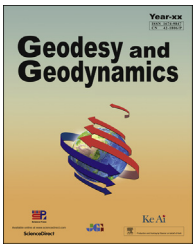
Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/en/journals/geog;
http://www.jgg09.com/jweb_ddcl_en/EN/volumn/home.shtml



Interevent times estimation of major and continuous earthquakes in Hormozgan region based on radial basis function neural network

M.R. Mosavi[a,*], M. Kavei[b], M. Shabani[b], Y. Hatem Khani[c]

[a] Department of Electrical Engineering, Iran University of Science and Technology, Narmak, Tehran 13114-16846, Iran
[b] Department of Physics, University of Hormozgan, Bandar Abbas, Iran
[c] Department of Electrical Engineering, University of Imam Khomeini Marine Sciences, Noshahr, Iran

ARTICLE INFO

Article history:
Received 12 November 2015
Accepted 28 December 2015
Available online 7 April 2016

Keywords:
Interevent times
Radial basis function
Neural networks
Earthquakes

ABSTRACT

This paper presents a new method to estimate the time of important earthquakes in Hormozgan region with magnitude greater than 5.5 based on the Radial Basis Function (RBF) Neural Network (NN) models. Input vector to the network is composed of different seismicity rates between main events that are calculated in convenient and reliable way to create optimized training methods. It helps network with a limited number of training data to estimation. It is common for earthquakes modeling by data-driven methods in this case. In addition, the proposed method is combined with Rosenberg cluster method to remove aftershocks events from the history of catalog for NN to better process the data. The results show that created RBF model successfully estimates the interevent times between large and sequence earthquakes that can be used as a tool to predict earthquake, so that comparison with other NN structure, for example Multi-Layer Perceptron (MLP) NN, reveals the superiority of the proposed method. Because of superiority proposed method has higher accuracy, lower costs and simpler network structure.

* *Corresponding author.*
E-mail address: M_Mosavi@iust.ac.ir (M.R. Mosavi).
Peer review under responsibility of Institute of Seismology, China Earthquake Administration.

Production and Hosting by Elsevier on behalf of KeAi

http://dx.doi.org/10.1016/j.geog.2016.03.004

1. Introduction

The earthquake is a natural phenomenon caused by the sudden release of energy stored in the ground created by seismic waves. Earthquakes occur naturally due to the nature of tectonic motions [1,2]. Earthquakes cause ground displacement and in some cases trigger a tsunami, it should be noted that in some cases, human activities are effective in occurrence of earthquakes. Forecasting and prediction in many cases are synonymous, but have the subtle difference. The present work is prediction [3]. Over time there have been significant efforts in earthquake prediction. Earthquake prediction is an interdisciplinary field of research in seismology, physics, geology, mathematics, computer science, engineering, and even social sciences. The US National Academy of Science defines earthquake prediction as the estimation of any one or more parameters of a future earthquake, namely, time of occurrence, epicentral location, and magnitude. Earthquake prediction studies can be categorized into two types, one is based on the analysis of earthquake precursor data, and the other is based on the analysis of historic earthquake data. Important earthquake precursors include changes in the earth's electromagnetic field [4], abnormal animal behavior [5,6], seismic quiescence [7], fault creep and continuous strain [8–10], and anomalous geochemical observations [11,12]. Changes in seismicity patterns are the most successful long-term precursors. Studies based on historical earthquake data often attempt to establish a magnitude–frequency relationship. The most popular distributions are the Gutenberg-Richter inverse power law distribution [13]. One of the hardest but best searching methods is the use of new and emerging accounting principles, such as Neural Networks (NNs) and evolutionary computation are particularly suited to solving complex problems. In general, the time scale earthquake prediction according to interval of the impending earthquake is classified as short, medium and long-term. The long-term prediction of natural disaster occurrence is one of the most sought-after goals in geoscience. Succeeding in such a goal involves obviating a multitude of difficulties; not only the proper variables which will act as precursors should be recognized and measured, but also the correlations between those variables and disaster occurrence should be identified. In spite of the significant progress over the last 20 years [14–16], the determination of such correlations remains a difficult endeavor as the governing relationships are usually rather complex and nonlinear, and the mechanisms creating the respective correlations are only recently coming to be understood [17,18].

Seismicity databases (catalogs) are the most popular source of data for long-term prediction studies for a number of reasons (including: abundance, existence for almost all regions of the world, availability on a continuous basis). On the other hand, NNs are powerful mathematical tools [19] that simulate the way that the human brain deals with information and the procedure of learning. Recently, efforts have been made to investigate the potential of Artificial Neural Networks (ANNs) as a tool for system behavior simulation that are governed by nonlinear multivariate and generally unknown interconnections within a noisy, poorly-controllable physical environment. The choice of the ANN approach is motivated by the lack of clear causal relations between seismicity patterns and related crustal environments. In addition, the smart methods have higher precision, lower cost and easier calculations than traditional and classical methods. It seems attractive to us considering the seismicity rates, because many seismologists share the view that changes in seismicity rates can occur as part of the process of preparation for large earthquakes [19,20]. Therefore, during the early 1980's the impact of man-made effects on seismicity rates was demonstrated for the first time [7].

2. Geological setting of studied region

Hormozgan region is in north of the Strait of Hormuz in southern Iran, and is one of the most strategic parts of the world politically and economically. Coastal zones of the region are on the east of Oman. The historical record confirms that some areas, for example, Zagros and Makran, in terms of seismicity are active of 2500 years ago, and this reflects the long-term nature of seismicity areas [21]. Hormozgan region as shown in Fig. 1 is located inside the interface between the geographical coordinates 25°24′N to 28°57′N and 53°41′E to 59°15′E from Greenwich meridian [22].

3. Neural networks

Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) NNs are briefly explained in this section.

3.1. *RBF Neural networks*

RBF network is a kind of forward NNs composed of three layers including input layer, hidden layer and output layer. Each of these layers has different roles, respectively [23]. In RBF networks, outputs are determined by calculating the distance between network inputs and the centers of the hidden layer. The second layer is a hidden linear layer, and outputs of this layer weight bearing samples from the outputs of the input layer. Each hidden layer neuron with a vector parameter called the center. Therefore, a general description of the network is given by equation (1) [24]:

$$\hat{y} = \sum_{i=1}^{I} wi\phi(\|\mathbf{x} - ci\|) + \beta \quad (1)$$

Standard mode is usually the Euclidean distance, and RBF is intended with Gaussian function as equation (2):

$$\varphi(r) = \exp\left(-\alpha i\|\mathbf{x} - ci\|^2\right) \quad (2)$$

In equations (1) and (2), the following definitions are considered: $i \in \{1, 2, 3, \ldots, I\}$, so I is the number of neurons in the hidden layer; w_i, weight between neuron in the hidden layer and output; φ, Gaussian function; α_i, spread parameter (amount of variance) neuron; $\mathbf{x}$, input data vector; c_i, center vector of neuron; β, bias of output; $\hat{y}$, output of the network.

Fig. 2 shows the schematic overview of a RBF network. M-dimensional inputs (x_1,…, x_m) are placed in the input layer.

