https://doi.org/10.20965/jaciii.2014.p0701

Paper:

Artificial Neural Networks for Earthquake Anomaly Detection

Aditya Sriram, Shahryar Rahanamayan, and Farid Bourennani

University of Ontario Institute of Technology (UOIT)
2000 Simcoe Street North, Oshawa, Ontario L1H 7K4, Canada
E-mail: shahryar.rahnamayan@uoit.ca

[Received October 29, 2013; accepted March 6, 2014]

Earthquakes are natural disasters caused by an unexpected release of seismic energy from extreme levels of stress within the earth's crust. Over the years, earthquake prediction has been a controversial research subject that has challenged even the smartest of minds. Because numerous seismic precursors and other factors exist that may indicate the potential of an earthquake occurring, it is extremely difficult to predict the exact time, location, and magnitude of an impending quake. Nevertheless, evaluating a combination of these precursors through advances in Artificial Intelligence (AI) can certainly increase the possibility of predicting an earthquake. The sole purpose for predicting a seismic event at a pre-determined locality is to provide substantial time for the citizens to take precautionary measures. With this in mind, Artificial Neural Networks (ANNs) have been promising techniques for the detection and prediction of locally impending earthquakes based on valid seismic information. To highlight the recent trends in earthquake abnormality detection, including various ideas and applications, in the field of Neural Networks, valid papers related to ANNs are reviewed and presented herein.

Keywords: artificial neural networks, ANN, earthquake anomalies detection, precursor, earthquake prediction

1. Introduction

An earthquake is a natural disaster caused by an unexpected release of seismic energy due to extreme stress within the earth's crust. Such energy is released because of aggressive movements of the tectonic plates in active fault zones. The accumulated energy, containing immense pressure, is transferred from the earth's crust to its surface in the form of seismic waves. These waves can either roll or travel parallel to the surface causing the destruction of anything that falls within its path. Earthquakes can create severe structural damages, irretrievable financial ruin, and irrecoverable loss of human life.

Over the years, earthquake prediction has been a controversial subject that has challenged even the brightest researchers. Because numerous seismic precursors and other factors indicating a potential earthquake exist, it is extremely difficult to predict the exact time, location, and magnitude of an impending quake. Nevertheless, evaluating a combination of these precursors through advances in Artificial Intelligence (AI) can increase the possibility of an earthquake prediction.

In this direction, Neural Networks (NNs) have been utilized to translate seismic information and provide a valid detection and prediction of locally impending earthquakes. A Neural Network is an AI method inspired from the functionality of the human brain. A NN consists of interconnected neurons, weights, links, activation functions, and a training set through which the system "learns" from experience by corresponding with output errors [1]. The accuracy of NNs predictions depends highly on the network's output uncertainties; a network adjusts itself using provided learning method to minimize output errors [2]. NNs have the aptitude to deduce patterns and detect trends that are nearly impossible for humans to recognize, and hence are a valuable commodity for detecting seismic activity.

For example, on February 4, 1975, a 7.3 magnitude earthquake struck the city of Haicheng in Northeast China, resulting in over 2,041 casualties, leaving thousands of people homeless, and destroying various structures in its path [3]. Chinese officials confirmed that an earthquake warning was announced only hours before the main shock occurred [4, 5]. The impending earthquake was successful predicted based on various seismic precursors observed by seismologists and other scientists [6]. The most important precursor was a sequence of foreshocks, although other precursors such as abnormal animal behavior, radon activity, changes in land and ground water elevations, and altered chemical properties each played a vital role prior to the evacuation [4, 6]. This specific case shows that earthquakes may provide multiple precursors. When these different precursors are integrated through an NN analogy, they can increase the probability of predicting earthquakes with a higher accuracy [7].

This paper describes background information on earthquake properties, and surveys the scientific possibilities of earthquake prediction using NNs. In addition, this survey provides a detailed layout of different seismic precursors such as peak ground acceleration, liquefaction, radon detection, and aftershocks. Moreover, this paper discusses how these seismic precursors are utilized by NN computations for earthquake prediction and detection. A respective network analysis is presented for each type of seismic precursor, along with the type of NN used. In addition, a detailed explanation of the objective of the network, a description of its input and output neurons, and a review of the training and testing phases are provided. Finally, this paper surveys recent trends in detecting earthquake abnormalities and current NN applications for seismic prediction.

2. Strong Ground Motion Analysis as a Seismic Precursor

Strong ground motion is a sudden violent tremble on the surface of the earth that occurs before an imminent earthquake. Seismic instruments such as accelerometers are widely used in earthquake-prone areas to monitor and collect such data. Located approximately 30 m below the surface of the earth, accelerometers are among the most essential tools used to acquire input parameter readings for various types of ground motion analyses [8]. However, as one of the many tools providing impending earthquake information, accelerometers are ineffective as primary tools for seismic detection because they only observe vibrations at high frequencies. Over the years, ground motion prediction using various NN approaches and methodologies has been a highly reviewed subject. Idriss [9] conducted an extensive ground motion analysis using related content periodically collected until 1978. Boore and Joyner [10] followed up by incorporating important ground motion prediction equations in 1981; their studies laid the foundations for earthquake prediction using a network analogy. Later, Campbell [11] conducted a wider range of ground-motion analyses up to 1985 that contained vital equations and innovative precursory analogies.

In this section, we blend critical ground motion applications into network architecture to increase the likelihood of a strong ground motion prediction. The ground motion applications taken into consideration as precursors to an earthquake are a Peak Ground Acceleration (PGA) and the potential liquefaction. The following subsection features a thorough survey on various approaches to predicting a PGA and the possibility of liquefaction using ANNs.

2.1. Predicting Peak Ground Acceleration Using Artificial Neural Networks

A PGA is a measure of earthquake acceleration with respect to extensive ground-shaking movements [12]. As a seismic precursor, a PGA is induced through an intense release of energy from an earthquake, causing ground deformations such as liquefaction, landslides, and surface fault ruptures [13, 14].

Derras and Bekkouche [15] introduced a comprehensive approach to estimating the maximum PGA using a Feed-forward Back-propagation Neural Network (FFBP-ANN). The outcome of the network was then compared to two Ground Motion Prediction Equations (GMPE) mod-

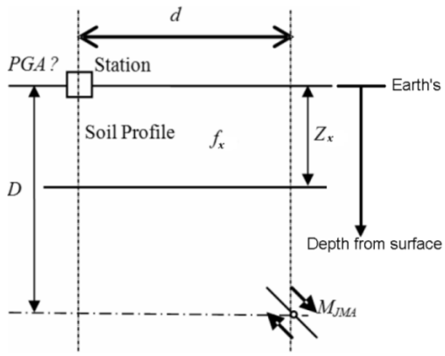

Fig. 1. Indicated site seismic parameters utilized in a PGA evaluation.

eled by Ambraseys and Takahashi [15], respectively. The GMPE models were used as an alternative approach for the estimation of the PGA values where accelerometric monitoring stations are not present. Such an approach requires a large volume of data on the site coefficients as well as pre-recorded PGA values. In return, the GMPE theory has been observed to be relatively weaker in terms of prediction compared to an FFBP-ANN owing to its inability to cope with non-linear expressions and complex data types. An FFBP-ANN was designed with a total selected set of 1,000 epochs and a tangential-hyperbolic sigmoid/linear activation function, and consists of five input parameters: the locally measured meteorological agency magnitude, the depth of focus at which an earthquake is triggered, the epicenter distance, the thickness of the sedimentary layers (Z_x), and the corresponding resonant frequency (f_x). Importantly, both the sedimentary thickness and the resonant frequency have a constant shear wave velocity of $x = 800$ m/s, as depicted in [16]. **Fig. 1** [15] provides a visual representation of the site parameters utilized for evaluation.

The ANN configuration utilized in **Fig. 1** consists of 326 training and 1,850 testing records extracted from KiK-net data. A comparison between an FFBP-ANN, Ambraseys's GMPE model, and Takahashi's GMPE model shows that the performance of the NN (FFBP-ANN) is far superior to the two GMPEs. The coefficient of determination (R^2) for the PGA estimated by an NN is 0.94 as compared to those of the GMPE approaches, which are 0.76 and 0.82, respectively. In the same venue, the NMRSE for the NN approach is considerably smaller, at 0.11%, as compared to the GMPE models, which shows a respective NMRSE of 0.25% and 0.17% confirming that a PGA approximation using an NN surpasses that of the GMPE models in terms of both performance and accuracy. Specifically, the results show that the epicentral distance parameter heavily influences the outcome of the PGA value, obtaining the best R and MSE values of 0.51/0.48 and 0.075/0.076 for the training and testing phases, respectively. In contrast, the focal depth and site parameters have the least influence on the outcome of the PGA value. Furthermore, a combination of all five parameters is observed to return the optimal results, retaining an $R-score$ of 0.85 and 0.84, and an MSE value of 0.0203 and 0.0205 correspondingly, for the training and testing phases. The $R-score$ is a correlational coefficient that indicates the degree of independence between an input and output. An R value closer to ± 1 indicates a strong correlation; however, when $R=0$, the prediction is considered inaccurate.

