

# APPROACH OF THE HUNGARIAN GEOID SURFACE WITH SEQUENCE OF NEURAL NETWORKS

P. Zaletnyik<sup>a,\*</sup>, L. Völgyesi<sup>a</sup>, B. Paláncz<sup>b</sup>

<sup>a</sup> Department of Geodesy and Surveying, Budapest University of Technology and Economics, H-1521 Budapest, Hungary, e-mail: volgyesi@eik.bme.hu, zaletnyikp@hotmail.com

<sup>b</sup> Department of Photogrammetry and Geoinformatics, Budapest University of Technology and Economics, H-1521 Budapest, Hungary, e-mail: palancz@epito.bme.hu

**KEY WORDS:** Neural, Artificial Intelligence, Surface, Data Mining, Geodesy

## ABSTRACT

For global data representation, like the approximation of a surface, algebraic or trigonometric polynomials may be used. However, polynomial approaches are limited concerning their accuracy. In the last decade neural networks were applied very successfully in many fields of data mining and representation.

In this research sequence of neural networks has been employed to high accuracy regression in 3D as data representation in form  $z = f(x,y)$ . The first term of this series of networks estimates the values of the dependent variable as it is usual, while the second term estimates the error of the first network, the third term estimates the error of the second network and so on. Assuming that the relative error of every network in this sequence is less than 100%, the sum of the estimated error can be reduced very significantly and effectively. To illustrate this method the geoid of Hungary was estimated. To approach this surface, a RBF neural network has been employed with 35 neurons having Gaussian activation functions. We used this type of network, because the radial basis type activation function proved to be the most efficient in case of function approximation problems. According to our experience, the iteration process is converging rapidly, and after 3-4 iteration steps there were no further significant change in the values.

Comparing the results of the first network with the fourth network the value of standard deviations was reduced with about 30 percents. And comparing these results with the polynomial approach the improvement is more significant, it is about 60 percents. These computations were carried out with the symbolic-numeric integrated system *Mathematica*.

## 1. INTRODUCTION

Nowadays the GPS measurements are one of the most frequently used technique in geodesy. With this technique ellipsoidal height can be reckoned. However in the engineering practice orthometric heights (height above sea level) are used. The orthometric heights are determined by levelling. Transforming the GPS-derived ellipsoidal heights to orthometric heights it is important to know the distance between the ellipsoidal and the geoid surface, called the geoid height or geoid undulation.

In geodesy two types of Earth's figure are distinguished. One of them is the physical or topographic earth's surface, that is the real surface of the Earth with the mountains, seas and plains. The other is the mathematical or the theoretical figure of the Earth. This is the shape of the free water surface be balanced by the gravity field only, this surface is one of the equipotential surfaces of the gravity. The equipotential surface at the mean sea level (MSL) is called geoid.

The geoid is the reference surface of the orthometric heights, but the reference surface of the GPS measurements is an ellipsoid, the WGS-84. The geoid height can be computed with a simple subtraction:

$$N = h - H \quad (1)$$

where  $N$  = geoid height  
 $h$  = ellipsoidal height  
 $H$  = orthometric height.

The transformation of ellipsoidal heights to orthometric ones requires that the geoid height must refer to the same reference ellipsoid (to the WGS-84 ellipsoid).

With the spreading of the GPS measurements the need for a good geoid model has been increased. The geoid can be calculated from different types of input data. The simplest method is to use GPS and levelling points, where both the ellipsoidal and orthometric heights are given. Another possibility and the most commonly used technique for precise determination of geoid is using the gravimetric solution, carried out by the Stokes-integral.

In Hungary considerable investigations are in progress for the determination of the separation of the geoid: lithospheric geoid solution (Papp and Kalmár 1996), gravimetric solution HGR97 (Kenyeres 1999), HGTUB98 and HGTUB2000 solution (Tóth and Rózsa 2000).