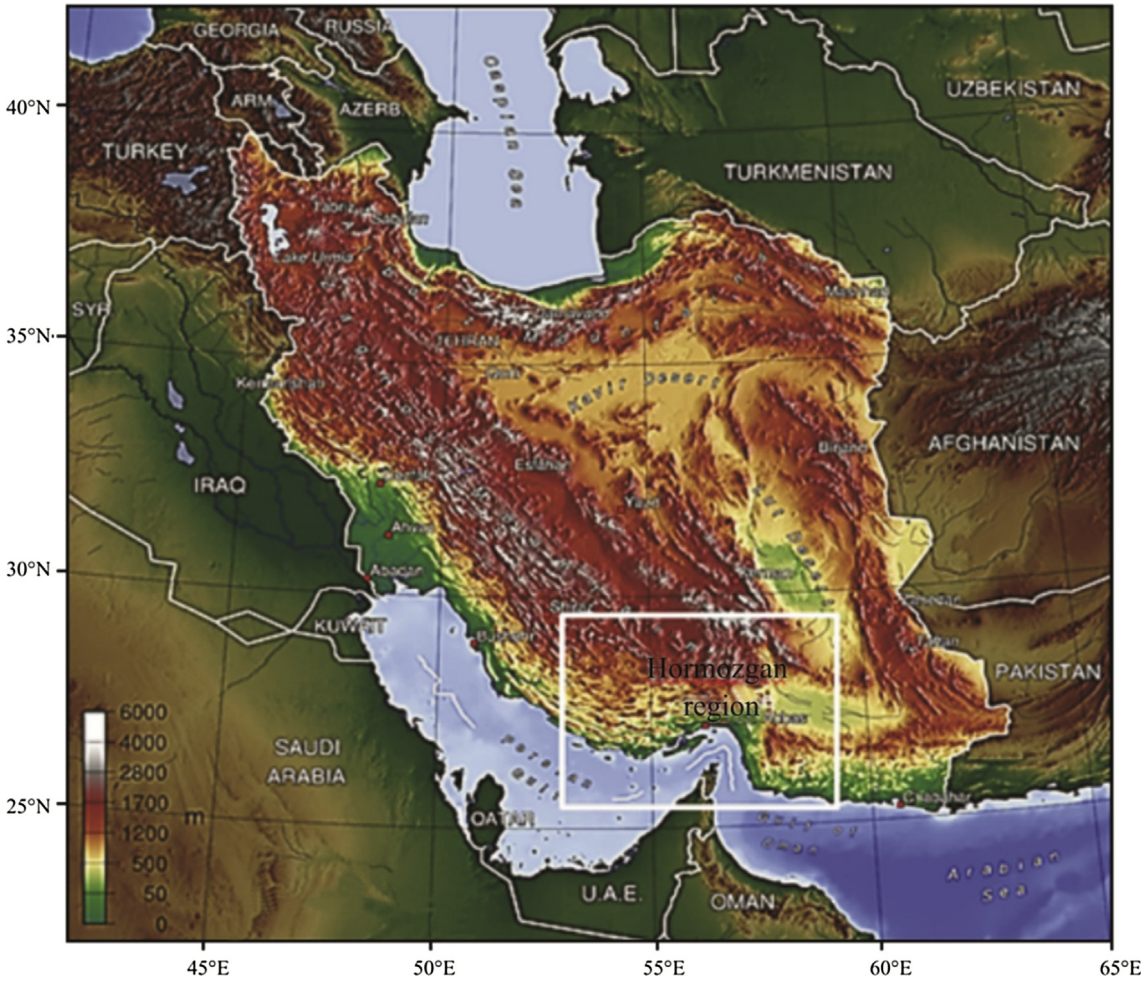


Fig. 1 – Geographical location of Hormozgan region.

Hidden layer include I neurons and each neuron in this layer is calculated in terms of the Euclidean distance between the centers and inputs. Hidden layer's neuron has an activation function that is called RBF function. RBF function is a selected Gaussian function which encompasses the spread parameters (α_1, …, α_i) to adjust the shape of the curve. Outputs of hidden layers are transmitted to the output layer through weights (w_1, …,w_i). Here I show the number of hidden layer neurons. Output layer is linear combination of the hidden layer's outputs and bias parameter β. Finally, $\hat{y}$ is obtained as RBF's output.

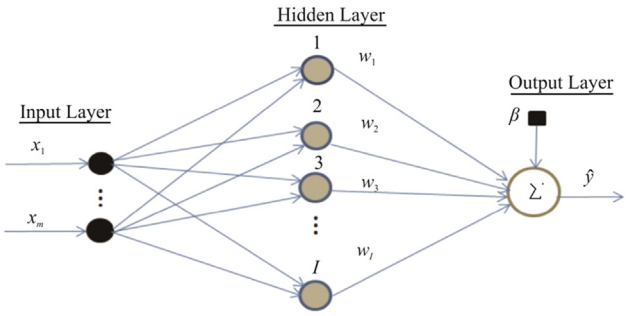


Fig. 2 – Schematic overview of a RBF NN.

RBF NN design method should determine the number of neurons in the hidden layer. α, c and β must be set correctly, in order to obtain the desired output of RBF NN parameters w. Root Mean Square Error (RMSE) can be used to evaluate network performance. The error for RBF network can be defined as:

$$E_{RMSE} = \sqrt{\frac{\sum_{i=1}^{I}(y-\hat{y})^2}{I}} \quad (3)$$

Here y is the desired output and $\hat{y}$ shows output of RBF NN. And minimizing the error function is the RBF NN training method [24,25].

3.2. *Multi-layer perceptron neural network*

MLP NN is able to simulate various issues. In most important part of MLP networks, the number of middle layers and number of neurons should be determined by the input and output parameters. MLP can do complex classifications by using sufficient perceptron layers in the network. These classifications by sigmoid transfer functions including LogSig, TanSig and PureLin are related to the hidden layer MLP NN.

Fig. 3 is a view of MLP network. In this network, neurons of each layer are connected to neurons the previous layer. Output of each layer after influence function, is the input of next layer, and this process will continue until the result to be obtained [26–28].

4. Methods and process

In this section, we review and analyze the work process in a rundown look.

4.1. *Earthquakes catalog of study region*

In this paper, we obtained earthquakes catalog from International Institute of Seismology and Engineering. According to Fig. 4, time period (time window) is years between 1900 and late 2014, and geographical coordinates (the location window) is 25°24′N–28°57′N, and 53°41′E–59°15′E. Number of events in this window is 3840. The selection reason for this spatial and temporal window is it covering all earthquakes of the Hormozgan region. Hormozgan region and its earthquakes has been separated by using ARCGIS software carefully (Fig. 5).

The first step for estimation the earthquakes in the region, is providing a valid catalog of all historical and devices events. In the following, the evaluation of the uncertainty of earthquake parameters is important. Obvious uncertainty in earthquake parameters is based on the number of local and regional stations, stations distribution and velocity models. By comparing the location, depth and magnitude reported by different seismology centers can be evaluated the uncertainties for each earthquake and its center. Determination of the location of an earthquake for the Middle East before the early 1970s had little accuracy. Based on the quality and quantity of available information, determined uncertainty of earthquake parameter for Iran in three periods of time for the historical period (pre 1900) uncertainty in depth, location and magnitude represent in 30 km, 100 km, 0.4 to 0.8, respectively; for the early instrumental period (1900–1963), 20 km and 0.3 to 0.5 was calculated for location and magnitude; and for the modern instrumental period (1964–1994), location errors for moderate and major earthquake are about 15 km and 10 km, and direct assignment of the surface wave magnitude *Ms* may contain 0.2 to 0.4 magnitude, units of error and *Ms* values from the conversion of, the body wave magnitude *Mb*, suffer 0.45 to 0.67 magnitude units of uncertainty [21]. So that amount of uncertainty is decreased along the time. Therefore, in this research IT is considered a valid catalog from 1964 onwards and for the 1609 earthquake.

