Using geoelectric field skewness and kurtosis to forecast the 2016/2/6, M_L 6.6 Meinong, Taiwan Earthquake

Hong-Jia Chen[1,3], Chien-Chih Chen[1,*], Guy Ouillon[2], and Didier Sornette[3]

[1] Department of Earth Science, National Central University, Taoyuan City, Taiwan
[2] Lithophyse, Nice, France
[3] Department of Management, Technology and Economics, ETH Zürich, Zürich, Switzerland

ABSTRACT

The earthquake-alarm model developed by Chen and Chen (Nat. Hazards 2016) is investigated to validate its forecasting performance for the 2016/2/6, M_L 6.6 Meinong, Taiwan earthquake. This alarm model is based on geoelectric field skewness and kurtosis anomalies. The model parameters, such as the detection range and predicted time window, allow us to estimate the empirical relationships between geoelectric anomalies and large earthquakes. As a result, the skewness and kurtosis anomalies are shown to appear before the Meinong earthquake on the four neighboring stations (LIOQ, WANL, KAOH, and CHCH). According to the model analysis a time lag exists between anomaly clusters and earthquakes, depending on local geological features, as well as the durations over which anomalies are continuously observed, which might also display time dependence. In conclusion, this alarm model is able to correlate earthquakes and geoelectrical anomalies, with promising usefulness in forecasting large earthquakes.

1. INTRODUCTION

This work investigates an earthquake-forecasting algorithm based on geoelectric field skewness and kurtosis (Chen and Chen 2016). The M_L 6.6 Meinong earthquake hit the southern part of Taiwan (120.54°E, 22.92°N) at 03:57 AM, 2016/2/6 (UTC+8) at a depth of 14.64 km, causing the death of over one hundred people, the collapse of tens of buildings, and other significant infrastructure damage (cf. the listed Website links: GEER Association 2016; NCDR 2016; Wikipedia 2016). Earthquake forecasting, therefore, is a critical task. The effective earthquake forecasting should provide information that includes the time, location, magnitude and probability for an impending earthquake. Effective forecasting should also help scientists simulate shaking hazard maps, as well as warn the public in areas that are likely to be seriously damaged if a large earthquake strikes.

The physical mechanisms of earthquake rupture precursors are still debated, but possible models have been proposed. Those models include positive hole charge carriers (Freund 2007; Freund et al. 2009; Freund and Pilorz 2012), the piezoelectric effect of quartz (Nitsan 1977), the electrokinetic effect due to water filtration (Ishido and Mizutani 1981), pressure stimulated currents (Varotsos and Alexopoulos 1984b) arising from the (re)orientation of electric dipoles due to lattice defects (Londos et al. 1996) formed mainly in the ionic constituents of rocks, and so forth. Of the many kinds of earthquake precursors, earthquake forecasting using geoelectromagnetic fields is claimed to have some promising potential (Varotsos and Alexopoulos 1984a, b; Kawase et al. 1993; Varotsos et al. 1993, 2006a, b, 2008, 2009, 2011a, 2013; Uyeda et al. 2000, 2002; Eftaxias et al. 2004, 2007, 2013; Telesca et al. 2004, 2005, 2008, 2009; Orihara et al. 2009, 2012; Eftaxias 2010; Ramirez-Rojas et al. 2013; Chen and Chen 2016). Chen and Chen (2016) built their forecasting algorithm on the previous analogy between large earthquakes and critical points, suggesting that both mechanical and electromagnetic signals should feature generic symptoms, including increased autocorrelation,

* Corresponding author
  E-mail: chencc@earth.ncu.edu.tw

increased variance, varied skewness, increased kurtosis, etc. (Sornette and Sornette 1990; Sornette and Sammis 1995; Bowman et al. 1998; Sammis and Sornette 2002; Varotsos et al. 2003; Scheffer 2009; Scheffer et al. 2009; Dakos et al. 2012). Varotsos et al. (2006a, b, 2008) developed the natural time analysis (NTA), applied it to their seismic electric signals (SESs), and suggested that long-range temporal correlations increase prior to large earthquakes. Eftaxias and his team analyzed the power law distributions in power spectral densities (Eftaxias et al. 2004, 2007) and used non-extensive statistical physics (Eftaxias 2010; Eftaxias et al. 2013) to investigate seismogenic electromagnetic emissions. They concluded that those emissions are associated with the earthquake preparation process. Telesca and his team applied multifractal detrended fluctuation analysis (MFDFA) (Telesca et al. 2004, 2005) and the Fisher Information Measure (FIM) (Telesca et al. 2009) to geoelectric datasets in different seismic areas. They found clear correlations with the largest earthquakes that occurred within the monitored area during the observation period. Before the 2011/3/11 M_w 9.0 Tohoku, Japan earthquake, geoelectromagnetic anomalies and many other precursors had been reported (Ouzounov et al. 2011; Varotsos et al. 2011b; Sarlis et al. 2013; Xu et al. 2013; Skordas and Sarlis 2014; Han et al. 2015; Kamiyama et al. 2016).

Evidence of the precursory nature of these indexes proposed in these previous works was provided via dependence measures, but without tests of their forecasting skills or within systematic analyses on full time series. To solve the ambiguity, Chen and Chen (2016) developed a model to examine the empirical relationship between earthquakes and anomalous geoelectric field statistics. They proposed the "Geoelectric Monitoring System's Time of Increased Probability" (GEMSTIP) model and applied this model to estimate the occurrence time of the anomalies, the predicted time window for the earthquakes and other parameters associated with the anomalies. The GEMSTIP model, as a ruler measuring the distance between two points, helps us understanding the empirical connection between earthquakes and geoelectric fields. This paper provides a validation step for this model using the available datasets from 2013/1/16 to 2015/12/31 as the training set, and the datasets from 2016/1/1 to 2016/3/31 as the forecasting set. We expect that the best parameters obtained from the training period allow us to forecast the M_L 6.6 Meinong earthquake.

Chen and Chen (2016) found time lags between two earthquakes with $M_L \geq 6$ and their preceding clustered anomalies. A time lag is the elapsed time between the end time of the clustered precursory anomalies and an earthquake occurrence. We modify their model in this paper by taking the time-lag effect into account. Other questions addressed in this paper are as follows: (1) What is the best model, the original one or the modified one? (2) Why does a time-lag effect exist? (3) What should be the optimal training window size for

better forecasting performances? Answering those questions may help to obtain a deeper understanding of the potential links between earthquakes and the geoelectric field.

2. DATA

The earthquake catalog is routinely processed and released by the Central Weather Bureau (CWB) of Taiwan. The geoelectric field is registered by the Geoelectric Monitoring System (GEMS), which is maintained by Prof. Chien-Chih Chen and his team. We used all $M_L \geq 5$ earthquakes and calculated the skewness and kurtosis of the daily geoelectric field distribution, using a sampling rate of 1 Hz (cf. Chen and Chen 2016). Figure 1 shows the geoelectric station spatial distribution and the earthquakes with $M_L \geq 5$. The solid magenta star is the 2016/2/6 M_L 6.6 Meinong earthquake, and LIOQ is the nearest station to this event. We used both the earthquake catalog and geoelectric data from 2013/1/16 to 2015/12/31 as the training set, and data from 2016/1/1 to 2016/3/31 as the forecasting set. We also selected different training windows in the training period (2013/1/16 to 2015/12/31) to discuss the effect of the training window size on the forecasting performances. The training windows thus span from 180 days to 1080 days, with a step of 90 days, simply denoted as (180:90:1080) (days). The first training set spans from 2015/7/5 to 2015/12/31 with a training window of 180 days. The second training set spans from 2015/4/6 to 2015/12/31 with a training window of 270 days. The third training set spans from 2015/1/6 to 2015/12/31 with a training window of 360 days, and so forth. The end time for all training sets is 2015/12/31. In the forecasting period spans from 2016/1/1 to 2016/3/31. Seven target earthquakes are denoted as colorful stars in Fig. 1. Four of them are inland and around LIOQ, the others are located on the eastern coast.

3. METHODS

This section is composed of (i) the "Geoelectric Monitoring System's Time of Increased Probability" (GEMSTIP) analysis algorithm workflow, (ii) the original GEMSTIP model proposed in Chen and Chen (2016), (iii) the modified GEMSTIP model considering the time-lag effect, (iv) the priority-search area concept (PSA), and (iv) model ranking used to select the best model in similar performance cases.

3.1 GEMSTIP Analysis Algorithm Workflow

Chen and Chen (2016) developed the "Geoelectric Monitoring System's Time of Increased Probability" (GEMSTIP) model and tested the relationship between earthquakes with $M_L \geq 5$ and the anomalous skewness and kurtosis of geoelectric fields. Figure 2 shows the methodology workflow proposed in Chen and Chen (2016), which we
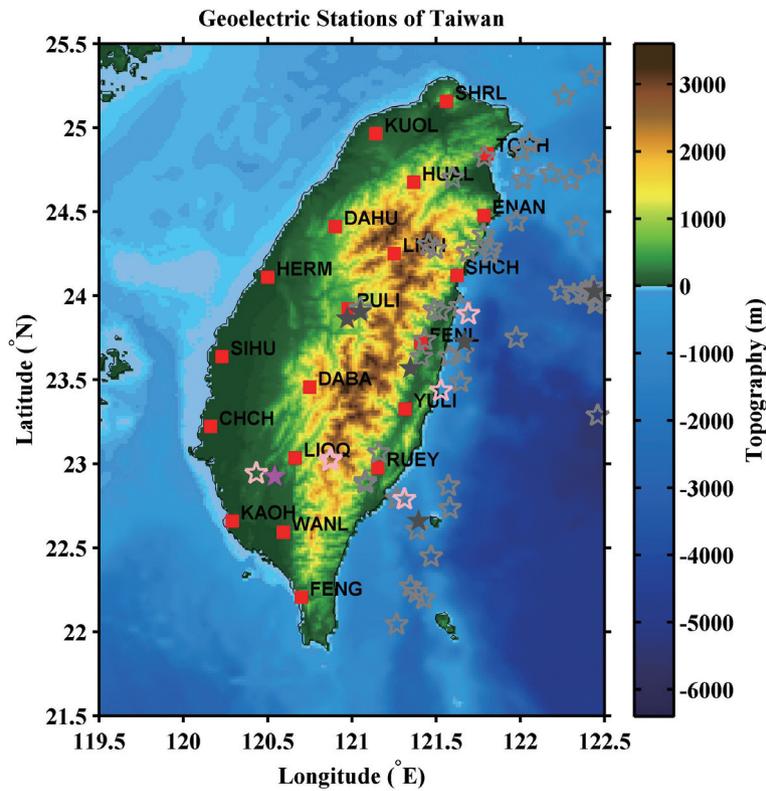
Fig. 1. Spatial distribution of the geoelectric stations and the earthquakes with $M_L \geq 5$ in Taiwan. Open stars are the earthquakes with $M_L \in [5, 6)$, and solid stars are the earthquakes with $M_L \geq 6$. Gray and dark stars are the earthquakes before 2015/12/31 within the training set, while pink and magenta stars are the earthquakes from 2016/1/1 to 2016/3/31 within the forecasting set. (Color online only)
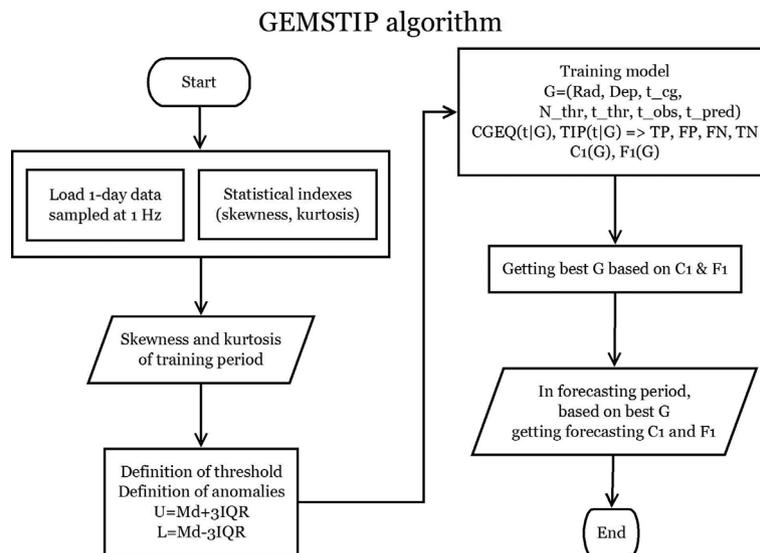
Fig. 2. Flowchart of the GEMSTIP analysis algorithm.

shall also follow in this paper. The skewness and kurtosis of the daily geoelectric field distributions for North-South (NS) and East-West (EW) components are first calculated. There are four indexes per day. The start and end time of a training period is selected next. Chen and Chen (2016) used the training period from 2012/1/1 to 2014/12/31, whereas we use the training period from 2013/1/16 to 2015/12/31 and also discuss the effect of different training windows between the start and end times. The upper and lower thresholds used to detect anomalies are defined for each index as the median ±3 times the interquartile range in the selected time window (shown in the upper panel of Fig. 3). An anomalous index is detected if the observed index value is beyond one of the thresholds. The anomalous index number (AIN) is then counted and summed according to the anomalous indexes of the skewness and kurtosis per component (shown in the lower panel of Fig. 3). Note that AIN is a discrete value. The GEMSTIP model is finally applied to the AIN series and the target earthquakes. The best model parameters are then determined for the considered training window.

3.2 Original GEMSTIP Model

We consider the Chen and Chen model (2016) as the original GEMSTIP model. In the original model the parameter vector G^o is: $G^o = (Rad, Dep, t_{cg}, N_{thr}, t_{thr}, t_{obs}, t_{pred})$, where the superscript O means the original GEMSTIP model. The meaning of each parameter is shown in Fig. 3. The parameter Rad is the radius for selecting earthquakes within the region centered at a given station. The parameter Dep is the depth above which we select earthquakes. The parameter t_{cg}

is the coarse-grained time of an earthquake. The parameter N_{thr} is a threshold number to label a day as anomalous if $AIN \geq N_{thr}$ for that day. The parameters t_{thr}, t_{obs}, and t_{pred} are time windows, which mean that if the number of anomalous days within the observation time window t_{obs} is greater than or equal to the threshold duration t_{thr}, then the immediate future time window t_{pred} is alarmed as a time of increased probability (TIP) (cf. Chen and Chen 2016). Note that the parameter t_{cg} is equal to t_{pred} hereafter.

We then define the coarse-grained earthquake (CGEQ) time function as follows:

$$CGEQ(t \mid Rad, Dep, t_{cg}) =$$
$$\begin{cases} 1, & t_{EQ}(Mag, Rad, Dep) - \dfrac{t_{cg}}{2} \leq t \leq t_{EQ}(Mag, Rad, Dep) + \dfrac{t_{cg}}{2} \\ 0, & others \end{cases} \quad (1)$$

where Mag stands for the magnitude of the earthquake. We herein use $Mag \geq 5$. We also define the prediction index time function, i.e., time of increased probability (TIP) as follows:

$$TIP(t \mid N, t_{thr}, t_{obs}, t_{pred}) =$$
$$\begin{cases} 1, & \begin{aligned} & t_e < t \leq t_e + t_{pred} \, conditioned \, to \, N(AIN \geq N_{thr}) \geq t_{thr}, \\ & t_e - t_{obs} \leq t \leq t_e \end{aligned} \\ 0, & others \end{cases} \quad (2)$$

where $N(.)$ is a number-counting operator and t_e means the end time point of the observation time window t_{obs}. Equation (2) means that when the number of $AIN \geq N_{thr}$ is
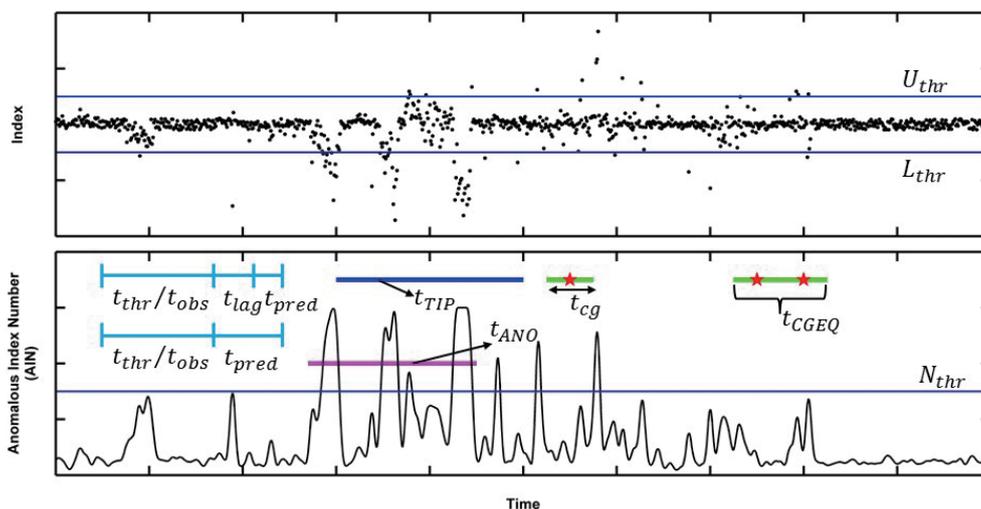
Fig. 3. Schematic diagram of the GEMSTIP model. The upper threshold U_{thr} and the lower threshold L_{thr} are the median ±3 times of interquartile range of this index time series. An anomalous index number (AIN) is counted if the index is out of U_{thr} and L_{thr}. Note that AIN is integer in practice and the AIN series here is an exaggeration for easy explanation. An anomalous day is defined if AIN $\geq N_{thr}$. If anomalous days in t_{obs} are greater than or equal to t_{thr}, then the time window t_{pred} will be issued as TIP (time of increased probability). Note that it is not necessary that the anomalous days be consecutive. A large earthquake occurrence (red star) is coarse-grained by $\pm t_{cg}/2$. A continuous segment of TIP is denoted as t_{TIP}, while a continuous segment of coarse-grained earthquake (CGEQ) is denoted as t_{CGEQ}. A continuous segment of anomalies is denoted as t_{ANO}. (Color online only)

greater than or equal to the value t_{thr} in the period $t_e - t_{obs} \leq t \leq t_e$, the TIP is declared to be 1 in the period $t_e \leq t \leq t_e + t_{pred}$. Note that it is not necessary that the anomalous days are consecutive.

