Proceedings on Big Data Analytics & Innovation (Peer-Reviewed), Volume 1, 2016, pp.165-175

Predicting the Time Lag between Primary and Secondary Waves for Earthquakes using Artificial Neural Network

OGBOLE C. I.
Federal University of Technology Minna, Nigeria

MUHAMMED S.
Eastern Mediterranean University, Famagusta, North Cyprus

MUHAMMAD E.B., FOLORUNSO T. A. & NUHU B.K.
Federal University of Technology Minna, Nigeria

ABSTRACT Seismic waves experienced prior to earthquake are the primary and the secondary waves. This paper investigates the time lag after the primary wave before the occurrence of the secondary (destructive) wave. The aim is to allow for necessary warning signals and safety steps to be taken prior to the impending disaster. Seismometer records from previous earthquakes were used in this investigation, putting into consideration the time lag between the primary and secondary waves. Other parameters considered include the magnitude, the epicenter distance, the seismic station's distance and the direction (in azimuths). Consequently, a prediction model was developed from the derived data using Artificial Neural Network (ANN). Data obtained from earthquakes of magnitude 6.0 to 7.0, based on Richter's scale, was used to train the ANN. The results therein showed high performance, with regression values greater than 0.9 and root mean squared errors of 0.1003-0.1148 for the most satisfactory architecture. The final results showed that the developed ANN model achieved a high performance, hence, adequate for this type of application.

Keywords: Earthquake, seismic waves, fault lines, hypocenter, epicenter, neural networks

1. Introduction

Earthquakes are the result of plate tectonics, and it occurs when energy is released in the earth crust resulting in seismic waves. Earthquake occurrence varies spatially and its prediction has been a goal of mankind for millennia (Wiemer 2000). The basics in earthquake prediction begins with measuring the changes in distance (geodetic), also a creep-meter can be used; this is a device to measure movement across a fault line. In (Liu et al. 2013), a measure of the change in slope on earth's surface using a tilt-meter is considered, this inclinometer measures small changes on the ground and on physical structures. Other changes in the properties of physical structures can also be measured; solid rocks are highly resistive but under excessive strain, they develop cracks and shatter, thus allowing water to percolate through, resulting to increase in its conductivity (Furen 2010). Seismologist uses various tools for earthquake prediction analysis, the most common of this is the seismograph machine (or seismometer) which detects and records seismic waves. For a region's seismicity factors considered includes: the air ionization around rock surface which increases prior to earthquakes (Freund et al. 2009), the geology of the area, location of faults, the earthquake history of the area, the previous earthquake intensities and evidence for recent fault movement.

Further useful works give reviews on various animal behavioral anomaly (Bhargava et al. 2009), the possibility of prediction from physical climatic elements using neural networks (Maitha et al. 2011), evidence on the relationship between seismic electric signal (SES) with earthquake focal mechanisms (Varotos & Alexopoulos 1982), and the use of wireless sensor networks in earthquake monitoring (Azzam et al. 2011).

In addition, (Kirschvink 2000; Adi & Schnytzer 2011; Grant & Halliday 2010) considered behavioral activities (seismic-escape response) put up by some animals in response to the precursor to be helpful in earthquake prediction. These animals through natural selection are forced to develop anticipatory mechanism for predicting possible natural disasters. Although, the issue with the belief that certain animal do anticipate earthquakes is that it is poorly supported by evidence (Grant & Halliday 2010).

In another related approach, called the VAN method, coined from the initials of the three Greek scientist (Varotsos, Alexopoulos and Nomicos). They found out that seismic electric signals which results to variations in the earth's electric field occurs prior to an earthquake. Depending on this SES's types, the earthquake can be predicted to occur within days to weeks (Varotsos et al. 2006). This has attracted a high level of debate, which has been majorly on how to distinguish between similar electric signals from natural occurrence like thunderstorms and other man-made disturbances (Moustra et al. 2011). In (Kai Tan & Xiushan Cai 2010), data from earthquakes of magnitude 3.5 and greater collected from 1970 to 2008 in Yunnan region (22-28^{o}N, 98-104^{o}E) were used to predict earthquakes in 1999-2008 and verified using the Support Vector Machine (SVM).

