

# A Study on 2D similarity transformation using multilayer perceptron neural networks and a performance comparison with conventional and robust outlier detection methods

Berkant Konakoğlu<sup>1</sup> and Ertan Gökalp<sup>2</sup>

*One of such transformation methods, the two-dimensional similarity transformation, is widely used in geodetic studies. The outliers in the measure group should be detected so that the model established during the transformation process gives accurate results. In transformation practices, conventional outlier detection test procedures based on the least squares estimation (LSE) and robust estimation methods are widely used for the detection of outliers. The aim of this study was to compare performances of the result data obtained by multilayer perceptron artificial neural networks (MLPNNs) including various activation functions and training algorithms of artificial neural networks (ANNs), which has recently begun to be widely used in scientific studies and engineering fields, and of the result data obtained using various methods to detect outliers in two-dimensional similarity transformation process between two different coordinate systems. In ANNs consisting of three layers, eight different network configurations were generated using different activation functions and training algorithms. The coordinates of the control points calculated by two-dimensional similarity and ANNs methods were compared with known coordinate values. Differences between the coordinates calculated using two-dimensional similarity transformation and eight different network configurations and the coordinates of control points were examined in terms of the root mean square error, and network configuration which uses a combination of 'tansig-purelin' activation functions and Bayesian regulation algorithm provided the best result.*

**Key words:** similarity transformation, outliers, conventional methods, robust estimation, artificial neural networks

## 1. Introduction

In addition to exploration of metallic mines, industrial raw materials, natural gas, coal, oil and underground water, maps produced at various scales serve as a basis for various engineering projects such as railways, highways, subway, dams and tunnels. It is necessary to produce geological maps responding to the needs to monitor the geological features and underground resources of a country. Coordinate systems are important for maps to gain numerical meaning. Transformation between coordinate systems is a necessity for the association of geological maps produced in local or national coordinate systems (such as NAD27 and ED50) and global coordinate systems (such as ITRF-yy and WGS84). Selection of the transformation method depends on the purpose of transformation and the number of known common points in both systems (Ghilani & Wolf, 2006). In two-dimensional systems, mostly similarity, affine and projective transformation methods are used as the transformation method. In geodetic studies, excessive measurements are made to increase the accuracy of the measurements. The relationship between a large number of measurements is established by adjustment theory. Precise values and sensitivity of the measurements are calculated by least squares estimation (LSE). In addition to gross and systematic errors in the measurement group, there can also be unidentifiable measurement errors in the measurement group. Measurements containing such measurement errors in the measurement group are called outliers. Detection of outliers is usually performed by conventional outlier detection test procedures based on LSE. Disadvantages of conventional outlier detection test procedures led to the search for other methods. In literature, there are many studies on conventional outlier detection test procedures (Berberan, 1995; Kok, 1984; Pope, 1976; Schwarz & Kok, 1993; Knight & Wang, 2009). An alternative method for detecting outliers in the measurement group is robust estimation (Hampel, Ronchetti, Rousseeuw & Stahel, 1986; Huber, 1981; Jurecková & Sen, 1996). Artificial neural networks (ANNs) have recently begun to be widely used in scientific studies and engineering fields. ANNs began to be employed in Geodesy, and it is possible to see them in various studies (Schuh, Ulrich, Egger, Müller & Schwegmann, 2002; Zaletnyik, 2004; Kavzoglu & Saka, 2005; Lin, 2007; Tierra, Dalazoana & De Freitas, 2008; Gullu & Yilmaz, 2010; Turgut, 2010).

The aim of this study was to compare performances of the result data obtained by multilayer perceptron artificial neural networks (MLPNNs) including various activation functions and training algorithms, and of the result data obtained using various methods to detect outliers in two-dimensional similarity transformation process between two different coordinate systems. In the two-dimensional similarity transformation, t-test, one

<sup>1</sup> Berkant Konakoğlu, Ph.D., Karadeniz Technical University, Department of Geomatic Engineering, Division of Surveying, 61080 Trabzon/TURKEY, [bkonakoglu@ktu.edu.tr](mailto:bkonakoglu@ktu.edu.tr)

<sup>2</sup> Professor Ertan Gökalp, Karadeniz Technical University, Department of Geomatic Engineering, Division of Surveying, 61080 Trabzon/TURKEY, [ertan@ktu.edu.tr](mailto:ertan@ktu.edu.tr)

of conventional outlier tests, and Andrews, Beaton-Tukey and Danish M-Estimations, which are robust estimation methods, were employed to carry out transformation process. In ANN application, tanjant-sigmoid (tansig), logarithmic-sigmoid (logsig) and linear (purelin) were used as activation functions, while Levenberg-Marquardt and Bayesian regulation were used as training algorithms. The resulting values were analyzed in terms of the root mean square error (RMSE) of the coordinate differences.

