Study on Techniques of Earthquake Prediction

G.Preethi
Assistant Professor
Department of CA/SOC
SASTRA University,
Thanjavur, India

B.Santhi
Professor
School of Computing
SASTRA University,
Thanjavur, India

ABSTRACT
An event called prediction in a time series is more important for geophysics and economy problems. The time series data mining is a combination field of time series and data mining techniques. The historical data are collected which has follow the time series methodology, combine the data mining for preprocessing and finally apply the fuzzy logic rules to predict the impact of earthquake. Earthquake prediction has done by historical earthquake time series to investigating the method at first step ago. Huge data sets are preprocessed using data mining techniques. Based on this process data prediction is possible. This paper is focused on statistics and soft computing techniques to analyze the earthquake data.

Keywords – Event, time series analysis, data mining, time series data mining and soft computing.

1. INTRODUCTION
The time series is acquired by the determined time interval from any system. For example, Price change of the daily trade market, Gold price rise and fall changes, stock market variation and growth of population according to the year in a country. A time series includes a huge number of observations are well-defined data items ordered in equal time or equal space. The collected data are irregular or happens only once are not a time series data. An observed time series are categorized into three types as follows:

1. The Seasonal component
2. The Trend component and
3. The Irregular component

The seasonal data are systematic or regular movements of data. The Trend data mean by Long-term fluctuations and the irregular data mean by unsystematic or short-term fluctuations. The major utility area of the time series model is statistical forecasting. The available prediction approaches are regression, time series and chaotic approaches. Each and every method has its own advantage and disadvantage. The historical sequences of data are to be used for forecasting purposes, because of this reason the time series models are used to predict the future values. The prediction will show what will happen but won't why it happens. Time series values are transformed to phase space by using a nonlinear method and then apply the fuzzy logic to predict optimum value. The time series data are derived from the time interval of any system. Traditional stationary time series models are Autoregressive Integrated Moving Average (ARIMA) and Minimum Mean Square Error forecasting methods. Data mining is used to extract useful and more relevant information from the huge database. The author Han describes an artificial intelligence and pattern recognition methods for prediction [7]. The time series data mining is used for prediction of earthquake [5]. Fuzzy logic methods are used to predict the earthquake, stock market changes, weather forecasting and gold price changes. The similarities of patterns are selected as a fuzzy set; these sets are specified in membership function. The fuzzy logic is accurately predicted the prediction event. The concepts of time series data mining had been used for clustering and natural event prediction. Povinelli, [8] suggested an application for the event prediction. In their study, they focus the different applications such as earthquake prediction, fall of stock price and gold price prediction.

The major advantage of time series is possible to predict the future value based on the previous historical data. The study of the past sequence of historical data may be more valuable. The time series is helpful to predict the next sequence of future values. The utility of time series method is specifically for Trend analysis, Trade market, Finance, Climatologic and earthquake prediction.

This paper is ordered as follows; Section 2 describes the related work of the prediction. Section 3 is describing the prediction method of time series. Section 4 discusses the statistical analysis of earthquake. Section 5 illustrates the experimental result of reviewed data. Finally, conclusion and future work are given in Section 6.

2. RELATED WORK
Morales-Esteban [10] discusses the behavior of seismic temporal data and that was used to predict the medium and large earthquakes. The known patterns are determined when the medium and large earthquake happen.

Chris gray [9] discuss the five geophysical precursors of earthquake. The precursors are used to predict the earthquake. Here five stages are used to predict the earthquake and give advanced warning system of the earthquake.

The major earthquake killed the several million people during the period from 2000 to 2010 in the area of Indonesia. Most of the scientists are currently looking to find a way to predict the earthquake so that the life and properties can be saved. The statistical analysis will be described by reviewing the case study of Indonesia region. Earthquakes are categorized by the depth and slip of the earthquake. The comparison of magnitude and years are possible using statistical data. The major objective of the paper is to review the methods of predicting earthquake and predicting the future using the time series model. The statistical analysis is used for long-term predictions of the earthquake. The reviews of the geophysical precursors are dependence on the particular region. The precursors are not suitable for the other region. The impacts of the earthquake are identified through the ground water- level, ground acceleration, time interval and P- wave changes.

3. PREDICTION METHODS

3.1 Linear Model:

A mathematical model is an equation that represents the behavior of a system. All the parameters are appearing equally in linear model. The following examples of linear models are:

1. **Multiple Line Regression:**

$$Y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n + \varepsilon$$

Here Y is the response variable, x_i are the predictor variable and ε is the experimental error term.

