

# A comparative study for the estimation of geodetic point velocity by artificial neural networks

M YILMAZ\* and M GULLU

*Department of Geomatics, Faculty of Engineering, Afyon Kocatepe University, ANS Campus, Gazligol Road, TR-03200, Afyonkarahisar, Turkey.*

*\*Corresponding author. e-mail: mustafayilmaz@aku.edu.tr*

Space geodesy era provides velocity information which results in the positioning of geodetic points by considering the time evolution. The geodetic point positions on the Earth's surface change over time due to plate tectonics, and these changes have to be accounted for geodetic purposes. The velocity field of geodetic network is determined from GPS sessions. Velocities of the new structured geodetic points within the geodetic network are estimated from this velocity field by the interpolation methods. In this study, the utility of Artificial Neural Networks (ANN) widely applied in diverse fields of science is investigated in order to estimate the geodetic point velocities. Back Propagation Artificial Neural Network (BPANN) and Radial Basis Function Neural Network (RBFNN) are used to estimate the geodetic point velocities. In order to evaluate the performance of ANNs, the velocities are also interpolated by Kriging (KRIG) method. The results are compared in terms of the root mean square error (RMSE) over five different geodetic networks. It was concluded that the estimation of geodetic point velocity by BPANN is more effective and accurate than by KRIG when the points to be estimated are more than the points known.

---

## 1. Introduction

Geodesy was defined by Helmert (1880) as the science devoted to measuring and mapping the Earth's surface. Further to this effectual definition, the scope of geodesy has been extended, especially due to space-based techniques allowing geodesy to determine the parameters of Earth system with high accuracy. Today, geodesy is the science of determining the geometry, gravity field, and rotation of the Earth and their evolution in time. This understanding of modern geodesy is based on the definition of the three pillars of geodesy: (1) geokinematics, (2) Earth rotation and (3) gravity field (Plag *et al.* 2009). Geokinematics (geometry and kinematics) refers to determine and monitor with utmost precision the geometric shape of the Earth (land, ice and ocean surface) as well as

its variations with time. This pillar of geodesy addresses the problems of the determination of precise three-dimensional object positions from global to local spatial scale and their changes in time. In recent decades, comprehensive efforts have been made to determine these time variations which have become possible owing to the accuracy of new space-geodetic methods, and also owing to a truly global reference system that only space geodesy can realise (Altamimi *et al.* 2002). Defining a suitable reference system and its realisation as a reference frame is a demanding endeavour to consider the time variation in geokinematics. The geodetic reference frames are the basis for providing means to assign three-dimensional coordinates to points as a function of time in global, regional and national geodetic reference networks. One of the major functions of geodesy is the establishment and

**Keywords.** Geodetic point velocity; artificial neural networks; back propagation; radial basis function; Kriging.

maintenance of geodetic reference networks for the presentation of geospatial positioning. Geodetic reference networks are comprised of a set of properly-defined and constructed points distributed on the surface of the Earth to materialise the reference systems to support sub-millimetre global change measurements over space, time and evolving technologies (Pearlman *et al.* 2006).

According to the theory of plate tectonics, the Earth's surface is in constant motion. As such, the points in the geodetic networks are not static entities. The crustal motion of the Earth's surface emanated from tectonic plate movements and displacements associated with earthquakes can cause the geodetic points to move at predictable rates from year to year. These kinematic effects on the geodetic points can be defined by a phenomenon called point velocity (Mathews and Biediger 2012). The most demanding scientific and non-scientific requirements concerning positioning are not only the increase in accuracy and temporal stability, but also high spatial and temporal resolution and low latency (Gross *et al.* 2007). In order to meet these demands, it is vital to determine velocities for all geodetic points on the Earth's surface. The widespread availability of GPS equipment provides accurate velocity information which results in the determination of precise three-dimensional point coordinates in geodetic networks by considering the time evolution. The spatial positions of geodetic points on the Earth's surface change over time due to the plate tectonics, and therefore, they are dependent on the epoch of their determination. All these spatial changes have paramount importance in geodetic applications. If we have GPS measurements at least in two epochs, it is possible to compute the change of geodetic point coordinates. Otherwise, a continuous contemporary velocity field of geodetic network is essential. The velocity field of geodetic network is determined from the periodic GPS measurements. The velocities of the new structured geodetic points (e.g., densification points) within the geodetic network are estimated from the velocity field of geodetic network or from the velocities of the existing geodetic points determined by campaign type repeated GPS sessions (static GPS surveying for 8–24 h) by the interpolation methods. The geodetic point velocities derived from measured time series of positions are used as the basic parameter in geodetic and geophysical applications including velocity field determination of geodetic networks, kinematic modelling of crustal movements, understanding plate boundary dynamics, and monitoring global sea level change (Yilmaz 2012). The estimation of an accurate geodetic point

velocity has, therefore, great importance in geosciences.

The velocity field determination has been investigated by several researchers (e.g., Demir and Acikgoz 2000; Nocquet and Calais 2003; Perez *et al.* 2003; D'Anastasio *et al.* 2006; Hefty 2008; Novotny and Kostecky 2008; Aktug *et al.* 2011). The velocity information has been used by several researchers in crustal movements, plate boundary dynamics, seismic site characterization and deformation kinematics (e.g., McClusky *et al.* 2000; Delacou *et al.* 2008; Hackl *et al.* 2009; Kanli 2009; Perez-Pena *et al.* 2010; Foti *et al.* 2011; Pinna *et al.* 2011).

Artificial neural network (ANN) can be viewed as a computational method that is a highly simplified model of learning, interpretation and decision-making processes presented in human biological nerve systems, and it is formed by layers of interconnected artificial neurons which transform the input data into associated output data. ANNs have been applied in diverse fields of science and engineering for several types of functions such as estimation, modelling, classification, prediction, filtering, and optimisation because of their major advantages (i.e., non-parametric nature, tolerance to noisy data, applicability to complex data, arbitrary decision making capabilities and incorporation of different types of data) (Yilmaz 2012). Many geophysical phenomena are described as self-affine fractals characterized by coefficients that can be calculated by various methods such as wavelet transform, power spectrum and rescaled range analysis, etc. In practice, the geophysical data are of limited duration with gaps or noises and non-stationary, and the calculation of coefficient is not reliable for short or noisy time series (Chamoli *et al.* 2007). ANNs gave a new dimension for solving complex geophysical problems. Neural-based methods are well equipped to deal with the real world problem of non-stationarity and non-linearity (Dimri and Chamoli 2008). ANNs have been found to be effective in identifying the complex behaviour of most geophysical data which, by their very nature, exhibits extreme variability (Shahin *et al.* 2008) and have the ability to analyse non-stationary geophysical data like wavelet transforms. In geophysical and geodetic applications, the data are assumed to have a normal distribution, but in real life problems this assumption is not always realistic as dataset might show abnormal and highly skewed distribution. ANNs provide linear and non-linear mapping between input and output spatial data by its non-parametric nature which assumes no *a priori* knowledge (as in traditional regression

models), particularly of the frequency distribution of the data. This provides ANN a unique advantage over other statistical and conventional prediction techniques such as regression and interpolation methods (i.e., Pariente 1994; Kumar 2005; Singh *et al.* 2005; Karabork *et al.* 2008; Erol and Erol 2013).

ANN applications in geophysical and geodetic velocity modelling problems have increased in the last decade. Calderon-Macias *et al.* (2000) have used ANN in order to obtain 1D velocity models from seismic waveform data. Eskandari *et al.* (2004) compared multiple regression and ANN to predict shear wave velocity in the seismic exploration. Baronian *et al.* (2007) described an ANN approach for seismic velocity analysis. Peak ground velocities are used as input data in ANN application in the seismic design of deep tunnels by Ornthammarath *et al.* (2008). Moghtased-Azar and Zaletnyik (2009) have compared the ability of ANNs and polynomials for modelling the crustal velocity field. Gullu *et al.* (2011) have applied ANN for the velocity estimation of the points in a local geodetic network.

The main objective of this study is to evaluate the utility of ANN in order to estimate the velocities of the points as an alternative tool for the conventional methods in a regional geodetic network. The development and optimisation of ANN are searched to obtain the best model configuration for the geodetic velocity estimation prediction. There are numerous kinds of neural networks. However, two different types of ANN that have been more widely applied among all other ANN applications are back propagation artificial neural networks (BPANN) and radial basis function neural networks (RBFNN), which are used to estimate the geodetic point velocities in this study. In order to evaluate the performance of BPANN and RBFNN, the point velocities are also estimated by Kriging (KRIG) interpolation method, and the results are compared in terms of the root mean square error (RMSE) over five different geodetic networks in the study area. The general scheme of this paper is organised as follows: In section 2, the theoretical aspects of BPANN, RBFNN, and KRIG are described. Section 3 outlines the study area, the structured geodetic networks, the point velocity data used and the evaluation methodology. The detailed information about the design and the optimisation of ANN is given in section 4. Section 5 is concerned with the case study. The results and conclusions of ANN's utility for geodetic point velocity estimation are presented in section 6 to motivate further studies.

## 2. Theoretical aspects

Two supervised and feed-forward ANN types, BPANN and RBFNN were used in artificial neural approach of this study. A commonly applied method for spatial data, KRIG, was used in interpolation approach. The detailed theoretical information about these methods is given below.