## 2. Methods

### 2.1. Two-Dimensional Similarity Transformation

The relationship between two coordinate systems can be expressed geometrically with the aid of points known as common to both systems. Two-dimensional similarity transformation has a total of 4 parameters, including 1 scale, 1 rotation, and 2 translations. Therefore, the relationship between the two coordinate systems requires minimum 2 points known in both systems (Eq. 1). The two-dimensional similarity transformation is applied to two-dimensional coordinates between two coordinate systems by;

$$\begin{aligned} X &= ax - by + c \\ Y &= ay + bx + d \end{aligned} \quad (1)$$

where,  $x, y$  is the first coordinate system,  $X, Y$  is the second coordinate system,  $a, b, c$  and  $d$  coefficients are transformation parameters.

If the number of common points in both systems is high, transformation parameters are calculated by the method of LSE. The matrix representation can be written as (Eq. 2);

$$A = \begin{bmatrix} x_1 & -y_1 & 1 & 0 \\ y_1 & x_1 & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_n & -y_n & 1 & 0 \\ y_n & x_n & 0 & 1 \end{bmatrix} X = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} L = \begin{bmatrix} X_1 \\ Y_1 \\ \cdot \\ X_n \\ Y_n \end{bmatrix} V = \begin{bmatrix} V_{X_1} \\ V_{Y_1} \\ \cdot \\ V_{X_n} \\ V_{Y_n} \end{bmatrix} \quad (2)$$

where,  $X$  is the vector of the unknown parameters and calculated by the equation  $X = (A^T P A)^{-1} A^T P L$ ,  $A$  is the coefficients matrix,  $P$  is the weight matrix,  $V$  is the vector of residuals and  $L$  is the translation vector.

### 2.2. Conventional Outlier Detection Methods

In geodetic studies, outliers are usually detected by a conventional outlier detection test procedures based on LSE. Reasons to prefer this method include simple calculation algorithm, no need to know the statistical distribution of observations, and stability of mathematical and functional models throughout the calculation. Outliers are detected iteratively and removed from the measurement group, which means ignoring the information contained in the measurement. Due to the propagation effect of LSE, an error in any of the measurements can also be reflected in the adjustment of other measurements. In order to test whether there is any outlier in the measurement group as a result of balancing calculation;

$$\begin{aligned} H_0 &= E\{\nabla l_i\} = 0 \\ H_S &= E\{\nabla l_i\} \neq 0 \end{aligned} \quad (3)$$

zero and differ from zero as the alternative hypothesis is established for gross error (Eq. 3). Hypothesis testing is performed using adjustment values introduced to the measurements. A test value is calculated for each adjustment value. This value is compared with the limit value determined according to the degree of freedom from the table which the distribution of the measurements matches. Any measurement with a test value larger than the limit value is considered an outlier and removed from the measurement group. This process continues iteratively until there is no outlier in the measurement group. Three different approaches are used in conventional solution methods: Data-Snooping (DS), Tau and t (student) tests (Gökalp, Güngör & Boz, 2008).

If the a priori variance is not known or a value cannot be assigned to it before adjustment, the a priori variance  $m_0^2$  is calculated after adjustment.  $\bar{m}_o^2$  is the a posteriori variance calculated from the residuals free from the model errors.

Tab. 1. Test statistic, test distribution and critical value.

Test	Test Statistic	Test Distribution	Critical Value
Data-Snooping	$\frac{(P V_i)}{s_0 \sqrt{(P Q_{VV} P)_{ii}}}$	$\sim N(0,1)$	$N_{(1-a_0/2)}$
Tau	$\frac{(P V_i)}{m_0 \sqrt{(P Q_{VV} P)_{ii}}}$	$\sim \tau_{f,(1-a_0/2)}$	$\tau_{f,(1-a_0/2)}$
t	$\frac{(P V_i)}{\bar{m}_0 \sqrt{(P Q_{VV} P)_{ii}}}$	$\sim t_{f-1,(1-a_0/2)}$	$t_{f-1,(1-a_0/2)}$

In Table 1,  $P$  is the weight matrix,  $V_i$  is the residuals vector,  $s_0$  is the a priori standard deviation of unit weight,  $f$  is the degree of freedom,  $Q_{VV}$  is the cofactor matrix of the residuals,  $a_0$  is the significance level,  $N$  represents normal distribution,  $\tau$  represents tau test and  $t$  represents student test. The significance level is usually chosen as 5 % (Ghilani & Wolf, 2006; Biacs, Krakiwsky & Lapucha, 1990).