The mitigation to earthquakes is in its prediction which could be long term, medium or short term. In short-term prediction, specific information of the earthquakes time and location is given within minutes, weeks, or months and are therefore very useful for public safety and evacuation (Uyeda et al. 2009). This method of prediction has attracted extensive research lately. Most earthquake studies in the past were basically on understanding the basics; its occurrence and extent of damage. Prediction study on this area only started in the 1980's (Zuniga & Wyss 1995).

Seismograph is used to detect and record seismic waves and the seismic measurements are the basis for short-term prediction (Uyeda et al. 2009). There are two basic types of seismic waves; the primary wave (P-wave) and secondary wave (S-wave). Though a third wave exists that is called the surface wave (this is the resulting wave formed when the P & S-waves combines at the surface). This research work presents a novel approach which focuses on the prediction of the arrival time of the S-wave after a P-wave has been detected. It further shows how the time lag between these two wave forms can be computed using Artificial Neural Network (ANN).

This work uses a supervised learning strategy in ANN. The supervised learning can also be termed 'learning with a teacher'. Illustration for this kind of learning uses a teacher. The teacher is believed to have full knowledge of the system, this knowledge is given in a set of input-output mapping, but the neural network does not know this. Thus, the knowledge of the system is transferred from the teacher to the neural network to a certain degree measured with statistical tools. While in contrast, unsupervised learning looks at how systems learn to represent input pattern in ways to reflect the structure of the entire collection of the inputs. For this method of learning, there are no explicit target outputs associated with the inputs.
The remainder of this paper is organized as follows: Dataset description in section II, section III shows the ANN Model Development. The result and its discussion is presented in section IV. Finally, in section V the conclusion and limitations are presented.

2. Dataset Description

A number of factors are known to influence the occurrence of earthquake and these factors have varying effect on the strength as well as impact of the quake. The sampled seismic data set used for this study was obtained from the World Data Center for Seismology in China, measured from January, 2012 to August, 2014. The dataset contains varying values of five key parameters namely; the distance, the azimuth, the measured magnitude, the depth and the time lag between the primary and the secondary waves. The range of values for the parameters are as depicted in Table 1.

Table 1: Dataset description and their value ranges

Parameters	Value range
Distance (degrees)	8.8-164.6
Azimuth (degrees)	0-358
Measured Magnitude (Richter)	6.0-7.0
Depth (meters)	5,000-80,000

Characteristically, the distance (D) measured in degrees, is the representation of the distance from the earthquake's source and the seismological station (point of observation); the Azimuth (Az) is a clockwise measurement referenced from the earth's true north, also in degrees; the Magnitude (M) is the measure for the earthquake's primary wave as recorded by the seismograph at the station; the Depth (D) is the distance from the earthquake's hypocenter (wave origin) to the epicenter, and finally, the time lag (Ts-Tp) is the time difference between the arrival of the first primary wave and the first secondary wave signals.

A total of 86 earthquake cases were sampled from 1,478 stations, and they were all of the magnitude range of 6.0 to 7.0 (Strong earthquakes) on the magnitude scale. The choice of this magnitude range is because of the frequency of occurrence amongst the earthquake classes that poses serious threats to lives and properties. Table 2 shows the classes of earthquakes base on the magnitude and their effects.

Table 2: Earthquake magnitude, effect and annual frequency (source UPSeis)

Magnitude	Earthquake Effect	Annual Frequency
8.0 or more	Can totally destroy communities near the epicenter	One in 5-10 years
7.0 – 7.9	Causes serious damage	20
6.1 – 6.9	May cause a lot of damage in very populated areas	100
5.5 – 6.0	Slight damage to buildings and other structures	500
2.5 – 5.4	Often felt, but only causes minor damage	30,000
2.5 or less	Usually not felt, but can be recorded by a seismograph	900,000

3. Artificial Neural Network (ANN) Model Development

In this work, the design of an ANN model that will give a significantly high level of generalization (prediction) for the time lag between the primary and secondary earthquake waves is presented.