### 2.1 Back propagation artificial neural network

BPANN (Werbos 1974; Rumelhart *et al.* 1986) is a widely used and effective multilayer perceptron (MLP) model due to their simple implementation and flexibility for a wide spectrum of problems in many application areas varying from military purposes to finance, medicine, engineering and space sciences. BPANN consists of (i) an input layer with neurons representing input variables to the problem, (ii) one or more hidden layers containing neurons to help capture the nonlinearity in the data and (iii) an output layer with neurons representing the dependent variables. The architecture of a simple BPANN is shown in figure 1. All inter-neuron connections have been associated by means of synaptic weights that are adjusted by an iterative back propagation algorithm known as training process. The introduction of back propagation algorithm has overcome the drawback of previous ANN algorithm of 1970s where the single layer perceptron failed to solve a simple XOR (Exclusive OR) problem. After the training procedure, an activation function is applied to all neurons to generate the output information (Leandro

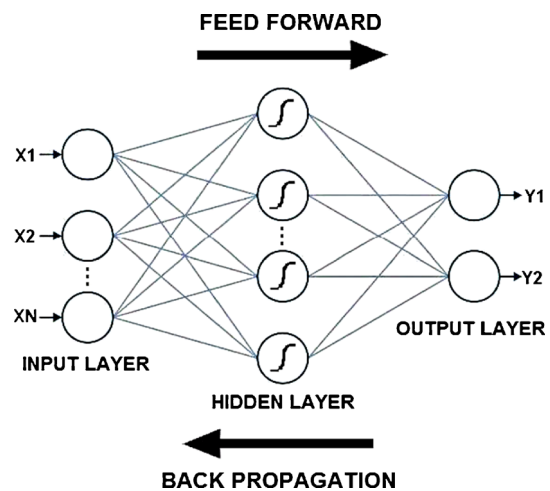


Figure 1. The BPANN architecture.

and Santos 2007) within a permissible amplitude range.

The output of BPANN with a single output neuron (output layer represented by only one neuron, i.e.,  $n = 1$ ) can be expressed according to Nørgaard (1997) by:

$$y = f \left( \sum_{j=1}^q W_j f \left( \sum_{l=1}^N w_{j,l} x_l + w_{j,0} \right) + W_0 \right), \quad (1)$$

where  $N$  is the number of inputs,  $q$  is the number of hidden neurons,  $W_j$  is the weight between the  $j$ th hidden neuron and the output neuron,  $w_{j,l}$  is the weight between the  $l$ th input neuron and the  $j$ th hidden neuron,  $x_l$  is the  $l$ th input parameter,  $w_{j,0}$  is the weight between a fixed input equal to 1 and  $j$ th hidden neuron, and  $W_0$  is the weight between a fixed input equal to 1 and the output neuron (Valach *et al.* 2007). The sigmoid function is the most commonly used activation function satisfying the approximation conditions of BPANN (Haykin 1999; Beale *et al.* 2010) and is represented by:

$$f(z) = \frac{1}{(1 + e^{-z})}, \quad (2)$$

where  $z$  is the input information of the neuron and  $f(z) \in [0, 1]$ . The input and output values of BPANN have to be scaled in this range.

The back propagation algorithm based on squared error minimization corresponds to an adjustment of the weights between the hidden layer and the output layer. This iterative process updates the weights in order to decrease the residuals of the predicted output of the neural network. It requires the estimation of the network parameters that lead to the global minimum of a cost function  $E$ . Typically, this cost function is chosen to be the sum of the squared discrepancies between computed and target output over all samples  $N$  and all output units  $K$ :

$$E = \frac{1}{N} \sum_{i=1}^N \left[ \frac{1}{2} \sum_{k=1}^K (y'_i(k) - y_i(k))^2 \right], \quad (3)$$

where  $y'_i(k)$  is the BPANN output and  $y_i(k)$  is the target response of each output neuron  $k$ .

The new weights are estimated by modifying them in the opposite direction of the gradient of the cost function in the point of actual estimation. The weight-update at iteration ' $t$ ' is given by:

$$\Delta w_{j,l}(t) = -\eta \frac{\partial E}{\partial w_{j,l}} + \alpha \cdot \Delta w_{j,l}(t-1), \quad (4)$$

where the parameter  $\eta$  denotes learning rate and  $\alpha$  is the momentum term.

## 2.2 Radial basis function neural network

RBFNN (Powell 1987) is known from the approximation theory as it is applied to the real multivariate interpolation problem. RBFNN is popularized by Moody and Darken (1989), and many researchers suggested it as an alternative ANN structure to MLP. RBFNN is very useful for function approximation and classification problems because of its more compact topology and faster learning speed. RBFNN is configured with three layers (figure 2). An input layer consists of source neurons (sensory units) and distributes input vectors to each of the neurons in the hidden layer without any multiplicative factors. The single hidden layer has receptive field units (hidden neurons) each of which represents a nonlinear transfer function called a basis function. The output layer produces a linear weighted sum of hidden neuron outputs and supplies the response of RBFNN.

The output of  $i$ th output neuron can be described in a general expression as follows:

$$y_i = \sum_{j=1}^q w_{ij} \phi_j(x) + w_0, \quad (5)$$

where  $q$  is the number of hidden neurons,  $w_{ij}$  is the weight between the  $j$ th hidden neuron and the  $i$ th output neuron, and  $w_0$  is the bias value. The basic function  $\phi(\cdot)$  is a nonlinear transformation from the input layer to hidden layer of high dimensionality, and it plays the role of the activation function in MLPs. The most common form of basic function in RBFNN is the Gaussian function (Bishop 2005; Yeung *et al.* 2010) and it is defined by:

$$\phi_j(x) = \exp \left( -\frac{\|x - u_j\|^2}{2\sigma_j^2} \right), \quad (6)$$

where  $x \in R^d$  is the input vector and  $u_j \in R^d$  and  $\sigma_j$  are the centre value and the width parameter of

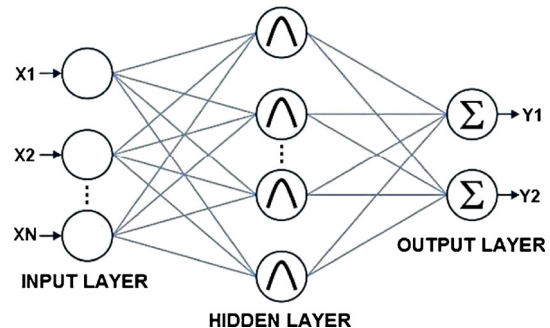


Figure 2. The RBFNN architecture.

the basic function, respectively, associated with the  $j$ th hidden neuron.  $\|\cdot\|$  denotes the Euclidean distance. The hidden neuron is activated whenever  $x$  is close enough to its corresponding  $u$  in RBFNN. The location of neurons, the weight coefficients, and the bias are defined during the training process of RBFNN. The training of RBFNN requires a set of data samples for which the corresponding network outputs are known. Mathematically, the training can be considered as an optimization problem where the network parameters are to be solved while the error of the neural network must be minimal.

### 2.3 Kriging interpolation method

KRIG (Krige 1951) is a geostatistical and flexible interpolation method which has been extensively used in diverse fields of mathematics, earth sciences, geography and engineering and has proved to be powerful and accurate in its fields of use. According to KRIG, both the distance and the degree of variation between reference points are taken into account for optimal spatial prediction (Joseph 2006). KRIG assigns a mathematical function to a certain number of points or all the points located within a certain area of effect in order to determine the output values for each location (Chaplot *et al.* 2006). KRIG uses the semivariogram which measures the average degree of dissimilarity between unsampled values and nearby values to define the weights that determine the contribution of each data point to the prediction of new values at unsampled locations (Krivoruchko and Gotway 2004). KRIG is based on a constant mean  $\mu$  for the data and random errors  $\varepsilon$  with spatial dependence as follows:

$$Z(x_0) = \mu(x_0) + \varepsilon(x_0), \quad (7)$$

where  $Z(x_0)$  is the variable of interest,  $\mu(x_0)$  is the deterministic trend, and  $\varepsilon(x_0)$  is the correlated error (Erdogan 2010). In the ordinary algorithm of KRIG, equation (7) can be given as follows:

$$Z(x_0) = \mu(x_0) + \sum_{i=1}^n \lambda_i [z(x_i) - \mu(x_0)], \quad (8)$$

where  $n$  is the number of sampled points used for the estimation,  $\lambda_i$  is the weight assigned to the sampled point ( $x_i$ ), and  $\sum_{i=1}^n \lambda_i = 1$  is forced (Li and Heap 2008). KRIG is the most appropriate interpolation method when a spatially correlated distance or directional bias in the data is known.

### 3. Data acquisition and evaluation methodology

The estimation of the geodetic point velocities is performed over a study area located in central and western Anatolian parts of Turkey. The study area is limited by 36.95°–40.50°N in latitude and, 27.10°–32.75°E in longitude, and it defines a total area of ~182,500 km<sup>2</sup>. Its span is approximately 380 km in the north–south direction and 480 km in the east–west direction.

The evaluating procedure of the geodetic points' velocity refers to a source dataset in the study area that comprises 125 control points belonging to Turkish National Fundamental GPS Network (TNFGN) (Ayhan *et al.* 2002). The positional accuracies of the TNFGN stations are about 1–3 cm whereas the relative accuracies are within the range of 0.01–0.1 ppm. For each TNFGN station, time-dependent 3D coordinates and their associated velocities were computed in ITRF2000 (reference epoch 2005.00) with repeated GPS observations (Caglar 2006). Velocity solution of TNFGN over the interval 1992–2004 was obtained by the procession of campaign type GPS measurements of 366 TNFGN points. TNFGN velocities with 1 $\sigma$  standard deviations used in this study are given in table 1.

The source dataset (125 TNFGN points) is classified into two groups as a reference dataset for the training (modelling) process and a test dataset for the controlling process. In ANN approach, the reference points are used to train BPANN and RBFNN, and the test points are used to evaluate the performance of ANNs. In KRIG approach, the reference points are used to generate a surface model of the study area, and the test points are used to check the estimation accuracy of KRIG. Five different geodetic networks are generated to assess the impact of the point density on the velocity estimation results. The reference points are selected to cover the study area from outside, and the test points are selected as densification points of the geodetic network formed by the reference points. The point classification and density information about the geodetic networks are summarized in table 2, and the spatial distribution of the reference and test points in geodetic networks within the study area is plotted in figure 3.

Table 1. Standard deviations (1 $\sigma$ ) of TNFGN velocities (units in mm/year).

	$\sigma V_X$	$\sigma V_Y$	$\sigma V_Z$
Minimum	0.60	0.40	0.60
Maximum	11.10	6.00	10.20
Mean	2.00	1.17	1.86

Table 2. Point classification and density of geodetic networks.