Various novel approaches to an ANN have been used to approximate PGA values, leading to different accuracy levels and performance ratings. Kemal and Ayten [17] introduced a comparative approach to PGA prediction by evaluating three types of ANNs, namely, Feed-Forward Back-Propagation (FFBP), a Radial Basis Function (RBF), and a Generalized Regression Neural Network (GRNN). All these versions of ANNs taken into consideration were trained using a Back-Propagation (BP) methodology with only one hidden layer, and were evaluated using four input parameters, namely, the earthquake moment magnitude, hypo-central distance, focal depth, and site conditions. These input parameters, along with 95 records from 15 accelerometers, were captured from three wave directions (i.e., up–down, north–south, and east–west). From the 95 records selected, 72 datasets were used for training the network, leaving 23 sets of data for testing purposes. The performance results show that the FFBP-ANN achieved the highest $R-score$ and lowest $RMSE$ and MAE scores of 0.856, 44.45, and 18.46 in the east-west (horizontal) direction. To obtain a high level of accuracy, the predicted values generated by the FFBP-ANN were further modified through a linear regression analysis to make the prediction more sensitive. On the other hand, the GRNN provided a better performance (for PGA $>$20 cm/s^2) in the vertical direction compared to the other two ANNs with an $R-score$, $RMSE$, and MAE of 0.933, 14.22, and 5.27 respectively. Above all, the RBF-ANN is the poorest performer in both the vertical and horizontal directions. A detailed comparison between the NN efficiencies can be found in [17]. To summarize, for a horizontal PGA prediction, the FFBP-ANN is the ideal choice, whereas for predicting vertical PGA values, a GRNN-ANN is preferable.

Ambraseys and Douglas [18] performed a similar analysis on near-field horizontal and vertical ground motions triggered by an earthquake. The results suggest that a vertical PGA decays faster with respect to distance compared to a horizontal PGA value. Specifically, for a magnitude of 6.0, a vertical PGA drops from approximately 0.33 to 0.22 g in 5 km, whereas a horizontal PGA only drops from 0.39 to 0.37 g. This suggests that vertical PGAs have a higher frequency, and therefore attenuate more rapidly compared to horizontal ground motions. In addition, an investigation on this subject also showed that vertical motions contain much less energy compared to horizontal motions. Arguably, it would be effective to obtain more vertical motion data from the accelerometers because this component is perceived to influence the PGA values exclusively.

Kerh and Chu [19] explored a different direction in predicting PGA values at a given site by comparing an ANN approach with micro-tremor measurements. Micro-

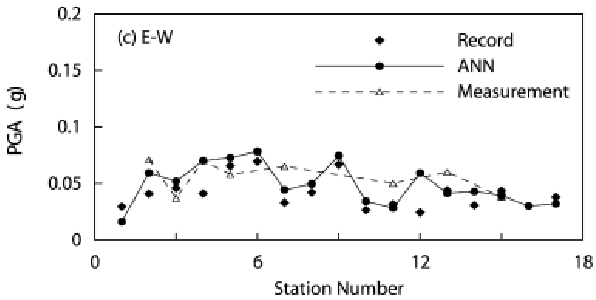

Fig. 2. A comparison of PGA prediction using an ANN and micro-tremor measurements.

tremor measurements, on the other hand, are an empirical method for collecting all ground vibrations within a very short span of time, thereby having a clear advantage over a classical seismic record database. Their ANN approach was developed by employing pre-defined data points for training, and validating the network using a BP analogy. Specifically, three input neurons were evaluated: the epicentral distance, focal depth, and magnitude of the earthquake. The results showed that using three input parameters achieved an $R-score$ of 0.972, which is significantly higher compared to two input parameters (averaging 0.6 to 0.9) and a single input parameter (averaging below 0.6). A comparison of this subject has shown clear results: micro-tremor measurements are definitively useful when time is constrained, but are lacking in performance ability. In contrast, an ANN is remarkable in its performance, but as a weakness is highly dependent on pre-defined data. **Fig. 2** shows a comparison of the PGA estimation between the use of an ANN and micro-tremor measurements [19].

Kerh and Ting [20] also brought attention to the use of PGA estimation, confirming that increasing the number of input neurons has an enhancing effect on the correlational coefficient. Their results showed that the Nakamura transformation method using micro-tremor measurements closely resembles the ANN results described in [19]. In this context, the estimation of the PGA value using traditional seismic data embedded into an ANN is fairly smoother, providing an 80% more accurate performance as compared to micro-tremor measurements. In brief, this concept reveals an area for further research by integrating micro-tremor measurements as an NN parameter under the application of PGA prediction. Granted, although this approach may have a higher chance of uncertainty in comparison to an ANN with pre-defined seismic datasets, its collaboration may result in a more accurate performance if the data retrieved are minimal and time is constrained.

2.2. Predicting Liquefaction Potential Using Artificial Neural Networks

Soil liquefaction is a phenomenon that occurs when saturated soil changes its state into liquid-like behavior after losing its strength, density, and stiffness in response to sudden ground motions caused by seismic events. In the event of liquefaction, various infrastructures will be severely damaged, and liquefaction should therefore be considered a vital precursor. A response spectrum is an earthquake-engineering tool utilized for analysis of a structure quality and performance by evaluating its natural frequency. The purpose of introducing a response spectrum is to put forward the importance of liquefaction, and provide the basic information on the true effects that this precursor can cause on various applications. Although response spectra are not discussed comprehensively in regard to the subject of an ANN in this survey, the intension here is to relate the effects of liquefaction on a response spectrum application.

Derras et al. [21] proposed an MLP-ANN approach in combination with a retro-propagation learning methodology for determining a response spectrum by utilizing appropriate seismic data on a given site. To train, test, and validate the network, KiK-net accelerographs and soil dynamics were utilized to simulate the response spectrum on the surface. The behavior of the network prediction reveals that spectral acceleration is highly dependent on the behavior of the soil conditions, and is hence used as input. Under such circumstances where the input parameters are highly sensitive and non-linear, a reliable input variable is required to reduce network uncertainties. The ANN considered is composed of two input neurons: the response spectrum of acceleration, and its time periods. In addition, the NN is composed of two hidden layers with a non-linear hyperbolic tangent activation function. The validation of the network involves comparing the response spectra estimated by the NN to that estimated by TOTTORI, which are datasets obtained from accelerograms. Overall, the estimated response spectra value using the NN is identical to the accelerogram readings, and a graphical representation on their relationship can be found in [21].

Similarly, Bojorquez et al. [22] used an FF-MLP ANN along with a BP training methodology to predict inelastic response spectra. In addition, five different types of soils were taken into consideration for testing purposes. To train the network, 50 ground motion records were extracted from the NGA database; specifically, moment magnitudes ranging from 5.9 to 7.7 were considered, covering most of the moderate to large seismic events. The results show that it becomes increasingly difficult to estimate the response spectra as the soil softens. Although the test results obtained in this study were acceptable, the overall error increased as the ductility parameter grew larger. A graphical representation of the comparison between the actual data and the ANN earthquake response spectra can be found in [22].

Liquefaction is determined by integrating the major ground parameters, which involves PGA values as an application for analysis. Goh [14] conducted a study on the evaluation of seismic liquefaction using a Probabilistic Neural Network (PNN) based on the Bayesian classifier method through two different analyses. The first analysis was based on Cone Penetration (CPT) datasets, and the second involved data on the shear wave velocity. The liquefaction potential is commonly examined using in situ testing methods. However, owing to the complex analysis of non-linear relationships such as seismic and soil properties, a composite structure such as an NN is desirable. In particular, the Bayesian classifier method is an appropriate complement to a PNN for pattern-recognition when predicting a sensitive parameter such as the liquefaction potential. The architecture of the PNN consists of four layers and six input neurons, namely, the earthquake magnitude, surface-level PGA, cumulative vertical overburden stress, the effective vertical overburden stress, the measured CPT tip resistance, and the mean grain size. During the training phase of the network, it was found that the PNN can discover the difference among the soil and seismic parameters and the liquefaction potential, and it is therefore not mandatory to scale or normalize the parameters. This phenomenon is specific to a PNN, which makes a PNN an advantageous architecture compared to any other NN structures. Moreover, CPT prediction using a PNN resulted in an overall success rate of 100%, with a standard deviation of 0.2235, and without errors generated during the training and testing phases. In contrast, the conventional (in situ) method using the same dataset resulted in 17 errors and an overall success rate of 90%. On the other hand, the shear wave velocity approach generated an overall output rate of 98%, with a standard deviation of 0.1265, and outputting two errors during both the training phase and the testing phase. In contrast, the traditional in situ method outputted roughly 60 errors with an overall success rate of 68%. The comparative results obtained are truly astounding, and illustrate that NNs possess high-quality pattern-detection ability when specifically considering the prediction of sensitive precursors.