2. **Polynomial models as follows:**

$$Y = a + bX + \varepsilon$$

$$Y = a + bX + cX^2 + \varepsilon$$

First equation is representing First-Order model and the second is Second-order or Quadratic Model. These models are widely used in earthquake prediction, Agriculture and Medicine, etc.

3.2 Non Linear Model:

At least one of the parameter appears nonlinear. The following nonlinear model examples are:

$$Y(t) = \exp(at+bt^2)$$

$$Y(t) = at + \exp(-bt)$$

At least one derivative parameter should involve in nonlinear model. The nonlinear models are used in time series analysis for predicting future values.

4. STATISTICAL ANALYSIS OF EARTHQUAKES

4.1 Review of Magnitude and number of earthquake occurrence using time series:

Using Indonesia earthquake data, the relationship between number of occurrence and magnitude are analyzed and plotted in the fig.1. An earthquake is defined as an event in the seismic time series. An event is including raise and fall of a gold price. One of the predicting methods is statistical analysis. In the analysis of case study of the particular region, a persistent or cyclic pattern of the earthquakes are identified. The sample dataset of the Indonesia region includes depth and magnitude details. The prediction is possible by identifying the probability of occurrence, time interval, magnitude, longitude, and latitude range of the earthquakes.

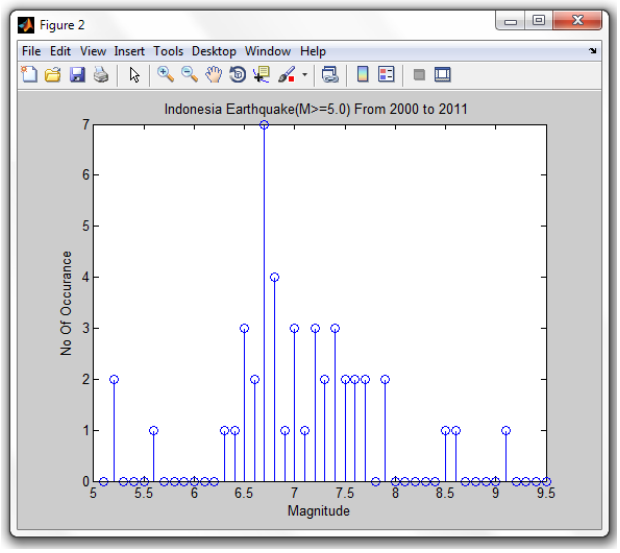

Fig. 1. Magnitude and No. of occurrence of earthquake [USGS database]

The collection of Indonesia earthquake data based on magnitude is greater than 5.0. The ranges of 6.5 to 8.0 magnitudes were heavily damaged the Indonesia region during the year of 2000 to 2011[1]. Using a statistical data the experts can able to predict the magnitude and depth of the earthquakes. Earthquake magnitude distribution has been observed from the beginning of 21st century [2].

4.2 The review of magnitude and depth of earthquakes:

The reviews of the depths are used to identify the fault of the earthquake. There are three types of faults are identified through the magnitude and depth of the earthquake. The faults are namely normal (or detachment) fault, reverse (or thrust) fault and strike-slip fault. Also the plate motions are identified by historical data. The plate motions and the faults are related to the magnitude and depths of earthquake. This information is retrieved from statistical earthquake data. More than 5.0 magnitude earthquakes are affected the Indonesia area. The

