IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 11, 2018, accepted September 10, 2018, date of publication September 17, 2018, date of current version October 12, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2870189

Determining Neuronal Number in Each Hidden Layer Using Earthquake Catalogues as Training Data in Training an Embedded Back Propagation Neural Network for Predicting Earthquake Magnitude

JYH-WOEI LIN, CHUN-TANG CHAO, AND JUING-SHIAN CHIOU
Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan
Corresponding author: Juing-Shian Chiou (jschiou@stust.edu.tw)

This work was supported by the Ministry of Science and Technology, Taiwan, under Grant MOST 106-2221-E-218-001-MY2.

ABSTRACT In this paper, the 2000–2010 earthquake catalogue with a Richter magnitude (ML) of 5 and a depth of 300 km in the study region, located at 21°–26° N and 119°–123° E, was used as a training data to construct an initial earthquake Richter magnitude (ML) prediction backpropagation neural network (first IEMPBPNN) model with two hidden layers. By using final weights and biases of IEMPBPNN as initials for an embedded earthquake Richter magnitude (ML) prediction backpropagation neural network (EEMPBPNN) from the 1990–1999 and 2011–2014 earthquake catalogues (ML $\geqq$ 5 and depth $\leqq$ 300 km) for the same region, the IEMPBPNN was updated to EEMPBPNN with 10 neurons in each hidden layer. The predicted Richter magnitude (ML) errors could be reduced with EEMPBPNN, and the data from 2000 to 2010 as the outside test and data from 2011 to 2014 as the inside test were compared with the predicted Richter magnitude (ML) under the EEMPBPNN model, which exhibited high accuracy due to the lower standard deviation (SDV), lower mean squared error (mse), and higher correlation coefficient. The accuracy of the second IEMPBPNN, as trained with the 1990–2014 earthquake catalogue under the same proceeding of the first IEMPBPNN, could not be improved with the accuracy of EEMPBPNN. Moreover, the training process of the second IEMPBPNN consumed significant computing time due to massive amount of training data. In predicting the Richter magnitudes of five earthquakes in 2016 and 2018 (TST), lower SDV, lower mse, and higher correlation coefficients were illustrated with reliable prediction accuracy using EEMPBPNN. The objective of this procedure was to determine the neuronal number in each hidden layer using the earthquake catalogue and the slip rate of the Philippine Sea Plate related to the Eurasian plate as the training data, where the number of neurons has not been determined by the training data in previous works.

INDEX TERMS Earthquake catalogue, initial earthquake Richter magnitude (ML) prediction back propagation neural network (IEMPBPNN), embedded earthquake Richter magnitude (ML) prediction back propagation neural network (EEMPBPNN), standard deviation (SDV), mean squared error (MSE).

I. INTRODUCTION

The greatest achievement in seismic prediction for the decade was the discovery that earthquakes do not occur everywhere in the world, but are only concentrated in a few regions. The global standard seismic observation network [18], [45] was established as part of Taiwan's Central Weather Bureau (CWB) in the 1960s, which allowed easier estimation of the locations and extents of earthquakes worldwide.

There are three major seismic zones in the world, namely, the circum-Pacific seismic belt, the Eurasian seismic belt and the mid-ocean ridge seismic belt [62], [81], Taiwan is an island in one of these great earthquake belts. Taiwan is located in one of these great earthquake belts, and, according to past earthquake catalogues [7], [88], the island has suffered many major earthquakes that have caused serious loss of life and property. One such earthquake occurred in 1999, known as

the 921 earthquake [37]. It is inevitable that another major earthquake will occur in Taiwan, therefore, it is necessary to research the prediction of upcoming earthquakes. However, previous attempts at predicting earthquakes indicated that it is a daunting and almost impossible task. Statistical methods are a feasible option if major earthquakes are a regular occurrence [65]. According to this assumption, the regularity of earthquakes could be estimated from a sufficient number of records. Taiwan is located at the collision between the Eurasian and Philippine Sea Olates and exhibits the typical structure of a continental margin island arising from such a collision. Many major and minor earthquakes occur in Taiwan annually. Approximately six or seven million years ago, the Philippine Sea Plate collided with the Eurasian Plate, and the resulting compression of the Eurasian Plate formed the island. Land reclamation and orogeny processes still actively continue to this date [43], [89], [90]. As a result, the island is subject to intense seismic activity. Earthquake prediction in seismology often involves the analysis of earthquake time series and the spatial distribution of seismic variability. When a foreshock occurs, the change in seismic activity can be used to predict the mainshock [60]. However, when there is no significant foreshock, prediction depends on the seismic characteristics prior to the mainshock. The earthquake swarms [26] that occur in the eastern region are relatively significant, but they are not typically accompanied by large earthquakes. Following the 921 earthquake, the National Science Council and relevant research institutions invested a large amount of money and manpower into seismic research, and this effort has generated some useful information in recent years. The Taiwan University team found that at least four similar-sized earthquakes caused by the Chelungpu Fault occurred prior to the 921 earthquake. Large-scale earthquakes with a Richter magnitude (ML) [72] of approximately 7 have occurred every 300 to 400 years on this fault. Active fault recurrence data are important because the analysis of paleoseismicity can aid in determining particular fault characteristics. Ditch digging, drilling, and fault outcrop observation can all obtain fault recurrence data and historical seismic activity trends, but they require significant investment and have shown limited progress [47].

Prior to an earthquake, small and slow crustal deformation will occur as energy accumulates within the crust, and data of such deformation can be a valuable precursor to earthquake prediction [35]. Satellite positioning technology allows the continuous observation of changes in the Earth's crust; a continuous Global Positioning System (GPS) observation network has been deployed in Taiwan for recording crustal deformation information. There are currently several GPS network stations (CWB) in Taiwan, which are mostly distributed across the Chelungpu fault, but related research has achieved limited success. Current observations have shown that large earthquakes are generally accompanied by some coseismic phenomena [87], therefore, the crustal deformation recorded by GPS observation may not necessarily represent true crustal deformation. Observations could also include instrument noise, as well as anomalies arising from the atmosphere or distant earthquakes; therefore, complex preprocessing is required to remove these interferences [27].

When rock is under stress, surface deformation occurs due to dilatancy [15]. Stress can also cause transfer [29], leading to variations in the surface gravity and magnetic field, therefore, a small geomagnetic anomaly appears in the crust due to small ruptures before an earthquake [69]. Gravitational and magnetic anomalies can also be used as earthquake precursors [8], despite the interference that variations in the geomagnetic field face as a result of solar flare-induced geomagnetic storms [32], [82], which can complicate the identification of the geomagnetic anomaly precursors. Intense activity on the surface of the sun, especially during maximum sunspot periods, will increase the number of geomagnetic storms and X-ray, ultraviolet, visible light, and high-energy proton and electron beam flare radiation [32], [82]. Occasionally, rock near the earthquake's epicentre is fractured before an earthquake due to the accumulation of energy in the Earth's crust, which can be recorded by electric and magnetic fields, and electromagnetic waves, as confirmed by laboratory experiments. Electric and magnetic fields and electromagnetic waves may affect the outer surface of the ionosphere [69]. A study on earthquakes of ML$\geq$5.0 in Taiwan investigated these relationships and found that the total number of ions in the ionosphere decreased approximately five days before the occurrence of an earthquake [42]. Although reports of this phenomenon remain controversial [92], such results are currently insufficient for earthquake prediction as changes in the total number of ions in the ionosphere cannot indicate time, ML, or the location of an earthquake. Stress changes in the Earth's crust can also lead to changes in the groundwater level and fluids within the rock, resulting in geochemical and hydrologic variation that can be taken as earthquake precursors. The most commonly used method for measuring these precursors is analyzing the radon levels in hot spring water samples to explore the relationship between geochemical and hydrologic changes and seismic activity [24], [44]. However, some difficulties remain as specialized facilities are required for these measurements; for example, hot water can corrode instruments, which makes it difficult to obtain accurate results. Following the Chi-Chi earthquake, a Taiwan University research team funded by the Central Geological Survey successfully developed more advanced observation facilities and obtained some basic results. However, earthquakes of ML 3 cannot be accurately predicted and it is still difficult to apply geochemical and hydrologic changes to earthquake prediction. Earthquake light [19], clouds [22], and sound [66] are natural phenomena that may indicate the release of both electric and magnetic energy stored in the Earth's crust prior to an earthquake, but further studies are required to obtain a deeper understanding. Some studies have observed abnormal animal behaviors before earthquakes [20], although this suggestion is also controversial.

An artificial neural network (ANN) with artificial intelligence (AI) has been used to efficiently resolve a series of

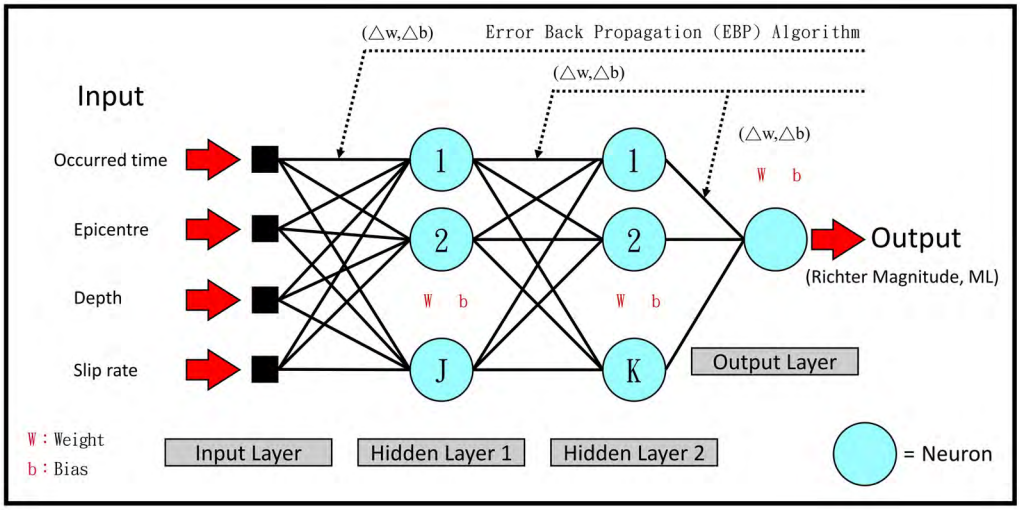


FIGURE 1. The framework of a BPNN includes two hidden layers. The first hidden layer has J neurons and the second hidden layer has K neurons. The LMA is used for the error back propagation (EBP) algorithm of BPNN. The EBP algorithm estimates the errors (ΔW, Δb) in terms of weights and biases between each hidden layer and minimizes the errors.

complex and nonlinear problems, such as fatigue behavior prediction, the properties of high carbon steel, and material data validation and imputation [54], [55], [83]. The objective of this study was to instruct the training process of an ANN to establish a network model with two hidden layers and determine the optimal neuronal number in each hidden layer for predicting the magnitude of an upcoming earthquake. This involved the re-training of a network model using a new earthquake catalogue without requiring a long significant computing time and relying on more initial inputs from earthquake catalogues. It was expected that the prediction error would be reduced when a long computing time was no longer required owing to the use of an earthquake catalogue covering a long time period. This was implemented without the use of localized geological features and could be accomplished by determining the optimal number of neurons through an intelligent optimization training algorithm using an inversion method to construct the ANN.