Geodetic network	Classification		Density (km <sup>2</sup> /point)	
	Reference	Test	Reference	Test
GN-1	20	105	~9000	~1750
GN-2	40	85	~4500	~2150
GN-3	60	65	~3000	~2800
GN-4	80	45	~2250	~4050
GN-5	100	25	~1825	~7300

The evaluation of the geodetic point velocity estimation by BPANN, RBFNN, and KRIG is focused on the differences between the known and estimated point velocities using the equation below:

$$\Delta V_{X,Y,Z} = V_{X,Y,Z}(\text{TNFGN}) - V_{X,Y,Z}(\text{BPANN, RBFNN, KRIG}), \quad (9)$$

where  $\Delta V_{X,Y,Z}$  is the geodetic point velocity residual,  $V_{X,Y,Z}$  (TNFGN) is the point velocity known through repeated GPS measurements within TNFGN and  $V_{X,Y,Z}$  (BPANN, RBFNN, KRIG) is the point velocity based on BPANN, RBFNN, and KRIG.

For the statistical analysis of geodetic point velocity residuals ( $\Delta V_{X,Y,Z}$ ) minimum, maximum, and mean values were determined and investigated by RMSE value because RMSEs are sensitive to even small errors to measure the deviations between known and estimated discharges on models (Gullu *et al.* 2011), RMSEs are global measures for comparing interpolation techniques (Erdogan 2010), and are effective tools for evaluating the results of ANN applications (Schroeder *et al.* 2009). RMSE is always positive and it is defined by:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\Delta V_{X,Y,Z})^2}, \quad (10)$$

where  $n$  is the number of test points used in the geodetic network.

#### 4. ANN design and optimisation

The main goal of ANNs is to find a solution to generalize the multidimensional input–output mapping problems. In other words, ANNs perform well when they do not extrapolate beyond the range of the (training) data used for the estimation of their parameters. In order to do so, ANNs have to capture the functional relationship that leads to the mapping of the input data into the output data. ANN structure that is chosen to be too complex in relation to the functional relationship that has

to be captured also memorises its free coefficients, the noise contained in the data. This occurrence is called overfitting. Such a model will perform well in approximating the data used in order to estimate its parameters (ANN has memorised the training data) but it will be extremely poor on new data (ANN has not learnt to generalize). To allow proper generalization capabilities, ANN overfitting of the training data must be avoided (i.e., model should be fitted only to the signal present in the training sample, not to the noise). A number of techniques have been developed to further improve ANN generalization capabilities including different variants of cross-validation (Haykin 1999), noise injection (Holmstrom and Koistinen 1992), error regularization, weight decay (Poggio and Girosi 1990; Haykin 1999), and the optimized approximation algorithm (Liu *et al.* 2008). A number of cross-validation variants exist, and some of them are of special attention when data are very scarce, i.e., multifold cross-validation or leave-one-out (Haykin 1999). But probably the most popular in practical applications (Liu *et al.* 2008) is the so-called early stopping (Piotrowski and Napiorkowski 2013). To use the early stopping approach in this study, the available dataset is divided into two subsets: (i) training (reference) data used during ANN optimization and (ii) the test data (not presented to ANN during optimization) used to define stopping criteria to prevent overfitting thus ensuring a generalised solution. The division is done keeping in mind that training dataset should be extensive and comprehensive (representative of all possible variations of the data on which ANN will be tested). The mean square error (MSE) (Graupe 2007; Hsieh 2009) is used as the model evaluation indicator. For a given set of  $N$  inputs, MSE is defined by:

$$\text{MSE} = \sum_1^N (y_{\text{act}} - y_{\text{pred}})^2 / N, \quad (11)$$

where  $y_{\text{act}}$  denotes the given actual output value, and  $y_{\text{pred}}$  denotes the ANN (predicted) output. The performance of ANN during the training and testing process is monitored in the form of MSE. The testing error normally decreased during the initial phase of training, as did the training error. However, when overfitting occurred, the testing error

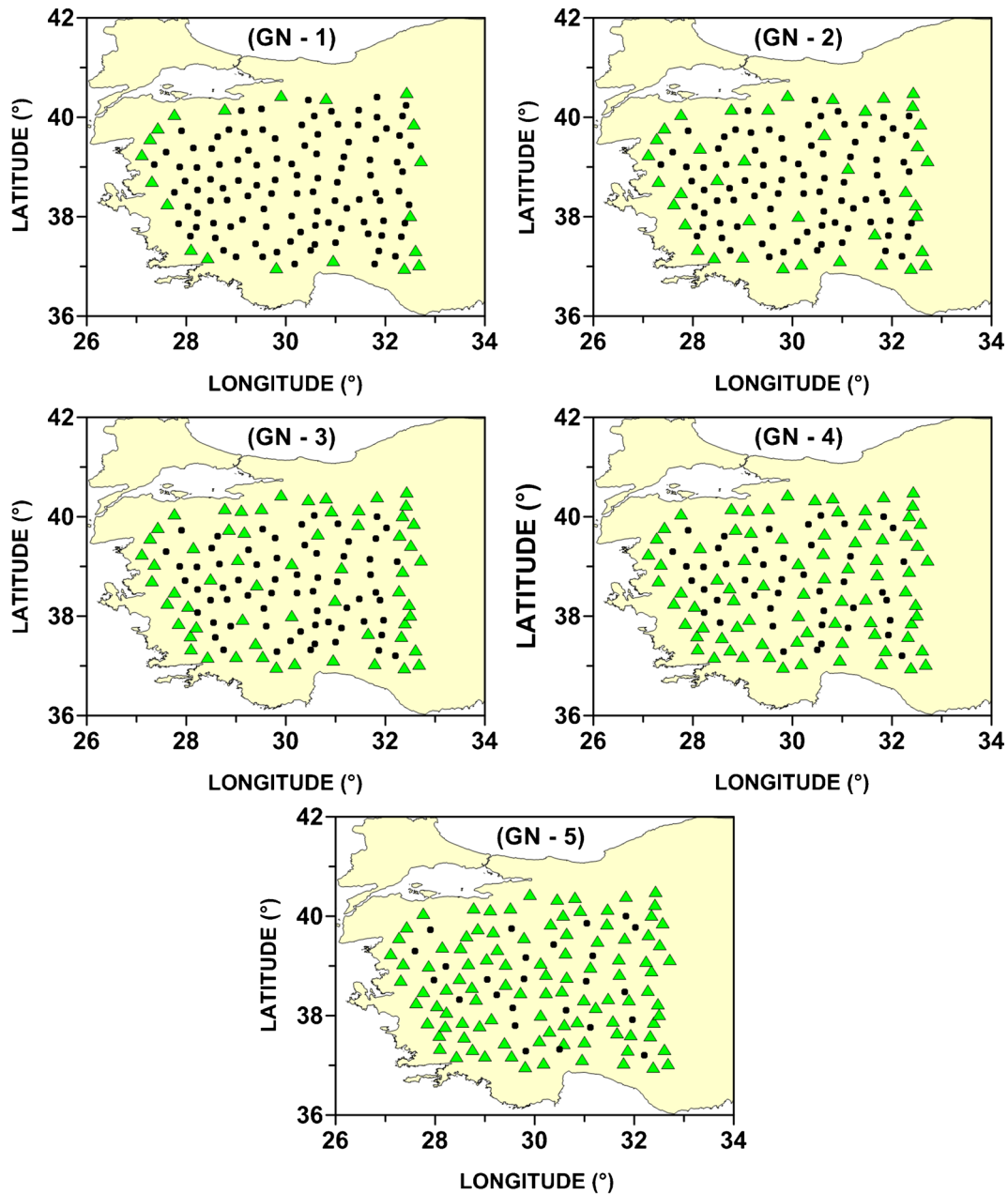


Figure 3. Geodetic networks ( $\blacktriangle \rightarrow$  reference;  $\bullet \rightarrow$  test).

typically began to rise. When the testing error increased, the training process was stopped, and it is assumed that optimal ANN parameters were reached. In table 3, the MSEs of ANNs obtained by early stopping to avoid overfitting are shown for training and test dataset on Geodetic Network (1).

Once the available data have been divided into training (reference) and testing subsets, ANN training can be made more efficient by pre-processing the data in a suitable form before they are applied to ANN. Data pre-processing is necessary to ensure that all inputs receive equal attention during the training process and to give numerical stability to ANN. Moreover, pre-processing usually speeds up

the learning procedure and minimizes the prediction error (Boukhrissa *et al.* 2013). Pre-processing can be in the form of data scaling, normalization and transformation (Shahin *et al.* 2008). In this study, the minimum–maximum normalization (a linear transformation that preserves exactly all relationships of the original data) is used for scaling the inputs and outputs to commensurate within the specified range of the activation function used for ANN. The associated normalization is expressed by:

$$P_n(i) = \frac{(P_i - P_{\min})}{(P_{\max} - P_{\min})} \quad (12)$$

Table 3. *MSEs of the datasets (units in mm/year).*

MSE	BPANN			RBFNN		
	$V_X$	$V_Y$	$V_Z$	$V_X$	$V_Y$	$V_Z$
Training	3.15	2.12	3.90	3.98	3.26	3.36
Test	6.51	3.73	5.04	9.76	5.04	6.67

where  $P_n(i)$  is the normalized parameter,  $P$  is either input or output parameter,  $P_{\min}$  and  $P_{\max}$  refers to the minimum and maximum values of the parameters, respectively.

The architecture of ANN determines the number of parameters to be calibrated. This architecture should always be adapted to the problem in question (Zhang *et al.* 1998), as it depends on the number of input and output variables. Generally, the number of neurons in the input layer depends on the number of possible inputs (independent variables) that we used in ANN. However, the number of neurons in the output layer depends on the number of desired (target) outputs. For this study, ANNs are proposed with two neurons in the input layer and one neuron in the output layer. The geographical coordinates (latitude and longitude) of the geodetic point are selected as input quantities, and the velocity component of the point ( $V_{X,Y,Z}$ ) is used as output quantity for training and testing procedure of BPANN and RBFNN.

