Expert Systems with Applications 38 (2011) 15032–15039

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals

Maria Moustra, Marios Avraamides, Chris Christodoulou *

Department of Computer Science, University of Cyprus, 75 Kallipoleos Avenue, P.O. Box 20537, 1678 Nicosia, Cyprus

ARTICLE INFO

Keywords:
Neural networks
Earthquake prediction
Seismic Electric Signals

ABSTRACT

The aim of this study is to evaluate the performance of artificial neural networks in predicting earthquakes occurring in the region of Greece with the use of different types of input data. More specifically, two different case studies are considered: the first concerns the prediction of the earthquake magnitude (M) of the following day and the second the prediction of the magnitude of the impending seismic event following the occurrence of pre-seismic signals, the so-called Seismic Electric Signals (SES), which are believed to occur prior to an earthquake, as well as the time lag between the SES and the seismic event itself. The neural network developed for the first case study used only time series magnitude data as input with the output being the magnitude of the following day. The resulting accuracy rate was 80.55% for all seismic events, but only 58.02% for the major seismic events ($M \geqslant 5.2$ on the Richter scale). Our second case study for earthquake prediction uses SES as input data to the neural networks developed. This case study is separated in two parts with the differentiating element being the way of constructing the missing SES. In the first part, where the missing SES were constructed randomly for all the seismic events, the resulting accuracy rates for the magnitude of upcoming seismic events were just over 60%. In the second part, where the missing SES were constructed for the major seismic events ($M \geqslant 5.0$ on the Richter scale) only by the use of neural networks reversely, the resulting accuracy rate by predicting only the magnitude was 84.01%, and by predicting both the magnitude and time lag was 83.56% for the magnitude and 92.96% for the time lag. Based on the results we conclude that, when the neural networks are trained by using the appropriate data they are able to generalise and predict unknown seismic events relatively accurately.

1. Introduction

Earthquakes are one of the most costly natural hazards faced by the nation in which they occur without an explicit warning and may cause serious injuries or loss of human lives as a result of damages to buildings or other rigid structures. During the last decades there has been an increasing interest and academic research on predicting seismic events. In the effort to predict earthquakes, people and researchers have tried to associate an impending earthquake with such varied phenomena as seismicity patterns, electromagnetic fields, weather conditions and unusual clouds, radon or hydrogen gas content of soil or ground water, water level in wells and animal behaviour (like for example the recently reported pre-seismic anticipatory behaviour in the common toad *Bufo bufo*, see Grant and Halliday (2010) and references therein for reports of seismic activity responses of other species). Earthquake prediction, which aims to specify three elements, namely when, where and how large the impending earthquake will be, constitutes the most important unsolved problem of seismology.

In the past, various efforts have been made to solve this particular problem. These efforts led to the construction of models, which attempted to comprehend the nature of seismic phenomena and predict high magnitude (M) seismic events based on different approaches. Some of the most important efforts which are related to this study are reviewed below. For a detailed survey of earthquake prediction efforts, see Adeli and Panakkat (2008).

One of the most debated methods for earthquake prediction is a method called "VAN", after the initials of three Greek scientists from the University of Athens, Varotsos, Alexopoulos and Nomicos (Uyeda, 1997). These scientists found that transient variations of the earth's electric field, known as Seismic Electric Signals (SES) are observed before an earthquake. The SES are used to determine the location and the magnitude of the impending earthquake. After a SES is recorded, an earthquake occurs within several days to several weeks based on the SES's type (Varotsos & Alexopoulos, 1984a, 1984b). To determine the epicentre of an impending earthquake, a process of elimination of the possible seismic areas is applied, including the selectivity effect, the polarity effect and the ratio of

* Corresponding author. Tel.: +357 22892752; fax: +357 22892701.
E-mail addresses: cs04mm2@cs.ucy.ac.cy (M. Moustra), cs03am1@cs.ucy.ac.cy (M. Avraamides), cchrist@cs.ucy.ac.cy (C. Christodoulou).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.05.043

the two components of the SES. The magnitude of the earthquake is estimated using the data from the specific station which was collected for the same seismic area in the past (Varotsos & Lazaridou, 1991). The VAN group scores their prediction "successful" when the actual earthquake occurred within several days to several weeks after the precursor SES is recorded, within ca. 100 km from the predicted epicentre and within ca. 0.7 units of the predicted magnitude on the Richter scale (RS). Based on these criteria, about 60% of their predictions are successful and about 60% of Greek earthquakes of $M > 5.3$ on the RS are successfully predicted (Uyeda, 1997). However, this success has not been widely recognised by the scientific community, including the Greek seismological community. Some argue for example, that the VAN's SES are all noise unrelated to earthquakes, and others persist that the success is attained by chance (Uyeda, 1997).