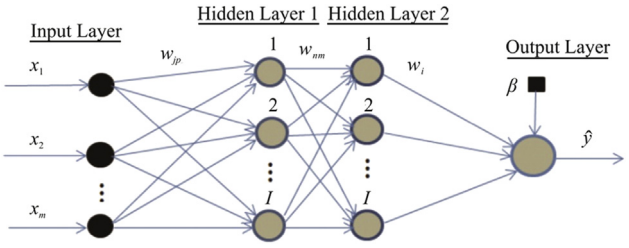


Fig. 3 – Schematic overview of a MLP NN.

4.1.1. *Catalog magnitude scale*

Since International Institute of Seismology and Earthquake Engineering reported earthquake magnitude on different scales, the scales need to be unified to prepare a catalog with uniform magnitude scale [21]. In this study, we convert all magnitudes to *Mb*, because most earthquakes have been reported with *Mb* scale, the following relationships is used for conversion [22,29].

$$Mb = 2.41 + 0.558Ms \quad 4.0 \leq Mb \leq 6.2 \tag{4}$$

$$Mw = 0.85Mb + 1.03 \quad 3.5 \leq Mb \leq 6.2 \tag{5}$$

There are many studies about correlation between the local magnitude *Ml* and the moment magnitude *Mw*. But it is

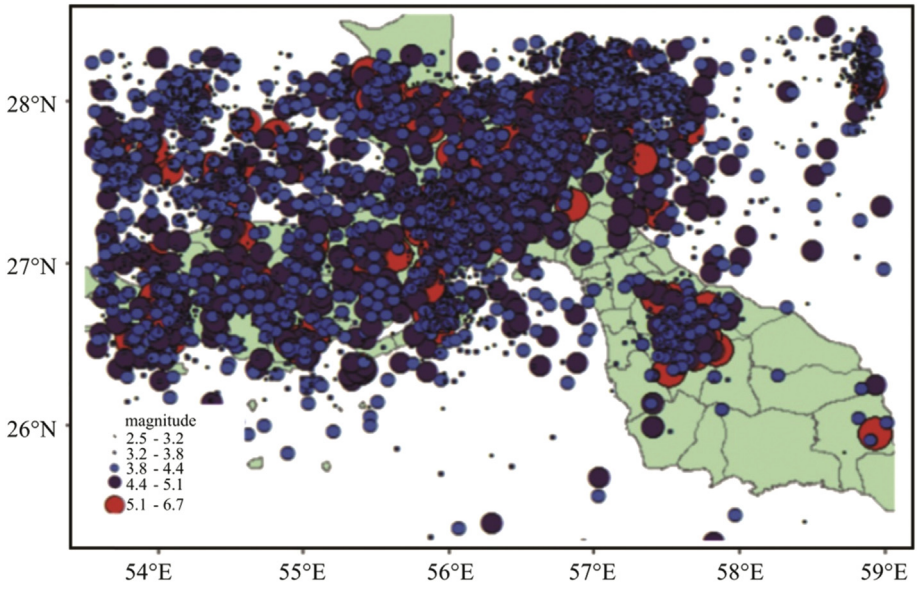


Fig. 4 – Map of geographical coordinates (the location window) and earthquakes distribution in time period of study.

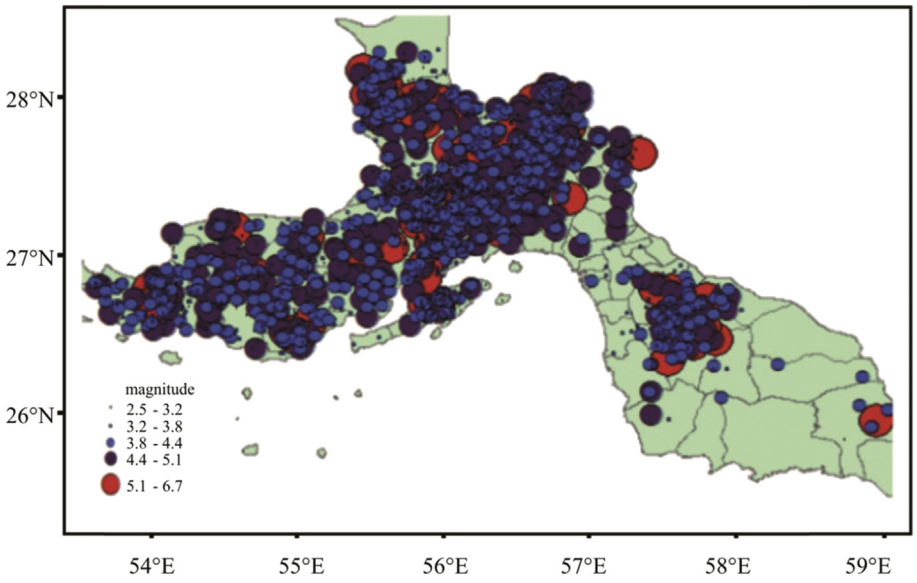


Fig. 5 – Separating earthquakes of Hormozgan region using ARCGIS software.

impossible to define unique global relations between Ml and Mb or other magnitude scales. Therefore, we could only trust the internal correlation which exists in about 2000 records with both Mw and Ml. This correlation is obtained from 2271 records in the time period of 1975–2010, shown in Fig. 6. The divergence or distribution of the points is minimum, and the $R^2 = 0.98$ is also convincing to get this equation at this stage [21].

$$Mw = 1.0136Ml - 0.0502, \quad 4.0 \leq Ml \leq 8.3 \tag{6}$$

Thus, according to equations (4)–(6), all related magnitudes are converted to scale Mb and are evaluated.

4.1.2. *Completeness magnitude variations through ZMAP*

The magnitude of completeness Mc is theoretically defined as the lowest magnitude at which 100% of the earthquakes in a specific space-time range are detected. Mc determination of instrumental earthquake catalogs is an essential and compulsory step for any seismicity analysis. There are several methods to determine Mc. Two main methods are the traditional and common methods for estimating completeness of a catalog. One is using the cumulative frequency–magnitude distribution (i.e., $\log N = a - bM$, where N counts the number of earthquakes with magnitude greater than or equal to magnitude M; and a and b are seismicity and zone-dependent constants). Another Mc determination method is using the frequency–magnitude distribution to apply the seismological analysis in ZMAP software [30]. For instrumental earthquakes, depending on the region and accuracy of seismogram recorded, Mc can be presented by recorded events which have different value [21]. If different time periods in this study are considered for raw catalog, magnitude of completeness variations can be studied clearly. As is shown in Fig. 7 for years before 1997 the biggest magnitude recorded is 4 or more, but in the following years to 2005, the amount declined to about 3. A reason of these changes is possibly related to increase or decrease in the number of stations in that period [31].