To determine vital precursors, such as an occurrence of liquefaction at a given site, it is extremely essential to obtain a healthy set of data from the ground. When the samples obtained from the ground are limited, it may jeopardize the overall outcome of a network, and therefore, an alternative approach should be examined to deal with such constrained situations. Kumar et al. [23] introduced an innovative approach to predicting the possibility of liquefaction of alluvial soil using situ measurement systems based on a Standard Penetration Test (SPT) value and concatenating it in terms of an FFBP-ANN architecture. An SPT is a soil strength parameter used as a guide to obtain the ground conditions when inadequate borehole samples of sand, gravel, weak rocks, and clay are present. The established FFBP-ANN was then compared with the traditional Idriss and Boulanger (I&B) [23] method, which also uses an SPT analogy. A SPT is relatively inexpensive to conduct and the results are simpler to comprehend as compared to an ANN, and it is used specifically to determine the liquefaction potential. Adversely, an SPT only provides useful results when the soil density is low, and yields a poor outcome when highly dense soil conditions such as clay or gravel are present. Embedding an SPT as an input neuron complements an NN by simulating specific soil input properties, thereby increasing the likelihood of predicting a complex non-linear precursor, such as the liquefaction potential. The ANN architecture comprises six input parameters: the depth, SPT value, classification of the soil, natural moisture content, angle of internal friction, and particle size (which is slightly less than 2 mm). The training phase includes 160 datasets obtained from field and laboratory tests, 133 of which were used to develop the ANN, with the other 27 used to validate the network. A network analysis revealed that the regression for validation varies from 0.967 to 0.9969, and the average absolute error ranges from 0.929 to 2.687%. In addition, the Pearson's correlational coefficient for the ANN models exceeded 0.96, which is a healthy correlation among the parameters. Overall, the FFBP-ANN exceeded the performance of the I&B methodology. Comprehensive details on its liquefaction outcome can be found in [23]. Therefore, the use of an FFBP-ANN with SPT datasets provides accurate results overall, and should be considered when the ground condition samples are inadequate for evaluation.

In brief, the gap between a NN analogy and a ground motion analysis was successfully bridged. In particular, the PGA and liquefaction potential precursors were surveyed in adequate detail. In addition to the implementation of a PGA on an ANN, various types of NNs and alternative datasets were reviewed. Moreover, the effects of liquefaction on the response spectra were also discussed to stress the importance of the precursor. Network analyses along with the ground motion precursors show that an FFBP-ANN is the ideal choice for predicting the PGA values; furthermore, it is also evident that a PNN will have superior performance when predicting the liquefaction potential. The strategies exposed in this subsection show that an NN is exceptional in its performance with regard to determining a ground motion analysis. For instance, a PGA is a vital precursor that can be determined traditionally using accelerometers; however, there are limitations to this technology, such as inaccuracy in reading certain frequencies, and sensitivity to directional vibrations. These uncertainties are refined when using an NN analysis, owing to the immense training that a network undergoes. It has been established that accelerometers are critical to determining the PGA values, but integrating the datasets obtained by accelerometers and embedding them as input neurons, along with other sensitive parameters, can provide phenomenal levels of performance and accuracy. This in-turn can address greater insight into ground motion analytics.

3. Artificial Neural Networks on Large-Magnitude Earthquakes

The purpose of this subsection is to provide a survey on predicting large-magnitude main-shock and seismic aftershock events using an NN analogy. Scientists around the world have strived over the years to predict earthquake magnitudes with precision and accuracy. Despite the number of scientific attempts and amount of precursory information, the accuracy of such prediction remains strained. Large earthquake magnitudes cannot be determined using a specific tool or parameter; they are highly dependent on various parameters, and require extensive training to obtain an accurate prediction. Consequently, an analysis of the right type of precursors using an appropriate NN can significantly increase the chances of prediction and the overall accuracy of the network.

3.1. Predicting Large-Magnitude Main Shocks

In 2007, Panakkat and Adeli [24] conducted an analysis by comparing three different ANN models, namely, a Feed-Forward Levenberg-Marquardt Back-propagation Neural Network (FF-LMBP), a Recurrent Neural Network (RNN), and a Radial Basis Function (RBF) NN, to ultimately predict a largest earthquake one month in advance. The input parameters were similar to the neurons depicted in [25]. Comparative results among the networks showed that for magnitudes ranging from 5.0 to 6.0, the RNN obtained an $R-score$ between 0.20 and 0.51, whereas, RBF and LMBP networks acquired an $R-score$ of 0.12 to 0.37 and 0.01 to 0.14, respectively. Therefore, it is clear that an RNN has the ability to predict large earthquakes (above 6.0) faster and more accurately than LMBP and RBF networks.

Two years later, Adeli and Panakkat [25] presented an innovative method by introducing both an RNN and a PNN to predict the magnitude of the largest earthquake that is likely to occur in a pre-defined time frame in the near future. Their parameters, methodology, and process are the same as in [24]. The network models considered do not require a training phase, and hence are significantly faster than a BP-NN. The testing phase revealed that the PNN model obtains accurate predictions only when the magnitude ranges between 4.5 and 6.0, which yields an $R-score$ of 0.62 to 0.78; however, the prediction of a magnitude above 6.0 is very poor, returning an $R-score$ of 0 to 0.5. An RNN, on the other hand, outperforms a PNN for magnitudes ranging from 6.0 to 7.5, yielding an $R-score$ of 0.5 to 1.0; however, an RNN obtains a minimal performance when the magnitude is less than 6.0, returning an $R-score$ between 0.36 and 0.51. From the analysis held by Panakkat and Adeli [25], it is apparent that an RNN is the preferable choice for predicting large-magnitude earthquakes as compared to other NN analogies. **Fig. 3** [24] shows the layout of an RNN.

Simultaneously, Panakkat and Adeli [26] implemented an RNN using the Levenberg-Marquardt training algorithm for approximating the time occurrence of an earthquake and its corresponding geological location. They then compared the network to actual datasets generated in sub-regions of California. The prediction accuracies were evaluated through statistical measurements: the Probability Of Detection (POD), False Alarm Ratio (FAR), frequency bias, and $R-score$. The RNN was able to predict the epicentral distance and time of occurrence of four major earthquakes with a magnitude of greater than 5.5. However, the outcome of the network yielded location errors that ranged from 15 to 39 miles. On the other hand,

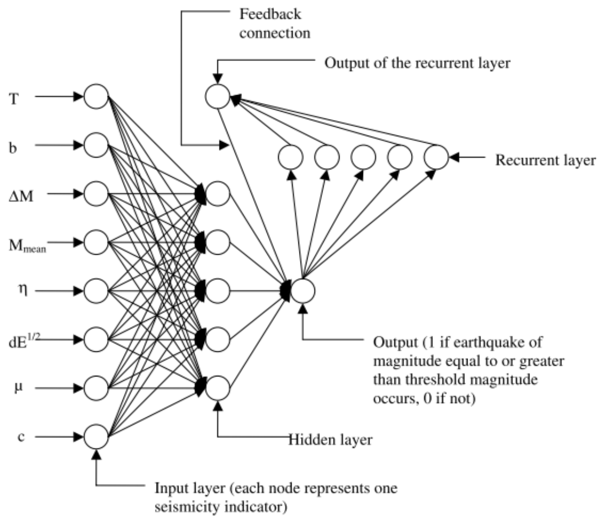

Fig. 3. Architecture of an RNN developed to predict large-magnitude earthquakes.

the errors in time accuracy ranged from 75 to 94 days for the main shock, and 5 to 16 days for aftershocks. The results obtained on the locality of the main-shock were not impressive; however, the time predictions for the aftershocks were reasonably accurate when compared to actual data records. After immense training, the RNN yielded a high $R-score$ of 1.0 for magnitudes greater than 6.5, and a modest $R-score$ of 0.4 for a magnitude of 4.5, leading to the conclusion that an RNN is ideal for high-seismic zones, but not as effective when used at a region with low seismicity. Furthermore, an analysis on RNN behavior suggests that the $R-score$ deteriorates when the time of prediction is less than 13 days; this indicates that an RNN is weak when predicting seismic events within a short time span. On the contrary, the $R-score$ is strong when the time of prediction has a window of two weeks. A detailed comparison between the predicted and actual values for the epicentral distance and time of occurrence can be found in [26].

