Computer-Aided Civil and Infrastructure Engineering **24** (2009) 280–292

Recurrent Neural Network for Approximate Earthquake Time and Location Prediction Using Multiple Seismicity Indicators

Ashif Panakkat*

M-Engineering, Inc., 10 W. Hubbard Street, Suite 3N, Chicago, IL 60610, USA

&

Hojjat Adeli

Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, 470 Hitchcock Hall, 2070 Neil Avenue, Columbus, OH 43210, USA

Abstract: *A computational approach is presented for predicting the location and time of occurrence of future moderate-to-large earthquakes in an approximate sense based on neural network modeling and using a vector of eight seismicity indicators as input. Two different methods are explored. In the first method, a large seismic region is subdivided into several small subregions and the temporal historical earthquake record is divided into a number of small equal time periods. Seismicity indicators are computed for each subregion for each time period and their relationship to the magnitude of the largest earthquake occurring in that subregion during the following time-period is studied using a recurrent neural network. In the second more direct approach, the temporal historical earthquake record is divided into a number of unequal time periods where each period is defined as the time between large earthquakes. Seismicity indicators are computed for each time-period and their relationship to the latitude and longitude of the epicentral location, and time of occurrence of the following major earthquake is studied using a recurrent neural network.*

*To whom correspondence should be addressed. E-mail: apanakkat@m-corporations.com.

1 INTRODUCTION

Most efforts in earthquake parameter prediction focus on monitoring the overall activity in a large seismic region (Varotsos and Alexopoulos, 1984; Keilis-Borok et al., 1990; Kagan and Jackson, 1991; Agnew and Jones, 1991; Ma et al., 1999; Poitrasson et al., 1999; Dey and Singh, 2003; Rydelek and Pujol, 2004; Sharma and Arora, 2005). In such studies, the time of occurrence and epicentral location are predicted (in most cases indirectly) only in the range of hundreds of miles in location and a few years in time. A review of recent efforts in earthquake prediction is presented in Panakkat and Adeli (2008).

In the past two decades neural networks have been used successfully to solve complicated pattern recognition and classification problems in many fields such as image analysis and recognition (Adeli and Hung, 1993, 1994; Hung and Adeli, 1993, 1994a; Raftopoulos et al., 2007; Rigatos and Tzafestas, 2007; Gopych, 2008; Khashman and Sekeroglu, 2008), speech recognition (Yau et al., 2007), robotics and computer vision (Sabourin et al., 2007; Taylor et al., 2007; Villaverde et al., 2007; Jorgensen et al., 2008), natural language

DOI: 10.1111/j.1467-8667.2009.00595.x

processing (Ruiz-Pinales et al., 2008), biomedical engineering and medical diagnosis (Adeli et al., 2005, 2008; Ghosh-Dastidar et al., 2007; Gil-Pita and Yao, 2008; Alexandre et al., 2008; Huynh et al., 2008; Khashman, 2008; Nemissi et al., 2008), neuroscience (Ghosh-Dastidar and Adeli, 2007; Iglesias and Villa, 2008), construction engineering (Adeli and Wu, 1998; Senouci and Adeli, 2001), environmental engineering (Adeli, 2001), structural engineering (Hung and Adeli, 1994b; Adeli and Park, 1995a,b,c, 1996; Park and Adeli, 1995, 1997; Adeli and Karim, 1997a,b; Sirca and Adeli, 2001, 2005; Adeli and Kim, 2001; Tashakori and Adeli, 2002; Ahmadkhanlou and Adeli, 2005; Jiang and Adeli, 2005a, 2007, 2008a,b; Adeli and Jiang, 2006), transportation engineering (Adeli and Samant, 2000; Adeli and Karim, 2000; Samant and Adeli, 2001; Karim and Adeli, 2002; Ghosh-Dastidar and Adeli, 2003, 2006; Adeli and Jiang, 2003; Dharia and Adeli, 2003; Hooshdar and Adeli, 2004; Jiang and Adeli, 2005b; Cyganek, 2007, 2008), video and audio analysis (Fyfe et al., 2008), computer networking (Kimura and Ikeguchi, 2007), control (Rigatos, 2008; Liu and Zhang, 2008), computer security (Neumann et al., 2007), communication analysis (Chen et al., 2008), reservoir engineering (Banchs et al., 2007), air traffic control (Christodoulou and Kontogeorgou, 2008), solar activity (Ni and Yin, 2008), and financial forecasting (Schneidr and Graupe, 2008). If neural networks can be used for so many different complicated and seemingly intractable problems, why not earthquake prediction?

Recently the authors investigated the application of three multilayered neural fnetworks, a feed-forward Levenberg–Marquardt backpropagation (LMBP) neural network, a recurrent neural network, and a radial basis function (RBF) neural network, for predicting the magnitude of the largest seismic event in the following month based on the analysis of eight mathematically computed parameters known as seismicity indicators (Panakkat and Adeli, 2007). The indicators are selected based on the Gutenberg–Richter (Gutenberg and Richter, 1956) and characteristic earthquake magnitude distributions and also on the conclusions drawn by recent earthquake prediction studies. They are the time elapsed during a particular number (say n) of significant seismic events before the month in question (T), the slope of the Gutenberg–Richter inverse power law curve for the n events (b), the mean square deviation about the regression line based on the Gutenberg–Richter inverse power law for the n events (η), the average magnitude of the last n events (M_{mean}), the difference between the observed maximum magnitude among the last n events and that expected through the Gutenberg–Richter relationship known as the magnitude deficit (ΔM), the rate of square root of seismic energy released during time T ($\mathrm{d}E^{1/2}$), the mean time or period between characteristic events (μ value), and the coefficient of variation of the mean time (c).

Prediction accuracies of the models were evaluated using four different statistical measures: the probability of detection (POD), the false alarm ratio (FAR), the frequency bias, and the true skill score or R score (Panakkat and Adeli, 2007). The models were trained and tested using data from two different seismic regions: Southern California and the San Francisco Bay region. The study showed that in the great majority of the cases the recurrent neural network model yielded the best prediction accuracies compared with the BP and RBF networks. However, even the recurrent neural network did not predict moderate (with Richter magnitude between 5 and 6.5) to large (with magnitude 6.5 or greater) earthquakes in about 10% of the cases and falsely predicted a moderate-to-large earthquake in 15% of the cases.

How can we improve the earthquake prediction accuracy of the recurrent neural network? This question is studied in the first part of the current article by (1) reducing the prediction time frame from one month to 15 days and one week and (2) dividing the seismic region into smaller areas and performing a parametric study. This exercise of reducing the time and space considered for the study can result in prediction of the time of occurrence and epicentral location of the earthquake, a subject not studied by the authors previously, but of great importance for emergency management agencies.

In the second part of this article, a more direct method for the prediction of epicentral location and time of occurrence is presented based on several well-known theories in geophysics such as the elastic rebound theory (Reid, 1910), the gap theory (Nishenko, 1991), and the characteristic earthquake distribution theory (Kagan and Jackson, 1991). These theories are based on the same basic hypothesis that the earthquake occurrence process in many seismically active regions follows a temporal pattern where large or major earthquakes are punctuated by approximately equal time intervals. In the second part of the article, the eight seismicity indicators are computed for a large seismic region (such as the entire southern California) using significant events (defined as those of a predefined threshold magnitude or greater) occurring between two major earthquakes (defined as having a certain minimum magnitude). The values of the indicators are used as input in a recurrent neural network model to predict the epicentral location and time of occurrence of the succeeding major earthquake in the region.