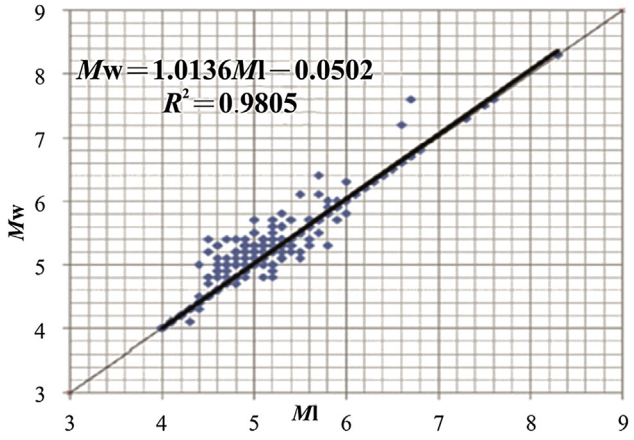


Fig. 6 – Correlation between Ml with Mw in the catalog [21].

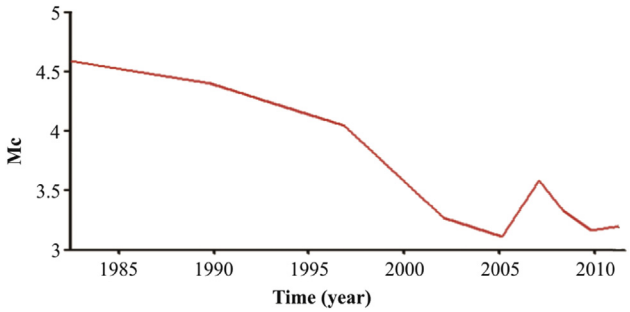


Fig. 7 – Temporal variations magnitude of completeness in raw catalog.

4.1.3. *Magnitude of completeness raw catalog in Hormozgan province*

According to the available catalog for provinces of magnitude greater than or equal to 3.1 earthquakes in the region (Fig. 8), the magnitude of completeness value is obtained equal to 3 in this region by using ZMAP software.

4.2. *Major and continuous earthquakes interevent times*

This paper aims to estimate the occurrence of large earthquakes, the output (target) variable of the RBF network was selected to be the interevent time between two consecutive main events. Thus, a threshold M was defined and each event exceeding M in magnitude was considered as a main event. We define the interevent time T_M^k as follows:

$$T_M^k = t^k - t^{k-1} \tag{7}$$

where t^k denotes the occurrence time of the kth event, and t^{k-1} denotes the occurrence time of the $(k-1)$th event in the catalog, both of which are greater in magnitude than the value M [32]. Threshold M in this research is set equal to 5.5 for the region according to investigated data and based on the modified Mercalli and reviews. Finally, interevent times are calculated for Hormozgan region according to equation (7).

In this study, real-time or origin-time is the occurrence time of the recorded earthquake, according to the International Institute of Earthquake Engineering and Seismology [33].

4.3. *Seismicity patterns*

Seismicity pattern studies gather data on the distribution of earthquakes in space, time and size (hereafter referred to as the first-order moment), examining seismicity rate, location and migration of foreshocks and aftershocks, gaps, donuts and variations in b value. In some cases, subjectively recognizable patterns (e.g., foreshock cluster, gap, b value anomaly) are observed which closely resemble those predicted for various theories of seismogenesis in reference [31]. Changes in seismicity activity are stresses reflector and preparation stages of earthquakes [34–37]. In any case, uncertainty physical description generation of earthquake and lack of relationships between seismicity patterns and the skin environments are limitation factors for the development of prediction. Data-driven nets as alternative to classical approximation methods to predict earthquakes are so important that during the last few years it has been taken into consideration [32].

4.3.1. *Seismicity rates*

Change in seismicity rate is important as it shows the preparation of large earthquakes. The necessary catalog parameters for seismicity rate analysis are earthquake origin times and magnitudes. Seismicity rate variations can be well illustrated by the cumulative number curves. Artificial changes in seismicity rate changes arise mainly from three categories: (1) detection changes arising from increased capability of a network to recognize and locate events due to installation of new stations, (2) systematic changes in magnitudes of events (magnitude shifts), caused by changes in the station distribution chosen to determine magnitudes, and (3) reporting changes which are related to lack of reporting magnitudes for detected events. As these changes will lead to ineffective seismicity rates [20]. Selecting the proper input variables is of paramount importance for the prediction abilities of the resulting NN model. In this paper, the inputs were selected as the seismicity rates between main and continuous events. There are many reasons for this selection: (1) seismicity rates can characterize the strain accumulation and release process on the average, (2) they are very easy to calculate and (3) their calculation is reliable [32]. Seismicity rate as the rate of earthquake occurrence in the time interval between the main events numbered k and $k-1$ is:

$$R_{M_1,M_2}^k = \frac{N_{M_1,M_2}^k}{T_M^k} \tag{8}$$

where N_{M_1,M_2}^k is the number of earthquake events that are larger in magnitude than M_1 and smaller than M_2 and have occurred between main events numbered k and $k-1$. Previous studies showed that estimation accuracy increases when more than one seismicity rate is given as input, attributing the fact to the existence of clear differences in the variations of seismic rates with the size of events. As the results of our experiments were consistent with this hypothesis, we

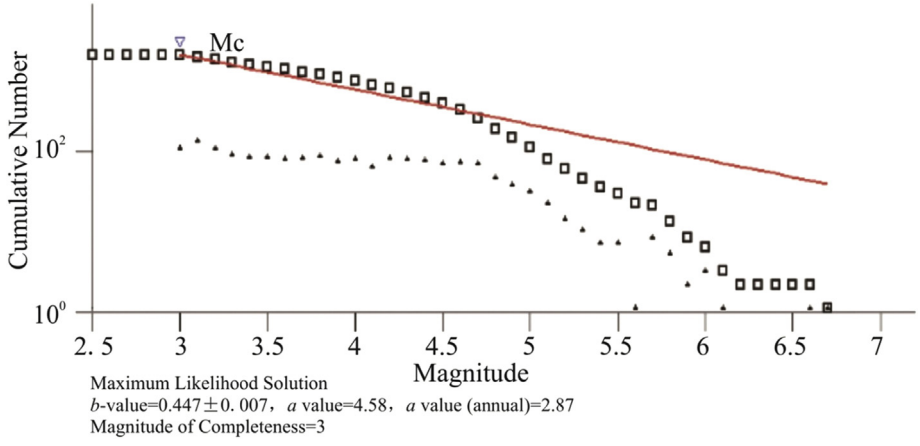


Fig. 8 – Completeness magnitude of raw catalog.

allowed the input vector to contain several values of different seismicity rates, which were calculated for different values of M_1, M_2 and k. The exact number of seismicity rates to be chosen as input variables was optimized based on numerical experiments, as will be discussed later. It would be intuitive to include as input seismicity rates of the type $R^k_{M_1,M_2}$ for the prediction of the corresponding interevent time T^k, as they contain the most recent information preceding the main event numbered k. However, such an inclusion would significantly diminish the practical value of the resulting NN as a predictive tool. The reason is that, in order to make a prediction for the interevent time T^k between main events $k-1$ and k, one would have to wait until main event k has already occurred, so that the rate $R^k_{M_1,M_2}$ is available. In this case, prediction of the interevent time T^k would be of no practical use. Thus, the input vector for prediction of the interevent time T^k should contain past seismicity rates from the period between main events $k-1$ and $k-2$, i.e., $R^k_{M_1,M_2}$. More than one period between past main events could be utilized to produce additional input variables. Fig. 9 depicts a visual example for the definition of the seismicity rate $R^k_{M_1,M_2}$, in relation with main events numbered $k-1$ and $k-2$.