The structure of the neutral network is as shown in Figure 1, this depicts the perceptron process, with ,.. representing the input parameters.

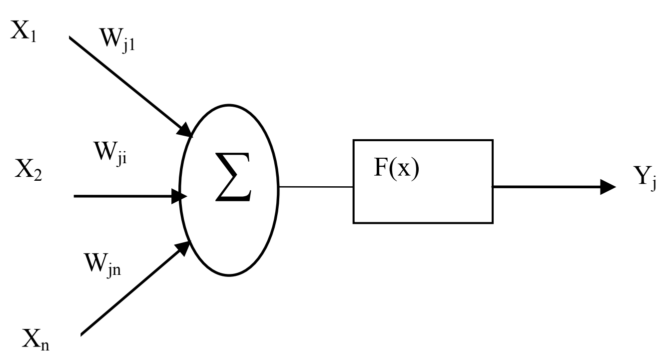


Figure 1: A detailed perceptron process

The input neurons buffers the inputs $\times_i$ ($\times_1$, $\times_2$,..$\times_n$) to the neurons in the hidden layer. Summation of inputs is done in each neuron j of the hidden layer, where these inputs are weighted with the inter-neuron connection weights w_{ji} and the output y_{ji} is computed as a threshold function of the sum using equation (1).

$$y_{ji} = f\left(\sum_{i=1}^{n} w_{ji}\, x_i\right) \tag{1}$$

In the Multilayer Perceptron (MLP) structure, the threshold function is a continuous derivative. The goal is to minimize the error function, which is achieved by finding the squared error of the network. Equation 2 gives how the training weights are adapted:

$$\Delta w_{ji} = \eta \delta_j x_i \tag{2}$$

Where η is the η learning rate; it determines the level of modification to the link weights (w) and node biases base on the change rate and direction.
A “momentum” term (μ) is added to help the network skip over the local minima and successfully reach the global minimum, while still maintaining the change rate and direction. This is adopted into the weight update equation as shown in Equation (3), while the change rate for both the output and hidden neurons are computed as given in Equations (4) and (5) respectively:

$$\Delta w_{ji}(t+1) = \eta \delta_j x_i + \mu \Delta w_{ji}(t) \tag{3}$$

For the output neurons,

$$\delta_o = \left(\frac{\partial f}{\partial netj}\right)\left(y_j^{(t)} - y_j\right) \tag{4}$$

For the hidden neurons,

$$\delta_{h} = (\partial f/\partial netj)(\textstyle\sum_{q} w_{ji}\ \partial_{j}) \tag{5}$$

And training continues until the error function reaches a target minimum value.

The parameters considered for this prediction work are; the distance (D), the azimuths (Az), the measured magnitude (M), the depth (Ep), the measured time lag (Ts-Tp).

The design for a neural network on MATLAB adopts certain systematic procedures. In general, these five basic steps are followed; importing the data, preprocessing data, building the network, training the network, testing the network and evaluating the system's performance. The data is first grouped in two sets; the training set and the testing set. In the preprocessing stage, normalization of the data set is applied. This is necessary considering the range of values of the parameters which largely varies. The design program for this work follows the flow chart presented in Figure 2.

