How credible are earthquake predictions that are based on TEC variation?

Ryoya Ikuta[1] and Ryoto Oba[1]

[1]Faculty of Science, Shizuoka University

November 24, 2022

Abstract

We conduct numerical experiments to examine two studies that reported preseismic anomalies in the ionospheric total electron content (TEC) and argued for the significance of their respective analyses based on statistical evaluations. The first study is Liu et al. (2018), who statistically studied the relationship between 62 M[?]6 earthquakes in the Chinese interior over an 18-year period and the TEC, which was deduced from the Global Ionospheric Map. The TEC showed anomalies with specific polarities at set times during certain days that preceded the earthquakes. They defined alarms based on this and drew receiver operating characteristic curves, which yielded a significantly better performance (higher area under the curve (AUC) and lower p-value) than random alarms. We conduct this analysis using random synthetic earthquakes. The resulting AUC and p-values are very similar to those for real earthquakes, indicating that the high performance of the Liu et al. (2018) alarm is an artifact. The second study is Le et al. (2011), who classified the TEC time series into anomalous and non-anomalous days based on the TEC perturbation. They found that the anomalous day rate increased as the nucleation time of the earthquakes was approached, especially for larger and shallower earthquakes. We conduct the same analysis using random synthetic earthquakes. The anomalous day rate that is comparable to their result occurs in ~40 % of the 1,000 random trials, thereby suggesting that their result may also be an artifact.

How credible are earthquake predictions that are based on TEC variations?

[1] Ikuta, R. and [2] Oba, R.

Key Points:

- Pre-seismic ionospheric electron anomalies detected based on GNSS and Global Ionospheric Map are tested from a statistical point of view.
- Numerical experiments based on randomly generated earthquakes reproduced the statistical features of the "significant precursor."
- The statistical significance inferred by previous studies may yield false predictions due to arbitrarily defined "precursor" conditions.

Abstract

We conduct numerical experiments to examine two studies that reported preseismic anomalies in the ionospheric total electron content (TEC) and argued for the significance of their respective analyses based on statistical evaluations. The first study is Liu et al. (2018), who statistically studied the relationship between 62 M≥6 earthquakes in the Chinese interior over an 18-year period and the TEC, which was deduced from the Global Ionospheric Map. The TEC showed anomalies with

specific polarities at set times during certain days that preceded the earthquakes. They defined alarms based on this and drew receiver operating characteristic curves, which yielded a significantly better performance (higher area under the curve (AUC) and lower *p*-value) than random alarms. We conduct this analysis using random synthetic earthquakes. The resulting AUC and *p*-values are very similar to those for real earthquakes, indicating that the high performance of the Liu et al. (2018) alarm is an artifact. The second study is Le et al. (2011), who classified the TEC time series into anomalous and non-anomalous days based on the TEC perturbation. They found that the anomalous day rate increased as the nucleation time of the earthquakes was approached, especially for larger and shallower earthquakes. We conduct the same analysis using random synthetic earthquakes. The anomalous day rate that is comparable to their result occurs in ~40 % of the 1,000 random trials, thereby suggesting that their result may also be an artifact.

1. Introduction

Numerous studies have reported that pre-earthquake processes can induce ionospheric fluctuations. Heki (2011) reported an increase in the ionospheric total electron content (TEC) above the epicenter of the 2011 Tohoku-Oki mainshock ~40 minutes before the earthquake based on the phase delays in the global navigation satellite system (GNSS) signals. GNSS TEC observations have since been employed to characterize the three-dimensional distribution of the increase in ionospheric electron density preceding the 2010 Chilean earthquakes (He & Heki, 2016) and the 2015 Illapel earthquake (He & Heki, 2018), as well as the TEC increase preceding global M7-8 earthquakes (He & Heki, 2017). However, it has been suggested that the observed variations in Heki (2011) may be due to space weather (Utada & Shimizu, 2014), such that the method adopted in He & Heki (2016, 2017, 2018) may have produced artifacts that were influenced by post-earthquake ionospheric disturbances (Kamogawa & Kakinami, 2013; Masci et al., 2015; Eisenbeis&Occhipinti, 2021). A number of recent studies have claimed to overcome these artifacts by not including post-earthquake data in their respective analyses (Heki & Enomoto, 2013, 2015; Iwata & Umeno, 2017, 2018; Goto et al., 2019). However, the statistical significance of these papers has also been questioned and it has been suggested that the observed variability may be due to space weather (Ikuta et al., 2020; Tozzi et al., 2020). Furthermore, Ikuta et al. (2021) have pointed out that the precursor criteria in Iwata & Umeno (2017, 2018) and Goto et al. (2019) contradict each other, and that the reported precursors are not statistically significant.

Liu et al. (2000) and subsequent studies from this research group have also reported ionospheric TEC anomalies prior to a number of earthquakes across East Asia. Ionospheric TEC anomalies preceding the 1999 Chi-Chi earthquake in Taiwan (Liu et al., 2001), the 2004 Sumatra earthquake (Liu et al., 2010a), and the 2010 Heiti earthquake (Liu et al., 2011) have been detected. Statistical studies of the relationship between the long-term electron content time series and earthquakes include the relationship between a six-year time series of M ≥ 6 earthquakes in Taiwan and the corresponding F2-layer critical frequency (foF2) that was measured by an ionosonde (Liu et al., 2000), the relationship between a two-year M≥5 earthquakes in Taiwan and GNSS TEC/foF2 (Liu et al., 2004), and the relationship between two-year time series of M ≥ 5 earthquakes in Taiwan and the GNSS-TEC (Liu et al., 2010b). The above mentioned studies reported that the earthquakes were associated with a decrease in the critical frequency (decrease in electron density), a decrease in the TEC values, and a north–south shift in the TEC peak, respectively. The Global Ionospheric Map (GIM) that is provided by the Center for Orbit Determination in Europe (CODE), instead of ionosonde and GNSS phase data, has been used to study the relationship between TEC and earthquakes that have occurred outside Taiwan, with targeted analyses around Japan (Liu et al., 2013a) and China (Liu et al., 2009, 2013b; Chen et al., 2015). However, the polarity of the increase or decrease in ionospheric TEC and the preceding time of the precursor differ according to the locations and earthquake magnitudes. These ionospheric TEC trends do not appear to follow a consistent law for the same physical phenomena. Furthermore, the statistical significance of the results has not been sufficiently evaluated.

Liu et al. (2018) introduced receiver operating characteristic (ROC) curves to assess the earthquake prediction performance of the TEC variations in a statistical sense. They calculated ROC curves for various precursor thresholds of the TEC variations for 62 M ≥ 6.0 earthquakes across

China over an 18-year period. The ROC curves for the real earthquakes, which differ significantly from those based on random synthetic earthquakes, suggested a good prediction performance of their criteria. Liu et al. (2018) dealt only with earthquakes in China; however, Le et al. (2011) extracted the common properties of the TEC variations for global earthquakes via a comparison of the GIM TEC for 736 M ≥ 6.0 earthquakes that occurred during the 2002–2010 period. Le et al. (2011) statistically showed that the M ≥ 7 earthquakes with hypocenters above 20 km depth are associated with an increase/decrease in TEC (regardless of polarity) above/below a given threshold that peaks within one day of earthquake nucleation.

Here we conduct a follow-up test on the Liu et al. (2018) assessment to test the statistical significance of the "precursors." We then conduct a follow up test on the Le et al. (2011) assessment to investigate whether the TEC fluctuations identified by Le et al. (2011) in the days preceding the analyzed earthquakes are also statistically significant.

