THE EARTHQUAKE MAGNITUDE PREDICTION USED SEISMIC TIME SERIES AND MACHINE LEARNING METHODS

ENES CELIK, MUHAMMET ATALAY, ADIL KONDILOGLU
Enes Celik, Lecturer, Kirklareli University, Turkey,
Muhammet Atalay, Assist. Prof. Dr., Kirklareli University, Turkey, Adil Kondiloglu, Lecturer, Beykent University, Turkey

ABSTRACT
The occurrence of earthquakes, seismic wave propagation in the form of the earth's crust, to be measured, a lot of depends on variables such as the evaluation records obtained from measurement methods and metrics. Early to predict earthquakes are been very important to minimize the damage. Expert decision systems can be developed only using seismic time series analysis. In this study, seismic time series of earthquake are used received from Turkey Bogazici University Kandilli Observatory Earthquake Research Institute, Regional Earthquake Tsunami Monitoring Center. At this point, seismic time series comparative results are analyzed at prediction stage and test stage.
Keywords: Seismic time series, earthquake prediction, machine learning.

INTRODUCTION
Earthquake, at a depth of accumulated energy drain caused by elastic waves are vibrations caused by the movement of the ground surface. Earthquakes, mainly due to the Earth's crust and upper mantle are composed of fewer and larger faults occur with them. They are linked to many major earthquake active fault detected or observed. Occurs on both sides of the fault line of elastic deformation in deformation of rock mass due to accumulate energy. This energy by wearing the elastic rock mass fracture or breakage resistance comes once a shift occurs along existing fractures. Meanwhile discharged from the elastic deformation energy is spread as earthquake seismic waves. The fault line formation, which causes earthquakes and energy, the deformation energy that is collected over time both sides of the fault. The value of this energy in a big earthquake can be 1024 ergs (Ketin, 2005).
Earthquakes are the result of complex geological disasters developing events. Therefore, to understand the behaviour of earthquakes accurately, estimating, modelling is required a high level of knowledge and technological possibilities. The earthquake early prediction depends on the time, magnitude and location. These three units, compared to other parts of the mantle and the displacement of part of the earth's crust, is a variable speed loading of tension in the environment and the atmosphere rocks of crack resistance. Even if it knows the total ergs or elements controlling purging in any region, initial and boundary conditions of the properties of rocks and very well known in the environment of any jolt makes it difficult to know previously (Ercan, 2001).
Today on earthquake prediction, although some hard work and determination are clear of the earthquake until today to determine the exact location, time and magnitude of a technique cannot be developed. Before the formation of the earthquake, loss of life and property is important to minimize. It relies on a lot of data on earthquake prediction. These strain in the shell, geochemical changes in rocks, changes in underground water, drying of water sources, magnetic, electric, deformation changes, physical, chemical and biological changes occurring in the region, behavioural changes in living things, monitoring radon gas of source waters changes can be considered (Karaman, 2006).
Seismic activity is the earthquake occurred in the area, type, frequency and size. Earthquakes are measured by seismograph. Seismology is also known as the scientific study of earthquakes. The violence of the earthquake is determined by the magnitude scale. Magnitude is defined energy liberated during earthquake as a measure. There is no possibility to measure directly the energy, C. Richter, which is a measure of an instrumental method available earthquake "magnitude" is defined about 1930 (2016).
According to this scale 3 and under violent earthquakes usually is imperceptible, 7 and above violent earthquakes can be devastating. The depth of the point where the earthquake occurred is effective to force the destruction and earthquakes that occur at the point nearest the earth give more damage. Large violent earthquakes occur less frequently. For example, 10 times a day that most of the violence of the earthquake is more likely than 4 to

5 intensity. Furthermore, in England violent earthquakes between 3,7-4,6 every year, violent earthquakes between 4,7-5,5 10 years and violent earthquakes 5,6 and over a 100 years can be seen (2015).

LITERATURE REVIEW

Neuro-fuzzy classifier using data recorded with the seismograph application for short-term earthquake prediction accuracy rate with 82% predicting earthquakes have been estimated five minutes once. In addition, statistical entropy, Discrete Wavelet Transform, Fast Fourier Transform, Chaotic Properties, Power Spectral Density is used (L. Dehbozorgi, 2010).

Based on the radial basis function artificial neural network model is presented a new proposal for earthquake prediction. Artificial neural network training leave-one-out procedure was performed with a modified fuzzy tool to internal. In addition, the training algorithm is combined with the Reasenberg clustering technique used to remove data sets for processing data from aftershocks (A. Alexandridis, 2013).