In Fig. 3, we see a continuous TIP segment and a continuous CGEQ segment. We denote them as t_{TIP} and t_{CGEQ}, respectively. We then define a continuously anomalous time t_{ANO} as follows:

$$t_{ANO} = t_{TIP} - t_{pred} + t_{obs} \qquad (3)$$

Imagine that the window set of t_{obs} and t_{pred} find anomalies in the t_{obs} only one time, so that $t_{TIP} = t_{pred}$ and $t_{ANO} = t_{obs}$. Hence, the start time point of t_{ANO} is that of t_{TIP} minus t_{obs}, and the end time point of t_{ANO} is that of t_{TIP} minus t_{pred}.

Performing a binary classification, we compare CGEQ to TIP time series, and obtain:

$$
\begin{aligned}
TP &= N(TIP = 1 \cap CGEQ = 1) \\
FP &= N(TIP = 1 \cap CGEQ = 0) \\
FN &= N(TIP = 0 \cap CGEQ = 1) \\
TN &= N(TIP = 0 \cap CGEQ = 0)
\end{aligned}
\qquad (4)
$$

In Eq. (4), TP means the number of true positives, FP the number of false positives, FN the number of false negatives, and TN the number of true negatives. For example, CGEQ $= (1, 0, 0, 0, 1)$ and TIP $= (1, 1, 0, 0, 0)$. In this case, TP $= 1$ based on the first elements of CGEQ and TIP, FP $= 1$ based on the second, FN $= 1$ based on the fifth, and TN $= 2$ based on the third and the fourth. We then define the model performances as follows:

$$
\begin{aligned}
C1(G) &= \frac{2TN}{2TN + FP + FN'} \\
F1(G) &= \frac{2TP}{2TP + FP + FN'} \\
R(G) &= \sqrt{C1^2 + F1^2}
\end{aligned}
\qquad (5)
$$

In Eq. (5), C1 is the TN performance, F1 is the TP performance, and R is the ensemble performance.

We set the parameters $Rad = (30:5:60, 70:10:100)$ (km), $Dep = (10:10:50, 100:50:300)$ (km), $N_{thr} = (1:1:4)$, $t_{thr} = (1:1:\frac{2}{3}t_{obs})$ (day), $t_{obs} = (5:5:30, 40:10:100)$ (day), $t_{pred} = (5:5:30, 40:10:100)$ (day). Note that the data format is (initial value : increment : end value) which means that the values of one parameter are from an initial value to an end value with an incremental step. The number of original tested models thus amounts to 1361360.

3.3 Modified GEMSTIP Model

Chen and Chen (2016) found that there is a time lag between a large earthquake and the clusters of geoelectric anomalies. We take the time-lag effect into consideration in this paper. We, hereafter, called the GEMSTIP model considering the time-lag effect as the modified GEMSTIP model.

The parameter vector G^M is: $G^M = (Rad, Dep, t_{cg}, N_{thr}, t_{thr}, t_{obs}, t_{pred}, t_{lag})$, where the superscript M means the modified GEMSTIP model. The time lag parameter t_{lag} is the elapsed time between the end of the observation time window t_{obs} and the beginning of the predicted time window t_{pred} (shown in the lower panel of Fig. 3), whereas the other parameters have the same meanings as in the original model. The parameter t_{cg} is also equal to t_{pred} hereafter.

The only difference between the modified model and the original one is:

$$
TIP(t \mid N, t_{thr}, t_{obs}, t_{pred}, t_{lag}) = \\
\begin{cases}
1, & \begin{array}{l} t_e + t_{lag} < t \leq t_e + t_{lag} + t_{pred} \, conditioned \, to \, N(AIN \geq N_{thr}) \geq t_{thr}, \\ t_e - t_{obs} \leq t \leq t_e \end{array} \\
0, & \qquad\qquad others
\end{cases}
\qquad (6)
$$

Comparing Eqs. (2) and (6), we see that the original GEMSTIP model is a specific case of the modified GEMSTIP model when $t_{lag} = 0$. The relationship between t_{ANO} and t_{TIP} in the modified model is the same as in Eq. (3). However, imagine that the window set of t_{obs}, t_{pred}, and t_{lag} find anomalies in the t_{obs} only one time, so that $t_{TIP} = t_{pred}$ and $t_{ANO} = t_{obs}$, but there is a time gap t_{lag} between t_{ANO} and t_{TIP}. Hence, the start time point of t_{ANO} is that of t_{TIP} minus $(t_{obs} + t_{lag})$, and the end time point of t_{ANO} is that of t_{TIP} minus $(t_{pred} + t_{lag})$.

We set the parameters $Rad = (30:5:60, 70:10:100)$ (km), $Dep = (10:10:50, 100:100:300)$ (km), $N_{thr} = (1:1:4)$, $t_{thr} = (1:1:\frac{2}{3}t_{obs})$ (days), $t_{obs} = (5:5:20, 30:10:100)$ (days), $t_{pred} = (5:5:20, 30:10:60)$ (days), $t_{lag} = (0:10:90)$ (days). The total number of tested modified models is 5603840.

3.4 Priority-Searching Area

The priority-search area concept (PSA) is now introduced. Chen and Chen (2016) plotted the C1-F1 diagrams to observe the model performance patterns and the location of the best model. However, it is not convenient to compare the C1-F1 patterns when the number of the models with different tested conditions is strongly increased. For the sake of simplifying the C1-F1 diagram comparisons, we divided the C1-F1 area into 7 PSAs, as shown in Fig. 4. PSA 1 corresponds to R > 1, C1 > 0.5, and F1 > 0.5; PSA 2 to R ≤ 1, C1 > 0.5, and F1 > 0.5; PSA 3 to R > 1 and F1 ≤ 0.5; PSA 4 to R > 1 and C1 ≤ 0.5; PSA 5 to R ≤ 1, C1 > 0.5, and F1 ≤ 0.5; PSA 6 to R ≤ 1, C1 ≤ 0.5, and F1 > 0.5; and finally PSA 7 to C1 ≤ 0.5 and F1 ≤ 0.5. Intuitively, PSA 1 is the best region for the model performance, while PSA 7 is the worst region. Hence, PSAs 1 to 7 correspond to the model

performances in descending order. We looked for the corresponding PSA of each model and eventually determined the best model parameters for each PSA. Based on the PSA number, we are able to image the C1-F1 patterns without plotting the C1-F1 diagrams (cf. Figs. 4 and S3 of Chen and Chen 2016), which is rather helpful when comparing the enormous tested model sets. Moreover, when the best model occurred in PSA 5, 6, or 7, we did not use this as the best model.

3.5 Model Ranking

Consider the original GEMSTIP model as an example and assume a M_L 6 earthquake occurred 40 km from a given station, with a depth of 39 km. If the models $G_1^O = (50, 40, 2, 4, 10, 15)$ and $G_2^O = (60, 40, 2, 4, 10, 15)$ have the same ensemble performance R, we prefer model G_1^O because the parameter $Rad = 60$ in G_2^O is overestimated in this case. Or if models $G_3^O = (50, 40, 2, 4, 10, 15)$ and $G_4^O = (50, 50, 2, 4, 10, 15)$ have the same ensemble performance R, we choose model G_3^O because the parameter $Dep = 50$ in the model G_4^O is overestimated. One can imagine similar considerations for the other parameters. This amounts to a kind of LASSO (least absolute shrinkage and selection operator) penalization of the parameters (Tibshirani 1996). We solve this problem using the spatial LASSO and temporal LASSO in order to rank the models. The spatial LASSO is the detected volume:

$$L_S = Dep \cdot Rad^2 \tag{7}$$

The temporal LASSO is the 'volume' of anomalies:

$$L_T = N_{thr} \cdot t_{thr} \tag{8}$$

If there are models with the same best ensemble performance R, we first choose the model with the minimum of the spatial LASSO L_S. If models with the same R value still occur, we then choose the model with the minimum temporal LASSO L_T. In practice, we observed that the spatial LASSO is sufficient to rank models with similar R values most of the time.

4. RESULTS

Figure 5 shows the skewness and kurtosis series at 4 stations, which are nearest to the epicenter of the 2016/2/6 M_L 6.6 Meinong earthquake. LIOQ, the closest station (at 17.30 km), exhibits anomalous skewness and kurtosis clusters within 2015/6/18 to 2015/11/13, corresponding apparently to the Meinong earthquake. There is thus a time lag of approximately 86 days between the end time of the anomalous clusters and the earthquake occurrence. At WANL, the

second nearest station (at 37.17 km), the skewness and kurtosis start to deviate from their thresholds from 2015/7/10. Yet, before, during and after the Meinong earthquake, the anomalies persist. Note that WANL started recording on 2012/2/2, and thus features a gap in the data at the beginning of its time series (Fig. 5b). A similar situation holds at KAOH, the third nearest station (at 39.30 km). KAOH's skewness and kurtosis start to deviate from its thresholds from 2015/5/31, and keeps on up to the end of the dataset. CHCH, the fourth nearest station (at 51.2 km), displays kurtosis deviating from its thresholds from 2015/9/16, but its skewness behaves relatively quiescently. For WANL and KAOH, we also observe different continuous anomaly durations from 2015/10/14 to 2016/3/31 and from 2015/6/1 to 2016/3/31, respectively, before and after the Meinong earthquake occurrence. For other periods, at those 4 stations, there are other anomalies corresponding apparently to other large earthquakes. Because it is hard to determine the relationship using the naked eye, we applied the "Geoelectric Monitoring System's Time of Increased Probability" (GEMSTIP) model to quantitatively measure the relationship between those anomalies and earthquakes.

After the GEMSTIP analysis, Fig. 6 shows the CGEQ-TIP matching diagrams for the best original and modified models. We selected the training datasets from 2015/1/6 to 2015/12/31 with a training window of 360 days to optimize the models. The training window of 360 days is the best option among those different training windows, which is discussed in section 5. The best model parameters are listed
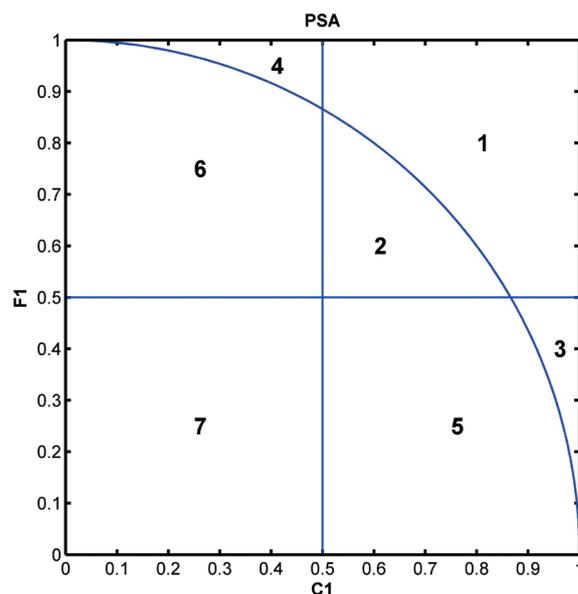
Fig. 4. Schematic diagram of priority-searching areas (PSA) of the C1-F1 diagram. The vertical line is the TN performance C1 = 0.5, the horizontal line is the TP performance F1 = 0.5, and the quarter circle curve is the ensemble performance R = 1. PSAs 1 through 7 are ranked in descending order of performance. If the best model occurs in PSAs 5, 6, or 7, it should not be used for forecasting. (Color online only)
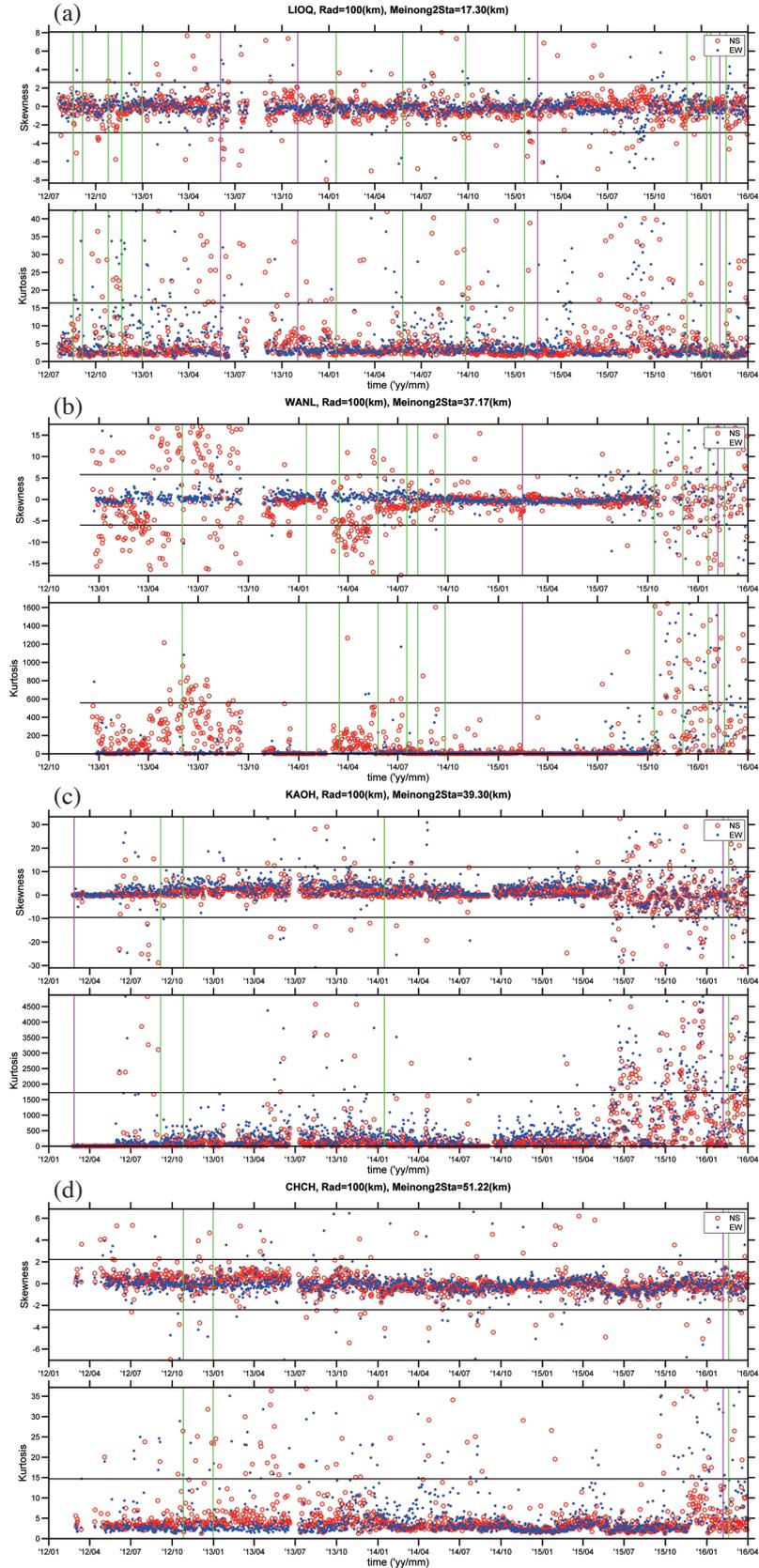
Fig. 5. Time series of skewness and kurtosis in the daily geoelectric field distribution at the 4 stations (a) LIOQ, (b) WANL, (c) KAOH, and (d) CHCH, which are close to the 2016/2/6 M_L 6.6 Meinong earthquake. Red circles are NS component and blue dots are EW component. Green vertical lines are the earthquakes with $M_L \in [5, 6)$ within a radius of 100 km from a given station, and magenta vertical lines are the earthquakes with $M_L \geq 6$. Note that the datasets begin from their own starting times to 2016/3/31. (Color online only)
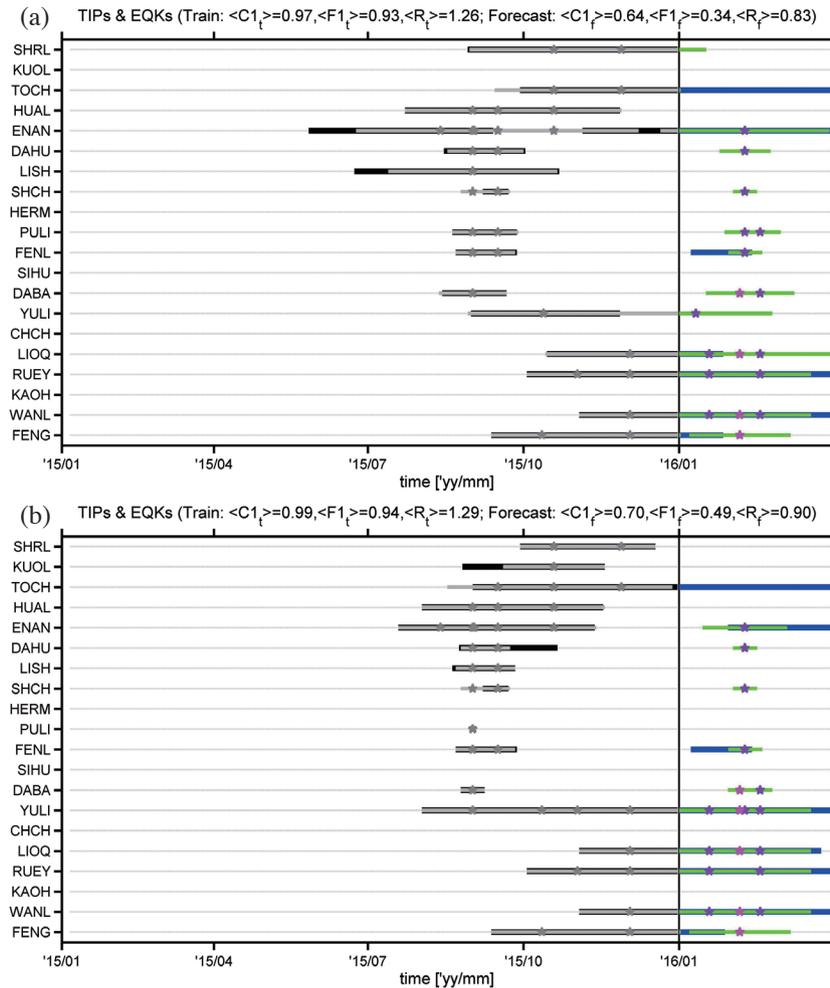
Fig. 6. CGEQ-TIP matching diagrams of (a) the best original model and (b) the best modified model. The training set is from 2015/1/6 to 2015/12/31 with a training window of 360 days, and the forecasting set is from 2016/1/1 to 2016/3/31. Black and blue horizontal lines are TIPs. Gray and green horizontals are CGEQs, the time expansions of earthquake occurrence. Open stars are the earthquakes with $M_L \in [5, 6)$, and solid stars are the earthquakes with $M_L \geq 6$. The gray-level part on the left concerns the training set, while the colorful part on the right deals with the forecasting set. Gray dotted lines are auxiliary guides for the eye. (Color online only)