### 2.3. Robust Estimation Methods

LSE is mostly criticized due to its sensitivity to measurements with a gross error. In this case, robust estimation methods capable of detecting outliers in the measurement group in a more effective way are preferred. Robust estimator functions can be typically classified as M-, R- and L-Estimators used in statistical analyses. As M-Estimators, there are many methods in the literature. Normal distribution of geodetic methods facilitates widespread use of M-Estimators, which require the knowledge of premise distributions (Ince, Konak & Arslan, 2008).

In robust estimation, rather than the score function  $V^T P V = \min$  used in the LSE, another  $\rho(V)$  which is less affected by the errors of corrections is selected as the score function. A solution that makes the selected score function minimum is sought. Solution process is performed as in equation (Eq. 4).

$$\sum_{i=1}^n \rho(V_i) = \min \tag{4}$$

The condition in which the score function is minimum is obtained by taking the derivative with respect to  $V$  values in the equation. The solution of resulting non-linear equation is iterative. In robust estimation,  $\Psi(V_i) = \partial \rho(V_i) / \partial V_i$  influence function is obtained by taking the derivative of the score function  $\rho(V_i)$  with respect to corrections. Weight function  $W(V) = \Psi(V) / V$  is obtained by dividing the influence function by corrections. The estimation process is performed as in equation (Eq. 5);

$$\begin{aligned} \hat{X}_i &= (A^T \bar{W}_i A)^{-1} A^T W_i L \\ V_i &= A \hat{X}_i - L \\ \bar{W}_i &= P W_{i-1}, \quad i = 1, 2, \dots, n \end{aligned} \tag{5}$$

where,  $\bar{W}_i$  is the equivalent weight matrix,  $P$  is the first weight matrix of the measurements,  $W_i$  is the robust weight factor,  $n$  is the number of measurements and  $i$  is the iteration number. The weight matrix  $P$  in the first iteration can either be a unit matrix or estimated premise weight matrix. The iterations continue until the difference between  $\hat{X}_{i+1}$  and  $\hat{X}_i$  are negligible. As a result of the iterations, equivalent weights of outliers become smaller, or even reduced to zero. This means that the negative effect of the outliers in the measurement group of the unknowns decreases (Gökalp et al., 2008). The reliability of the results obtained from robust estimations depends on accurate selection as much as possible of the critical value with which the correction values of the measurements are compared (Tab. 2). In the literature search, it was found that the limit parameter value  $c$  is not a fixed number, and it was concluded that the limit value is selected based on experiences.

Tab. 2. The weight function of the M-Estimations.

M-Estimation	Weight Function	Critical Value
Andrews	$W_i = \begin{cases} \sin \frac{V_i}{c} \left(\frac{V_i}{c}\right)^{-1} \\ 0 \end{cases}$	$ V_i  \leq c\pi$ $ V_i  > c\pi$
Beaton-Tukey	$W_i = \begin{cases} \left[1 - \left(\frac{V_i}{c}\right)^2\right]^2 \\ 0 \end{cases}$	$ V_i  \leq c$ $ V_i  > c$
Danish	$W_i = \begin{cases} \exp\left(-\frac{V_i^2}{c^2}\right) \\ 1 \end{cases}$	$ V_i  > c$ $ V_i  \leq c$

## 2.4. Artificial Neural Networks

Artificial neural networks can be defined as computer programs that mimic the human brain, which has the ability to process even non-linear information at a high-performance. In addition, ANNs can be considered as a parallel information processing system in a sense. In contrast to the complexity of the human brain, they have a much simpler structure. Nerve cells are grouped into layers, creating ANNs.

Multilayer perceptron neural networks (MLPNNs) are the most commonly used feed-forward neural networks. A MLPNN consists of three types of layers: the input layer, one or more hidden layers and the output layer, as shown in Fig.1. The input layer is the layer where data is presented to the network. The number of neurons in this layer equals to the number of input data. In this layer, the data obtained from outside are not processed; the layer is only obliged to write the data and communicate it to the hidden layer. Hidden layer is the layer where the data from the input layer is processed. In some networks, the number of hidden layers and the number of neurons in the hidden layer may change depending on the problem. Output layer is the final layer of the network. It provides output by processing the data received from the hidden layer with the function used. The number of neurons equals the number of data presented in input layer at the output. It is the layer that generates the result value using the data from the hidden layer (Hornik, Stinchcombe & White, 1989).