The bounding parameters M_1 and M_2 define the range of magnitudes used for calculating the corresponding seismicity rate $R^k_{M_1,M_2}$. The selection of the lower bounding parameter M_1, which acts as a cutoff magnitude, is important in order to avoid producing a non-homogenous set of training data. A sufficiently high value for M_1 should be applied, guaranteeing that variations in the calculated seismicity rate are not due to differences in the way which earthquakes are recorded (e.g., changes in the sensitivity of earthquake recorders), but in fact reflect a change in seismicity dynamics. The securing of homogeneity data set by properly selecting the cutoff magnitude is important. It is also possible to introduce additional bounding parameters, M_3, M_4 … M_z, thus segmenting the magnitude range in more than one zone. In this case, the total number of seismicity rates used as input variables for each period between older main events is equal to $z-1$. Assuming that p periods between older main events are taken into account, a total of $p(z-1)$ possible input variables to the NN is produced [32].

Based on what was described here will calculate the seismicity rates for region. To do this work, according to the interevent times of the major events in the catalog, border parameters M_1 and M_2 is selected to calculate the seismicity rates magnitude range $R^{k-1}_{M_1,M_2}$. According to the province raw catalog reviews and magnitude of changes, at first M_1 equals to 3 and the M_2 equals to 4.2; and then M_1 equals to 4.2 and the M_2 equal to 6.7. It should be noted that because of incomplete catalog in some intervals, and few earthquakes recorded by research and study, border parameters are the best choice. By selecting border parameters M_1 and M_2, two input vector seismicity rate for Hormozgan region is defined as follows:

$$R^{k-1}_{3,4.2} = \frac{N^{k-1}_{3,4.2}}{T^{k-1}_{5.5}} \tag{9}$$

$$R^{k-1}_{4.2,6.7} = \frac{N^{k-1}_{4.2,6.7}}{T^{k-1}_{5.5}} \tag{10}$$

Thus the input vector for a NN is:

$$R^{k-1}_{M_1,M_2} = \left[\frac{N^{k-1}_{3,4.2}}{T^{k-1}_{5.5}} \; \frac{N^{k-1}_{4.2,6.7}}{T^{k-1}_{5.5}}\right] \tag{11}$$

According to equation (11) seismicity rates are calculated for successive interevent times and major events in catalog.

4.3.2. *Declustering*

Seismicity declustering is the process of removing foreshocks and aftershocks from the mainshock. The identification of background earthquakes is important for many applications in seismology with regard to seismic hazard assessment, development of clustered seismicity models, earthquake prediction research, and seismicity rate change estimation [32]. In this paper, the ZMAP software was used for declustering these events. There are four algorithms to do decluster in this software. Each algorithm considers different time and distance ranges for declustering. Reasenberg is used widely in linked-window method and Reasenberg's method windows are designed larger in space but shorter in time for larger shock. Up to now, most users

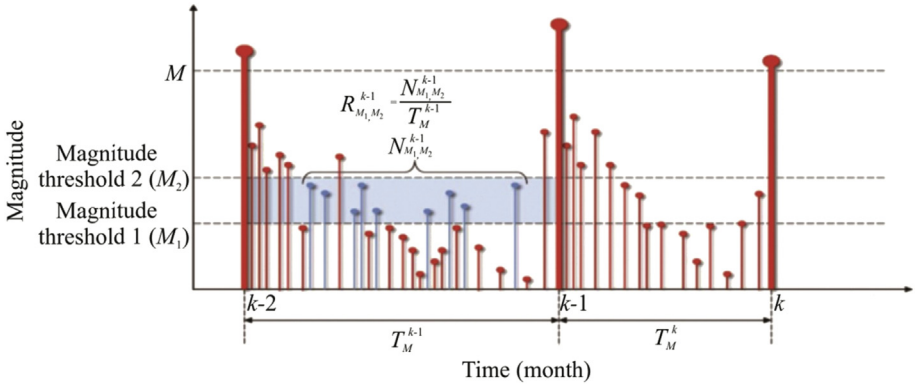


Fig. 9 – Visual example for the definition of the seismicity rate $R^k_{M_1,M_2}$, in relation with main events numbered $k-1$ and $k-2$ [32].

have applied either the algorithm of Gardner and Knopoff (1974) or Reasenberg (1985), mainly because of the availability of the source codes and the simplicity of the algorithms. Each catalog needs to omit aftershocks and foreshocks; on the other hand, main shocks should be separated for use in the final catalog. Declustering method is one of the most popular methods of seismology society that it has become a routine procedure using standard parameter values Table 1 [21].

In Table 1, τ_{min} is the minimum value of the look-ahead time for building clusters when the first event is not clustered, τ_{max} is the maximum value of the look-ahead time for building clusters, P is the probability of detecting the next clustered event used to compute the look-ahead time, τ and x_k is the increase of the lower cut-off magnitude during clusters.

$$x_{meff} = x_{meff} + x_k M \tag{12}$$

where M is the magnitude of the largest event in the cluster, x_{meff} is the effective cutoff magnitude for catalog, r_{fact} is the number of crack radii surrounding each earthquake within new events considered to be part of the cluster (unit: Km). Rosenberg cluster method [31] is organized to remove the aftershocks and declustering the data, since the existence of these things will cause problems to the implementation of the network [32]. Due to the earthquake events, there is general assumption in seismicity studies [38]. Aftershocks are detected, in declustering methods, based on the spatio-temporal approximation and the higher average seismicity rates than previous earthquakes. In this study, we use the cluster approach introduced by Rosenberg [31].

This method identifies the aftershocks by linking earthquakes to clusters according to spatial and temporal interaction zones. Moreover, Reasenberg's procedure is free from assumptions with regard to the spatial aftershock distribution, and it describes their migration given that the background seismicity is low. The spatial extent of the interaction zone is chosen according to stress distribution near the mainshock area. Reasenberg's local nearness of events is described by the spatial threshold d, depending on magnitude according to:

$$\log(d) = 0.4M_0 - 1.93 + k \tag{13}$$

where M_0 is the magnitude, k is equal to 1 for the distance to the largest earthquake and equal to 0 for the distance to the last event, and d is in kilometers. The model is a simple circular fault model with radius d. The Keilis-Borok formula defines the seismic moment for static cracks as $\frac{16}{7}\Delta\sigma d^3$, where $\Delta\sigma$ is the stress drop. Its temporal extension is based on Omori's law. Each subsequent event is linked with the largest event or with the last one in each cluster, which has formed until current time. It should be also mentioned that overlapping clusters are joined [39].

In order to obtain a confidence of probability p_1 for observing the next sequence event, the time interval τ is:

$$\tau = \frac{-\ln(1-p_1)t}{10^{\frac{2(\Delta M-1)}{3}}} \tag{14}$$

$$\Delta M = M_{mainshock} - Mc \tag{15}$$

where Mc is completeness magnitude. It has become common practice in seismological studies to use standard parameters provided by ZMAP software packages [32]. In this study, to achieve the primary goals, declustering catalog is used only with standard parameters mentioned in Table 1. In ZMAP software is adjustable, without calculation of the equations (12)–(15) for the region. Thus the use of Rosenberg cluster algorithm for the 1609 earthquake in Hormozgan region led to the identification of 68 clusters of 17.03% of the catalog. The remaining events were not associated with any cluster, and the number of which is included the 1335 earthquake.

Table 1 – The standard input parameters for declustering algorithm by Reasenberg method [21].