The movement patterns of earthquake prediction data points back propagation artificial neural network method to analyse the seismic data movement model has been tried to find the formula (Wah, 2012).

Likely to be repeated in the future of earthquakes and periods are calculated using Poisson method. Firstly, the parameters are found in the Gutenberg-Richter relationship to the work area. These parameters, the probabilities of occurrence of different magnitude earthquakes using the Poisson method and the return period was estimated (Bayrak, 2015).

After the chaotic analysis of earthquakes using artificial neural Network is estimated short term prediction. Calculating the characteristics of time series is added to the artificial neural network architecture (Tzanaki, 2001).

Practical methods developed for non-fixed time series. Soft computer programming techniques is done with group methods, chaos theory, data processing, fuzzy model and neural network (Nomura, 2003).

The similarity matching in the same time zone the earthquake region of the model is presented. Based on the sequence similarity model is given earthquake similarity (L. Wei, 2010).

Owing to characteristics of time series used the chaotic character of the seismic time series. This seismic time series was estimated with Radial Basis Function neural network (Y. Chen, 2010).

Formal statistical techniques are foresees in time-series forecasting. Formal systems are predicted earthquakes structure and parameters (Mohsin, 2012).

METHODOLOGY

Dataset

Seismic time series is used to receive from Bogazici University, Kandilli Observatory and Earthquake Research Institute, Regional Earthquake Tsunami Monitoring Center. These series include the date, local time, latitude, longitude, depth and local magnitude. These series consists of the last 3000 measurements in Turkey. The data used in this study the local magnitude consist from 1,0 to 5,0 values. Therefore, there is not big earthquake data (Bölgesel Deprem-Tsunami İzleme ve Değerlendirme Merkezi, 2015).

Artificial Neural Network

Artificial neural networks (ANN) have emerged as the simulation of biological neural system. Operating system of the computer analogy to working of brain's has been developed neural network model. Artificial neural network learning algorithms depend on previously acquired experience. Artificial neural networks are composed of neurons. These neurons are able to connect even in very complex shapes into one, in a real nervous system. There is different weights inputs and one output of each neuron. It is expressed as in equation 3.1 of the total input of different weights (Haykin, 1994).

$$f(n) = f\left(\sum_{i=1}^{P} W_i X_i + b\right) \quad (1)$$

P is the number of input, W is weight of input, X is input and b is bias in neural network parameters. The sum weighted of inputs and the value of each neuron bias, it is passed through the activation function to obtain the result that it connected to the output neuron. According to the structure of the system the activation function (ϕ) may be sigmoid function, hyperbolic tangent function or threshold function. After the multi-layered system obtained the output can be output (y) of a neuron another input (x) of a neuron. A multilayer neural network model shown in Figure 1. (H.M. Ertunç, 2012).

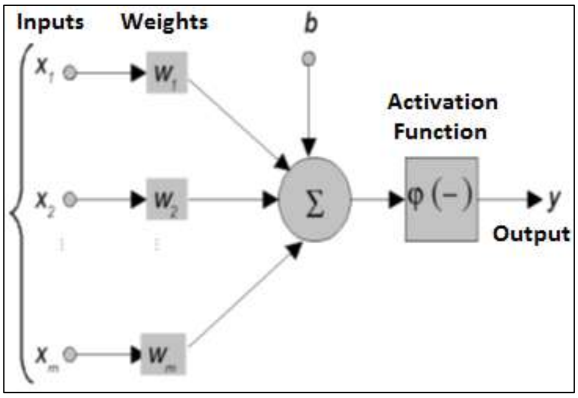

Figure 1. The general structure of artificial neural network (Kınacı, 2006)

Artificial neural network model is generally composed of three parts: the input layer, hidden layer and output layer. Each layer may consist of a lot of neurons. The information passes the activation function and the input layer. The output of the input layer continues as the input of the hidden layer. At this stage, once again evaluating the final result of activation functions for hidden layer output layer is obtained after the final output. For example, giving a 1 or 0 cells responsive to input patterns of a neural network capable of classification, it is deemed to have decided by assigning values of 1 or 0 pattern. The basic building blocks of the learning process are created decision and classify (Fausett, 1994).

Time Series Analysis

Time series which has a dynamic system of which contains a lot of information. Time series analysis means to determine the characteristics of the analysed data. It is called time series analysis for that applied to the methods. When the time series of a single variable, a one-dimensional time series analysis should be done. Some of these methods; decomposition methods, exponential smoothing methods, autoregressive (AR) methods, moving average (MA) methods and mixed autoregressive moving average (ARMA) methods (Orhunbilge, 1999).