In ANN approach, there are two major challenges regarding the hidden layers: the number of hidden layers and how many neurons will be in each of these hidden layers. Two hidden layers are required for modelling data with discontinuities such as a sawtooth wave pattern. Actually, one hidden layer is sufficient for nearly all problems (Panchal *et al.* 2011). In the present study, the proposed BPANN and RBFNN are composed of one hidden layer. ANN with one hidden layer can approximate any continuous function given a sufficient number of hidden neurons (Cybenko 1989; Funahashi 1989; Hornik *et al.* 1989; Bishop 2005). Essentially, the number of neurons in the hidden layer defines the complexity and power of ANN to delineate the underlying relationships and structures inherent in a dataset. The number of hidden layer neurons has a considerable effect on both classification accuracy and training time requirements. The accuracy that can be produced by ANN relates to the generalisation capabilities. Basically, the number of neurons in the hidden layer should be large enough for the correct representation of the problem, but at the same time low enough to have adequate generalisation capabilities (Kavzoglu and Mather 2003). Several strategies and heuristics (destructive, constructive and hybrid methods) have been suggested to estimate the optimum

number of hidden layer nodes (i.e., Hecht-Nielsen 1987; Ripley 1993; Kaastra and Boyd 1996; Kanellopoulos and Wilkinson 1997; Witten and Frank 2005). Another way of determining the optimal number of hidden neurons that can result in good generalization and avoid overfitting is to relate the number of hidden neurons to the number of training samples (i.e., Masters 1993; Rogers and Dowla 1994; Amari *et al.* 1997). A number of systematic approaches have also been proposed to obtain the optimal ANN architecture (i.e., Ghaboussi and Sidarta 1998; Chakraverty *et al.* 2006; Chakraverty 2007; Kingston *et al.* 2008). However, none of these suggestions has been universally accepted or used. There is no direct and precise way of determining the best number of nodes in each hidden layer. In most of the reported applications, the number of hidden neurons is determined from the experience of individuals using trial-and-error strategies. Hence, a trial-and-error strategy with 30 neurons in the hidden layer of ANN is applied in this study, and ANN is pruned by gradually decreasing the hidden neurons. Consequently, the optimal number of neurons in the hidden layer was selected as 20 for BPANN and 17 for RBFNN which produced the smallest MSE. Thus, the optimum structure of BPANN [2:20:1] and of RBFNN [2:17:1] was determined by MATLAB ANN module that allows changing the learning algorithm parameters dynamically, monitoring error values and generating digital data about sufficient learning rates. The significance of pruning away hidden neurons in ANN architecture on Geodetic Network (5) is represented in figure 4.

The main disadvantage of ANNs that use back propagation algorithm is its slow convergence to the global minimum. It is also likely to become trapped into a local minimum. The learning rate ( $\eta$ ), also referred to as the step size, is used to control the degree of the change in the weights in response to errors in the output during each training cycle. The learning rate determines the size of the steps taken towards the global minimum error throughout the training process. It can be considered as the key parameter for a successful ANN application because it controls the learning process (Kavzoglu and Saka 2005). If the learning rate is set too high, then large changes are allowed in the weight and no learning occurs. Conversely, if the learning rate is set too low, only small changes



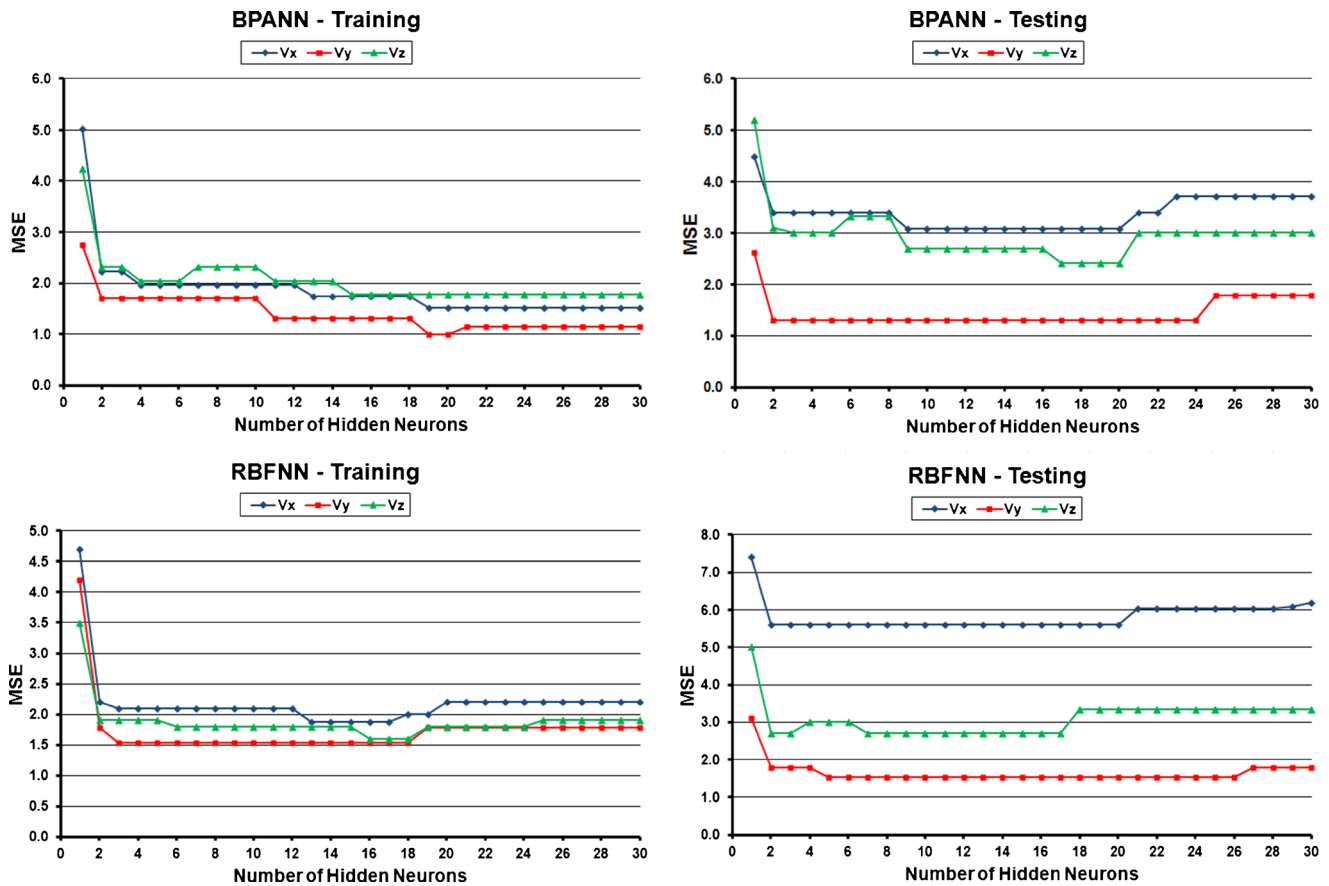


Figure 4. The significance of pruning away hidden neurons in ANN architecture.

are allowed, which can increase the learning time. The momentum term ( $\alpha$ ) dampens the amount of weight change by adding in a portion of the weight change from the previous iteration. The momentum term is credited with smoothing out large changes in the weights and with helping the network converge faster when the error is changing in the correct direction (Neyamadpour *et al.* 2010). In this study, the estimation is started with a learning rate of 0.3 according to the guidelines suggested by Neuner (2010). Due to the fact that only an adaptive learning rate ensures the convergence (Bishop 2005), the learning rate is decreased by a factor of 0.5 if the cost function decreases and it is increased by a factor of 1.05 when the cost function increases during the training procedure. Also, the momentum term is fixed to 0.6 for weight update process. According to the general guidelines from ANN literature (i.e., Gallagher and Downs 1997; Graupe 2007), the initial values of the inter-neuron weights were set to a range  $[-0.25, 0.25]$  (suitable for the activation function) at the beginning of the training process for converging to global minimum quickly without getting stuck in a local minimum. The design and optimization parameters of ANNs of this study are summarized in table 4.

## 5. Case study

BPANN and RBFNN are trained in Geodetic Network (5) (maximum reference points), and the velocities of the test points are estimated via the trained ANNs for the controlling process. The ANN parameters obtained in the training procedure in Geodetic Network (5) are fixed and used as constants in the training process of ANNs for the other geodetic networks.

In KRIG approach, the reference velocity field vectors of the study area are generated from the reference dataset by Surfer 11 surface modelling program that is used widely for contour mapping, terrain modelling, and 3D surface mapping. These vector maps are overlaid on the velocity field contour maps (figure 5) to describe the directional dependence. Figure 5 reveals that the reference velocity fields used in this study are consistent with the horizontal and vertical velocity fields of Turkey computed by Aktug *et al.* (2011). The KRIG defaults of the software were accepted for modelling the velocity fields which were point Kriging type, non-drift type (ordinary Kriging), and linear variogram model. The reference velocity fields were checked by cross-validation technique

Table 4. Design and optimization parameters of ANNs.

Parameters	Settings
Training algorithm	Gradient descent back propagation
Activation function	Sigmoid (BPANN); Gaussian (RBFNN)
Number of input neurons	2
Number of output neurons	1
Number of hidden neurons	20 (BPANN); 17 (RBFNN)
Early stopping criterion for training process	Test dataset
Data pre-processing	Min-max normalization
Initial weight range	$[-0.25, 0.25]$
Learning rate	0.3
Learning rate decrease	0.5
Learning rate increase	1.05
Momentum term	0.6
Performance function	MSE

and the velocity residuals of the test points were computed from these fields.

The results of the test dataset are significant in the evaluation procedure of BPANN, RBFNN and KRIG. Therefore, velocity residual maps are produced with regard to the velocity differences of the test points computed by equation (9) in the geodetic networks. The velocity residual maps of the test points associated with  $\Delta V_{X,Y,Z}$  are given in figures 6–8, respectively. The contour lines are drawn at 2 mm intervals on the velocity residual maps.

## 6. Results and conclusions

The analysis of the velocity residuals plotted in figures 6–8 shows that the point velocity residuals are getting smaller depending on the increase in the number of the reference points in geodetic networks. BPANN's point velocity estimation is better than RBFNN's estimations in all geodetic networks for ANN approach.

The statistical values of the test dataset's velocity residuals are presented in table 5, and the velocity residual RMSEs of the test points based on BPANN, RBFNN and KRIG are shown in figure 9.