2. *Z*-test and ROC curve assessments

We tested the Liu et al. (2018) results to assess the validity of the GIM TEC fluctuations being concentrated before the analyzed earthquakes. Specifically, Liu et al. (2018) searched for the common preceding day and time of day when a TEC increase/decrease occurred for multiple earthquakes and defined the preceding day, time of day, and polarity of the TEC increase/decrease as the earthquake precursors, which were determined significant via a *z*-test. They then reanalyzed the TEC records; if an anomaly with the defined time of day and polarity was detected, then it was declared positive for an earthquake after the defined preceding day. They finally constructed an ROC curve using the true positive rate (TPR) and false positive rate (FPR) of the alarm with varying anomaly thresholds and demonstrated that the alarm was significant based on the large area under the curve (AUC). We first replicated the Liu et al. (2018) procedure, then investigated their significance validation using *z*-tests and ROC curves via a series of numerical experiments that assumed a random occurrence of earthquakes that were independent of the TEC variations.

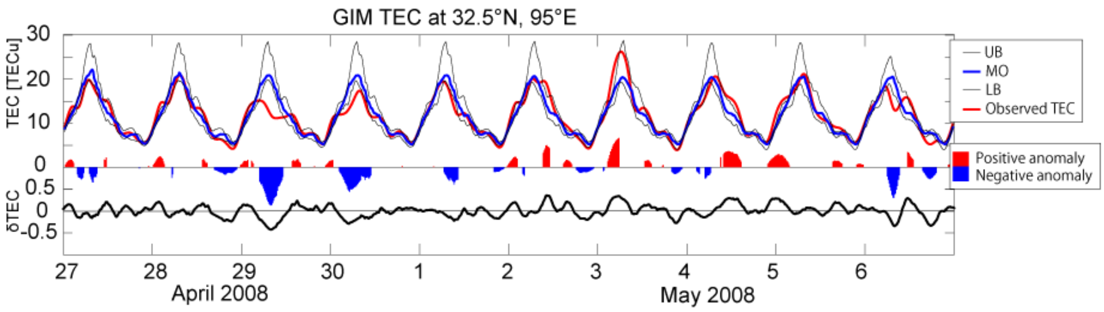

Figure 1. Reproduction of part of fig. 3 in Liu et al. (2018). Upper: TEC time series from 27 April to 7 May 2008 that was deduced from GIM. The red line is the observed TEC at 32.5°N, 95°E. Blue and black lines show the median (MO) and upper/lower bounds (UB/LB). Red and blue shaded areas show the positive and negative anomalies, respectively. Bottom: δTEC time series, which is the difference between the observed TEC and MO normalized by MO.

2-1. Reproducing the Liu et al. (2018) results

We first reproduced the Liu et al. (2018) results using the same data. Liu et al. (2018) obtained an interpolated TEC time series (15-min interval) at 32.5°N, 95°E from the 2-h GIM time series (28 March 1998–31 December 2015 period), with the TEC anomalies for all dates and times of interest defined on the basis of TEC values at the same time of day during the preceding 15 days. Liu et al. (2018) defined the upper boundary (UB) and lower boundary (LB) based on the lower quartile (LQ), upper quartile (UQ), and median (MO) during the same time of day over the 15-day period as follows, and defined a TEC anomaly as a value above/below the UB/LB. UB and LB are defined as $\mathrm{UB} = \mathrm{MO} + k(\mathrm{UQ} - \mathrm{MO})$ and $\mathrm{LB} = \mathrm{MO} - k(\mathrm{MO} - \mathrm{LQ})$, respectively, where k is a threshold coefficient, which controls the frequency of the anomalies. Figure 1 shows the time series of the TEC anomalies when $k = 1.5$, following Liu et al. (2018). They also defined $\delta\mathrm{TEC} = (\mathrm{TEC} - \mathrm{MO})/\mathrm{MO}$, which is the normalized deviation of the TEC from the median, arranged it into a time series with time 0 denoting the earthquake nucleation time, and calculated the median for each earthquake group ($6.0 \leq M < 6.5$, $6.5 \leq M < 7.0$, and $M \geq 7.0$). The median of the 15-day-long δTEC is shown in Figure 2, which is identical to the Liu et al. (2018) result, thereby demonstrating that we reproduced their calculation exactly.

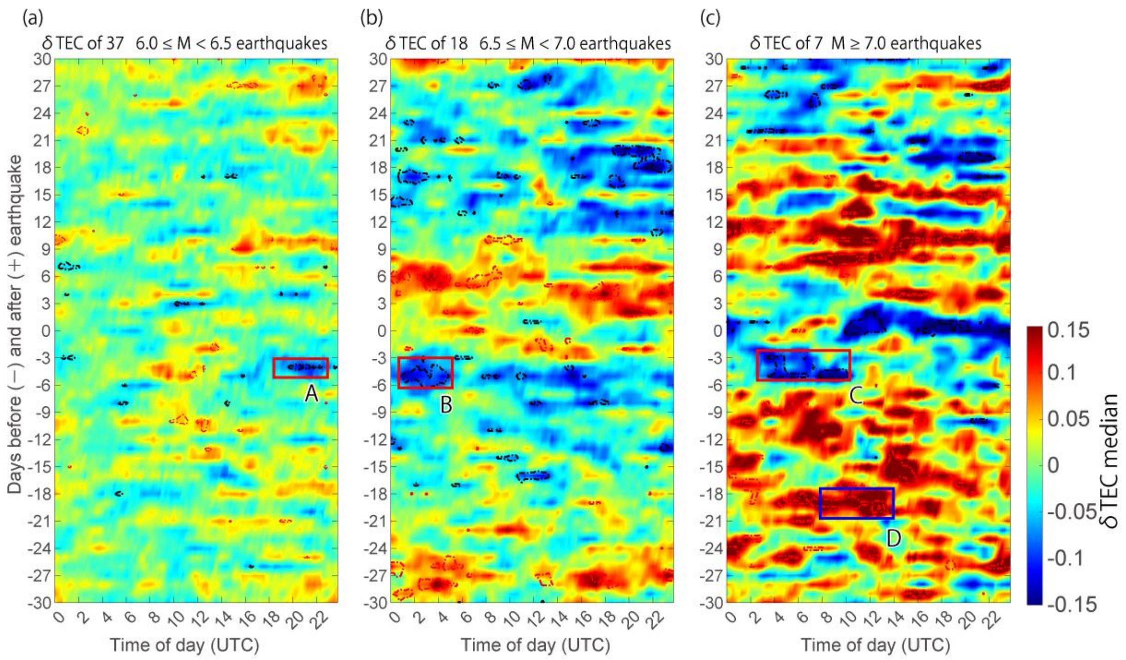

Figure 2. Reproduction of fig. 5 in Liu et al. (2018) using the same data. Median δTEC values at 32.5°N, 95°E over a 60-day period: 30 days before and after (a) 37 6.0 ≤ M < 6.5, (b) 18 6.5 ≤ M < 7.0, and (c) 7 M ≥ 7.0 earthquakes. Contours denote significant *z*-test results at a significance level of 0.05. Zones A, B, and C are the TEC values with negative polarity anomalies during the 18:00–22:00 UTC timeframe and 4–5 days before the group (a) earthquakes; during the 01:00–04:00 UTC timeframe and 3–6 days before the group (b) earthquakes; and during the 04:00–10:00 UTC timeframe and 3–5 days before the group (c) earthquakes. Zone D is the TEC with positive polarity anomalies during the 08:00–12:00 UTC timeframe and 18–20 days before the group (c) earthquakes.