In the same context, Nuannin explored seismic behaviors when trying to predict large main-shock events using the $b-value$ as a precursor [27]. The $b-value$ is a slope extracted from a historical law, and was designed specifically to predict aftershocks. A detailed description of this expression is provided in the following subsections. For a main shock, it is important to acknowledge that the $b-value$ yields the optimal results when the seismic event is from days to weeks in the future. This phenomenon shows that the RNN and $b-value$ may have similar features that have not been explored yet. As the $b-value$ is more lenient in predicting aftershocks, it may be a good idea to evaluate the $b-value$ using an RNN when predicting large-magnitude aftershocks.

Li et al. [7] proposed the use of a Genetic Algorithm (GA) as a training algorithm complimenting an NN when predicting a large earthquake within one year before the event. The network considered uses a large volume of historic metadata extracted from northern and southern China for roughly 28 years, along with six different seismic precursory indicators including the reciprocating frequency, the $b-value$ derived from the Gutenberg-Richter law, the average magnitude in the region of interest, the magnitude deficit, the rate of strain release, and the Mean Square Deviation (MSD), which is based on a Gutenberg-Richter logarithmic-frequency plot. For a reliable prediction, a categorical data analysis (AIC) was implemented to provide the best model by assessing its quality and accuracy of fit. The training and testing phases of the prediction were classified as reasonably accurate when compared to the recorded datasets, with $R-scores$ ranging from 0.02 to 0.55. Utsu [28] studied an earthquake size distribution, and suggested that the Gutenberg-Richter relation is ideal when the AIC value is low; this is true for earthquake applications unless a modified Gutenberg-Richter equation is implemented. A comprehensive exploration of the Gutenberg-Richter relationship can be found in their study.

In theory, the performance of an NN is highly dependent on the training process and hidden-layer design. Perez [29] showed that the GA has a self-searching method that shadows the concept of evolution. This means that a GA network has the ability to generate new values as possible solutions undetectable by humans. A comparison between the GA and BP training methods indicates the importance of the training data precision. The GA training methodology produces a 92% accuracy rating, which is substantially higher when compared to the BP training (which is only 72% accurate). With regard to the time consumption, the BP outperforms the GA training approach, providing 86% to 72% performance rating.

To complement this argument, Montana and Davis [30] used a GA to train an FF-NN, and compared the results to those using the BP learning algorithm. Their exploration on the subject suggests that the GA enhances its performance when constraining the domain to contain only filtered and valuable data. On the other hand, one major drawback of the BP is its poor performance when the network complexity and data volume increase. On the contrary, one advantage of the BP architecture is its ability to work if the network being dealt with has a simple training problem. As future work, it will be very interesting to remove the BP learning mechanism in favor of the GA, as discussed in [7], and evaluate the $R-score$ of the NN when predicting seismic magnitudes.

A follow-up on aftershock and main-shock estimations would be quite interesting; although they are categorized as different entities in nature, it is important to realize that they sustain similar properties. Esteban et al. [31] used an unsupervised clustering method called K-means, to obtain patterns and predict large earthquakes. The seismic input indicator taken into consideration is the slope from the Gutenberg-Richter law ($b-value$), which is analyzed as a network stress-meter. Predicting large earthquakes using the pioneered relationship between magnitude and total number of earthquakes, also known as the Gutenberg-Richter law, is a classical approach that is robust in nature [32]. Of course, this relationship is inexpensive and easier to comprehend, but it does have its flaws when the data obtained are non-linear, as is the case in reality, and hence this law can be made definitive if implemented as an input neuron. A major advantage of integrating the $b-value$ with an NN is an improvement in efficiency and overall accuracy, which is possible from extensive training and testing, which such a network is designed to withstand. The measure of quality of the results obtained from the NN yields a respective sensitivity value of 79.31% and 90%, and a respective specificity value of 90.38% and 90%, in particular Spanish-speaking areas. Insight into the K-means approach yields a good overall performance, especially when the uncertainty of an earthquake occurrence is taken into consideration. Schorlemmer et al. [33], on the other hand, showed that the $b-value$ is higher in normal fault zones as compared to thrust zones, which is unexpected because the thrust zones are higher in stress compared to normal zones. Nevertheless, an examination on this topic has revealed that the $b-value$ is in-fact a stress-meter inversely dependent on the differential stress.

With regard to this same topic, Htwe and WenBin [34] conducted a study on evaluating the Gutenberg-Richter recurrence law and analyzing the mean rate of large earthquakes occurring annually. The study involves a collection of database records from 1975 to 2006 around the Mayanmar area. In addition, the Gutenberg-Richter law contains a pre-defined magnitude threshold of 4.0 to 7.5 on the Richter scale. After a thorough analysis on this subject, the author concluded that the bounded Gutenberg-Richter law provides a smaller mean annual-rate of magnitude threshold, and a longer time frame for large seismic events. Theoretically, this means that increasing the maximum magnitude of an earthquake will require a dramatic decrease in the mean annual-rate of a low-magnitude threshold. In addition, a study by Aldo et al. [35] showed that a $b-value$ of $\cong 2$ indicates a volcanic region; on the other hand, a $b-value$ of $\cong 1$ represents a highly seismic area that is likely to produce a large-magnitude earthquake. Furthermore, other studies [35], [36] have suggested an error equation associated with the $b-value$, which could prove useful when implementing $b-values$ directly as an input parameter into the network. Such seismic laws have greatly influenced the outcome of a network, as demonstrated by Esteban et al. [31].

Reyes et al. [37] recently developed a ground-breaking method for predicting earthquakes in Chile by implementing a three-layered BP-ANN utilizing the $b-value$, Bath's law, and Omori-Utsu's law. The outcome of such a network is two-fold. First, it indicates the probability of an earthquake magnitude exceeding a pre-defined threshold. Second, the network will output a probable magnitude that may occur within the following five days. A comparison between different NNs such as an ANN, a KNN, an SVM, and K-means clustering showed that an ANN, a KNN, and K-means clustering produce better prediction accuracies. To validate the network used, 500 epochs were used in four different regions in Chile. A comparison between the NNs showed that the overall performance measured based on the specificity and sensitivity values is highly dependent on the location. For instance, within the region of Santiago, the ANN performed slightly better than the KNN and K-means clustering, leaving the SVM far behind in terms of prediction accuracy. In particular, the sensitivity was 35.7% for the ANN, 42.9% for the KNN, and 50% for the K-means clustering, with indeterminate results shown for the SVM. A further comparison of the NN performances in other regions within Chile can be found in [37].

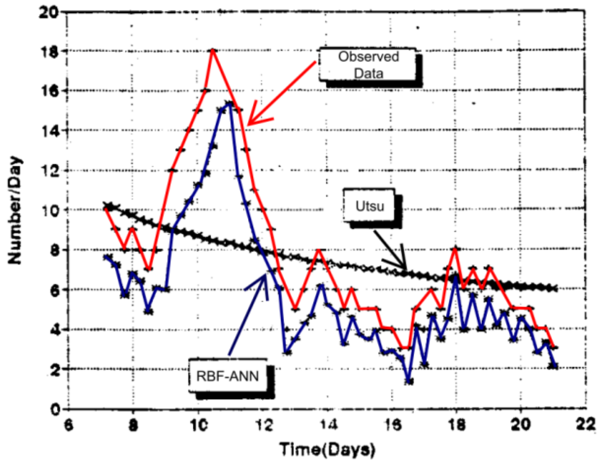

Fig. 4. A comparison between an RBF-NN and modified Omori function in predicting large aftershocks.

3.2. Predicting Large Aftershock Sequences

The occurrence of a large main-shock earthquake is typically found to trigger a secondary seismic shock, known as an aftershock. Aftershocks resemble the behavior of a decaying probability model, and are said to be just as dangerous as main-shock sequences [38]. The magnitude of an aftershock depends significantly on the time of occurrence and the magnitude of the main-shock.