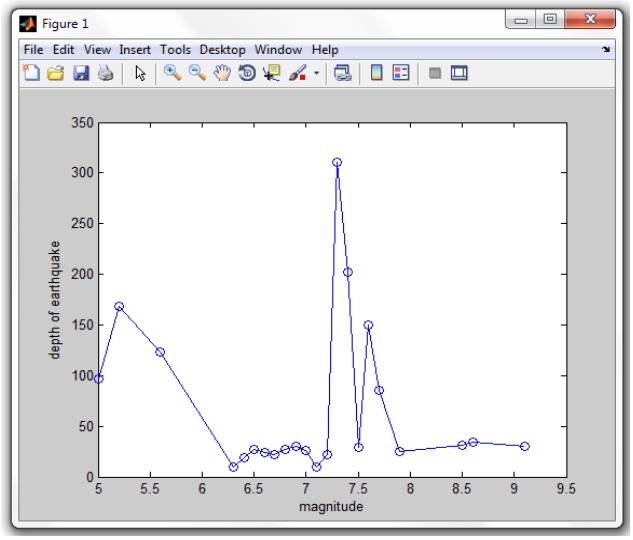

Fig 2. Magnitude & depth of earthquake of the same magnitude

Depth and magnitude of earthquakes are identified and shown in fig.2. Drastic effects are identified in between the ranges of 7.0 to 8.0. Review of the magnitude and depth of earthquakes has been found which are used to identify the distance from the origin. This information is more capable to classify the types of earthquakes. The shallow types of earthquakes are occurred at depth down to 300kms [9]. Other types of earthquakes are deep earthquake that may be occurred at the depth starts from 300kms to 680kms [9]. Majority of the shallow of earthquakes had been occurred rather than deep earthquakes. Hence the shallow of earthquakes are producing more damages in an environment. This paper refers only for the shallow earthquake. The case study identifies the Indonesia is mostly affected by the shallow earthquakes.

4.3 Feature Extraction

Department of Geography or Geographic Resources Center (GRC) at university of Missouri- Columbia [3] develops feature extraction methods. The investigation of earthquake prediction can be possible by precursor information and related regional factors [5]. The feature extraction involves a preselecting process of statistical properties of data and a set of seismic parameter, which are linearly independent in feature space. The seismic parameters are in the form of time series could be analyzed by various pattern recognition techniques. The statistical or soft computing methodology is usually performs this predictions [5], to measure the depth of the earthquake. The distance from the origin and magnitude level of the earthquake should be extracted from the scattered information.

5. EXPERIMENTAL RESULT

The prediction is based on fuzzy logic. The MATLAB software is used to estimate the impact of earthquake. The fuzzy rules are given in fig.3. The sample data are taken from United States of Geological Survey (USGS) database. Fifty records are used in Fuzzy inference System (FIS) and results are shown in fig.4. The rules are generated with the assumption that Magnitude (M), Depth (Dt) and Impact (I) are linguistic variables. The possible values for linguistic variables are as follows:

Magnitude (M) - Low, Medium and High

Depth (Dt) – Shallow (S) and Deep (D)

Impact (I) – Low (L), Medium (M) and High (H)

Fuzzy rules:

If M = Low and Dt = Shallow then I = Medium
If M = Low and Dt = Deep then I = High
If M = Medium and Dt = shallow then I = High
If M = Medium and Dt = Deep then I = Medium
If M = High and Dt = Shallow then I = high
If M = High and Dt = Deep then I = High

Fig. 3. Generation of fuzzy rules to predict the impact of earthquake

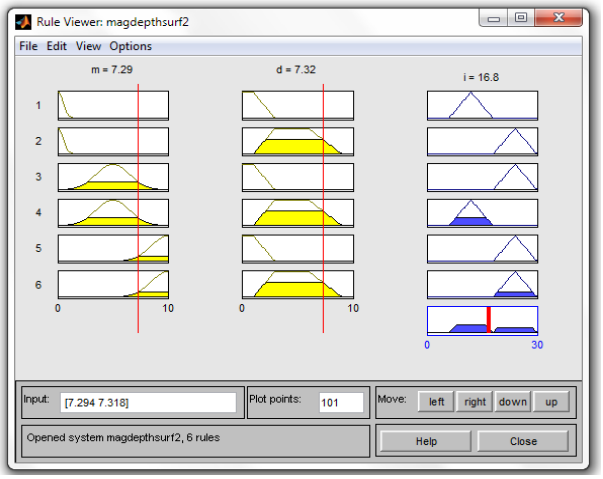

Fig. 4. Fuzzy Rule for impact of earthquake

The results of the fuzzy rule are shown in the table.1.

Table .1. Prediction Rule based on the fuzzy logic

Dt / M	Shallow	Deep
Low	M	H
Medium	H	M
High	H	H

The graphical output of the prediction is given in fig 5 as the Fuzzy surf diagram. The impacts of earthquakes are described in this diagram. The levels are changed from low to high. If M magnitude is low and Dt depth is shallow the impact of earthquake is medium. Likewise if M is high and Dt is shallow then impact of the earthquake is also high. If M is low and Dt is shallow, then impact of earthquake is medium. The maximum impact of the earthquake is shown in surf diagram. Therefore the fuzzy logic could predict the maximum number of occurrence.

