

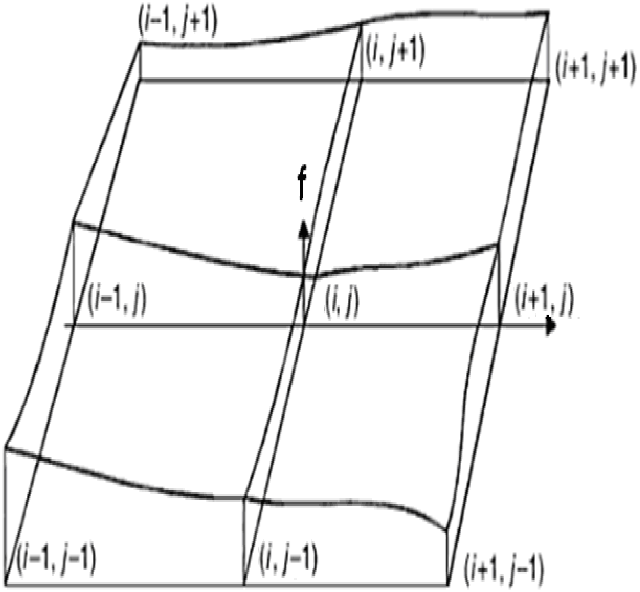

Fig.1 Regular grid interpolation [10]

An irregular network is a spatial representation by simple shapes like triangles or tetrahedral, with an irregular shape shown in Fig.2. Irregular grids may be used in the finite element analysis where the input data are the areas that have an irregular shape. Compared to regular networks, irregular networks require connectivity list or rules on the representation of the data set in irregular network. In addition for the representation by triangles or tetrahedral generating another form of surface elements is quadrilateral (4 knots) or hexahedron (8 nodes).

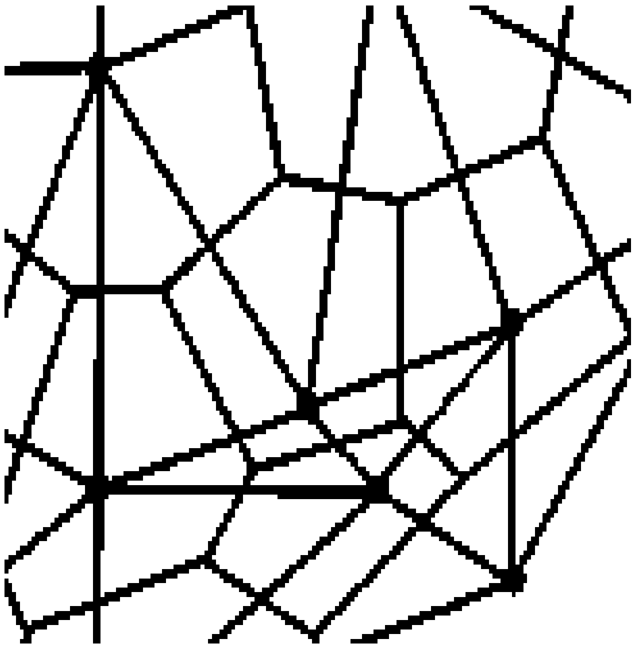
Fig. 2 Irregular network

On such surfaces form equation generated by the above forms or algorithms used to generate such networks nothing can be concluded, each interpolated or generated surface following its typical characteristics.
In the case of interpolation in a regular grid are several methods known in the function used. Thus we have:

- bilinear interpolation
- Bicubic Interpolation
- Interpolation method nearest neighbor.

Interpolation in irregular networks can be achieved using the following methods:

- nearest neighbor method
- natural neighbor method
- Delaunay triangulation with linear interpolation method
- the inverse distance weighted method
- Kriging method
- radial basis function method
- polyharmonic spline functions method.