Farahbod and Allamehzadeh [39] developed a RBF-NN used for prediction of large-magnitude aftershocks in eastern and central Iran. Their approach involves comparison of the modified Omori function with the self-learning RBF-NN to determine which concept is superior when predicting large aftershocks. The results showed that an RBF-NN provides a better prediction ability as compared to the modified Omori relation; however, incorporating a modified version of Omori's law as an input parameter, as suggested in [37, 40], was shown to deliver aftershock predictions with higher accuracy and reliability. **Fig. 4** [39] provides a graphical representation of the comparison between an RBF-NN and modified Omori function, and their accuracy against the observed data.

Similarly, Barrile et al. [40] used an RBF-NN to evaluate a series of Omori aftershock temporal aftershocks in Colfiorito, Italy. In this case, the network used consisted of a pre-defined magnitude threshold of greater than 7.0 on the Richter scale. In particular, the RBF-NN was considered over other NNs because an RBF has been scientifically recognized for its ability to resolve complex time-series functions accurately. A detailed schematic of the

RBF-NN layout used for this application can be found in [40]. For testing purposes, 153 datasets for magnitudes greater than 7 were collected from all over the world for the period of 1973 to 2004. An analysis of the network output showed that the probability of a large aftershock is 81% provided that it occurs within ten days after the main-shock. Moreover, the RMSE value obtained was 3.47%, which is reasonable. It is thereby conclusive that predicting an aftershock using an NN along with a modified Omori function is quite possible; however, this methodology does not support predicting the locality of an aftershock.

Wiemer et al. studied the properties of an aftershock by analyzing the 1999 Hector Mine aftershock sequence without considering any type of NN [41]. The findings of this phenomenon concluded that the earthquake size factor ($b-value$), aftershock decay rate ($p-value$), and seismic activity rate ($a-value$) played vital roles in the seismic event of the Hector Mine [41]. An analysis of this subject demonstrated that the $b-value$ is observed to be higher in the rupture zone and increases with respect to time for the first two months after the main shock. In addition, the highest aftershock reading was observed away from the ruptured main shock, suggesting that the transfer of stress from the main shock to the aftershock is static. Ranalli [42] examined a statistical study on the aftershock sequences using the Omori-Utsu law, the magnitude stability law, and the Gutenberg-Richter distribution. The results show that aftershocks are generally surface-oriented events, although magnitudes below the surface of the earth have been observed. It was also observed that using the least squares method to evaluate the Omori law delivers satisfactory results for uncorrelated values even when the theoretical conditions are not met. Using any of these techniques ($b-value$, $p-value$, $a-value$, or Omori-Utsu function) in an isolated fashion can provide seismologists and researchers with a rough estimate on impeding earthquakes. As mentioned before, these expressions are time-oriented and strictly follow the behavior of the function itself, which is not the case when predicting earthquakes in the real world. The need to resolve complex, non-linear, and unfriendly patterns demands a self-learning architecture, which is supported by NN strategies.

In conclusion, this subsection provided a detailed assessment of large-magnitude main-shock and aftershock predictions using different NN approaches. It is clear that an RNN is currently the ideal choice for predicting large-magnitude main-shocks. Although the accuracy of main-shock prediction has not always been at a consistently high level, an evaluation of this topic has illustrated promising results when incorporating a combination of seismic parameters into the proper type of NN. With regard to aftershock prediction, it is apparent that aftershocks are time-dependent in a functional format; in such cases, an RBF-NN is the ideal choice given its strengths and versatility in analyzing functional approximations and time-series prediction.

4. Detecting Radon Anomalies Using Artificial Neural Networks

Radon is a radioactive gas transmitted from layers beneath the earth's surface owing to rock micro-fracturing. It is colorless, odorless, and tasteless, making it virtually undetectable by humans [43]. Radon gas is present on land and in water bodies, and is released above the earth's surface gradually over time. This phenomenon has been observed by many authors in [44–48].

The concept of radon concentration as an influential precursor in predicting earthquakes has been brought to light by seismologists around the world. Observations led by Loomis [49] and King et al. [50] have shown that the radon concentration varies highly in different types of rock. It has also been shown that the radon concentrations in sedimentary rocks are relatively lower than those in granite or crystalline. As mentioned by Igarashi et al. [51], the presence of radon anomalies accumulated from cavities and cracks on the surface is a vital earthquake precursor. Radon release has been observed to promote intense degassing fluxes when the earth's crust is strained prior to a sudden slip from an earthquake. Ramola et al. [52] recently put forward valuable information on seismo-tectonic indicators on soil-gas radon. The results of their study suggest that radon concentrations may be affected by the magnitude, epicentral distance, precursory time, micro-structural changes, and focal depth of an earthquake, as well as non-linear functions including the type of rock, stress and strain, and the transport/diffusion measures within the local vicinity.

As an earthquake precursor, Gregoric et al. [53] conducted a study on radon concentration for detecting possible anomalies using NNs. The architecture of the utilized ANN consisted of five environmental input indicators, namely, the soil temperature, air temperature, soil air pressure, air pressure, and rainfall using data obtained in the Krsko basin site located in Slovenia for network testing and validation. To accurately identify the existence of radon anomalies caused by an increase in seismicity, four approaches were considered: the standard deviation related to the mean value, the relation between barometric pressure and radon concentration, an ANN using a BP learning approach, and decision tress. The results obtained from the performance tests show that using the standard deviation and barometric pressure as threshold values yields better accuracy than the measured-to-predicted ratio thresholds used by both an ANN and decision trees. However, the barometric pressure approach is believed to output many more false anomalies from environmental disturbances. A graphical comparison of these approaches can be found in [53].

Based on Gregoric et al. [53] results, Zmazek et al. [54] monitored the radon concentration in Krsko basin soil and found that its concentration is inversely proportional to the barometric pressure; hence, a decrease in barometric pressure causes an increase in the release of radon in soil-gas from the surface, and vice-versa. This relationship shows that an earthquake can be anticipated to occur

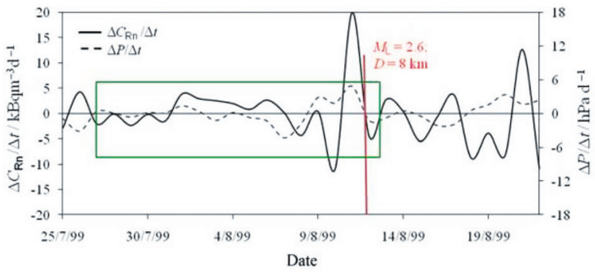

Fig. 5. Relationship between radon gas and barometric pressure during a seismically active period.

when the barometric pressure and soil-gas concentration are proportional; under this circumstance, the relation between the radon concentration and barometric pressure is considered a precursor. It is thereby evident that a traditional standard-deviation strategy, along with the barometric pressure, provides modest outcomes; owing to its simple mechanics, ease-of-use, and inexpensive nature, this is an effective approach to predicting radon anomalies. In this venue, incorporating these traditional tactics with NNs is a fair proposition for further enhancing the certainty of prediction. **Fig. 5** shows the relationship between barometric pressure and radon gas release during a seismically active period [53].

In 2005, Zmazek et al. [55] proposed a related approach to identify radon anomalies in soil gas; this approach includes the input of environmental parameters using decision trees and an MLP-type ANN combined with a conjugate gradient learning algorithm. The proposed network was trained twice using over 45,000 pre-defined epochs, and the results obtained show that ten random anomalies were detected for 12 earthquakes using the ANN, giving a performance rating of 83%. On the other hand, decision trees found radon anomalies for every earthquake, even during those periods where no seismicity occurred. Observations on the Pearson's correlational coefficient from the test results convey that, for seismically active data, the radon concentration fluctuates drastically seven days before or after a seismic event. Arguably, both Zmazek [55] and Gregoric [53] agree that during a seismically active period, the measured and predicted ratios are significantly weaker; however, during a non-seismically active period, the measured and predicted threshold decisions work well. The evaluation results of these approaches show that the radon anomalies are likely caused from seismic phenomena and not entirely by environmental parameters. It is now clear that radon anomalies can naturally be false positives incurred from environmental misconceptions, and that this differentiation should be resolved when using an NN strategy.