Bodri (2001) attempted to relate neural network ideas to seismic activity patterns in the Carpathian–Pannonian area of Hungary, and the Pelloponnesos–Aegean region of Greece. A three-layer feed-forward multilayer perceptron neural network model using error backpropagation as learning algorithm was developed for the prediction of the origin times of large earthquakes ($M \geqslant 6.0$ on the RS). The network used as input the mean seismicity rates (number of earthquakes per unit time) in selected time intervals, and more specifically within the time intervals between the recorded $M \geqslant 6.0$ (RS) earthquakes. The results of this effort were particularly satisfactory despite the fact that the training set was inadequate because of the infrequent occurrence of large earthquakes. The neural networks managed to predict the origin times of such events with a deviation of ±6 months. The impressive performance of the neural networks revealed the usefulness of such tools in the problem of earthquake prediction.

Lakshmi and Tiwari (2007) examined the temporal evolution of seismicity of the Central Himalaya (CH), Western Himalaya (WH) and Northeast Himalaya (NEH). A multilayer feed-forward artificial neural network (ANN) model was developed to simulate monthly resolution earthquake frequency time series for the three regions. The learning algorithm used was the backpropagation with gradient descent optimisation technique. Cross-validation was also utilised to test the networks generalisation ability. The data used concerned the seismic events which occurred in the period of 1960–2003 and for magnitude of $M \geqslant 4$ (RS). The earthquake monthly frequency data was used as input values to the neural network. More specifically, a temporal sequence of the previous five monthly frequency data was selected as input. The frequency value of the next month was used as output of the network. According to the sum-squared error that was calculated for each region, in order to measure the differences between the actual and predicted values, the results obtained by the ANN model were described as reasonably good. Furthermore, the results showed that the earthquake dynamics in the regions of WH and NEH are better "organised" than in the CH, since the earthquake processes of WH and NEH have a higher predictive correlation coefficient, at 50–55%, in contrast to the CH which has 30%.

Lakkos, Hadjiprocopis, Comley, and Smith (1994) used a feed-forward neural network which was simulated using the XERION software package (van Camp, 1993) and the Delta-Bar-Delta as training algorithm (van Camp, 1993) to predict the geographical location (longitude and latitude) and the magnitude of an impending earthquake. The input data presented to the network consisted of the two components of the SES of the VAN method (East–West, North–South). The data used for training was collected by the monitoring station of Ioannina in North Western Greece, but the data quantity was not sufficient. Thus they expanded the original data set by a factor of five by adding a small amount of Gaussian noise. After testing the neural network using data that was not part of the training data set, the results showed that the network was able to give more accurate predictions for the geographical area of 20.0°E–21.5°E and 37.5°N–40.0°N since the majority of training data was associated with this area. In addition, the epicentre location can be predicted with an error of less than 0.3° and the magnitude with an error of less than 0.5 on the RS (Lakkos et al., 1994). However, the authors do not clarify the magnitude range of the data used for training and testing.

Another example is the probabilistic neural network (PNN) that was implemented for the magnitude prediction of the largest earthquake in a pre-defined future time period (Adeli & Panakkat, 2009). More specifically, the future time period that they considered was the following fifteen days and the region examined was Southern California. The PNN takes as input eight mathematical earthquake parameters called seismicity indicators (Gutenberg & Richter, 1956) and classifies the predicted magnitude in one of the several output classes. The indicators are: the time elapsed during a particular number (n) of significant seismic events before the month in question, the slope of the Gutenberg-Richter inverse power law curve for the n events, the mean square deviation about the regression line based on the Gutenberg-Richter inverse power law for the n events, the average magnitude of the last n events, the difference between the observed maximum magnitude among the last n events and the expected ones using the Gutenberg–Richter relationship (Gutenberg & Richter, 1956) known as the magnitude deficit, the rate of square root of seismic energy released during the n events, the mean time or period between characteristic events, and the coefficient of variation of the mean time. Three different statistical measures have been used for the model's evaluation: the probability of detection, the false alarm ratio and the true skill score. According to the results based on these three metrics, the PNN gave good prediction accuracies for magnitudes between 4.5 and 6.0 (RS), but not for magnitudes greater than 6.0 (RS).

Artificial neural networks (ANNs) have been used by many researchers to investigate their potential as a tool for simulation of the behaviour of systems that are governed by nonlinear multivariate data and generally unknown interconnections within a noisy, poorly-controllable physical environment. The advantage of this framework is that the ANN provides a black-box approach and the user does not need to know much about the nature of the process being simulated. Considering this advantage, in the present paper we describe two test cases for earthquake prediction for the region of Greece by applying ANNs. A number of inherent features of ANNs make them suitable for such a task and for a huge number of other applications (see for example Velido, Lisboa, and Vaughan (1999) and Paliwal and Kumar (2009), for surveys). More specifically, in contrast to model-based techniques, ANNs are data-driven self-adaptive (through learning) systems, as they do not need any *a priori* assumptions with regards to the models of the scenarios being investigated, or if they do, they are minimal. ANNs can usually generalise pretty well after they are trained with a sample of the data, which could even be noisy (Haykin, 2009).

The remainder of the paper is structured as follows: Section 2 describes the first case study which concerns earthquake prediction using only time series magnitude data, Section 3 describes the second case study which concerns earthquake prediction using the SES of the VAN method discussed above, and finally Section 4 gives a discussion and conclusions.

2. Case study I: earthquake prediction using only time series magnitude data

2.1. Methodology and data preprocessing

The first case study outlines a methodology for predicting the magnitude of the most important seismic event for the following