Estimates are expressed only as a function of time in the first two of these methods. Whereas if another approach in to estimate is the regression analysis which based on relationships between variables. Here autoregressive (AR) model is the application of regression analysis of time series. As it is known that a dependent variable in the regression analysis is represented as a function with one or more independent variables. Because here is the only one available time series, because here is the only one available time series, the dependent variable y_t, its own past values function is written $y_{t-1}, y_{t-2}, \dots y_{t-p}$

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + e_t \quad (2)$$

The method of moving averages (MA) is based on the writing of as an error function of y_t the estimated value of the previous period. In generally, these models are shown as follows:

$$y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (3)$$

ARMA model is a combination of these two methods and developed by George Box and Gwilym Jenkins (Jenkins, 1970) which is known autoregressive moving average method or compound as Box-Jenkins methods:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (4)$$

In these models, p is degrees of autoregressive model as q is moving average models. These degrees is equal to the number of partial autocorrelation coefficient significantly different from zero. Naturally, for example if anyone from $\varphi_1, \varphi_2, \dots, \varphi_p$ in AR model is not different from zero, model carries a features of autoregressive.

Autocorrelation is relationship which is moving together in any period between with value of a term or after of term value. Because error in autocorrelation is relationship between the values of successive terms, usually it found positive autocorrelation in time series, so the series does not happen very sudden decline and rising. If the resulting time series is not stationary, autocorrelation significantly deviates from zero or delays or increases in counterfeit scratch out a time series (Goebel, 1979).

Autocorrelation coefficients shown that between same variable's degree of previous one, two, three, etc. value of dependence and this dependence is taken into account in AR model. When term value it's that term value (k) of between of shifted variables (y_{t-k}) in any term is increased with AR model variable (y_t), autocorrelation values will decrease exponentially. In MA model, degree of interdependence (q) limited between values of y_t, autocorrelation coefficients is significantly different of only first of q units, other approaches zero. The partial autocorrelation coefficients shown dependencies which is between variable (y_t) with shifted variable (y_{t-k})

any time in any term, when others held constantly. Partial autocorrelation coefficients higher before in MA model, when shifted term value increase, it decreases exponentially. Whereas the partial correlation coefficients is closer to zero immediately after a certain number of shift in the AR model (Orhunbilge, 1999).
In generally, time-series properties of the data are examined under two headings in the form of deterministic and stochastic characteristics. Deterministic properties of the series, the series is usually fixed, the presence or absence of trend and seasonality compounds. The stochastic properties are about more variables stationary is concerned with whether (Tarı, 2010).
A still of a time series is the correct approach to a specific value over time; in other words, a constant average is to have constant variance and covariance depends on the level of delay. More simple words; common variance between the two periods and unchanging over time, the average variance, not to the period of calculation of the common variance is still only be called a stochastic process, which depends on the distance between the two periods (Gujarati, 2011).
In order to obtain significant relationships for the variable and reproducible predictions made, the analysis must be made series of stationary series. A regression model was explored to use the relationship between static variables and any of the variables in the regression equation breaks down theory is not static in the sense described above (Emec, 2016).
If there is tendency in variable's time series, relationship may occur in the form too spurious regression than real. Therefore, whether they get a real relationship is related to whether stationary time series data (Tarı, 2010). Otherwise made t, F etc. the test will be invalid.
Investigation of stability in a time series is done the formal level by looking at whether the unit root of the time series. For this purpose, Dickey Fuller (DF) and Augmented Dickey Fuller (ADF) tests are used. Another method is based on the autocorrelation function and is tested with test autocorrelation showing autocorrelation of the various delays in the time series. Various time intervals (k) to calculate the autocorrelation and obtained by the partial autocorrelation coefficient k associated with graphics are referred to as autocorrelation (Tarı, 2010).
The graph shows quickly thinning of stationary time series, namely the autocorrelation coefficients of two or three semesters begin to reset scroll results, and it is slowly thinning in non-stationary series. In a series of completely random to one or more delays this autocorrelation is zero (Gujarati, 2011).
A non-stationary series is called a random walk series and even a long-term shift in the autocorrelation coefficients continue their significantly different from zero. For this series can be immobilized, is taken algorithm or need to obtain difference range in order of first, second etc. In taking the difference is taken of consecutive difference values (Orhunbilge, 1999). But in this case, it would not have tried to cool himself with difference values should not be forgotten.