Historical seismic data recorded in southern California dating back to 1930 are archived by the Southern

California Earthquake Data Center (SCEC) and are available for free download through the center's website at www.data.scec.org. Data obtained from the SCEC website are used to train and test the neural network model developed in this research.

2 IMPROVING PREDICTION ACCURACY OF THE RECURRENT NEURAL NETWORK MODEL

2.1 Seismic subregions

The recurrent neural network for earthquake magnitude prediction presented in Panakkat and Adeli (2007) is applied to smaller seismic areas in the Southern California region defined by the Working Group on California Earthquake Probabilities (2004) as an area between 114.75°W and 119.25°W longitude and 33.8°N and 35.4°N latitude. This region is divided into six small subregions each separated by 1.5° longitude and 0.8° latitude representing a surface area of less than 30 square miles each. The subregions are denoted by SC1 to SC6, as indicated in Figure 1, and their geographic coordinates are given in Table 1.

2.2 Computation of seismicity indicators

For each subregion, the magnitude of the largest earthquake occurring in the following predefined time period (15 days or one week) is predicted using the recurrent neural network model. The input to the neural network

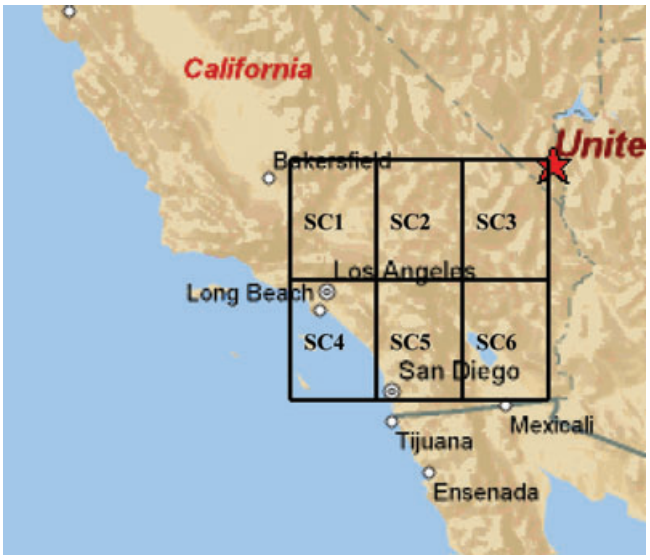

Fig. 1. Southern California divided into six seismic subregions SC1 to SC6 for determining the epicentral location of future earthquakes.

Table 1
Coordinate ranges for the six seismic subregions of southern California shown in Figure 1

Subregion	*Longitude range (W)*	*Latitude range (N)*
SC1	114.75–116.25	35.4–34.2
SC2	116.25–117.5	35.4–34.2
SC3	117.75–119.25	35.4–34.2
SC4	114.75–116.25	34.2–33.8
SC5	116.25–117.5	34.2–33.8
SC6	117.75–119.2	34.2–33.8

is a vector of seismicity indicators computed for each time period of the same duration (15 days or one week) using a predefined number of significant seismic events (defined as those of a predefined threshold magnitude or greater) occurring prior to that time period.

2.3 Neural network modeling

Recurrent neural networks are particularly suited to model time-series data such as temporal magnitude recordings because they incorporate a time delay in their operations through a feedback connection between the output layer and the hidden layer(s). In a recurrent neural network, during every iteration the network output is passed through a recurrent layer and the output of the recurrent layer is added to the output of the hidden layer and the sum is used as the argument of the transfer function to obtain the network output in the succeeding iteration. Network output is obtained as

$$\mathbf{o}_{pi} = \sum_{j=1}^{n} f[\mathbf{s}_i.\mathbf{w}_j + \mathbf{o}_{p_{(i-1)}}.\mathbf{w}_r] \qquad (1)$$

where $\mathbf{o}_{pi}$ is the single element output representing the predicted occurrence (1) or nonoccurrence (0) of an earthquake of threshold magnitude or greater during the ith time period, $\mathbf{s}_i$ is the 8×1 input vector of seismicity indicators for the ith time period, $\mathbf{w}_j$ is the vector of weights of the links connecting the nodes in the input layer to the nodes in the jth hidden layer, f is the transfer function, n is the total number of hidden layers, and $\mathbf{w}_r$ is the vector of weights of the links connecting the nodes in the recurrent layer to the node in the output layer.

The recurrent neural network model for predicting the occurrence of an earthquake of threshold magnitude or greater during the following time period consists of an input layer with eight nodes representing the eight aforementioned seismicity indicators. The number of hidden layers and the number of nodes in the hidden and recurrent layers are determined by numerical

experimentation to obtain the best results. The tan-sigmoid function is used as the transfer function.

The neural network output is either 1 or 0 indicating the occurrence or nonoccurrence of an earthquake of a certain threshold magnitude (different from the threshold magnitude used to define significant seismic events) or greater, respectively. The magnitude of the largest earthquake in the following time-period is determined by gradually increasing the threshold magnitude in increments of 0.5 Richter until the network output changes from 0 to 1 (In this article, an earthquake of a given magnitude m means all seismic events in the magnitude range $m \leq \mathrm{M} < (m + 0.5)$).

Training the recurrent neural network is performed by comparing the network output to the actual occurrence or nonoccurrence of an earthquake of threshold magnitude or greater during the following time-period (desired output). The mean square error between the network output and the desired output is minimized using the Levenberg–Marquardt training algorithm (Hagan et al., 1996). Convergence criteria include limiting the value of the mean square error to 0.001 and the number of training iterations to 1,000.

2.4 Prediction verification

For each subregion, the predicted magnitudes are evaluated by computing the corresponding Hanssen–Kuiper skill score, also known as the real skill or R score. It is defined as the difference between the POD and FAR for each predicted magnitude

$$R = \mathrm{POD} - \mathrm{FAR} = \frac{N_{pc}}{N_{pc} + N_{pi}} - \frac{N_{pi}}{N_{pi} + N_{nc}} \quad (2)$$

where N_{pc} (predicted-correct) is the number of months (or 15 days or weeks, depending on the selected period) during which an earthquake of threshold magnitude or greater occurred and was predicted, N_{pi} (predicted-incorrect) is the number of months during which an earthquake of threshold magnitude or greater did not occur but was predicted, and N_{nc} (not predicted-correct) is the number of months during which an earthquake of threshold magnitude or greater did not occur and was not predicted. The R score is -1 if no correct predictions are made and $+1$ if all predictions are correct. This score is considered advantageous over POD and FAR because it includes an equal representation of both correct and incorrect predictions.