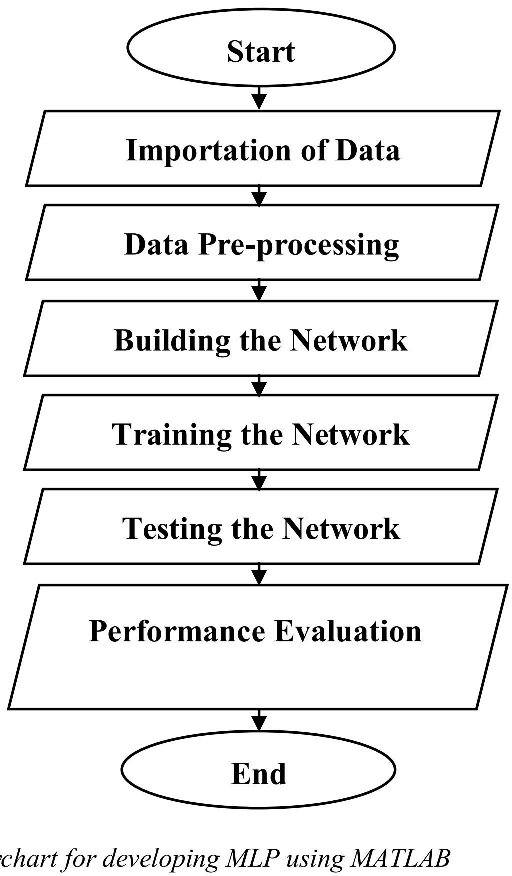


Figure 2. Flowchart for developing MLP using MATLAB

In the training, the network is taught how to generalize for the presented data set. These data sets consist of input-output pairs. The neural network learns from the input and updates its weight; this is why it is termed a supervised learning, since the neural network is taught what the output should be from the input set introduced to it. Figure 3 shows the neural network MLP architecture, presenting the network flow from input to the output for the design.

The architecture used has four (4) inputs neuron, one (1) hidden layer with hidden number of neurons varied from 3-7, 10 & 20. Each of this architecture was trained and tested with a learning rate (η) of 0.1 to 0.9. The network was observed while varying the number of neurons in the hidden layer, the momentum constant and also the learning rate for the 9 different structures, and the best performing structures are selected. The training is stopped whenever any of the network's performance parameter is met.

After training is completed, the network is tested with unseen data and the output compared with the target (measured result). This is to check how well the network can generalize (predict output from the unseen inputs). Checking the performance is carried out using statistical measures on the obtained results: the root mean square error (RMSE), the mean absolute error (MAE) and the mean bias error (MBE) are computed with the formulas as given in Equations (6), (7), and (8) respectively:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n}(t-O)^2} \qquad (6)$$

Where, *n* is the number of samples. *t* is the target output (measured value), and *O* is the network output (predicted value).

$$\mathbf{MAE} = \frac{1}{n}\sum_{i=1}^{n}|t-O| \qquad (7)$$

If MAE=RMSE, it means all the errors in the sample are of the same magnitude.

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(t-O) \qquad (8)$$

For every simulation, the computed values are recorded in excel and used to evaluate the system's performance. The performance of the trained network on testing data is the focus; it is most important measure of the training success.

4. Results and Discussion

Several network variations were investigated. The computed data for the network architecture with satisfactory performance are given in Tables 3 and 4 for this section. Table 3 gives the best result for the different architectures while varying both momentum constant and the learning rate for the eight different architectures.

Architecture	*Test error statistics at μ=0.01*			*Test error statistics at μ=0.001*		
	RMSE	*MAE*	*MBE*	*RMSE*	*MAE*	*MBE*
4-2-1	*0.1241*	*0.0955*	*0.1241*	*0.2139*	*0.154*	*-0.0247*
4-3-1	*0.1126*	*0.0925*	*-0.0243*	*0.2009*	*0.1474*	*-0.0024*
4-4-1	*0.1072*	*0.0937*	*-0.0079*	*0.2132*	*0.1526*	*-0.0236*
4-5-1	*0.1003*	*0.0867*	*-0.0661*	*0.2173*	*0.1555*	*-0.0222*
4-6-1	*0.1065*	*0.0899*	*-0.0054*	*0.2193*	*0.1468*	*-0.0054*
4-7-1	*0.1049*	*0.0915*	*-0.0553*	*0.2175*	*0.1387*	*-0.0318*
4-10-1	*0.1188*	*0.0998*	*-0.0439*	*0.2142*	*0.1497*	*-0.0252*
4-20-1	*0.1132*	*0.0971*	*-0.0414*	*0.224*	*0.1576*	*-0.0047*

Table 3 Statistical performance evaluation for the different ANN architectures

In Table 3, the notation 4-2-1 implies 4 input neurons, 2 hidden neurons and 1 output. The RMSE and MAE values are found to be better using a momentum constant of 0.01 (μ=0.01). Also from the table, we can deduce that the best overall RMSE value was at five (5) hidden neurons with a momentum constant of 0.01, while the test using the three (3) hidden neurons gave the best RMSE with the momentum constant of 0.001.