2.3. Bilinear interpolation

Bilinear interpolation or interpolation of 1st order is used for rectangular networks where nodes know its coordinates, aiming to find the value within the cell (Fig. 3)

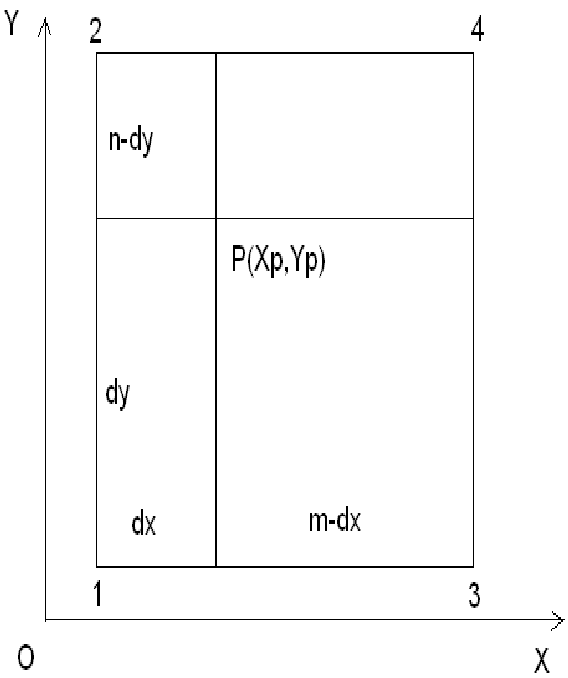

Fig.3 Bilinear interpolation in regular grid

For a point P belonging to the cell formed by the four points of the network and which has known planimetric coordinates x_p and y_p, the altitude H_p needs to be determined depending on the values of the altitudes of nodes by a bilinear interpolation.
Then the shape of the equation used for the interpolation of H_p value is (2):

$$H_P = (m - dx)(n - dy)H_1 + (m - dx)dyH_2 + dx(n - dy)H_3 + dxdyH_4 \quad (2)$$

where $dx = \frac{X_P - X_1}{X_3 - X_1}$ şi $dy = \frac{Y_P - Y_1}{Y_2 - Y_1}$ (3)

2.4. Bicubic spline interpolation

Bicubic interpolation is used to interpolate the regular grid points using a two-dimensional space. Unlike surfaces obtained by bilinear interpolation or nearest neighbor interpolated surface the method obtain smoother surfaces.
For this method it is be consider a cell in the network of 16 nodes and the interpolated point P (Fig.4):

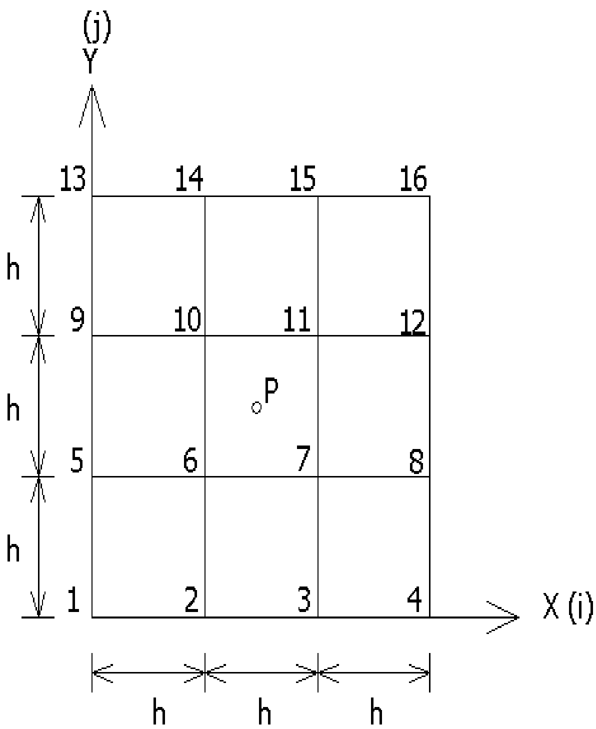

Fig.4 Bicubic spline interpolation [11]

The values of f = p (x, y) as well as the partial derivatives of the function f_x, f_y and f_{xy} are known in the corners of the unit square (h = 1, up) defined by the coordinate points 6,7,11,10 6 (0.0) 7 (1.0) 10 (0.1), and 11 (1,1).
Bicubic spline interpolation is performed using the relation (4):

$$f(x,y) = p(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} x^i y^j \quad (4)$$

ISBN: 978-960-474-335-3

2.5. Delaunay triangulation with linear interpolation

Triangulation with linear interpolation method is a method well known, being one of the first methods used before the development of the intensive computing. It is based on dividing the domain D in the R^n space into triangles.