2.5 Difficulty of modeling seismic data for small subregions

The biggest hurdle in modeling seismic data for small subregions is that a large number of significant seismic events may not be present in the historical record prior to all time periods used in training the network. In such regions, the computed values of seismicity indicators may not be an accurate reflection of the seismic activity because of the lack of significant events that define the Gutenberg–Richter distribution, characteristic earthquake distribution, and seismic quiescence. For example, if the threshold magnitude for determining significant seismic events is set to $\mathrm{M} = 4.5$ and the number of significant events used for computing seismicity indicators is set to 100, there are four subregions of southern California without a full set of events prior to all periods. Table 2 summarizes the number of time periods for which sufficient data are not available in each subregion. The shortage of data will affect the network's ability to model seismic activity and the method may not yield accurate results in regions with low seismicity. It can be concluded from Table 2 that there is sufficient seismic data recorded for regions SC1, SC2, and SC4 whereas there is insufficient data recorded for regions SC3, SC5, and SC6 to optimally model seismic activity.

Table 2
Number of periods for which there is incomplete data to compute seismicity indicators for the different subregions. Larger number of periods implies lower prediction accuracies

Subregion	*Number of periods for which 100 events of $M > 4.5$ are not available in the historical record*	*Minimum number of events of magnitude 4.5 or greater available prior to any period in the historical record*
SC1	3	94
SC2	0	110
SC3	11	77
SC4	0	121
SC5	26	40
SC6	29	43

2.6 Example application

As an example, it is attempted to predict the magnitude of the largest earthquake occurring in each subregion of southern California in the following two periods: (1) 15 day and (2) one week. Network operation and results for 15-day time periods are discussed in some detail in the following paragraphs and those for one-week time periods are discussed briefly in the conclusion section.

2.6.1 Network training. For each sub-region, eight seismicity indicators corresponding to each of the 997

15-day periods between January 1, 1950 and December 13, 1990 (997 8 × 1 input vectors) are computed and stored in respective two-dimensional arrays. The indicators are computed from 100 events of magnitude 4.5 or greater prior to the beginning of each time period. If 100 recorded events of magnitude 4.5 or greater are not available prior to the beginning of a certain period, the available number of such events is used. The desired network output (the recorded occurrence or nonoccurrence of an event of magnitude 4.5 or greater) for each period is also stored. As such, for each subregion, there are 997 training data sets. Chronological order of input data are maintained in the recurrent neural network, and network weights are adjusted after each training data set is applied.

2.6.2 Best network architecture. The recurrent neural network is initially modeled with one hidden layer of four nodes and one recurrent layer of one node. Once trained using the 997 training data sets for subregion SC4 (the most seismically active subregion per Table 2) the network with this base architecture is used to predict the occurrence of an earthquake of magnitude 4.5 or greater in SC4 for each of the 12 months of 2005 and the corresponding *R* score is computed for the 12 predictions. The best network architecture is obtained by gradually increasing the number of hidden layers and the number of nodes in each hidden layer and the recurrent layer until the increase in *R* score by doing so is less than 0.01. It was determined that the best network architecture comprised of one hidden layer of eight nodes and four nodes in the recurrent layer.

2.6.3 Prediction results. For each subregion, the trained network with the best architecture is used to predict the occurrence of an earthquake of threshold magnitude or greater for each of the 383 periods between January 1, 1990 and September 24, 2005 (testing data set). Network training and testing is repeated by increasing the threshold magnitude in increments of 0.5 starting with M = 4.5 up to M = 7.0. The computed values of the *R* score for different predicted magnitudes are shown in Table 3.

For the three subregions with adequate modeling data (SC1, SC2, and SC4), for all magnitudes except M = 4.5, the recurrent neural network yields equal or larger *R* score than that obtained for the time period of one month reported in Panakkat and Adeli (2007) (Table 3). And for M = 4.5, the *R* score value drops only slightly.

For subregions with inadequate modeling data (SC3, SC5, and SC6), the *R* score values obtained for the 15-day time period are significantly lower than those computed for the entire Southern California region as well as the seismically active regions (SC1, SC2, and SC4).

Table 3
Computed values of *R* score for the different subregions of California for different predicted magnitudes during the following 15-day time period

	Predicted magnitude					
Subregion	*4.5*	*5.0*	*5.5*	*6.0*	*6.5*	*7.0*
SC1	0.55	0.36	0.72	0.50	0.00	0.00
SC2	0.27	0.34	0.80	0.00	0.00	0.00
SC3	0.36	0.37	0.54	0.50	0.00	0.00
SC4	0.60	0.35	0.89	0.72	1.00	1.00
SC5	−0.13	0.07	0.33	−1.00	−1.00	−1.00
SC6	−0.04	−0.05	−1.00	−1.00	−1.00	−1.00

To determine the statistical significance of the predictions, the POD for each predicted magnitude is compared to the probability of occurrence (p_0) computed using the Poisson's null hypothesis and the results are presented in Table 4. The computed POD is significantly larger than the corresponding p_0 in all instances in the seismically active subregions (except when there is no recorded events of the magnitude in the testing data set for that subregion in which case the POD is 0) establishing that the results obtained are not based on pure chance.

The conclusion of the research so far is reducing the future prediction time period from one month to 15 days by and large improves the accuracy of the prediction for seismically active regions where adequate modeling data are available. Will the prediction accuracy be improved further by reducing the time window for prediction? To answer this question, the prediction time window was reduced to 13, 11, 9, and 7 days. The *R* score results computed for different predicted magnitudes for the most seismically active region, SC4, versus the prediction time window are plotted in Figure 2. The most accurate results are obtained for a prediction time window of about two weeks. In general, the prediction accuracy deteriorates when the prediction time window is reduced below 13 days. It should be noted that this conclusion is only for the region considered with its own particular data set. The methodology presented, however, can be used for any other seismically active region.

The recurrent neural network is found to be more effective in predicting large earthquakes occurring at average intervals of 10–15 years compared to small-to-moderate earthquakes occurring more frequently. This is true for both large seismic regions such as southern California and the San Francisco Bay region (Panakkat and Adeli, 2007) and small seismic regions such as

Table 4
Computed values of POD and p_0 for different predicted magnitudes for each subregion. Where the p_0 values are not given, there were no recorded events of the corresponding magnitude or the computed value of p_0 is less than 0.001

	Predicted magnitude											
	4.5		*5.0*		*5.5*		*6.0*		*6.5*		*7.0*	
Subregion	*POD*	p_0	*POD*	p_0	*POD*	p_0	*POD*	p_0	*POD*	p_0	*POD*	p_0
SC1	0.77	0.39	0.59	0.21	0.72	0.11	0.82	0.06	0.00	0.009	0.00	0.003
SC2	0.59	0.31	0.55	0.13	0.69	0.09	0.77	0.08	0.00	0.010	0.00	0.001
SC3	0.29	0.19	0.38	0.06	0.18	0.02	0.31	0.009	0.00	-	-	-
SC4	0.72	0.34	0.47	0.31	0.71	0.16	0.91	0.11	1.00	0.006	1.00	0.006
SC5	0.31	0.21	0.10	0.19	0.29	0.09	0.29	0.002	0.00	-	0.00	-
SC6	0.39	0.26	0.31	0.08	0.15	0.03	0.11	0.006	0.00	-	0.00	-

those considered in this article. The prediction accuracies computed for subregions SC1, SC2, and SC4 (subregions with adequate modeling data) are significantly improved compared with those computed for the entire Southern California region as reported in Panakkat and Adeli (2007). This could be due to the fact that there is de-clustering of data that form separate temporal and spatial trends when the size of the region under consideration is reduced. Predictions over a smaller spatial and temporal range are more helpful for emergency management efforts. However, in SC3, SC5, and SC6 (subregions with low seismic activity or inadequate modeling data), the recurrent network yields a modest POD and high FAR for all predicted magnitudes leading to the conclusion that the neural network approach is not effective for predicting earthquake magnitude in regions with low seismicity.