in Table 1. We then used the determined parameters to forecast the earthquakes with $M_L \geq 5$ occurring from 2016/1/1 to 2016/3/31. We first compared the CGEQ-TIP matching results in the training set. Overall, the original and modified models have a similar pattern for the CGEQ-TIP matchings. However, comparing their performances, the modified model ($\langle C1_t \rangle = 0.99$, $\langle F1_t \rangle = 0.94$, $\langle R_t \rangle = 1.29$) slightly outperforms the original model ($\langle C1_t \rangle = 0.97$, $\langle F1_t \rangle = 0.93$, $\langle R_t \rangle = 1.26$). Note that, hereinafter, the operator $\langle . \rangle$ means the average of a certain parameter over the 20 stations. The subscript t means the training set, while the subscript f means the forecasting set. For the original model, we obtained $\langle Rad \rangle = 64$ (km), $\langle Dep \rangle = 44.5$ (km), $\langle N_{thr} \rangle = 2$, $\langle t_{thr} \rangle = 10.85$ (days), $\langle t_{obs} \rangle = 43.5$ (days), $\langle t_{pred} \rangle = 48.75$ (days) (averaged from Table 1a). For the modified model, we obtained $\langle Rad \rangle = 69$ (km), $\langle Dep \rangle = 49$ (km), $\langle N_{thr} \rangle = 2$, $\langle t_{thr} \rangle = 10.65$ (days), $\langle t_{obs} \rangle = 49.5$ (days), $\langle t_{pred} \rangle = 34$ (days), $\langle t_{lag} \rangle$

= 22.5 (days) (averaged from Table 1b). Those averages are summarized in Table 2. The average predicted time window $\langle t_{pred} \rangle$ in the modified model is smaller than that in the original one, which allows us to point out more precisely when an earthquake will occur. Focusing on the forecasting set, the performances of the original model are $\langle C1_f \rangle = 0.64$, $\langle F1_f \rangle = 0.34$, and $\langle R_f \rangle = 0.83$, while those of the modified model are slightly better with $\langle C1_f \rangle = 0.70$, $\langle F1_f \rangle = 0.49$, and $\langle R_f \rangle = 0.90$. WANL forecasted the M_L 6.6 Meinong earthquake and the other two earthquakes with $M_L \geq 5$ in both the original and modified models. On the other hand, LIOQ and YULI do not forecast the Meinong earthquake in the original model, but forecast the earthquake in the modified one. In the supplementary materials, Figs. S1 and S2 show the CGEQ-TIP matching results of the original and the modified models with different training windows, and Tables S1 and S2 list their best model parameters.

Table 1. Best GEMSTIP parameters of (a) the original model and (b) the modified model. The training set is from 2015/1/6 to 2015/12/31 with a training window of 360 days.

(a)

| StaN	Rad (km)	Dep (km)	N_{thr}	t_{thr} (day)	t_{obs} (day)	t_{pred} (day)	C1	F1	R	PSA
SHRL	60	20	3	13	90	30	0.99	0.98	1.4	1
KUOL	90	20	2	3	5	10	1	1	1.41	1
TOCH	60	250	2	3	5	30	0.9	0.84	1.23	1
HUAL	70	20	2	5	70	40	0.97	0.97	1.38	1
ENAN	100	250	3	2	10	15	0.54	0.55	0.77	2
DAHU	100	20	1	5	25	30	1	0.99	1.4	1
LISH	45	20	1	8	15	30	1	0.98	1.4	1
SHCH	30	20	2	5	30	15	1	0.98	1.4	1
HERM	30	10	2	3	5	5	1	0	1	5
PULI	100	20	2	21	70	5	1	1	1.41	1
FENL	80	20	3	1	20	20	1	0.99	1.4	1
SIHU	30	10	4	4	5	5	1	0	1	5
DABA	100	20	1	4	5	40	1	0.99	1.4	1
YULI	80	50	4	1	5	15	0.92	0.72	1.17	1
CHCH	30	10	3	2	5	5	1	0	1	5
LIOQ	30	10	3	2	5	5	1	0	1	5
RUEY	55	20	1	11	30	60	0.99	0.99	1.4	1
KAOH	30	10	1	3	5	5	1	0	1	5
WANL	90	30	4	8	15	60	1	1	1.41	1
FENG	90	30	1	23	70	5	0.96	0.56	1.11	1

(b)

StaN	Rad (km)	Dep (km)	N_{thr}	t_{thr} (day)	t_{obs} (day)	t_{pred} (day)	t_{lag} (day)	C1	F1	R	PSA
SHRL	80	300	4	5	50	40	40	1	1	1.41	1
KUOL	90	20	3	35	100	60	90	0.94	0.84	1.25	1
TOCH	70	300	1	7	90	60	10	0.99	0.99	1.4	1
HUAL	90	20	1	13	90	60	10	1	1	1.41	1
ENAN	80	30	1	11	50	50	80	1	1	1.41	1
DAHU	100	20	1	7	50	15	10	1	1	1.41	1
LISH	60	20	1	19	50	20	60	1	1	1.41	1
SHCH	30	20	1	5	15	15	0	1	0.98	1.4	1
HERM	30	10	1	4	5	5	0	1	0	1	5
PULI	55	20	2	11	30	5	30	1	1	1.41	1
FENL	80	20	4	1	20	20	0	1	0.99	1.4	1
SIHU	30	10	4	13	20	5	0	1	0	1	5
DABA	100	20	3	7	20	15	10	1	1	1.41	1
YULI	100	30	1	17	100	60	30	1	1	1.41	1
CHCH	30	10	3	2	5	5	0	1	0	1	5
LIOQ	90	30	2	15	70	60	70	1	1	1.41	1
RUEY	55	30	2	15	80	60	0	1	1	1.41	1
KAOH	30	10	1	4	5	5	0	1	0	1	5
WANL	90	30	3	15	70	60	0	1	1	1.41	1
FENG	90	30	2	7	70	60	10	1	1	1.41	1

Table 2. Averages of (a) best parameters and (b) performances of the original and modified models. The training set is from 2015/1/6 to 2015/12/31 with a training window of 360 days.

(a)

Model Type	Rad (km)	Dep (km)	N_{thr}	t_{thr} (day)	t_{obs} (day)	t_{pred} (day)	t_{lag} (day)
Original	64	44.5	2	10.85	43.5	48.75	0
Modified	69	49	2	10.65	49.5	34	22.5

(b)

Model Type	$C1_t$	$F1_t$	R_t	$C1_f$	$F1_f$	R_f	Retrained $C1_f$	Retrained $F1_f$	Retrained R_f
Original	0.97	0.93	1.26	0.64	0.34	0.83	0.56	0.39	0.81
Modified	0.99	0.94	1.29	0.7	0.49	0.9	0.62	0.53	0.88

Note: * The term "retrained" means that the datasets of CHCH and KAOH are reselected to re-optimize the GEMSTIP models. The (re)training set of CHCH is from 2012/4/26 to 2013/4/20 with a training window of 360 days, while the (re)training set of KAOH is from 2013/4/1 to 2014/3/26 with a training window of 360 days.

Figure 7 shows plots of averages of the best parameters versus the size of the training windows. For the parameters $\langle Rad \rangle$, $\langle Dep \rangle$, $\langle N_{thr} \rangle$, $\langle t_{thr} \rangle$, $\langle t_{obs} \rangle$, $\langle t_{pred} \rangle$ in both the original and modified models, and the parameter $\langle t_{lag} \rangle$ in the modified model, we see that changing the training window size changes slightly the average parameters. These results suggest that the inverted models are rather robust. The predicted time windows $\langle t_{pred} \rangle$ in the modified models are smaller than in the original models for all training window sizes, except for a training window of 180 days. The other average parameters have similar values in both the original and modified models. The original model proposed in Chen and Chen (2016) gives a predicted time window for an earthquake that is overestimated because of the time-lag effect between earthquake occurrence and geoelectric field anomalies. We truly improved the original model by considering the time-lag effect. Due to the smaller predicted time windows, we could forecast the earthquake occurrence time more accurately.

We next compared the average performances of the original models to those of the modified models when using different training windows. Figure 8 shows the plots of the average performances ($\langle C1 \rangle$, $\langle F1 \rangle$, $\langle R \rangle$) versus the training windows (180:90:1080) (days) for both the original and modified models. We observe in Fig. 8a that the TN performances $\langle C1_t^o \rangle$ and $\langle C1_t^m \rangle$ in the training sets (subscript t) of the original model (superscript o) and the modified model (superscript m) are mostly above 0.9, and slightly decrease with the increasing training windows size beyond 540 days. The TN performances $\langle C1_f^o \rangle$ and $\langle C1_f^m \rangle$ in the forecasting sets (subscript f) are all decreased by approximately 0.25 when compared to the training TN performances $\langle C1_t^o \rangle$ and $\langle C1_t^m \rangle$. Additionally, the forecasting TN performances $\langle C1_f^m \rangle$ are mostly larger than $\langle C1_f^o \rangle$ except for training windows of 180 and 450 days. Focusing on Fig. 8b, the training TP performances $\langle F1_t^o \rangle$ and $\langle F1_t^m \rangle$ decrease approximately

from 0.9 - 0.5 with increasing training window sizes. The forecasting TP performances $\langle F1_f^o \rangle$ and $\langle F1_f^m \rangle$ are approximately between 0.1 and 0.5. There are maxima of $\langle F1_f^o \rangle$ and $\langle F1_f^m \rangle$ with a training window of 360 days. In Fig. 8c, the training ensemble performances $\langle R_t^o \rangle$ and $\langle R_t^m \rangle$ decrease approximately from 1.30 - 1.05 as the training window size increases. The forecasting ensemble performances $\langle R_f^o \rangle$ and $\langle R_f^m \rangle$ are around 0.8.