Liu et al. (2018) evaluated the statistical significance of the earthquake-related TEC anomalies for the obtained TEC time series by applying the *z*-test using the following equation:

$$z = \frac{\pi - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}, \quad (1)$$

where π is the observed fraction of the earthquake-related anomalies with k = 1.5 in a particular time of day and preceding day with respect to the analyzed earthquakes; π_0 is the background fraction of the anomalies observed at the same time of day over an 18-year period (6473 days); and n is the number of earthquakes. The positive and negative polarities are evaluated separately. The z = 1.96 threshold corresponds to a significance level of 0.05, whereby a z-value above 1.96 means that the observed fraction of the anomalies will realize by chance on less than 5% of the day based on the time of day matrix. The contours for z = 1.96 are shown in Figure 2. They defined the time of day and preceding date where the z-values that are larger than 1.96 are clustered (zones A, B, C, and D in Figure 2) as earthquake precursors. They issued an earthquake alarm when more than a third of the defined time range was occupied by TEC anomalies, which was based on this definition of precursors by the time of day and polarity. For example, if one third of the 18:00–22:00 UTC time range, which corresponds to zone A, is occupied by negative anomalies, then the period 4–5 days after these anomalies are considered earthquake alarm days. If an earthquake of the specified magnitude range occurs within the alarm days, then the alarm is considered true positive; otherwise, the alarm is considered false positive. They also conducted ROC tests by varying the threshold coefficient k from 0 to 10. The tests evaluated the balance between the TPR, where TPR = (true positive days)/(all earthquakes), and FPR, where FPR = (false positive days)/(all non-alarmed or false positive days).

We applied the same method and obtained results that were almost identical to those in Liu et al. (2018) (Figure 3). The gray lines in the figure show the results of 1,000 simulations that employ the same criteria but with random earthquake nucleation times. There is an upward shift in the ROC curve for the alarms for real earthquakes, which indicates a higher TPR (= sensitivity) and/or lower FPR (= 1 − specificity) than the alarms for random earthquakes. The significance of the alarm is expressed as the AUC, which is the ratio of the area under the ROC curve to the total area. An AUC of 0.5 is expected for the alarms generated from random earthquakes, whereas the AUC is 1 for an ideal alarm system. Note that the ROC curve in Figure 3, which is drawn in the same way as in Liu et al. (2018), does not equal 0.5 for the random prediction case. The deviation from 0.5 arises because the graph is staircase shaped and the corners of the staircase are added above the true value, resulting in a slight overestimation (~0.05) of the true AUC.

Liu et al. (2018) defined the *p*-value as the fraction of the AUC for the random cases that exceeds the AUC for the real earthquake case. Although their *p*-values were 0% for all of the cases, our calculations yield *p*-values of 2.4%, 1.8%, 5.0%, and 0.4% for zones A, B, C, and D, respectively. The differences between our results and those in Liu et al. (2018) are shown in Table 1. Our random simulation generates larger maximum AUC values, ranging from 0.10 (zone A) to 0.27 (zone D), than Liu et al. (2018) for all of the criteria zones. Although we cannot speculate on the reason for this difference, our definition of the AUC is consistent between the actual and simulated earthquakes, as shown above, so we adopt these values as our criteria for the alarm.

Table 1.

Zone	Polarity	Appearance (UTC)	Alarm day (–)	AUC	$AUC_{max}1^*$	$AUC_{max}2^*$ (Liu)	p-value [%]	Youden index		
								k value	TPR	FPR
A	negative	18:00–22:00	4 – 5	0.6196	0.6617	0.5590	2.4	1.8	0.5676	0.3578
B	negative	01:00–04:00	3 – 6	0.6494	0.7182	0.5581	1.8	2.4	0.7222	0.3527
C	negative	04:00–10:00	3 – 5	0.6919	0.8547	0.5978	5	2.1	0.7143	0.3476
D	positive	08:00–12:00	18 – 20	0.8080	0.8363	0.5624	0.4	2.1	1.0000	0.4209

Note: $AUC_{max}1^*$ and $AUC_{max}2^*$ denote the maximum AUC values of the 1,000 random simulations that were conducted in this study and Liu et al. (2018), respectively.