Another observation from Table 3 is that the RMSE and MAE values were considerably consistent from N=2 to N=7. This indicates that optimal performance of the network was within this range and then much increase from N=10. The performance plot of the training also indicates a good training. The performance plots selected for N=5 for μ=0.01 is seen in Figure 3.

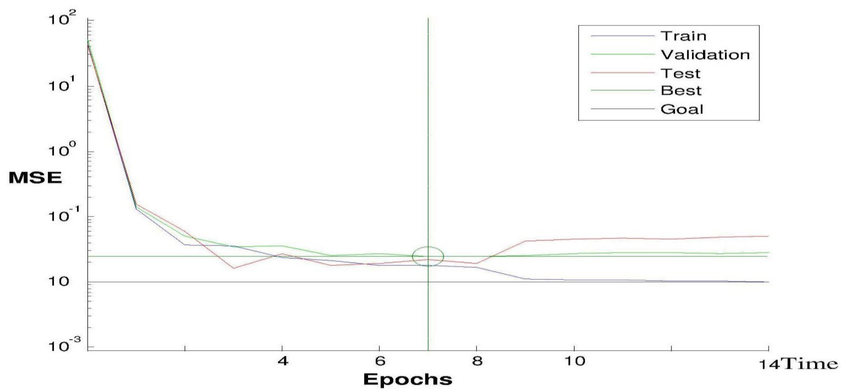
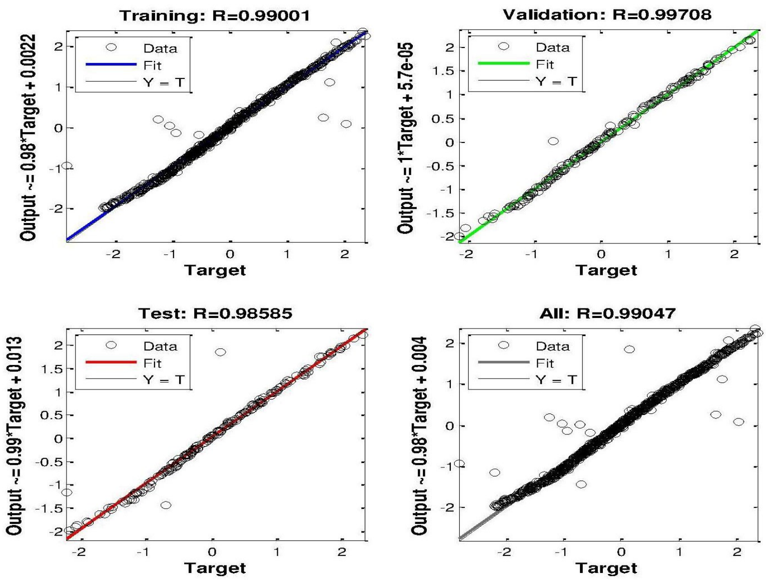


Figure 3. The performance plot for N=5 for μ=0.01



The plot as seen in Figure 3 shows that the best validation was reached at epoch seven (7), even though the training still proceeded for 7 more epochs and with a mean square error of 0.024839. Also the test plot is noticed to be similar with that of the validation. Figure 4 gives the linear regression between the output and the target for the training, validation, test and all three results combined.

Figure 4. Schematic of the regression plot at N=5 for μ=0.01

The R-values shown in Figure 4 are all greater than 0.9. This is an indication of a very good fit for the training data and it shows how very close the output of the network and the target (measured) values are. The training stops whenever any of the performance goal is met. Also in Figure 5, it is observed that the network gave a better result (i.e. a lower root mean square value), for the testing when a momentum constant (μ) of 0.01 than when using 0.001. For the different training using varying number of hidden neurons, it is observed that this trend is maintained.