Torkar et al. [56] introduced another ANN-BP application to simulate the Radon concentration trends in soil-gas from three boreholes at the Orilica fault zone. For the network architecture, the same input parameters from [53] were used; however, five different types of thresholds were utilized, namely, the total time of an anomaly, the seismic activity period, the separation of anomalies, the anomaly integral, and once again, the measured-to-predicted threshold value. The results show that the percent of specificity of the soil temperature and air pressure increases ten days before and after the occurrence of a radon anomaly. Research on this topic shows an ANN as the ideal choice to predict the radon level in soil-gas as compared to other data mining techniques; in particular, the proposed prediction strategy was successful in ten out of 13 cases (77%). In addition, comparative threshold results show that the ANN measured-to-predicted ratio produces a high $R-score$ of 0.95 for the training set, 0.78 for the cross-validation, and 0.79 for the testing set. These performance results are more accurate in comparison to decision trees and regression models, which predict the radon concentration with an $R-score$ of 0.83 during an NSA and 0.69 during an SA period. Furthermore, Planinic et al. [57] studied temporal variations of radon in soil and found that the temperature variation only influences the radon concentration in well water; this indicates that temperature variations are not as heavily weighted as soil parameters.

Xia-Ting and Masahiro developed an ANN-BP that predicts rock micro-fracturing when under tri-axial compressive stress [58]. Their network modeling consists of Acoustic Emissions (AEs) and time distributions as the primary input parameters for characterizing the rock micro-fracturing. AE is specifically utilized for determination of the internal changes in an object under stress using sound waves, and is hence an ideal component when analyzing minor changes in environmental behaviors. Over the years, researchers have attempted to adapt AE models with non-linear characteristics. However, through various tests and observations, such models are unable to resemble to actual radon data because the datasets are non-linear in nature. Therefore, the use of a prediction tool, such as an NN, may significantly increase the prediction ability when complying with the random, dynamic, and non-linear data produced by rock micro-fractures. The results extracted from AE event patterns before and after a failure show that the MSE values change from 0.035 to 0.0419, and that an error in an extrapolated prediction changes from 0.223 to 0.149, respectively. These performance results obtained are certainly not the best, nor are they the most accurate; however, the idea behind the prediction of rock micro-fracturing, as an NN application, is the increased likelihood of radon emission detection. This topic should be investigated in greater detail because the prediction of rock micro-fracturing may act as a threshold, filtering the differences between environmental issues and actual radon anomalies detected from a seismic event.

Negarestani et al. [59] brought forward a creative approach for the prediction of radon concentration in soil using environmental parameters, and assessing the soil using a Layered NN Back-Propagation (LNN-BP) architecture. This network consists of data analytics obtained from a site in Thailand that are roughly 40 weeks old and trained using 400 iterations. The input neurons evaluated by the NN are the soil temperature, soil pressure, rainfall, and weekly radon concentration data obtained at 100 and 50 cm below the surface. An investigation on this subject showed that the best-fit between the estimated and measured time-series for radon recognition is between the 15th and 18th weeks. The authors also concluded that the relationship between rainfall and radon concentration is not linear; however, using this type of network, the radon concentration formed by a seismic event can be differentiated from false radon detection. A detailed graphical comparison between the NN prediction and measured radon concentration can be found in [59]. The evidence of rainfall and its effect on radon anomalies strongly agree with the findings of Pinault and Baubron [60]. They showed that the kinetics of water in soil causes a reaction that gently opens the cracks in the rock precisely 11 days after a rainfall. A comprehensive representation of the signal processing of soil radon can be found in [60].

Negarestani et al. [61] later introduced a similar method for predicting the radon concentration in soil using environmental variables, however, this time using an Adaptive Linear Neural Network (Adaline). For the training phase, 263 iterations of the input parameters, similar to the study in [59], were taken into consideration. The test results of the Adaline network show the same time-series as in [59]. Although the testing phase reveals a similar time-series estimate for Radon detection, a closer comparison of Adaline and LNN affirms that an Adaline NN is a better fit for predicting the Radon concentration owing to the fact that it resembles the actual data more vividly. A graphical representation of an Adaline NN performance can be found in [61]. In addition, Negarestani also stated that the proposed radon prediction method using Adaline is considerably more accurate, easier to use, and faster in terms of application processing than other methods.

Gupta and Shahani [62] projected an RBF-type ANN approach in combination with a time-reliant Fast Fourier Transformation (FFT) for predicting radon emissions. An RBF network was explicitly chosen over an MLP-ANN for two reasons. First, an RBF can model any complex non-linear equation using a single hidden layer, thereby increasing the overall performance. Second, a single-layered transformation located in the output neuron is optimized using a linear modeling strategy to overcome the local minima problems encountered by an MLP-ANN. The network evaluation includes database records from 1994 to 1996, as well as the temperature, pressure, wind velocity, rainfall, and humidity from the soil-gas as input parameters. Testing of this methodology resulted in a significant 87.8% prediction rate, which could be increased if the number of false alarms is reduced. In addition, it was observed that the barometric pressure and rainfall parameters heavily influence the radon emission levels. Moreover, optimal results were obtained from the system when using the mean and standard deviation thresholds, which were processed using an FFT to train the network.

In brief, the detection of radon emissions as an earthquake anomaly is possible when using an FFBP-ANN; however, additional work needs to be carried out on this subject to accurately predict the radon concentration level. Toutain and Baubron [63] and Ramola [64] studied gas-geochemistry and its relation to seismological events. One observation of their studies showed the advisability of recording other gases such as helium, carbon-dioxide, and chlorine along with the radon emissions when conducting an earthquake-like prediction. Additionally, Toutain and Baubron [63] suggested combining various other chemical and physical parameters such as water chemistry, water temperature and level, tilt, V_p/V_s ratio, low-level seismicity, and electric conductivity to increase the accuracy of Radon prediction. Therefore, it appears that the prediction of a Radon anomaly is quite possible when using an FFBP-ANN architecture.

5. Results and Discussion

Table 1 below shows a comprehensive overview of the NNs discussed in this paper, along with the training mechanism used and the respective authors. An FF-ANN, along with a BP learning methodology, is the most common type of network used to predict earthquake precursors because this network architecture is simple, reliable, and can adapt to various types of data for providing a valid prediction. However, certain networks have been observed to perform better when embedded with specific types of precursors; for instance, a PGA and radon anomaly prediction yield better accuracy when paired with an FFBP-ANN methodology, a prediction of the potential liquefaction provides optimal results when evaluated using a PNN, large-magnitude main-shocks are better predicted using an RNN, and aftershock sequences are more accurate when evaluated using an RBF-NN. To summarize, each of the precursors mentioned above can be depicted using an NN. Since this type of network is designed to deal with real-world (non-linear) data types, embedding precursory data as input neurons can provide seismologists and researchers with accurate and concise prediction models.

6. Conclusion

As this paper has shown, most earthquake anomalies that are currently detected by other systems can be covered through ANNs, making the outcome more precise and accurate. The goal of this paper was to present NN applications through the integration of earthquake anomalies, namely, a PGA, the liquefaction potential, large magnitudes of main-shock and aftershock events, and finally, predicting radon anomalies as precursors. When using a network approach, it is vital to collect adequate seismic data on the particular region of interest, as a greater amount of data will make the predictions more accurate. With reference to the seismicity, NNs are utilized as an important aid in classifying seismic windows and constraints, and in predicting earthquake precursors. This methodology demonstrates the value of adding expert knowledge to machine learning algorithms, and provides

Table 1. Recapitation of the precursors and types of NN/learning methodology used for predicting seismic events.

Researcher	Year	Analysis Topic	Seismic Precursor	Feed Forward Artificial Neural Network	Recurrent Neural Network	K-means Clustering	K-nearest Neighbor Technique	Support Vector Machines	Radial Basis Function	Adaptive Linear Neural Network	Probability Neural Network	Generalized Regression Neural Network	Training Technique
Derras and Bekkouche	2011	Ground Motion Prediction	PGA	×									BP
Kemal and Ayten	2008			×					×			×	
Kerh and Chu	2002			×									
Derras et al.	2013		Liquefaction Potential	×									Retro-Propagation
Bojorquez et al.	2012			×									BP
Goh	2002										×		Bayesian classifier
Kumar et al.	-			×									BP
Panakkat and Adeli	2007	Large Magnitude Prediction	Main-shock	×	×				×				Levenberg-Marquardt BP
Adeli and Panakkat	2009				×						×		-
Panakkat and Adeli	2009				×								Levenberg-Marquardt
Li et al.	1998			×									GA
Montana and Davis	1989			×									GA and BP
Esteban et al.	2010					×							-
Reyes et al	2012			×		×	×	×					BP
Farahbod and Allamehzadeh	1999		Aftershock						×				-
Barrile et al.	2006								×				-
Gregoric et al.	2011	Radon Anomalies		×									BP
Zmazek et al.	-			×									Conjugate gradient
Torkar et al.	2010			×									BP
Xia-Ting and Masahiro	1998			×									BP
Negarestani et al.	2002			×									BP
Negarestani et al.	2003									×			-
Gupta and Shahani	2011								×				-

evidence showing that this knowledge may further increase the accuracy of the machine learner.