TEST RESULT

The data in the study of Bogazici University Kandilli Observatory and Earthquake Research Institute covering local magnitude values from the time series are 3000 measurements of data. This time series autocorrelation and partial autocorrelation coefficients of the delay be investigated. The measurements obtained with E-Views application shown in Figure 2.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.167	0.167	55.996	0.000
		2	0.105	0.079	77.883	0.000
		3	0.085	0.058	92.337	0.000
		4	0.067	0.039	101.38	0.000
		5	0.081	0.055	114.46	0.000
		6	0.061	0.029	121.85	0.000
		7	0.030	0.000	123.64	0.000
		8	0.068	0.049	132.90	0.000
		9	0.083	0.056	146.72	0.000
		10	0.031	-0.006	148.67	0.000
		11	0.048	0.023	153.32	0.000
		12	0.045	0.020	157.43	0.000
		13	0.059	0.033	164.41	0.000
		14	0.027	-0.007	165.87	0.000
		15	0.037	0.017	168.70	0.000
		16	0.031	0.009	170.63	0.000
		17	0.031	0.007	172.62	0.000
		18	0.067	0.047	181.73	0.000
		19	0.043	0.015	185.47	0.000
		20	0.044	0.016	189.31	0.000
		21	0.024	-0.005	190.52	0.000
		22	0.041	0.020	193.98	0.000
		23	0.031	0.006	195.92	0.000
		24	0.054	0.030	201.75	0.000
		25	0.035	0.008	204.19	0.000
		26	0.017	-0.010	204.79	0.000
		27	0.022	0.000	205.81	0.000
		28	-0.015	-0.037	206.29	0.000
		29	0.024	0.016	207.44	0.000
		30	0.051	0.039	212.81	0.000
		31	0.038	0.014	215.73	0.000
		32	0.030	0.007	217.54	0.000
		33	0.017	-0.006	218.15	0.000
		34	0.023	0.008	219.19	0.000
		35	0.035	0.015	221.72	0.000
		36	0.047	0.028	226.16	0.000

Figure 2. Correlogram measurements

Figure 2. has been seen, according to the correlogram measurements time series is stationary. Because there is no confidence intervals exceeding 1% autocorrelation and partial autocorrelation coefficient. In addition, the autocorrelation coefficient decreases exponentially and the partial autocorrelation coefficient is significantly closer to zero after a delay. Therefore, the data structure of the AR (1) conforms to the model that has a time delay.

The analysis have been made in Weka application forecasting tools. This application, comparative analysis is made that latitude, longitude, depth and local magnitude features are estimated. The training results are expressed as regression. Errors are calculated as the mean squared error (MSE). In the test phase, the relationship between predicted results and actual results are expressed as correlation coefficient. In the prediction process, as core function is made Linear Regression, Additive Regression, Decision Tree, Rep Tree and Multilayer Perceptron (ANN). Total results shown in Table 1.

		Linear Regression	Additive Regression	Decision Tree	Rep Tree	Multilayer Perceptron
Latitude	**Regression**	72%	62%	59%	58%	65%
	Mean Square Error	0,84	2,31	1,56	2,12	1,56
	Correlation Coefficient	0,52	0,47	0,5	0,44	0,48
Longitude	**Regression**	71%	61%	63%	59%	66%
	Mean Square Error	0,96	1,98	1,76	2,25	1,78
	Correlation Coefficient	0,51	0,49	0,47	0,44	0,43
Depth	**Regression**	51%	52%	54%	49%	69%
	Mean Square Error	3,49	4,53	4,24	4,68	1,88
	Correlation Coefficient	0,22	0,23	0,33	0,24	0,21
Local Magnitude	**Regression**	71%	68%	73%	70%	83%
	Mean Square Error	0,42	0,45	0,4	0,43	0,34
	Correlation Coefficient	0,63	0,59	0,67	0,62	0,71

Table 1. Prediction Analysis

According to Table 1, for local magnitude of earthquake Multilayer Perceptron is been the highest regression rate, the lowest mean square error rate and the highest correlation coefficient rate. In addition to, for depth of earthquake Multilayer Perceptron is been the highest regression rate, the lowest mean square error rate and the highest correlation coefficient rate.

According to Table 1, for latitude of earthquake Linear Regression is been the highest regression rate, the lowest mean square error rate and the highest correlation coefficient rate. In addition to, for longitude of earthquake Linear Regression is been the highest regression rate, the lowest mean square error rate and the highest correlation coefficient rate.