Weights are the parameters indicating the importance and effect in the cell of data entering the network. Each input data has a weight of its own. Hence, each input data is multiplied by its own weight; the product is added to error value and sent to the activation function in order to calculate the output value (Eq. 6);

$$N_i = \left( \sum_{j=1}^{n_{k-1}} W_{ij}^k x_i^{k-1} + W_0 \right) \quad (6)$$

where,  $x_i$  is the input neuron,  $W_{ij}^k$  is the weight coefficient of each input neuron,  $W_0$  is the bias,  $i$  is the unit number of  $(k - 1)^{\text{th}}$  layer and  $j$  is the unit number of  $k^{\text{th}}$  layer.

The back-propagation (BP) algorithm reduces the errors from output to input, which is why it is called “back-propagation algorithm”. It is an easy to understand and preferred learning algorithm. The algorithm is used to recalculate the weights in each layer according to the existing level of error in the network output. BP algorithm is generally used for multilayer networks, rather than single layer networks (Fig. 1).

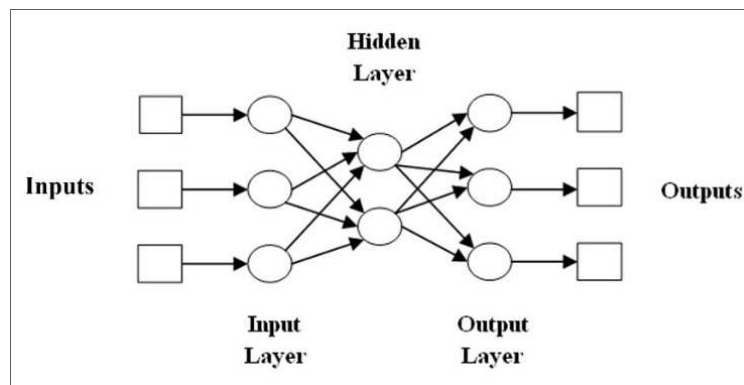


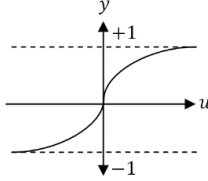
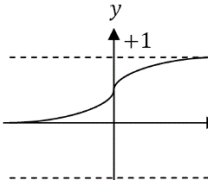
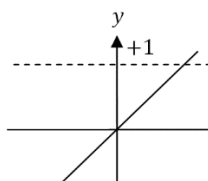
Fig.1. A typical example of a multilayer neural network.

The relationship between input data and output data is based on training the network and determining the optimum values of the weight values of the network. Training algorithms calculate the weights which will minimize differences between output values produced by the network  $N_i$  and the actual values  $N_{actual}$ . The error of the network is calculated by Eq. 7 and then weights and bias are updated iteratively until the error approaches the given fair value (Basheer & Hajmeer, 2000; Abdalla, Zakaria, Sulaiman & Ahmad, 2010).

$$E = \frac{1}{2} \sum_{i=1}^{n_0} (N_{actual} - N_i) \quad (7)$$

where,  $n_0$  is the number of output units. Activation functions frequently used in the literature include hyperbolic tangent sigmoid, logarithmic sigmoid and linear activation functions (Haykin, 1999) (see Tab. 3).

Tab. 3. Some activation functions commonly used in artificial neural network applications.

Function Name	Function Type	Mathematical Expression	Explanation
Hyperbolic tangent sigmoid function		$f(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}}$	The output values of the hyperbolic tangent function range from -1 to 1.
Logarithmic sigmoid function		$f(u) = \frac{1}{1 + e^{-\beta u}}$	The output values of the logarithmic sigmoid function range from 0 to 1.
Linear function		$f(u) = au$	With $a$ a fixed number, the output values of linear function range from -1 to 1.

### 3. Materials and Methodology

In this study, the new geodetic network has been created to examine the above-described methods. The coordinates of points on a network created for this purpose were obtained using the RTK-GPS method in the ED50 and ITRF96 coordinate systems. In the created geodetic network, geodetic points with known coordinates in both systems were divided into two groups. Of these, 47 geodetic points, which were evenly distributed over the study area, used as reference points, while remaining 12 geodetic points were used as control points to be used for performance analysis (Fig. 2).