When the results summarized in table 5 are evaluated, it can be seen from figure 9 that BPANN estimated the point velocities more accurately in Geodetic Networks (1), (2), and (3), with respect to KRIG, in terms of RMSE as compared to the others. In Geodetic Networks (4) and (5), KRIG is more useful than BPANN for the point velocity estimation. RBFNN's estimation accuracy is approximately at the same level with KRIG's estimation only in Geodetic Network (1). On the other

geodetic networks, KRIG's results are better than RBFNN's results.

Based on the experimental results of evaluating the utility of ANN for the velocity estimation in regional geodetic networks, the following conclusions can be drawn from this study:

- The employment of BPANN is an alternative tool to KRIG for the geodetic point velocity estimation, in practice.
- BPANN can be used effectively with a small reference point density in geodetic networks. A rough guideline of point density can be introduced as  $\leq \sim 3000 \text{ km}^2/\text{point}$  with respect to the accuracy of the result. When the number of the points that will be estimated (test) is smaller than the number of the points known (reference), the estimation of geodetic point velocity with the use of KRIG is evaluated as more powerful than using BPANN.
- The main advantage of BPANN considered as a velocity estimator is model-free estimation of its flexible structure. The properly trained BPANN can be used in the geodetic velocity estimation for additional points, whereas KRIG is re-determining weights for each additional point in geodetic network.
- The combination of ANN models (with diverse architecture; e.g., different training algorithms and activation functions, additional hidden layers and neurons) with interpolation methods would be an appealing tool in geodetic velocity field applications where one has little or incomplete point velocity data, because of ANN's adaptive feature that 'learning by example' replaces 'programming' and extrapolation ability in estimation problems (for boundary or outside of the geodetic networks).

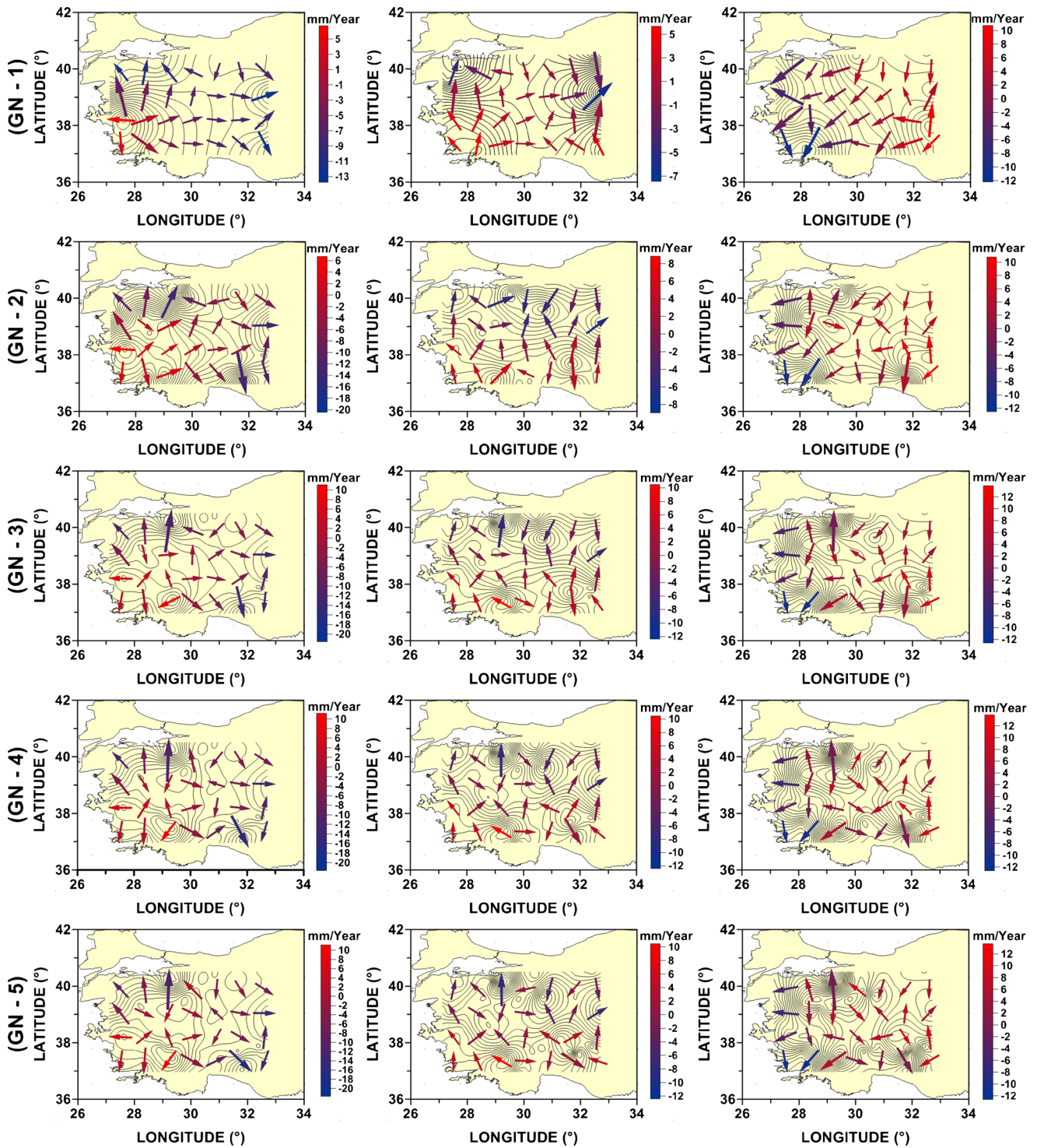


Figure 5. Reference velocity fields of geodetic networks:  $V_X$  (left),  $V_Y$  (middle), and  $V_Z$  (right).

ANN is a data-driven approach in which the model can be trained by input-output data to determine the structure (parameters) of the model. For ANN, there is no need to either simplify the physical complexity of the problem or incorporate any assumptions about the frequency distribution of the data. Besides, ANN

can always be updated with new training data to obtain better results. In this regard, ANN outperforms the conventional methods and can be used as a powerful modelling tool in geodetic and geophysical problems. The results of this study reflect that the application of ANN has the ability to estimate the point velocity estimation

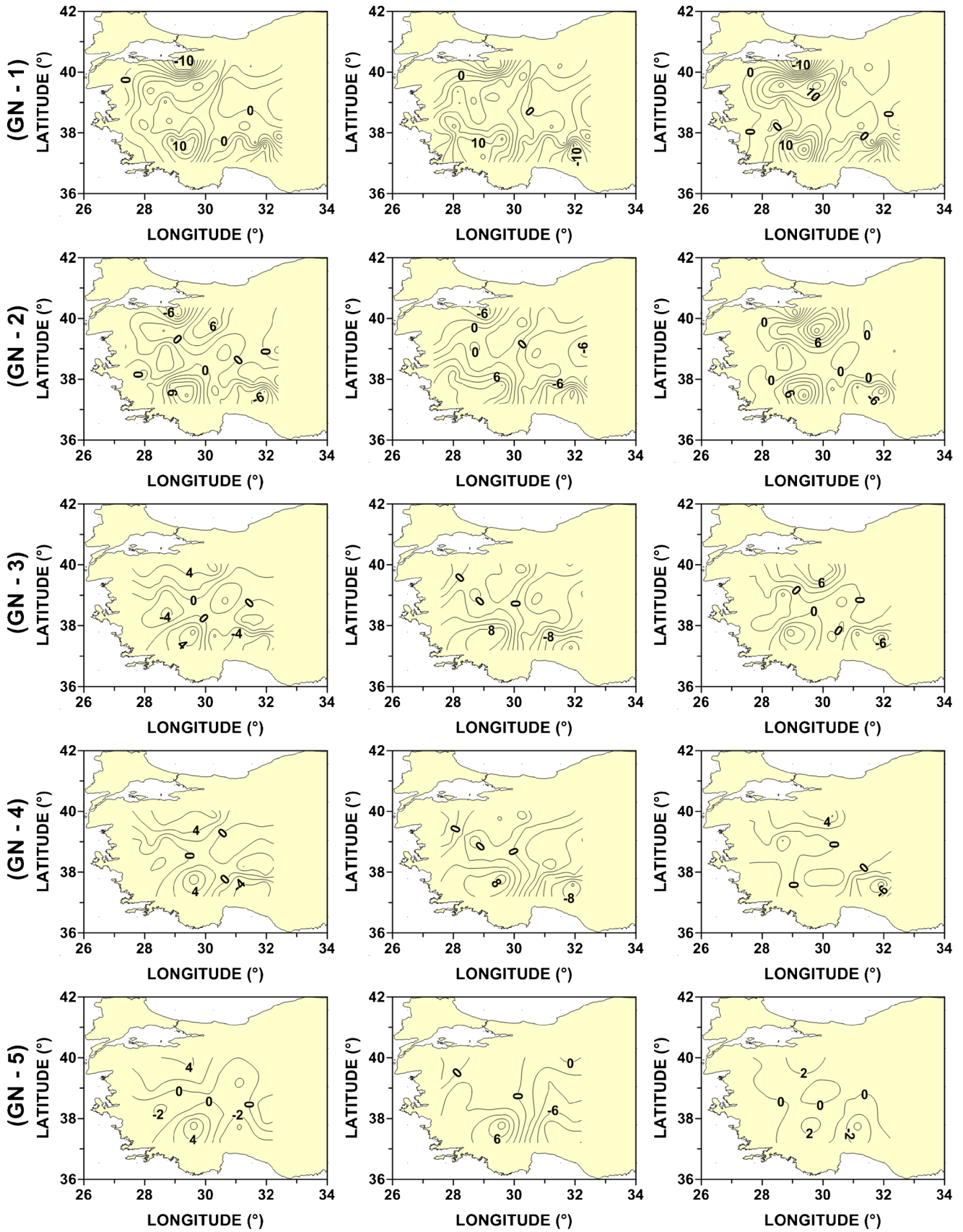


Figure 6.  $\Delta V_X$  velocity residual maps: BPANN (left), RBFNN (middle), and KRIG (right).