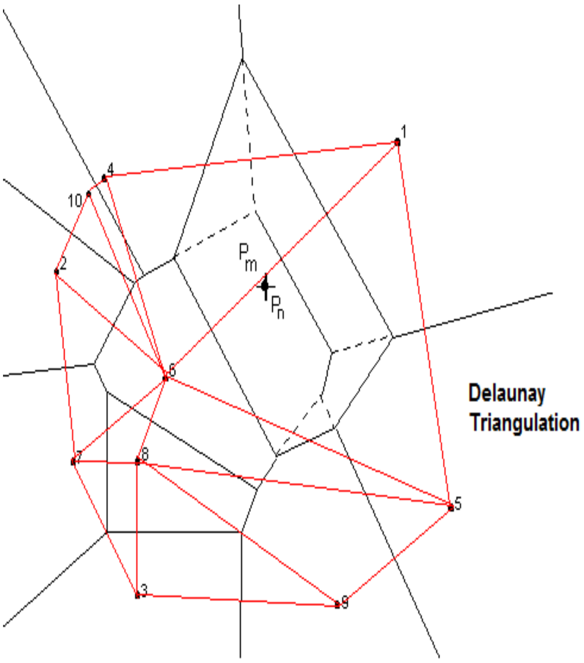

Fig. 5 Delaunay Triangulation

Then each triangle defined, thru it corners, creates a plane surface from linear parts [1].
In analytic geometry, Delaunay triangulation for a set P of points in a plane is a triangulation characterized by none of the points from P is in the circle circumscribing of the triangle (Fig.5).

2.6. Nearest-neighbor interpolation method

Nearest neighbor method was used for the first time by JG Skell and then by PJ Clark and FC Evans who introduced a statistical test to determine the significance of nearest neighbor in order to calculate the deviation from the general trend. This method is a subclass of Delaunay triangulation methods type.
Interpolation nearest neighbor method (also known as the maximum proximity interpolation method) is a simple variant of the method of multivariate interpolation in one or more dimensions [2]. Nearest neighbor algorithm selects the nearest point value and does not take into account the values of other neighboring points, producing a constant interpolation.
Nearest neighbor method is based on a comparison of the distribution of distances between a point and the nearest neighboring points of a set of randomly distributed data (5):

$$d(x,y)=\|x-y\|=\sqrt{(x-y)(x-y)}=\left(\sum_i (x_i-y_i)^2\right)^{1/2} \quad (5)$$

2.7. Inverse distance weighting (IDW) interpolation method

The inverse distance weighted method is a multivariate interpolation method, ie a process of assigning values to unknown points by using values of randomly distributed points. The method of inverse distance weighted interpolation means that the result is influenced by the nearby points and ignores unknown distant points [3].
Interpolation result is the weighted average of the values of the set of random points, the weight of each random point is reduced as the distance from the point interpolated random point increase.
The best known method of interpolation using the inverse distance weighted method is the method of Shepard, the simplest form of interpolation [3].
The form of an interpolation equation for Shepard method is represented by the following relation (6):

$$F(x,y)=\sum_{i=1}^{n} w_i f_i \quad (6)$$

where n is the number of random points in the data set, f_i the value function at points interpolated, and w_i is the weight attached to each random point. Classical formula of the calculation of the weight is given by (7):

$$w_i=\frac{h_i^{-p}}{\sum_{j=1}^{n} h_j^{-p}} \quad (7)$$

where p is a positive real number chosen randomly and called the power parameter (typically p = 2) and h_i is the distance from point to point interpolated from the original data set.
This can be determined by the relationship (8):