5. DISCUSSION

In this work, we found that of the four nearest stations to the M_L 6.6 Meinong earthquake, the clustered anomalous skewness and kurtosis in LIOQ became relatively quiescent prior to the Meinong earthquake. We consider this as the time-lag effect. Chen and Chen (2016) also observed time lags for two earthquakes with $M_L \geq 6$ in 2013 near PULI. Varotsos and his team also found time lags between SESs and earthquakes in Greece (Varotsos et al. 1993, 2009) and Japan (Varotsos et al. 2013). In rock experiments, Triantis et al. (2008) found the pressure stimulated current recordings (PSC) featured a peak before failure. They speculated that the time lags were produced by the formation of the fracture plane. This plane forms at high stresses and the extensive cracking process that is associated with it limits the available conductive path in the sample bulk, inducing a consequent obstacle on the emitted PSC. We applied the modified GEMSTIP model to quantify the time lag between the defined anomalies and earthquakes. However, not all stations show such a time-lag effect. When observing the column corresponding to the time lag t_{lag} in Table 1b, almost all stations in the coastal region display no time-lag, such as SHCH, HERM, FENL, SIHU, CHCH, RUEY, KAOH, and WANL, while stations in the central mountain region possess one. We suspect that the time lag depends on the local geology. The observed low resistivity under the central
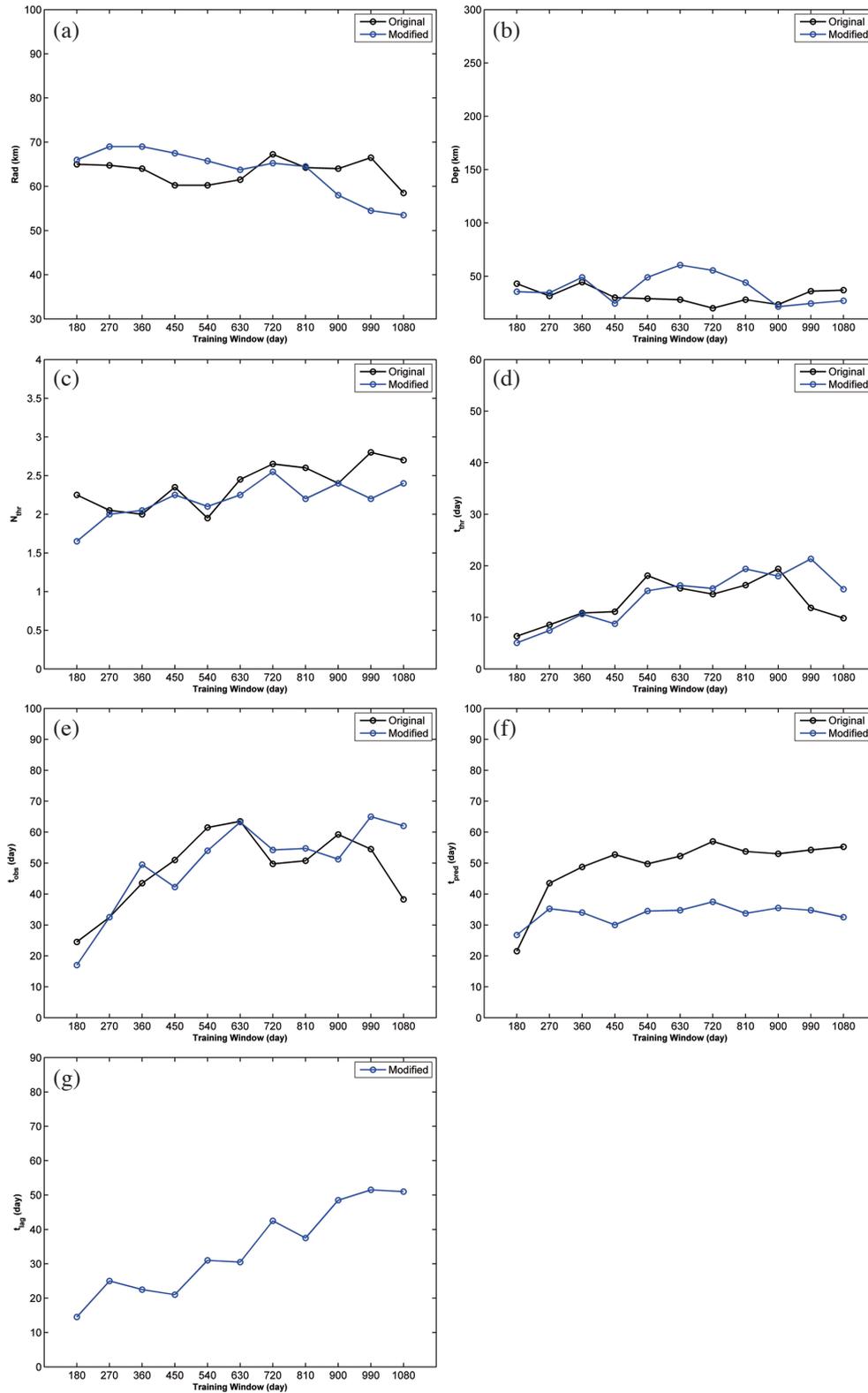
Fig. 7. Line charts of averages of best parameters versus training windows of the original and modified models for (a) the average *Rad*, (b) the average *Dep*, (c) the average N_{thr}, (d) the average t_{thr}, (e) the average t_{obs}, (f) the average t_{pred}, and (g) the average t_{lag}. (Color online only)
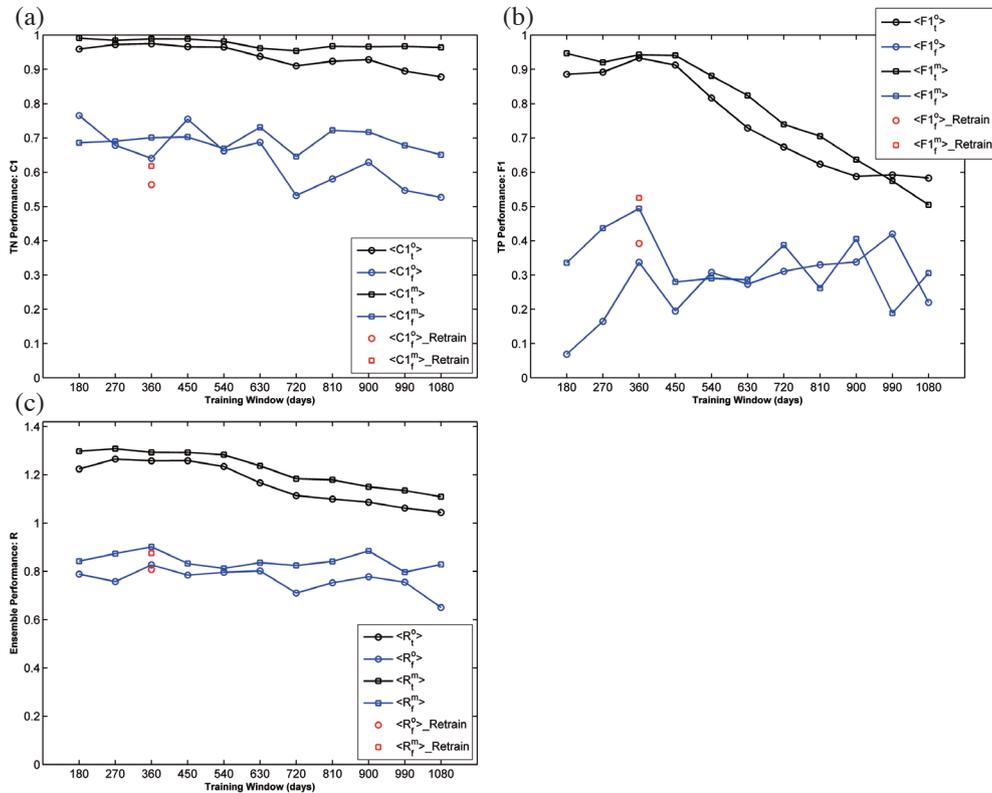
Fig. 8. Line charts of average model performances versus training windows, (a) the TN performance C1, (b) the TP performance F1, and (c) the ensemble performance R. Black lines with circles are the training performances of the original models, and blue lines with circles are for the forecasting performances. Black lines with squares are the training performances of the modified models, and blue lines with squares are for the forecasting performances. The red circle is for the forecasting performances of the retrained original model, and the red square is the forecasting performances of the retrained modified model. The operator $\langle . \rangle$ means the average of parameters for the 20 stations. The superscript o means the original model, the superscript m means the modified model, the subscript t means the training set, and the subscript f means the forecasting set. (Color online only)

mountain region suggests the presence of abundant ground water (Chen and Chen 1998, 2000; Chen et al. 1998; Bertrand et al. 2009, 2012). In the initial stage, small-scale irreversible deformations and fractures occur that generate electromagnetic radiation. Hence, the skewness and kurtosis of the geoelectric field vary largely. However, the abundant ground water starts to saturate those newly formed (or reactivated) cracks. It may also interfere with currents generated by the irreversible deformations. The generated electromagnetic radiation can then decay significantly when passing through the crack-saturated area. This scenario is depicted in Fig. 9. This might explain why there is a relative quiescence in skewness and kurtosis in the geoelectric field just prior to large earthquakes at stations deployed in the saturated area.

Besides the time-lag effect, we found that the duration of anomalous skewness and kurtosis periods vary. Consider SHRL, KUOL, TOCH, and HUAL for instance (in Fig. S2e of the supplementary materials): the continuous TIP times t_{TIP} have different durations. Based on Eq. (3), it is suggested that the durations of lasting anomalies t_{ANO} are also different. The varied lasting anomalous times seem

to rely on the preparation process of different earthquakes. This phenomenon suggests that the earthquake preparation processes are different even within the same region. Therefore, when analyzing geoelectric fields in order to forecast an earthquake, the varied lasting anomalous times largely affect the forecasting performance. The different durations of lasting anomalies not only have a spatial dependence but also a temporal dependence within the same region. This issue might be solved by relating it to the crustal parameters that mainly influence the electrical structure, such as permeability, porosity and so on.

The forecasting TP performance $\langle F1_f^m \rangle = 0.49$ with a training window of 360 days is larger than $\langle F1_f^o \rangle = 0.34$. For training windows larger 360 days, $\langle F1_f^o \rangle$ and $\langle F1_f^m \rangle$ have similar values. The better forecasting performance is produced by smaller predicted time window values t_{pred} in the modified model. YULI and LIOQ in the modified model can identify earthquakes in the forecasting set, whereas they cannot in the original model. On the other hand, the larger t_{pred} of the original model easily leads to longer TIP times t_{TIP}, and to longer coarse-grained earthquake times t_{CGEQ}. This situation leads to unreasonable forecasts. Therefore, it
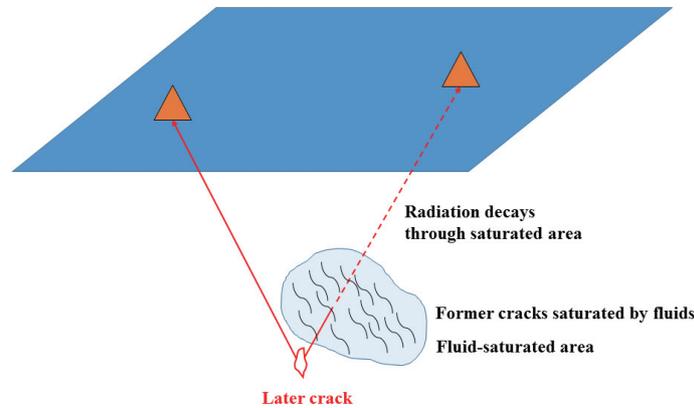
Fig. 9. Model of the time-lag effect between geoelectric signals and earthquakes. The former cracks are saturated with fluids and form a fluid-saturated area. The electromagnetic radiations generated by later cracking decay significantly when passing through the fluid-saturated area. The station above the fluid-saturated area (right side) registers less geoelectric anomalies by later cracking, whereas the other station keeps recording geoelectric anomalies. (Color online only)

is necessary to consider the performances (C1, F1, R) and also the size of the predicted time window t_{pred}. The training ensemble performances $\langle R_t^o \rangle$ and $\langle R_t^m \rangle$ are nearly equal for the models with 180 through 540 days training windows, and start decreasing for the models with training windows of 630 days and above. The forecasting ensemble performances $\langle R_f^o \rangle$ and $\langle R_f^m \rangle$ for the 360 day training window are higher in the 180 through 540 day training windows. Consequently, we conclude that the modified model truly outperforms the original one, and assume the modified model with a training window of 360 days as the best option.

Focusing on the KAOH and CHCH stations in Fig. 6, the CGEQ-TIP matchings are all empty. They also do not give any alarm for the Meinong earthquake. This is because their best models are located within PSA 5, and their performances C1 = 1 and F1 = 0. Hence, their best models have minimum parameters in the given ranges. For models located within PSA 5, 6, and 7, there is less relationship between earthquakes and anomalies skewness and kurtosis in the training periods. When meeting this situation, we moved the training window back to the previous period in which there are large earthquakes and anomalies. We used the reselected training datasets to re-optimize the models, and obtain the best model, called the historic best model. This historic model is substituted for the original best model. Based on Figs. 5c and d, we selected the retraining periods from 2013/4/1 to 2014/3/26 for KAOH and from 2012/4/26 to 2013/4/20 for CHCH. The historic best parameters for KAOH and CHCH are listed in Table 3. Figure 10 shows the new GEMSTIP matching diagrams for the original and modified models. We found that KAOH and CHCH can give alarms and identify an earthquake with $M_L \geq 5$, but they cannot identify the Meinong earthquake within their detection ranges. They share the common parameter $Dep = 10$, but the depth of the Meinong earthquake is 14.64 km. In Fig. 8, we found that the retrained TP performances $\langle F1_f^o \rangle$

= 0.39 and $\langle F1_f^m \rangle$ = 0.53 (marked as red symbols) are all increased compared to their original values, which suggests that substituting the historic best parameters solved this no TIPs and no CGEQs problem in the training set. Although the retrained TN performances $\langle C1_f^o \rangle$ = 0.56 and $\langle C1_f^m \rangle$ = 0.62 both decrease, the retrained ensemble performances $\langle R_f^o \rangle$ = 0.81 and $\langle R_f^m \rangle$ = 0.88 remained almost the same.

The modified GEMSTIP model may help to forecast earthquakes, but some problems remain. For instance, optimizing the large number of models is extremely time-consuming and increasing the execution efficiency will be one important task. The abovementioned issue of the lasting anomalous time t_{ANO} also should be carefully investigated to increase the forecasting accuracy of large earthquakes. The selection of ways to quantify the success rate of a given model should be considered. We choose a C1-F1 diagram to optimize the GEMSTIP models. However, there are other performance scores in a binary classification, such as a ROC diagram, a Molchan diagram, etc. We have to understand the advantages and disadvantages of those performance scores and select the suitable ones for the GEMSTIP models.

In summary, the GEMSTIP model is useful in estimating the empirical relationship between earthquakes and anomalous statistical geoelectric field indexes. We first showed that the anomalous skewness and kurtosis of the geoelectric field appearing before large earthquakes. We next showed that the time-lag effect can exist between the end of clustered anomalies and a large earthquake, and that this effect depends on the local geological features. Third, we found that the durations of continuous anomalies are different, which means that the preparation process of each earthquake is different and shows spatial and temporal dependences. Finally, the GEMSTIP model identified the M_L 6.6 Meinong earthquake and other large earthquakes. The GEMSTIP model helps us to understand which factor affects the connection between earthquakes and geoelectric

Table 3. Best GESMTIP parameters of (a) the original model and (b) the modified model with retrained CHCH and KAOH. The (re)training set of CHCH is from 2012/4/26 to 2013/4/20 with a training window of 360 days, while the (re)training set of KAOH is from 2013/4/1 to 2014/3/26 with a training window of 360 days.

(a)

StaN	Rad (km)	Dep (km)	N_{thr}	t_{thr} (day)	t_{obs} (day)	t_{pred} (day)	C1	F1	R	PSA
CHCH	90	10	1	11	20	50	1	0.98	1.4	1
KAOH	90	10	1	11	50	60	1	0.98	1.4	1

(b)

StaN	Rad (km)	Dep (km)	N_{thr}	t_{thr} (day)	t_{obs} (day)	t_{pred} (day)	t_{lag} (day)	C1	F1	R	PSA
CHCH	90	10	1	11	20	50	0	1	0.98	1.4	1
KAOH	90	10	1	11	50	60	0	1	0.98	1.4	1
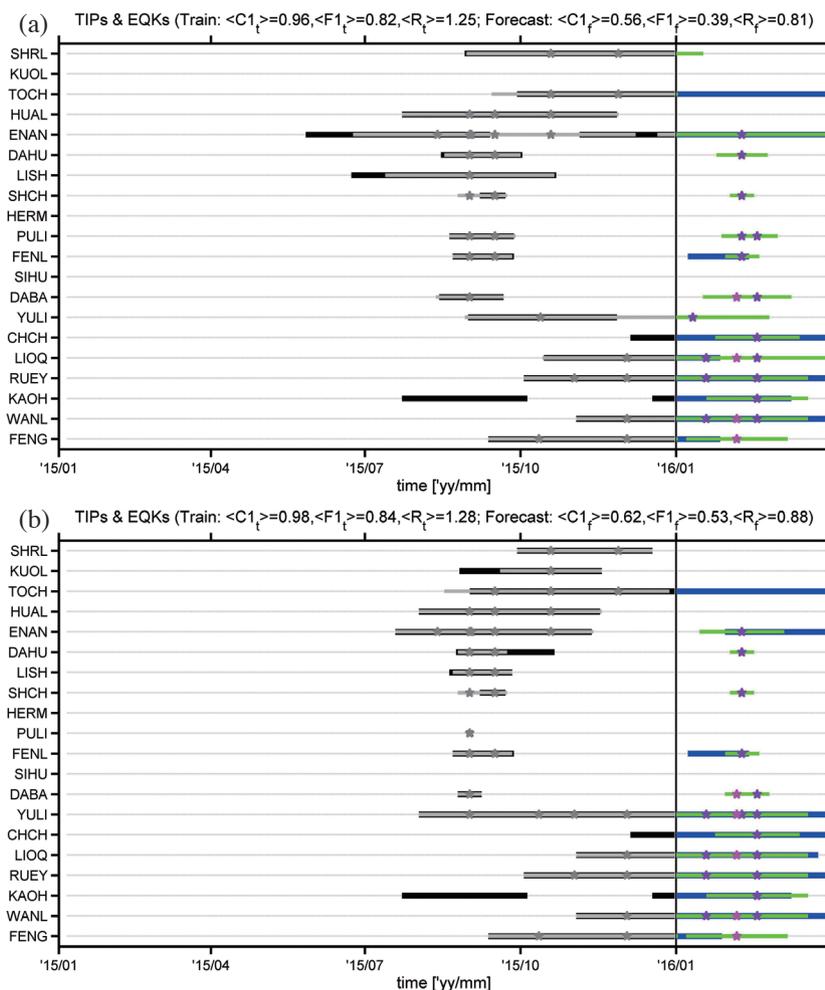
Fig. 10. CGEQ-TIP matching diagrams of (a) the original model and (b) the modified model with retrained KAOH and CHCH stations. The training set is from 2015/1/6 to 2015/12/31 with a training window of 360 days, and the forecasting set is from 2016/1/1 to 2016/3/31. Note that only KAOH and CHCH were retrained using the periods of 2013/4/1 through 2014/3/26 and 2012/4/26 through 2013/4/20, respectively; hence, one should neglect the TIP indexes of KAOH and CHCH in the training set from 2015/1/6 to 2015/12/31. (Color online only)

fields (which even could be extended to any geophysical data), and also forecast large earthquakes. Although some problems remain to be resolved, this model is promising for building an earthquake-forecasting system and make it practical in the future.