Computed *R* scores tend to reduce when the size of the time period is reduced beyond a certain point. The authors deduce this to be a result of the fact that the beginning of a particular time period may fall within an ongoing seismic trend. In other words, when the size of the time period is reduced beyond a certain point, a number of events that comprise the same temporal sequence get distributed into separate time periods. Although this is true for any time period, it becomes more and more probable as the size of the time period grows shorter. The authors conclude that the most accurate prediction time period has to do with the duration of a typical earthquake sequence and the data distribution

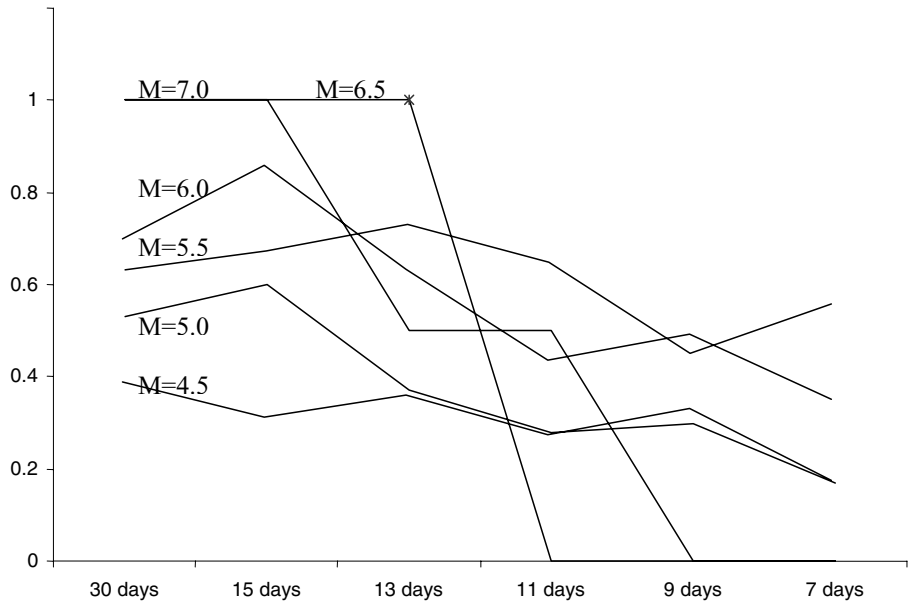

Fig. 2. Computed values of *R* score (Y axis) versus prediction time period (X axis) for different magnitudes for subregion SC4.

in the historical catalog and will be different for different seismic regions.

3 PREDICTION OF TIME OF OCCURRENCE AND EPICENTRAL LOCATION OF MODERATE AND LARGE EARTHQUAKES

3.1 Major earthquakes and intervening elapsed time periods

In this part of the article, a method for predicting the time of occurrence and epicentral location of the following major earthquake (defined as having a certain minimum magnitude) in a large seismic region is presented. The time elapsed between any two successive major earthquakes recorded in the historical earthquake catalog of a region is defined as an elapsed time period. Therefore, the historical earthquake record consists of a number of unequal elapsed time periods.

3.2 Computation of seismicity indicators

For each elapsed time period, the eight seismicity indicators are computed using all significant seismic events (defined as having a certain threshold magnitude or greater) occurring during that period. In contrast to the earlier research reported in Panakkat and Adeli (2007) and the first part of the current article where seismicity indicators are computed using a predefined number of events occurring prior to the beginning of a certain time period (which may include previous large earthquakes), seismicity indicators are computed using all significant events occurring between major earthquakes. Several earthquake occurrence theories presented in the literature dating back to the elastic rebound theory (Reid, 1910) explain the earthquake occurrence phenomenon as following a repetitious temporal pattern. This method may be more appropriate as it attempts to identify temporal recurrence patterns in the occurrence of large earthquakes in a seismic region and use such trends to predict the time of occurrence and epicentral location of the following major earthquake.

3.3 Neural network modeling

The architecture of the recurrent neural network model for predicting the epicentral location of the major earthquake following a particular elapsed time period is shown in Figure 3. The architecture of the recurrent neural network for predicting the following elapsed time period (which gives the time of occurrence of the

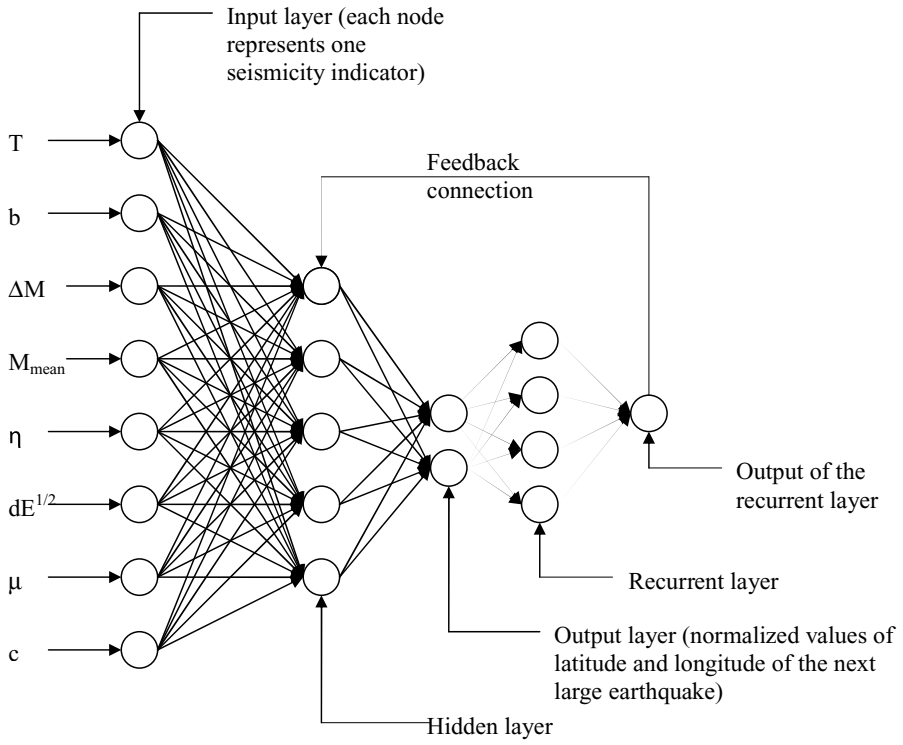

Fig. 3. Architecture of the recurrent neural network for predicting the latitude and longitude of the epicentral location of the following major earthquake.