Performance reliability was assessed by the differences (Eq. 8) between known values of the coordinates and actual coordinate values of the results obtained from the transformations performed using t-test, M-Estimation methods, and ANNs.

$$\Delta_{X,Y} = (X, Y)_{known} - (X, Y)_{calculated} \tag{8}$$

Root mean square error values (Eq. 9) were calculated for the statistical analysis ( $n$ : The number of measurements).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\Delta_{X,Y})^2} \tag{9}$$

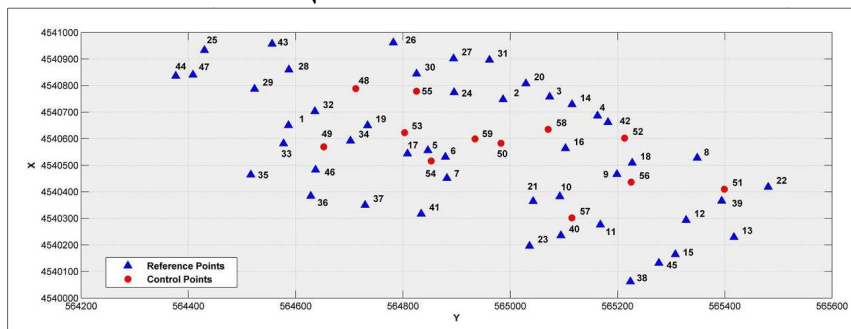


Fig. 2. Distribution of geodetic points.

The two-dimensional similarity transformation of the coordinates common in two coordinate systems was performed using LSE. Outliers in the measurement group were detected iteratively using t-test and removed from geodetic network. During detection of outliers, when  $\alpha_0 = 0.05$  was selected as the significance level, outliers were detected at the end of 8 iterations and 12 measurements were removed from the measurement group. When  $\alpha_0 = 0.01$  was selected as the significance level, outliers were detected as a result of 4 iterations and 6 measurements were removed from the measurement group. Selecting 0.01 as the significance level detected 6 measurements as not outlier, as compared to selecting 0.05 as the significance level (Tab. 4).

Tab. 4. Analysis of outliers using conventional solution methods.

Number of Iterations	Significance Level=0.05			Significance Level=0.01		
	Max Test Statistic	Critical Value	Point ID (Outlier)	Max Test Statistic	Critical Value	Point ID (Outlier)
I.	7.561572	1.661961	7,21	7.561572	2.368497	7,21
II.	5.275576	1.662765	18,19,20	5.275576	2.370493	18,19,20
III.	7.505862	1.664124	6	7.505862	2.373868	6
IV.	1.847396	1.664625	13,29,41	1.847396	2.375111	-
V.	1.903764	1.666294	15	-	-	-
VI.	1.708086	1.666914	23	-	-	-
VII.	1.750083	1.667572	8	-	-	-
VIII.	1.608736	1.668271	-	-	-	-

Efforts were made to iteratively detecting outliers in the measurement group using robust estimation methods. Andrews, Beaton-Tukey and Danish M-Estimations were used as estimation methods. The results show that all three robust methods iteratively detected point 7 and point 21 as outliers, whereas Beaton-Tukey and Denmark M-Estimations established that the weight of point 6 has changed but detected it as suspicious, rather than an outlier (Tab. 5).

Tab. 5. Outliers analysis by robust estimation methods.

Point ID	$V_i$	Robust Estimations			Result
		Andrew $W_i$	Beaton-Tukey $W_i$	Danish $W_i$	
7	5.650927	0.157	0.021	0.000	Outlier
21	8.395877	0.111	0.003	0.000	Outlier
6	1.078722	0.916	0.503	0.312	Suspicious

There are different training algorithms in the literature. In this article, two high-performance training algorithms that improve performance were studied using different optimization techniques. These algorithms were Levenberg-Marquardt and Bayesian regulation algorithms. Levenberg-Marquardt algorithm is based on the maximum neighborly least squares calculation method. Gauss-Newton and gradient descent algorithms containing Levenberg-Marquardt algorithm in the best properties not affected by the slow convergence problem. Also, it is faster than other methods in terms of processing time (Hagan & Menhaj, 1994). Bayesian regulation algorithm derived from the Levenberg-Marquardt algorithm, the Bayesian regulation algorithm updates weight and bias values. This method minimizes the combination of square error and weight. Also, it determines the correct combination to produce network (Mackay, 1992).

In MLPNNs, back-propagation algorithm was employed, with ED50 coordinate system in input layer as datum and ITRF96 coordinate system in the output layer. In an application network of three layers, various activation functions and training algorithms were utilized. Eight different network configurations were created by activation functions and training algorithms (Tab. 6). Concerning training parameters, MATLAB software was used, and learning rate was accepted as 0.1 and momentum coefficient as 0.9, and an error of 0.00001 was accepted during the training phase of the network. As a result of the trial-and-errors made in all network configurations, it was decided to create a network architecture with 2 neurons in the input layer, 8 neurons in the hidden layer and 2 neurons in the output layer [2:8:2].

Tab. 6. Network configurations created for training.