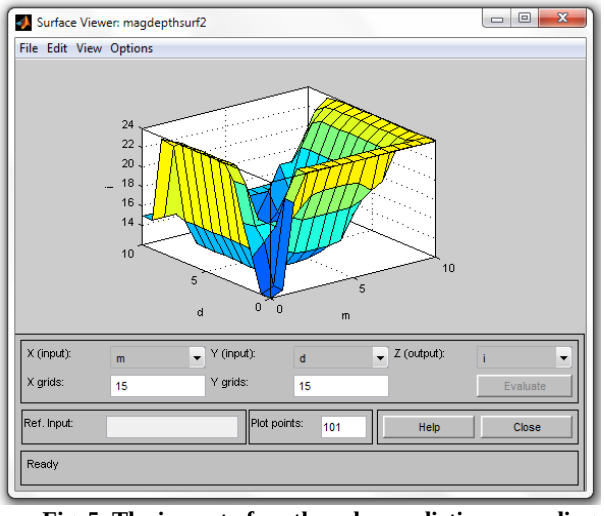

Fig. 5. The impact of earthquake prediction according to fuzzy inputs values

Based on the fuzzy rules the Indonesia datasets are analyzed and predicted the maximum of depth in the Indonesia area.

6. CONCLUSION AND FUTURE WORK

A time series is more important in the view of geophysics and economy problems. The shallow type earthquakes are majorly occurred in the Indonesia area. Most shallow earthquakes depths are start from 0km to 300kms down side from origin. The shallow earthquakes depths are identified using available methods. The obvious understanding of earthquake and precursors of earthquakes are used to create a new method for predicting earthquake. This will help to find the accurate impact of earthquake and distance from the origin point. Future work of identifying the ground water level, acceleration, intensity and p-wave changes. The fuzzy optimization algorithm should be applied to different time series data and predict the future values.

7. REFERENCES

[1]. United States of Geological Survey http://earthquake.usgs.gov

[2]. A. Morales-Esteban, F. Martinez- Alvarez, A. Troncoso, J.L. Justo, C. Rubio-Escudero, Pattern recognition to forecast seismic time series, Elsevier (2010).

[3]. Feature extraction, ArcView GIS Version 3.1.

[4]. J. S. R. Jang, C. T. Sun and E. Mizutani. Neuro-Fuzzy and Soft Computing. Prentice Hall, 1997. ,

[5]. I.Aydin, M. Karakose and E.Akin, The Prediction Algorithm Based on Fuzzy Logic Using Time Series Data Mining Method.

[6]. H. Kantz and T. Schreiber, Nonlinear Time Series Analysis, Cambridge: Cambridge University Press, 388 p., 1997.

[7]. J. Han and M. Kamber, Data Mining: Concepts and Techniques, San Francisco: Academic Press, 800 p., 2005

[8]. R. J. Povinelli, "Time Series Data Mining: Identifying Temporal Patterns for Characterization and Prediction of Time Series Events", Ph.D. Dissertation, Marquette University, 180 p., 1999.

[9]. Chris gray, A Review of Two Methods of Predicting Earthquakes.

[10]. A. Morales- Esteban, F. Martinez-Alvarez, A. Troncoso, J.L.Justo, C. Rubio-Escudero, Pattern recognition to forecast seismic time series, Elsevier, Experts System with Application 37(2010) 8333 – 8342

[11]. Muhammad Ardalani-Farsa, Saeed Zolfaghari, Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks, Elsevier, Neuro computing 73 (2010) 2540-2553.

[12]. Scott J. Goetz, Gregory J. Fiske, Andrew G. Bunn, Using satellite time-series data sets to analyze fire disturbance and forest recovery across Canada, Remote Sensing of Environment 101 (2006) 352–365.

[13]. Hao-Tien Liu, Mao-Len Wei, An improved fuzzy forecasting method for seasonal time series, Expert Systems with Applications 37 (2010) 6310–6318.

[14]. G. Peter Zhang, Time series forecasting using a hybrid ARIMA and neural network model, Neuro computing 50 (2003) 159 – 175.

[15]. T. Takagi, M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," IEEE Trans. Syst., Man Cyber, vol. 15, pp 116-132, 1985

[16]. Arash Andalib, Mehdi Zare, Farid Atry, a fuzzy expert system for earthquake prediction, case study: the zagros range, ICMSAO 2009

[17]. Neeti Bhargava, V. K. Katiyar, M. L. Sharma and P. Pradhan, Earthquake Prediction through Animal Behavior: A Review, Indian Journal of Biomechanics: Special Issue (NCBM 7-8 March 2009).

[18]. J Sajjad Mohsin, Faisal Azam, Computational seismic algorithm comparison for earthquake prediction. International Journal of Geology, Issue 3, Volume 5, 2011.