REFERENCES

Bertrand, E., M. Unsworth, C. W. Chiang, C. S. Chen, C. C. Chen, F. Wu, E. Türkoğlu, H. L. Hsu, and G. Hill, 2009: Magnetotelluric evidence for thick-skinned tectonics in central Taiwan. *Geology*, **37**, 711-714, doi:10.1130/G25755A.1. [Link]

Bertrand, E. A., M. J. Unsworth, C. W. Chiang, C. S. Chen, C. C. Chen, F. T. Wu, E. Türkoğlu, H. L. Hsu, and G. J. Hill, 2012: Magnetotelluric imaging beneath the Taiwan orogen: An arc-continent collision. *J. Geophys. Res.*, **117**, B01402, doi: 10.1029/2011JB008688. [Link]

Bowman, D. D., G. Ouillon, C. G. Sammis, A. Sornette, and D. Sornette, 1998: An observational test of the critical earthquake concept. *J. Geophys. Res.*, **103**, 24359-24372, doi: 10.1029/98JB00792. [Link]

Chen, C. and C. Chen, 1998: Preliminary result of magnetotelluric soundings in the fold-thrust belt of Taiwan and possible detection of dehydration. *Tectonophysics*, **292**, 101-117, doi: 10.1016/S0040-1951(98)00059-6. [Link]

Chen, C. C., C. Chen, and L. C. Sun, 1998: The possible causes of the crustal low resistive zone for the western foothills, Taiwan. *Terr. Atmos. Ocean. Sci.*, **9**, 279-285, doi: 10.3319/TAO.1998.9.2.279(T). [Link]

Chen, C. S. and C. C. Chen, 2000: Magnetotelluric soundings of the source area of the 1999 Chi-Chi Earthquake in Taiwan: Evidence of fluids at the hypocenter. *Terr. Atmos. Ocean. Sci.*, **11**, 679-688, doi: 10.3319/TAO.2000.11.3.679(CCE). [Link]

Chen, H. J. and C. C. Chen, 2016: Testing the correlations between anomalies of statistical indexes of the geoelectric system and earthquakes. *Nat. Hazards*, **84**, 877-895, doi: 10.1007/s11069-016-2460-4. [Link]

Dakos, V., S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kéfi, V. Livina, D. A. Seekell, E. H. van Nes, and M. Scheffer, 2012: Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE*, **7**, e41010, doi: 10.1371/journal.pone.0041010. [Link]

Eftaxias, K., 2010: Footprints of nonextensive Tsallis statistics, selfaffinity and universality in the preparation of the L'Aquila earthquake hidden in a pre-seismic EM emission. *Phys. Stat. Mech. Appl.*, **389**, 133-140, doi: 10.1016/j.physa.2009.08.034. [Link]

Eftaxias, K., P. Frangos, P. Kapiris, J. Polygiannakis, J. Kopanas, A. Peratzakis, P. Skountzos, and D. Jaggard, 2004: Review and a model of pre-seismic electromagnetic emissions in terms of fractal electrodynamics. *Fractals*, **12**, 243-273, doi: 10.1142/S0218348X04002501. [Link]

Eftaxias, K., V. E. Panin, and Y. Y. Deryugin, 2007: Evolution-EM signals before earthquakes in terms of mesomechanics and complexity. *Tectonophysics*, **431**, 273-300, doi: 10.1016/j.tecto.2006.05.041. [Link]

Eftaxias, K., G. Minadakis, S. M. Potirakis, and G. Balasis, 2013: Dynamical analogy between epileptic seizures and seismogenic electromagnetic emissions by means of nonextensive statistical mechanics. *Phys. Stat. Mech. Appl.*, **392**, 497-509, doi: 10.1016/j.physa.2012.09.015. [Link]

Freund, F. T., 2007: Pre-earthquake signals – Part I: Deviatoric stresses turn rocks into a source of electric currents. *Nat. Hazards Earth Syst. Sci.*, **7**, 535-541, doi: 10.5194/nhess-7-535-2007. [Link]

Freund, F. T. and S. Pilorz, 2012: Electric currents in the Earth crust and the generation of pre-earthquake ULF signals. In: Hayakawa, M. (Ed.), The Frontier of Earthquake Prediction Studies, Nihon Senmontosho Shuppan, Tokyo, 468-508.

Freund, F. T., I. G. Kulahci, G. Cyr, J. Ling, M. Winnick, J. Tregloan-Reed, and M. M. Freund, 2009: Air ionization at rock surfaces and pre-earthquake signals. *J. Atmos. Sol.-Terr. Phys.*, **71**, 1824-1834, doi: 10.1016/j.jastp.2009.07.013. [Link]

GEER Association, 2016: 2016 Meinong, Taiwan Earthquake, Geotechnical Extreme Events Reconnaissance (GEER) Association. Available at http://www.geerassociation.org/component/geer_reports/?view=geerreports&id=73.

Han, P., K. Hattori, G. Xu, R. Ashida, C. H. Chen, F. Febriani, and H. Yamaguchi, 2015: Further investigations of geomagnetic diurnal variations associated with the 2011 off the Pacific coast of Tohoku earthquake (Mw 9.0). *J. Asian Earth Sci.*, **114**, 321-326, doi: 10.1016/j.jseaes.2015.02.022. [Link]

Ishido, T. and H. Mizutani, 1981: Experimental and theoretical basis of electrokinetic phenomena in rock-water systems and its applications to geophysics. *J. Geophys. Res.*, **86**, 1763-1775, doi: 10.1029/JB086iB03p01763. [Link]

Kamiyama, M., M. Sugito, M. Kuse, A. Schekotov, and M. Hayakawa, 2016: On the precursors to the 2011 Tohoku earthquake: crustal movements and electromagnetic signatures. *Geomatics Natual Hazards and Risk*, **7**, 471-492, doi: 10.1080/19475705.2014.937773. [Link]

Kawase, T., S. Uyeda, M. Uyeshima, and M. Kinoshita, 1993: Possible correlation between geoelectric potential change in Izu-Oshima Island and the earthquake swarm off the east Izu Peninsula, Japan. *Tectonophysics*, **224**, 83-93, doi: 10.1016/0040-1951(93)90059-S. [Link]

Londos, C. A., N. Sarlis, L. G. Fytros, and K. Papastergiou, 1996: Precursor defect to the vacancy-dioxygen center in Si. *Phys. Rev. B*, **53**, 6900-6903, doi: 10.1103/PhysRevB.53.6900. [Link]

NCDR, 2016: 0206 Meinong Earthquake, National Science and Technology Center for Disaster Reduction. Available at http://www.ncdr.nat.gov.tw/EarthquakeMeinong1050206.aspx.

Nitsan, U., 1977: Electromagnetic emission accompanying fracture of quartz-bearing rocks. *Geophys. Res. Lett.*, **4**, 333-336, doi: 10.1029/GL004i008p00333. [Link]

Orihara, Y., M. Kamogawa, T. Nagao, and S. Uyeda, 2009: Independent component analysis of geoelectric field data in the northern Nagano, Japan. *Proc. Jpn. Acad. B Phys. Biol. Sci.*, **85**, 435-442, doi: 10.2183/pjab.85.435. [Link]

Orihara, Y., M. Kamogawa, T. Nagao, and S. Uyeda, 2012: Variations of geoelectric potential differences associated with an anomalous volumetric strain change in the region of expected Tokai Earthquake, Japan. *Nat. Hazards Earth Syst. Sci.*, **12**, 121-127, doi: 10.5194/nhess-12-121-2012. [Link]

Ouzounov, D., S. Pulinets, A. Romanov, A. Romanov, K. Tsybulya, D. Davidenko, M. Kafatos, and P. Taylor, 2011: Atmosphere-ionosphere response to the M9 Tohoku earthquake revealed by multi-instrument spaceborne and ground observations: Preliminary results. *Earthquake Sci.*, **24**, 557-564, doi: 10.1007/s11589-011-0817-z. [Link]

Ramirez-Rojas, A., R. T. Paez-Hernandez, and J. R. Luevano, 2013: Cross-correlation analysis for geoelectric time series associated with an earthquake by means of mutual information theory. *Revista Mexicana de Física S*, **59**, 14-17.

Sammis, C. G. and D. Sornette, 2002: Positive feedback, memory, and the predictability of earthquakes. *Proc. Natl. Acad. Sci. USA*, **99**, 2501-2508, doi: 10.1073/pnas.012580999. [Link]

Sarlis, N. V., E. S. Skordas, P. A. Varotsos, T. Nagao, M. Kamogawa, H. Tanaka, and S. Uyeda, 2013: Minimum of the order parameter fluctuations of seismicity before major earthquakes in Japan. *Proc. Natl. Acad. Sci. USA*, **110**, 13734-13738, doi: 10.1073/pnas.1312740110. [Link]

Scheffer, M., 2009: Critical Transitions in Nature and Society, Princeton University Press, New Jersey.

Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara, 2009: Early-warning signals for critical transitions. *Nature*, **461**, 53-59, doi: 10.1038/nature08227. [Link]

Skordas, E. S. and N. V. Sarlis, 2014: On the anomalous changes of seismicity and geomagnetic field prior to the 2011 9.0 Tohoku earthquake. *J. Asian Earth Sci.*, **80**, 161-164, doi: 10.1016/j.jseaes.2013.11.008. [Link]

Sornette, A. and D. Sornette, 1990: Earthquake rupture as a critical point: consequences for telluric precursors. *Tectonophysics*, **179**, 327-334, doi: 10.1016/0040-1951(90)90298-M. [Link]

Sornette, D. and C. G. Sammis, 1995: Complex critical exponents from renormalization group theory of earthquakes: Implications for earthquake predictions. *J. Phys.*, **5**, 607-619, doi: 10.1051/jp1:1995154. [Link]

Telesca, L., G. Colangelo, V. Lapenna, and M. Macchiato, 2004: Fluctuation dynamics in geoelectrical data: An investigation by using multifractal detrended fluctuation analysis. *Phys. Lett.*, **332**, 398-404, doi: 10.1016/j.physleta.2004.10.011. [Link]

Telesca, L., V. Lapenna, and M. Macchiato, 2005: Multifractal fluctuations in earthquake-related geoelectrical signals. *New J. Phys.*, **7**, doi: 10.1088/1367-2630/7/1/214. [Link]

Telesca, L., V. Lapenna, M. Macchiato, and K. Hattori, 2008: Investigating non-uniform scaling behavior in Ultra Low Frequency (ULF) earthquake-related geomagnetic signals. *Earth Planet. Sci. Lett.*, **268**, 219-224, doi: 10.1016/j.epsl.2008.01.033. [Link]

Telesca, L., M. Lovallo, A. Ramirez-Rojas, and F. Angulo-Brown, 2009: A nonlinear strategy to reveal seismic precursory signatures in earthquake-related self-potential signals. *Phys. Stat. Mech. Appl.*, **388**, 2036-2040, doi: 10.1016/j.physa.2009.01.035. [Link]

Tibshirani, R., 1996: Regression shrinkage and selection via the lasso: A retrospective. *J. Roy. Stat. Soc. B Stat. Meth.*, **58**, 267-288, doi: 10.1111/j.1467-9868.2011.00771.x. [Link]

Triantis, D., C. Anastasiadis, and I. Stavrakas, 2008: The correlation of electrical charge with strain on stressed rock samples. *Nat. Hazards Earth Syst. Sci.*, **8**, 1243-1248, doi: 10.5194/nhess-8-1243-2008. [Link]

Uyeda, S., T. Nagao, Y. Orihara, T. Yamaguchi, and I. Takahashi, 2000: Geoelectric potential changes: Possible

precursors to earthquakes in Japan. *Proc. Natl. Acad. Sci. USA*, **97**, 4561-4566, doi: 10.1073/pnas.97.9.4561. [Link]

Uyeda, S., M. Hayakawa, T. Nagao, O. Molchanov, K. Hattori, Y. Orihara, K. Gotoh, Y. Akinaga, and H. Tanaka, 2002: Electric and magnetic phenomena observed before the volcano-seismic activity in 2000 in the Izu Island Region, Japan. *Proc. Natl. Acad. Sci. USA*, **99**, 7352-7355, doi: 10.1073/pnas.072208499. [Link]

Varotsos, P. A. and K. Alexopoulos, 1984a: Physical properties of the variations of the electric field of the earth preceding earthquakes, I. *Tectonophysics*, **110**, 73-98, doi: 10.1016/0040-1951(84)90059-3. [Link]

Varotsos, P. A. and K. Alexopoulos, 1984b: Physical properties of the variations of the electric field of the earth preceding earthquakes. II. Determination of epicenter and magnitude. *Tectonophysics*, **110**, 99-125, doi: 10.1016/0040-1951(84)90060-X. [Link]

Varotsos, P. A., K. Alexopoulos, and M. Lazaridou, 1993: Latest aspects of earthquake prediction in Greece based on seismic electric signals, II. *Tectonophysics*, **224**, 1-37, doi: 10.1016/0040-1951(93)90055-O. [Link]

Varotsos, P. A., N. V. Sarlis, and E. S. Skordas, 2003: Attempt to distinguish electric signals of a dichotomous nature. *Phys. Rev. E*, **68**, doi: 10.1103/PhysRevE.68.031106. [Link]

Varotsos, P. A., N. V. Sarlis, E. S. Skordas, H. K. Tanaka, and M. S. Lazaridou, 2006a: Attempt to distinguish long-range temporal correlations from the statistics of the increments by natural time analysis. *Phys. Rev. E*, **74**, doi: 10.1103/PhysRevE.74.021123. [Link]

Varotsos, P. A., N. V. Sarlis, E. S. Skordas, H. K. Tanaka, and M. S. Lazaridou, 2006b: Entropy of seismic electric signals: Analysis in natural time under time reversal. *Phys. Rev. E*, **73**, doi: 10.1103/PhysRevE.73.031114. [Link]

Varotsos, P. A., N. V. Sarlis, E. S. Skordas, and M. S. Lazaridou, 2008: Fluctuations, under time reversal, of the natural time and the entropy distinguish similar looking electric signals of different dynamics. *J. Appl. Phys.*, **103**, doi: 10.1063/1.2827363. [Link]

Varotsos, P. A., N. V. Sarlis, and E. S. Skordas, 2009: Detrended fluctuation analysis of the magnetic and electric field variations that precede rupture. *Chaos*, **19**, doi: 10.1063/1.3130931. [Link]

Varotsos, P. A., N. V. Sarlis, and E. S. Skordas, 2011a: Natural Time Analysis: The New View of Time, Springer Berlin Heidelberg, 449 pp, doi: 10.1007/978-3-642-16449-1. [Link]

Varotsos, P. A., N. V. Sarlis, and E. S. Skordas, 2011b: Scale-specific order parameter fluctuations of seismicity in natural time before mainshocks. *Europhys. Lett.*, **96**, doi: 10.1209/0295-5075/96/59002. [Link]

Varotsos, P. A., N. V. Sarlis, E. S. Skordas, and M. S. Lazaridou, 2013: Seismic Electric Signals: An additional fact showing their physical interconnection with seismicity. *Tectonophysics*, **589**, 116-125, doi: 10.1016/j.tecto.2012.12.020. [Link]

Wikipedia, 2016: 2016 Taiwan Earthquake. Available at https://en.wikipedia.org/wiki/2016_Taiwan_earthquake.