Network Configuration	Training Algorithm	I. Function	II. Function
I.	Levenberg-Marquardt	tansig	purelin
II.	Bayesian Regulation	tansig	purelin
III.	Levenberg-Marquardt	logsig	purelin
IV.	Bayesian Regulation	logsig	purelin
V.	Levenberg-Marquardt	tansig	logsig
VI.	Bayesian Regulation	tansig	logsig
VII.	Levenberg-Marquardt	logsig	tansig
VIII.	Bayesian Regulation	logsig	tansig

RMSE values (Fig. 3) were calculated using the differences between the coordinate values (Tab. 7, 8) obtained as a result of the transformation and the calculated coordinate values of control points.

Tab. 7. Differences between the coordinate values obtained as a result of the transformation and the known X coordinate values of control points (meters).

$\Delta X$	Point ID											
	48	49	50	51	52	53	54	55	56	57	58	59
<b>2D Similarity Transformation (t)</b>	0.012	0.011	0.015	0.014	-0.020	0.014	0.023	0.013	-0.033	0.017	0.013	0.010
<b>2D Similarity Transformation (Andrew)</b>	0.008	0.005	0.016	0.022	-0.015	0.011	0.021	0.011	-0.028	0.018	0.016	0.009
<b>2D Similarity Transformation (Beaton-Tukey)</b>	0.008	0.005	0.016	0.022	-0.015	0.011	0.021	0.011	-0.028	0.018	0.016	0.009
<b>2D Similarity Transformation (Danish)</b>	0.008	0.005	0.016	0.022	-0.015	0.011	0.021	0.011	-0.028	0.018	0.016	0.009
<b>I.</b>	-0.004	0.019	-0.006	0.015	-0.015	-0.012	-0.011	0.006	-0.006	0.020	0.021	0.000
<b>II.</b>	0.008	0.015	0.011	0.018	0.015	0.017	-0.015	0.010	0.009	0.014	0.010	0.009
<b>III.</b>	0.011	0.005	0.027	0.021	-0.026	0.018	0.017	0.011	0.015	-0.029	0.018	0.003
<b>IV.</b>	-0.006	0.023	0.017	0.019	-0.019	0.015	-0.009	-0.007	-0.011	0.019	0.007	0.010
<b>V.</b>	0.001	0.013	-0.001	0.019	-0.022	0.031	-0.020	0.006	-0.016	0.028	0.007	-0.008
<b>VI.</b>	0.022	0.007	0.021	-0.009	-0.013	0.016	-0.009	0.013	-0.010	0.014	0.008	0.014
<b>VII.</b>	-0.034	-0.030	0.026	-0.033	-0.024	0.030	0.044	0.058	-0.021	0.032	-0.029	0.025
<b>VIII.</b>	0.036	0.015	0.012	0.023	-0.032	0.048	-0.022	0.022	0.034	0.024	0.018	0.024

Tab. 8. Differences between the coordinate values obtained as a result of the transformation and the known Y coordinate values of control points (meters).

$\Delta Y$	Point ID											
	48	49	50	51	52	53	54	55	56	57	58	59
<b>2D Similarity Transformation (t)</b>	0.010	0.013	0.009	0.011	-0.015	0.019	0.011	0.013	-0.017	0.024	0.016	0.013
<b>2D Similarity Transformation (Andrew)</b>	-0.002	0.008	0.005	0.013	-0.017	0.014	0.009	0.005	-0.023	0.024	0.012	0.009
<b>2D Similarity Transformation (Beaton-Tukey)</b>	-0.002	0.008	0.005	0.013	-0.017	0.014	0.009	0.005	-0.023	0.024	0.012	0.009
<b>2D Similarity Transformation (Danish)</b>	-0.002	0.008	0.005	0.013	-0.017	0.014	0.009	0.005	-0.023	0.024	0.012	0.009
<b>I.</b>	0.013	0.014	0.012	0.015	-0.018	-0.010	-0.019	0.012	-0.011	0.014	0.015	0.010
<b>II.</b>	-0.004	0.019	-0.006	0.015	-0.015	-0.012	-0.011	0.006	-0.006	0.013	0.021	0.007
<b>III.</b>	0.012	0.013	0.004	0.019	0.031	0.023	0.013	0.006	-0.014	0.017	0.026	0.007
<b>IV.</b>	0.014	0.001	0.018	0.022	-0.016	0.023	0.008	0.010	-0.015	0.013	0.018	0.022
<b>V.</b>	0.020	0.005	0.022	0.026	-0.038	0.022	0.009	0.022	-0.030	0.022	-0.005	0.016
<b>VI.</b>	0.002	0.017	0.008	0.012	-0.024	0.015	-0.005	0.017	-0.004	0.017	0.016	0.018
<b>VII.</b>	0.009	-0.037	-0.012	-0.025	-0.028	0.033	-0.036	0.021	0.006	0.027	0.032	-0.025
<b>VIII.</b>	0.037	0.019	0.016	0.023	-0.023	0.030	-0.015	0.021	0.034	0.045	0.006	0.033