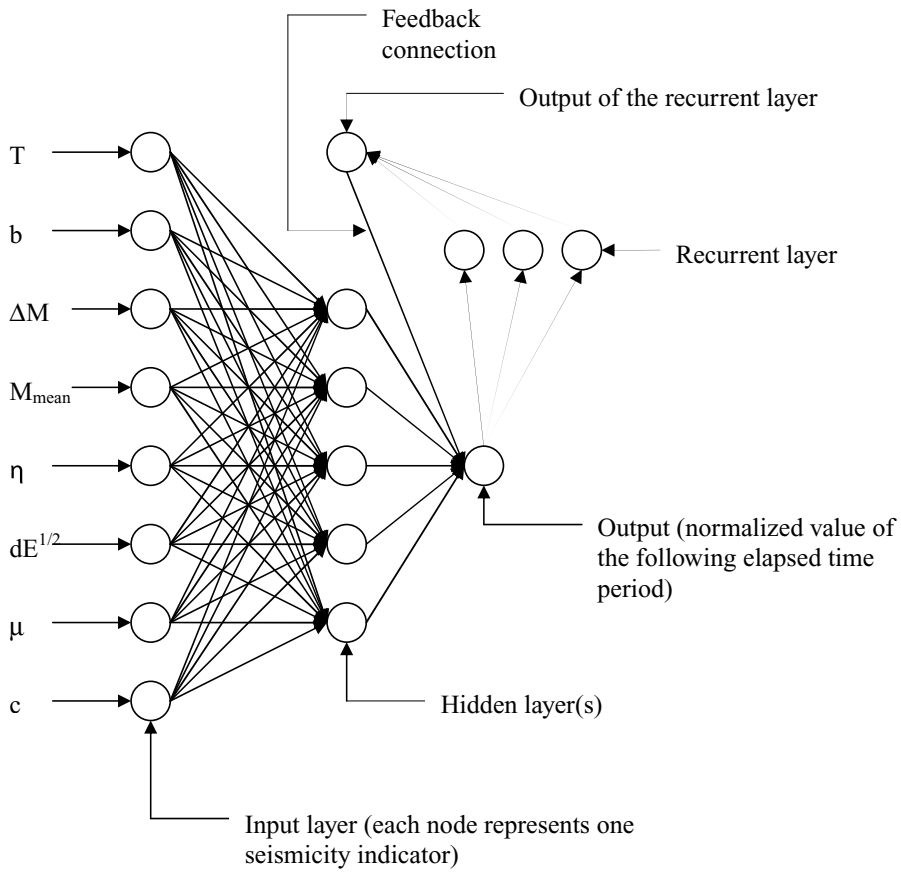

Fig. 4. Architecture of the recurrent neural network for predicting the time of occurrence of the following major earthquake.

following major earthquake) is shown in Figure 4. In the case of epicentral location prediction, the output vector ($\mathbf{o}_i$) in Equation (1) has two elements representing the normalized latitude and longitude of the epicentral location of the major earthquake following the *i*th elapsed time period. In the case of time of occurrence prediction, the output vector has a single element representing the normalized value of the $(i + 1)$st elapsed time period. Normalization is done to utilize the tan-sigmoid transfer function that restricts its output to $(-1, 1)$. Normalization is used also to eliminate the effect of data magnitudes on the classification process (Karim and Adeli, 2003).

Output parameter normalization is done using the following equation:

$$P_{ni} = \frac{P_i}{\sum_{i=1}^{N} \sqrt{P_i^2}} \tag{3}$$

where P_{ni} is the normalized value of the earthquake parameter (latitude, longitude, and the following elapsed time period) corresponding to the *i*th elapsed time period, P_i is the actual value of that parameter, and N is the number of training instances. The mean square error between the network output and the desired output is minimized using the Levenberg–Marquardt training algorithm. Convergence criteria include limiting the value of the mean-square error to 0.001 and the number of training iterations to be performed to 1,000.

3.4 Prediction verification

Once trained, the recurrent neural network is used to predict the epicentral location and time of occurrence of those major earthquakes that are recorded in the historical catalog, but are not part of the training data set (testing data set). In the case of epicentral location prediction, the error in prediction is computed as the geographic distance between the recorded location and the predicted location of the earthquakes in the testing data set. The approximate geographic distance along the surface of the earth between two points is computed from the latitudes and longitudes of the two points as follows

Table 5
Sample training data sets for the 10 time periods between July 31, 1952 and April 26, 1981 for southern California showing eight-element input vectors and the corresponding desired outputs for the prediction of epicentral location and time

	Input vector								*Desired output (normalized)*		
									Location		
Elapsed time period	*T (Days)*	*b*	η	ΔM	M_{mean}	$dE^{1/2}$ ($\times 10^{19} ergs$)	μ *(Days)*	*c*	*Latitude*	*Longitude*	*Elapsed time*
7/31/52–8/22/52	22	0.82	0.34	0.58	4.87	0.005	9	0.69	0.3535	0.1189	0.00091
8/23/52–11/22/52	91	0.71	0.28	0.87	5.16	0.045	6	0.10	0.3578	0.1213	0.00203
11/23/52–6/14/53	203	0.79	0.15	2.67	4.37	0.073	22	0.82	0.3277	0.1155	0.00277
6/15/53–3/19/54	277	0.69	0.56	0.71	4.98	0.004	34	0.51	0.3329	0.1161	0.01717
3/20/54–12/1/58	1717	0.81	0.48	1.13	5.88	0.098	33	0.31	0.3218	0.1157	0.02507
12/2/58–7/9/68	2507	0.95	0.18	0.88	4.65	0.087	57	0.34	0.3317	0.1161	0
7/10/68–2/9/71	0	0.72	0.67	1.60	5.30	0.112	0	0.87	0.3442	0.1184	0
2/9/71–2/9/71	0	0.87	0.59	2.55	5.12	0.101	0	0.91	0.3442	0.1184	0
2/9/71–2/9/71	0	0.80	0.46	0.98	5.40	0.119	0	0.89	0.3310	0.1156	0.03363
2/10/72–4/26/81	3363	0.75	0.80	0.37	4.88	0.037	79	0.45	0.3629	0.1204	0.11021

(Campbell, 2001):

$$D = 68.9722\,[\cos^{-1}\{(\sin a)\,(\sin b) + (\cos a)(\cos b)(\cos P)\}] \quad (4)$$

where D is the distance between the two points in miles, a is the latitude of the first point, b is the latitude of the second point, and P is the difference in the longitudes of the two points.

In the case of time of occurrence prediction, the error in prediction is computed as the actual difference (in days) between the recorded time of occurrence and predicted time of occurrence of the earthquakes in the testing data set.

3.5 Example application

As an example, it is attempted to predict the epicentral location and time of occurrence of major earthquakes in southern California. In this example, major earthquakes are defined as those events with a Richter magnitude of 5.5 or greater. With this definition, there are 41 major earthquakes recorded in southern California between January 24, 1951 and February 22, 2002 leading to 40 elapsed time periods. The threshold magnitude for determining significant events used in computing seismicity indicators is set at 4.5.