Xu, G., P. Han, Q. Huang, K. Hattori, F. Febriani, and H. Yamaguchi, 2013: Anomalous behaviors of geomagnetic diurnal variations prior to the 2011 off the Pacific coast of Tohoku earthquake (Mw9.0). *J. Asian Earth Sci.*, **77**, 59-65, doi: 10.1016/j.jseaes.2013.08.011. [Link]
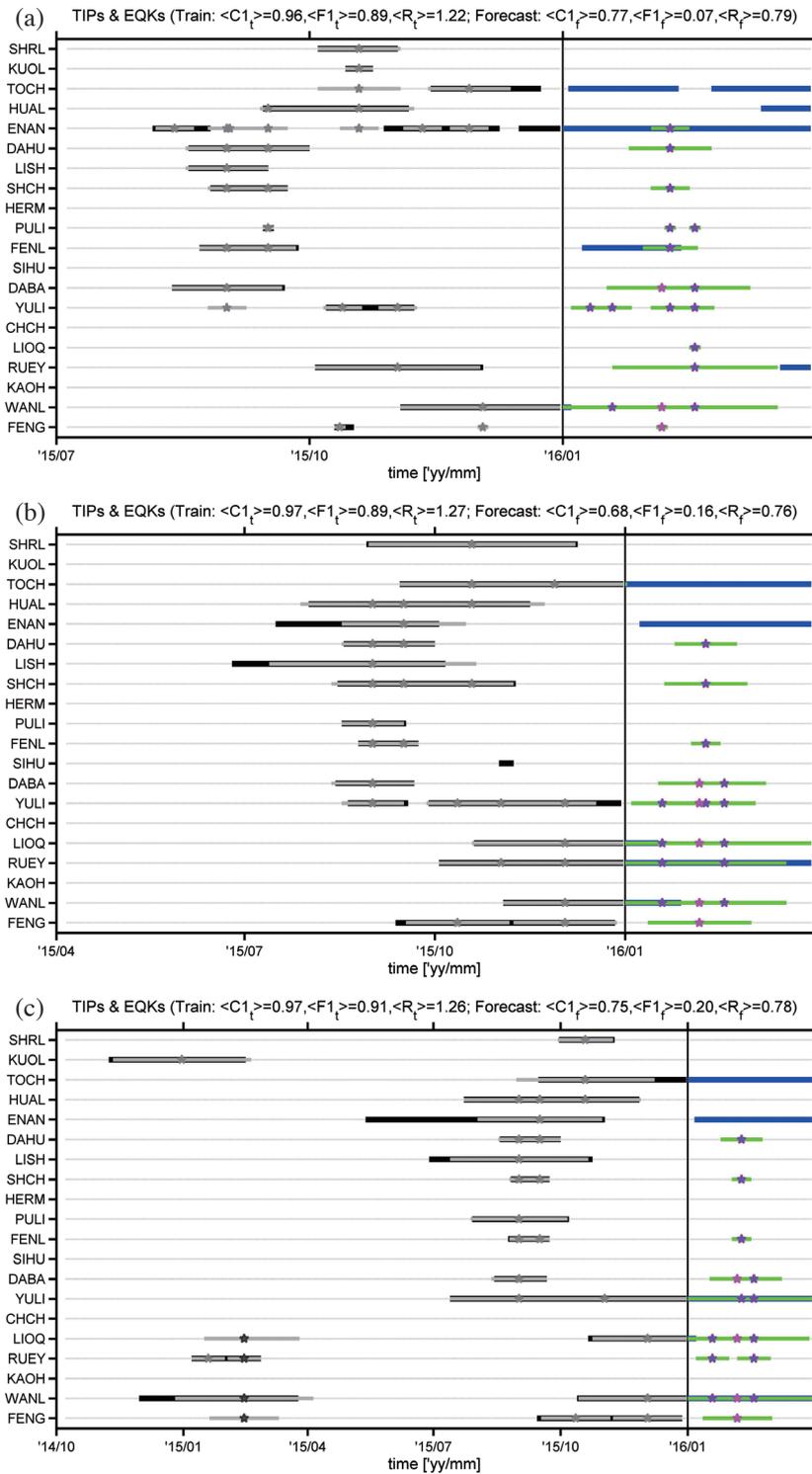
Fig. S1. CGEQ-TIP matching diagrams of the best original models with different training windows (T_w). (a) The training set is from 2015/7/5 to 2015/12/31 (T_w = 180 days), (b) the training set from 2015/4/6 to 2015/12/31 (T_w = 270 days), (c) the training set from 2014/10/8 to 2015/12/31 (T_w = 450 days), (d) the training set from 2014/7/10 to 2015/12/31 (T_w = 540 days), (e) the training set from 2014/4/11 to 2015/12/31 (T_w = 630 days), (f) the training set from 2014/1/11 to 2015/12/31 (T_w = 720 days), (g) the training set from 2013/10/13 to 2015/12/31 (T_w = 810 days), (h) the training set from 2013/7/15 to 2015/12/31 (T_w = 900 days), (i) the training set from 2013/4/16 to 2015/12/31 (T_w = 990 days), and (j) the training set from 2013/1/16 to 2015/12/31 (T_w = 1080 days). Black and blue horizontal lines are TIPs, while gray and green ones are CGEQs. Open stars are earthquakes with $M_L \in [5, 6)$, while solid ones are earthquakes with $M_L \geq 6$. Gray-level part is the training set, while colorful part is the forecasting set. Gray dotted lines are auxiliary.

Fig. S1. (Continued)
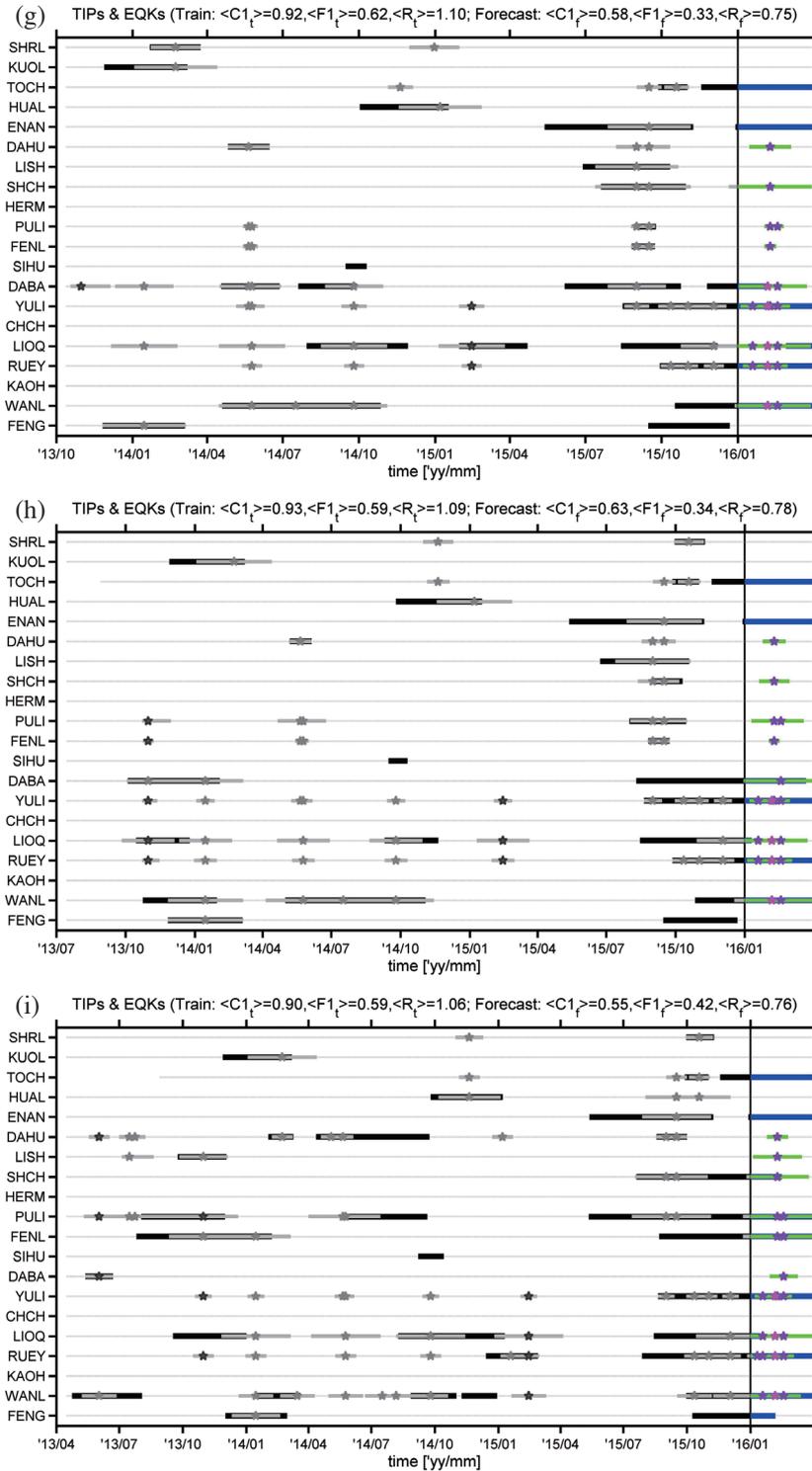
Fig. S1. (Continued)
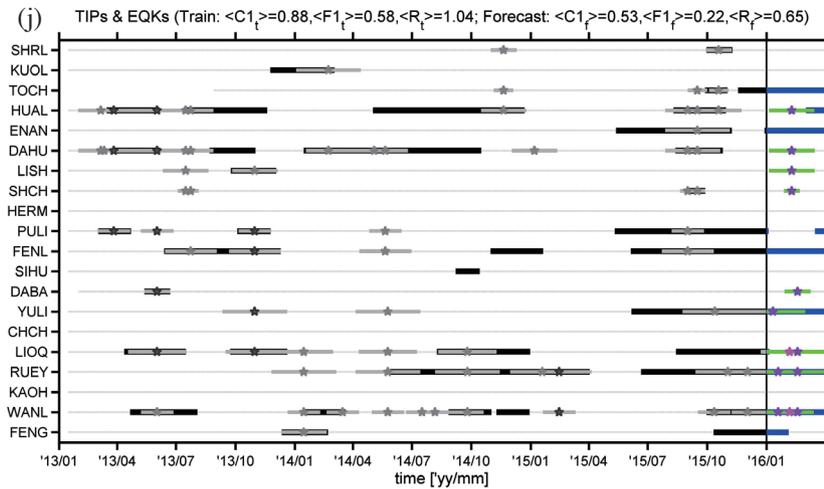
Fig. S1. (Continued)
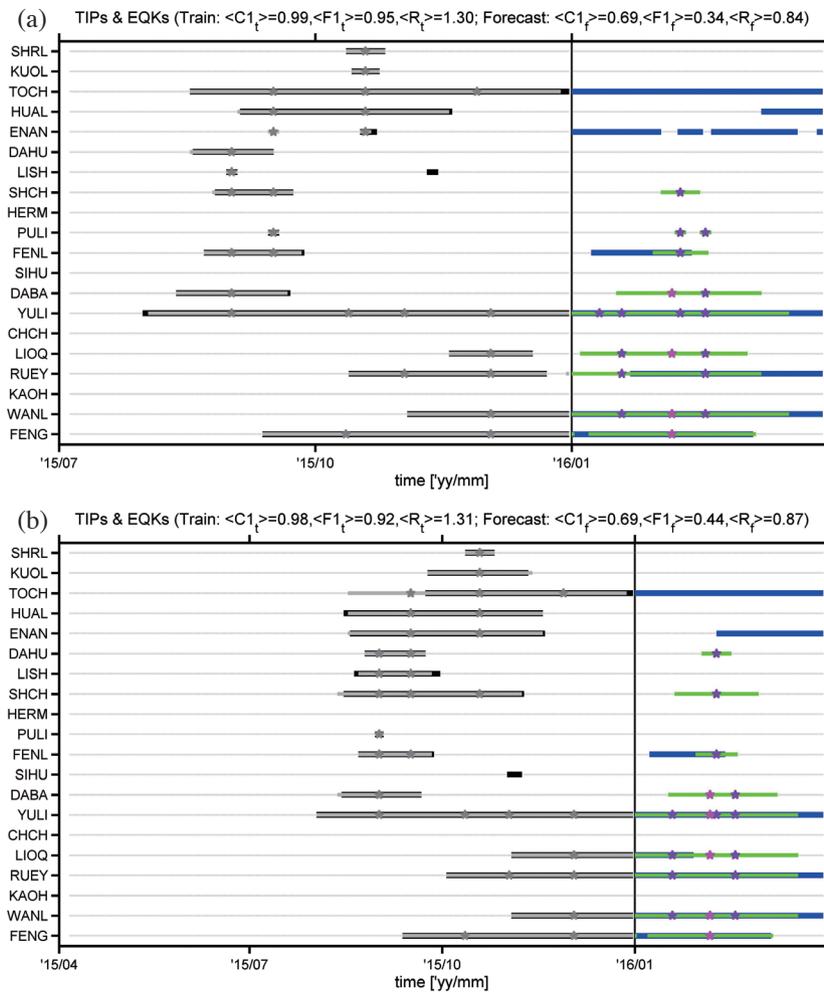
Fig. S2. CGEQ-TIP matching diagrams of the best modified models with different training windows (T_w). (a) The training set is from 2015/7/5 to 2015/12/31 (T_w = 180 days), (b) the training set from 2015/4/6 to 2015/12/31 (T_w = 270 days), (c) the training set from 2014/10/8 to 2015/12/31 (T_w = 450 days), (d) the training set from 2014/7/10 to 2015/12/31 (T_w = 540 days), (e) the training set from 2014/4/11 to 2015/12/31 (T_w = 630 days), (f) the training set from 2014/1/11 to 2015/12/31 (T_w = 720 days), (g) the training set from 2013/10/13 to 2015/12/31 (T_w = 810 days), (h) the training set from 2013/7/15 to 2015/12/31 (T_w = 900 days), (i) the training set from 2013/4/16 to 2015/12/31 (T_w = 990 days), and (j) the training set from 2013/1/16 to 2015/12/31 (T_w = 1080 days). Black and blue horizontal lines are TIPs, while gray and green ones are CGEQs. Open stars are earthquakes with $M_L \in [5, 6)$, while solid ones are earthquakes with $M_L \geq 6$. Gray-level part is the training set, while colorful part is the forecasting set. Gray dotted lines are auxiliary.

Fig. S2. (Continued)

(f) TIPs & EQKs (Train: $<C1_t>=0.95, <F1_t>=0.74, <R_t>=1.18$; Forecast: $<C1_f>=0.65, <F1_f>=0.39, <R_f>=0.82$)

(g) TIPs & EQKs (Train: $<C1_t>=0.97, <F1_t>=0.71, <R_t>=1.18$; Forecast: $<C1_f>=0.72, <F1_f>=0.26, <R_f>=0.84$)

(h) TIPs & EQKs (Train: $<C1_t>=0.97, <F1_t>=0.64, <R_t>=1.15$; Forecast: $<C1_f>=0.72, <F1_f>=0.41, <R_f>=0.88$)

Fig. S2. (Continued)

(i) TIPs & EQKs (Train: $<C1_t>=0.97, <F1_t>=0.57, <R_t>=1.13$; Forecast: $<C1_f>=0.68, <F1_f>=0.19, <R_f>=0.80$)

(j) TIPs & EQKs (Train: $<C1_t>=0.96, <F1_t>=0.50, <R_t>=1.11$; Forecast: $<C1_f>=0.65, <F1_f>=0.31, <R_f>=0.83$)

Fig. S2. (Continued)

Chen et al.
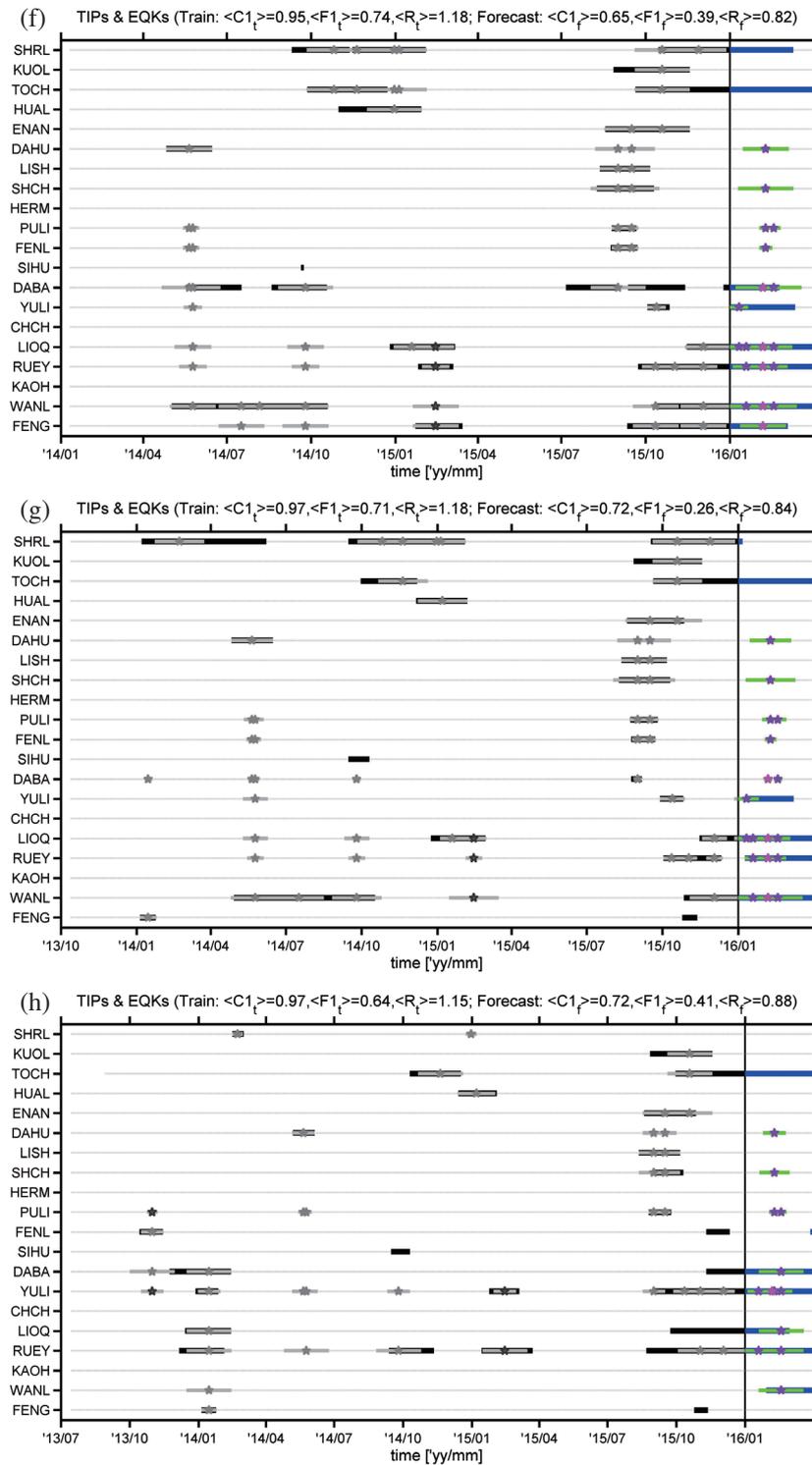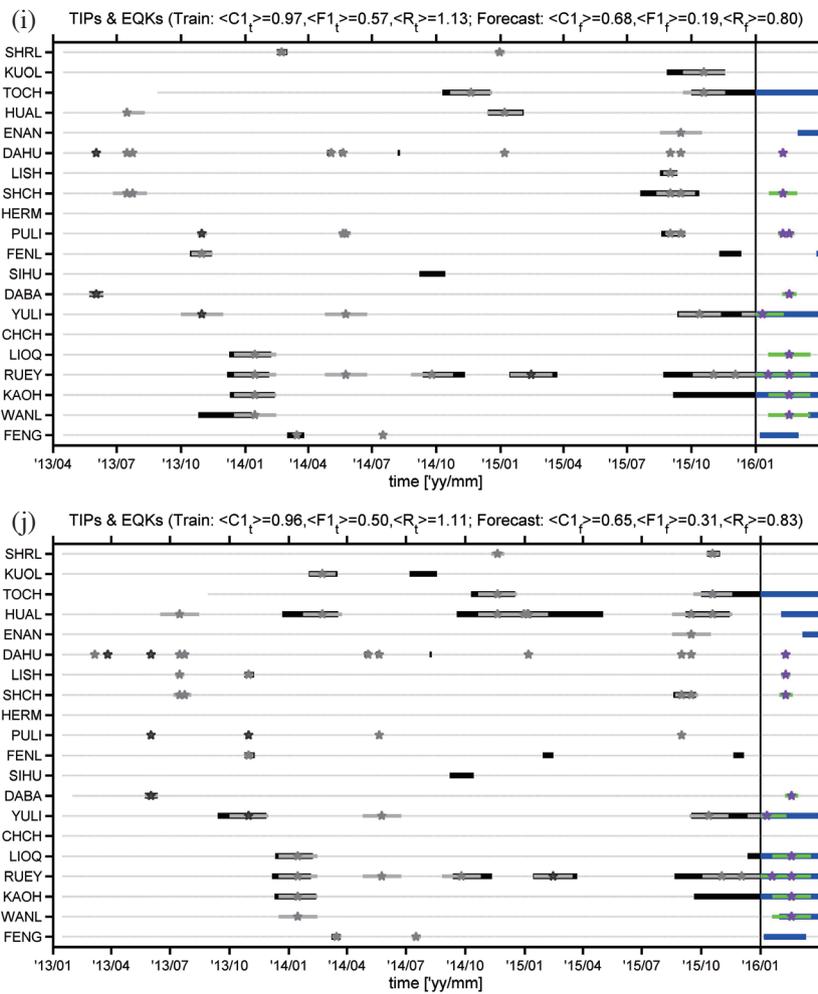
Table S1. Best parameters of the original GEMSTIP models with different training windows (T_w). (a) The training set is from 2015/7/5 to 2015/12/31 (T_w = 180 days), (b) the training set from 2015/4/6 to 2015/12/31 (T_w = 270 days), (c) the training set from 2014/10/8 to 2015/12/31 (T_w = 450 days), (d) the training set from 2014/7/10 to 2015/12/31 (T_w = 540 days), (e) the training set from 2014/4/11 to 2015/12/31 (T_w = 630 days), (f) the training set from 2014/1/11 to 2015/12/31 (T_w = 720 days), (g) the training set from 2013/10/13 to 2015/12/31 (T_w = 810 days), (h) the training set from 2013/7/15 to 2015/12/31 (T_w = 900 days), (i) the training set from 2013/4/16 to 2015/12/31 (T_w = 990 days), and (j) the training set from 2013/1/16 to 2015/12/31 (T_w = 1080 days).