II. BACK PROPAGATION NEURAL NETWORK (BPNN)

A back propagation neural network (BPNN) was proposed by Rumelhart and McClelland in 1986 and has been used to model pt/c cathode degradation in PEM fuel cells and detect Catechol in water [50], [53]. The BPNN is based on the multilayer perceptron (MLP) framework and uses the error back propagation (EBP) algorithm. BPNNs are widely used for supervising to accurately solve several nonlinear fitting problems [2], [9], [75]. The operation of an artificial neuron includes a non-linear element with an activation function and two parameters, i.e., the weight (W) and bias (b) [25]. The Levenberg–Marquardt algorithm (LMA), which is used to obtain optimum weights and biases in the EBP algorithm, is frequently applied to irregular data patterns [3]. The framework of a BPNN was introduced with four layers, two of which are hidden, as shown in Fig. 1, and its corresponding algorithm without EBP was defined with four inputs. This produced an input vector and a final output as follows [23]:

$$Y^3 = f^3(W_1^3 f^2(W_K^2 f^1(W_J^1 R + b_J^1) + b_K^2) + b_1^3) \quad (1)$$

where R is an input vector; and W, b, and f denote the weight matrix, bias vector, and activation function, respectively. The number of the layer is appended as a superscript to these variables, and the subscript of these variables indicates the number of neurons.

The two hidden layers were then linked to the final output, as an output could pass through a neuron's activation function in the connections of an ANN model [6], [57]. In this study, the final output is Richter magnitude (ML), and four components are used in the input vector (R) to train the BPNN model, including the occurred time, epicentre and depth of each earthquake from an earthquake catalogue without considering their magnitudes (ML), and the slip rate of the Philippine Sea Plate in relation to the Eurasian plate, as illustrated in Fig. 2. Corresponding demonstrations and explanations of the matrix are included in Figs 2(a) and (b), respectively.

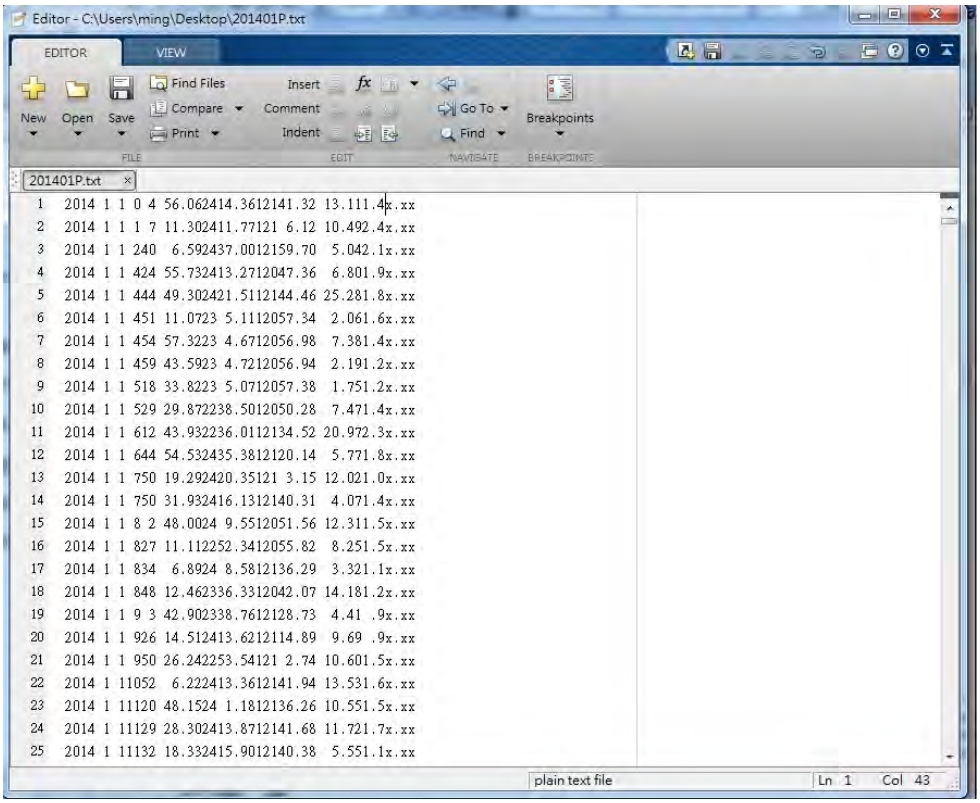
The sigmoid function [31] used as an activation function in this study is presented below.

$$f(t) = \frac{1}{1 + e^{-t}} \quad (2)$$

In this study, the power of the BPNN will be increased by training the network with additional earthquake catalogues without re-training the original BPNN.

III. PAST RESEARCH USING BPNN (ANN) IN SEISMOLOGY

Several studies have used ANN and BPNN in seismology. Reference [80] changed the levels of BPNN training in relation to variations in the horizontal component of the



(a)

Field Explanation:

Position		Column	Meaning
Start	End	Number	
1	4	4	Year
5	6	2	Month
7	8	2	Day
9	10	2	Hour
11	12	2	Minute
13	18	6	Second
19	20	2	Latitude（degree）
21	25	5	Latitude（minute）
26	28	3	Longitude（degree）
29	33	5	Longitude（minute）
34	39	6	Depth（KM）
40	43	4	Richter Magnitude Scale（M_L）

(b)

FIGURE 2. The matrix form of earthquake catalogue, which belongs a time series, on January 1, 2014 with 25 earthquakes with four parameters, which are occurred time, epicentre, depth and Richter magnitude(ML), is shown in Figure (a). Richter magnitude(ML) is not used as input. Four parameters From column 43 to 47, the character x.xx denotes a value with an integer and two decimals, which is the slip rate of the Philippine Sea Plate, relative to the Eurasian plate. The explanation of the format without the slip rate for the matrix form is shown in Figure (b). The size in this case has the dimension of 25x5 (Data Source, CWB).

geomagnetic field (data source: Tehran Geophysics Research Center in 1970-1976), hourly relative moisture, ground temperature, daily rainfall rate, and the average daily rainfall duration to predict the ML of an earthquake occurring two days prior to another earthquake. However, the measurement data used to train the model were vulnerable to climatic factors and accuracy issues, and their density values were dependent on the temporal and spatial resolution of sampling. The Earth's magnetic field is particularly vulnerable to space weather, and these data were assumed to be associated with earthquakes. Reference [49] used a BPNN model to create an automatic forecasting model for predicting earthquakes using strong seismic motion data, but prediction relied heavily on accurate, reliable, and timely data. Reference [1] used a BPNN model to predict the ML of earthquakes in the northern Red Sea, the Gulf of Aqaba, the Gulf of Suez, and the Sinai Peninsula. They also used other forecasting methods, including the moving average, normal distributed random predictor, and the uniform distributed random predictor. They applied statistical methods, such as linear quadratic and cubic regression, to the same region, and used and compared different evaluation metrics with the BPNN model, finding that the accuracy of the BPNN model was at least 32% as it is more suitable for identifying nonlinear relationships. Reference [68] applied the BPNN model to analyses abnormal VLF/LF signals and utilized and automatic forecasting system to identify VLF/LF signal anomalies, which could serve as activity precursors to earthquakes owing to their sensitivity to seismic activity, and established automatic VLF/LF detection. However, the VLF/LF signal is vulnerable to disruption by geomagnetic storms. Reference [11] used the BPNN model to predict the earthquake with the largest magnitude that would occur in the following month using a BPNN trained with earthquake catalogues from 1994-2011. The prediction success rate for the subsequent month's largest earthquake magnitude was approximately 47%. Reference [71] used the BPNN model to predict the ML of an earthquake in Chile by dividing the country into four zones, further divided into 0.5-1.0 degrees of latitude and longitude in order to build an earthquake prediction model. The b values of the Gutenberg-Richter theorem served as inputs, but the predicted outcome was susceptible to calculation error. Reference [79] used a BPNN model to analyses ionosphere dynamics and simulated their earthquake-induced patterns and characteristics. Ionosphere conditions cannot be easily observed or used as earthquake precursors as they are affected by the sun [39]–[41], [56]. If the errors in precursor patterns cannot be resolved during BPNN training, then simulated ionosphere characteristics related to earthquakes will exhibit the same errors.

Reference [61] used a MLP neural network with two hidden layers to predict the magnitude of earthquakes with 128 network models. The inputs to train these models included the latitudes, longitudes, depths, and times, as well as the soil types and fault mechanisms of 4099 earthquakes that occurred in Iran from the ground motion database of the International Institute of Earthquake Engineering and Seismology (IIEES). The data from the catalogue were first revised and declustered, and aftershocks and foreshocks were omitted using the Uhrhammer method [21]. A depth of 33 km was assumed for some earthquakes with an unknown depth. The soil type and fault mechanism classification depended on regional geological environmental features, which were not easy to identify, and there was no method to directly determine the optimal number of hidden nodes, indicating the difficulties of this method. Therefore, four different numbers of hidden neurons, i.e., 8, 12, 16 and 20, were used for each layer. They categorized the magnitude of earthquakes into four classes, denoted as by A, B, C, and D, that represented ML ranges of 4-5, 5-6, 6-7, and greater than 7, respectively. Four different learning epochs, including 1, 2, 4, and 8, have been used to resolve this problem, which was referred to as overtraining [73]. Therefore, 128 different groups of parameters and models were set for complicated analysis, especially during data mining.

Reference [59] used a BPNN to predict earthquake magnitude in the Himalayas, India. The earthquakes were all above 2.5 in magnitude for all events up to 2014, and, although it is necessary, the Gutenberg-Richter inverse power law curve for the earthquakes was not easy to determine as its components included the rate of the square root of seismic energy released during the earthquakes, coefficient of variation for the average time, average magnitude, and the mean square deviation of the regression line. Reference [91] used ANNs and stochastic techniques, including Gamma, Lognormal, Weibull, and Log-logistic possibility models, to predict the possible time of occurrence using an updated earthquake catalogue with magnitude Mw $\geq$ 6.0 from 1737 to 2015 in the study area between latitudes of 19.345° N and 29.431° N and longitudes of 87.590° E and 98.461° E. They estimated earthquake occurrence probabilities above 0.8 with probable future magnitudes of Mw 6.6 in the Churachandpur-Mao fault (CMF) region from 2014 to 2017, Mw 6.8 in the Myanmar Central Basin (MCB) region from 2013 to 2016 and Mw 6.5 in the Eastern Boundary Thrust (EBT) and Kabaw regions from 2015 to 2018. The prediction accuracy of certain earthquakes, including the Manipur earthquake on January 4, 2016 (Mw=6.7), Myanmar earthquake on April 13, 2016 (M 6.9), and Myanmar earthquake on August 24, 2016 (M 6.8) in the Churachandpur-Mao fault (CMF), Myanmar Central Basin (MCB), EBT, and Kabaw regions, respectively, with the ANN model was high. However, Gamma, Lognormal, Weibull and Log-logistic possibility models always required a parameter called return period for selecting different estimation model for occurrence time of earthquake was not always easy. Moreover, the return period was not constant for extreme events, such as earthquakes and tropical cyclones, which further complicated the work [70].