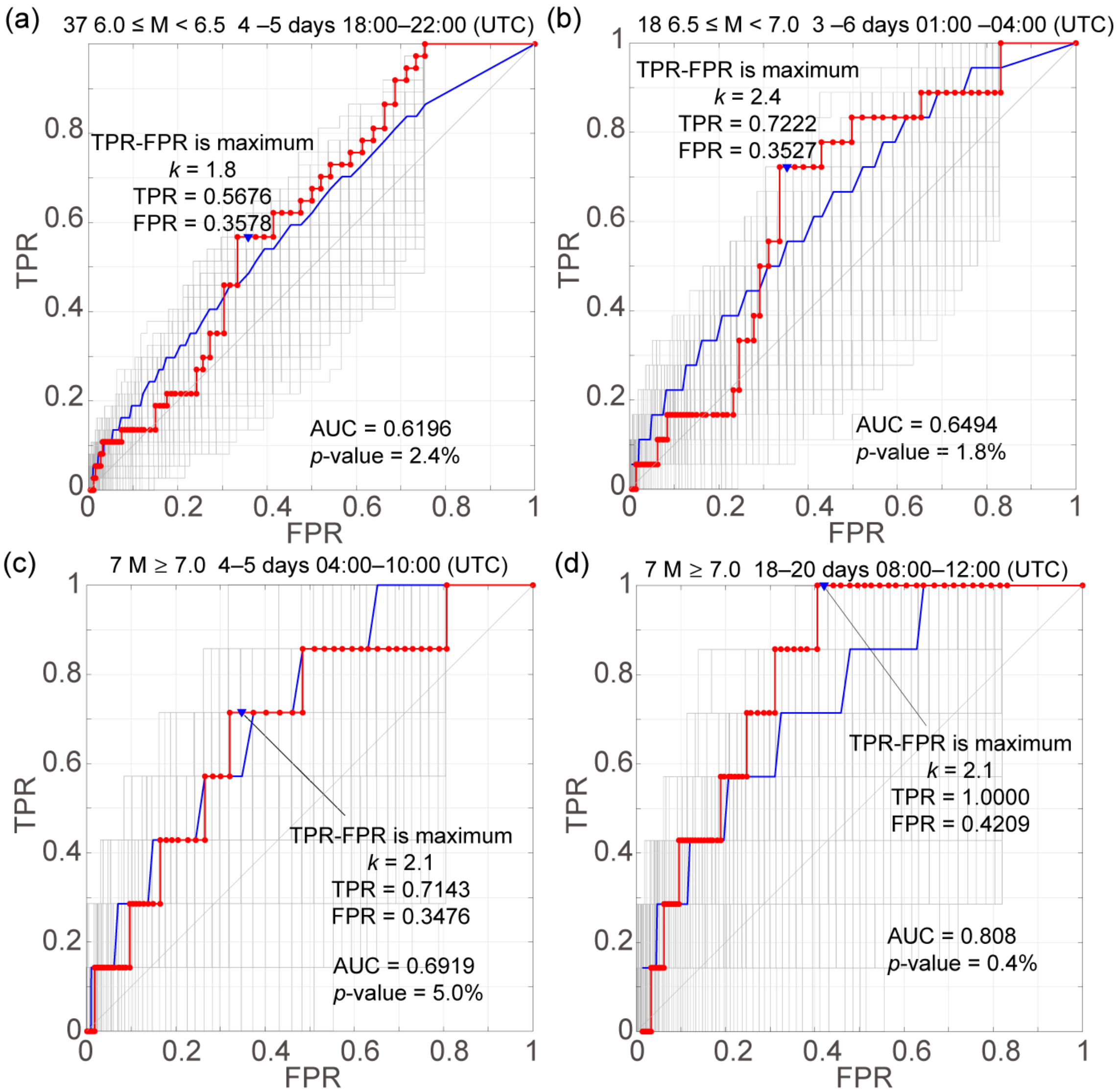

Figure 3. Reproduction of Fig. 7 in Liu et al. (2018). The ROC curves are deduced from the same data set used in Liu et al. (2018), with the same vertical TEC and earthquakes analyzed. Red, gray, and blue curves are the ROC curves for the real observations and 1,000 random simulations, and the 95% line for the simulations, respectively. The blue triangle in each plot denotes the point that yields the maximum R score (= TPR − FPR) for $1 \leq k \leq 10$.

2-2. Liu et al. (2018) analysis with random synthetic earthquakes

Liu et al. (2018) evaluated the significance of the relationship between TEC anomalies and earthquakes via a *z*-test, defined the alarm criteria based on this significance, and then issued alarms and evaluated the significance of these alarms using ROC curves. However, if the criteria, which are based on actual earthquakes, are re-applied to predict the occurrence of the same earthquakes, it would therefore seem natural that the predictive ability of the Liu et al. (2018) analysis would be high, even if the TEC anomalies are not related to earthquakes. Here we apply the same method to randomly generated earthquakes that are independent of the TEC, starting from the construction of the criteria for the alarm, to determine reliability of the Liu et al. (2018) approach. The synthesized earthquakes are of the same magnitude as real ones; only the nucleation times are randomly varied. Figure 4 shows the median δTEC for each randomly generated earthquake group ($6.0 \leq M < 6.5$, $6.5 \leq M < 7.0$, $M \geq 7.0$). The contours for $z = 1.96$ are also shown, which is very similar to the real earthquake scenario in Figure 2, including the spatial scale of the texture. Similar to Liu et al. (2018), we define the TEC variations for the range of dates and times when the *z*-value exceeds 1.96 consecutively (regions A, B, and C in Figure 4) as earthquake precursors.

The resultant date and time of day values indicate the following. If positive anomalies occupy more than one-third of the time of day range corresponding to zone A (2:00–6:00 UTC), then the alarm occurs 2–4 days later for $6.0 \leq M < 6.5$ earthquakes. Similarly, the alarm occurs 10–12 days later for positive anomalies within the 19:00–24:00 UTC timeframe and 6.5 ≤ M < 7.0 earthquakes, and 24–25 days later for negative anomalies within the 4:00–9:00 UTC timeframe and M ≥ 7.0 earthquakes.

ROC curves are drawn based on whether the predicted earthquakes occurred within the alarm timeframes (Figure 5). The ROC curve (red curve) is shifted upward relative to the ROC curves drawn for a set of 1,000 randomly generated earthquakes that did not depend on the definition of the alarm, with AUCs of 0.603, 0.662, and 0.684 and with *p*-values of 3.6%, 1.5%, and 6.4% obtained for alarm zones A, B, and C, respectively. These values are comparable to those calculated for the actual earthquakes, despite the fact that this experience has been carried out on random synthetic earthquakes.

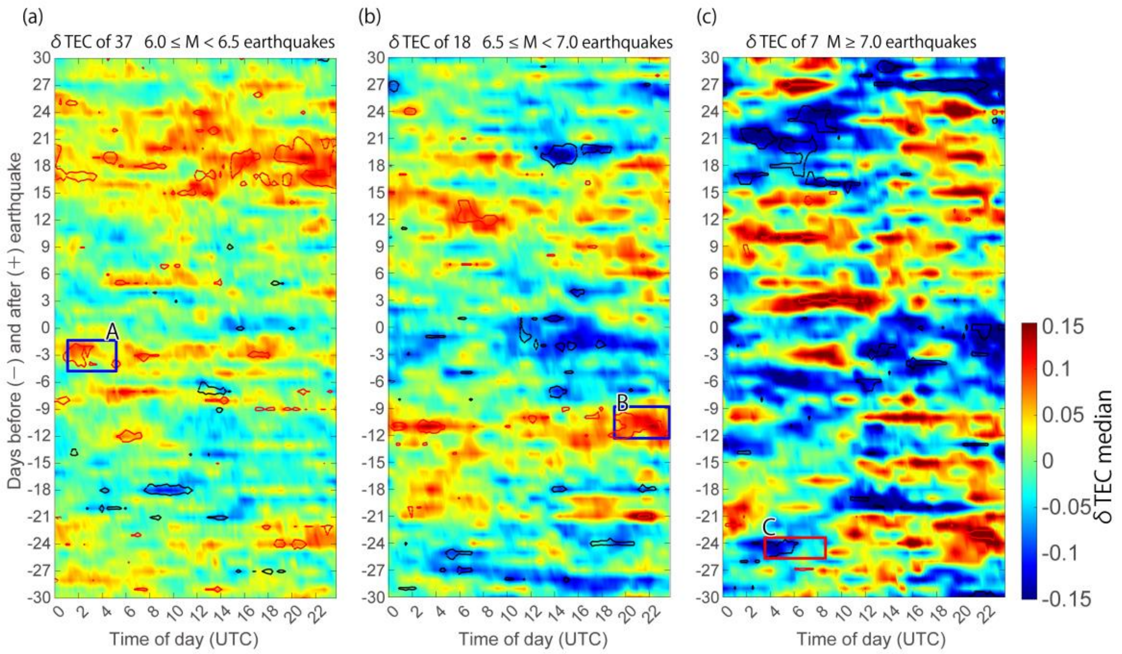

Figure 4. Median δTEC values that were calculated at 32.5°N, 95°E for the 60-day period surrounding each randomly synthesized earthquake. The median values are calculated for (a) 37

earthquakes assuming 6.0 ≤ M < 6.5, (b) 18 earthquakes assuming 6.5 ≤ M < 7.0, and (c) 7 earthquakes assuming M ≥ 7.0. Zones A and B are positive TEC anomalies that occurred during the 02:00–06:00 UTC timeframe and 2–4 days before the group (a) earthquakes, and during the 19:00–24:00 UTC timeframe and 10–12 days before the group (b) earthquakes, respectively. Zone C is a negative TEC anomaly that occurred during the 04:00–09:00 UTC timeframe and 24–25 days before the group (c) earthquakes.