$$h_i=\sqrt{(x-x_i)^2+(y-y_i)^2} \quad (8)$$

The pair (x, y) represents the coordinates of the interpolated and (x_i, y_i) coordinates of the original data set. The calculation of the weight varies from unit value randomly chosen point to a value close to zero as the distance between random points increases. The calculation of the weight is normalized so that their sum is close to the value of the unit. The effect of the weighting function is that the surface interpolates each point of the data set being influenced by the distance between points and points from the interpolated data set. Relation (7) is the general form for calculating the weights in practice using the following relationship (9):

$$w_i = \frac{\left[\frac{R-h_i}{Rh_i}\right]^2}{\sum_{j=1}^{n}\left[\frac{R-h_j}{Rhj}\right]^2} \quad (9)$$

where h_i is the distance from the interpolated point to point i of the set of random data R is the distance from the interpolated point to the farthest point of the set of data and the number of random points in the data set. Equation (9) has been developed to drive superior results compared to its classical form (Franke & Nielson, 1980) [6].

Through the global method, each of the random point in the data set is searched for each point is from the interpolated for determining point closest to the point of interpolation. This method is quick sets of data with a relatively small number of points distributed randomly, but can be slow data sets with a larger number of points.

The local method, random points are temporarily in the form of triangles arranged to form an irregular triangular grid before the interpolation begins. To calculate the closest points, determine the triangle that contains the point of interpolating by Delaunay triangulation, aiming to find a systematic way to the nearest points. Local method is much faster than the global large sets of random points.

2.8. Kriging interpolation method

Kriging is a method named after a South African mining engineer named DG Krige, who developed this method in an attempt to predict more accurately the surface of mineral reserves [7][8]. In recent decades, Kriging method has become a fundamental tool in geostatistics. Kriging method is a group of geostatic techniques to interpret the value of a random field such altitude determined by geographical location.

Kriging method is based on calculating the variogram. Variogram analysis is to calculate the experimental variogram (Fig.6) of the available data and to compare it with the variogram model. Variogram model is chosen from a set of mathematical functions that describe the spatial relationship of the point data set. They can be linear functions, exponential or spherical chosen in such a way that the graphical representation of curves to fit experimental variogram of a mathematical function.

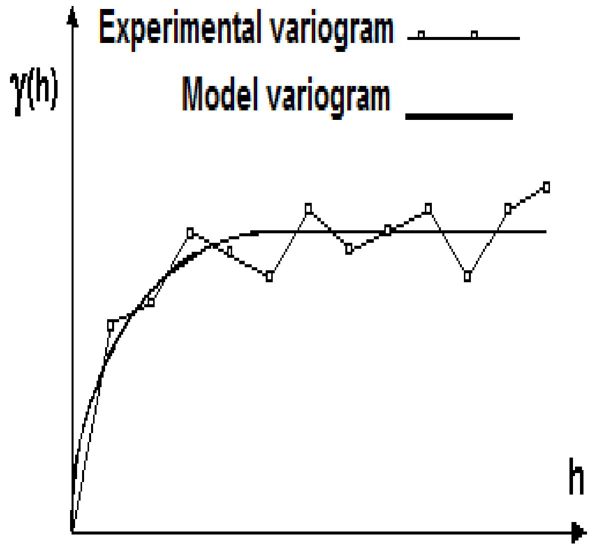

Fig.6 Experimental and model variogram for Kriging method [12]

Experimental variogram is calculated by averaging the differences in values of a function f (x, y) calculated on the basis of directions and distances between pairs of points (x, y). Variogram represents the degree of spatial dependence and is denoted by 2γ (h). The expression used to represent the variogram is given by the following (10):

$$2\gamma(h) = \frac{1}{n}\sum\left[f(x) - f(x+h)\right]^2 \quad (10)$$

where: h - the distance between two points

γ (h) - is semi-varoigram

n - number of points in the data set

Once the experimental variogram is calculated, the next step is to define a model of it. Variogram model is represented by a simple mathematical function that models the experimental variogram trends. This function can be linear, cubic, spherical, exponential, logarithmic or any other function that describes how the trend nearest experimental variogram. As can be seen in Fig. 6, indicating small differences form variogram model and experiment, proving that the variance function variogram model is small[7][8].