The HGTUB2000 geoid heights were used for our investigations. This gravimetric solution was based on terrestrial gravity data, height data and the EGM96 geopotential model, and was computed with the 1D Spherical FFT method

---

\* Corresponding author.

(Tóth and Rózsa 2000). The accuracy of HGTUB2000 geoid heights is about  $\pm 3\text{-}4$  cm. The used geoid heights cover the area of  $45^\circ 30' \leq \varphi \leq 49^\circ$ ,  $16^\circ \leq \lambda < 23^\circ$ ; the resolution of the grid is  $\Delta\varphi = 0' 30'' \times \Delta\lambda = 0' 50''$ . So the actual geoid heights are known in 211680 points. The geoid heights in the area vary between 37.0 and 47.1 m. Figure 1 shows the geoid surface in Hungary.

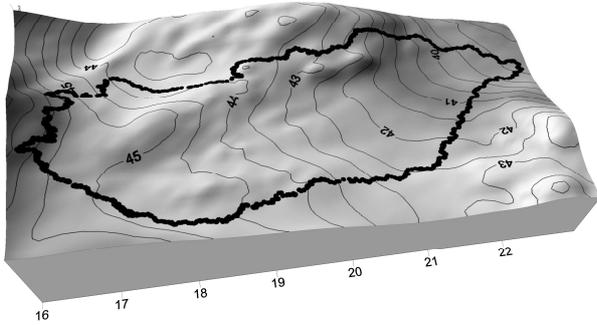


Figure 1. The HGTUB2000 geoid surface in Hungary

Instead of the application of this huge geoid database for practical purposes we tried to find a simple mathematical formula (an equation of surface of geoid forms in Hungary). Using this mathematical formula to compute geoid heights in arbitrary points in Hungary would be simpler than interpolating the geoid heights between known points, especially if it should be implemented in a computational procedure.

## 2. POLYNOMIAL FITTING

For global data representation, like the approximation of a surface, algebraic or trigonometric polynomials, least squares collocation or weighted linear interpolation may be applied. The interpolation or regression methods can be considered not only for computing unmeasured values but for compressing the data, too. In this case, with 211680 points, the data compression is a very important viewpoint.

As a classical approximation model polynomial fitting was used to approximate the geoid heights as a function of geographic coordinates  $\varphi, \lambda$ .

The formula of the used 6<sup>th</sup> order fitting polynomial is the following:

$$N = a_0 + a_1 \cdot \varphi + a_2 \cdot \lambda + a_3 \cdot \varphi^2 + a_4 \cdot \varphi \cdot \lambda + a_5 \cdot \lambda^2 + a_6 \cdot \varphi^3 + a_7 \cdot \varphi^2 \cdot \lambda + a_8 \cdot \varphi \cdot \lambda^2 + a_9 \cdot \lambda^3 + a_{10} \cdot \varphi^4 + a_{11} \cdot \varphi^3 \cdot \lambda + a_{12} \cdot \varphi^2 \cdot \lambda^2 + a_{13} \cdot \varphi \cdot \lambda^3 + a_{14} \cdot \lambda^4 + a_{15} \cdot \varphi^5 + a_{16} \cdot \varphi^4 \cdot \lambda + a_{17} \cdot \varphi^3 \cdot \lambda^2 + a_{18} \cdot \varphi^2 \cdot \lambda^3 + a_{19} \cdot \varphi \cdot \lambda^4 + a_{20} \cdot \lambda^5 + a_{21} \cdot \varphi^6 + a_{22} \cdot \varphi^5 \cdot \lambda + a_{23} \cdot \varphi^4 \cdot \lambda^2 + a_{24} \cdot \varphi^3 \cdot \lambda^3 + a_{25} \cdot \varphi^2 \cdot \lambda^4 + a_{26} \cdot \varphi \cdot \lambda^5 + a_{27} \cdot \lambda^6 \quad (2)$$

where  $a_i$  = coefficients of the polynomial  
 $N$  = geoid height  
 $\varphi, \lambda$  = geodetic latitude, longitude.

Differences between known geoid heights and approximated values are characteristic of accuracy of geoid heights computed by polynomials. Increasing the degree of polynomials, first

accuracy was increased, then decreased above the sixth degree, because of the deterioration of conditions of equations.

The most important statistical data describing the quality of the estimation are the followings: maximum, minimum error, mean value, standard deviation. These statistical data of the polynomial fitting are summarized in Table 1.

Min [m]	Max [m]	Mean [m]	St. dev. [m]
-0.812	0.722	0.000	0.180

Table 1. Quality of the polynomial fitting

The maximum accuracy resulted by applying 6<sup>th</sup> order polynomial was not enough for our purposes therefore a new method was needed to look for. As an alternative to the classical polynomial fitting a series of neural networks has been applied to approximate geoid heights.

## 3. APPROXIMATION WITH SEQUENCE OF NEURAL NETWORKS

### 3.1 Approximation with RBF neural network

To estimate the geoid, a RBF (Radial Basis Function) neural network has been employed with 35 neurons having Gaussian activation functions. We used this type of network, because the radial basis type activation function proved to be the most efficient in case of function approximation problems. Figure 2 illustrates the applied RBF network with input  $\varphi, \lambda$  (geodetic latitude, longitude) and output  $N$  (geoid height). The RBF network consists of one hidden layer of activation functions, or neurons.

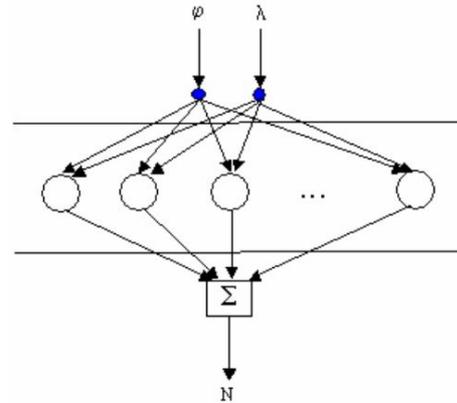


Figure 2. Applied RBF network with one output

The basis or activation function is a Gaussian bell-shaped curve with two parameters:

$$f(x) = e^{-\lambda^2(x-c)^2} \quad (3)$$

where  $\lambda$  = parameter of the function's width  
 $c$  = centre of the function  
 $x$  = input data

In case of the geoid approximation the function has two input data, two variables. The RBF network output is formed by a weighted sum of the outputs of neurons:

$$f(x_1, x_2) = \sum_{i=1}^n w_i e^{-\lambda_i^2 [(x-c_{1,i})^2 + (y-c_{2,i})^2]} \quad (4)$$

where  $x, y$  = input data  
 $\lambda, c_1, c_2$  = parameters of the activation function  
 $n$  = number of applied neurons  
 $w$  = weights of the neuron's outputs

The parameters ( $\lambda, c_1, c_2, w$ ) are determined during a supervised learning algorithm, using a teaching set to minimize the deviation between the known geoid heights and the outputs of the network.

The geoid heights are known in 211680 points, from these database 8484 points were selected for training the RBF network at a grid of  $\Delta\varphi=2'30'' \times \Delta\lambda= 4'10''$  resolution. (The original grid's resolution was  $\Delta\varphi=0'30'' \times \Delta\lambda= 0'50''$ ). The training procedure was executed with different numbers of neurons. The best configuration was using 35 neurons (Zaletnyik 2003). After the training procedure the network was tested in the whole database with the 211680 points. The summarized statistical data of the training set (teaching points) and the testing set are in Table 2.

	Min [m]	Max [m]	Mean [m]	St. dev. [m]
Teaching set (8484 points, RBF network)	-0.367	0.585	0.000	0.098
Testing set (211680 points, RBF network)	-0.416	0.600	0.000	0.099

Table 2. Quality of the estimation with RBF neural network

The results of the testing set and the teaching set are very similar, between the two standard deviations the difference is 1 mm and the maximum, minimum values are also very close to each other. All things considered can be declared that the training set with the 8484 points can represent quite well the whole database of the known geoid heights.

For our purposes the accuracy of the results was not enough. Generally the accuracy can be improved with increasing the number of the neurons, but in this case with more neurons the efficiency of the network decreased, the training procedure was slower and the improvement of the accuracy was not significant. Therefore to improve the estimation of the network we had to look for a new method.

### 3.2 Sequence of neural networks

To improve the approximation a sequence of neural networks has been applied. The first term of this series of networks estimates the values of the geoid heights, while the second term estimates the error of the first network, the third term estimates the error of the second network and so on. Assuming that the relative error of every network in this sequence is less than 100%, the sum of the estimated error can be reduced very significantly and efficiently (Paláncz, Völgyesi 2003).

According to our experience, the iteration process is converging rapidly, and after 3-4 iteration steps there was no further significant change in the values. Therefore in this study 4 networks were used. The first was a RBF neural network, and then the later used neural networks had saturated line activation function. The network learned fairly well. The results of the 4<sup>th</sup> order network are summarized in Table 3.

	Min [m]	Max [m]	Mean [m]	St. dev. [m]
Teaching set (8484 points, 4 <sup>th</sup> order)	-0.367	0.362	0.000	0.066
Testing set (211680 points, 4 <sup>th</sup> order)	-0.506	0.433	0.000	0.068

Table 3. Quality of the 4<sup>th</sup> order network

Comparing the results of the first network with the fourth network the value of standard deviations was reduced with about 30 percents. And comparing these results with the polynomial approach the improvement is more significant, about 60 percents. However the maximum errors are still too big. Figure 3 shows differences between the estimated and the original geoid heights.

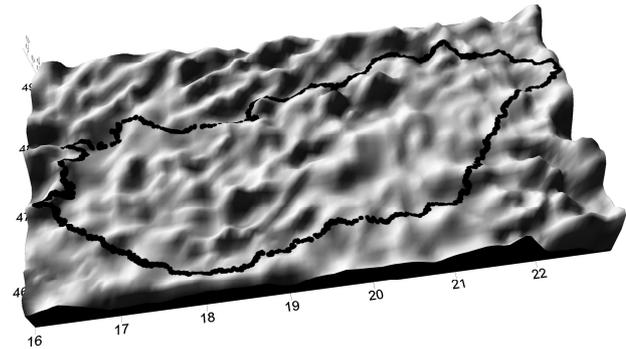


Figure 3. Differences between the estimated and the original geoid heights

Examining the distribution of the errors it was noticed that the greatest errors are outside of Hungary, in the south-east region, in Romania. In that region the quality of the input data of the geoid solution was not reliable. This could be the reason of these big errors. For our purposes these data are not necessary, because we only try to find a good geoid approximation in the region of Hungary, so they can be left out cutting them along a line. The equation of this line is very simple:  $\varphi=\lambda+25$ . Figure 4 shows this cutting line.

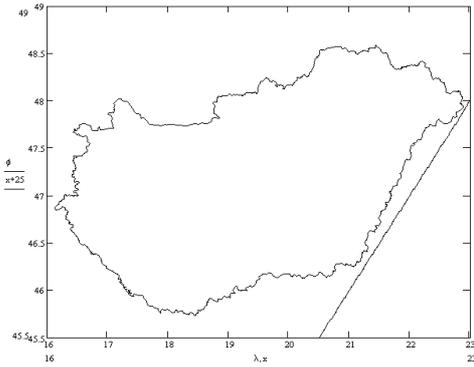


Figure 4. Cutting out the unreliable data of the south east region

Cutting out these data the number of points of the teaching set was reduced to 7438 and that of the testing set was reduced to 184910. Executing again the teaching and testing procedures the results of these networks were significantly better.