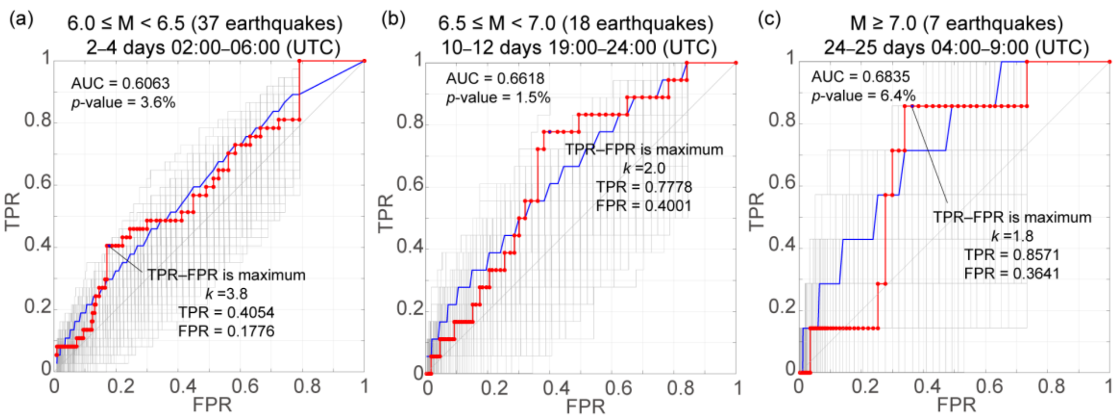

Figure 5. Same with Figure 3, except that the ROC curves are calculated for random synthetic earthquakes instead of actual earthquakes; the ROC curves are based on the precursor criteria defined by zones A, B, and C in Figure 4.

3. Testing the TEC anomaly concentration a few days before earthquakes

We next test the validity of the Le et al. (2011) assessment. We first review the Le et al. (2011) methodology, and then test their method via numerical experiments, which employ random synthetic earthquakes that are unrelated to the TEC variations.

3.1. Anomalous day rate in Le et al. (2011)

Le et al. (2011) compared 736 global M ≥ 6.0 earthquakes that occurred during the 2002–2010 period with the TEC time series that was deduced from the GIM directly above the epicenters via the following procedure. The 2-h sampling TEC time series was interpolated to a grid (2.5° × 5°; latitude × longitude) at 1-h sampling interval. Anomalies can be defined for any given date and grid point based on the distribution of the TEC values at the same time of day over the preceding 15 days. Specifically, UB and LB are defined as UB = $m + \sigma$ and LB = $m - \sigma$, respectively, with these two boundaries based on the median m and standard deviation σ of the same 15-day period. Le et al. (2011) defined an anomaly as a TEC value either above UB or below LB. If anomaly occurs for more than six consecutive samples (hours) on a given day and the largest deviation from the median is larger than (1 + R) σ in which R is a deviation level such as 60%, 80%, or 100%, this day would be considered as an anomalous day with the level R. Note that a TEC anomaly is less likely to occur as R increases.

Le et al. (2011) then calculated the anomalous day rates (anomalous days/total period) over various total periods that range from 1 to 21 days before the earthquake at the grid point closest to the epicenter of a given earthquake. The average of the anomalous day rates for all of the earthquakes of interest $P_E(T,R)$ is calculated as:

$$P_E(T,R) = \frac{1}{K}\sum_{n=1}^{K}\frac{N_{R,T}^{n}}{T-\Delta S} \times 100\%, \qquad (2)$$

where $N_{R,T}^{n}$ is the anomalous day number within T days before n-th earthquake when the deviation level is set to R; and K is the number of earthquakes that have occurred at the size and depth range of interest. Note that a given day with large magnetic storm disturbances and the following couple of days are removed from the calculation based on the D_{st} index. We quote the calculated $P_E s$ values from Le et al. (2011) for each three different R values (60%, 80%, and 100%) and depths (shallower than 20, 30, and 40 km depth; Figure 2 from Le et al. (2011)) in Figure 6. The horizontal axis is T and the vertical axis is the lower limit of the interested earthquake magnitude range. For example, the grid point at M = 6.0 and T = 10 represents the average of the anomalous day rates during 10 days before all M ≥ 6.0 earthquakes. Higher anomalous day rates are observed for the larger and shallower earthquake groups (Le et al., 2011). Higher anomalous day rates are also observed as the earthquake days are approached.

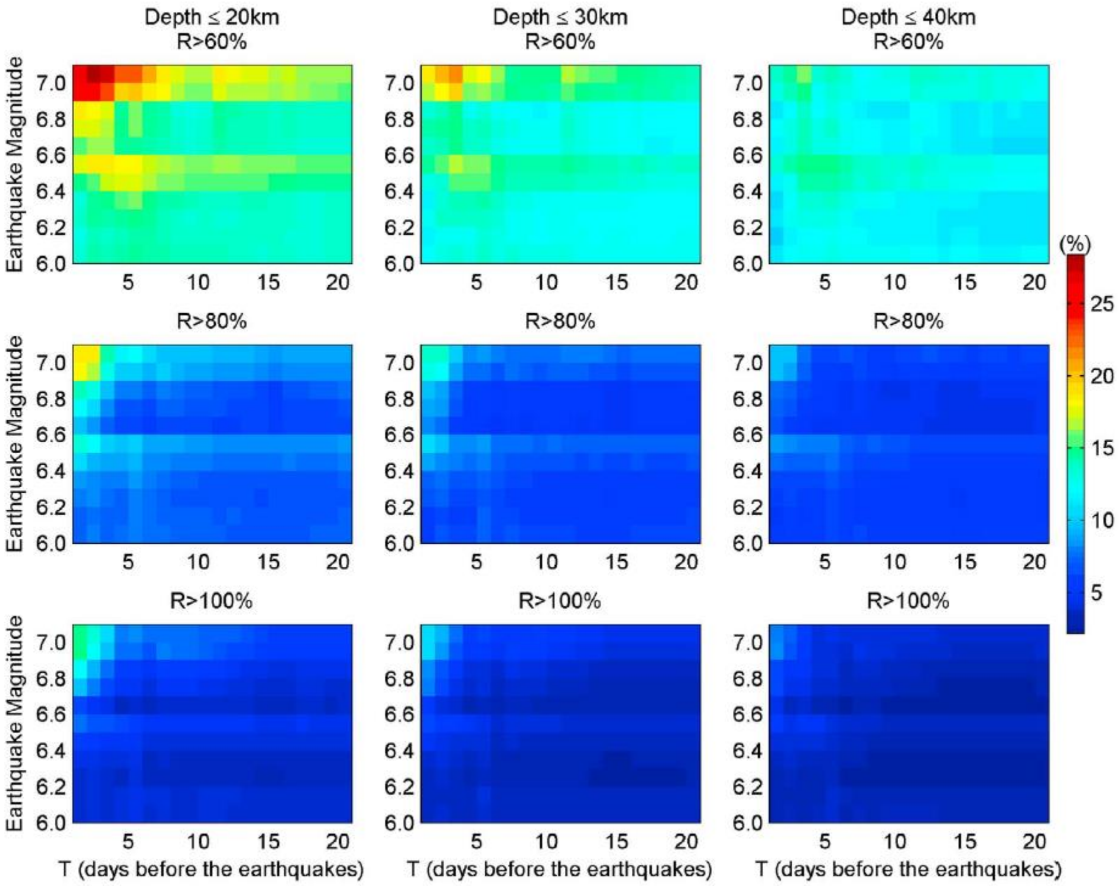

Figure 6. Adapted from Le et al. (2011). P_E for real earthquakes. The anomalous day rates when R > 60%, 80%, and 100% within T days before the earthquakes (P_E) that occur at ≤ 20, ≤ 30, and ≤ 40 km depth are shown.