Parameter	Standard	Simulation range	
		Min	Max
τ_{min} (days)	1	0.5	2.5
τ_{max} (days)	10	3	15
P	0.95	0.9	0.99
x_{meff}	4.0	0	1
x_k	0.5	1.6	1.8
r_{fact}	10	5	20

4.4. *Calculation of interevent times major and continuous events for declustered catalog*

In this section, interevent times calculation between two major and continuous event using equation (7) for declustered catalog has been investigated. After declustering and remove aftershocks in the raw catalog which included the 1609 earthquakes in the region, thus results declustered catalog includes 1335 earthquakes. The number of large earthquakes during the period 1965 to late 2014 is equal to 26 earthquakes. Analysis of the raw catalog reveals that there are four earthquakes with Mb > 5.5 as aftershocks in raw catalog which are removed by declustering. 22 earthquakes remain as the main earthquake. In this study, major and continuous earthquakes interevent times based on earthquake real-times is very important for earthquake prediction analysis. This graph increase visual analysis of targets and help users research and understand it. Earthquake occurrence real-time chart in terms of interevent times is shown in Fig. 10.

4.5. *Calculation seismicity rates by declustered catalog*

According to what described in Section 4.3, to calculate the seismicity rates, first major and continuous earthquakes interevent times in the catalog are carefully considered, and catalog boundary parameters M_1 and M_2 is selected to calculate $R^{k-1}_{M_1,M_2}$. It should be noted that the optimal parameters of the boundary for declustered catalog must be the previous values of the boundary parameters M_1, M_2 and so M. With selection boundary parameters M_1, M_2, two input vector for Hormozgan region for declustered catalog based on equations (9) and (10), and NN input vector by equation (11) is defined.

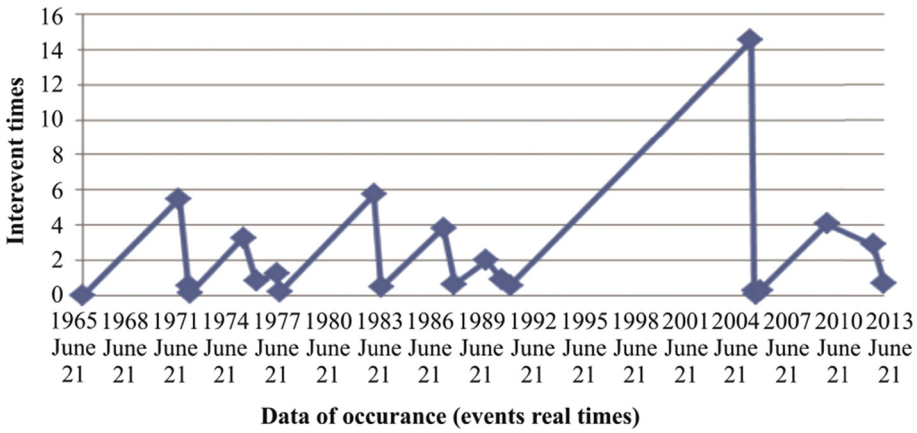


Fig. 10 – Graph of major and continuous earthquakes interevent times based on earthquake real-times.

Thus, seismicity rates are calculated for all intervals in declustered and raw catalog.

5. Evaluation of proposed approach

In this section, we analyze the results and discuss its simulations.

5.1. *Catalog data as NN input*

NN input data, including seismicity rates, major and continuous earthquakes interevent times are defined as matrices. Every NN is consisted of two phases "training" and "test". About 70 percent of the data randomly and optimized is chosen for training, and the remaining is for testing.

5.2. *Determination the optimal parameters in RBF network*

RBF network effective parameters are listed in Table 2, and their optimum values are obtained by trial and error with regard to the input raw data and declustered catalog. This means that by fixing one parameter, another parameter has been checked by different values to ensure optimal value. Therefore, determination of a parameter value assessed another parameter value. Finally, the optimum parameters for the data listed in Table 2 are considered.

5.3. *Determination the optimal parameters MLP network*

Effective parameters in a MLP network are showed in Table 3. Their optimum value is obtained with respect to input raw data and declustered catalog by trial and error. This means that by fixing one parameter, another parameter has been checked by different values to ensure optimal value. Therefore, determination of a parameter value assessed another parameter value. Finally, the optimum parameters listed in the Table 3 are considered for this study.

5.4. *Analysis of prediction results*

To evaluate the performance of network prediction, the results is compared with each other.

5.4.1. *Results of MLP network prediction*

Prediction results of the MLP network is shown in Figs. 11–13, respectively. Fig. 11 show the MLP NN output with declustered catalog data, Fig. 12 shows

Table 2 – The standard input parameters for RBF network.

Parameter	Standard	Simulation range		
		Min	Max	Optimum
Goal	0.0	–	–	0.01
Spread	1.0	0.001	2	0.02
MN	Default is matrix of Q input	2	14	10
DF	25	1	15	2
Function	Gaussian	–	–	Gaussian

Table 3 – The standard input parameters for MLP network.

Parameter	Simulation range		
	Min	Max	Optimum
S_i	1	10	2
Number of neurons	2	30	15
TF_i	–	–	TanSig
BTF	–	–	Trainlm
BLF	–	–	Learngdm

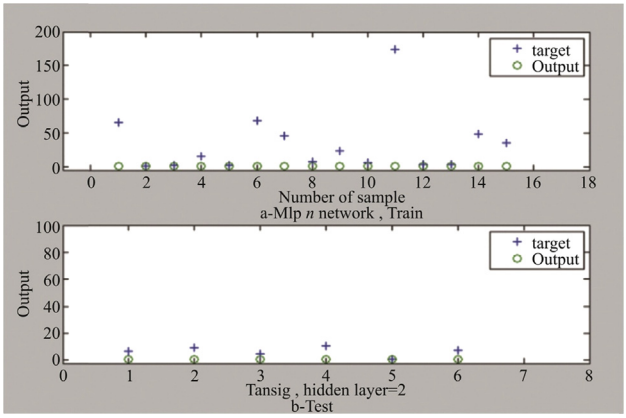


Fig. 11 – Network output with transfer function Tansig and 2 hidden layers.

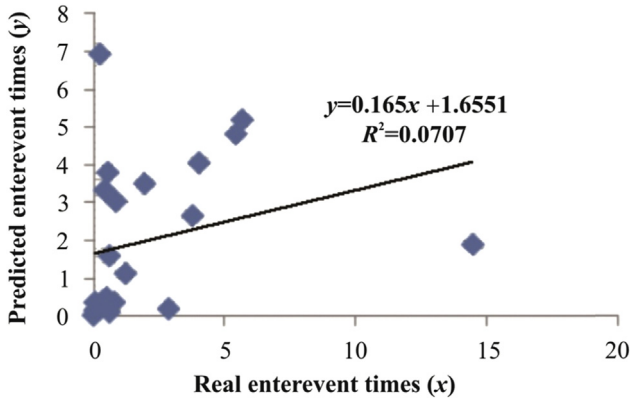


Fig. 12 – Correlation between predicted and real interevent times.

correlation between predicted and real interevent times, and Fig. 13 shows important and continuous earthquakes predicted interevent times versus the real time events.

It should be noted that MLP NN is trained by declustered and raw catalog optimized input vectors, so it is considered as a standard structure with transfer function Tansig and 2 hidden layer with 15 neuron in each layer and the number of layers range varies from 1 to 10. All possible combinations are tested and RMSE of the best mode for declustered and raw catalog equal to 6.66 and 17.79, respectively.