Once the variogram model is built it will be used to calculate weights equations for Kriging interpolation method.

Basic equation for Kriging method, given that it is based on determining the distances between points of the data set is represented by (6). This equation is just as important as the equation used in the method of inverse distance weighted interpolation except that in this case the weights are not associated with

ISBN: 978-960-474-335-3

random functions which are referenced functions to the creation of the variogram model.
To interpolate a point P which is surrounded by points P_1, P_2 and P_3 must first determine the weights w_1, w_2 and w_3. They are determined by solving the following system of equations:

$$w_1S(d_{11}) + w_2S(d_{12}) + w_3S(d_{13}) = S(d_{1p})$$
$$w_1S(d_{12}) + w_2S(d_{22}) + w_3S(d_{23}) = S(d_{2p})$$
$$w_1S(d_{13}) + w_2S(d_{23}) + w_3S(d_{33}) = S(d_{3p}) \quad (11)$$

where S (d_{ij}) is variogram model measured at a distance equal to the distance between points i and j. For example, S (d_{1p}) variogram model is measured at a distance equal to the distance between the points P_1 and P. Since it is necessary that the sum of weights is equal to the unit will enter the fourth equation:

$$w_1 + w_2 + w_3 = 1.0 \quad (12)$$

Because now there are four equations and three unknowns, it also introduces another variable to the final set of equations. The unknown is the Lagrange multiplier λ and aims to minimize or maximize the function of the variogram model defined in the sense of introducing constraints in the model. In this case the unknowns determined will define a function that will pass through the original data set points. What the sum of weights of the unknowns is equal to unity, the equation following the introduction of the condition λ becomes:

$$w_1 + w_2 + w_3 + 0 = 1.0 \quad (13)$$

This equation is added to the relations (11):

$$w_1S(d_{11}) + w_2S(d_{12}) + w_3S(d_{13}) + \lambda = S(d_{1p})$$
$$w_1S(d_{12}) + w_2S(d_{22}) + w_3S(d_{23}) + \lambda = S(d_{2p})$$
$$w_1S(d_{13}) + w_2S(d_{23}) + w_3S(d_{33}) + \lambda = S(d_{3p})$$
$$w_1 + w_2 + w_3 + 0 = 1.0 \quad (14)$$

The linear system of equations will be solved to determine the three values of the weights w1, w2 and w3. The value of the interpolation function fp of a point P is calculated by the formula:

$$F(x_p, y_p) = w_1f_1 + w_2f_2 + w_3f_3 \quad (15)$$

Using the method of least squares to solve the system of equations (16) for determining the unknowns the error is minimized.
There are two types of Kriging interpolation method known as Kriging punctual and block Kriging. Both methods generate an interpolated surface form a rectangular network. Punctual Kriging method estimates the values of points in the rectangular lattice. Block Kriging method estimates the average block centered rectangular lattice and generates smooth surfaces. The blocks have the size and shape of a rectangular grid of cells. Since the block Kriging interpolation method does not estimate the value of a point, this is not a solution for the generation of interpolated surfaces containing the points on which the surface model was generated. Even if the values that are intended to be interpolated are in a node of a rectangular network, block Kriging method estimates for that node a value close to the original [7][8].
The Kriging method involves the following techniques for interpolation approach:

- ordinary Kriging method
- simple Kriging method
- universal Kriging method

2.9. Natural neighbor interpolation method

The method of natural neighbor interpolation is based on Voronoi mosaic. Voronoi mosaic may be defined as the partition of a level of n points in each convex polygon so that contains exactly one point and each point in a given polygon is closer to the center than to any other point.
Using natural neighbor interpolation method has many beneficial features. It can be used both for interpolation and for extrapolation and works well in conjunction with a set of random locations in different cell types. The method has been used for the first time by Sibson (1981). A more detailed description of natural neighbor interpolation method using the multi-dimensional space has been performed by Owen (1992) [5].
The basic equation used in natural neighbor interpolation method is identical to that used in the interpolation using the inverse distance weighted method.