Reference [4] used ANNs to predict earthquake magnitude in the Hindukush region, which is one of the world's most seismically active regions, especially at depths of 70 to $\leqq$ 300 km. In this study, the Gutenberg–Richter curve

equation was required to predict earthquake magnitude. The b value from the Gutenberg–Richter inverse power law is an important parameter to build the Gutenberg–Richter curve (Rundle, 1989). Therefore, with this analysis method, earthquake parameters, such as energy release, could be estimated and then used by the ANN to predict the magnitude of an upcoming earthquake. However, these parameters are not easily estimated, so the training of ANNs is still complicated. If one of these parameters was not accurate, the predicted earthquake magnitude would be incorrect. Reference [36] detected the wave frequency by following a seismic detection method using a heterodyne laser interferometer. By applying an ANN with a STA/LTA algorithm to the measured seismic data, they could detect seismic signals with a certain degree of sensitivity. However, this method relied on the detection of a true seismic wave signal to predict earthquakes in real-time. However, a heterodyne laser interferometer must be operated correctly, because additional noise would be introduced as distortion to quality when the seismic signals were converted according to their frequency. It was also important to develop earthquake early warning (EEW) as it was already associated with earthquake prediction.

Reference [17] developed a ground motion prediction equation based on an ANN to predict shallow earthquakes. They used 13,552 ground motion records from 288 earthquakes provided by the Pacific Engineering Research Center. The input parameters for training the ANN included the moment magnitude (M_w), the closest distance to rupture plane, shear wave velocity in the region, and focal mechanism. These parameters must be precisely calculated. This analysis method was very complicated, because these parameters were obtained from 13,552 ground motion records, and their accuracy affected the prediction result, causing serious concern. In previous studies, the decision-making neuron network architecture was subjective. The Taguchi Method is a decision-making neuronal network architecture with corresponding parameters [57]. However, the properties of this method are statistical [5], so it is not an intelligent optimization training algorithm. The number of neurons was passively determined, but not decided, from training data, which was the proposition and purpose of this study.

IV. STUDY REGION AND DATA SOURCE

The region for this study is located between 21° N–26° N and 119° E–123° E. The foremost collected training earthquake catalogue in matrix form (Real-time Seismic Monitoring Network, CWB) for training the initial BPNN covered an 11-year period from 2000 to 2010 (266 earthquakes, Fig. 3). The data from the real-time Seismic Monitoring Network stations of the CWB were used to determine the occurrence time, epicentre, and depth of earthquakes, which formed the initial collection for the later construction of the earthquake catalogue (Fig.2).

In this study, earthquakes with ML of 5 and depth of 300 km were used as inputs to train **the initial BPNN** for the observation of 11 years from 2000 to 2010 at 21° N–26° N and 119° E–123° E. Therefore, the trained BPNN was used to predict earthquakes with ML of 5 and depths of 300 km.

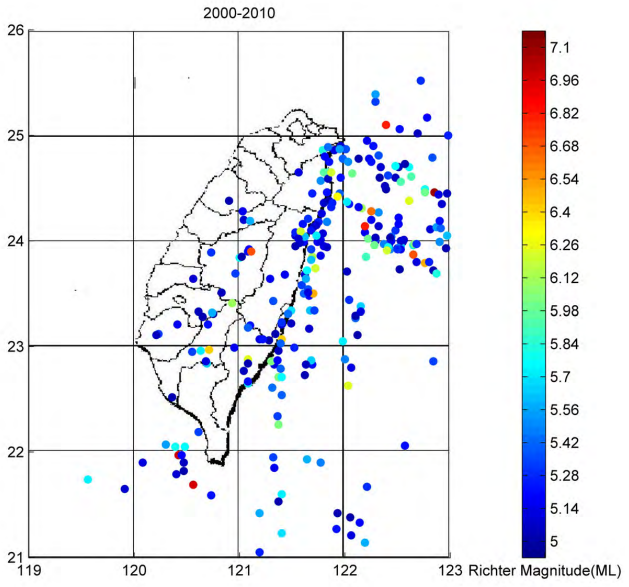


FIGURE 3. Data of 266 earthquakes from 2000 to 2010 for training the BPNN (IEMPBPNN) (Richter magnitude (ML) 5 and depth300 km).

The geological background was not necessary for training the ANN model, and the number of neurons with two hidden layers was decided using these training data as two hidden layers were used in this study. Several studies and engineering applications have indicated that a network of two hidden layers with small numbers of neurons could be used in place of a network with a large number of neurons in a hidden layer [28], [33], [34], [64], [85].

V. TRAINING PROCEDURE OF BPNN

The BPNN model as a type of forecasting method based on time series [78] is trained to predict the Richter magnitude (ML) of earthquakes, and the training procedure is a standard approach to prepare BPNN for upcoming earthquakes.

S_{plate} is the slip rate of the Philippine Sea Plate relative to the Eurasian Plate. This value ranges from 7–8 cm/y, and is randomly assigned with value of 0 to 1 after the feature scaling [64]. The initial weights and biases are also random variables set with a value of 0 to 1 because the sigmoid function output is between 0 and 1. The training epoch is set as 1000, and the learning rate is adaptive at 0 to 1 with an increment of 0.01. A target output [74] with a value of 1 is accepted. This error value is defined as the absolute value of the difference between the target output and the output of training the BPNN. Moreover, the error for these training patterns is defined as the training error.

To further explain the slip rate, the Philippine Sea Plate was assumed to move from the southeast to the northwest relative to the Eurasian plate at a rate of 7-8 cm per year as a long-term effect of the geological structure that has defined the behavior and information of past faults and seismic

activity, excluding the large plate movement. The slip rate also controls the earthquakes in eastern Taiwan that are caused by plate movement, rather than faults. A rate of 7-8 cm/y has been widely accepted by several studies [12], [13], [16], [43], [76]. However, the time period of the training data generated by the earthquake catalogue is too short. Thus, variations in slip rate must be included in the neural network input for long-term risk prediction and future seismic activity, as it would be illogical to predict events beyond 100 years without continuous slip rate data to indicate plate activity. Through these movements, the hidden geological conditions for training the BPNN will allow long-term future prediction. The logic behind such data processing to train the BPNN is similar to the theory of information processing. Predicted outputs for a short time period can be compared to the short-term memory of the information processing theory. Meanwhile, the predicted outputs of the BPNN over a long-term period can be compared to the long-term memory (LTM) concept for information processing [14], [46], [48]. When training the BPNN, the input matrix defined in this study was as follows:

$$R = [EC_{11} + S_{plate}] \tag{3}$$

where $[EC_{11} + S_{plate}]$ represents the collected training earthquake catalogue from 2000 to 2010, including S_{plate}, as training data to build the initial BPNN. Small prediction errors of the Richter magnitudes were expected when using an earthquake catalogue covering a short period. Training the BPNN using the results of Fig. 4 presented the initial earthquake Richter magnitude (ML) prediction back propagation neural network (IEMPBPNN) with a blue curve as the first training step. The learning rate was approximately 0.83. The training error was high when 10 neurons were present in each hidden layer; therefore, the error of the outputs should be relatively high as the earthquake catalogue required to train the BPNN was insufficient. The first ANN was built to model the first IEMPBPNN.

The second training step was then processed. Earthquake catalogues with the same defined criteria (ML $\geqq$ 5 and depth $\leqq$ 300 km) from 1990 to 1999 (316 earthquakes) and 2011 to 2014 (85 earthquakes) with S_{plate} were used as the second training data to retrain a new BPNN and determine the predicted Richter magnitudes errors. The initial weights and biases for re-training the BPNN were the final weights and biases of the IEMPBPNN to retain its features to build a new BPNN model. This was also expected to improve the Richter magnitude prediction accuracy. The second ANN model was then built. In Fig. 4, over 10 neurons were used in each hidden layer, and the training errors were almost identical. A learning rate of 0.25 resulted in the best accuracy. Therefore, the retrained BPNN was defined as the embedded earthquake Richter magnitude (ML) prediction back propagation neural network (EEMPBPNN). Moreover, using more neurons to train the BPNN would not be effective to reduce the degree of error due to the issue of over-fitting [30].

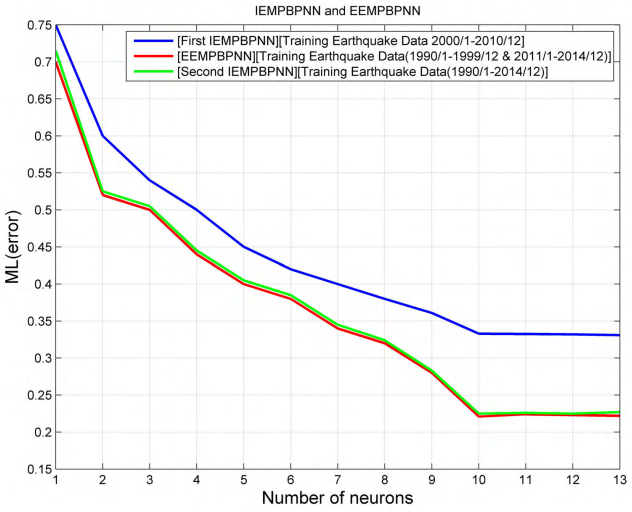


FIGURE 4. The training errors of first IEMPBPNN (first ANN model) as the number of neurons increases are indicated via a blue curve during the period from 2000 to 2010. The best learning rate is 0.83, which shows the best accuracy. Errors of EEMPBPNN are indicated by a red curve as the number of neurons in each hidden layer increases. When the number of neurons is equal to 10, the training error of EEMPBPNN is within the average error of $\Delta ML = \pm 0.3$, while the training error of IEMPBPNN is greater than the average error of $\Delta ML = \pm 0.3$. Therefore, this EEMPBPNN is expected relatively high prediction error in Richter magnitude (ML). Thus, to update IEMPBPNN to EEMPBPNN (second ANN model), the earthquake catalogues from 1990 to 1999 and from 2011 to 2014 are used, where results are indicated by the red curve, and the best learning rate is 0.25. For comparison with this EEMPBPNN, the training errors of second IEMPBPNN as the number of neurons increases are indicated by a green curve for the 1990–2014 earthquake catalogue, and the best learning rate is 0.33, which also shows the best accuracy. However, the accuracy of second IEMPBPNN is not enhanced (the third ANN model).

The first and second hidden layers both contained 10 neurons, according to Fig. 4, in which the red curve represents the variations in the training error as the number of neurons increases. This error is relatively small and can be expected to provide a more accurate Richter magnitude (ML). When the new earthquake catalogue with the same defined criteria (ML $\geqq$ 5 and depth $\leqq$ 300 km) as those in section four was used to retrain the BPNN, it was subsequently updated to EEMPBPNN, which was previously IEMPBPNN.