5.4.2. *Results of RBF network prediction*

Prediction results of the RBF network are shown in Figs. 14–16, respectively. Fig. 14 shows RBF NN output with declustered catalog data. Fig. 15 shows correlation between predicted and real interevent time. Fig. 16 shows important and continuous earthquakes predicted interevent times versus the real time events.

The entire process for raw catalog data with RBF NN is similar. By comparing the results of the RBF network output forms with raw data and declustered data, it is quite evident that the network with declustered data with optimize input data has better performance, as experimental tests have fewer error and better predictions than the raw data. So its RMSE for declustered and raw catalog input is 1.48 and 3.17, respectively.

Table 4 shows the results of both networks. Despite optimized parameters and input vector, MLP NN can't make meaningful and significant relationship between the input values and the interevent times with important and successive earthquakes, on the contrary, RBF NN provides a better prediction.

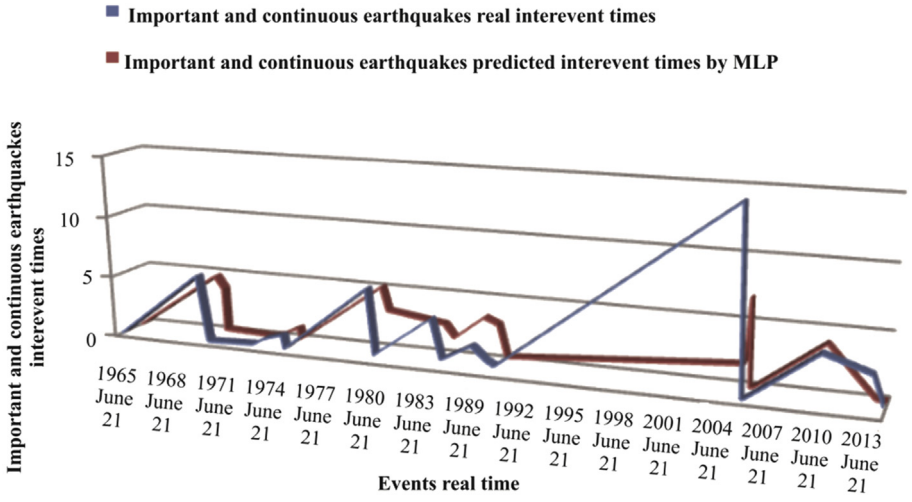


Fig. 13 – MLP network-important and continuous earthquakes predicted interevent times versus the real time event (declustered catalog).

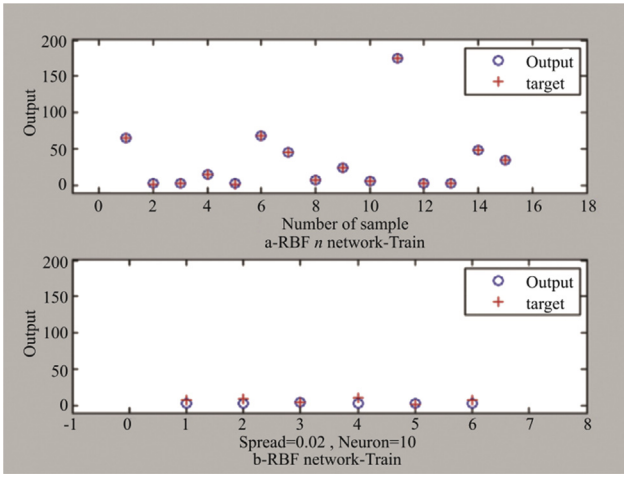


Fig. 14 – RBF NN output data by declustered earthquakes catalog data with neuron number = 10 and spread = 0.02.

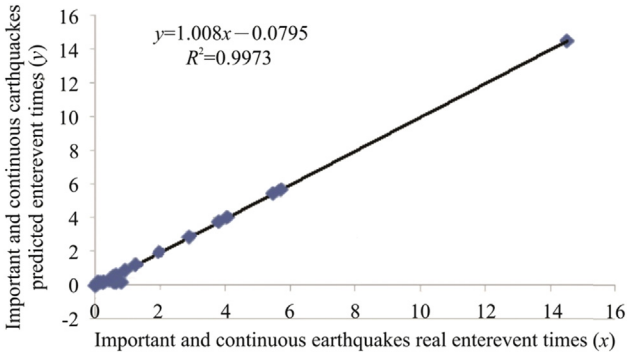


Fig. 15 – Correlation between predicted and real interevent times.

6. Conclusion

This paper introduced a new methodology to estimate the interevent times of significant earthquake events based on RBF NNs, given a database of historic seismicity data. In order to formulate the models, the input variables were selected among different seismicity rates, whereas the output variables were the interevent times between significant seismic events. The methodology was applied to the Hormozgan earthquakes catalog. The results stressed the importance of removing the aftershock events since a reliable estimation cannot be obtained using the raw catalog. To this end, the Reasenberg technique was applied to decluster the raw catalog. Following the application of the Reasenberg technique, the input vector to the NN model was optimized. The resulting predictions revealed a strong correlation of the input variables with the interevent times, thus was confirmed the applicability of the proposed approach is to successfully estimate large earthquakes interevent times. A different NN structure, namely, the MLP structure, was also used to analyze the same data. After repeating training method for both MLP and RBF network was revealed that the RBF NN to predict is best network with RMSE equal 1.48 for declustered catalog because of improved input data (seismicity rates). A comparison between the two techniques highlighted the superiority of the RBF network models, in terms of higher estimation accuracy and simpler network structure. Finally, results confirmed superiority of both the declustering method and the RBF NN to predict interevent times between major and continuous earthquakes in Hormozgan region.

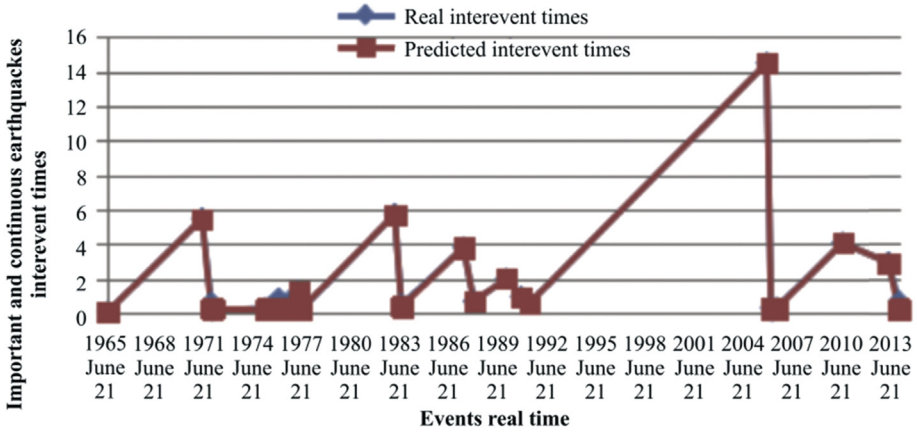


Fig. 16 – RBF network-important and continuous earthquakes predicted interevent times versus the real time event with remove aftershocks.

Table 4 – Comparison of the results for different modes.