(a)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	3	13	90	30	0.99	0.98	1.4	1
KUOL	90	20	2	3	5	10	1	1	1.41	1
TOCH	60	250	2	3	5	30	0.9	0.84	1.23	1
HUAL	70	20	2	5	70	40	0.97	0.97	1.38	1
ENAN	100	250	3	2	10	15	0.54	0.55	0.77	2
DAHU	100	20	1	5	25	30	1	0.99	1.4	1
LISH	45	20	1	8	15	30	1	0.98	1.4	1
SHCH	30	20	2	5	30	15	1	0.98	1.4	1
HERM	30	10	2	3	5	5	1	0	1	5
PULI	100	20	2	21	70	5	1	1	1.41	1
FENL	80	20	3	1	20	20	1	0.99	1.4	1
SIHU	30	10	4	4	5	5	1	0	1	5
DABA	100	20	1	4	5	40	1	0.99	1.4	1
YULI	80	50	4	1	5	15	0.92	0.72	1.17	1
CHCH	30	10	3	2	5	5	1	0	1	5
LIOQ	30	10	3	2	5	5	1	0	1	5
RUEY	55	20	1	11	30	60	0.99	0.99	1.4	1
KAOH	30	10	1	3	5	5	1	0	1	5
WANL	90	30	4	8	15	60	1	1	1.41	1
FENG	90	30	1	23	70	5	0.96	0.56	1.11	1

(b)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	3	13	60	100	0.99	0.99	1.4	1
KUOL	30	10	4	4	5	5	1	0	1	5
TOCH	60	250	1	2	5	70	1	1	1.41	1
HUAL	90	20	2	5	50	70	0.95	0.95	1.34	1
ENAN	30	20	3	15	100	60	0.77	0.68	1.03	1
DAHU	100	20	1	5	25	30	1	0.99	1.4	1
LISH	45	20	1	10	20	100	0.9	0.86	1.24	1
SHCH	100	20	1	15	80	40	0.98	0.98	1.38	1
HERM	30	10	2	3	5	5	1	0	1	5
PULI	55	20	4	3	40	30	1	0.98	1.4	1
FENL	80	20	2	8	30	15	1	1	1.41	1
SIHU	30	10	4	7	10	5	1	0	1	5
DABA	100	20	1	4	5	40	1	0.97	1.39	1
YULI	100	30	1	5	10	30	0.94	0.92	1.32	1
CHCH	30	10	3	2	5	5	1	0	1	5
LIOQ	90	30	1	33	80	90	1	0.99	1.41	1
RUEY	55	30	2	7	40	60	1	1	1.41	1
KAOH	30	10	1	3	5	5	1	0	1	5
WANL	90	30	3	10	15	60	1	1	1.41	1
FENG	90	30	1	17	60	50	0.96	0.96	1.35	1

Table S1. (Continued)

(c)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	4	7	90	40	1	0.98	1.39	1
KUOL	70	100	1	19	30	100	0.99	0.97	1.38	1
TOCH	30	20	1	21	100	100	0.95	0.9	1.31	1
HUAL	90	20	1	13	90	80	1	1	1.41	1
ENAN	30	20	3	15	100	90	0.81	0.69	1.06	1
DAHU	100	20	2	5	25	30	1	0.99	1.41	1
LISH	45	20	2	25	100	100	0.96	0.92	1.33	1
SHCH	30	20	2	5	30	15	1	0.98	1.4	1
HERM	30	10	1	4	5	5	1	0	1	5
PULI	55	20	4	2	20	70	1	0.99	1.4	1
FENL	80	20	3	3	30	15	1	0.98	1.4	1
SIHU	30	10	4	10	15	5	1	0	1	5
DABA	100	20	2	4	5	40	1	0.97	1.39	1
YULI	80	20	1	29	70	100	0.98	0.97	1.38	1
CHCH	30	10	3	2	5	5	1	0	1	5
LIOQ	90	30	2	39	100	80	0.85	0.65	1.07	1
RUEY	45	150	4	3	90	25	1	0.99	1.41	1
KAOH	30	10	2	4	5	5	1	0	1	5
WANL	90	30	3	3	40	100	0.91	0.9	1.28	1
FENG	90	30	2	9	70	50	0.89	0.79	1.19	1

(d)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	1	9	20	100	0.98	0.97	1.37	1
KUOL	70	100	1	45	100	100	0.93	0.78	1.22	1
TOCH	70	20	1	8	15	30	0.9	0.58	1.07	1
HUAL	90	100	2	7	100	60	0.97	0.97	1.38	1
ENAN	30	20	4	11	100	100	0.84	0.68	1.08	1
DAHU	100	20	2	11	90	15	0.97	0.72	1.21	1
LISH	45	20	2	25	100	100	0.97	0.92	1.34	1
SHCH	80	30	2	5	40	50	1	0.98	1.4	1
HERM	30	10	1	4	5	5	1	0	1	5
PULI	55	20	1	43	80	70	1	0.99	1.41	1
FENL	30	20	2	6	15	50	1	0.99	1.41	1
SIHU	30	10	4	53	80	5	1	0	1	5
DABA	100	20	1	29	90	10	1	0.95	1.38	1
YULI	40	50	1	51	90	100	0.96	0.87	1.29	1
CHCH	30	10	3	3	5	5	1	0	1	5
LIOQ	70	20	3	3	5	30	1	0.97	1.39	1
RUEY	55	20	3	13	100	60	0.97	0.84	1.28	1
KAOH	30	10	3	4	5	5	1	0	1	5
WANL	100	30	1	21	100	50	0.89	0.74	1.16	1
FENG	90	30	1	11	90	50	0.89	0.83	1.22	1

Table S1. (Continued)

(e)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	1	8	20	100	0.98	0.96	1.37	1
KUOL	90	20	3	9	100	100	0.78	0.53	0.94	2
TOCH	70	20	1	17	70	80	0.85	0.67	1.08	1
HUAL	90	100	1	7	100	40	0.79	0.74	1.08	1
ENAN	30	20	4	11	100	100	0.9	0.72	1.15	1
DAHU	100	20	4	6	10	50	0.94	0.61	1.12	1
LISH	45	20	2	25	100	100	0.97	0.9	1.32	1
SHCH	30	20	2	11	90	30	0.99	0.93	1.36	1
HERM	30	10	1	4	5	5	1	0	1	5
PULI	55	20	2	35	80	70	1	0.99	1.4	1
FENL	80	20	3	5	100	15	1	0.97	1.39	1
SIHU	30	10	4	61	100	5	1	0	1	5
DABA	100	20	4	7	80	30	0.97	0.64	1.16	1
YULI	40	50	3	27	100	100	0.96	0.87	1.3	1
CHCH	30	10	4	4	5	5	1	0	1	5
LIOQ	70	20	3	3	5	30	0.97	0.64	1.16	1
RUEY	70	30	2	23	70	30	0.91	0.69	1.15	1
KAOH	30	10	3	4	5	5	1	0	1	5
WANL	90	20	1	41	100	100	0.92	0.77	1.2	1
FENG	90	100	1	5	30	50	0.82	0.79	1.14	1

(f)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	2	9	80	100	0.94	0.89	1.3	1
KUOL	90	20	3	9	100	100	0.82	0.53	0.98	2
TOCH	70	20	1	17	70	80	0.88	0.67	1.1	1
HUAL	80	20	2	1	5	100	0.78	0.69	1.04	1
ENAN	30	20	4	11	100	100	0.91	0.71	1.16	1
DAHU	100	20	4	6	10	50	0.95	0.61	1.13	1
LISH	45	20	2	29	100	100	0.98	0.9	1.33	1
SHCH	30	20	3	3	30	100	0.99	0.94	1.36	1
HERM	30	10	3	3	5	5	1	0	1	5
PULI	100	20	4	5	40	25	0.98	0.72	1.21	1
FENL	80	20	3	4	30	15	0.98	0.74	1.23	1
SIHU	30	10	4	67	100	5	1	0	1	5
DABA	100	20	4	2	5	80	0.73	0.67	0.99	2
YULI	100	30	2	49	80	30	0.87	0.65	1.09	1
CHCH	30	10	4	4	5	5	1	0	1	5
LIOQ	90	30	1	15	50	80	0.7	0.55	0.89	2
RUEY	70	30	2	12	30	20	0.92	0.59	1.1	1
KAOH	30	10	3	4	5	5	1	0	1	5
WANL	90	20	1	33	100	100	0.89	0.84	1.23	1
FENG	90	30	1	7	50	40	0.79	0.7	1.06	1

Table S1. (Continued)

(g)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	55	100	2	51	100	60	0.95	0.66	1.16	1
KUOL	55	100	1	16	25	100	0.95	0.64	1.15	1
TOCH	70	20	1	7	10	30	0.94	0.61	1.12	1
HUAL	60	40	1	17	100	100	0.93	0.58	1.1	1
ENAN	30	20	4	11	100	100	0.93	0.71	1.17	1
DAHU	100	20	4	5	10	50	0.95	0.61	1.13	1
LISH	45	20	2	29	100	100	0.98	0.9	1.33	1
SHCH	30	20	3	3	30	100	0.99	0.94	1.37	1
HERM	30	10	2	4	5	5	1	0	1	5
PULI	100	20	4	9	80	15	0.98	0.68	1.2	1
FENL	80	20	3	4	30	15	0.99	0.74	1.24	1
SIHU	30	10	4	67	100	5	0.98	0	0.98	5
DABA	100	20	4	2	5	70	0.69	0.55	0.88	2
YULI	100	30	4	23	100	30	0.88	0.63	1.08	1
CHCH	30	10	4	4	5	5	1	0	1	5
LIOQ	90	30	2	3	5	80	0.61	0.51	0.79	2
RUEY	70	30	2	33	100	25	0.91	0.57	1.07	1
KAOH	30	10	3	4	5	5	1	0	1	5
WANL	90	20	1	29	100	80	0.88	0.79	1.18	1
FENG	90	10	1	4	5	100	0.92	0.67	1.14	1

(h)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	4	7	90	40	0.97	0.65	1.17	1
KUOL	55	100	1	16	25	100	0.95	0.64	1.15	1
TOCH	70	20	1	7	10	30	0.95	0.61	1.13	1
HUAL	60	40	1	17	100	100	0.93	0.56	1.09	1
ENAN	30	20	4	11	100	100	0.94	0.71	1.18	1
DAHU	100	20	4	7	20	30	0.97	0.54	1.11	1
LISH	45	20	2	27	90	100	0.98	0.9	1.33	1
SHCH	30	20	3	5	70	40	1	0.96	1.38	1
HERM	30	10	2	4	5	5	1	0	1	5
PULI	100	20	2	63	100	60	0.92	0.59	1.09	1
FENL	80	20	3	4	30	15	0.98	0.62	1.16	1
SIHU	30	10	4	67	100	5	0.98	0	0.98	5
DABA	80	10	3	7	80	100	0.86	0.58	1.04	1
YULI	100	30	1	61	100	25	0.87	0.54	1.02	1
CHCH	30	10	4	4	5	5	1	0	1	5
LIOQ	100	30	2	11	50	70	0.66	0.52	0.84	2
RUEY	70	30	2	31	100	30	0.87	0.51	1.01	1
KAOH	30	10	3	4	5	5	1	0	1	5
WANL	90	20	1	31	100	100	0.81	0.77	1.12	1
FENG	90	10	1	4	5	100	0.93	0.67	1.15	1

Table S1. (Continued)

(i)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	4	7	90	40	0.97	0.65	1.17	1
KUOL	55	100	1	16	25	100	0.96	0.64	1.15	1
TOCH	70	20	1	7	10	30	0.95	0.61	1.13	1
HUAL	80	20	1	17	100	90	0.88	0.51	1.01	1
ENAN	30	20	4	11	100	100	0.95	0.71	1.18	1
DAHU	100	100	1	5	25	30	0.85	0.52	0.99	2
LISH	80	10	3	13	80	70	0.97	0.73	1.21	1
SHCH	30	20	3	5	100	90	0.93	0.66	1.14	1
HERM	30	10	2	4	5	5	1	0	1	5
PULI	100	20	4	1	25	100	0.66	0.63	0.92	2
FENL	100	10	4	1	100	100	0.87	0.7	1.11	1
SIHU	30	10	3	67	100	5	0.98	0	0.98	5
DABA	55	20	4	6	10	40	1	1	1.41	1
YULI	100	30	4	25	100	25	0.87	0.52	1.01	1
CHCH	30	10	4	4	5	5	1	0	1	5
LIOQ	90	30	2	7	25	100	0.6	0.52	0.79	2
RUEY	70	150	3	9	90	30	0.83	0.56	1	2
KAOH	30	10	3	4	5	5	1	0	1	5
WANL	100	100	4	1	5	50	0.71	0.69	0.99	2
FENG	90	10	1	27	90	70	0.94	0.62	1.13	1

(j)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	C1	F1	R	PSA
SHRL	60	20	4	7	90	40	0.98	0.65	1.17	1
KUOL	55	100	1	14	20	100	0.96	0.6	1.13	1
TOCH	70	20	1	7	10	30	0.95	0.61	1.13	1
HUAL	100	20	2	1	15	70	0.69	0.6	0.92	2
ENAN	30	20	4	11	100	100	0.94	0.68	1.17	1
DAHU	100	100	1	2	5	70	0.65	0.6	0.88	2
LISH	80	10	3	13	80	70	0.96	0.64	1.15	1
SHCH	30	20	3	7	90	25	0.98	0.57	1.13	1
HERM	30	10	2	4	5	5	1	0	1	5
PULI	55	20	3	2	5	50	0.81	0.51	0.96	2
FENL	30	20	3	1	5	80	0.77	0.61	0.98	2
SIHU	30	10	4	67	100	5	0.98	0	0.98	5
DABA	55	20	4	6	10	40	1	1	1.41	1
YULI	40	50	3	5	50	100	0.83	0.63	1.04	1
CHCH	30	10	4	4	5	5	1	0	1	5
LIOQ	100	20	2	7	25	90	0.69	0.59	0.91	2
RUEY	55	150	2	5	50	100	0.73	0.73	1.03	1
KAOH	30	10	3	4	5	5	1	0	1	5
WANL	100	100	4	1	5	50	0.75	0.69	1.02	1
FENG	90	10	1	29	90	70	0.95	0.62	1.14	1

Table S2. Best parameters of the modified GEMSTIP models with different training windows (T_w). (a) The training set is from 2015/7/5 to 2015/12/31 (T_w = 180 days), (b) the training set from 2015/4/6 to 2015/12/31 (T_w = 270 days), (c) the training set from 2014/10/8 to 2015/12/31 (T_w = 450 days), (d) the training set from 2014/7/10 to 2015/12/31 (T_w = 540 days), (e) the training set from 2014/4/11 to 2015/12/31 (T_w = 630 days), (f) the training set from 2014/1/11 to 2015/12/31 (T_w = 720 days), (g) the training set from 2013/10/13 to 2015/12/31 (T_w = 810 days), (h) the training set from 2013/7/15 to 2015/12/31 (T_w = 900 days), (i) the training set from 2013/4/16 to 2015/12/31 (T_w = 990 days), and (j) the training set from 2013/1/16 to 2015/12/31 (T_w = 1080 days).