Finally, the third training step was processed. To confirm the efficiency of the EEMPBPNN and continue the third training process, the training errors of the second IEMPBPNN for a long time period (covering the earthquake catalogues for the periods 2000-2010, 1990-1999, and 2011-2014 for first IEMPBPNN and then EEMPBPNN) with an increase in the number of neurons through the same processing of the IEMPBPNN from 2000 to 2010 were indicated by a green curve in the 1990–2014 earthquake catalogue. The best learning rate was 0.33, which was also the best accuracy. The third ANN was built as a model of the second IEMPBPNN. The accuracy of the second IEMPBPNN was, therefore, insufficient in comparison to that of the EEMPBPNN, which required an extremely long computing time and training with a massive amount of data. The complexity and the number of iterations did not increase using EEMPBPNN (Figure. 4). However, as previously

stated, when more than 10 neurons were used in each hidden layer, the training errors were almost identical. Therefore 10 neurons in each hidden layer is ideal and good for a specific data to build EEMPBPNN. The results prove that EEMPBPNN can be used instead of Two IEMPBPNN

The red and green curves in Fig. 4 are nearly identical at first glance. However, when both curves were examined in detail, they were not parallel and only exhibited the similar, yet true, behaviors of both curves in this study. Furthermore, earthquake catalogues for other regions or countries are essential to develop a generalized method in the future. However, a new method for this has been provided in this study.

Therefore, separate BPNN models were constructed, which included two IEMPBPNNs trained with the 2000–2010 and 1990-2014 earthquake catalogues (667 earthquakes), and an EEMPBPNN trained with the 2011–2014 and the 1990-1999 earthquake catalogues. Both hidden layers contained 10 neurons (J=K=10) in Fig. 1. For each earthquake of the three previous earthquake catalogues, four inputs, including occurred time, epicentre, depth, and slip rate, were used to generate an input vector. The final output is Richter magnitude (ML). The model functions of the two IEMPBPNNs and EEMPBPNN from Eq. 1 could be determined as follows [23]:

$$(ML)_{1x1} = f^3([W_1^3]_{1x10} f^2([W_{10}^2]_{10x10} f^1([W_{10}^1]_{10x4}[R]_{4x1} + [b_{10}^1]_{10x1}) + [b_{10}^2]_{10x1}) + b_{1x1}^3). \quad (4)$$

VI. RESULTS OF INSIDE AND OUTSIDE TESTS

The predicted ML were obtained from the EEMPBPNN with 10 neurons in each hidden layer (red curve) in Fig. 4. However, to confirm the accuracy of the EEMPBPNN, both inside and outside tests were performed. The ML were predicted from 2000 to 2010 in the outside test and from 2011 to 2014 in the inside test, and found that, under the same defined criteria (ML $\geqq$ 5 and depth 300 km) presented in section four, the ML of the earthquakes from 2000 to 2014 confirmed the accuracy of EEMPBPNN when they were predicted again and compared with the true values.

For earthquakes with same defined criteria (ML $\geqq$ 5 and depth $\leqq$ 300 km) between 2000 and 2014, there were matches between the time of occurrence, epicentre, and depth, as well as their ML determined by the EEMPBPNN and the actual values.

Fig. 5 presents the errors for the predicted ML of earthquakes that occurred from 2000 to 2014. For assessing the reliability of the outputs from training the ANN, the results were statistically evaluated. Reference [51] used ANNs to predict the shot-peening effects on the residual stress and hardness composite.

The performance was statistically evaluated using four prediction score metrics from the test data, including the Pearson coefficient of correlation (PCC), root mean square error (RMSE), mean relative error (MRE), and mean absolute error (MAE). Reference [52] used the coefficient of correlation (R^2), RMSE, MRE, and MAE to statistically evaluate the training of an ANN that was developed to determine the effects of the hardened nickel coating on the fatigue behavior of CK45 mild alloy steel. In this study, the training data were earthquake catalogues with specific features, such as the discrete degree of the predicted ML relative to the true value. Therefore, the standard deviation (SDV), mean square error (MSE), and correlation coefficient [84] were suggested as statistical approaches to evaluate the predicted errors from the inside and outside tests, differing from those studies [51], [52]. A low SDV, low MSE, and high correlation coefficient confirmed the accuracy for the reliable application of the EEMPBPNN, as shown in Fig. 6.

VII. PREDICTION OF FIVE EARTHQUAKES USING EEMPBPNN

In this section, the ML of five earthquakes that in Taiwan occurred after 2015 were predicted using the EEMPBPNN, which had occurred in the study region with ML of 5.0 and depth of 300 km. This prediction does not depend on the inside and outside tests stated previously. (1) The first earthquake occurred at 03:57:26 on February 6, 2016 (TST) at a depth of 14.64 km, with a ML of 6.60. The epicentre was located at 22.92° N and 120.54° E. The predicted ML was 6.23. (2) The second earthquake occurred at 23:50:42 on February 6, 2018 (TST) at a depth of 6.31 km, with a ML of 6.26. The epicentre was located at 24.10° N and 121.73° E. The predicted ML was 6.01. (3) The third earthquake occurred at 22:47:01 on February 19, 2018 (TST) at a depth of 52.68 km, with a ML of 5.53. The epicentre was located at 24.59° N and 121.61° E. The predicted ML was 5.30. (4) The fourth earthquake occurred at 07:10:41 on February 22, 2018 (TST) at a depth of 35.57 km, with a ML of 5.66. The epicentre was located at 23.41° N and 121.53° E. The predicted ML was 5.45. (5) The fifth earthquake occurred at 07:10:41 on May 2, 2018 (TST) at a depth of 24.75 km, with a ML of 5.59. The epicentre was located at 24.02° N and 122.36° E, and the predicted ML was 5.80.

Four earthquakes that occurred in 2018 were selected to demonstrate the prediction capability of the EEMPBPNN for earthquake magnitude in the same year, because the EEMPBPNN was trained with data from before 2014. The predicted errors are shown in Fig. 7. A low SDV, low MSE, and high correlation coefficient also confirmed the reliability of prediction accuracy, as shown in Fig. 8. The results indicated that the proposed EEMPBPNN method could predict the magnitudes of two earthquakes that occurred in 2016 and 2018, and may be a feasible approach for predicting upcoming earthquakes two to four years in advance.

VIII. DISCUSSION

The average error of the ML in Taiwan was $\Delta ML = \pm 0.3$ [10], [88]. In Fig. 4, the training errors of the EEMPBPNN (red curve) (<0.25) were below the average ML error ML ($\Delta ML = \pm 0.3$); therefore, the error in the predictive ML of the EEMPBPNN should be lower for

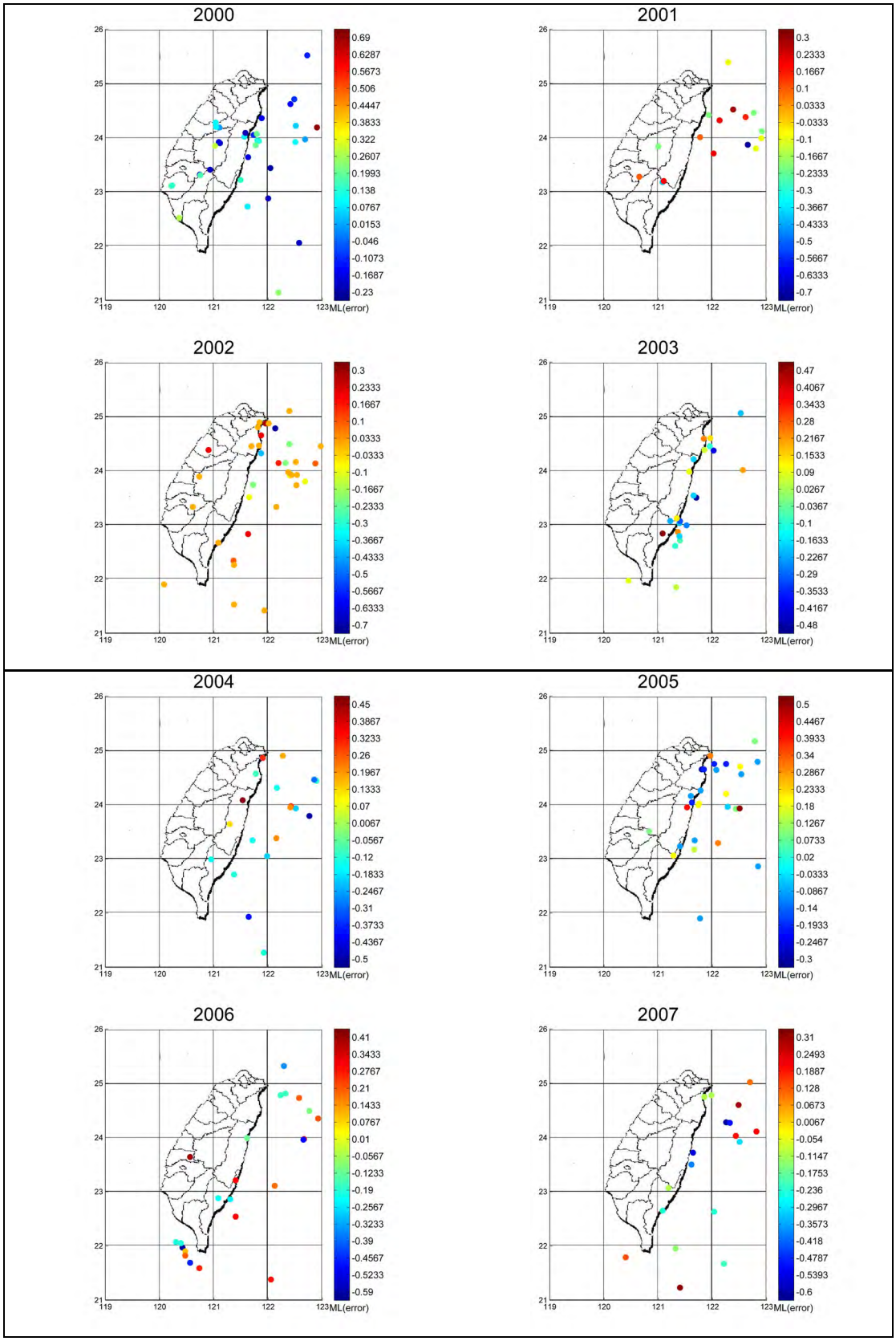


FIGURE 5. Prediction errors in Richter magnitude (ML) of earthquakes (ML $\geq$ 5, depth300 km) from 2000 to 2014 using EEMPBPNN. The difference between the predicted value and the real value is defined as the error.