3.2. Le et al. (2011) with random synthetic earthquakes

Le et al. (2011) defined the anomalies from the TEC time series at the grid point closest to the epicenter of each earthquake and calculated the anomalous day rate P_E. They then assessed the significance of P_E by estimating the background anomalous day rate P_N of the TEC time series over a 240-day period that ranged from 60 to 300 days before each earthquake for comparison with

$P_E s$:

$$P_N(R) = \frac{1}{K \times 240 - \Delta W} \sum_{n=1}^{K} N_R^n \times 100\%, \qquad (3)$$

where N is the number of anomalous days, excluding the magnetic disturbance days, as in Equation 2; and ΔW is the total number of excluded days. However, $P_N(R)$ is a function that is independent of T and the denominator is constant at 240, whereas the denominator in the $P_E(R,T)$ calculation for each earthquake varies between 1 and 20 and is dependent on T. It is expected that the small denominator in $P_E(R,T)$ magnifies the stochastic fluctuations of N_R^n and makes the result unstable.

We also apply the same method to randomly generated earthquakes that are independent of the TEC to investigate whether the results of Le et al. (2011) also appear when there is no correlation between the TEC anomalies and earthquakes. We ignore the spatial distribution of the TEC anomalies and earthquakes, for simplicity, and focus only on the number of earthquakes and the TEC anomalous day rate in our simulation. We first assume the background average anomalous day rates for the $R \geq 60\%$, $R \geq 80\%$, and $R \geq 100\%$ deviation levels as 15%, 8%, and 4%, respectively, based on Figure 3 in Le et al. (2011), and apply these anomalous day rates to the 8-year TEC time series. We then generate earthquakes for each depth and magnitude group after synthesizing the time series of the anomalous days. We generate all 736 M ≥ 6 earthquakes with hypocenters at ≤ 40 km depth at random times based on the earthquake magnitude and depth distribution in table 1 of Le et al. (2011). Then we select 602 earthquakes as of hypocenters at ≤ 30 km depth, of which 490 are at ≤ 20 km depth. In terms of size, out of the total 736 events, we randomly select 573 events as M ≥ 6.1, of which 454 as M ≥ 6.2, of which 362 as M ≥ 6.3, of which 273 as M ≥ 6.4, of which 221 as M ≥ 6.5, of which 176 as M ≥ 6.6, of which 130 as M ≥ 6.7, of which 104 as M ≥ 6.8, of which 79 as M ≥ 6.9, of which 66 as M ≥ 7.0, and of which 53 as M ≥ 7.1. This approach effectively reproduces the earthquakes used in Le et al. (2011). Figure 7 shows an example where P_E is calculated using the same method as in Le et al. (2011), but from our random earthquake sequence. The results are very similar to those in Figure 6, even though the earthquakes are randomly generated and independent of the TEC.

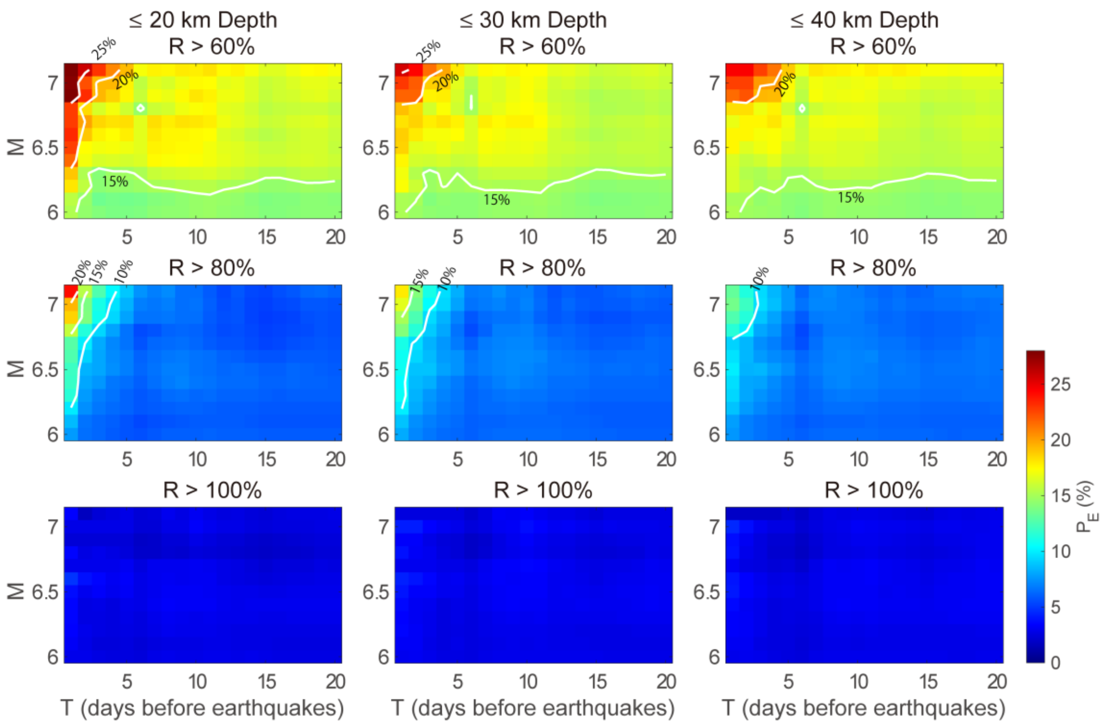

Figure 7. Calculated P_E from both the synthetic TEC anomaly days and earthquake catalog. The occurrence rate of the TEC anomaly days and the number for each earthquake magnitude range are

the same as those used in Le et al. (2011), with the contours plotted at constant 5% intervals.

4. Discussion

We find that the same results in Le et al. (2011) and Liu et al. (2018) can be obtained for randomly generated earthquakes that are independent of the TEC. We discuss the probability that the Le et al. (2011) and Liu et al. (2018) results occurred by chance, even for earthquakes that exhibited no relationship to the TEC.

Liu et al. (2018) first used the *z*-test to assess whether the TEC anomalies that occurred before the earthquakes were significantly related to the earthquakes. However, their choice of $z = 1.96$ is a statistical criterion that is inevitably exceeded by 5% of the samples by chance. It is not possible to say that exceeding this value has anything to do with the earthquake. Therefore, the probability that a "precursor" is defined for a specific number of preceding days and time of day for each selected earthquake group is almost 100%.

We then repeat the random earthquake prediction test performed in section 2-2 100 times to test the probability of the large AUCs and small *p*-values in Liu et al. (2018). The "precursor" day and time ranges for each simulation are defined by choosing the largest area (day × hour) that is enclosed by the $z = 1.96$ contour in the *z*-map, as shown in Figure 4. A rectangular zone that includes the "precursor" area is extracted. Figure 8 shows the distribution of the *p*-values and ΔAUCs that are obtained by the repeated simulations. ΔAUC is the difference between the obtained AUC from that expected for random earthquakes, which is estimated as the median of the 100 random experiments for each simulation. The obtained ΔAUCs and *p*-values are similar to those obtained for the real earthquakes. The ΔAUCs for real earthquakes were 0.082, 0.132, 0.164, and 0.282 for alarm zones A (37 earthquakes), B (18 earthquakes), C (7 earthquakes), and D (7 earthquakes), respectively whereas the median ΔAUCs for 100 simulation experiments are 0.082, 0.114, and 0.175 for alarm zones for 37, 18, and 7 random earthquakes, respectively. These values for real earthquakes are very similar to the median of random earthquakes except for zone D, which assigns the same earthquake group in zone C. The *p*-values were 2.4%, 1.8%, 5.0%, and 0.4% of the real earthquakes in alarm zones A, B, C, and D, respectively. These *p*-values, although very small, are larger than the 96%, 98%, 99%, and 69% values for all of the simulations, respectively. These results suggest that the *p*-values that were obtained for the real earthquake groups can be obtained with a probability of almost 100%, even if the TECs and earthquakes are not correlated with each other. The large ΔAUCs and small *p*-values for real earthquakes in Liu et al. (2018) do not support the claim that TEC variations precede earthquakes. Their results could therefore be typical artifacts that are caused by the misapplication of the employed statistical tests.