3.5.1 Network training. For training the network, eight seismicity indicators corresponding to each of the 40 elapsed time periods between January 24, 1951 and September 20, 1995 (40 8 $\times$ 1 input vectors) are used as input. The recorded values of latitude and longitude of the epicentral location of the major earthquake and the following elapsed time period form the 2 $\times$ 1 and the single element output of their respective neural network. As such, there are 40 training data sets for each neural network.

Sample training data sets for the 10 elapsed time periods between July 31, 1952 and April 26, 1981 showing eight-element input vectors and the corresponding desired outputs are presented in Table 5.

3.5.2 Best Network architecture.

(1) Prediction of Epicentral Location

The recurrent neural network is initially modeled with one hidden layer of four nodes and one node in the recurrent layer. Once trained with the 40 training data sets for epicentral location prediction, the network with this base architecture is used to predict the latitude and longitude of the epicentral location of the major earthquake following the last elapsed time period of the training data set. The geographic distance between the predicted location and the recorded location is computed. The best network architecture is obtained by increasing the number of hidden layers and the number of nodes in the hidden and recurrent layers until the reduction in the computed distance is less than one mile. The best network architecture thus obtained has two hidden layers of 10 nodes and two nodes in the recurrent layer.

(2) Prediction of Time of Occurrence

The recurrent neural network is initially modeled with one hidden layer of four nodes and one node in the recurrent layer. Once trained with the 40 training data sets for time of occurrence prediction, the network with this base architecture is used to predict the following elapsed time period. The difference (in days) between the predicted and actual elapsed time periods is

Table 6
Predicted and actual values for the epicentral location and time of occurrence of the four major earthquakes in southern California between September 21, 1995 and February 22, 2002

		Epicentral location					*Time of occurrence*		
Seismic event	*Magnitude (Richter)*	*Predicted latitude (N)*	*Actual latitude (N)*	*Predicted longitude (W)*	*Actual longitude (W)*	*Error in location (miles)*	*Predicted date*	*Actual date*	*Error in time (days)*
Event 1	7.10	33.39	34.59	115.9	116.3	28.3	07/14/99	10/16/99	94
Event 2	5.77	34.04	34.68	116.1	116.3	38.8	10/21/99	10/16/99	5
Event 3	5.62	34.89	34.44	117.2	116.3	15.2	10/31/99	10/16/99	16
Event 4	5.70	35.68	32.32	114.8	115.3	24.6	05/08/02	02/22/02	75

computed. The best network architecture is obtained by increasing the number of hidden layers and the number of nodes in the hidden and recurrent layers until there is no reduction in the computed difference. The best network architecture has two hidden layers of 10 nodes each and one node in the recurrent layer.

3.5.3 Prediction results. The network with the best architecture is used to predict the epicentral location and time of occurrence of the four major earthquakes with magnitude 5.5 or greater in southern California that occurred after the end of the last elapsed time period in the training data set (September 20, 1995). As noted in Table 6, they include three earthquakes of Richter magnitude 7.10, 5.77, and 5.62 on October 16, 1999 and one earthquake of magnitude 5.70 on February 22, 2002. The predicted and actual locations and time of occurrence of the four events are compared in Table 6. The recurrent neural network model predicted the epicentral location with an error in the range of 15–39 miles, which is useful for emergency management and planning.

Events 1 and 4 in Table 6 are the main shocks and events 2 and 3 are aftershocks of event 1. The accuracy of time of occurrence prediction for main shocks is in the range of 75–94 days and for aftershocks is in the range of 5–16 days. Although in this research there was no direct attempt to differentiate between aftershocks and the main shocks the recurrent neural network appears to automatically learn the occurrence of aftershocks. Considering the fact that large earthquakes in southern California have a recurrence interval of 15–20 years, an earthquake prediction with an error of $\pm$ three months can be considered of value.

Network operation is repeated by changing the threshold magnitude used to define major earthquakes from 5.5 to 6.0. With this definition, there is only one recorded event in the testing data set (event 1). The recurrent neural network yielded an error of 60 days in time and 22.0 miles in distance when used to predict the time and location of this event. When the threshold magnitude was increased to 6.5, the network yielded an error of 56 days in time and 15.7 miles in location. There is an apparent improvement in prediction accuracy on increasing the threshold magnitude. However, due to the relative scarcity of major seismic events in the historical earthquake records of most regions, especially in a time period suitable for retrospective predictions, it is difficult to measure the level of statistical success in the prediction of the location and time of occurrence of very large earthquakes.

4 FINAL COMMENTS

Earthquake prediction is a highly complex and arguably intractable problem. Some purists may argue earthquakes cannot be predicted scientifically. However, earthquake magnitude, epicentral location, and time prediction, even in an approximate sense, can be of great value for an emergency-management and hazard-preparedness perspective. Prediction studies can be classified based on the basic approach that varies from purely theoretical geophysics, to genetic mutations and biology, to statistical, mathematical, and computational modeling of historical earthquake data (Panakkat and Adeli, 2008). The authors take the view that historical data provide some basis and direction for predicting the three earthquake parameters namely, the time of occurrence, epicentral location, and the magnitude of future earthquakes. This article provides a computational model to predict the future-based historical data. The model provides more accurate results in regions where more seismic data is available. It may not yield accurate results in regions with low seismicity.

REFERENCES

Adeli, H. (2001), Neural networks in civil engineering: 1989–2000, *Computer-Aided Civil and Infrastructure Engineering*, **16**(2), 126–42.

Adeli, H., Ghosh-Dastidar, S. & Dadmehr, N. (2005), Alzheimer's disease and models of computation: imaging,

classification, and neural models, *Journal of Alzheimer's Disease*, **7**(3), 187–99.
Adeli, H., Ghosh-Dastidar, S. & Dadmehr, N. (2008), A spatio-temporal wavelet-chaos methodology for EEG-based diagnosis of Alzheimer's disease, *Neuroscience Letters*, **444**(2), 190–4.
Adeli, H. & Hung, S. L. (1993), A concurrent adaptive conjugate gradient learning algorithm on MIMD machines, *Journal of Supercomputer Applications, MIT Press*, **7**(2), 155–66.
Adeli, H. & Hung, S. L. (1994), An adaptive conjugate gradient learning algorithm for effective training of multilayer neural networks, *Applied Mathematics and Computation*, **62**(1), 81–102.
Adeli, H. & Jiang, X. (2003), Neuro-fuzzy logic model for freeway work zone capacity estimation, *Journal of Transportation Engineering, ASCE*, **129**(5), 484–93.
Adeli, H. & Jiang, X. (2006), Dynamic fuzzy wavelet neural network model for structural system identification, *Journal of Structural Engineering, ASCE*, **132**(1), 102–11.
Adeli, H. & Karim, A. (1997a), Neural dynamics model for optimization of cold-formed steel beams, *Journal of Structural Engineering, ASCE*, **123**(11), 1535–43.
Adeli, H. & Karim, A. (1997b), Scheduling/cost optimization and neural dynamics model for construction, *Journal of Construction Management and Engineering, ASCE*, **123**(4), 450–8.
Adeli, H. & Karim, A. (2000), Fuzzy-wavelet RBFNN model for freeway incident detection, *Journal of Transportation Engineering, ASCE*, **126**(6), 464–71.
Adeli, H. & Kim, H. (2001), Cost optimization of composite floors using the neural dynamics model, *Communications in Numerical Methods in Engineering*, **17**, 771–87.
Adeli, H. & Park, H. S. (1995a), Counter propagation neural network in structural engineering, *Journal of Structural Engineering, ASCE*, **121**(8), 1205–12.
Adeli, H. & Park, H. S. (1995b), A neural dynamics model for structural optimization—theory, *Computers and Structures*, **57**(3), 383–90.
Adeli, H. & Park, H. S. (1995c), Optimization of space structures by neural dynamics, *Neural Networks*, **8**(5), 769–81.
Adeli, H. & Park, H. S. (1996), Fully automated design of superhighrise building structure by a hybrid AI model on a massively parallel machine, *AI Magazine*, **17**(3), 87–93.
Adeli, H. & Samant, A. (2000), An adaptive conjugate gradient neural network—wavelet model for traffic incident detection, *Computer-Aided Civil and Infrastructure Engineering*, **15**(4), 251–60.
Adeli, H. & Wu, M. (1998), Regularization neural network for construction cost estimation, *Journal of Construction Engineering and Management, ASCE*, **124**(1), 18–24.
Agnew, D. & Jones, L. (1991), Prediction probabilities from foreshocks, *Journal of Geophysical Research*, **96**(B7), 11959–971.
Ahmadkhanlou, F. & Adeli, H. (2005), Optimum cost design of reinforced concrete slabs using neural dynamics model, *Engineering Applications of Artificial Intelligence*, **18**(1), 65–72.
Alexandre, A., Cuadra, L., Alvarez, L., Rosa-Zurera, M. & Lopez-Ferreras, F. (2008), Two-layer automatic sound classification for conversation enhancement in hearing aids, *Integrated Computer-Aided Engineering*, **15**(1), 85–94.
Banchs, R., Klie, H., Rodriguez, A., Thomas, S. G. & Wheeler, M. F. (2007), A neural stochastic multiscale optimization framework for sensor-based parameter estimation, *Integrated Computer-Aided Engineering*, **14**(3), 213–23.
Campbell, J. (2001), *Map Use and Analysis*, McGraw-Hill, New York, NY.
Chen, H. C., Goldberg, M., Magdon-Ismail, M. & Wallace, W. A. (2008), Reverse engineering a social agent-based hidden Markov model, *International Journal of Neural Systems*, **18**(6), 491–526.
Christodoulou, M. A. & Kontogeorgou, C. (2008), Collision avoidance in commercial aircraft free fight, via neural networks and non-linear programming, *International Journal of Neural Systems*, **18**(5), 371–87.
Cyganek, B. (2007), Circular road signs recognition with soft classifiers, *Integrated Computer-Aided Engineering*, **14**(4), 323–43.
Cyganek, B. (2008), Color image segmentation with support vector machines: applications to road signs detection, **18**(4), 339–45.
Dey, S. & Singh, R. (2003), Surface latent heat flux as an earthquake precursor, *Natural Hazards and Earth Systems*, **3**(1), 749–55.
Dharia, A. & Adeli, H. (2003), Neural network model for rapid forecasting of freeway link travel time, *Engineering Applications of Artificial Intelligence*, **16**(7–8), 607–13.
Fyfe, C., Barbakh, W., Ooi, W. C. & Ko, H. (2008), Topological mappings of video and audio data, *International Journal of Neural Systems*, **18**(6), 481–89.
Ghosh-Dastidar, S. & Adeli, H. (2003), Wavelet-clustering-neural network model for freeway incident detection, *Computer-Aided Civil and Infrastructure Engineering*, **18**(5), 325–38.
Ghosh-Dastidar, S. & Adeli, H. (2006), Neural network-wavelet micro-simulation model for delay and queue length estimation at freeway work zones, *Journal of Transportation Engineering, ASCE*, **132**(4), 331–41.
Ghosh-Dastidar, S. & Adeli, H. (2007), Improved spiking neural networks for EEG classification and epilepsy and seizure detection, *Integrated Computer-Aided Engineering*, **14**(3), 187–212.
Ghosh-Dastidar, S., Adeli, H. & Dadmehr, N. (2007), Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection, *IEEE Transactions on Biomedical Engineering*, **54**(9), 1545–51.
Gil-Pita, R. & Yao, X. (2008), Evolving edited k-nearest neighbour classifiers, *International Journal of Neural Systems*, **18**(6), 459–67.
Gopych, P. (2008), Biologically plausible BSDT recognition of complex images: the case of human faces, *International Journal of Neural Systems*, **18**(6), 527–45.
Gutenberg, B. & Richter, C. F. (1956), Earthquake magnitude, intensity, energy and acceleration, *Bulletin of the Seismological Society of America*, **46**(1), 105–46.
Hagan, M. T., Demuth, H. B. & Beale, M. (1996), *Neural Network Design*, PWS Publishing Company, Boston, MA.
Hooshdar, S. & Adeli, H. (2004), Toward intelligent variable message signs in freeway work zones: a neural network approach, *Journal of Transportation Engineering, ASCE*, **130**(1), 83–93.
Hung, S. L. & Adeli, H. (1993), Parallel backpropagation learning algorithms on Cray Y-MP8/864 supercomputer, *Neurocomputing*, **5**(6), 287–302.
Hung, S. L. & Adeli, H. (1994a), A parallel genetic/neural network learning algorithm for MIMD shared memory

machines, *IEEE Transactions on Neural Networks*, **5**(6), 900–9.
Hung, S. L. & Adeli, H. (1994b), Object-oriented back propagation and its application to structural design, *Neurocomputing*, **6**(1), 45–55.
Huynh, H. T., Won, Y. & Kim, J. J. (2008), An improvement of extreme learning machine for compact single-hidden-layer feedforward neural networks, *International Journal of Neural Systems*, **18**(5), 433–41.
Iglesias, J. & Villa, A. E. P.(2008), Emergence of preferred firing sequences in large spiking neural networks during simulated neuronal development, **18**(4), 267–77.
Jiang, X. & Adeli, H. (2005a), Dynamic wavelet neural network for nonlinear identification of highrise buildings, *Computer-Aided Civil and Infrastructure Engineering*, **20**(5), 316–30.
Jiang, X. & Adeli, H. (2005b), Dynamic wavelet neural network model for traffic flow forecasting, *Journal of Transportation Engineering, ASCE*, **131**(10), 771–9.
Jiang, X. & Adeli, H. (2007), Pseudospectra, MUSIC, and dynamic wavelet neural network for damage detection of highrise buildings, *International Journal for Numerical Methods in Engineering*, **71**(5), 606–29.
Jiang, X. & Adeli, H. (2008a), Dynamic fuzzy wavelet neuroemulator for nonlinear control of irregular highrise building structures, *International Journal for Numerical Methods in Engineering*, **74**(7), 1045–66.
Jiang, X. & Adeli, H. (2008b), Neuro-genetic algorithm for nonlinear active control of highrise buildings, *International Journal for Numerical Methods in Engineering*, **75**(8), 770–86.
Jorgensen, T. D., Haynes, B. P. & Norlund, C. C. F. (2008), Pruning artificial neural networks using neural complexity measures, *International Journal of Neural Systems*, **18**(5), 389–403.
Kagan, Y. Y. & Jackson, D. (1991), Long-term earthquake clustering, *Geophysical Journal International*, **104**, 117–33.
Karim, A. & Adeli, H. (2002), Comparison of the fuzzy–wavelet RBFNN freeway incident detection model with the California algorithm, *Journal of Transportation Engineering, ASCE*, **128**(1), 21–30.
Karim, A. & Adeli, H. (2003), Radial basis function neural network for work zone capacity and queue estimation, *Journal of Transportation Engineering, ASCE*, **129**(5), 494–502
Kellis-Borok, V., Knopoff, L., Kossobokov, V. & Rotvain, I. (1990), Intermediate-term prediction in advance of the Loma Prieta earthquake, *Geophysical Research Letters*, **17**(9), 1461–64.
Khashman, A. (2008), Blood cell identification using a simple neural network, *International Journal of Neural Systems*, **18**(5), 453–8.
Khashman, A. & Sekeroglu, B. (2008), Document image binarisation using a supervised neural network, *International Journal of Neural Systems*, **18**(5), 405–18.
Kimura, T. & Ikeguchi, T. (2007), An optimum strategy for dynamic and stochastic packet routing problems by chaotic neurodynamics, *Integrated Computer-Aided Engineering*, **14**(4), 307–22.
Liu, M. & Zhang, S. (2008), An LMI approach to design H_∞ controllers for discrete-time nonlinear systems based on unified models, *International Journal of Neural Systems*, **18**(5), 443–52.
Ma, L., Zhu, L. & Shi, Y. (1999), Attempts at using seismicity indicators for the prediction of large earthquakes by genetic algorithm-neural network method, *Asia-Pacific Economic Cooperation for Earthquake Simulation*, January 31–February 5, Brisbane, Australia.
Nemissi, M., Seridi, H. & Akdag, H. (2008), The labeled systems of multiple neural networks, **18**(4), 321–30.
Neumann, D., Eckmiller, R. & Baruth, O. (2007), Combination of biometric data and learning algorithms for both generation and application of a secure communication link, *Integrated Computer-Aided Engineering*, **14**(4), 345–52.
Ni, H. & Yin, H. (2008), Self-organizing mixture autoregressive model for non-stationary time series prediction, *International Journal of Neural Systems*, **18**(6), 469–80.
Nishenko, S. (1991), Circum-Pacific seismic potential: 1989–1999, *Pure and Applied Geophysics*, **135**(1), 169–259.
Panakkat, A. & Adeli, H. (2007), Neural network models for earthquake magnitude prediction using multiple seismicity indicators, *International Journal of Neural Systems*, **17**(1), 2007, 13–33.
Panakkat, A. & Adeli, H. (2008), Recent efforts in earthquake prediction (1990–2007), *Natural Hazards Review*, **9**(2), 70–80.
Park, H. S. & Adeli, H. (1995), A neural dynamics model for structural optimization—application to plastic design of structures, *Computers and Structures*, **57**(3), 391–99.
Park, H. S. & Adeli, H. (1997), Distributed neural dynamics algorithms for optimization of large steel structures, *Journal of Structural Engineering, ASCE*, **123**(7), 880–8.
Poitrasson, F., Dundas, S., Toutain, J., Munoz, M. & Rigo, A. (1999), Earthquake-related elemental and isotopic lead anomaly in a springwater, *Earth and Planetary Science Letters*, **169**(1), 269–76.
Raftopoulos, K., Papadakis, N., Ntalianis, K. & Kollias, S. (2007), Shape-based invariant classification of gray scale images, *Integrated Computer-Aided Engineering*, **14**(4), 365–78.
Reid, H. (1910), The mechanism of the earthquake: the California earthquake of April, 18, 1906, *Report of the state earthquake investigation commission*, Carnegie Institute of Washington, Washington, D.C., vol. **2**, pp. 16–28.
Rigatos, G. G. (2008), Adaptive fuzzy control with output feedback for H-infinity tracking of SISI nonlinear systems, **18**(4), 305–20.
Rigatos, G. G. & Tzafestas, S. G. (2007), Neurodynamics and attractors in quantum associative memories, *Integrated Computer-Aided Engineering*, **14**(3), 224–42.
Ruiz-Pinales, J., Jaime-Rivas, R., Lecolinet, E. & Castro-Bleda, M. J. (2008), Cursive word recognition based on interactive activation and early visual processing models, *International Journal of Neural Systems*, **18**(5), 419–31.
Rydelek, P. & Pujol, J. (2004), Real-time seismic warning using a two-station sub-array, *Bulletin of the Seismological Society of America*, **94**(4), 1546–50.
Sabourin, C., Madani, K. & Bruneau, O. (2007), Autonomous biped gait pattern based on fuzzy-CMAC neural networks, *Integrated Computer-Aided Engineering*, **14**(2), 173–86.
Samant, A. & Adeli, H. (2001), Enhancing neural network incident detection algorithms using wavelets, *Computer-Aided Civil and Infrastructure Engineering*, **16**(4), 239–45.
Schneidr, N. C. & Graupe, D. (2008), A modified LAMSTAR neural network and its applications, **18**(4), 331–37.
Senouci, A. B. & Adeli, H. (2001), Resource scheduling using neural dynamics model of Adeli and Park, *Journal of*

Construction Engineering and Management, ASCE, **127**(1), 28–34.

Sharma, M. & Arora, M. (2005), Prediction of seismicity cycles in the Himalayas using artificial neural networks, *Acta Geophysica Polonica*, **53**(3), 299–309.

Sirca, G. & Adeli, H. (2001), Neural network model for uplift load capacity of metal roof panels, *Journal of Structural Engineering, ASCE*, **127**(11), 1276–85.

Sirca, G. & Adeli, H. (2005), Cost optimization of prestressed concrete bridges, *Journal of Structural Engineering, ASCE*, **131**(3), 380–88.

Tashakori, A. R. & Adeli, H. (2002), Optimum design of cold-formed steel space structures using neural dynamic model, *Journal of Constructional Steel Research*, **58**(12), 1545–66.

Taylor, N. R., Panchev, C., Hartley, M., Kasderidis, S. & Taylor, J. G. (2007), Occlusion, attention and object representations, *Integrated Computer-Aided Engineering*, **14**(4), 283–306.

Varotsos, P. & Alexopoulos, K. (1984), Physical properties of the variations of the electric field of the earth preceding earthquakes, *Tectonophysics*, **110**(1), 73–98.

Villaverde, I., Grana, M. & d'Anjou, A. (2007), Morphological neural networks and vision based simultaneous localization and mapping, *Integrated Computer-Aided Engineering*, **14**(4), 355–63.

Working Group on California Earthquake Probabilities. (2004), Earthquake probabilities in the San Fransisco Bay region, *United States Geological Survey Open-File Report*, 03-214.

Yau, W. C., Kumar, D. K. & Arjunan, S. P. (2007), Visual recognition of speech consonants using facial movement features, *Integrated Computer-Aided Engineering*, **14**(1), 49–61.