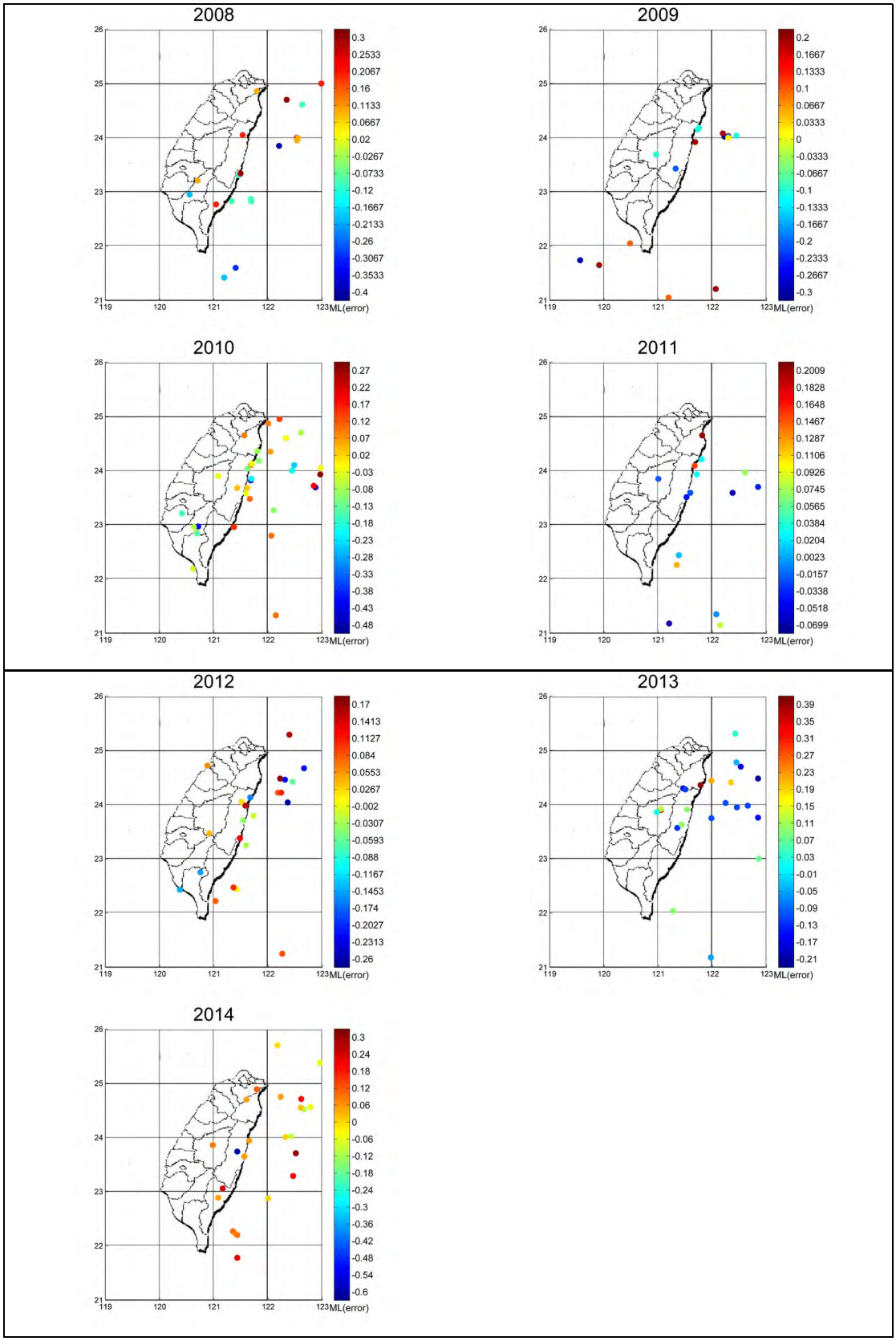


FIGURE 5. ***(Continued.)*** **Prediction errors in Richter magnitude (ML) of earthquakes (ML $\geq$ 5, depth300 km) from 2000 to 2014 using EEMPBPNN. The difference between the predicted value and the real value is defined as the error.**

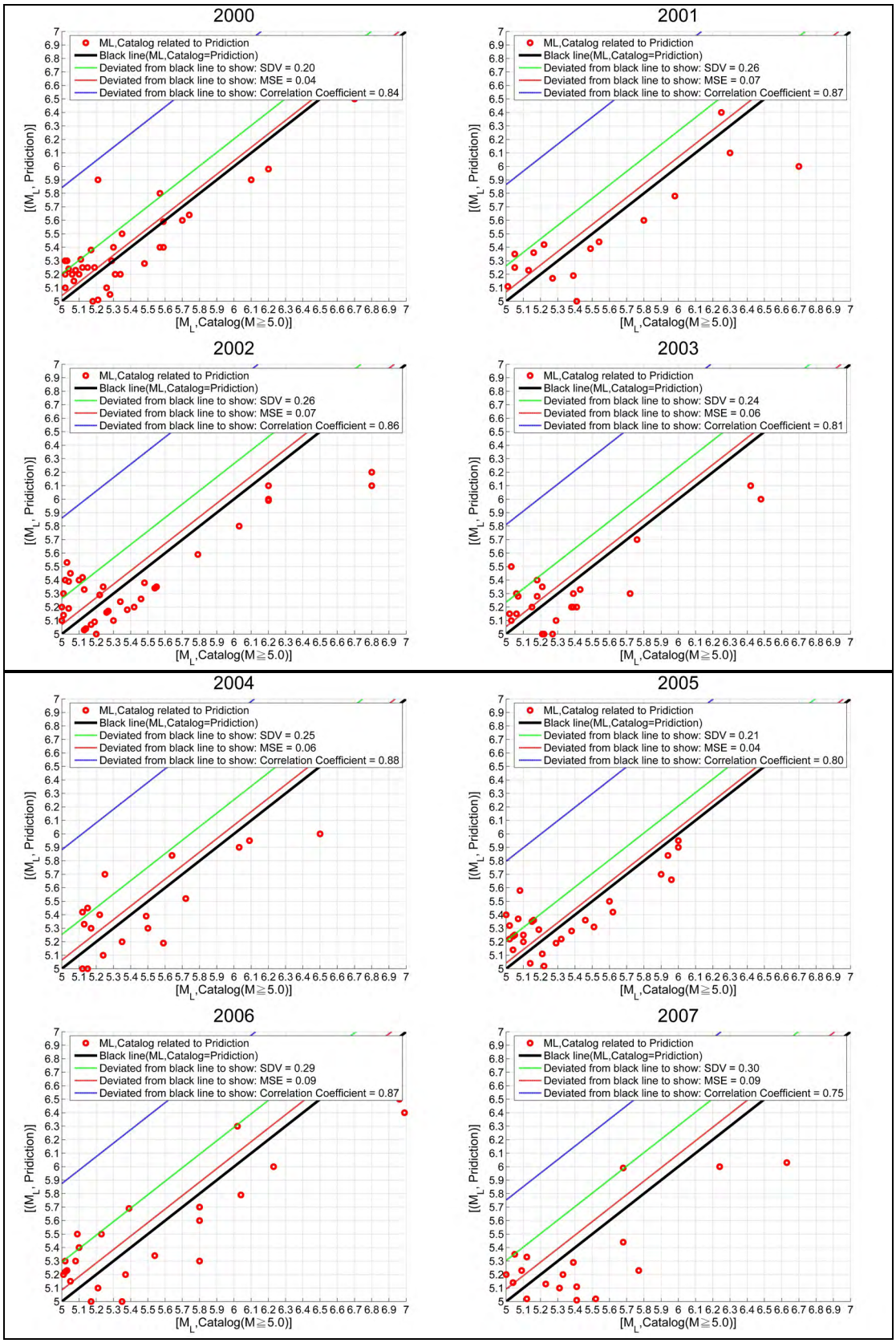


FIGURE 6. The relationship of correlation coefficient, the standard deviation (SDV) and the mean squared error (MSE) between the measured and the predicted Richter magnitudes of earthquakes that occurred from 2000 to 2014 using EEMPBPNN.

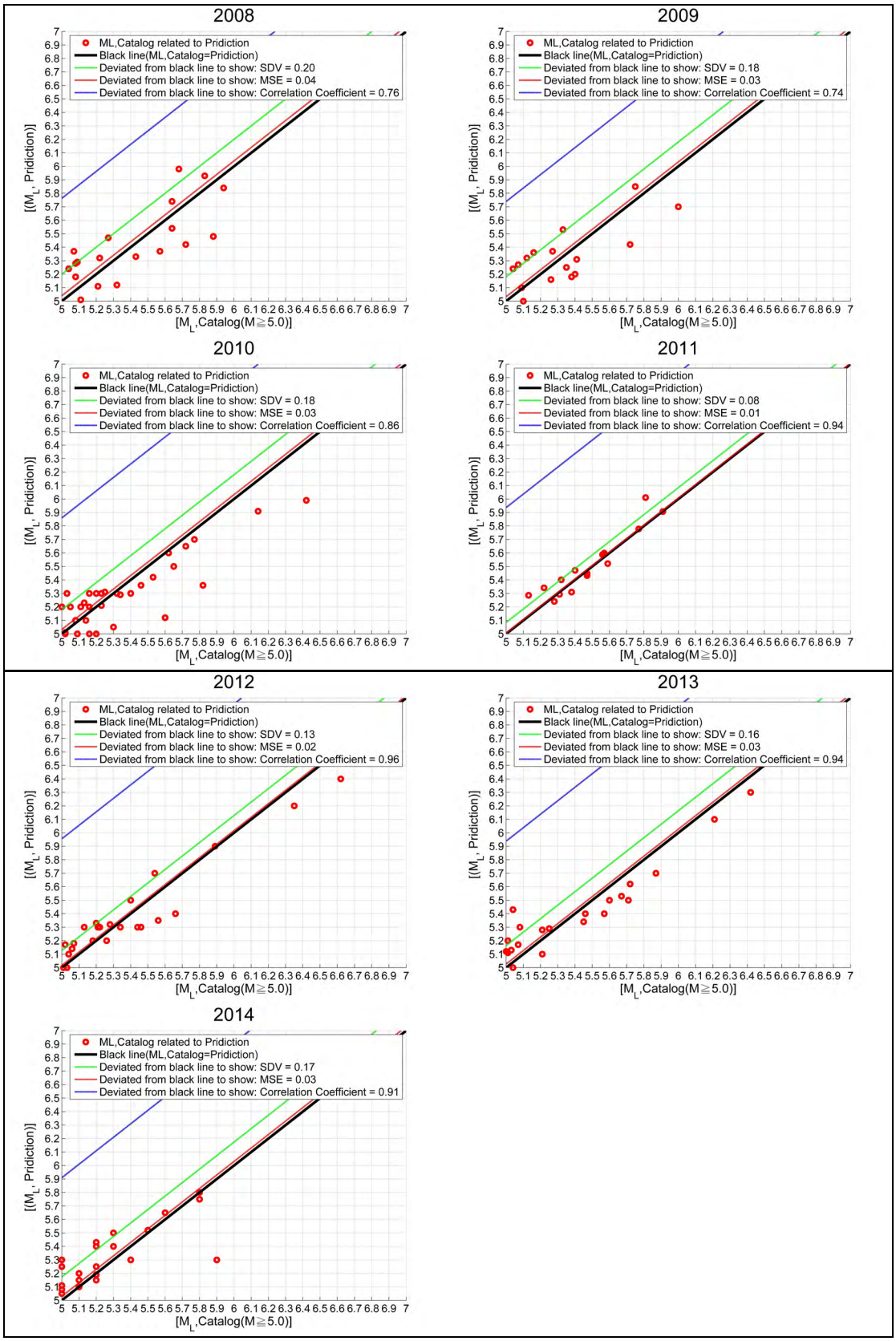


FIGURE 6. *(Continued.)* **The relationship of correlation coefficient, the standard deviation (SDV) and the mean squared error (MSE) between the measured and the predicted Richter magnitudes of earthquakes that occurred from 2000 to 2014 using EEMPBPNN.**

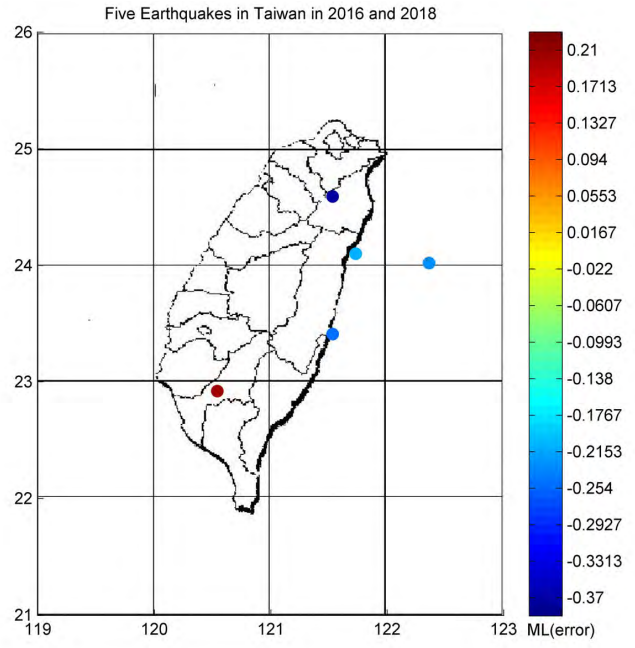


FIGURE 7. Errors in the predicted Richter magnitudes (ML) of five earthquakes using EEMPBPNN. The difference between the predicted value and the measured value is defined as the error.

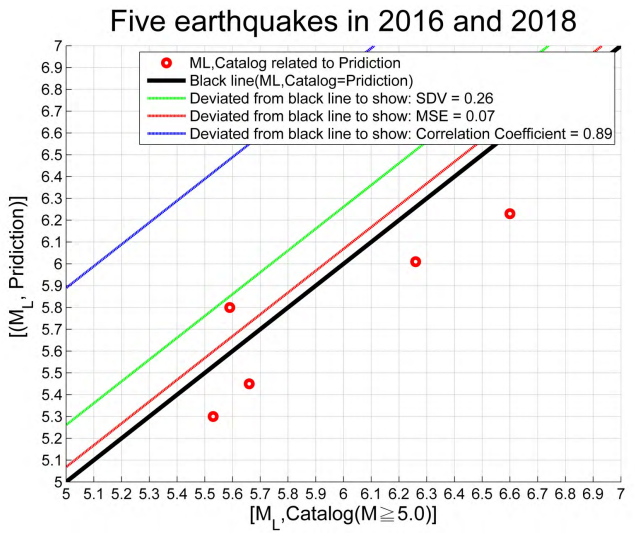


FIGURE 8. The relationship of correlation coefficients, the standard deviations (SDV), and the mean squared errors (MSE) between the measured and the predicted Richter magnitudes (ML) of two earthquakes from Figure 7.

upcoming earthquakes. Thus, in Fig. 6 and Table 1, relatively lower SDV, lower MSE, and higher correlation coefficients confirmed the predictive capability of the EEMPBPNN based on the inside and outside tests. A low SDV, low MSE, and high correlation coefficients indicated the accuracy of the predicted ML of five earthquakes that occurred in 2016 and 2018 (TST), as illustrated in Fig. 8. The results were rounded to the second decimal place. The earthquake positioning system in Taiwan used GPS and was initiated in 1990 [38]. After the 1999 Chi-Chi earthquake, the CWB constructed a GPS

TABLE 1. The list of correlation coefficient, the standard deviation (SDV), and the mean squared error (MSE) between the measured and the predicted Richter magnitude (ML) of earthquakes that occurred from 2000 to 2014 using EEMPBPNN.

Year	Standard Deviation (SDV)	Correlation Coefficient	Mean Squared Error (MSE)
2000	0.20	0.84	0.04
2001	0.26	0.87	0.07
2002	0.26	0.86	0.07
2003	0.24	0.81	0.06
2004	0.25	0.88	0.06
2005	0.21	0.80	0.04
2006	0.29	0.87	0.09
2007	0.30	0.75	0.09
2008	0.20	0.76	0.04
2009	0.18	0.74	0.03
2010	0.18	0.86	0.03
2011	0.08	0.94	0.01
2012	0.13	0.96	0.02
2013	0.16	0.94	0.03
2014	0.17	0.91	0.03

array to enhance its positioning accuracy [77]. Therefore, the earthquake catalogue after 1990 could be used to retrain the BPNN, which resulted in accurate predicted ML, even though the CWB begun to gather earthquake information in catalogues from 1900. As a result, the S_{plate} value was imported as training data as the short-term earthquake catalogue was reliable.

The EEMPBPNN was used to predict upcoming earthquakes without the 2000-2010 earthquake catalogue, therefore, large amounts of training data were not necessary and the computing time could be reduced. In previous studies, when new training data were imported, the BPNN must be retrained with all training data, including the original set, which resulted in a long computing time. Another advantage of this method was that it could determine the number of neurons in each hidden layer using earthquake catalogue without relying on geological environmental features, differing from previous studies, however, the numbers of neurons in the two hidden layers were equally limited in this study. If different numbers of neurons in each hidden layer and increasing numbers of hidden layers are required, further research is necessary to develop a more complicated ANN structure and a method independent of training data, such as seismic strong motion data, to yield results as consistent as those in this study. Even the number of neurons in each hidden layer and the number of hidden layers could be determined through an intelligent optimization training algorithm. Moreover, the number of neurons does not need to be identical. This paper has already reported a new method of using earthquake catalogue to predict the magnitude of upcoming earthquakes without complicated analysis. Reference [67] found that the P wave phase needed to be determined to predict the magnitude of an earthquake when the characterization of a P wave phase may not be easy, and a subnetwork be required [86].

IX. CONCLUSION

An IEMPBPNN with two hidden layers was first constructed using the 2000–2010 earthquake catalogue (ML $\geqq$ 5, depth 300 km). By using the 1990–1999 and

2011–2014 earthquake catalogues as training data (ML $\geqq 5$, depth $\leqq 300$ km), as well as final weights and biases of the IEMPBPNN as initial inputs, an EEMPBPNN was constructed without the 2000–2010 earthquake catalogue. The error in the prediction of the ML by the EEMPBPNN was reduced. Ten neurons were included in each hidden layer. Based on the low SDV, low MSE, and high correlation coefficients, the ML was predicted with the 2000–2010 catalogue as the outside test and the 2011–2014 catalogue as the inside test by the EEMPBPNN. The predicted ML was highly accurate in comparison to a real-time measurement. As a result, this procedure could successfully determine a suitable neuronal number for each hidden layer using earthquake catalogues. The accuracy of the second IEMPBPNN by training with data from the 1990–2014 earthquake catalogue could not be improved with long computing times. When predicting the ML of the five earthquakes in 2016 and 2018 (TST), a low SDV, low MSE, and high correlation coefficient further confirmed reliability of the EEMPBPNN.

ACKNOWLEDGMENT

The authors are grateful to the support of data source of from Central Weather Bureau, Taiwan (CWB). Especially, in Memory of death to the mother of first author, " Lo, Yu-Mei" (8 Nov 1943–9 Oct 2016). Also thank you for the Lover Yung of first author.

REFERENCES

[1] A. S. N. Alarifi, N. S. N. Alarifi, and S. Al-Humidan, "Earthquakes magnitude predication using artificial neural network in northern Red Sea area," *J. King Saud Univ. Sci.*, vol. 24, no. 4, pp. 301–313, 2012.

[2] H. Adeli and A. Panakkat, "A probabilistic neural network for earthquake magnitude prediction," *Neural Netw.*, vol. 22, no. 7, pp. 1018–1024, 2009.

[3] C. Alippi, *Intelligence for Embedded Systems: A Methodological Approach*. Springer, 2014, p. 283. [Online]. Available: https://www.springer.com/la/book/9783319052779, doi: 10.1007/978-3-319-05278-6.

[4] K. M. Asim, F. Martínez-Álvarez, A. Basit, and T. Iqbal, "Earthquake magnitude prediction in Hindukush region using machine learning techniques," *Nat Hazards*, vol. 85, no. 1, pp. 471–486, 2017, doi: 10.1007/s11069-016-2579-3.

[5] A. Atkinson, A. Donev, and R. Tobias, *Optimum Experimental Designs, With SAS*. London, U.K.: Oxford Univ. Press, 2007, p. 528.

[6] M. Caudill, "Neural networks primer, part I," *AI Expert*, vol. 2, no. 12, pp. 46–52, 1987.

[7] W.-Y. Chang, K.-P. Chen, and Y.-B. Tsai, "An updated and refined catalog of earthquakes in Taiwan (1900–2014) with homogenized M_ω magnitudes," *Earth, Planets Space*, vol. 68, p. 45, Mar. 2016, doi: 10.1186/s40623-016-0414-4.

[8] C.-H. Chen *et al.*, "Pre-seismic geomagnetic anomaly and earthquake location," *Tectonophysics*, vol. 489, nos. 1–4, pp. 240–247, 2010, doi: 10.1016/j.tecto.2010.04.018.

[9] S. Chen, F. Cerda, P. Rizzo, J. Bielak, J. H. Garrett, and J. Kovacevic, "Semi-supervised multiresolution classification using adaptive graph filtering with application to indirect bridge structural health monitoring," *IEEE Trans. Signal Process.*, vol. 62, no. 11, pp. 2879–2893, Jun. 2014.

[10] D.-Y. Chen, Y.-M. Wu, and T.-L. Chin, "Incorporating low-cost seismometers into the central weather bureau seismic network for earthquake early warning in Taiwan," *Terr., Atmos., Ocean. Sci.*, vol. 26, no. 5, pp. 503–513, 2015, doi: 10.3319/TAO.2015.04.17.01(T).

[11] Y. C. Chen, "Earthquake prediction via back propagation neural network," M.S. thesis, Dept. Comput. Sci., Nat. Taipei Univ. Educ., Taipei, Taiwan, 2013.

[12] W.-S. Chen *et al.* "Paleoseismology of the Chelungpu fault during the past 1900 years," *Quaternary Int.*, vols. 115–116, pp. 167–176, Jan. 2004.

[13] C.-T. Cheng, S.-J. Chiou, C.-T. Lee, and Y.-B. Tsai, "Study on probabilistic seismic hazard maps of Taiwan after Chi–Chi earthquake," *J. GeoEng.*, vol. 2, no. 1, pp. 19–28, 2007.

[14] M. T. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006, p. 776.

[15] N. Cristescu, "Rock dilatancy in uniaxial tests," *Rock Mech.*, vol. 15, no. 3, pp. 133–144, 1982, doi: 10.1007/BF01238260.

[16] S. D'Amico, *Earthquake Research and Analysis—Seismology, Seismotectonic and Earthquake Geology*. Rijeka, Croatia: InTech, 2012, pp. 125–142, doi: 10.5772/1117.

[17] J. Dhanya and S. T. G. Raghukanth, "Ground motion prediction model using artificial neural network," *Pure Appl. Geophys.*, vol. 175, no. 3, pp. 1035–1064, 2018, doi: 10.1007/s00024-017-1751-3.

[18] J. R. Elliott, R. J. Walters, and T. J. Wright, "The role of space-based observation in understanding and responding to active tectonics and earthquakes," *Nature Commun.*, vol. 7, Dec. 2016, Art. no. 13844, doi: 10.1038/ncomms13844.

[19] C. Fidani, "Statistical and spectral properties of the L'Aquila EQL in 2009," *Bollettino Geofisica Teorica Applicata*, vol. 53, no. 1, pp. 135–146, 2012, doi: 10.4430/bgta0034.

[20] C. Fidani, "Biological anomalies around the 2009 L'Aquila earthquake," *Animals*, vol. 3, no. 3, pp. 693–721, 2013, doi: 10.3390/ani3030693.

[21] J. K. Gardner and L. Knopoff, "Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian," *Bull. Seismological Soc. Amer.*, vol. 64, no. 5, pp. 1363–1367, 1974.

[22] G. Guangmeng and Y. Jie, "Three attempts of earthquake prediction with satellite cloud images," *Nat. Hazards Earth Syst. Sci.*, vol. 13, pp. 91–95, Jan. 2013, doi: 10.5194/nhess-13-91-2013.2013.

[23] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. De Jesús, *Neural Network Design*, 2nd ed. 2014, p. 800. [Online]. Available: http://hagan.okstate.edu/nnd.html

[24] J. Hartmann and J. K. Levy, "Hydrogeological and gasgeochemical earthquake precursors—A review for application," *Natural Hazards*, vol. 34, no. 3, pp. 279–304, 2005, doi: 10.1007/s11069-004-2072-2.

[25] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2008, p. 936.

[26] S. G. Holtkamp, M. E. Pritchard, and R. B. Lohman, "Earthquake swarms in South America," *Geophys. J. Int.*, vol. 187, no. 1, pp. 128–146, 2001, doi: 10.1111/j.1365-246X.2011.05137.x.

[27] S. Hu, S. Xu, D. Wang, and A. Zhang, "Optimization algorithm for Kalman filter exploiting the numerical characteristics of SINS/GPS integrated navigation systems," *Sensors*, vol. 15, no. 11, pp. 28402–28420, 2015, doi: 10.3390/s151128402.

[28] D. R. Hush and B. G. Horne, "Progress in supervised neural networks," *IEEE Signal Process. Mag.*, vol. 10, no. 1, pp. 8–39, Jan. 1993, doi: 10.1109/79.180705.

[29] F. P. Incropera, D. P. DeWitt, T. L. Bergman, and A. S. Lavine, *Fundamentals of Heat and Mass Transfer*, 6th ed. Hoboken, NJ, USA: Wiley, 2006, p. 1024.

[30] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*, 1st ed. Springer, 2013. [Online]. Available: https://www.springer.com/us/book/9781461471370

[31] H. Kim, J. Son, and J. Lee, "A high-speed sliding-mode observer for the sensorless speed control of a PMSM," *IEEE Trans. Ind. Electron.*, vol. 58, no. 9, pp. 4069–4077, Sep. 2011, doi: 10.1109/TIE.2010.2098357.

[32] G. Kopp, G. Lawrence, and G. Rottman, "The total irradiance monitor (TIM): Science results," *Solar Phys.*, vol. 230, nos. 1–2, pp. 129–139, 2005, doi: 10.1007/s11207-005-7433-9.

[33] V. Kecman, *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models*. Cambridge, MA, USA: MIT Press, 2001, p. 608.

[34] H. Larochelle, Y. Bengio, J. Louradour, and P. Lamblin, "Exploring strategies for training deep neural networks," *J. Mach. Learn. Res.*, vol. 1, pp. 1–40, Jan. 2009.

[35] S.-C. Lan, T.-T. Yu, C. Hwang, and R. Kao, "An analysis of mechanical constraints when using superconducting gravimeters for far-field preseismic anomaly detection," *Terr. Atmos. Ocean. Sci.*, vol. 22, no. 3, pp. 271–282, 2011, doi: 10.3319/TAO.2010.11.12.01(T).

[36] K. Lee, H. Kwon, and K. You, "Neural network-based laser interferometer compensation for seismic signal detection," *J. Sensors*, vol. 2018, Oct. 2018, Art. no. 6490861, doi: 10.1155/2018/6490861.

[37] C.-Y. Lin, H.-M. Lo, W.-C. Chou, and W.-T. Lin, "Vegetation recovery assessment at the Jou–Jou mountain landslide area caused by the 921 earthquake in central Taiwan," *Ecological Modelling*, vol. 176, nos. 1–2, pp. 75–81, doi: 10.1016/j.ecolmodel.2003.12.037.

[38] K.-C. Lin *et al.* "GPS crustal deformation, strain rate, and seismic activity after the 1999 Chi–Chi earthquake in Taiwan," *J. Geophys. Res.*, vol. 115, no. B7, p. B07404, 2010, doi: 10.1029/2009JB006417.

[39] J.-W. Lin, "Ionospheric total electron content (TEC) anomalies associated with earthquakes through Karhunen-Loéve transform (KLT)," *Terr. Atmos. Ocean. Sci.*, vol. 21, no. 2, pp. 253–265, 2010.

[40] J.-W. Lin, "Is it possible to trace an impending earthquake's occurrence from Seismo-ionospheric disturbance using principal component analysis? A study of Japan's Iwate–Miyagi Nairiku earthquake on 13 June 2008," *Comput. Geosci.*, vol. 37, no. 7, pp. 855–860, 2011.

[41] J.-W. Lin, "Nonlinear principal component analysis in the detection of ionospheric electron content anomalies related to a deep earthquake (>300 km, M 7.0) on 1 January 2012, Izu Islands, Japan," *J. Geophys. Res.*, vol. 117, no. A6, p. A06314, 2012, doi: 10.1029/2012JA017614.

[42] J. Y. Liu, C. H. Chen, Y. I. Chen, H. Y. Yen, K. Hattori, and K. Yumoto, "Seismo-geomagnetic anomalies and $M \geqslant 5.0$ earthquakes observed in Taiwan during 1988–2001," *Phys. Chem. Earth*, vol. 31, nos. 4–9, pp. 215–222, 2006, doi: 10.1016/j.pce.2006.02.009.

[43] S. Lallemand, T. Theunissen, P. Schnürle, C.-S. Lee, C. S. Liu, and Y. Font, "Indentation of the Philippine Sea plate by the Eurasia plate in Taiwan: Details from recent marine seismological experiments," *Tectonophysics*, vol. 594, pp. 60–79, May 2013, doi: 10.1016/j.tecto.2013.03.020.

[44] T. Kuo, K. Fan, H. Kuochen, Y. Han, H. Chu, and Y. Lee, "Anomalous decrease in groundwater radon before the Taiwan M6.8 Chengkung earthquake," *J. Environ. Radioactivity*, vol. 88, no. 1, pp. 101–106, 2006, doi: 10.1016/j.jenvrad.2006.01.005.

[45] J. J. Litehiser, *Observatory Seismology: A Centennial Symposium for the Berkeley Seismographic Stations*. Berkeley, CA, USA: Univ. California Press, 2018, p. 392.

[46] S. Lutz and W. Huitt, "Information processing and memory: Theory and applications," in *Educational Psychology Interactive*. Valdosta, GA, USA: Valdosta State Univ., 2003.

[47] K.-F. Ma, Y.-Y. Lin, S.-J. Lee, J. Mori, and E. E. Brodsky, "Isotropic events observed with a borehole array in the Chelungpu fault zone, Taiwan," *Science*, vol. 337, no. 6093, pp. 459–463, 2012.

[48] K. E. MacDuffie, A. S. Atkins, K. E. Flegal, C. M. Clark, and P. A. Reuter-Lorenz, "Memory distortion in Alzheimer's disease: Deficient monitoring of short- and long-term memory," *Neuropsychology*, vol. 26, no. 4, pp. 509–516, 2012, doi: 10.1037/a0028684.

[49] G. Madureira and A. E. Ruano, "A neural network seismic detector," *Acta Technica Jaurinensis*, vol. 2, no. 2, pp. 159–170, 2009.

[50] E. Maleki and N. Maleki, "Artificial neural network modeling of Pt/C cathode degradation in PEM fuel cells," *J. Electron. Mater.*, vol. 45, no. 8, pp. 3822–3834, 2016, doi: 10.1007/s11664-016-4718-8.

[51] E. Maleki and A. Zabihollah, "Modeling of shot peening effects on the surface properties of (TiB + TiC)/Ti–6 Al–4 V composite employing artificial neural networks," *Mater. Technol.*, vol. 50, no. 6, pp. 851–860, 2016, doi: 10.17222/mit.2015.140.

[52] E. Maleki and K. R. Kashyzadeh, "Effects of the hardened nickel coating on the fatigue behavior of CK45 steel: Experimental, finite element method, and artificial neural network modeling," *Iranian J. Mater. Sci. Eng.*, vol. 14, no. 4, pp. 81–99, 2017, doi: 10.22068/ijmse.14.4.81.

[53] N. Maleki, S. Kashanian, E. Maleki, and M. Nazari, "A novel enzyme based biosensor for catechol detection in water samples using artificial neural network," *Biochem. Eng. J.*, vol. 128, no. 15, pp. 1–11, 2017, doi: 10.1016/j.bej.2017.09.005.

[54] E. Maleki, O. Unal, and K. R. Kashyzadeh, "Fatigue behavior prediction and analysis of shot peened mild carbon steels," *Int. J. Fatigue*, vol. 116, pp. 48–67, Nov. 2018, doi: 10.1016/j.ijfatigue.2018.06.004.

[55] E. Maleki and G. H. Farrahi, "Modelling of conventional and severe shot peening influence on properties of high carbon steel via artificial neural network," *IJE Trans. B, Appl.*, vol. 31, no. 2, pp. 382–393, 2018.

[56] J. J. Makela *et al.* "Imaging and modeling the ionospheric airglow response over Hawaii to the tsunami generated by the Tohoku earthquake of 11 March 2011," *Geophys. Res. Lett.*, vol. 38, no. 24, p. L00G02, Dec. 2011, doi: 10.1029/2011GL047860.