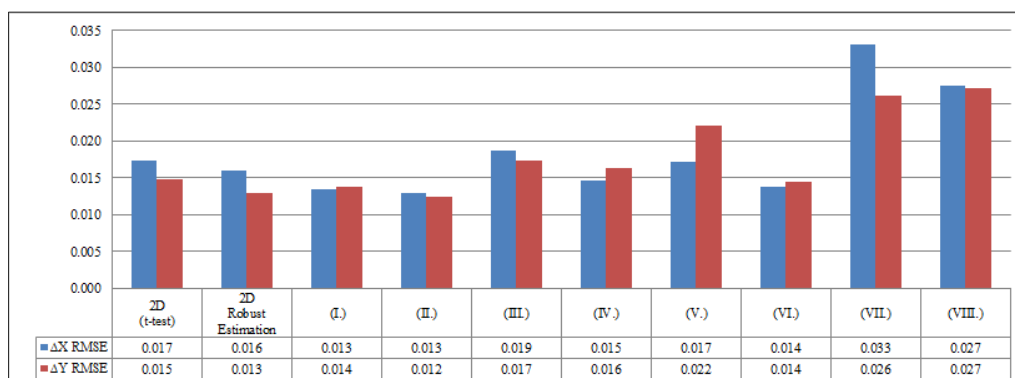


Fig. 3. RMSE values of the coordinate differences of control points (meters).

The RMSE of the coordinate differences revealed that, in both directions ( $X, Y$ ), the best result was obtained for the network configuration (II) using a combination of 'tansig-purelin' activation functions and Bayesian regulation algorithm. The worst result was obtained for the network configuration (VII) using a combination of 'logsig-tansig' activation functions and Levenberg-Marquardt algorithm (Fig. 3).

#### 4. Conclusion

Coordinate transformation is a necessity when there is a need for the collocation of different geological maps with different coordinate systems. Different transformation methods can be used depending on the shape of the calculation surface, purpose of transformation and the number of common points in both systems. This study use two-dimensional coordinate transformation method was used. In two-dimensional coordinate transformation applications, outliers are detected by a conventional outlier detection test procedures based on LSE due to its advantages. Disadvantages of LSE, including the effect of propagation and iterative removal of measurements from the measurement group have led to the search for other statistical methods. Alternatively, M-estimation methods are used for outlier detection. In this study, the applicability of ANN models was tested by creating network configurations using various training algorithms and activation functions, in addition to the methods mentioned.

A study area with 59 geodetic points with known coordinates in ED50 and ITRF96 coordinate systems were employed for the study. Of these points, 47 were selected as reference points, and remaining 12 as control points were used in performance analysis. In this context, according to two-dimensional similarity transformation, RMSE values obtained for ( $X, Y$ ) using t-test, a method for detection of outliers, were  $\pm 0.017$  m.,  $\pm 0.015$  m, respectively, whereas those obtained using robust estimation, which localizes and even eliminates disruptive effects of outliers, were  $\pm 0.016$  m and  $\pm 0.013$  m, respectively. According to the transformation conducted using ANNs, the best result was obtained for the network configuration (II) using a combination of 'tansig-purelin' activation functions and Bayesian regulation algorithm, compared to other network configurations and two-dimensional similarity transformation methods. The results are valid for the data sets used in the study so generalizations cannot be made.

The advantages of the ANNs model used in the study: 1) No need for any mathematical model, 2) Ability to immediately reveal the relationship based on the sample data presented to ANNs, 3) Ability to be trained iteratively with the required sensitivity to produce solutions suitable for the system. The disadvantages of the ANN model: 1) That the network structure applicable to the problem is made by trial-and-error method, 2) Lack of a rule for determining learning coefficient and the number of layers, which are parameters of the network.

In conclusion, the ANNs were able to demonstrate the relationship between two systems according to the two-dimensional similarity transformation method. Therefore, in addition to geodetic coordinate transformation methods, utilization of an ANN can be recommended to researchers for further similar studies. Moreover, in further studies, the subject can be addressed in different coordinate transformation models and coordinate data with a wider distribution.

