A

Prediction of Major Earthquakes as Rare Events Using RF-Typed Polynomial Neural Networks

Simon Fong
Department of Computer and Information Science, University of Macau, China

Suash Deb
Cambridge Institute of Technology, India

INTRODUCTION

In the history of mankind, earthquakes inflicted greatest loss of life surpassing the events of world wars and plagues, devastation to civilization, and huge loss in economy of a nation. Unfortunately, as of now, there exists no best method in accurate and reliable earthquake prediction. In general there are three schools of thoughts pertaining to earthquake prediction: one being optimistic that earthquake could be predicted accurately one day given sufficient external data (Jordan, 2011); the other felt that certain trending may be possible by assuming and abstracting all the factors behind – the historical results as earthquakes that happened have already embraced all the environmental factors which contributed to the events. So similar to trend-following strategies in stock markets (Fong, 2012), by observing just the tending data from the earthquake time-series it is possible to forecast the forthcoming events. The last group of researchers disbelieves the possibility and reliability of such forecasts as the variations of earthquakes are seemed to be too high (Wang, 2006). They are nothing more than a random process.

Technically earthquake data when viewed as a time series over a long time, exhibits a complex pattern that is composed of a mix of statistical features. The magnitude variation is high, the occurrence of extreme values is sporadic; hence the white noise component is dominating backed by the fact that many minor earthquakes are occurring in almost every minute somewhere around the world. This characterizes a situation of rare-event prediction whose difficulty has been well recognized in data mining and statistics research communities.

In this article, we explore the outlook of applying a highly non-linear prediction model, namely Polynomial Neural Network (PNN) as an alternative to traditional time-series forecasting algorithm (TTF), on major earthquake prediction. In particular, major earthquakes are treated as rare-events. Thus we opt to select only the major earthquakes, converting the univariate time-series to multivariate dataset with relevant extra inputs. Providing relevant multivariate inputs to PNN is important because the strength of neural network is on predicting outcomes from multivariate data. PNN, as its name indicates, is a well-known ensemble type of prediction method that is capable of modeling highly non-linear relations, and achieving an optimal accuracy by inducing through all possible structures of polynomial forecasting models.

In TTF, the prediction models are usually based on the input of univariate time series. PNN is extended with residual-feedback (RF) and Hurst factors; hence, the name RF-typed PNN. We investigate the qualities of different prediction models by TTF and PNN and search for prediction results that offer the lowest error of curve fitting. This article contributes a step forward in finding an accurate prediction model from the perspective of rare-event forecasting.

BACKGROUND

Decades ago, classical methods such as Omori law (Omori & Utsu, 1961) and Gutenberg-Richter law (Gutenberg & Richter, 1944), laid a mathematical foundation in modeling temporal pattern of aftershocks

DOI: 10.4018/978-1-4666-5888-2.ch022

and frequency and magnitude respectively. Later on the time series that were used in forecasting was believed to be composed individually from components: trend, seasonality, cycle, long-term memory and white noise. Since 1970's temporal forecasting techniques became prevalent such as AR, MA, ARIMA, SARIMA and ARIMAX, etc. At about the same time, in the 1970s an intense optimism arose amongst scientists advocating that some of these methods and/or their variants are able to predict earthquakes to certain accuracy. A famous claim of a successful prediction is that alleged for the 1975 Haicheng earthquake (Jackson, 2004). However, continuing failure by the 1990s arouse questions on the efficacy of the prediction algorithms (Geller, 1997). A latter study refuted the efficacy that there was no valid short-term prediction scientifically possible (Wang, 2006). The controversy continued. While many scientists still hold that, given sufficient resources, prediction might be possible, many others maintain that earthquake prediction is inherently impossible (Kagan, 1997) like a random guess. With more than 13,000 earthquakes of magnitude 4+ happen around us each year, anybody could easily and luckily score a correct prediction with broad variables in some season of some year, some regions of a hemisphere, or some approximate scale of magnitude (Mabey, 2001). It was accused that such so-called "successful predictions" are vague and not helpful at all, as we generally know that at least one or two significant earthquake will usually take place at some point of time in a year. What we are concerned about are useful predictions of disastrous earthquakes such are those with magnitude 7 or greater, with a predicted time-frame preferably down to week or month of a year, and a variance +/- 0.5 in magnitude.

Reviewing the earthquake prediction methods being used pertaining to the forecast mean and acceptable level of variance, they may carry several limitations at large. These techniques are generally referred to traditional time-series forecasting techniques (TTF) (Montgomery & Johnson, 1976) in the context of this article. Firstly the forecast is produced usually in the form of a mean value because these algorithms are founded upon the mechanisms of estimating moving average values and exponential smoothing over a time series. In some special cases such as predicting disasters, rare events rather than average values or normal happenings are of interest. The second shortcoming is the manual selection of parameters for calibrating traditional forecasting techniques (TTFs). It is well known that a slight change in the values of these parameters will lead to extremely poor fitting of the forecast model to the actual data.

On the other hand, rare event forecasting has been studied extensively in past years (Vilalta & Ma, 2002), targeting at solving the abovementioned fundamental problems. Two sources of problems are being identified in contributing to the difficulty of rare-event forecasting. The first lies in data collection. In rare events data, researchers may be overwhelmed with very large numbers of observations but with few samples of rare predictable variables. In this article we adopt a simple way of data collection approach where only the event of maximum earthquake happened in each year is retained. Secondly, as it was pointed out (King & Zeng, 2001) in rare events the biases in probabilities can be substantively significant in a predictable direction. Similarly in TTFs that are based on regression methods, the probabilities of predicted rare events may reside sub-optimally in finite samples of rare events data, leading to errors in the same direction as biases in the coefficients.

An alternative approach to traditional forecasting techniques (TTFs) is neural network that has completely different underlying mechanisms from forecasting models of moving average. The strength of neural network is centered in capturing non-linear relationships and the structural characteristics of a time series that are fuzzy and chaotic in nature. By training a neural network with suitable architecture of neurons and layers with past data, the irregular patterns of the time series could be "memorized" by adjusting the weights of the neurons; and it is able to produce a prediction of a future value or values that exhibit a similar underlying temporal pattern mirroring from the past. Nevertheless, the same problem of finding the optimal parameters of neurons and layers for a neural network for the best prediction result still persists. The selection range is very wide and any careless configuration will lead to very different results. To solve this problem, an alternative type of forecasting techniques, called Polynomial Neural Network (PNN) (Madala & Ivakhnenko, 1994) that is based on the 1971 principle of Group Method Data Handling (GMDH) (Ivakhnenko, 1970) was proposed.

PNN has the advantages of generating an ideally fit model by incrementally evolving the structure of the neural network model from simple to complex, while fine-tuning the underlying neuron structure that optimally model the non-linear and varying behavior