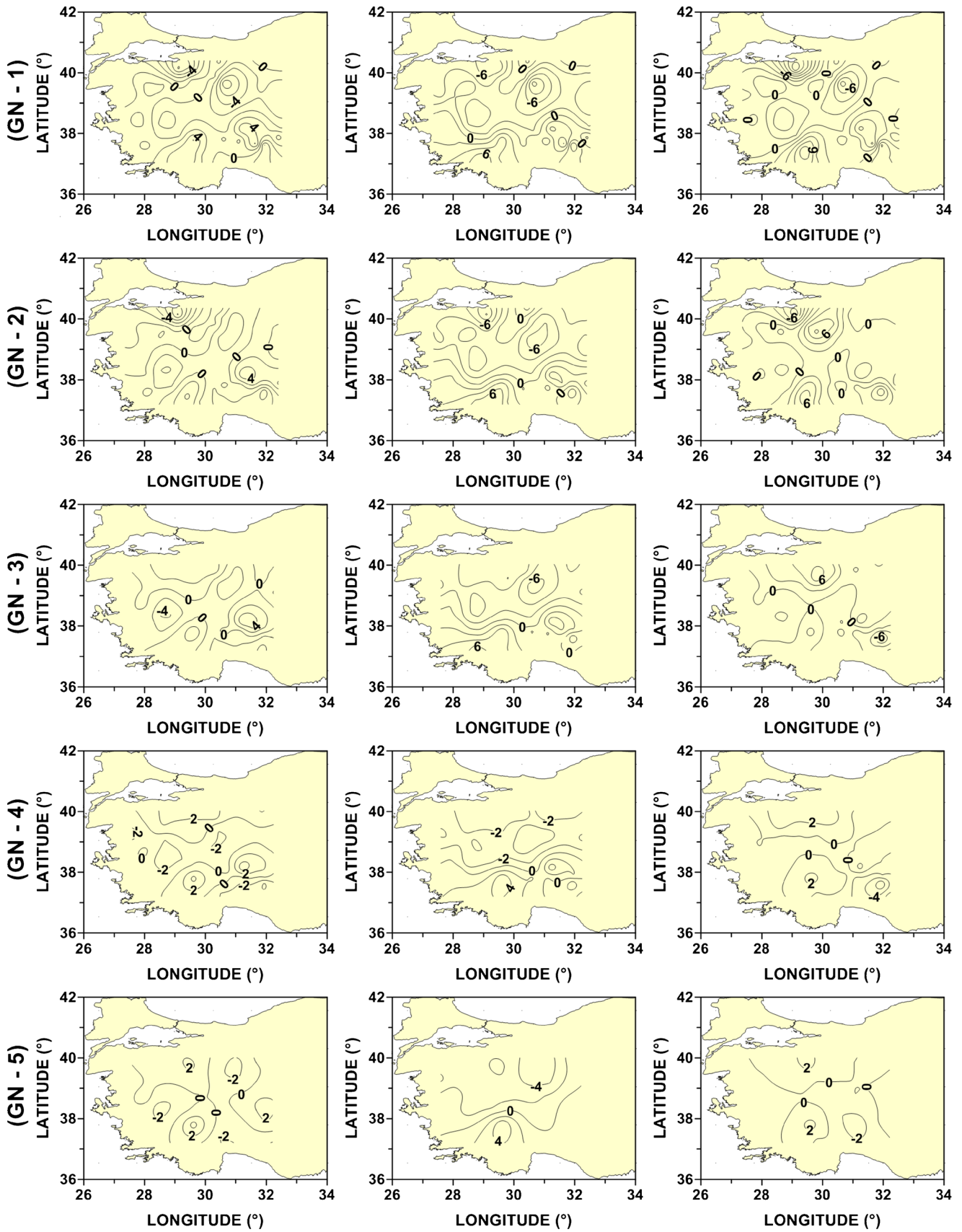


Figure 7.  $\Delta V_Y$  velocity residual maps: BPANN (left), RBFNN (middle), and KRIG (right).

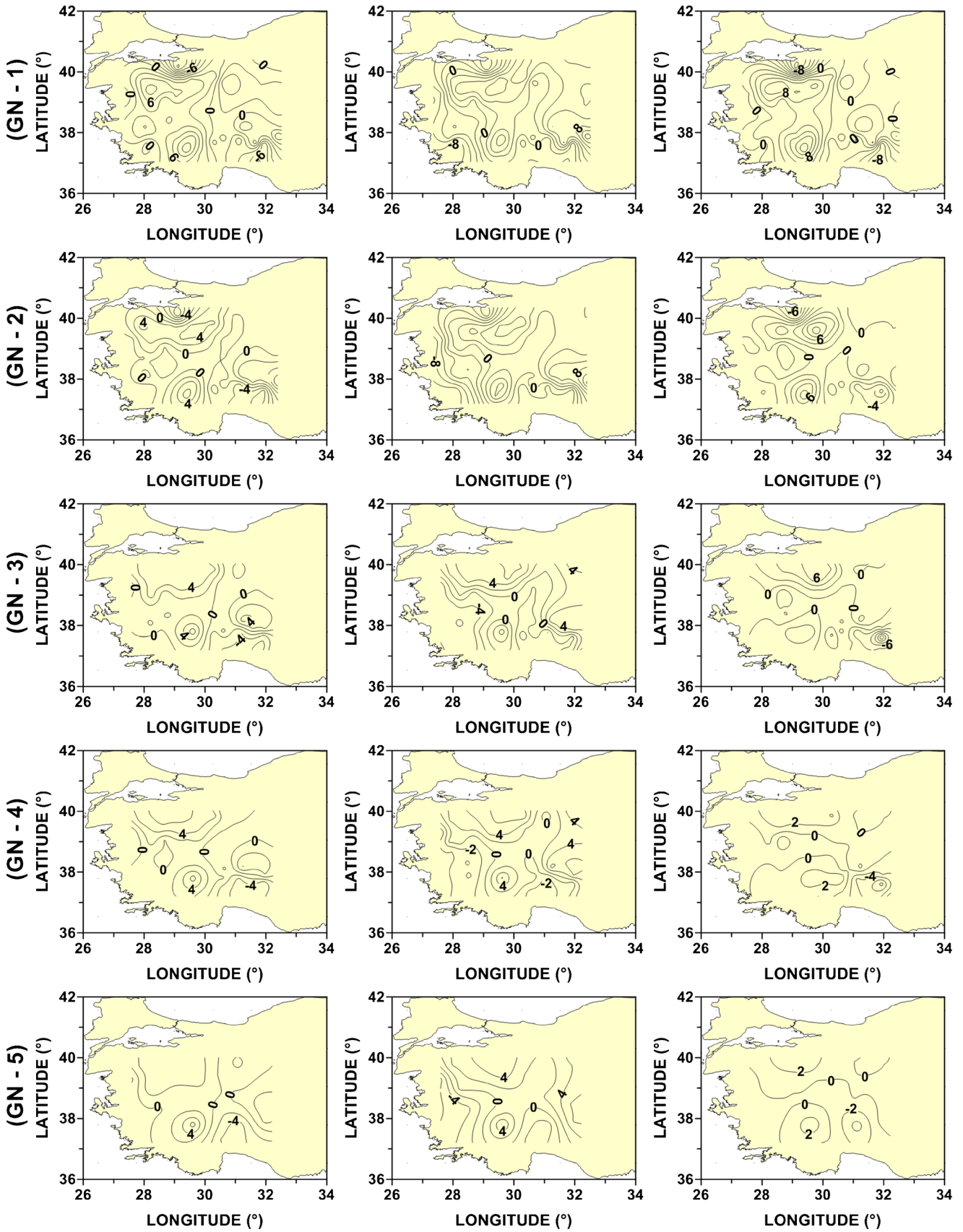


Figure 8.  $\Delta V_z$  velocity residual maps: BPANN (left), RBFNN (middle), and KRIG (right).

Table 5. Statistics of test dataset's velocity residuals over the geodetic networks (units in mm/year).

	BPANN			RBFNN			KRIG		
	$\Delta V_X$	$\Delta V_Y$	$\Delta V_Z$	$\Delta V_X$	$\Delta V_Y$	$\Delta V_Z$	$\Delta V_X$	$\Delta V_Y$	$\Delta V_Z$
Geodetic network (1)									
Minimum	-11.0	-10.3	-12.5	-12.7	-11.0	-10.6	-11.8	-16.0	-12.7
Maximum	12.1	6.5	11.9	12.7	7.6	13.6	17.2	10.2	14.1
Mean	2.0	-0.3	0.9	2.0	-1.5	2.0	2.7	-0.5	2.1
RMSE	4.8	3.0	4.1	5.4	3.7	4.6	5.4	3.5	4.5
Geodetic network (2)									
Minimum	-10.4	-10.4	-10.1	-14.8	-10.4	-11.7	-10.1	-9.7	-9.9
Maximum	10.5	5.7	8.2	10.2	8.2	12.5	15.3	8.5	11.5
Mean	0.1	0.3	-0.3	-0.2	-1.2	1.0	1.0	0.5	0.5
RMSE	3.5	2.5	3.2	4.8	3.6	4.4	4.0	2.8	3.3
Geodetic network (3)									
Minimum	-8.0	-4.5	-6.9	-14.5	-7.0	-5.7	-10.0	-9.6	-14.2
Maximum	7.1	6.0	7.3	10.0	6.5	6.6	9.9	7.4	8.4
Mean	-0.5	-0.3	0.6	-0.5	-1.1	-0.1	0.0	0.0	-0.2
RMSE	3.1	2.2	2.9	4.7	3.1	3.5	3.1	2.3	3.0
Geodetic network (4)									
Minimum	-6.0	-4.4	-5.9	-14.5	-5.8	-5.8	-9.6	-8.0	-8.9
Maximum	7.8	5.7	7.0	9.7	5.9	7.1	6.3	3.7	4.3
Mean	1.0	0.1	0.9	-0.3	-1.2	1.3	0.1	0.0	0.0
RMSE	3.1	2.1	2.9	4.5	2.9	3.6	2.4	1.8	2.2
Geodetic network (5)									
Minimum	-4.4	-3.1	-5.2	-10.1	-5.7	-6.5	-4.9	-3.6	-5.1
Maximum	7.3	5.0	7.0	9.4	6.1	6.2	3.7	3.0	3.9
Mean	0.5	0.0	0.4	-0.7	-1.2	1.1	0.2	-0.1	0.0
RMSE	2.7	1.9	2.7	4.2	2.9	3.6	1.5	1.3	1.6