CONCLUSIONS

The occurrence of earthquakes, seismic wave propagation in the form of the earth's crust, to be measured, a lot of depends on variables such as the evaluation records obtained from measurement methods and metrics. Early to predict earthquakes are been very important to minimize the damage. Expert decision systems can be developed only using seismic time series analysis. In this study, the latest 3000 measurement earthquake in Turkey using seismic time series data, are developed decision system that estimate local magnitude of earthquake, coordinates and depth center in underground. Weka application is used for the prediction process. In the local magnitude was successful artificial neural networks with the 83% regression rate and the 0.34 error (MSE) rate. In the depth was successful artificial neural networks with the 69% regression rate and the 1.88 error (MSE) rate. In the latitude and longitude was successful linear regression with respectively 72%, 71% regression rate and respectively 0.84, 0.96 error (MSE) rate. Emergence of high error rate, indicates the need to use long-period seismic time series. The value of delay can be changed for to increase the regression value and to reduce the error rate.

REFERENCES

(2015), http://www.quakes.bgs.ac.uk/hazard/Hazard_UK.htm/

(2016), http://www.crustal.ucsb.edu/outreach/faq.php/

A. Alexandridis, E. C. (2013). Large Earthquake Occurrence Estimation Based on Radial Basis Function Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing.*

Bayrak, T. T. (2015). DoğuAnadoluBölgesiveCivarının Poisson YöntemiileDepremTehlikeTahmini. *III. Türkiye Deprem Mühendisliği ve Sismoloji Konferansı.* İzmir.

Bölgesel Deprem-Tsunami İzleme ve Değerlendirme Merkezi. (2015). http://www.koeri.boun.edu.tr/scripts/lst1.asp

Emec, H. (2016). *Zaman Serileri Ekonometrisi*. http://www.deu.edu.tr/userweb/hamdi.emec/zamanserileri.pdf
Ercan, A. (2001). *Deprem Olacağı Nasıl Bilinir.* Ankara: TMMOB Jeofizik Mühendisleri Odası Kozan Ofset.
Fausett, L. (1994). *Fundamentals of Neural Networks.*
Goebel, A. G. (1979). *Nonlinear Regression with Autocorrelated Errors.* Jasa volume 71.
Gujarati, D. (2011). *Temel Ekonometri.* İstanbul: Literatür Yayıncılık.
H.M. Ertunç, H. O. (2012). ANN and ANFIS based multi-staged decision algorithm for the detection and diagnosis of bearing faults. *Neural Computer and Application*.
Haykin, S. (1994). *Neural networks a comprehensive foundation.*
Jenkins, G. B. (1970). *Time Series Analysis, Forecasting and Control.* California: Holden Day.
Karaman, E. (2006). *Yapısal Jeoloji ve Uygulamaları.* Ankara.
Ketin, İ. (2005). *Genel Jeoloji Yer Bilimlerine Giriş.* Ankara: İTÜ Vakfı.
Kınacı, A. U. (2006). Yapay Zeka Teknikleri ve Yapay Sinir Ağları Kullanılarak Web Sayfalarının Sınıflandırılması. *11. İnternet Konferansı.*
L. Dehbozorgi, F. F. (2010). Effective Feature Selection for Short-Term Earthquake Prediction Using Neuro-Fuzzy Classifier. *IEEE II. International Conference on Geoscience and Remote Sensin.*
L. Wei, Z. H. (2010). Based on time series similarity matching algorithm for earthquake prediction research. *3rd International Conference on Advanced Computer Theory and Engineering.*
Mohsin, F. A. (2012). Agent Based Prediction of Seismic Time Series Data. *10th International Conference on Frontiers of Information Technology.*
Nomura, H. F. (2003). Time series prediction of earthquake input by using soft computing. *Fourth International Symposium on Uncertainty Modeling and Analysis.*
Orhunbilge, N. (1999). *Zaman Serileri Analizi Tahmin ve Fiyat İndeksleri.* İstanbul: İstanbul Üniversitesi İşletme Fakültesi Yayınları.
Tarı, R. (2010). *Ekonometri.* Kocaeli: Umuttepe Yayınları .
Tzanaki, V. P. (2001). Chaotic analysis of seismic time series and short term forecasting using neural networks. *Proceedings, International Joint Conference on Neural Networks*.
Wah, J. V. (2012). Earthquake Prediction Based on the Pattern of Points Seismic Motion. *International Conference on Advanced Computer Science Applications and Technologies.*
Y. Chen, J. Z. (2010). Research on application of earthquake prediction based on chaos theory. *International Conference on Intelligent Computing and Integrated Systems (ICISS).*