We also assess the probability of the Le et al. (2011) results. We repeated the simulations in Section 3-2 by creating 1,000 different random combinations of the anomalous days and earthquakes. There are many individual P_E distributions, which differ from those in Figure 7. Therefore, we plot the top 150th value of the P_E distribution in the 1,000 simulations in Figure 9. The P_E distribution pattern is still very similar to the Le et al. (2011) results, with this distribution showing that even when there is no correlation between the actual spatiotemporal distributions of the TECs and earthquakes, the probability of obtaining a result similar to that in Le et al. (2011) is about 15%. A probability of 15% may seem small, such that one could argue that the Le et al. (2011) results occurred by chance. However, we obtained 396 trials out of 1,000 simulations where the maximum P_E is greater than 20 for all magnitude and depth ranges (Figure 10a). Therefore, the probability that an artifact will cause the TEC to appear to fluctuate abnormally within a specific preceding time for a given earthquake of a specific size and depth range is about 40%. The larger anomalous day rates for the larger and shallower earthquakes in the Le et al. (2011) results should be due to the same principle as the larger anomalous day rates for shorter preceding time T, whereby the fluctuations increase as the denominators decrease. In fact, the calculated standard deviations of P_E for each T and M range for the 1,000 simulations (Figure 10) indicate that the P_E perturbation increases as the number of included earthquakes decreases (Figure 10b).

The correlations between the TEC variations and earthquakes shown in Liu et al. (2018) and Le et al. (2011) are considered methodological artifacts of their analyses, with probabilities of almost 100% and 40%, respectively. Therefore, the statistical analyses in Liu et al. (2018) and Le et al.

(2011) do not provide a valid reason for concluding that the observed TEC variations are influenced by pre-earthquake processes.

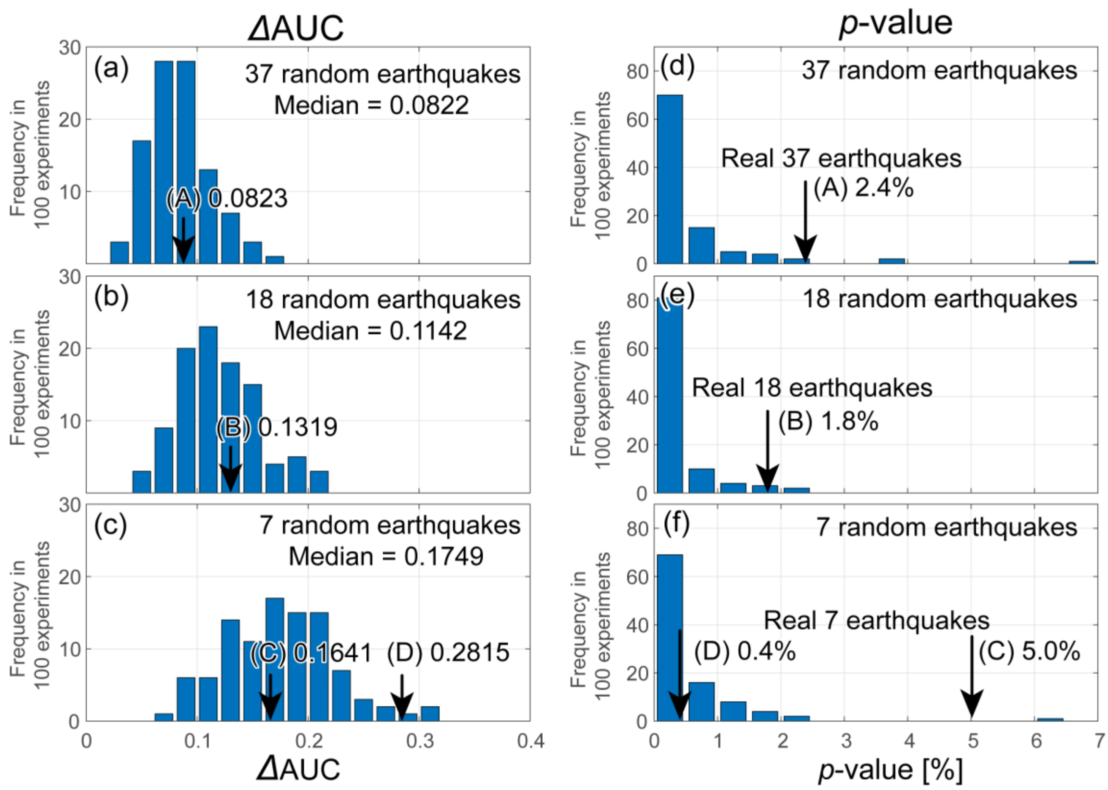

Figure 8. Histograms of ΔAUCs and *p*-values that were obtained by 100 simulation experiments. ΔAUCs for (a) 37, (b) 18, and (c) 7 random earthquake groups. The bin width is 0.02 ΔAUC. *p*-values for (d) 37, (e) 18, and (f) 7 random earthquake groups. Arrows show the values for the real earthquake groups that were predicted based on zones A, B, C, and D, which are shown in Figures 2 and 3.