#### References

Abdalla, O. A., Zakaria, M. N., Sulaiman, S., & Ahmad, W. F. W.: A comparison of feed-forward back-propagation and radial basis artificial neural networks: A Monte Carlo study. *2010 Information Technology*. doi:10.1109/itsim.20105561599



- Basheer, I. A., & Hajmeer, M.: Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43(1), 2000, 3-31. doi:10.1016/S0167-7012(00)00201-3
- Berberan, A.: Multiple Outlier Detection. A Real Case Study. *Survey review*, 33(255), 41-49. doi:10.1179/003962695791624849
- Biacs, Z. F., Krakiwsky, E. J., & Lapucha, D.: Reliability Analysis of Phase Observations in GPS Baseline Estimation. *Journal of Survey Engineering*, 116(4), 204-224. 1990, doi:10.1061/(asce)0733-9453(1990)116:4(204)
- Ghilani, C. D., & Wolf, P. R.: Adjustment computations: spatial data analysis. Hoboken: *John Wiley & Sons* 2006.
- Gökcalp, E., Güngör, O., & Boz, Y.: Evaluation of Different Outlier Detection Methods for GPS Networks. *Sensors*, 8(11), 7344-7358, 2008, doi: 10.3390/s8117344
- Gullu, M., & Yilmaz, I.: Outlier detection for geodetic nets using ADALINE learning algorithm. *Scientific Research and Essays*, 5(5), 440-447, 2010.
- Hagan, M. T., & Menhaj, M. B.: Training feedforward networks with the Marquardt algorithm. *IEEE transactions on Neural Networks*, 5(6), 989-993, 1994, doi:10.1109/72.329697
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W.A.: Robust Statistics: The approach based on influence functions. *New York: John Wiley and Sons*, 1986.
- Haykin, S.: Neural networks: A comprehensive foundation, *Prentice Hall PTR Upper Saddle River, NJ, USA*, 1999.
- Hornik, K., Stinchcombe, M., & White, H.: Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359-366, 1989, doi:10.1016/0893-6080(89)90020-8
- Huber, P. J.: Robust statistics. *Wiley Series in Probability and Mathematical Statistics*, 1981, doi:10.1002/0471725250
- Ince, C., Konak, H., & Arslan, O.: Improving of Spatial Data Obtained from Remote Sensing Images for Monitoring of Earthquake Damages, *International Earthquake Symposium*, 526-536, 2007.
- Mackay, D. J.: Bayesian Interpolation, *Neural computation*, 4(3), 415-447, 1992, doi:10.1162/neco.1992.4.3.415
- Jurecková, J., Sen, P. K.: Robust Statistical Procedures: Asymptotics and interrelations. *New York: Wiley*, 1996.
- Kavzoglu, T., Saka, M. H.: Modelling local GPS/levelling geoid undulations using artificial neural networks. *Journal of Geodesy*, 78(9), 520-527, 2005, doi:10.1007/s00190-004-0420-3
- Knight, N. L., & Wang, J.: A Comparison of Outlier Detection Procedures and Robust Estimation Methods in GPS Positioning, *Journal of Navigation*, 62(04), 699-709, 2009, doi: 10.1017/s0373463309990142
- Kok, J. J.: On data snooping and multiple outlier testing. Rockville, MD: U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, *National Ocean Service, Charting and Geodetic Services*, 1984.
- Lin, L. S.: Application of a Back-Propagation Artificial Neural Network to Regional Grid-Based Geoid Model Generation Using GPS and Leveling data. *Journal of Survey Engineering*, 133(2), 81-89. 2007, doi:10.1061/(asce)0733-9453(2007)133:2(81)
- Pope, A. J.: The statistics of residuals and the detection of outliers. Rockville, MD: Geodetic Research and Development Laboratory, *National Geodetic Survey, National Ocean Survey*, 1976.
- Schuh, H., Ulrich, M., Egger, D., Müller, J., & Schwegmann, W. (2002). Prediction of Earth orientation parameters by artificial neural networks. *Journal of Geodesy*, 76(5), 247-258. doi: 10.1007/s00190-001-0242-5
- Schwarz, C. R., Kok, J. J.: Blunder Detection and Data Snooping in LS and Robust Adjustments. *Journal of Surveying Engineering*, 119(4), 127-136. 1993, doi: 10.1061/(asce)0733-9453(1993)119:4(127)
- Tierra, A., Dalazoana, R., De Freitas, S.: Using artificial neural network to improve the transformation of coordinates between classical geodetic reference frames. *Computers & Geosciences*, 34(3), 181-189, 2008, doi: 10.1016/j.cageo.2007.03.011
- Turgut, B.: A back-propagation artificial neural network approach for three-dimensional coordinate transformation. *Scientific Research and Essays*. 5(21), 2010, 3330-3335.
- Zaletnyik, P.: Coordinate transformation with neural networks and with polynomials in Hungary. In *Proceedings of International Symposium on Modern Technologies, Education and Professional Practice in Geodesy and Related Fields, Sofia, Bulgaria*, 2004.