The studies discussed in this paper show clearly that certain networks in collaboration with a specific precursor yield more accurate results. An FFBP-ANN is the best choice when predicting the ground motion and radon concentration. Liquefaction prediction is better examined using a PNN. An RNN is more appropriate for the prediction of large-magnitude main-shocks; and finally, an RBF-NN is a suitable choice for a functional approach such as predicting aftershocks. Each of these NNs has its merits and demerits; however, they each excel in terms of accuracy and performance for their respective precursory applications.

As future research on AI under the application of seismicity, the following topics should be evaluated: the Seismic Gap Theory to measure the seismic energy release, IR thermal anomalies to detect temperature variations (hotspots) on the surface of the earth, and finally, Seismic Electric Signals (SES) to determine the transient changes in the electrotelluric field emitted by rocks under stress. In addition, an integration of the known proportionality into an NN is another area of work that may increase the accuracy given larger training and testing sets. In conclusion, a combination of different precursors with the implementation of a NN methodology cannot make earthquake predictions 100% accurate, but it can definitely improve, approaching an acceptable level of accuracy.

References:

[1] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," Computer, Vol.29, No.3, pp. 31-44, 1996.

[2] M. Negnevitsky, "Artificial intelligence: a guide to intelligent systems," Pearson Education, 2005.

[3] K. Wang, Q. F. Chen, S. Sun, and A. Wang, "Predicting the 1975 Haicheng earthquake," Bulletin of the Seismological Society of America, Vol.96, No.3, pp. 757-795, 2006.

[4] C. H. Scholz, "A physical interpretation of the Haicheng earthquake prediction," Nature, Vol.267, No.5607, pp. 121-124, 1977.

[5] F. Zhu and G. Wu, "Haicheng earthquake in 1975," 1982.

[6] A. Jin and K. Aki, "Temporal change in coda q before the Tangshan earthquake of 1976 and the Haicheng earthquake of 1975," J. of Geophysical Research: Solid Earth (1978-2012), Vol.91, No.B1, pp. 665-673, 1986.

[7] M. Li, Z. Lieyuan, and S. Yaolin, "Attempts at using seismicity indicators for the prediction of large earthquakes by genetic algorithm-neural network method," 1998.

[8] P. Lussou, P. Y. Bard, F. Cotton, and Y. Fukushima, "Seismic design regulation codes: contribution of k-net data to site effect evaluation," J. of Earthquake Engineering, Vol.5, No.1, pp. 13-33, 2001.

[9] I. M. Idriss, "Characteristics of earthquake ground motions," Proc. of the ASCE Geotechnical Engineering Division Speciality Conf.: Earthquake Engineering and Soil Dynamics, Vol.3, pp. 1151-1265, 1978.

[10] D. M. Boore and W. B. Joyner, "The empirical prediction of ground motion," Bulletin of the Seismological Society of America, Vol.72, No.6B, pp. 43-60, 1982.

[11] K. W. Campbell, "Strong motion attenuation relations: a ten-year perspective," Earthquake Spectra, Vol.1, No.4, pp. 759-804, 1985.

[12] J. Douglas, "Consistency of ground-motion predictions from the past four decades: peak ground velocity and displacement, arias intensity and relative significant duration," Bulletin of Earthquake Engineering, Vol.10, No.5, pp. 1339-1356, 2012.

[13] C. C. Lin and J. Ghaboussi, "Recent progress on neural network based methodology for generating artificial earthquake accelerograms," ICSSD 2000: 1st Structural Conf. on Structural Stability and Dynamics, pp. 793-798, 2000.

[14] A. T. Goh, "Probabilistic neural network for evaluating seismic liquefaction potential," Canadian Geotechnical J., Vol.39, No.1, pp. 219-232, 2002.

[15] B. Derras and A. Bekkouche, "Use of the artificial neural network for peak ground acceleration estimation," Lebanese Science J., Vol.12, No.2, p. 101, 2011.

[16] J. A. Meredith, R. H. Wilkens, and C. H. Cheng, "Evaluation and prediction of shear wave velocities in soft marine sediments," Technical report, Massachusetts Institute of Technology, Earth Resources Laboratory, 1989.

[17] K. Günaydn and A. Günaydn, "Peak ground acceleration prediction by artificial neural networks for northwestern turkey," Mathematical Problems in Engineering, 2008.

[18] N. N. Ambraseys and J. Douglas, "Near-field horizontal and vertical earthquake ground motions," Soil dynamics and earthquake engineering, Vol.23, No.1, pp. 1-18, 2003.

[19] T. Kerh and D. Chu, "Neural networks approach and microtremor measurements in estimating peak ground acceleration due to strong motion," Advances in Engineering Software, Vol.33, No.11, pp. 733-742, 2002.

[20] T. Kerh and S. B. Ting, "Neural network estimation of ground peak acceleration at stations along Taiwan high-speed rail system," Engineering Applications of Artificial Intelligence, Vol.18, No.7, pp. 857-866, 2005.

[21] B. Derras, A. Bekkouche, and D. Zendagui, "Neuronal approach and the use of kik-net network to generate response spectrum on the surface," 2013.

[22] E. Bojórquez, J. Bojórquez, S. E. Ruiz, and A. Reyes-Salazar, "Prediction of inelastic response spectra using artificial neural networks," Mathematical Problems in Engineering, 2012.

[23] V. Kumar, K. Venkatesh, and R. P. Tiwari, "Application of ANN to predict liquefaction potential."

[24] A. Panakkat and H. Adeli, "Neural network models for earthquake magnitude prediction using multiple seismicity indicators," Int. j. of neural systems, Vol.17, No.1, pp. 13-33, 2007.

[25] H. Adeli and A. Panakkat, "A probabilistic neural network for earthquake magnitude prediction," Neural Networks, Vol.22, No.7, pp. 1018-1024, 2009.

[26] A. Panakkat and H. Adeli, "Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators," Computer-Aided Civil and Infrastructure Engineering, Vol.24, No.4, pp. 280-292, 2009.

[27] P. Nuannin, "The potential of b-value variations as earthquake precursors for small and large events," Ph.D. thesis, Uppsala University, 2006.

[28] T. Utsu, "Representation and analysis of the earthquake size distribution: a historical review and some new approaches," Seismicity Patterns, their Statistical Significance and Physical Meaning, Springer, pp. 509-535, 1999.

[29] S. Perez, "Apply genetic algorithm to the learning phase of a neural network."

[30] D. J. Montana and L. Davis, "Training feedforward neural networks using genetic algorithms," IJCAI, Vol.89, pp. 762-767, 1989.

[31] A. M. Esteban, F. M. Álvarez, A. Troncoso, J. L. Justo, and C. R. Escudero, "Pattern recognition to forecast seismic time series," Expert Systems with Applications, Vol.37, No.12, pp. 8333-8342, 2010.

[32] C. F. Richter, "An instrumental earthquake magnitude scale," Bull. Seism. Soc. Am, Vol.25, No.1, pp. 1-32, 1935.

[33] D. Schorlemmer, S. Wiemer, and M. Wyss, "Variations in earthquake-size distribution across different stress regimes," Nature, Vol.437, No.7058, pp. 539-542, 2005.

[34] Y. M. Htwe and S. WenBin, "Gutenberg-Richter recurrence law to seismicity analysis of southern segment of the sagaing fault and its associate components."

[35] A. Zollo, W. Marzocchi, P. Capuano, A. Lomax, and G. Iannaccone, "Space and time behavior of seismic activity at Mt. Vesuvius volcano, Southern Italy," Bulletin of the Seismological Society of America, Vol.92, No.2, pp. 625-640, 2002.

[36] Y. Shi and B. A. Bolt, "The standard error of the magnitude-frequency b value," Bulletin of the Seismological Society of America, Vol.72, No.5, pp. 1677-1687, 1982.

[37] J. Reyes, A. M. Esteban, and F. M. Álvarez, "Neural networks to predict earthquakes in Chile," Applied Soft Computing, 2012.

[38] S. Hainzl and D. Marsan, "Dependence of the Omori-Utsu law parameters on main shock magnitude: Observations and modeling," J. of Geophysical Research: Solid Earth (1978-2012), Vol.113, No.B10, 2008.

[39] A. M. Farahbod and M. Allamehzadeh, "Large aftershocks prediction results in eastern and central Iran using artificial neural networks (ANN's)," 3rd Int. Conf. on Seismology and Earthquake Engineering, 1999.