[57] M. J. Madić and M. R. Radovanović, "Optimal selection of ANN training and architectural parameters using Taguchi method: A case study," *FME Trans.*, vol. 39, no. 2, pp. 79–86, 2011.

[58] K. Madsen, H. B. Nielsen, and O. Tingleff, "Methods for non-linear least squares problems," in *Informatics and Mathematical Modelling*, 2nd ed. Lyngby, Denmark: DTU, 2004, p. 60.

[59] S. Narayanakumar and K. Raja, "A BP artificial neural network model for earthquake magnitude prediction in Himalayas, India," *Circuits Syst.*, vol. 7, no. 11, pp. 3456–3468, 2016, doi: 10.4236/cs.2016.711294.

[60] A. Mignan, "The debate on the prognostic value of earthquake foreshocks: A meta-analysis," *Sci. Rep.*, vol. 4, Feb. 2014, Art. no. 4099, doi: 10.1038/srep04099.

[61] J. Mahmoudi, M. A. Arjomand, M. Rezaei, and M. H. Mohammadi, "Predicting the earthquake magnitude using the multilayer perceptron neural network with two hidden layers," *Civil Eng. J.*, vol. 2, no. 1, pp. 1–12, 2016.

[62] K. Mogi, "Active periods in the world's chief seismic belts," *Tectonophysics*, vol. 22, nos. 3–4, pp. 265–282, 1974, doi: 10.1016/0040-1951(74)90086-9.

[63] M. Moustra, C. Avraamides, and C. Christodoulou, "Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals," *Expert Syst.*, vol. 38, no. 12, pp. 15032–15039, 2011, doi: 10.1016/j.eswa.2011.05.043.

[64] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, San Diego, CA, USA, vol. 3, 1990, pp. 21–26. [Online]. Available: https://ieeexplore.ieee.org/document/5726777/, doi: 10.1109/IJCNN.1990.137819.

[65] Y. Ogata, "Statistical models for earthquake occurrences and residual analysis for point processes," *J. Amer. Statist. Assoc.*, vol. 83, no. 401, pp. 9–27, 1988, doi: 10.2307/2288914.

[66] Z. Peng, C. Aiken, D. K. David, R. Shelly, and B. Enescu, "Listening to the 2011 magnitude 9.0 Tohoku–Oki, Japan, earthquake," *Seismological Res. Lett.*, vol. 83, no. 2, pp. 287–293, 2012, doi: 10.1785/gssrl.83.2.287.

[67] M. Picozzi, D. Bindi, P. Brondi, D. Di Giacomo, S. Parolai, and A. Zollo, "Rapid determination of *P* wave-based energy magnitude: Insights on source parameter scaling of the 2016 Central Italy earthquake sequence," *Geophys. Res. Lett.*, vol. 44, no. 9, pp. 4036–4045, 2017, doi: 10.1002/2017GL073228.

[68] I. Popova, A. Rozhnoi, M. Solovieva, B. Levin, and M. Hayakawa, "Application of neural network methodology in analysis of VLF/LF signal," EMSEV, Gotemba, Japan, Tech. Rep., Oct. 2012. [Online]. Available: http://www.mdpi.com/1099-4300/20/9/691

[69] S. A. Pulinets, "Ionospheric precursors of earthquakes; recent advances in theory and practical applications," *Terr. Atmos. Ocean. Sci.*, vol. 15, no. 3, pp. 413–435, 2004, doi: 10.3319/TAO.2004.15.3.413(EP).

[70] L. K. Read and R. M. Vogel, "Reliability, return periods, and risk under nonstationarity," *Water Resour. Res.*, vol. 51, no. 8, pp. 6381–6398, 2015, doi: 10.1002/2015WR017089.

[71] J. Reyes, A. F. Morales-Esteban, and F. Martínez-Álvarezc, "Neural networks to predict earthquakes in Chile," *Appl. Soft Comput.*, vol. 13, no. 2, pp. 1314–1328, 2013.

[72] C. F. Richter, "An instrumental earthquake magnitude scale," *Bull. Seismological Soc. Amer.*, vol. 25, no. 1, pp. 1–32, 1935.

[73] R. J. Roiger, *Data Mining: A Tutorial-Based Primer*, 2nd ed. London, U.K.: Chapman & Hall, 2016, p. 487.

[74] D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*, vol. 1. Cambridge, MA, USA: MIT Press, 1986, p. 567.

[75] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986, doi: 10.1038/323533a0.

[76] T. Seno, S. Stein, and A. E. Gripp, "A model for the motion of the Philippine Sea plate consistent with NUVEL-1 and geological data," *J. Geophys. Res.*, vol. 101, no. B10, pp. 11305–11315, 1993.

[77] T.-C. Shin, K.-W. Kuo, P.-L. Leu, C.-H. Tsai, and J.-S. Jiang, "Continuous CWB GPS array in Taiwan and applications to monitoring seismic activity," *Terr. Atmos. Ocean. Sci.*, vol. 22, no. 5, pp. 521–533, 2011, doi: 10.3319/TAO.2011.05.18.01(T).

[78] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and its Applications: With R Examples*, 3rd ed. Springer, 2011, p. 596. [Online]. Available: https://www.springer.com/us/book/9783319524511

[79] A. F. Sompotan, N. T. Puspito, E. Joelianto, and K. Hattori, "Analysis of ionospheric precursor of earthquake using GIM-TEC, Kriging and neural network," *Asian J. Earth Sci.*, vol. 8, no. 2, pp. 32–44, 2015.

[80] A. A. Suratgar, F. Setoudeh, A. H. Salemi, and A. Negarestani, "Magnitude of earthquake prediction using neural network," in *Proc. 4th Int. Conf. Natural Comput. (ICNC)*, vol. 2, 2008, pp. 448–452.
[81] W. A. Thomas and C. A. Powell, "Necessary conditions for intraplate seismic zones in North America," *Tectonics*, vol. 36, no. 12, pp. 2903–2917, 2017, doi: 10.1002/2017TC004502.
[82] B. T. Tsurutani W. D. Gonzalez, G. S. Lakhina, and S. Alex, "The extreme magnetic storm of 1–2 September 1859," *J. Geophys. Res.*, vol. 108, no. A7, p. 1268, 2003, doi: 10.1029/2002JA009504.
[83] P. C. Verpoort, P. MacDonald, and G. J. Conduit, "Materials data validation and imputation with an artificial neural network," *Comput. Mater. Sci.*, vol. 147, pp. 176–185, May 2018, doi: 10.1016/j.commatsci.2018.02.002.
[84] D. Wackerly, W. Mendenhall, and R. L. Scheaffer, *Mathematical Statistics with Applications*, 7th ed. p. 944, Pacific Grove, CA, USA: Brooks/Cole, 2008.
[85] N. M. Wagarachchi and A. S. Karunananda, "Towards a theoretical basis for modeling of hidden layer architecture in artificial neural networks," in *Proc. 2nd Int. Conf. Adv. Comput., Electron. Commun. (ACEC)*, 2014, pp. 47–52, doi: 10.15224/.978-1-63248-029-3-73.
[86] Y. M. Wu and T.-I. Teng, "A virtual subnetwork approach to earthquake early warning," *Bull. Seismological Soc. Amer.*, vol. 92, no. 5, pp. 2008–2018, 2002, doi: 10.1785/0120010217.
[87] Y. M. Wu *et al.*, "Coseismic versus interseismic ground deformations, fault rupture inversion and segmentation revealed by 2003 $M\omega$ 6.8 Chengkung earthquake in eastern Taiwan," *Geophys. Res. Lett.*, vol. 33, no. 2, p. L02312, 2006, doi: 10.1029/2005GL024711.
[88] Y. M. Wu *et al.*, "A high-density seismic network for earthquake early warning in Taiwan based on low cost sensors," *Seismological Res. Lett.*, vol. 84, no. 6, pp. 1048–1054, 2013, doi: 10.1785/0220130085.
[89] S.-B. Yu and L.-C. Kuo, "Present-day crustal motion along the Longitudinal Valley fault, eastern Taiwan," *Tectonophysics*, vol. 333, nos. 1–2, pp. 199–217, 2001.
[90] S. X. Zang, Q. Y. Chen, J. Y. Ning, Z. K. Shen, and Y. G. Liu, "Motion of the Philippine Sea plate consistent with the NUVEL-1A model," *Geophys. J. Int.*, vol. 150, no. 3, pp. 809–819, 2002, doi: 10.1046/j.1365-246X.2002.01744.x.
[91] A. Zarola and A. Sil, "Artificial neural networks (ANN) and stochastic techniques to estimate earthquake occurrences in Northeast region of India," *Ann. Geophys.*, vol. 60, no. 6, p. S0658, 2017, doi: 10.4401/ag-7353.
[92] B. Zhao *et al.*, "Is an unusual large enhancement of ionospheric electron density linked with the 2008 great Wenchuan earthquake," *J. Geophys. Res.*, vol. 113, no. A11, p. A11304, 2008, doi: 10.1029/2008JA013613.

JYH-WOEI LIN received the B.Sc. degree from the Department of Physics, Chung Yuan Christian University, Chungli, Taiwan, in 1989, the M.Sc. degree from the Institute of Geophysics, National Central University, Chungli, in 1991, and the Ph.D. degree from the Institut für Geophysik, Technische Universität Clausthal, Clausthal-Zellerfeld, Germany, in 2000. Since 2016, he has been with the Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan. Until 2018, he has about 60 papers, including artificial intelligence, space physics, geophysics, and remote sensing and, especially, four books publications in Germany.

CHUN-TANG CHAO received the B.S. degree from the Department of Electrical Engineering, National Central University, Taiwan, in 1989, the M.S. degree from the Department of Electrical Engineering, National Taiwan University, Taiwan, in 1991, and the Ph.D. degree from the Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 1995.

In 2000, he joined the Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan, where he is currently an Associate Professor. In 2002, he did a short-term research at the Department of Electrical Engineering, University of Washington, Seattle. His research interests include robotic control, intelligent control system, artificial intelligence, and Geophysics.

Dr. Chao received the Annual Paper Award of the Chinese Engineer Society in 1993. During his service at the Industrial Technology Research Institute of Taiwan, his research on thermal transfer color printer received the Gold Medal of Outstanding Achievement Award in 1997.

JUING-SHIAN CHIOU received the B.S. degree from the Department of Electrical Engineering, Feng Chia University, Taiwan, in 1986, the M.S. degree from the Department of Electrical Engineering, National Central University, Taiwan, in 1990, and the Ph.D. degree from the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan, in 2000.

In 2000, he joined the Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan, where he is currently a Distinguished Professor and the Vice-Dean of the College of Engineering and also the Vice-President of the Taiwan Association for Academic Innovation and the Review Committee of the Ministry of Science and Technology, Taiwan.

He has over 60 articles. His research interests include artificial intelligence, fuzzy control, unmanned aerial vehicles, algorithm, large-scale systems, hybrid systems, and Geophysics.