Neuro – Fuzzy Networks for Identification of Mathematical Model Parameters of Geofield

A. Pashayev, R. Sadiqov, C. Ardil, F. Ildiz , and H. Karabork

Abstract**—**The new technology of fuzzy neural networks for identification of parameters for mathematical models of geofields is proposed and checked. The effectiveness of that soft computing technology is demonstrated, especially in the early stage of modeling, when the information is uncertain and limited.

Keywords**—**Identification, interpolation methods, neuro-fuzzy networks, geofield.

I. Introduction

For many problems in sciences on Earth (geodesy, geology, geophysics, cartography, photogrammetry, etc.) the problem of modeling the geofields surface (height, depth, pressure, temperature, pollution factor, etc.), wich is usually displayed on maps by means of isolines, is urgent. If representation of geofields surface is possible as function of two variables h=f (x, y), which has hi values at (x_i, y_i), (i $=\overline{1,n}$) peaks, the digital model of this function is required for computer processing and storage.

We are going to consider the digital model of geofield (DMG) as a set of digital values of continuous objects in cartography (e.g. height of a relief) for which their spatial coordinates and the mean of structural description are specified. It will allow calculating the values of geofield in the given area. The important part of any DMG is the method of interpolating of its surface. For this, various ways of interpolation yield various results which can be estimated only from the point of view of practical applications [1- 6].

Nowadays, more than ten methods of surface interpolation are known. They are as fellows algebraic and orthogonal polynoms, rational fractions; in some eases they take functions satisfying some apriori given conditions (e.g. positivity of f (x, y)) values; multi squadric function, at which approximation is reached bu means of square – law functions (squadric), representing hyperboles; splines; geostatic methods (kriging). However, none of them is completely universal. We shall consider widely used procedure of interpolation by algebraic polynoms

$$h(x,y)=\sum_{i=0}^{m}\sum_{j=0}^{n}A_{ij}x^iy^j$$

where $i=\overline{0,m}$; $j=\overline{0,n}$ - exponents; A - factors at decomposition members received on a method of least squares (LSM).

Realization of these methods is rather simple; therefore they have received a wide circulation [1-5]. This is the linear interpolation modeling of a surface as set of triangles. Thus the normal to a surface is constant along all surface of a triangle and sharply varies at transition through the sides separating triangles. Therefore, LSM constructed with use of linear interpolation, frequently insufficiently adequately represent the investigated phenomenon [2].

The much better result (absence of sharp differences of values of researched parameter, smoothness of isolines), is given by modeling with the use of polynomial to interpolation of higher degree. The general (common) expression for calculation of value, for example, heights h in a point of a surface with coordinates (x, y) looks like:

$$h(x,y)=\sum_{j=0}^{m}\sum_{k=0}^{m-j}C_{jk}x^jy^k \quad (1)$$

We shall consider a special case (1) at m=2, that is the equation of regress of the second order

$$H(x,y)=C_{00}+C_{10}x+C_{01}y+C_{20}x^2+C_{11}xy+C_{02}y^2 \quad (2)$$

The equation of measurements of target coordinate h for this case will be written down as:

$$Z_h=C_{00}+C_{10}x+C_{01}y+C_{20}x^2+C_{11}xy+C_{02}y^2+\delta_h$$

Then the model of an experimental material can be presented in the following matrix kind:

$$Z_h=X\theta+\delta_{h,}$$

where Z_h = || $z_{1h,}$ $z_{2h,\ldots,}$ z_{nh} || - a vector of measurements of target coordinate h; $\theta=||C_{00,}\ C_{10,}\ C_{01,}\ C_{20,}\ C_{11,}\ C_{02}||^T$ - a vector of required factors;

Manuscript received June 30, 2005. This work (R. Sadiqov) was supported by TUBITAK NATO-PC B program.