Catalog	Input vector	Number of data	NN type	RMSE
Raw ($M > 5.5$)	$\begin{bmatrix} R^{k-1}_{3,4.2} \\ R^{k-1}_{4.2,6.7} \end{bmatrix}$	25	RBF	3.17
			MLP	17.79
Declustered ($M > 5.5$)	$\begin{bmatrix} R^{k-1}_{3,4.2} \\ R^{k-1}_{4.2,6.7} \end{bmatrix}$	21	RBF	1.48
			MLP	6.66

REFERENCES

[1] Jin S, Occhipinti G, Jin R. GNSS ionospheric seismology: recent observation evidences and characteristics. Earth Sci Rev 2015;147:54–64.
[2] Jin S, Park PH, Zhu WY. Micro-plate tectonics and kinematics in Northeast Asia inferred from a dense set of GPS observations. Earth Planet Sci Lett 2007;257:486–96.
[3] Reyes J, Esteban AM, Álvarez FM. Neural networks to predict earthquakes in Chile. Appl Soft Comput 2013;13(2):1314–28.
[4] Serebryakova ON, Bilichenko SV, Chmyrev VM, Parrot M, Rauch JL, Lefeuvre F, et al. Electromagnetic ELF radiation from earthquake regions as observed by low-altitude satellites. Geophys Res Lett 1992;19:91–4.
[5] Tributsch H. When the snakes awake – animals and earthquake prediction. Boston: MIT Press; 1982.
[6] Kirschvink J. Earthquake prediction by animals, evolution and sensory perception. Bull Seismol Soc Am 2000;90(2):312–3.
[7] Habermann RE. Precursory seismic quiescence: past, present, and future. Pure Appl Geophys 1988;126:279–318.
[8] Zion YB, Eneva M, Liu Y. Large earthquake cycles and intermittent criticality on heterogeneous faults die to evolving stress and seismicity. J Geophys 2003;108:307–2310.
[9] Zion YB, Rice J, Dmowska R. Interaction of the San Andreas fault creeping segment with adjacent great rupture zones and earthquake recurrence at Parkfield. J Geophys 1993;98(B2):2135–44.
[10] Gabrielov A, Zaliapin I, Newman W, Borok VK. Colliding cascades model for earthquake prediction. J Geophys 2000;143(2):427–37.
[11] King CY. Episodic radon changes in subsurface soil gas along active faults and possible relation to earthquakes. J Geophys Res 1980;85:3065–78.
[12] Hartmann J, Levy K. Hydrogeological and gas geochemical earthquake precursors – a review for application. Nat Hazards 2005;34:279–304.
[13] Panakkat A, Adeli H. Recent efforts in earthquake prediction (1990–2007). Nat Hazards Rev 2008;9(2):70–80.
[14] Vallianatos F, Sammonds P. Evidence of non-extensive statistical physics of the lithospheric instability approaching the 2004 Sumatran-Andaman and 2011 Honshu mega-earthquakes. Tectonophysics 2013;590:52–8.
[15] Vallianatos F, Benson P, Meredith P, Sammonds P. Experimental evidence of a non-extensive statistical physics behaviour of fracture in triaxially deformed Etna basalt using acoustic emissions. Europhys Lett – EPL 2012;97(5). 58002-1-58002-6.
[16] Vallianatos F, Sammonds P. Is plate tectonics a case of nonextensive thermodynamics? Phys A Statist Mech Appl 2010;389(21):4989–93.
[17] Pliakis D, Papakostas T, Vallianatos F. A first principles approach to understand the physics of precursory accelerating seismicity. Ann Geophys 2012;55(1):165–70.
[18] Varotsos PA, Sarlis NV, Skordas ES. Natural time analysis the new view of time: precursory seismic electric signals, earthquakes and other complex time series, 76. Berlin, Germany: Springer-Verlag, Practica of Athens Academy; 2011. p. 294–321.
[19] Haykin S. Neural networks: a comprehensive foundation. 2nd ed. New Jersey: Prentice-Hall; 1999.
[20] Bodri B. A neural-network model for earthquake occurrence. J Geod 2001;32(3):289–310.
[21] Zare M, Amini H, Yazdi P, Sesetyan K, Demircioglu MB, Kalafat D, et al. Recent developments of the Middle East catalog. J Seismol 2014;18(4):749–72.
[22] Kavei M. Investigations on seismotectinics of parts of Hormzgan province in Southern Iran. Ph.D. Thesis. Pune, India: Pune University; 2008.
[23] Yu B, He X. Training radial basis function networks with differential evolution. In: IEEE Conference on Granular Computing; 2006. p. 934–41.
[24] Karaboga D. An idea based on honey bee swarm for numerical optimization. Techn Rep TR06. Erciyes: Erciyes Univ. Press; 2005. p. 1–10.
[25] Saraf M, Mosavi MR, Mohammadi K. Novel radial basis function neural networks based on probabilistic evolutionary and Gaussian mixture model for satellites optimum selection. J Aerosp Sci Technol 2012;9(2):71–81.
[26] Mosavi MR. A practical approach for accurate positioning with L1 GPS receivers using neural networks. J Intell Fuzzy Syst 2006;17(2):159–71.
[27] Nakhaei F, Mosavi MR, Sam A, Vaghei Y. Recovery and grade accurate prediction of pilot plant flotation column concentrate: neural network and statistical techniques. J Mineral Process 2012;110–111:140–54.
[28] Rafei M, Sorkhabi SE, Mosavi MR. Multi-objective optimization by means of multidimensional MLP neural networks. J Neural Netw World 2014;25(1):31–56.
[29] Department of Strategic Affairs of the Technical System. Assistance president of strategic planning and monitoring. Appl Guide Seism Risk Anal 2013;626.
[30] Wiemer S. A software package to analyze seismicity: ZMAP. Seismol Res Lett 2001;72(2):373–82.
[31] Reasenberg P. Second-order moment of central California seismicity, 1969–82. J Geophys Res 1985;90(B7):5479–95.
[32] Alexandridis A, Chondrodima E, Efthimiou E, Papadakis G, Vallianatos F, Triantis D. Large earthquake occurrence estimation based on radial basis function neural networks. IEEE Trans Geosciences Remote Sens 2014;52(9):5443–53.
[33] www.iiees.ac.ir.
[34] Matthews M, Reasenberg P. Statistical methods for investigating quiescence and other temporal seismicity patterns. Pure Appl Geophys 1988;126(2–4):357–72.
[35] Mariana E, Yehuda BZ. Techniques and parameters to analyze seismicity patterns associated with large earthquakes. J Geophys Res Solid Earth 1997;102(B8):17785–95.
[36] Bufe CG, Varnes DJ. Predictive modeling of the seismic cycle of the greater San Francisco Bay Region. J Geophys Res Solid Earth 1993;98(B6):9871–83.
[37] Eneva M, Ben-Zion Y. Application of pattern recognition techniques to earthquake catalogs generated by model of segmented fault systems in three-dimensional elastic solids. J Geophys Res Solid Earth 1997;102(B11):24513–28.
[38] van Stiphout Ô, Zhuang J, Marsan D. Seismicity declustering. In: Proc. Community online resource for statistical seismicity analysis; 2012. p. 1–25.
[39] Molchan G, Dmitrieva O. Aftershock identification: methods and new approaches. Geophys J Int 1992;109(3):501–16.

Mohammad-Reza Mosavi received his B.S., M.S., and Ph.D. degrees in Electronic Engineering from Iran University of Science and Technology (IUST), Tehran, Iran in 1997, 1998, and 2004, respectively. He is currently faculty member (full professor) of the Department of Electrical Engineering of IUST. He is the author of more than 250 scientific publications in journals and international conferences. His research interests include circuits and systems design.