3 Conclusion

After analyzing the known methods of multivariate interpolation they can be presented their main characteristics, the goal being to choose the best method of interpolation of the available data set.
Delaunay triangulation with linear interpolation method is a proven method of surfaces interpolation based on dividing a space domain into triangles. Each triangle is defined by the three vertices forming a plane surface resulting from the combined interpolated flat triangles. The method presents a fast interpolation algorithm but has the disadvantage that the interpolation function is limited to the area bounded by convex random set of data points. The resulting surface is not smoothed and isolines are represented by line segments [12]. Division into

triangles can be ambiguous given the simple example below by choosing different triangulation of the same points it can get different shapes (valley or hill) even if Delaunay criteria were met (Fig.7):

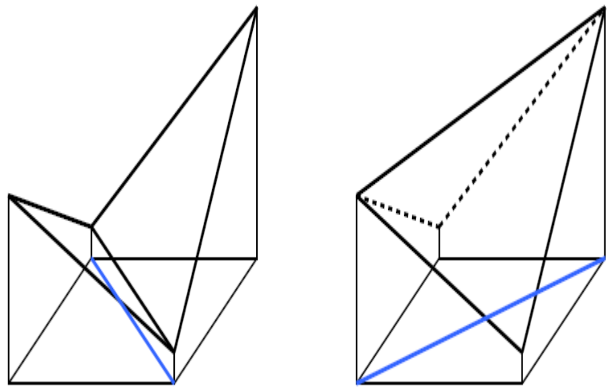

Fig. 7 The ambiguity of Delaunay triangulation [12]

Natural neighbor can be understood intuitively as all points adjacent to a given point P and forming a mosaic type Voronoi including point P. The area created using the natural neighbor interpolated surface is continuous at any point except the points that were the basis for interpolation. It presents a fast interpolation algorithm and the resulting surface is smoothed except the points in which it was made. It has the disadvantages that the interpolation function is limited to the area bounded by convex random set of data points and the resulted surface cannot be accepted in many fields such as geology or hydrology.
Inverse distance weighted method was among the first computer implemented method of interpolation, being easily programmed.

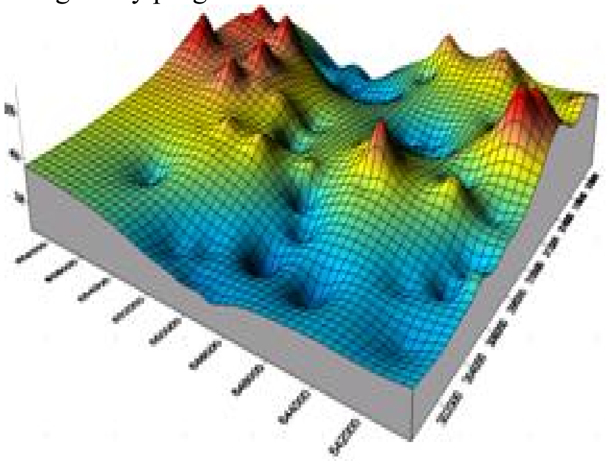

Fig. 8 Bull's eyes effect for IDW method [10]

The resulting surface is obtained by interpolation point that is located on a circle whose radius has previously been determined.
The advantages are: simple development and implementation of the algorithm easy programming and the method could be implemented in all the applications.
The disadvantages are: the calculation method is time consuming when the number of points in the data set is large. It increases the number of distances that have to be calculated. It generates so-called "bull's eyes" around a point in the interpolated field. For this reason, the resulting function is not supported for the vast majority of applications (Fig.8).
Kriging interpolation method, based on a statistical formulation of the best linear estimation, is the most used method to solve the problem of interpolation and approximation of the surfaces. An important concept derived from this method is the empirical or experimental evaluation using the variogram. Variogram is approximated using empirical or theoretical variogram model. In order to generate it the linear variogram model, Gaussian or exponential model could be used. The interpolation algorithm is based on a statistical linear formulation for the best estimation, which means that there is no other better interpolation methods in statistical terms but equations weights are calculated for each node of the network. If there is a large number of points than the number of equations to be solved will be large. For this reason the method is used to solve problems in small areas. The method produces grooves (pits) or circular isolines into the interpolated model.
As a final conclusion the table below (Table 1) presents the advantages and disadvantage of all interpolation methods presented in the current paper.