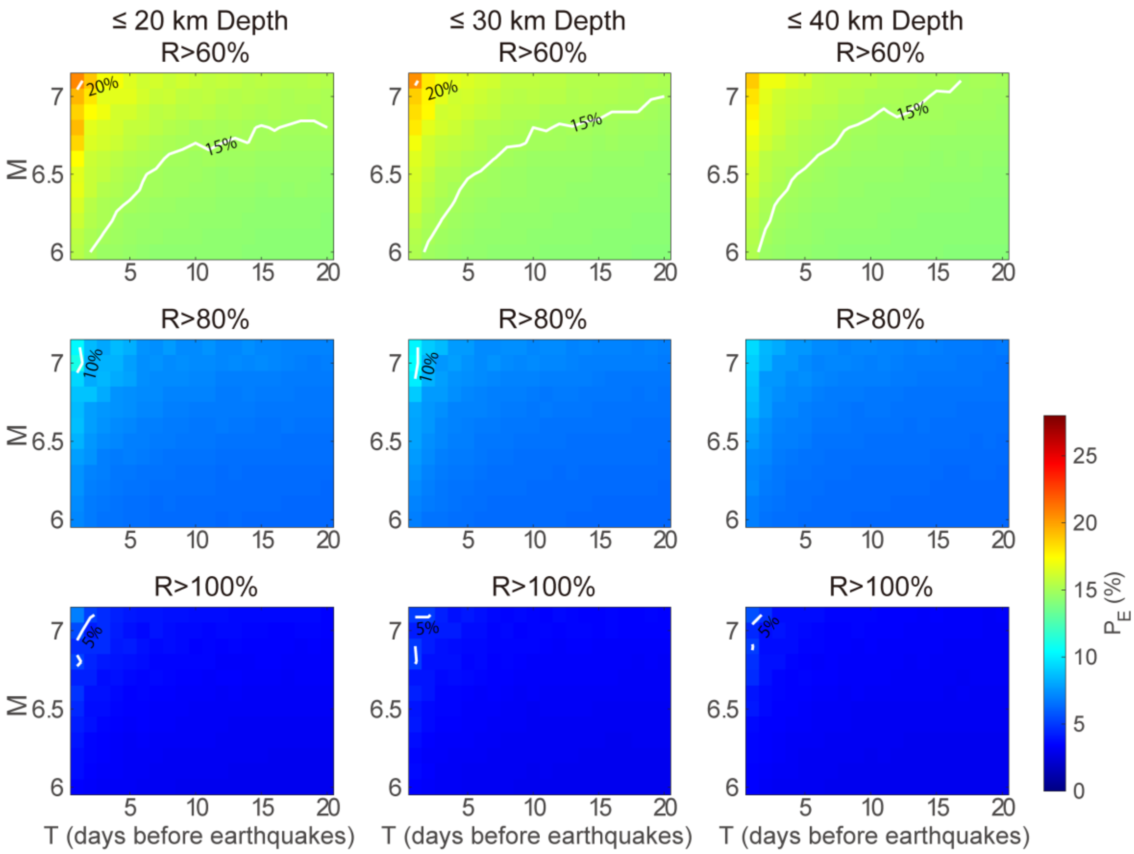

Figure 9. The top 15% $P_E(R,T)$ for the 1,000 numerical simulations.

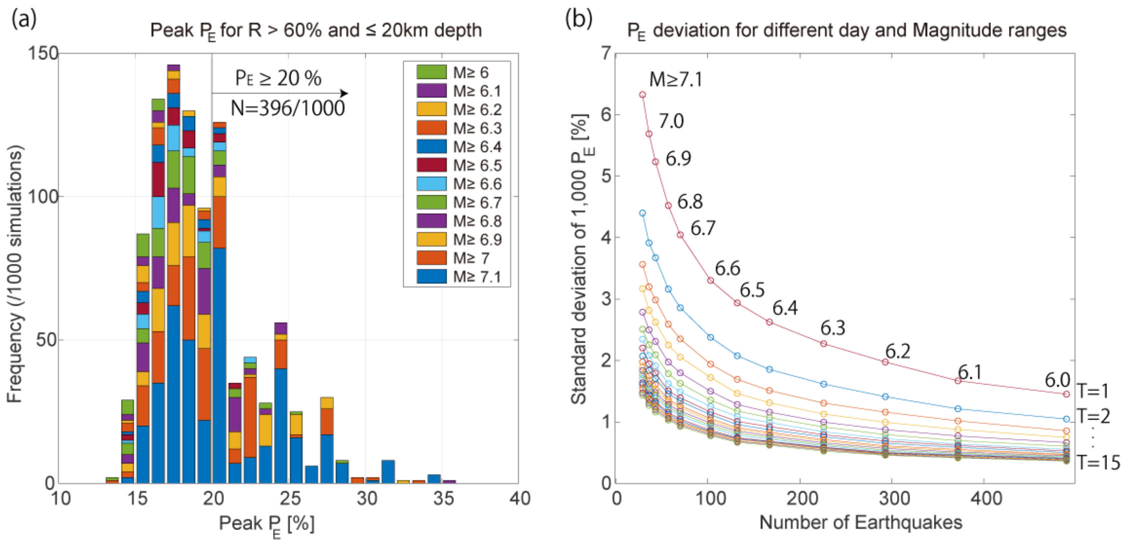

Figure 10. Relationship between the P_E values and different data sets. (a) Histogram of the peak P_E

for 1,000 different simulations, whereby a deviation level R of 60% was assumed for 490 synthetic earthquakes with a ≤20 km hypocenter depth. Colors correspond to the magnitude ranges that yield the largest P_E value in each experience. (b) Standard deviation of the P_E values among 1,000 experiments, whereby each T and M range was based on an assumed deviation level R of 60% and ≤ 20 km hypocenter depth.

5. Conclusion

We tested the applicability of ionospheric TEC variations as earthquake precursors that were reported in Liu et al. (2018) and Le et al. (2011) via random earthquake simulations. We found that their respective analyses reach the conclusion that there is a causal link between earthquakes and ionospheric TEC variations, even for random earthquakes, with probabilities of almost 100% and about 40% for the Liu et al. (2018) Le et al. (2011) assessments, respectively.

However, the Liu et al. (2018) and Le et al. (2011) studies both possessed key issues that hindered proper statistical assessments. The problem with the Liu et al. (2018) statistical analysis, which employs the ROC curve, is that the alarm criteria were first derived from actual earthquakes and then reapplied to predict the same earthquakes. The problem with the Le et al. (2011) statistical evaluation, which assessed the TEC anomalous day rates in the periods preceding the analyzed earthquakes, is that the assessment did not consider the fact that the variations in the anomalous day rate increase as both the period and total number of earthquakes evaluated become smaller. Therefore, the statistical significance proposed by Liu et al. (2018) and Le et al. (2011) cannot be considered conclusive evidence that the observed TEC variations are influenced by pre-earthquake processes.

Acknowledgements

References

Chen, Y. I., Huang, C. S., & Liu, J. Y. (2015). Statistical evidence of seismo-ionospheric precursors applying receiver operating characteristic (ROC) curve on the GPS total electron content in China. Journal of Asian Earth Sciences, 114, 393-402, https://doi.org/10.1016/j.jseaes.2015.05.028

Eisenbeis, J., & Occhipinti, G. (2021). The TEC enhancement before seismic events is an artifact. *Journal o fGeophysical Research: Space Physics*,*126*, e2020JA028733. https://doi.org/10.1029/2020JA028733

Goto, S.-I., Uchida, R., Igarashi, K., Chen, C.-H., Kao, M., & Umeno, K. (2019). Preseismic ionospheric anomalies detected before the 2016 Taiwan earthquake. *Journal of Geophysical Research: Space Physics*, *124*, 9239–9252. https://doi.org/10.1029/2019JA026640

He, L., & Heki, K. (2016). Three-dimensional distribution of ionospheric anomalies prior to three large earthquakes in Chile. *Geophysical Research Letters*, *43*, 7287–7293. https://doi.org/10.1002/2016GL069863

He, L., & Heki, K. (2017). Ionospheric anomalies immediately before Mw 7.0–8.0 earthquakes. *Journal of Geophysical Research: Space Physics*, *122*, 8659–8678. https://doi.org/10.1002/2017JA024012

He, L., & Heki, K. (2018). Three-dimensional tomography of ionospheric anomalies immediately before the 2015 Illapel earthquake, central Chile. *Journal of Geophysical Research: Space Physics*, *123*, 4015–4025. https://doi.org/10.1029/2017JA024871

Heki, K. (2011). Ionospheric electron enhancement preceding the 2011 Tohoku-Oki earthquake. *Geophysical Research Letters*, *38*, L17312. https://doi.org/10.1029/2011gl047908

Heki, K., & Enomoto, Y. (2013). Preseismic ionospheric electron enhancements revisited. *Journal of Geophysical Research: Space Physics*, *118*, 6618–6626. https://doi.org/10.1002/jgra.50578

Heki, K., & Enomoto, Y. (2015). Mw dependence of pre-seismic ionospheric electron enhancements. *Journal of Geophysical Research: Space Physics*, *120*, 7006–7020. https://doi.org/10.1002/2015ja021353

Ikuta, R., