Predicting ionospheric precursors before large earthquakes using neural network computing and the potential development of an earthquake early warning system

Jyh-Woei Lin1,2

Received: 18 February 2022 / Accepted: 7 April 2022 / Published online: 3 May 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract
Total electron content (TEC) precursors of the Chi-Chi Earthquake, which occurred in Taiwan at 01:47:15 on September 21, 1999 (Taiwan Standard Time, TST), with its epicenter at 23.85° N and 120.82° E, a Richter magnitude (ML) of 7.3 (a moment magnitude (Mw) of 7.6), and a focal depth of 8.00 km, were detected 1, 3, and 4 days before the earthquake using two back-propagation neural network (BPNN) models. These results are consistent with the analysis results of Liu et al. (2001) and Lin (2010). Another TEC precursor was detected on May 13, 2003 (TST), 2 days before the earthquake on May 15, 2003 (TST), with an ML of 5.21. Their precursors might be induced by the variations of the geomagnetic and electric fields near their hypocenter. The two BPNN models were to be verified for stability and reliability after performing cross-validation and evaluating the variance in the learning process. For this, the result of the analyzed method can serve as a real-time TEC predicted system for giving any future time as inputs of the two BPNN models. Hence, an earthquake early warning (EEW) system is performable.

Keywords Total electron content (TEC) precursors · Chi-Chi earthquake · Back-propagation neural network (BPNN) · Earthquake early warning (EEW) system

1 Introduction
Many recent studies have focused on the detection of ionospheric anomalies as precursors to large earthquakes according to the Lithosphere–Atmosphere–Ionosphere Coupling (LAIC) model (Pulinets and Ouzounov 2011; Marchetti et al. 2020). In this study, Total Electron Content (TEC) is the value obtained from the vertical integration of the electron density, and it is a critical indicator of the status of the ionosphere, also called vertical TEC (VTEC). The unit of the TEC is TECu (1 TECu = 10¹⁶ e/m²), and the symbols ‘e’ and ‘m’...
denote electrons and meters, respectively (Mannucci et al. 1998; Cooper et al. 2019). The calculation of the slant TEC from the vertical TEC is an essential research topic (Gautam et al. 2018; Cooper et al. 2019). However, it is not within the scope of this study.

In recent years, one concept has become particularly attractive: artificial intelligence (AI). Several researchers have already applied the concept of AI to detect TEC anomalies associated with earthquakes. Kalita et al. (2012) used a peak detection algorithm, an automatic pattern matching approach, and an artificial neural network (ANN) to process GPS data, generate the TEC record, and detect the TEC precursor of an impending earthquake in Guwahati, India, with an epicenter at 26.10° N and 91.45° E. Homam (2016) presented TEC anomalies related to earthquakes using an ANN, GPS ionospheric scintillation, and a TEC monitor (GISTM) receiver from 2005 to 2010 (Coordinated Universal Time, UTC) during low- to medium solar activity, with sunspot numbers (SSNs) of 0.0–42.6. To reduce the predicted errors, Okoh et al. (2016) used a type of ANN model for the Global Navigation Satellite System (GNSS) and the vertical TEC (GNSS-VTEC) in Nigeria. Akhoondzadeh (2019) applied a neural network to detect the TEC precursors related to the M_w = 8.2 Mexico earthquake at 04:49:19 (UTC) on September 8, 2017, with an epicenter at 15.022° N and 93.899° W and a focal depth of 47.40 km. Tariq et al. (2019) employed an International Reference Ionosphere (IRI-2016) TEC model with GPS-TEC measurements in different regions of Pakistan during solar cycle 24 (2015–2017) (UTC). The measured TEC diurnal variability was estimated from the maximum (06:00–12:00 UTC) to the minimum (during 21:00–24:00 UTC).

Based on the results of previous studies researchers have indicated complicated computational processing using a large number of ionospheric resources. Two factors that are not associated with earthquakes have been simultaneously considered. They are ionospheric anomalies, such as the equatorial ionization anomaly (EIA); and space weather information, such as seasonal variations, solar wind, diurnal variations, SSNs, and the geomagnetic storm (Homam 2016; Kumar et al. 2016; Tsagouri et al. 2018; Hudson et al. 2021). Separating the earthquake-associated TEC anomalies from the aforementioned two factors has been an important procedure. However, this has complicated the data processing and required a long computation time to exclude the two factors. Moreover, in some previous studies, two different methods have been simultaneously used to predict the TEC precursors, and thus, they have simultaneously combined two different methods with a large amount of data, making the analysis procedure complex and the economic cost high.

2 Overview of past related studies

Liu et al. (2001) examined TEC data to detect the TEC precursors 1, 3, and 4 days before the Chi-Chi Earthquake. Liu et al. (2006) found that the TEC precursors of earthquakes with a Richter magnitude (ML) ≥ 5.0 were detectable by examining the earthquakes in Taiwan from 1988 to 2001 (UTC) using a statistical approach. They found that the TEC precursors of earthquakes with an ML < 5.0 could not be detected. Lin (2010) examined the TEC data using principal component analysis (PCA) to detect the TEC precursors before the Chi-Chi Earthquake on September 17, 18, and 20, 1999, at 01:47:15 (Taiwan Standard Time, TST), with its epicenter at 23.85° N and 120.82° E, an ML of 7.3 (a moment magnitude (Mw) of 7.6), and a focal depth of 8.00 km on September 21, 1999. The results of previous studies were examined to verify the reliability and correctness of the analyzed method in this study.
3 Objectives of this study

This study mainly focused on researching the TEC precursors of earthquakes using two back-propagation neural network (BPNN) models before large earthquakes with an ML ≥ 5.0, with a few TEC datasets as the training datasets. The source of TEC datasets is introduced in the next section. The training datasets were the 24-day TEC records of different seasons and months owing to the TEC seasonal and monthly variants (Ogwala et al. 2019). The 24-day TEC records are called the short-period TEC records (SPTRs) in the plural form. Similarly, the abbreviation of the short-period TEC record is the SPTR in the singular form. They are the VTEC. There were only two earthquakes with an ML ≥ 5.0 in the periods examined, that is, the Chi-Chi Earthquake on September 21, 1999 (TST), and an earthquake on May 15, 2003 (TST).

Therefore, for the first objective, it is expected that the TEC precursors of these two earthquakes can be detected by predicting the TEC situation using the two BPNN models. A large predicted error can be an indicator for the TEC precursors, according to the results of the study of Lin et al. (2018). For the second objective, any TEC precursor is expected to be detected relating to the earthquakes with an ML ≥ 5.0. For the third objective, it is expected that the analyzed method can serve as a real-time TEC predicted system for giving any future time as inputs of the two BPNN models by controlling the computation time using appropriate computing hardware. Thus, an earthquake early warning (EEW) system is performable. The data processing in training the two BPNN models is not complicated, which reduces the economic cost owing to the use of a few SPTRs as training datasets. These are the distinguishing advantages of the analyzed method in this study, which can simplify the data processing and thereby overcome the shortcomings of methods used in previous studies.

The SPTRs are a set of small datasets. Zakeri et al. (2020) emphasized that the cross-validation is indispensable for ensuring the reliability and robustness of ANN models with small training datasets to avoid overfitting. Many studies have also focused on this issue (Pasini 2015; Bhadra et al. 2017; Baquirin and Fernandez 2018; Hemmerich et al. 2020; Seddiki et al. 2020; Zakeri et al. 2020; Jiang et al. 2021). Therefore, the cross-validation is an important step for meeting the objectives of this study. The SPTRs, which were the training datasets, were predicted to verify and validate the reliability and robustness of the two BPNN models (i.e., the inside test), and other SPTRs that were not training datasets (i.e., test datasets), were predicted to test the predicted ability of the two BPNN models (i.e., the outside test). Following the cross-validation, standard deviation (STD) and mean square error (MSE) were used as statistical approaches to evaluate the reliability of the two BPNN models. The predicted results were compared with those of Liu et al. (2001), Liu et al. (2006), and Lin (2010) to verify the consistency of the two BPNN models.

4 Data sources

In this study, the SPTRs were classified into five SPTRs with their corresponding earthquake catalogs. The five SPTRs with different time periods were selected due to their long-term seasonal variations. Therefore, they could serve as the training datasets to build the two suitable BNNN models. The corresponding earthquake catalogs consisted of the earthquakes with an ML ≥ 4.0 (CWB). Three SPTRs were the training datasets. Two SPTRs were the test datasets.
The SPTRs (sampling per 15 min) were acquired from Chung-Li GPS station (25.00° N and 121.20° E). These were the records in the F region of the Ionosphere, at a height of approximately 150–800 km. The features and scale of the examined ionospheric TEC records for Taiwan were measured at this station, called Chung-Li Ionosonde. The Chung-Li Ionosonde monitors the TEC variants within a radius of 500 km from the Chung-Li GPS station, which covers GPS network receiver stations in the region of 15–30° N and 110–130° E using the very-high-frequency (VHF) radar echoes emitted by the Chung-Li VHF radar station (24.91° N, 121.24° E) (refer to Fig. 1 of Liu et al. 2004). This means that the TEC records of the Chung-Li Ionosonde can reflect the TEC variants of Taiwan (Lin et al. 2016). The Chung-Li Ionosonde was built by the Center for Space and Remote Sensing Research (CSRSR), National Central University (NCU), Taiwan. Therefore, Chung-Li Ionosonde monitors the ionospheric situation in Taiwan. The five SPTRs are listed below.

1. The first SPTRs were from September 1 to 15 and from September 22 to 30, 1999 (Fig. 1a).
2. The second SPTRs were from January 1 to 24, 2003, and from February 1 to 24, 2003 (Fig. 1b, c, respectively).
3. The third SPTRs were from March 1 to 24, 2003, and from April 1 to 24, 2003 (Fig. 1d, e, respectively).
4. The fourth SPTR was from May 1 to 24, 2003 (Fig. 1f).
5. The fifth SPTR was from 01 to 24 October 2002 (Fig. 1g).

The visual representations of their corresponding earthquake catalogs in Taiwan with an ML ≥ 4.0 are shown in Figs. 2, 3, and 4.

5 Algorithm of the back-propagation neural network

Several studies and engineering applications have revealed that a neural network consisting of two hidden layers with few neurons can replace a network with numerous neurons in a hidden layer (Thomas et al. 2016). Therefore, in Fig. 5, a four-layered BPNN model with two hidden layers is employed similarly as that in the study of Lin et al. (2018). This BPNN model is based on the multilayer perceptron (MLP) neural network with an error back-propagation (EBP) algorithm. The EBP algorithm is a type of supervised learning algorithm that solves several nonlinear and complex problems (Rumelhart et al. 1986; Khoshgoftaar et al. 2010). Its training datasets are called label signals and include the training input and corresponding target output (Belo et al. 2017). A BPNN model includes an activation function and two parameters. The two parameters are the weight and bias.

A modified elementary Levenberg–Marquardt Algorithm (M-LMA) served as the EBP algorithm in this study (Lin et al. 2018). The M-LMA is the best method for general nonlinear problems, in which rank-deficient nonlinear least squares are encountered; that is, when ill-conditioned data are encountered, such as the nonlinear TEC and seismic data, the global minimum can be easily accessed after a single completed iteration. The BPNN model was constructed in the following equation:

\[
TEC^4_{Output} = f_i^3(W_i^3)j^2(W_j^2)k^1(R_i + b_j^1) + b_k^2) + b_1^3)
\]

(1)
Fig. 1 The five SPTRs for the following periods (TST): a From September 1 to 30, 1999. The red line indicates the timing of the Chi-Chi Earthquake at 01:47:15 on September 21, 1999. b From January 1 to 24, 2003. c From February 1 to 24, 2003. d From March to 24, 2003. e From April 1 to 24, 2003. f From May 1 to 24, 2003. g From October 1 to 24, 2002. The horizontal scale from 1 to 2 indicates the SPTR on September 1, 1999 as an example. The SPTRs were acquired from the Center for Space and Remote Sensing Research (CSRSR), National Central University (NCU), Taiwan.
Fig. 1 (continued)
Fig. 1 (continued)
where $R_I$ is the training input; and $W$, $b$, and $f$ are the weight, bias, and activation function, respectively. TEC\textsuperscript{\textsubscript{Output}} is the target output. According to the framework of the BPNN model in Fig. 5 and the superscripted variables in Eq. (1), following the input layer, layer 1 is the first hidden layer (i.e. Hidden Layer 1 in Fig. 5). Following the first hidden layer, layer 2 is the second hidden layer (i.e. Hidden Layer 2 in Fig. 5). Following the output layer, the output is described as the temporary output for each epoch using the EBP algorithm. The subscripts $J$ and $K$ indicate the number of neurons in each hidden layer. They were the true TEC values in this study. The superscripts indicate the number of layers.
Fig. 3 Visual presentation of the earthquake catalogs with an ML ≥ 4.0 for the examined period from January 1, 2003 to May 31, 2003 (TST). The numbers on the top of the orange rectangles (the right side of the vertical black line) indicate the number of earthquakes with an ML ≥ 5.0 (CWB).

Fig. 4 Visual presentation of the earthquake catalogs with a Richter magnitude (ML) ≥ 4.0 for the examined period from October 1 to 31, 2002 (TST). The numbers on the top of the orange rectangles (the left side of the vertical black line) indicate the number of earthquakes with an ML between 4.0 and 5.0 (CWB). There are no earthquakes with an ML ≥ 5.0 in this period.

Fig. 5 Framework of a four-layered BPNN model. Following the output layer, the output on the right side is described as the temporary output for each epoch using the EBP algorithm (Rumelhart et al. 1986).
sigmoid function was used as an activation function, as shown in the following equation (Schmidhuber 2015):

\[ f(t) = \frac{1}{1 + e^{-t}} \]  

(2)

6 Training procedure of the two back-propagation neural network models

The training procedure of the two BPNN models was as follows:

1. We let each hidden layer have 10 neurons \( (J = K = 10) \) in Fig. 5; then, Eq. (1) becomes:

\[
TEC_{\text{Output}}^d = f^3(W_{10}^3f^2(W_{10}^2f^1(W_{10}^1R_I + b_{10}^1) + b_{10}^2) + b_{10}^3)
\]

(3)

where the target output \( TEC_{\text{Output}}^d \) is the real TEC value (TECu) (Moses et al. 2020).

2. The SPTRs with the same dimension (i.e., 24 days) with three SPTRs (i.e., the first, second, and third SPTRs) were selected as the training inputs, which is a matrix, in training a BPNN model according to the concept of the algorithm for training a BPNN model (Rumelhart et al. 1986). Therefore, the training inputs \( R_I \) in Eq. (3) were as follows:

\[
R_I = \begin{bmatrix}
TEC - \text{Record} - \text{Time(first)} \\
TEC - \text{Record} - \text{Time(second)} \\
TEC - \text{Record} - \text{time(third)}
\end{bmatrix}
\]

(4)

3. The M-LMA and the sigmoid function in Eq. (2) were used as the EBP algorithm and the activation function in training a BPNN model. The training epoch was set as adaptive (Mocanu et al. 2018). In this study, after testing, we performed more than 1000 training epochs to achieve a similar training error. Therefore, performing more than 1000 training epochs was already meaningless, and more computing time was necessary.

4. The learning rate was adaptive based on the concept used by Ede and Beanland (2020).

5. The initial random weights and the initial biases (Schmidhuber 2015; Nadia et al. 2017) were within the range of 0 and 1.

6. The magnitudes of all of the components of \( R_I \) were normalized in the range of 0–1 via feature scaling (Bo et al. 2006).

7. The training error was defined as the target output minus the temporary output (described as output in Fig. 5) for each epoch using the EBP algorithm. The unit of this error is TECu.

8. A BPNN model was built at the best learning rate of 0.11, with the minimal training error occurring after finishing 1000 training epochs. This BPNN model is called the first BPNN model.

9. The predicted error was defined as the ‘predicted TEC value minus the target output’. The unit of this error is also TECu.
According to the same procedure in training the first BPNN model, a second BPNN model was built using the SPTRs with the same dimension (i.e., 24 days), as the training inputs, with three SPTRs (i.e., the second, third, and fifth SPTRs). Therefore, the training inputs $R_i$ in Eq. (3) were as follows:

$$R_i = \begin{bmatrix} \text{TEC} - \text{Record} - \text{Time(second)} \\ \text{TEC} - \text{Record} - \text{Time(third)} \\ \text{TEC} - \text{Record} - \text{time(fifth)} \end{bmatrix}$$

(5)

The second BPNN model was built at the best learning rate at 0.23, with the minimal training error occurring after finishing 1000 training epochs. After testing, we performed more than 1000 training epochs to achieve a similar training error. Therefore, performing more than 1000 training epochs was already meaningless, and more computing time was necessary. Building two BPNN models, the epoch was 1000; the similar data character of the training data might be the reason.

The training dataset SPTRs were predicted to verify and validate the reliability and robustness of the two BPNN models (inside test). The test dataset SPTRs were predicted to test the predicted ability of the two BPNN models (outside test). Therefore, the cross-validation of the two BPNN models was performed.

Following step 11, the flowchart for training the two BPNN models was created, as shown in Fig. 6.

### 7 Results

The five SPTRs were predicted using the first BPNN model. The predicted results are shown in Fig. 7. Figure 7a shows the predicted results of the SPTR for September 1999, and the predicted results for the period from September 16 to 21 are the outside test. Figure 7b shows the SPTR for January 2003, and Fig. 7c shows the SPTR for February 2003. Figure 7d shows the predicted results of the SPTR for March 2003. Figure 7e shows the predicted results for the SPTR for April 2003. Figure 7f shows the predicted results for the SPTR for May 2003. Figure 7g shows the predicted results for the SPTR for October 2002. From these predicted results, the large predicted errors occur on September 17, 18, and 20, 1999 in Fig. 7a and on May 13, 2003, in Fig. 7f.

The first and fourth SPTRs for September 1999 and October 2002 were predicted using the second BPNN model. The predicted results for September 1999 and May 2003 are shown in Figs. 8 and 9. From these predicted results, the large predicted errors occur on September 17, 18, and 20, 1999 in Fig. 8 and on May 13, 2003, in Fig. 9. Referring to the captions of Figs. 7, 8, and 9, the threshold of the predicted error is 11.91 TECu in Fig. 8, which is the maximum value of the predicted errors (Note: predicted errors for September 17, 18, and 20, 1999, and May 13, 2003, were not evaluated). In this study, under the condition that the predicted error is greater than and equal to 11.91 TECu, the predicted error was identified as a TEC precursor (Lin et al. 2018). The predicted results of the two BPNN models were compared with the results of past related studies in Sect. 2 to verify the reasonability of the selected threshold, that is, 11.91 TECu.
Fig. 6  Flowchart for training the two BPNN models

**Start**

- Determining the framework of two BPNN Models with Two hidden layers. Each hidden layer has 10 Neurons

- Determining the entire training data (A training input vector with a related target output)

- Modified elementary Levenberg–Marquardt (M-LMA) Algorithm serves as the error back propagation (EBP) algorithm. Sigmoid function is used as an activation function

- The training epoch is set as adaptive

- The learning rate is adaptive

- The initial weights and biases are set to random variables within the range of 0 and 1

- Normalized the training inputs in the range at 0–1 with feature scaling

- The training error with unit TECu is defined as the Target Output minus the temporary output for each epoch using the EBP algorithm

- The two BPNN models are built with best learning rate and minimal training error after finishing 1000 training epoch. The cross-validation is performed

- The predicted error with unit TECu is defined as “predicted TEC value subtracting Target Output”

**The large predicted error**

- **No**
  - Non-TEC precursors

- **Yes**
  - (1) TEC precursors.
  - (2) A real-time TEC predicted system
  - (3) An EEW system is preformable

**End**
8 Discussion

As stated previously, the STD and MSE (with the unit: TECu) were used as statistical approaches to evaluate the ability application of the two BPNN models. For an excellent and reasonable predicted BPNN model, the accuracy of the inside test should be better than that of the outside test (Jiang et al. 2021). Combining the predicted results of subfigures in Fig. 7, the results, which are inside tests, were entirely examined using a set of the STD and MSE values as an evaluation. In contrast, the results, which are outside tests, were entirely examined using another set of the STD and MSE values as an evaluation. Therefore, the training dataset SPTRs were predicted to verify and validate the reliability and robustness of the first BPNN model, with STD = 0.24 and MSE = 0.38 (inside test). The test dataset SPTRs were predicted to test the predicted ability of this BPNN model, with STD = 0.36 and MSE = 0.49 (outside test). Evaluating the results of the inside and outside tests using the previous STD and MSE values, the accuracy of the inside test was better than that of the outside test. The low STD and MSE values confirm that the first BPNN model was an excellent and reasonable predicted BPNN model. As shown in Fig. 8, the SPTRs (i.e., test datasets) were predicted to test the predicted ability of the second BPNN model with STD = 0.59 and MSE = 0.79 (outside test). In Fig. 9, the SPTRs (i.e., test datasets) were predicted to test the predicted ability of the second BPNN model with STD = 0.54 and MSE = 0.71 (outside test). The low STD and MSE values confirm the reliability of the second BPNN model. The results show the reasonable application of the cross-validation for the two BPNN models. The TEC precursors of the Chi-Chi Earthquake were successfully predicted using the two BPNN models. As was previously stated, when the predicted error was greater than 11.91 TECu, it was a TEC precursor. Therefore, the precursors of the 21 September Chi-Chi Earthquake were detectable 1, 3, and 4 days before the Chi-Chi Earthquake, that is, on September 17, 18, and 20, respectively. These results are consistent with the results of Liu et al. (2001) and Lin (2010), and they verify the reliability of the two BPNN models. The predicted errors were large on 13 May 2003, indicating a TEC precursor before the earthquake on May 15, 2003, with an ML = 5.21. Therefore, the two BPNN models had a reasonable predicted ability. They produced results that are consistent with the results of past related studies in Sect. 2. Thus, it was shown that determining the TEC precursors via a threshold, that is, 11.91 TECu, was reasonable.

As with the statistical approach of Liu et al. (2006), this study confirmed that the TEC precursors of earthquakes with an ML < 5.0 were not detectable. Therefore, the two BPNN models were reliable and suitable for predicting TEC precursors before large earthquakes. The predicted errors from September 22 to 24, 1999, were small; the possible reason for this may be that the variations in the TEC data were not associated with the aftershocks of the Chi-Chi Earthquake. By controlling the computation time using appropriate computing hardware, the analyzed method can serve as a real-time TEC predicted system for giving any future time as inputs of the two BPNN models. Thus, an EEW system is performable. Simultaneously, as stated previously, the data processing is not complicated, which reduces the economic cost. Naturally, the results of this study demonstrate that the duration of the TEC precursors, which were defined with large TEC predicted errors, was difficult to evaluate. However, under this condition, i.e., a large TEC predicted error that is larger than and equal to the selected threshold, the EEW system begins to monitor the variants of TEC, and then an alarm is broadcast. Thus, such EEW system is already well significant in Taiwan.
Finally, after performing cross-validation to avoid overfitting the two BPNN models, determining the variance in the learning process, i.e., the bias–variance tradeoff, was also another important step taken to validate the possibility of the overfitting or underfitting of the two BPNN models. Usually, a stable BPNN model with an appropriate tradeoff bias is used to let the predicted datasets of this BPNN model closely match its training datasets. In other words, the training error should be small, and the difference between the training error and the prediction error should also be small to avoid overfitting or underfitting (Belkin et al. 2019; Doroudi 2020). Therefore, we analyzed the accuracy of the training errors and predicted errors of the two BPNN models so as to validate the two BPNN models, which were not the overfitting and underfitting models. First, two training parameters—three ranges of the initial weight and bias—were selected. The ranges were 0–1, 1–2, and 2–3. Three pairs of BPNN models were built using initial weight and bias, which were random variables in the three ranges, using the same training datasets and procedures detailed in Sect. 6. Each pair included the first and second BPNN models. The MSE values were used to evaluate the training and predicted accuracy for the three pairs of BPNN models. Figure 10 shows the MSE variants of training errors with the initial weight and bias, which were randomly selected in the three different ranges, for the three pairs of BPNN models. The BPNN models with high training accuracy had low initial weight and bias. After predicting the training datasets with the three pairs of BPNN models, the MSE variants of predicted errors for the three pairs of BPNN models are shown in Fig. 11. The BPNN models with high predicted accuracy had low initial weight and bias. Based on Figs. 10 and 11, the two BPNN models trained in Sect. 6 were stable models because the training error was small, and the difference between the training error and the prediction error was also small as stated previously. The initial weight and bias, which were randomly selected within the range of 0–1, were at the optimal tradeoff. Therefore, these initial weight and bias, which fitted the value of the sigmoid function and were randomly selected within the range of 0–1, were at the optimal tradeoff.

It is noted that, although both models were stable after testing, the training and prediction errors of the second BPNN model were better than those of the first BPNN model. For this, the two BPNN models were built with the initial weight and bias, which were randomly selected within the range of 0–1 as in Sect. 6, with different training epochs. The MSE variants of training errors with different training epochs are shown in Fig. 12. The training accuracy of the second BPNN model was better that of the first model in training process. This should be the reasonable verification for the second BPNN model with more stability compared to the first BPNN model. More training epochs increased the accuracy of the training process. The accuracy when using 1250 epochs was a little better than that when using 1000 epochs. However, more computing time was necessary when using 1250 epochs for two BPNN models as stated previously in Sect. 6.
Fig. 7 (continued)
Fig. 7 (continued)
Predicted results using the second BPNN model (i.e., the outside test) for September 1999 (TST). The maximum value of the predicted error without September 17, 18, and 20 is 11.91 TECu
Fig. 9  Predicted results using the second BPNN model (i.e., the outside test) for May 2003 (TST). The maximum value of the predicted error without May 13 is 11.89 TECu. A TEC precursor was detected on May 13, 2003, before the May 15, 2003 earthquake

Fig. 10  The MSE variants of training errors with different ranges of the initial weight and bias. The initial weight and bias were randomly selected within different ranges in training the three pairs of BPNN models. Each pair includes the first and second BPNN models. The ranges of initial weight and bias were scaled in the horizontal coordinate. The scale at one indicates the range of initial weight and bias, from 0 to 1. The scale at two indicates the range of initial weight and bias, from 1 to 2. The scale at three indicates the range of initial weight and bias, from 2 to 3.
Regarding the mechanical principles that induce the ionospheric earthquake-associated TEC anomalies, relevant studies have described three possibilities, as follows. (1) The gravity and acoustic waves, for example, the co-seismic acoustic waves generated before earthquakes, cause ionospheric anomalies. When an earthquake occurs, the fine low-frequency vibrations of the ground surface induce the ionospheric TEC anomalies. (2) The

Fig. 11  The MSE variants of predicted errors for the three pairs of BPNN models. The scale in the horizontal coordinate has the same definition as that in Fig. 10

Fig. 12  The MSE variants of training errors with different training epochs for the two BPNN models

Regarding the mechanical principles that induce the ionospheric earthquake-associated TEC anomalies, relevant studies have described three possibilities, as follows. (1) The gravity and acoustic waves, for example, the co-seismic acoustic waves generated before earthquakes, cause ionospheric anomalies. When an earthquake occurs, the fine low-frequency vibrations of the ground surface induce the ionospheric TEC anomalies. (2) The
crustal chemistry theory suggests that before earthquakes, the ground vibrates and radon gases are released from the crust, causing ionospheric TEC anomalies. (3) The anomalous rock and magmatic activities cause variations of the geomagnetic and electric fields near the hypocenter. The rocks deform, and the magma flow slows before a large earthquake; thus, geomagnetic and electric field anomalies are caused, resulting in ionospheric anomalies (Liu et al. 2001; Freund 2003; Pulinets 2004; Kamogawa and Kakinami 2013; Oyama et al. 2016; Pulinets et al. 2016; Fu et al. 2017). However, the precursors of the two earthquakes, which were researched in this study, might be caused by the fluctuations of the geomagnetic and electric fields close to their hypocenter (Liu et al. 2001).

9 Conclusions

The first BPNN model was trained using three SPTRs, that is, the first, second, and third SPTRs. The second BPNN model was trained using three SPTRs, that is, the second, third, and fifth SPTRs. Cross-validation was performed to avoid the overfitting of the two BPNN models because their training data (i.e., SPTRs) were small datasets. The low STD and MSE values prove the predicted abilities of the two BPNN models. The results show that the precursors of the September 21 Chi-Chi Earthquake were detected 1, 3, and 4 days before the Chi-Chi Earthquake, that is, on September 17, 18, and 20, 1999. These results are consistent with the results of Liu et al. (2001) and Lin (2010). Another TEC precursor was detected on May 13, 2003, 2 days before the earthquake on May 15, 2003, with an ML = 5.21. The precursors of the two earthquakes might be caused by the differences in the geomagnetic and electric fields close to their hypocenter. No TEC precursors were detected for the earthquakes with an ML < 5.0, as with the statistical approach of Liu et al. (2006). After cross-validation and evaluation of the variance in the learning process, the two BPNN models were determined to be reliable and stable models. For this, this analyzed method can serve as a real-time TEC predicted system for giving any future time as inputs of the two BPNN models. Hence, an EEW system is performable. Simultaneously, the data processing is not complicated, which reduces the economic cost.

Acknowledgements The author Jyh-Woei Lin (German name: Stefan Lin) is grateful to the Central Weather Bureau of Taiwan (CWB) for their useful references Data. The author is also grateful to China and the Center for Space and Remote Sensing Research (CSRSR), National Central University (NCU). Taiwan. https://www1.csrsr.ncu.edu.tw/homepage.php.

Author contributions J-WL: designed the research, performed the research, analyzed the data, and wrote the paper.

Funding statement The authors have funding from Binjiang College, Nanjing University of Information Science & Technology, Wuxi, 214105, Jiangsu Province, China.

Data Availability The earthquake catalogs can be obtained from Geophysical Database Management System (GDMS) of Central Weather Bureau of Taiwan (CWB) (an account with a password is necessary, and the earthquake location and parameters are revised by experts, so after corrections, they are higher in accuracy than preliminary estimations). https://gdms.cwb.gov.tw/index.php?t=1648023901. The TEC data can be required from the Center for Space and Remote Sensing Research (CSRSR), National Central University (NCU), in Taiwan.

Declarations

Conflict of interest The author declare that no conflicts of interest.
References

Akhoondzadeh M (2019) Kalman filter and Neural Network methods for detecting irregular variations of TEC around the time of powerful Mexico (Mw=8.2) earthquake of September 08, 2017. J Earth Space Phys 44(4):87–97. https://doi.org/10.22059/jesphys.2018.258251.1007007

Baquirin RBM, Fernandez PL (2018) Artificial neural network (ANN) in a small dataset to determine neutrality in the pronunciation of English as a foreign language in filipino call center agents. Int Artif 21(62):134–144. https://doi.org/10.4114/intartif.vol21iss62pp134-144

Belkin M, Hsu D, Ma S, Mandal S (2019) Reconciling modern machine-learning practice and the classical bias–variance trade-off. Proc Natl Acad Sci USA 116(32):15849–15854. https://doi.org/10.1073/pnas.1903070116

Bhadra A, Bandyopadhyay A, Clakraborthy S, Kumar T (2017) Development and testing of an ANN model for estimation of runoff from a snow covered catchment. J Inst Eng India 98:29–39. https://doi.org/10.1007/s40030-017-0203-5

Bo L, Wang L, Jiao L (2006) Feature scaling for Kernel Fisher discriminant analysis using leave-one-out cross validation. Natural Hazards 44(4):961–978. https://doi.org/10.22059/jesphys.2018.258251.1007007

Belo D, Rodrigues J, Vaz JR, Pezarar-Correia P, Gamboa H (2017) Biosignals learning and synthesis using deep neural networks. BioMed Eng OnLine 16:115. https://doi.org/10.1186/s12938-017-0405-0

Chou YM, Aubourg C, Yeh EC, Song SR, Lin YK, Humbert F, Jiang X, Lee TQ (2020) The magnetic fabric of gouge mimics the coseismic focal mechanism of the Chi-Chi earthquake (1999, Mw 76). Geophys Res Lett 47(22):e2020GL090111. https://doi.org/10.1029/2020GL090111

Cooper C, Mitchell CN, Wright CJ, Jackson DR, Witvliet BA (2019) Measurement of ionospheric total electron content using single-frequency geostationary satellite observations. Radio Sci 54(1):10–19. https://doi.org/10.1029/2018RS006575

Doroudi S (2020) The bias-variance tradeoff: how data science can inform educational debates. AERA Open 6(4):1–18. https://doi.org/10.1177/2332858420977208

Ede JM, Beanland R (2020) Adaptive learning rate clipping stabilizes learning. Mach Learn 1(1):015011. https://doi.org/10.1088/2632-2153/ab81e2

Freund FT (2003) Rocks that crackle and sparkle and glow strange pre-earthquake phenomena. J Sci Explor 17(1):37–71

Fu CC, Walia V, Yang TF, Lee LC, Liu TK, Chen CH, Kumar A, Lin SJ, Lai TH, Wen KL (2017) Pre-seismic anomalies in soil-gas radon associated with 2016 M66 Meinong earthquake, southern Taiwan. Terrestr Atmos Oceanic Sci 28(5):787–798. https://doi.org/10.3319/TAO.2017.03.22.01

Gautam PK, Chauhan V, Sathyaseelan R, Kumar N, Pappachen JP (2018) Co-seismic ionospheric GPS-TEC disturbances from different source characteristic earthquakes in the Himalaya and the adjoining regions. NRIAG J Astron Geophys 7(2):237–246. https://doi.org/10.1016/j.nrjag.2018.05.009

Hemmerich J, Asilar E, Ecker GF (2020) COVER: conformational oversampling as data augmentation for molecules. J Cheminform 12:18. https://doi.org/10.1186/s13321-020-00420-z

Hawkins DM (2004) The problem of overfitting. J Chem Inf Comput Sci 44(1):1–12. https://doi.org/10.1021/ci0402472

Homaj MJ (2016) Prediction of total electron content of the ionosphere using neural network. Jurnal Teknologi 78:53–57. https://doi.org/10.11131/jt.v78.8750

Hu D, Simões PJA, Fletcher L, Hayes LA, Hannah IG (2021) Hot X-ray onsets of solar flares. Mon Not Astron Soc 501(1):1273–1281. https://doi.org/10.1093/mnras/staa3664

Jiang T, Hu XJ, Yao XH, Tu LP, Huang JB, Ma XX, Cui J, Wu QF, Xu JT (2021) Tongue image quality assessment based on a deep convolutional neural network. BMC Med Inform Decis Making 21:147. https://doi.org/10.1186/s12911-021-01508-8

Kalita S, Devi M, Barbara AK, Talukdar PH (2012) Soft computing technique for recognition of earthquake precursor from low latitude total electron content (TEC) profiles. Int J Comput Appl 44(17):11–14. https://doi.org/10.5120/6354-8775

Kamogawa M, Kakinami Y (2013) Is an ionospheric electron enhancement preceding the 2011 Tohoku-Oki earthquake a precursor? J Geophys Res Space Phys 118(4):1751–1754. https://doi.org/10.1002/jgra.50118

Khoshgoftaar TM, Hulse JV, Napolitano A (2010) Supervised neural network modeling: an empirical investigation into learning from imbalanced data with labeling errors. IEEE Trans Neural Netw 21(5):813–830. https://doi.org/10.1109/TNN.2010.2042730

Kumar S, Patel K, Singh AK (2016) TEC variation over an equatorial and anomaly crest region in India during 2012 and 2013. GPS Soluti 20(4):617–626. https://doi.org/10.1007/s10291-015-0470-4
Lin FF, Wang CY, Su CL, Shiokawa K, Saito S, Chu YH (2016) Coordinated observations of F region m field-aligned plasma irregularities associated with medium-scale traveling ionospheric disturbances. J Geophys Res Space Phys 121(4):3750–3766. https://doi.org/10.1002/2016JA022511

Lin JW (2010) Ionospheric total electron content (TEC) anomalies associated with earthquakes through Karhunen-Loève Trans form (KLT). Terrestr Atmos Oceanic Sci J 21(2):253–265. https://doi.org/10.3319/TAO.2009.06.11.01(T)

Lin JW, Chao CT, Chiu JS (2018) Backpropagation neural network as earthquake early warning tool using a new modified elementary Levenberg–Marquardt Algorithm to minimise backpropagation errors. Geosci Instrum Methods Data Syst 7(3):235–243. https://doi.org/10.5194/gi-7-235-2018

Liu JY, Chen YI, Chuo YJ, Tsai HF (2001) Variations of ionospheric foF2 and TEC anomalous days associated with M ≥ 5.0 earthquakes in Taiwan during 1997–1999. Terrestr Atmos Oceanic Sci J 15(3):371–383. https://doi.org/10.3319/TAO.2004.15.3.371(EP)

Liu, J. Y., Chen, C. H., Chen, Y. I., Yen, H. Y., Hattori, K., & Yamoto, K., 2006. Seismo-geophysical anomalies and M≥5.0 earthquakes observed in Taiwan during 1988–2001, Physics and Chemistry of the Earth Parts A/B/C, 31(4–9), 215–222. doi:https://doi.org/10.1016/j.pece.2006.02.009

Mannucci AJ, Wilson BD, Yuan DN, Ho CH, Lindqwister UJ, Runge TF (1998) A global mapping technique for GPS-derived ionospheric total electron content measurements. Radio Sci 33(3):565–582. https://doi.org/10.1029/97RS02707

Marchetti D, Santis AD, Shen X, Campuzano SA, Perrone L, Piscini A, Giovambattista RD, Jin S, Ippolito A, Cianchini G, Cesaroni C, Sabbagh D, Spogli L, Zhima Z, Huang J (2020) Possible lithosphere-atmosphere-ionosphere coupling effects prior to the 2018 Mw=7.5 Indonesia earthquake from seismic, atmospheric and ionospheric data. J Asian Earth Sci 188:104097. https://doi.org/10.1016/j.jseaes.2019.104097

Mocanu DC, Mocanu E, Stone P, Nguyen PH, Giesbusc M, Liotta A (2018) Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science. Nat Commun 9:2383. https://doi.org/10.1038/s41467-018-04316-3

Moses M, Dodo JD, Ojigi LM, Lawal K (2020) Regional TEC modelling over Africa using deep structured supervised neural network. Geodesy Geodynami 11(5):367–375. https://doi.org/10.1016/j.geog.2020.05.004

Nadia JT, Diedan R, Suryana Y (2017) Plate recognition using backpropagation neural network and genetic algorithm. Procedia Comput Sci 116:365–372. https://doi.org/10.1016/j.procs.2017.10.068

Okoh D, Owolabi O, Ekekchukwu C, Folarin O, Arhiwo G, Agbo J, Bolaji S, Rabiu B (2016) A regional GNSS-VTEC model over Nigeria using neural networks: a novel approach. Geodesy and Geodynamics 7(1):19–31. https://doi.org/10.1016/j.geog.2016.03.003

Ogwala A, Somoye EO, Ognumodimu O, Adeniji-Adele RA, Onori EO, Oyedokum O (2019) Diurnal, seasonal and solar cycle variation in total electron content and comparison with IRI-2016 model at Birnin Kebbi. Ann Geophys 37(5):775–789. https://doi.org/10.5194/angeo-37-775-2019

Oyama KL, Devi M, Ryu K, Chen CH, Liu JY, Liu H, Bankov L, Kodama T (2016) Modifications of the ionosphere prior to large earthquakes: report from the Ionosphere Precursor Study Group. Geosci Lett. https://doi.org/10.1186/s41062-016-0038-3

Pasini A (2015) Artificial neural networks for small dataset analysis. J Thorac Dis 7(5):953–960. https://doi.org/10.3978/j.issn.2072-1439.2015.04.61

Pulinets SA (2004) Ionospheric precursors of earthquakes: recent advances in theory and practical applications. Terrestr Atmos Oceanic Sci J 15(3):413–435. https://doi.org/10.3319/TAO.2004.15.3.413(EP)

Pulinets S, Ouzounov D (2011) Lithosphere–atmosphere–ionosphere coupling (LAIC) model—an unified concept for earthquake precursors validation. J Asian Earth Sci 41(4–5):371–382. https://doi.org/10.1016/j.jseaes.2010.03.005

Pulinets S, Ouzounov D, Davydenko D, Petrukhin A (2016) Multiparameter monitoring of short-term earthquake precursors and its physical basis. Implementation in the Kamchatka region, E3S Web Conf, VII International Conference “Solar-Terrestrial Relations and Physics of Earthquakes Precursors”, vol 11, Article Number: 00019, https://doi.org/10.1051/e3sconf/20161100019.

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323:533–536. https://doi.org/10.1038/323533a0

Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Seddiki K, Sautemont P, Precioso F, Ogircnic N, Salzet M, Fournier I, Droi A (2020) Cumulative learning enables convolutional neural network representations for small mass spectrometry data classification. Nat Commun 11:5595. https://doi.org/10.1038/s41467-020-19354-z
Tariq MA, Shah M, Ulukavak M, Iqbal T (2019) Comparison of TEC from GPS and IRI-2016 model over different regions of Pakistan during 2015–2017. Adv Space Res 64(3):707–718. https://doi.org/10.1016/j.asr.2019.05.019

Thomas A, Walters S, Ghaytassi MM, Morgan R, Petridis M (2016) On the optimal node ratio between hidden layers: a probabilistic study. Int J Mach Learn Comput 6(5):241–247. https://doi.org/10.18178/ijmlc.2016.6.5.605

Tsagouri I, Koutroumbas K, Elias P (2018) A new short-term forecasting model for the total electron content storm time disturbances. J Space Weather Space Climat 8:A33. https://doi.org/10.1051/swsc/2018019

Zakeri Z, Mansfield N, Sunderland C, Omurtag A (2020) Cross-validating models of continuous data from simulation and experiment by using linear regression and artificial neural networks. Inform Med Unlocked 21:100457. https://doi.org/10.1016/j.imu.2020.100457

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.