Table 1. Interpolation methods [9][10][12][13]

Method	Type of method	Advantages/Disadvantages
Inverse distance weighting (IDW)	Exact Smooth - when the smoothing parameter is known	Tends to generate the effect of "bull's eye". It is a simple method. Does not extrapolate. All values are interpolated from points used to generate the model. In this case is possible that it is not approximating fair the landforms (valleys or hills)
Natural neighbor	Exact	Interpolation is performed using the calculation of weights for neighboring points

Method	Type of method	Advantages/Disadvantages
		after the concept of Voronoi polygons. It is a good way for large data sets. A typical implementation is not extrapolated.
Nearest-neighbor	Exact	Mainly used for filling datasets (networks with missing values)
Kriging - Geostatistical models (stochastic)	Exact	High flexibility based on variogram modelling. Can extrapolate and estimate errors. Computing speed is affected by the number of points in the data set and the size of the interpolated surface.
Radial basis function	Exact Smooth - when the smoothing parameter is known	Similar to Kriging method and using the variogram models. Flexible but not give information on the statistical properties of items in the dataset.
Modified Shepard	Exact Smooth - when the smoothing parameter is known	Subclass method using an inverse distance weighted. It could not generate the correct forms of relief. Does extrapolation.
Triangulation with linear interpolation	Exact	It is a method based on Delaunay triangulation. Requires a medium or large number of data to generate acceptable results.
Triangulation with bicubic spline interpolation	Exact	It is a method based on Delaunay triangulation. Use cubic spline functions instead of linear functions. Generates very smooth surfaces.
Spline Functions	Accurately as possible and smoothed	It is available as a separate procedure and is incorporated in most interpolation methods Biharmonic bicubic spline functions or commonly used
Polynomial regression	Smooth	Produce an estimated surface that fits the points in the dataset. The most used are the linear (grade I) and the quadrics (grade II). Approximation edge can be a problem if there are insufficient data.

References:

[1] Lee DT, Schachter BJ. Two Algorithms for Constructing a Delaunay Triangulation, International Journal of Computer and Information Sciences [J].1980,9(3):219.

[2] Franke, R.: Scattered Data Interpolation: Test of Some Methods, Mathematics of Computations, 1982, 33(157):181.

[3] Watson, D. F., Philip, G. M: A refinement of inverse distance weighted interpolation. Geo-Processing 2: 315- 327, 1985.

[4] Ch. Yang, S. Kao, F. Lee, P. Hung, Twelve Different Interpolation methods: A Case Study of Surfer 8.0, Geo-Imagery Bridging Continents, XXth ISPRS Congress, 12-23 July 2004 Istanbul, 2006

[5] Steven J. Owen, An Implementation of Natural Neighbor Interpolation in Three Dimensions, Brigham Young University. Department of Engineering., 1992

[6] Franke, R. and Nielson, G., Smooth interpolation of large sets of scattered data. Int. J. Numer. Meth. Engng.,1980, 15: 1691–1704. doi: 10.1002/nme.1620151110

[7] Royle, A. G., F. L. Clausen, & P. Frederiksen, 1981, "Practical universal kriging and automatic contouring," Geo-Processing, Vol. 1, No. 4, pp. 377-394.

[8] Deutsch, C.V., & A.G. Journel, 1992, GSLIB: Geostatistical Software Library and User's Guide, Oxford University Press, New York, 340 p.

[9] www.goldensoftware.com (visit at 25.07.2013)

[10] www.mathworks.com (visit at 25.07.2013)

[11] www.ancpi.ro (visit at 25.07.2013)

[12] Mgr. Ph.D. Miroslav Dressler, Art of Surface Interpolation, Kunstat, 2009, http://m.dressler.sweb.cz/AOSIM.pdf

[13] http://www.spatialanalysisonline.com (visit at 25.07.2013)