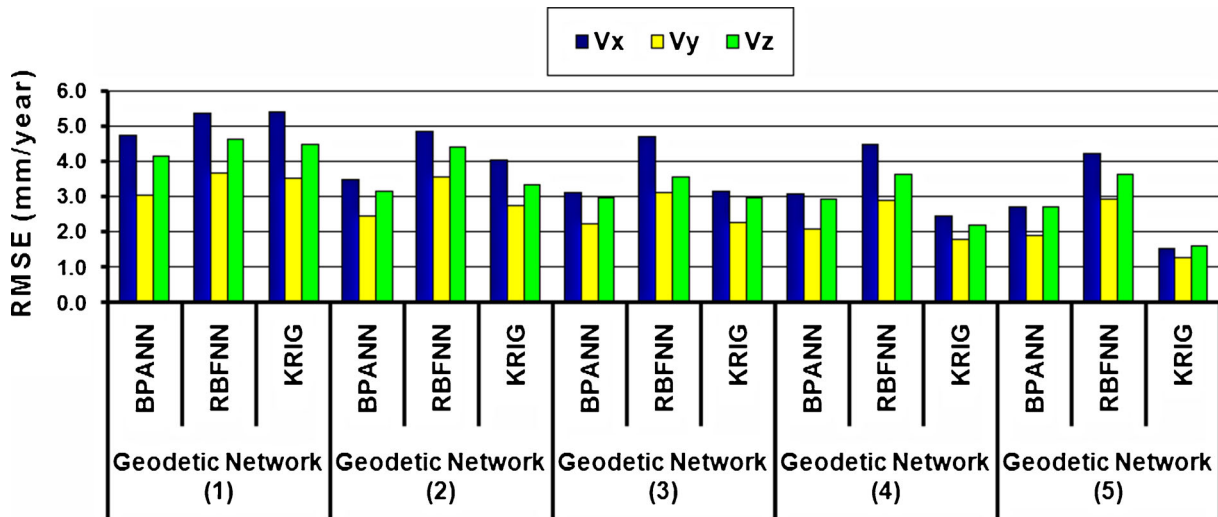


Figure 9. RMSEs of velocity residuals of the test dataset.

for regional geodetic networks. The diverse ANN architectures can be applied to other datasets for determining the velocity field of geodetic GPS networks which is an open research problem. Despite the feasibility of ANN in velocity field determination, future research should give further attention to ensuring robust models, improving extrapolation ability, and dealing with uncertainty.

### Acknowledgements

This study was supported by Afyon Kocatepe University Scientific Research Projects Coordination Department (Project No: 11.FEN.BIL.23). The authors thank the two anonymous reviewers whose constructive comments led to significant improvements on the original manuscript.

## References

- Aktug B, Sezer S, Ozdemir S, Lenk O and Kilicoglu A 2011 Computation of the actual coordinates and velocities of Turkish National Fundamental GPS Network; *J. Mapp.* **145** 1–14 (in Turkish).
- Altamimi Z, Sillard P and Boucher C 2002 ITRF2000: A new release of the international terrestrial reference frame for Earth science applications; *J. Geophys. Res. Solid Earth* **107**(B10) ETG 2-1–ETG 2-19, doi: [10.1029/2001JB000561](https://doi.org/10.1029/2001JB000561).
- Amari S I, Murata N, Muller K R, Finke M and Yang H H 1997 Asymptotic statistical theory of overtraining and cross-validation; *IEEE Trans. Neural Network* **8**(5) 985–996.
- Ayhan M E, Demir C, Lenk O, Kilicoglu A, Aktug B, Acikgoz M, Fırat O, Sengun Y S, Cingoz A, Gurdal M A, Kurt A I, Ocak M, Turkezer A, Yildiz H, Bayazit N, Ata M, Caglar Y and Ozerkan A 2002 Turkish National Fundamental GPS Network-1999A (TNGFN-99A); *J. Mapp. (Spec. Issue)* **16** 1–73 (in Turkish).
- Baronian C, Riahi M A, Lucas C and Mokhtari M 2007 A theoretical approach to applicability of artificial neural networks for seismic velocity analysis; *J. Appl. Sci.* **7**(23) 3659–3668.
- Beale M H, Hagan M T and Demuth H B 2010 *Neural Network Toolbox 7 Users Guide*; The MathWorks Inc., Natick, MA, 951p.
- Bishop C M 2005 *Neural Networks for Pattern Recognition*; Oxford University Press, New York, 504p.
- Boukhrissa Z A, Khanchoul K, Le Bissonnais Y and Tourki M 2013 Prediction of sediment load by sediment rating curve and neural network (ANN) in El Kebir catchment, Algeria; *J. Earth Syst. Sci.* **122**(5) 1303–1312.
- Caglar Y 2006 National report of Turkey–2005; *Mitt. Bundesamtes. Kartogr. Geod.* **38** 299–306.
- Calderon-Macias C, Sen M K and Stoffa P L 2000 Artificial neural networks for parameter estimation in geophysics; *Geophys. Prospect.* **48**(1) 21–47.
- Chakraverty S, Singh V P and Sharma R K 2006 Regression based weight generation algorithm in neural network for estimation of frequencies of vibrating plates; *J. Comput. Methods Appl. Mech. Eng.* **195** 4194–4202.
- Chakraverty S 2007 Identification of structural parameters of two-storey shear buildings by the iterative training of neural networks; *Archit. Sci. Rev.* **50**(4) 380–384.
- Chamoli A, Bansal A R and Dimri V P 2007 Wavelet and rescaled range approach for the Hurst coefficient for short and long time series; *Comput. Geosci.* **33** 83–93.
- Chaplot V, Darboux F, Bourennane H, Leguedois S, Silvera N and Phachomphon K 2006 Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landform types and data density; *Geomorphology* **77** 126–141.
- Cybenko G 1989 Approximations by superpositions of sigmoidal functions; *Math. Cont. Sig. Syst.* **2** 303–314.
- D'Anastasio E, De Martini P M, Selvaggi G, Pantosti D, Marchioni A and Maseroli R 2006 Short-term vertical velocity field in the Apennines (Italy) revealed by geodetic levelling data; *Tectonophysics.* **418** 219–234.
- Delacou B, Sue C, Nocquet J M, Champagnac J D, Allanic C and Burkhard M 2008 Quantification of strain rate in the Western Alps using geodesy: Comparisons with seismotectonics; *Swiss. J. Geosci.* **101**(2) 377–385.
- Demir C and Acikgoz M 2000 The estimation of long period coordinate changes (secular velocities) in Turkish National Fundamental GPS Network; *J. Mapp.* **123** 1–16 (in Turkish).
- Dimri V P and Chamoli A 2008 Development of computational geophysics in India; In: *Five Decades of Geophysics in India* (eds) Singh B and Dimri V P, *Geol. Soc. India Memoir* **68** 1–14.
- Erdogan S 2010 Modelling the spatial distribution of DEM error with geographically weighted regression: An experimental study; *Comput. Geosci.* **36** 34–43.
- Erol B and Erol S 2013 Learning-based computing techniques in geoid modeling for precise height transformation; *Comput. Geosci.* **52** 95–107.
- Eskandari H, Rezaee M R and Mohammadnia M 2004 Application of multiple regression and artificial neural network techniques to predict shear wave velocity from wireline log data for a carbonate reservoir, south-west Iran; *CSEG Recorder* **29** 40–48.
- Foti S, Parolai S, Albarello D and Picozzi M 2011 Application of surface-wave methods for seismic site characterization; *Surv. Geophys.* **32**(6) 777–825.
- Funahashi K 1989 On the approximate realization of continuous mappings by neural networks; *Neural Netw.* **2**(3) 183–192.
- Gallagher M and Downs T 1997 Visualisation of learning in neural networks using principal component analysis; In: *Proceedings of International Conference on Computational Intelligence and Multimedia Applications* (eds) Varma B and Yao X, Gold Coast, Australia, pp. 327–331.
- Ghaboussi J and Sidarta D E 1998 New nested adaptive neural networks (NANN) for constitutive modeling; *Comput. Geotech.* **22**(1) 29–52.
- Graupe D 2007 *Principles of Artificial Neural Networks*; World Scientific Publishing, Singapore, 320p.
- Gross R, Beutler G and Plag H-P 2007 Integrated scientific and societal user requirements and functional specifications for the GGOS; In: *Global Geodetic Observing System – Meeting the Requirements of a Global Society on a Changing Planet in 2020* (eds) Plag H-P and Pearlman M (Berlin Heidelberg: Springer-Verlag), pp. 126–134.
- Gullu M, Yilmaz I, Yilmaz M and Turgut B 2011 An alternative method for estimating densification point velocity based on back propagation artificial neural networks; *Stud. Geophys. Geod.* **55** 73–86.
- Hackl M, Malservisi R and Wdowinski S 2009 Strain rate patterns from dense GPS networks; *Nat. Hazards Earth Syst. Sci.* **9** 1177–1187.
- Haykin S 1999 *Neural Networks: A Comprehensive Foundation*; Prentice Hall, Upper Saddle River, 842p.
- Hecht-Nielsen R 1987 Kolmogorov's mapping neural network existence theorem; *Proceedings of the First IEEE International Conference on Neural Networks*, San Diego, CA, pp. 11–14.
- Hefty J 2008 Densification of the central Europe velocity field using velocities from local and regional geokinematical projects; *Geophys. Res. Abs.* **10** EGU2008-A-01735.
- Helmert F R 1880 *Die Mathematischen und Physikalischen Theorien der Höheren Geodäsie (Mathematical and Physical Theories of Higher Geodesy)*; Druck und Verlag von B.G. Teubner, Leipzig, 626p (in German).
- Holmstrom L and Koistinen P 1992 Using additive noise in back-propagation training; *IEEE Trans. Neural Netw.* **3**(1) 24–38.
- Hornik K, Stinchcombe M and White H 1989 Multilayer feedforward networks are universal approximators; *Neural Netw.* **2** 359–366.
- Hsieh W W 2009 *Machine Learning Methods in the Environmental Sciences*; Cambridge University Press, New York, 349p.
- Joseph V R 2006 Limit Kriging; *Technometrics* **48**(4) 458–466.