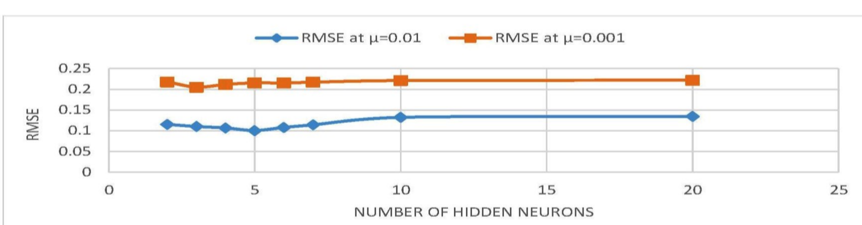


Figure 5. Plot of obtained RMSE values for the different number of hidden neurons.

5. Conclusion

From the result, this research work does not only present the possibility of earthquake prediction using the primary and secondary earthquake waves (P & S-waves), its performance evaluation also shows a high confidence for its adoption as a prediction model. The neural network was trained to a level that it was able to achieve a very good generalization from the input parameters (distance, azimuth, depth of the source and the magnitude of the received primary wave) & the output (time-lag between the P & S-waves). A limitation to this proposed model is the high amount of data sampling required for effective training. This will require a lot of computational resources. As such, for an effective and efficient implementation of this model on an industrial scale, it is recommended that large computing and processing machines be used.

Correspondence
Ogbole C. I.
Department of Computer Engineering
Federal University of Technology Minna, Nigeria
Email: ogbole.inalegwu@futminna.edu.ng

References

Adi, S. & Schnytzer, Y., 2011. *Animal modelling of earthquake and prediction markets*, Israel.

Azzam, R. et al., 2011. Monitoring of landslides and infrastructures with wireless sensor networks in an earthquake environment. *Memorias de la conferencia 5th International Conference on Earthquake Geotechnical Engineering*, pp.10–13.

Bhargava, N. et al., 2009. Earthquake prediction through animal behavior: A review. *Indian J. Biomech*, 78.

Freund, F.T. et al., 2009. Air ionization at rock surfaces and pre-earthquake signals. *Journal of Atmospheric and Solar-Terrestrial Physics*, 71(17–18), pp.1824–1834.

Furen, X., 2010. *Rock stress and earthquakes*, London.

Grant, R.A. & Halliday, T., 2010. Predicting the unpredictable; evidence of pre-seismic anticipatory behaviour in the common toad. *Journal of Zoology*, 281(4), pp.263–271.

Kai Tan & Xiushan Cai, 2010. Prediction of earthquake in Yunnan region based on the AHC over sampling. In *2010 Chinese Con-*

trol and Decision Conference. IEEE, pp. 2449–2452.

Kirschvink, J.L., 2000. Earthquake prediction by animals: evolution and sensory perception. *Bulletin of seismological society of America*, 90, pp.312–323.

Liu, G. et al., 2013. Volcanic earthquake timing using wireless sensor networks. In *Proceedings of the 12th international conference on Information processing in sensor networks - IPSN '13*. New York, New York, USA: ACM Press, p. 91.

Maitha, H., Ali, H. & Hassan, A., 2011. *Using MATLAB to Develop Artificial Neural Network Models for Predicting Global Solar Radiation in Al Ain city-UAE*,

Moustra, M., Avraamides, M. & Christodoulou, C., 2011. Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals. *Expert Systems with Applications*, 38(12), pp.15032–15039.

Uyeda, S., Nagao, T. & Kamogawa, M., 2009. Short-term earthquake prediction: Current status of seismo-electromagnetics. *Tectonophysics*, 470(3–4), pp.205–213.

Varotos, P. & Alexopoulos, K., 1982. Physical properties of the variation of electric field of the earth preceeding earthquakes. *Technophyics*, 110(1), pp.73–98.

Varotsos, P. et al., 2006. Additional evidence on some relationship between Seismic Electric Signals (SES) and earthquake focal mechanism. *Tectonophysics*, 412(3–4), pp.279–288.

Wiemer, S., 2000. *Earthquake statistics and earthquake prediction research*,

Zuniga, F.R. & Wyss, M., 1995. Inadvertent changes in magnitude reported in earthquake catalogues: Influence on b-value estimate. *Bulletin of seismological society of America*, 85(6), pp.1858–1866.