(a)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	60	20	2	9	15	15	50	1	1	1.41	1
KUOL	90	20	2	3	5	10	0	1	1	1.41	1
TOCH	70	300	1	1	5	60	10	0.96	0.99	1.38	1
HUAL	70	20	2	3	60	60	0	0.98	0.99	1.39	1
ENAN	50	20	1	2	5	5	70	0.98	0.59	1.14	1
DAHU	90	20	1	4	10	30	0	1	0.98	1.4	1
LISH	45	20	1	4	5	5	20	0.99	0.67	1.19	1
SHCH	30	20	2	5	30	15	0	1	0.98	1.4	1
HERM	30	10	2	3	5	5	0	1	0	1	5
PULI	100	20	2	21	70	5	0	1	1	1.41	1
FENL	80	20	3	1	20	20	0	1	0.99	1.4	1
SIHU	30	10	4	4	5	5	0	1	0	1	5
DABA	100	20	1	4	5	40	0	1	0.99	1.4	1
YULI	90	50	1	1	5	60	20	0.95	0.99	1.38	1
CHCH	30	10	3	2	5	5	0	1	0	1	5
LIOQ	90	30	1	17	40	30	60	1	1	1.41	1
RUEY	55	30	1	7	30	40	20	1	1	1.41	1
KAOH	30	10	1	3	5	5	0	1	0	1	5
WANL	90	30	1	4	10	60	10	1	1	1.41	1
FENG	90	30	1	3	5	60	30	1	1	1.41	1

(b)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	60	20	2	9	15	15	50	1	1	1.41	1
KUOL	90	20	4	21	80	50	90	0.99	0.98	1.4	1
TOCH	70	300	1	5	60	60	10	0.98	0.99	1.39	1
HUAL	70	20	2	7	90	60	20	0.99	0.99	1.4	1
ENAN	50	20	2	9	40	60	90	0.99	0.99	1.4	1
DAHU	100	20	1	7	50	15	10	1	1	1.41	1
LISH	60	20	1	15	50	20	60	1	1	1.41	1
SHCH	100	20	1	15	80	40	0	0.98	0.98	1.38	1
HERM	30	10	2	3	5	5	0	1	0	1	5
PULI	55	20	2	11	30	5	30	1	1	1.41	1
FENL	80	20	4	1	20	20	0	1	0.99	1.4	1
SIHU	30	10	4	7	10	5	0	1	0	1	5
DABA	100	20	1	4	5	40	0	1	0.97	1.39	1
YULI	100	30	1	7	20	60	30	1	1	1.41	1
CHCH	30	10	3	2	5	5	0	1	0	1	5
LIOQ	90	30	2	9	30	60	70	1	1	1.41	1
RUEY	55	30	2	7	40	60	0	1	1	1.41	1
KAOH	30	10	1	3	5	5	0	1	0	1	5
WANL	90	30	3	4	10	60	10	1	1	1.41	1
FENG	90	30	1	3	5	60	30	1	1	1.41	1

Table S2. (Continued)

(c)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	60	20	4	7	90	40	0	1	0.98	1.39	1
KUOL	70	100	1	19	30	50	30	1	0.97	1.39	1
TOCH	70	20	1	15	80	60	10	0.92	0.81	1.22	1
HUAL	90	20	1	13	90	60	10	1	1	1.41	1
ENAN	80	30	3	11	70	50	80	1	1	1.41	1
DAHU	100	20	2	7	50	15	10	1	1	1.41	1
LISH	60	20	2	19	70	30	60	1	1	1.41	1
SHCH	30	20	2	5	30	15	0	1	0.98	1.4	1
HERM	30	10	1	4	5	5	0	1	0	1	5
PULI	55	20	4	3	40	30	0	1	0.98	1.4	1
FENL	80	20	3	3	30	15	0	1	0.98	1.4	1
SIHU	30	10	4	13	20	5	0	1	0	1	5
DABA	100	20	3	7	20	15	10	1	1	1.41	1
YULI	100	30	2	3	40	60	60	0.99	0.99	1.4	1
CHCH	30	10	3	2	5	5	0	1	0	1	5
LIOQ	90	30	3	7	30	15	70	0.98	0.64	1.17	1
RUEY	55	20	1	19	40	15	10	1	1	1.41	1
KAOH	30	10	2	4	5	5	0	1	0	1	5
WANL	100	30	2	5	30	50	50	0.99	0.98	1.39	1
FENG	90	30	1	9	70	60	20	0.98	0.97	1.38	1

(d)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	80	300	1	2	5	60	70	0.98	0.98	1.38	1
KUOL	70	100	1	49	100	50	50	1	0.98	1.4	1
TOCH	70	20	1	5	50	60	40	0.86	0.78	1.16	1
HUAL	60	40	2	39	60	60	90	1	1	1.41	1
ENAN	50	20	4	13	90	60	90	1	0.99	1.41	1
DAHU	100	20	1	13	100	15	10	0.98	0.75	1.23	1
LISH	60	20	2	19	70	30	60	1	1	1.41	1
SHCH	80	30	2	5	40	50	0	1	0.98	1.4	1
HERM	30	10	1	4	5	5	0	1	0	1	5
PULI	55	20	2	13	30	5	30	1	1	1.41	1
FENL	80	20	2	5	15	50	0	1	0.99	1.4	1
SIHU	30	10	4	53	80	5	0	1	0	1	5
DABA	100	20	1	29	90	10	0	1	0.95	1.38	1
YULI	40	50	3	13	40	30	80	1	0.95	1.38	1
CHCH	30	10	3	3	5	5	0	1	0	1	5
LIOQ	70	20	3	3	5	30	0	1	0.97	1.39	1
RUEY	100	200	3	7	100	50	50	0.93	0.95	1.33	1
KAOH	30	10	3	4	5	5	0	1	0	1	5
WANL	90	30	2	13	100	50	50	0.94	0.85	1.27	1
FENG	90	30	1	11	90	60	0	0.91	0.88	1.27	1

Table S2. (Continued)

(e)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	60	20	2	5	20	50	30	0.98	0.91	1.34	1
KUOL	100	40	1	41	100	60	90	0.96	0.92	1.32	1
TOCH	70	300	1	7	100	60	70	0.9	0.92	1.29	1
HUAL	90	100	1	7	100	60	10	0.86	0.87	1.22	1
ENAN	30	20	1	25	100	60	60	1	0.99	1.41	1
DAHU	100	20	4	6	10	30	10	0.96	0.57	1.12	1
LISH	60	20	2	19	70	30	60	1	1	1.41	1
SHCH	30	20	2	11	90	30	0	0.99	0.93	1.36	1
HERM	30	10	1	4	5	5	0	1	0	1	5
PULI	55	20	2	35	80	50	10	1	0.98	1.4	1
FENL	80	20	3	5	100	15	0	1	0.97	1.39	1
SIHU	30	10	4	61	100	5	0	1	0	1	5
DABA	100	20	4	7	80	30	0	0.97	0.64	1.16	1
YULI	40	50	4	13	50	20	80	0.99	0.84	1.3	1
CHCH	30	10	4	4	5	5	0	1	0	1	5
LIOQ	100	200	1	19	90	50	90	0.92	0.81	1.23	1
RUEY	70	200	3	7	60	30	30	0.9	0.79	1.2	1
KAOH	30	10	3	4	5	5	0	1	0	1	5
WANL	80	20	1	39	70	50	70	1	0.99	1.41	1
FENG	90	100	1	5	30	50	0	0.82	0.79	1.14	1

(f)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	80	300	2	9	90	60	60	0.94	0.89	1.3	1
KUOL	90	20	2	37	100	60	90	0.98	0.84	1.29	1
TOCH	45	100	1	13	100	60	80	0.94	0.88	1.29	1
HUAL	50	100	1	19	80	60	90	0.97	0.79	1.25	1
ENAN	50	20	4	13	90	60	90	1	0.99	1.41	1
DAHU	100	20	4	6	10	50	0	0.95	0.61	1.13	1
LISH	60	20	1	19	40	40	60	1	1	1.41	1
SHCH	30	20	3	3	30	60	20	0.99	0.91	1.34	1
HERM	30	10	3	3	5	5	0	1	0	1	5
PULI	100	20	1	45	70	15	40	0.98	0.73	1.22	1
FENL	80	20	3	3	20	15	0	0.98	0.73	1.23	1
SIHU	30	10	4	67	100	5	0	1	0	1	5
DABA	100	20	4	2	5	60	30	0.81	0.62	1.02	1
YULI	40	50	4	15	50	20	80	0.98	0.63	1.16	1
CHCH	30	10	4	4	5	5	0	1	0	1	5
LIOQ	100	200	2	17	100	40	90	0.9	0.69	1.13	1
RUEY	70	30	2	5	15	30	80	0.92	0.71	1.16	1
KAOH	30	10	3	4	5	5	0	1	0	1	5
WANL	100	100	1	23	100	50	10	0.88	0.87	1.24	1
FENG	90	30	2	5	70	50	30	0.84	0.68	1.08	1

Chen et al.

Table S2. (Continued)

(g)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	80	300	1	5	15	60	70	0.88	0.84	1.22	1
KUOL	90	20	2	37	100	60	90	0.98	0.84	1.29	1
TOCH	30	20	1	13	90	60	80	0.94	0.77	1.21	1
HUAL	60	40	1	15	80	60	80	1	0.98	1.4	1
ENAN	50	20	2	15	60	60	90	1	0.99	1.41	1
DAHU	100	20	4	5	10	50	0	0.95	0.61	1.13	1
LISH	60	20	1	19	40	40	60	1	1	1.41	1
SHCH	30	20	3	3	30	60	20	0.99	0.91	1.34	1
HERM	30	10	2	4	5	5	0	1	0	1	5
PULI	100	20	2	47	90	20	30	0.98	0.72	1.21	1
FENL	80	20	3	3	20	15	0	0.98	0.68	1.2	1
SIHU	30	10	4	67	100	5	0	0.98	0	0.98	5
DABA	100	20	2	21	90	10	0	0.97	0.36	1.04	3
YULI	40	50	4	17	50	30	70	0.98	0.64	1.17	1
CHCH	30	10	4	4	5	5	0	1	0	1	5
LIOQ	100	200	2	17	100	30	90	0.91	0.64	1.11	1
RUEY	70	30	1	59	100	20	20	0.93	0.6	1.11	1
KAOH	30	10	3	4	5	5	0	1	0	1	5
WANL	90	30	1	29	100	60	10	0.88	0.83	1.21	1
FENG	90	10	1	4	5	20	40	0.99	0.67	1.19	1

(h)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	55	100	2	57	100	15	30	0.99	0.65	1.19	1
KUOL	90	20	2	37	100	60	90	0.98	0.84	1.29	1
TOCH	30	20	1	13	90	60	90	0.94	0.76	1.21	1
HUAL	60	40	1	17	90	50	80	1	0.97	1.39	1
ENAN	50	20	2	15	60	60	90	1	0.99	1.41	1
DAHU	100	20	4	7	20	30	0	0.97	0.54	1.11	1
LISH	60	20	1	19	40	40	60	1	1	1.41	1
SHCH	30	20	3	5	70	40	0	1	0.96	1.38	1
HERM	30	10	2	4	5	5	0	1	0	1	5
PULI	100	20	2	57	90	15	30	0.98	0.63	1.16	1
FENL	30	10	4	1	5	30	80	0.98	0.63	1.17	1
SIHU	30	10	4	67	100	5	0	0.98	0	0.98	5
DABA	80	10	3	5	40	60	90	0.92	0.53	1.06	1
YULI	100	30	4	3	60	30	90	0.88	0.68	1.11	1
CHCH	30	10	4	4	5	5	0	1	0	1	5
LIOQ	50	10	1	21	70	60	40	0.93	0.54	1.07	1
RUEY	55	30	2	3	10	60	70	0.85	0.75	1.14	1
KAOH	30	10	3	4	5	5	0	1	0	1	5
WANL	60	10	2	17	60	60	90	1	0.98	1.4	1
FENG	90	10	1	4	5	20	40	0.99	0.67	1.19	1

Table S2. (Continued)

(i)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	55	100	2	55	100	15	30	0.99	0.65	1.19	1
KUOL	90	20	2	37	100	60	90	0.97	0.73	1.22	1
TOCH	30	20	1	13	90	60	90	0.94	0.76	1.21	1
HUAL	60	40	1	17	90	50	80	0.98	0.77	1.25	1
ENAN	30	20	1	27	100	60	60	1	0.99	1.41	1
DAHU	100	40	2	7	15	5	0	0.98	0.32	1.03	3
LISH	45	20	2	29	100	20	50	1	0.86	1.31	1
SHCH	30	20	3	5	50	40	0	0.95	0.59	1.12	1
HERM	30	10	2	4	5	5	0	1	0	1	5
PULI	100	20	3	33	100	15	20	0.97	0.55	1.12	1
FENL	30	10	4	1	5	30	80	0.98	0.63	1.17	1
SIHU	30	10	3	67	100	5	0	0.98	0	0.98	5
DABA	55	20	4	6	10	20	10	1	1	1.41	1
YULI	40	50	1	15	60	60	90	0.9	0.6	1.09	1
CHCH	30	10	4	4	5	5	0	1	0	1	5
LIOQ	50	10	4	5	100	60	90	0.99	0.89	1.33	1
RUEY	55	30	2	3	10	60	70	0.87	0.75	1.15	1
KAOH	90	10	1	9	70	60	90	0.92	0.48	1.04	3
WANL	60	10	1	63	100	60	90	0.97	0.71	1.2	1
FENG	80	20	1	27	90	5	90	1	0.59	1.16	1

(j)

StaN	Rad	Dep	N_{thr}	t_{thr}	t_{obs}	t_{pred}	t_{lag}	C1	F1	R	PSA
SHRL	60	20	4	7	90	20	10	0.99	0.63	1.17	1
KUOL	55	100	1	13	20	40	70	0.98	0.64	1.16	1
TOCH	30	20	1	13	90	60	90	0.94	0.76	1.21	1
HUAL	80	100	1	5	100	60	90	0.84	0.7	1.09	1
ENAN	30	20	1	29	100	60	60	1	0.99	1.41	1
DAHU	100	40	2	7	15	5	0	0.97	0.28	1.01	3
LISH	80	10	3	13	80	15	30	0.99	0.65	1.19	1
SHCH	30	20	3	5	60	20	30	0.98	0.65	1.18	1
HERM	30	10	2	4	5	5	0	1	0	1	5
PULI	55	20	3	35	90	5	30	0.99	0.4	1.07	3
FENL	30	10	4	1	5	15	90	0.98	0.45	1.08	3
SIHU	30	10	4	67	100	5	0	0.98	0	0.98	5
DABA	55	20	4	6	10	20	10	1	1	1.41	1
YULI	40	50	1	21	90	60	90	0.92	0.64	1.12	1
CHCH	30	10	4	4	5	5	0	1	0	1	5
LIOQ	50	10	3	11	100	60	90	0.98	0.77	1.25	1
RUEY	55	30	2	3	10	60	70	0.89	0.75	1.16	1
KAOH	90	10	2	9	70	60	90	0.94	0.52	1.07	1
WANL	60	10	2	25	100	60	90	0.95	0.6	1.13	1
FENG	80	20	1	31	100	15	80	0.99	0.62	1.17	1