- Kaastra I and Boyd M 1996 Designing a neural network for forecasting financial and economic time series; *Neurocomputing* **10** 215–236.
- Kanellopoulos I and Wilkinson G G 1997 Strategies and best practice for neural network image classification; *Int. J. Remote Sens.* **18(4)** 711–725.
- Kanli A I 2009 Initial velocity model construction of seismic tomography in near-surface applications; *J. Appl. Geophys.* **67** 52–62.
- Karabork H, Baykan O K, Altuntas C and Yildiz F 2008 Estimation of unknown height with artificial neural network on digital terrain model; *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **37(1)** 115–118.
- Kavzoglu T and Mather P M 2003 The use of backpropagating artificial neural networks in land cover classification; *Int. J. Remote Sens.* **24(23)** 4907–4938.
- Kavzoglu T and Saka M H 2005 Modelling local GPS/levelling geoid undulations using artificial neural networks; *J. Geod.* **78(9)** 520–527.
- Kingston G B, Maier H R and Lambert M F 2008 Bayesian model selection applied to artificial neural networks used for water resources modeling; *Water Resour. Res.* **44(4)** W04419.
- Krige D G 1951 A statistical approach to some basic mine valuation problems on the Witwatersrand; *J. Chem. Metall. Min. Soc. S. Af.* **52(6)** 119–139.
- Krivoruchko K and Gotway C A 2004 Creating exposure maps using Kriging; *Public Health GIS News. Inf.* **56** 11–16.
- Kumar U A 2005 Comparison of neural networks and regression analysis: A new insight; *Expert Syst. Appl.* **29(2)** 424–430.
- Leandro R F and Santos M C 2007 A neural network approach for regional vertical total electron content modelling; *Stud. Geophys. Geod.* **51** 279–292.
- Li J and Heap A D 2008 *A Review of Spatial Interpolation Methods for Environmental Scientists*; Geoscience Australia Canberra, 137p.
- Liu Y, Starzyk J A and Zhu Z 2008 Optimized approximation algorithm in neural networks without overfitting; *IEEE Trans. Neural Netw.* **19(6)** 983–995.
- Masters T 1993 *Practical neural network recipes in C++*; Academic Press, San Diego, CA.
- Mathews L and Biediger J 2012 The impact of horizontal velocity on the APFO ground control database; [http://www.fsa.usda.gov/Internet/FSA\\_File/hv\\_paper\\_final.pdf](http://www.fsa.usda.gov/Internet/FSA_File/hv_paper_final.pdf).
- McClusky S, Balassanian S, Barka A, Demir C, Ergintav S, Georgiev I, Gurkan O, Hamburger M, Hurst K, Kahle H, Kastens, K, Kekelidze G, King R, Kotzev V, Lenk O, Mahmoud S, Mishin A, Nadariya M, Ouzounis A, Paradissis D, Peter Y, Prilepin M, Reilinger R, Sanli I, Seeger H, Tealeb A, Toksoz M N and Veis G 2000 GPS constraints on crustal movements and deformations in the Eastern Mediterranean (1988–1997): Implications for plate dynamics; *J. Geophys. Res.* **105(B3)** 5695–5719.
- Moghtased-Azar K and Zaletnyik P 2009 Crustal velocity field modelling with neural network and polynomials; In: *Observing Our Changing Earth* (ed.) Sideris M G (Berlin Heidelberg: Springer-Verlag), pp. 809–816.
- Moody J and Darken C 1989 Fast learning in networks of locally-tuned processing units; *Neural Comput.* **1(2)** 281–294.
- Neuner H 2010 Modelling deformations of a lock by means of artificial neural networks; In: *Proceedings of the 2nd International Workshop on Application of Artificial Intelligence and Innovations in Engineering Geodesy* (eds) Reiterer A, Egly U, Heinert M and Riedel Björn, Braunschweig, Germany, pp. 32–41.
- Neyamadpour A, Abdullah W W and Taib S 2010 Inversion of quasi-3D DC resistivity imaging data using artificial neural networks; *J. Earth Syst. Sci.* **119(1)** 27–40.
- Nocquet J M and Calais E 2003 Crustal velocity field of western Europe from permanent GPS array solutions, 1996–2001; *Geophys. J. Int.* **154** 72–88.
- Nørgaard M 1997 *Neural Network Based System Identification Toolbox*; Technical Report 97-E-51, Department of Automation, Technical University of Denmark, Copenhagen, Denmark, 37p.
- Novotny Z and Kostelecky J 2008 The estimated annual velocities of EUREF-EPN stations located in central European region; *Geophys. Res. Abs.* **10** EGU2008-A-02183.
- Ornthammarath T, Corigliano M and Lai C G 2008 Artificial neural networks applied to the seismic design of deep tunnels; [http://www.iitk.ac.in/nicee/wcee/article/14\\_04-01-0088.pdf](http://www.iitk.ac.in/nicee/wcee/article/14_04-01-0088.pdf).
- Panchal G, Ganatra A, Kosta Y P and Panchal D 2011 Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers; *Int. J. Comput. Theory Eng.* **3(2)** 332–337.
- Pariente D 1994 Geographic interpolation and extrapolation by means of neural networks; *Proceedings of EGIS/MARI'94* **1** 684–693.
- Pearlman M, Altamimi Z, Beck N, Forsberg R, Gurtner W, Kenyon S, Behrend D, Lemoine F G, Ma C, Noll C E, Pavlis E C, Malkin Z, Moore A W, Webb F H, Neilan R E, Ries J C, Rothacher M and Willis P 2006 Global Geodetic Observing System – considerations for the geodetic network infrastructure; *Geomatica* **60(2)** 193–204.
- Perez J A S, Monico J F G and Chaves J C 2003 Velocity field estimation using GPS precise point positioning: The south American plate case; *J. Glob. Position. Syst.* **2(2)** 90–99.
- Pérez-Pena A, Martín-Davila J, Gárate J, Berrocoso M and Buforn E 2010 Velocity field and tectonic strain in southern Spain and surrounding areas derived from GPS episodic measurements; *J. Geodyn.* **49** 232–240.
- Pinna G, Carcione J M and Poletto F 2011 Kerogen to oil conversion in source rocks. Pore-pressure build-up and effects on seismic velocities; *J. Appl. Geophys.* **74** 229–235.
- Piotrowski A P and Napiorkowski J J 2013 A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling; *J. Hydrol.* **476** 97–111.
- Plag H-P, Altamimi Z, Bettadpur S, Beutler G, Beyerle G, Cazenave A, Crossley D, Donnellan A, Forsberg R, Gross R, Hinderer J, Komjathy A, Ma C, Mannucci A J, Noll C, Nothnagel, A, Pavlis E C, Pearlman M, Poli P, Schreiber U, Senior K, Woodworth P L, Zerbini S and Zuffada C 2009 The goals, achievements, and tools of modern geodesy; In: *Global Geodetic Observing System – Meeting the Requirements of a Global Society on a Changing Planet in 2020* (eds) Plag H-P and Pearlman M (Berlin Heidelberg: Springer-Verlag), pp. 15–87.
- Poggio T and Girosi F 1990 Networks for approximation and learning; *Proc. IEEE* **78(9)** 1481–1497.
- Powell M J D 1987 Radial basis functions for multivariate interpolation: A review; In: *Algorithms for Approximation* (eds) Mason J and Cox M (Oxford: Clarendon Press), pp. 143–167.
- Ripley B D 1993 Statistical aspects of neural networks; In: *Networks and Chaos – Statistical and Probabilistic Aspects* (eds) Barndorff-Nielsen O E, Jensen J L and Kendall W S (London: Chapman & Hall), pp. 40–123.
- Rogers L L and Dowla F U 1994 Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling; *Water Resour. Res.* **30(2)** 457–481.

- Rumelhart D E, Hinton G E and Williams R J 1986 Learning internal representations by error propagation; In: *Parallel Distributed Processing*, vol. 1 (eds Rumelhart D, McClelland J and Group P R (Cambridge, MA: MIT Press), pp. 318–362.
- Schroeder M, Cornford D and Nabney I T 2009 Data visualisation and exploration with prior knowledge; In: *Engineering Applications of Neural Networks* (eds Palmer-Brown D, Draganova C, Pimenidis E and Mouratidis H (Berlin: Springer), pp. 131–142.
- Shahin M A, Jaksa M B and Maier R H 2008 State-of-the-art of artificial neural networks in geotechnical engineering; *Electron. J. Geotech. Eng.* **8** 1–26.
- Singh T N, Kanchan R K, Verma A K and Saigal K 2005 A comparative study of ANN and neuro-fuzzy for the prediction of dynamic constant of rockmass; *J. Earth Syst. Sci.* **114**(1) 75–86.
- Valach F, Hejda P and Bochník J 2007 Geoeffectiveness of XRA events associated with RSP II and/or RSP IV estimated using the artificial neural network; *Stud. Geophys. Geod.* **51** 551–562.
- Werbos P J 1974 *Beyond Regression: New Tools for Prediction and Analysis in the Behavioural Sciences*; Ph.D. Thesis, Applied Mathematics, Harvard University, Boston, MA.
- Witten I H and Frank E 2005 *Data mining: Practical machine learning tools and techniques*; (San Francisco, CA: Morgan Kaufmann), 560p.
- Yeung D S, Cloete I, Shi D and Ng W W Y 2010 *Sensitivity Analysis for Neural Networks*; Springer-Verlag, Berlin Heidelberg, 86p.
- Yilmaz M 2012 *The Utility of Artificial Neural Networks in Geodetic Point Velocity Estimation*; Ph.D. Thesis, Afyon Kocatepe University, Afyonkarahisar (in Turkish).
- Zhang G, Patuwo B E and Hu M Y 1998 Forecasting with artificial neural networks: The state-of-the-art; *Int. J. Forecast* **14** 35–62.

*MS received 14 May 2013; revised 31 October 2013; accepted 10 November 2013*