[40] V. Barrile, M. Cacciola, S. D'Amico, A. Greco, F. C. Morabito, and F. Parrillo, "Radial basis function neural networks to foresee aftershocks in seismic sequences related to large earthquakes," Neural Information Processing, Springer, pp. 909-916, 2006.

[41] S. Wiemer, M. Gerstenberger, and E. Hauksson, "Properties of the aftershock sequence of the 1999 mw 7.1 hector mine earthquake: Implications for aftershock hazard," Bulletin of the Seismological Society of America, Vol.92, No.4, pp. 1227-1240, 2002.

[42] G. Ranalli, "A statistical study of aftershock sequences," Annals of Geophysics, Vol.53, No.1, pp. 43-58, 2010.

[43] R. F. Holub and B. T. Brady, "The effect of stress on Radon emanation from rock," J. of Geophysical Research: Solid Earth (1978-2012), Vol.86, No.B3, pp. 1776-1784, 1981.

[44] V. I. Ulomov, A. I. Zakharova, and N. V. Ulomova, "Tashkent earthquake of April 26, 1966, and its aftershocks," Akad. Nauk. SSSR, Geophysic, Vol.177, pp. 567-570, 1967.

[45] M. Noguchi and H. Wakita, "A method for continuous measurement of Radon in groundwater for earthquake prediction," J. of Geophysical Research, Vol.82, No.8, pp. 1353-1357, 1977.

[46] C. Y. King, "Episodic Radon changes in subsurface soil gas along active faults and possible relation to earthquakes," J. of geophysical research, Vol.85, No.B6, pp. 3065-3078, 1980.

[47] R. C. Ramola, M. Singh, A. S. Sandhu, S. Singh, and H. S. Virk, "The use of Radon as an earthquake precursor," The Int. j. of radiation applications and instrumentation. Nuclear geophysics, Vol.4, No.2, pp. 275-287, 1990.

[48] B. Singh and H. S. Virk, "Investigation of Radon-222 in soil-gas as an earthquake precursor," The Int. j. of radiation applications and instrumentation. Part E. Nuclear geophysics, Vol.8, No.2, pp. 185-193, 1994.

[49] D. P. Loomis, "Radon-222 concentration and aquifer lithology in north Carolina," Ground Water Monitoring & Remediation, Vol.7, No.2, pp. 33-39, 1987.

[50] P. T. King, J. Michel, and W. S. Moore, "Ground water geochemistry of 228 Ra, 226 Ra and 222 Rn," Geochimica et Cosmochimica Acta, Vol.46, pp. 1173-1182, 1982.

[51] G. Igarashi, S. Saeki, N. Takahata, K. Sumikawa, S. Tasaka, Y. Sasaki, M. Takahashi, and Y. Sano, "Ground-water Radon anomaly before the Kobe earthquake in japan," Science-New York Then Washington, pp. 60-60, 1995.

[52] R. C. Ramola, Y. Prasad, G. Prasad, S. Kumar, and V. M. Choubey, "Soil-gas Radon as seismotectonic indicator in Garhwal Himalaya," Applied Radiation and Isotopes, Vol.66, No.10, pp. 1523-1530, 2008.

[53] A. Gregoric, B. Zmazek, S. Džeroski, D. Torkar, and J. Vaupotic, "Radon as earthquake precursor-methods for detecting anomalies," Earthquake Research and Analysis, 2011.

[54] B. Zmazek, M. Živcic, L. Todorovski, S. Džeroski, J. Vaupotic, and I. Kobal, "Radon in soil gas: how to identify anomalies caused by earthquakes," Applied geochemistry, Vol.20, No.6, pp. 1106-1119, 2005.

[55] B. Zmazek, S. Dzeroski, D. Torkar, J. Vaupotic, and I. Kobal, "Identification of Radon anomalies in soil gas using decision trees and neural networks."

[56] D. Torkar, B. Zmazek, J. Vaupotic, and I. Kobal, "Application of artificial neural networks in simulating Radon levels in soil gas," Chemical Geology, Vol.270, No.1, pp. 1-8, 2010.

[57] J. Planinic, V. Radolic, and Ž. Lazanin, "Temporal variations of Radon in soil related to earthquakes," Applied Radiation and isotopes, Vol.55, No.2, pp. 267-272, 2001.

[58] X. T. Feng and M. Seto, "Neural network dynamic modelling of rock microfracturing sequences under triaxial compressive stress conditions," Tectonophysics, Vol.292, No.3, pp. 293-309, 1998.

[59] A. Negarestani, S. Setayeshi, M. G. Maragheh, and B. Akashe, "Layered neural networks based analysis of Radon concentration and environmental parameters in earthquake prediction," J. of environmental radioactivity, Vol.62, No.3, pp. 225-233, 2002.

[60] J. L. Pinault and J. C. Baubron, "Signal processing of soil gas Radon, atmospheric pressure, moisture, and soil temperature data: A new approach for Radon concentration modeling," J. of geophysical research, Vol.101, No.B2, pp. 3157-3171, 1996.

[61] A. Negarestani, S. Setayeshi, M. G. Maragheh, and B. Akashe, "Estimation of the Radon concentration in soil related to the environmental parameters by a modified Adaline neural network," Applied radiation and isotopes, Vol.58, No.2, pp. 269-273, 2003.

[62] D. Gupta and D. T. Shahani, "Estimation of Radon as an earthquake precursor: A neural network approach," J. of the Geological Society of India, Vol.78, No.3, pp. 243-248, 2011.

[63] J. P. Toutain and J. C. Baubron, "Gas geochemistry and seismotectonics: a review," Tectonophysics, Vol.304, No.1, pp. 1-27, 1999.

[64] R. C. Ramola, "Relation between spring water Radon anomalies and seismic activity in Garhwal Himalaya," Acta Geophysica, Vol.58, No.5, pp. 814-827, 2010.

Name:
Aditya Sriram

Affiliation:
Undergraduate Student Researcher, University of Ontario Institute of Technology (UOIT)

Address:
2000 Simcoe Street North, Oshawa, Ontario L1H 7K4, Canada

Brief Biographical History:
2010- Admitted to Bachelor of Mechatronics Engineering at UOIT
2012- Began Conducting Research on applications under Artificial Intelligence with Dr. Shahryar Rahnamayan
2013- Joined IBM Toronto Software Laboratory as a DB2 LUW QA and Developer Intern

Membership in Academic Societies:
• American Society of Mechanical Engineers (ASME)

Name:
Shahryar Rahanamayan

Affiliation:
University of Ontario Institute of Technology (UOIT)

Address:
2000 Simcoe Street North, Oshawa, Ontario L1H 7K4, Canada

Brief Biographical History:
1998 Received the B.Sc. Degrees in software engineering
2001 Received the M.Sc. Degrees in software engineering
2007 Received the Ph.D. Degree in the field of evolutionary computation from the University of Waterloo, Waterloo, Canada
2007.8- Chief Research Manager at OMISA Inc. (Omni-Modality Intelligent Segmentation Assistant)
2008 Postdoctoral Fellow at Simon Fraser University, Vancouver, Canada
2008 Joining as a Core Faculty Member the Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Oshawa, Canada

Main Works:
• Metaheuristics, evolutionary computation, large-scale optimization, image processing, computer visio
• Received several prestigious awards such as the Ontario Graduate Scholarship, Presidents Graduate Scholarship, Natural Sciences and Engineering Research Council of Canada (NSERC)s Japan Society for the Promotion of Science Fellowship, NSERCs Industrial R&D Fellowship, NSERCs Visiting Fellowship in Canadian Government Laboratories, the Canadian Institute of Health Research Fellowship for two times, NSERC Discovery Grant, 2009

Name:
Farid Bourennani

Affiliation:
Lecturer, University of Ontario Institute of Technology (UOIT)

Address:
2000 Simcoe Street North, Oshawa, Ontario L1H 7K4, Canada

Brief Biographical History:
2002 Completed B.Sc. in Business Informatics
2009 Completed M.A.Sc. (Electrical & Computer Eng.) in Heterogeneous Data Mining
2013 Completed the Ph.D. (Electrical & Computer Eng.) in Multi-Objective Optimization

Main Works:
• Multi-objective optimization, optimal design of energy systems, applied data mining, heterogeneous data mining
• Twice and recipient the Jeffrey Boyce Engineering Award which is the highest and most prestigious award at UOIT
• Nominated for Ph.D. Alumni Gold Medal with a GPA of 4.3/4.3
• Recipient of Several Prestigious Awards such as Ontario Graduate Scholarship (OGS), Dean Scholarship, Graduate Excellence Scholarship, and Best Paper Award at the DBKDA'09 Conf.