Authors are with the National Academy of Aviation, AZ1045, Bina, 25th km, Baku, Azerbaijan (corresponding author to provide phone: 99412- 497-28-29, fax: 99412-497-28-29, e-mail : sadixov@mail.ru).

$$X = \begin{Vmatrix} 1 & x_1 & y_1 & x_1^2 & x_1 y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2 y_2 & y_2^2 \\ - & - & & - & - & - \\ 1 & x_n & y_n & x_n^2 & x_n y_n & y_n^2 \end{Vmatrix}$$

A structural matrix; n - quantity(amount) of points of supervision (measurements).

Usually for identification (estimation) of factors of a polynom (2) are used LSM of the following kind

$$\theta=(X^TX)^{-1}(X^TZ_h),$$
$$D_\theta=(X^TX)^{-1}\sigma^2,$$

where D_θ - dispersive matrix of mistakes of estimations.

The use of statistical probability methods, such as the least-squares method, requires preliminary analysis of the data for normality of the sample distribution. A normality check assumes that the following four conditions are satisfied.

1. The intervals $\bar{x} \pm \sigma, \bar{x} \pm 2\sigma\, x$ and $\bar{x} \pm 3\sigma$ must contain 68, 95, and 100%, respectively, of the sample values $\bar{x}$ is the mean and o is the standard deviation).
2. The coefficient of variation V must not exceed 33%.
3. The kurtosis E_x and the asymmetry coefficient S_k must be close to zero.
4. $x \approx M$. where M is the sample median.

The analysis [6] was used for modeling (2) showed that distribution contradicted the normality assumption (Table 1).

It must be noted that in the early stage modeling of geofield, the data are not only limited and uncertain but also fuzzy (the output and input coordinates of the system are measured in definite intervals and their values are measured with errors).

II. Problem Formulation and Solution

It is therefore necessary to identify the parameters of a mathematical model of a multivariate fuzzy object described by the regression equation

$$\tilde{H}_m = \sum_{j=0}^{m} \sum_{k=0}^{m-j} \tilde{c}_{jk} \otimes \tilde{x}^j \otimes \tilde{y}^k \qquad (3)$$

$$(j = \overline{0,m};\ k = \overline{0,m},\ j+k \le m)$$

where $\tilde{c}_{jk}$ are the desired fuzzy parameters.

We shall determine the fuzzy values of the parameters $\tilde{c}_{jk}$ of equation (3) using. experimental fuzzy statistical data of the process, i.e., the input $\tilde{x}, \tilde{y}$ and output $\tilde{H}$ coordinates of the model. Let us consider a solution of this problem using fuzzy logic and neural networks [7,8].

A neural network consists of interconnected sets of fuzzy neurons. When an neural network is used to solve equation (3), the input signals of the network are the fuzzy values of the variable $\tilde{B}=(\tilde{x}, \tilde{y})$, and the output is $\tilde{H}$. The fuzzy values of the parameters $\tilde{c}_{jk}$ are the network parameters. We present the fuzzy variables in triangular form, the membership functions of which are calculated by the formula

$$\mu(x) = \begin{cases} 1-(\bar{x}-x)/\alpha, & \text{if} \quad \bar{x}-\alpha < x < \bar{x}; \\ 1-(\bar{x}-x)/\beta, & \text{if} \quad \bar{x} < x < \bar{x}+\beta; \\ 0, & \text{otherwise} \end{cases}$$

Neural-network training is the principal task in solving the problem of identification of the parameters $\tilde{c}_{jk}$ of equation (3). An α-section is used to train the parameter values [7].

We assume the presence of experimentally obtained fuzzy statistical data. From the input and output data we compose training pairs for the network $(\tilde{B},\tilde{T})$. To construct a model of a process, the input signals $\tilde{B}$ are fed to the neural network input (Fig.1); the output signals are compared with standard output signals $\tilde{T}$.

After comparison, the deviation is calculated:

$$\tilde{E} = \frac{1}{2}\sum_{i=1}^{l}(\tilde{H}_i - \tilde{T}_i)^2$$

When an α-section is used. the deviations for the left and right parts are calculated by the formulas

$$E_1 = \frac{1}{2}\sum_{i=1}^{l}\left[h_{i1}(\alpha) - t_{i1}(\alpha)\right]^2,$$
$$E_2 = \frac{1}{2}\sum_{i=1}^{l}\left[h_{i2}(\alpha) - t_{i2}(\alpha)\right]^2,$$
$$E = E_1 + E_2,$$

where $\tilde{H}_i(\alpha) = [h_{i1}(\alpha), h_{i2}(\alpha)]$; $\tilde{T}_i(\alpha) = [t_{i1}(\alpha), t_{i2}(\alpha)]$

Training (correction) of the network parameters is concluded when the deviations E for all training pairs are less than the specified value (Fig. 2). Otherwise, it is continued until E is minimized.

The network parameters for the left and right parts are corrected a-s follows:

$$c_{jk1}^n = c_{jk1}^o + \gamma\frac{\partial E}{\partial c_{jk}}, \qquad c_{jk2}^n = c_{jk2}^o + \gamma\frac{\partial E}{\partial c_{jk}}, \qquad (4)$$

Here $c_{jk1}^o, c_{jk1}^n, c_{jk2}^o$ and c_{jk2}^n are the old and new values of the left and right pans of the neural network parameters $\tilde{c}_{jk} = [c_{jk1}, c_{jk2}]$, and γ is the training rate.

ISSN: 2517-9934
Vol:4, No:11, 2010

III. Numerical Example

Large Let us consider the mathematical model is described the equation of fuzzy a regression (consider a special case (3) at m=2):

$$\widetilde{H} = \widetilde{c}_{00} + \widetilde{c}_{10}\widetilde{x} + \widetilde{c}_{01}\widetilde{y} + \widetilde{c}_{20}\widetilde{x}^2 + \widetilde{c}_{11}\widetilde{x}\,\widetilde{y} + \widetilde{c}_{02}\widetilde{y}^2. \qquad (5)$$

We shall construct a neural structure for solution of (5) in which the network parameters are the coefficients $\widetilde{c}_{00}, \widetilde{c}_{10}, \widetilde{c}_{01}, \widetilde{c}_{20}, \widetilde{c}_{11}, \widetilde{c}_{02}$. The structure has four inputs and one output (Fig. 3).

Using a neuro-network structure, we employ (4) to train the network parameters. For $a = 0$, we obtain the following expressions:

$$\begin{aligned}
&\frac{\partial E_1}{\partial c_{001}} = \sum_{i=1}^{1}(h_{i1} - t_{i1}); && \frac{\partial E_2}{\partial c_{002}} = \sum_{i=1}^{1}(h_{i2} - t_{i2}); \\
&\frac{\partial E_1}{\partial c_{101}} = \sum_{i=1}^{1}(h_{i1} - t_{i1})x_1; && \frac{\partial E_2}{\partial c_{102}} = \sum_{i=1}^{1}(h_{i2} - t_{i2})x_2; \\
&\frac{\partial E_1}{\partial c_{011}} = \sum_{i=1}^{1}(h_{i1} - t_{i1})y_1; && \frac{\partial E_2}{\partial c_{102}} = \sum_{i=1}^{1}(h_{i2} - t_{i2})y_2; \\
&\frac{\partial E_1}{\partial c_{111}} = \sum_{i=1}^{1}(h_{i1} - t_{i1})x_1^2; && \frac{\partial E_2}{\partial c_{112}} = \sum_{i=1}^{1}(h_{i2} - t_{i2})x_2^2; \\
&\frac{\partial E_1}{\partial c_{201}} = \sum_{i=1}^{1}(h_{i1} - t_{i1})x_1y_1; && \frac{\partial E_2}{\partial c_{202}} = \sum_{i=1}^{1}(h_{i2} - t_{i2})x_2y_2 \qquad (6) \\
&\frac{\partial E_1}{\partial c_{021}} = \sum_{i=1}^{1}(h_{i1} - t_{i1})y_1^2; && \frac{\partial E_2}{\partial c_{202}} = \sum_{i=1}^{1}(h_{i2} - t_{i2})y_2^2
\end{aligned}$$

$$\frac{\partial E_1}{\partial c_{111}} = \sum_{i=1}^{1}(h_{i1} - t_{i1})x_2y_2; \qquad \frac{\partial E_2}{\partial c_{112}} = \sum_{i=1}^{1}(h_{i2} - t_{i2})x_1y_1;$$

For $a = 1$, we obtain

$$\begin{aligned}
&\frac{\partial E_3}{\partial c_{003}} = \sum_{i=1}^{1}(h_{i3} - t_{i3}); && \frac{\partial E_3}{\partial c_{113}} = \sum_{i=1}^{1}(h_{i3} - t_{i3})x_3y_3; \\
&\frac{\partial E_3}{\partial c_{103}} = \sum_{i=1}^{1}(h_{i3} - t_{i3})x_3; && \frac{\partial E_3}{\partial c_{203}} = \sum_{i=1}^{1}(h_{i3} - t_{i3})x_3^2 \\
&\frac{\partial E_3}{\partial c_{013}} = \sum_{i=1}^{1}(h_{i3} - t_{i3})y_3; && \frac{\partial E_3}{\partial c_{023}} = \sum_{i=1}^{1}(h_{i3} - t_{i3})y_3^2 \qquad (7)
\end{aligned}$$

As a result of training (6) and (7), we find network parameters that satisfy the knowledge base with the required training quality.

Fuzzy statistical data (see Table 2) were collected from experiments before the computer simulation It should be noted that for negative values of the parameter $\widetilde{c}_{jk}(\widetilde{c}_{jk} < 0)$, the formulas that include the parameter $\widetilde{c}_{jk}$ in (5) and the correction of that parameter in (6) will have changed forms. For example, if $\widetilde{c}_{jk} < 0$, the formula for the fifth expression, which includes $\widetilde{c}_{jk}$ in (5) will have the following form: $h_{51} = c_{111}x_2y_2$; $h_{52} = c_{112}x_1y_1$, and the correction formulas was performed.

The network parameters were thus trained using the described fuzzy-neural network structure and experimental data. As a result, network-parameter values that satisfied the experimental statistical data were found (see Table 2):

$$\begin{aligned}
&\widetilde{c}_{00} = (1.4124;\ 1.4223;\ 1.4275); \\
&\widetilde{c}_{10} = (1.9884;\ 2.1131;\ 2.2339); \\
&\widetilde{c}_{01} = (-2.5353;\ -2.5349;\ -2.5326); \\
&\widetilde{c}_{20} = (-1.1043;\ -1.1042;\ -1.1036) \\
&\widetilde{c}_{11} - (-0.8845;\ -0.8741;\ -0.8639); \\
&\widetilde{c}_{02} = (1.3158;\ 1.3162;\ 1.3166).
\end{aligned}$$

These data were obtained as a result of 20-minute training of the neural network.The coefficients $\widetilde{c}_{00}, \widetilde{c}_{10}, \widetilde{c}_{01}, \widetilde{c}_{20}, \widetilde{c}_{11}, \widetilde{c}_{02}$ regression equation (5) were evaluated by a program written in Turbo Pascal on an IBM PC.

IV. Conclusions

The use of fuzzy neural networks (Soft Computing) to solve problems that involve evaluation parameters of mathematical models of geofields advantages over traditional statistical-probability approaches. Primary is the fact that the proposed procedure can be used regardless of the type of distribution of the parameters geofield. The more so because, in the early stage of modeling, it is difficult to establish the type of parameter distribution, due to insufficient data.

References

[1] M. Yanalak, Height interpolation in digital terran models. Ankara, Harita dergisi, Temmuz 2002. Sayi: 128. p. 44 – 58.
[2] A.Berlyant, L. Ushakova, Cartographic animations. Moscow: Scientific World, 2000.
[3] M. Jukov, S. Serbenyuk and V. Tikunov, Mathematical – cartographig modeling in geography. Moscow: misl, 1980.
[4] O. Akima, P. Hiroshi, Bivariate interpolation and smooth surface fitting for irregulary distributed date points. ACM Transactions for Mathematical Software. June 1978. p. 148 – 159.
[5] J. Delhhome, Kriging in the hudro sciences // Adv. Water Res. 1978. Vol. 1.№5.
[6] L.Buryakovskii, I. Dzhafarov and Dzhefanshir, Modeling systems. Moscow: Nedra, 1990.
[7] R. Yager, L. Zadeh. (Eds.) Fuzzy sets, Neural Networks and Soft Computing. Van Nostrand Reinhold – New York, 1994.
[8] H. Mohamad. Fundamentals of Artificial Neural Networks, MIT Press, Cambridge, Mass., London, 1995

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:4, No:11, 2010

APPENDIX

TABLE I
NORMALITY ASSUMPTION

$\overline{x} \approx M$	68%	95%	100%	V<33%	$E_x \to 0$	$S_k \to 0$
$0.71 \neq 0.59$ non – exe – cution	77.7% execu - tion	91.6 % non – exe – cution	100% execu - tion	47 % non – exe – cution	0.45 non – exe – cution	1.14 non – exe - cution

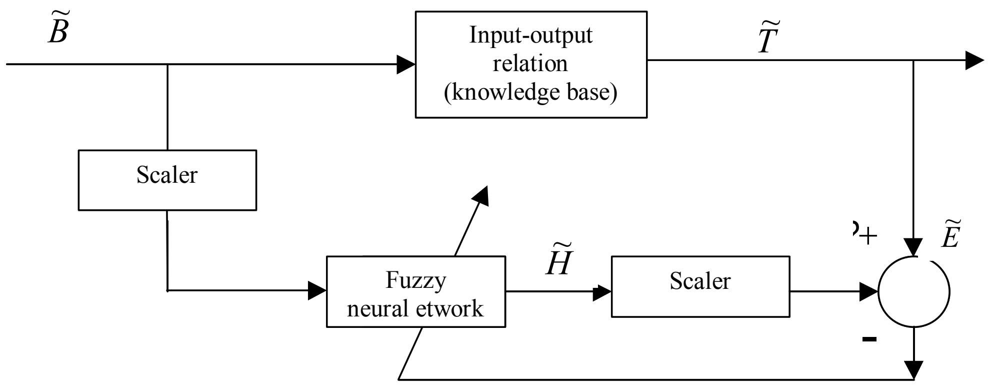

Fig. 1 Neural identification system

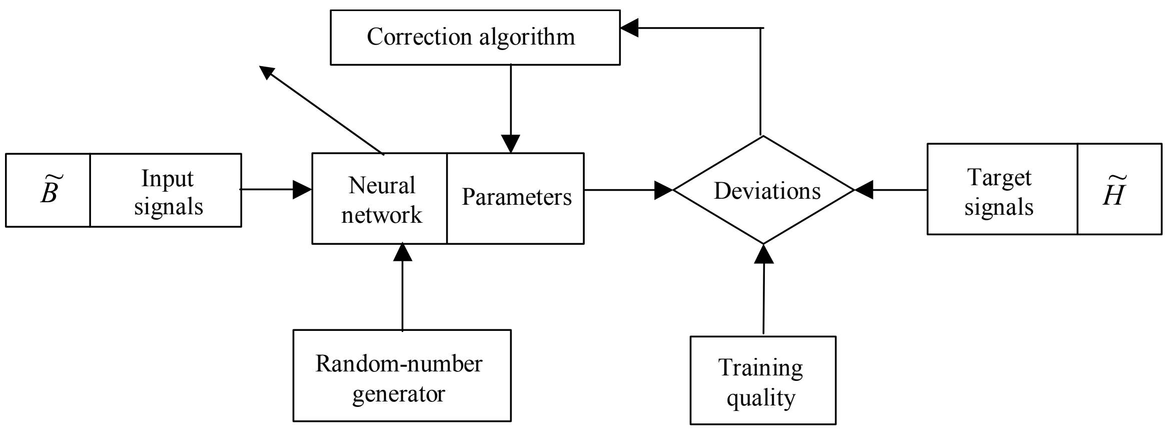

Fig. 2 System for network-parameter training (with backpropagation)

ISSN: 2517-9934
Vol:4, No:11, 2010

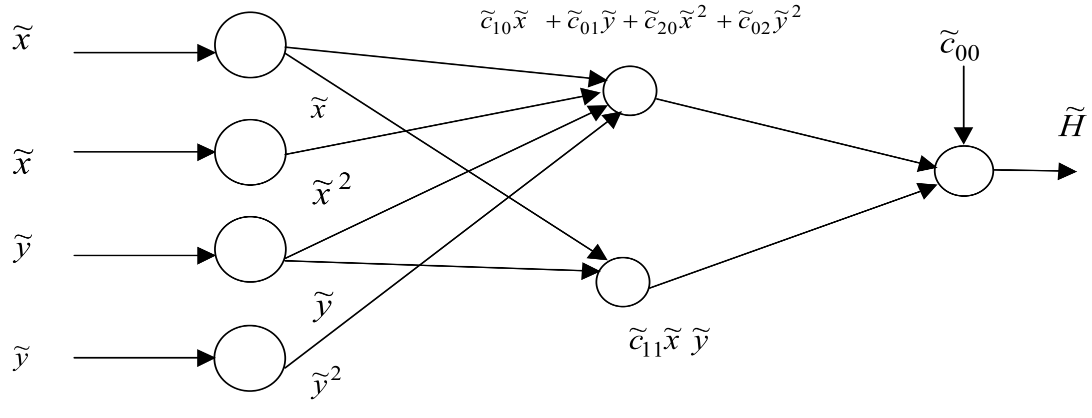

Fig. 3 Structure of neural network for second-order regression equation

TABLE II
THE EXPERIMENTAL STATISTICAL DATA

$\tilde{x}$ \ $\tilde{y}$	3,7,11	17,21,25	31,35,39	45,49,53	59,63,67	73,77,81
28,31,35	0.77,0.81,0.85	1.08,1.13,1.17	1.28,1.33,1.44	1.43,1.47,1.51	1.49,1.53,1.57	1.48,1.50,1.54
50,54,58	0.48,0.52,0.56	0.68,0.72,0.76	0.81,0.85,0.89	0.89,0.93,0.97	0.93,0.97,1.01	0.91,0.95,0.99
68,72,76	0.37,0.41,0.45	0.53,0.57,0.61	0.63,0.67,0.71	0.69,0.73,0.77	0.72,0.76,0.60	0.71,0.75,0.79
82,86,90	0.30,0.34,0.38	0.43,0.47,0.51	0.52,0.58,0.60	0.57,0.61,0.65	0.60,0.64,0.68	0.59,0.63,0.67
92,96,100	0.27,0.31,0.35	0.39,0.43,0.47	0.46,0.50,0.54	0.51,0.55,0.59	0.54,0.58,0.62	0.53,0.57,0.61
96,100,104	0.23,0.27,0.31	0.34,0.38,0.42	0.41,0.45,0.49	0.46,0.50,0.54	0.47,0.51,0.55	0.47,0.51,0.55