	Min [m]	Max [m]	Mean [m]	St. dev. [m]
Teaching set (7438 points, 4 <sup>th</sup> order)	-0.234	0.284	0.000	0.049
Testing set (184910 points, 4 <sup>th</sup> order)	-0.351	0.291	0.000	0.050

Table 4. Quality of the 4<sup>th</sup> order network cutting out the south east region

The standard deviation is reduced to 5 cm, and the maximum errors are lower with 10-15 cm than the maximum errors of the first sequence of neural networks. Figure 5 shows the differences between the estimated and the original geoid heights leaving out the south east region. The accuracy of the original geoid heights was about  $\pm 3-4$  cm. In Figure 5 the errors smaller than 4 cm are indicated with white color. According to Figure 5 in the greatest area of Hungary the errors of the estimation are of the same order as the errors of the original data. The estimation via sequence of neural networks provides a good approximation of the geoid heights in Hungary.

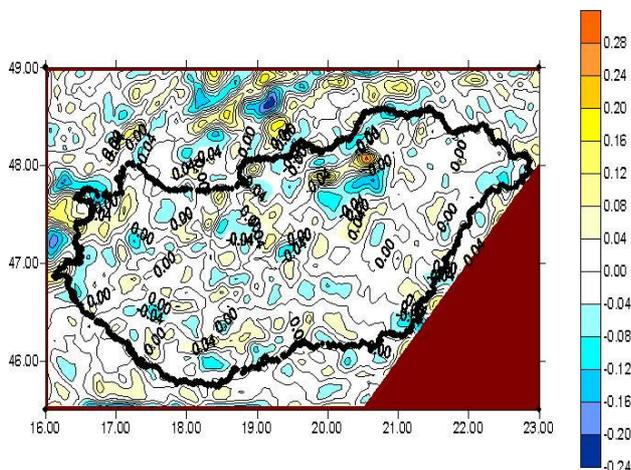


Figure 5. Differences between the estimated and the original geoid heights cutting out the south east region

#### 4. SUMMARY

In this research a sequence of neural networks was applied to approximate the geoid surface in the area of Hungary. To analyze the result, the errors of the estimation were compared with the errors of other approximation methods, with polynomial fitting and with a single RBF neural network. In both cases the sequence of neural networks proved to be better. On the basis of our research can be stated that using this method the error of the estimation can be reduced efficiently, even in the case of a morphologically so sophisticated data structure as a geoid.

For the approximation of the geoid surface a gravimetric geoid solution was used with 211680 known geoid heights in a regular grid. 8484 points were selected for the teaching set from the whole database, and the approximation method was tested in every known point. In accordance with the results the teaching set can represent quite well the whole database of the known geoid heights.

Cutting out an area with unreliable data outside of Hungary the estimation was improved significantly. The standard deviation of the errors of estimation was reduced to 5 cm and this accuracy is of the same order as the accuracy of the original data.

#### 5. REFERENCES

Kenyeres A. 1999: *Phys. Cem. Earth (A)*, 24, pp. 85-90.

Paláncz B., Völgyesi L. 2003: High accuracy data representation via sequence of neural networks, *Acta Geod. Geoph. Hung.*, Vol. 38 (3), pp. 337-343

Papp G., Kalmár J. 1996: In: Proceedings of the 7<sup>th</sup> International Meeting on Alpine Gravimetry, *Österreichische Beiträge zu Meteorologie und Geophysik*, pp. 95-96.

Tóth Gy., Rózsa Sz. 2000: New Datasets and Techniques – an Improvement in the Hungarian Geoid Solution, Paper presented at Gravity, Geoid and Geodynamics Conference, Banf, Alberta, Canada

Zaletnyik P. 2003: Neurális hálózatok alkalmazása a geodéziában, MSc. Thesis, Budapest University of Technology and Economics

#### 6. ACKNOWLEDGEMENTS

Our investigations were supported by the Hungarian National Research Fund (OTKA), contract No. T-046718.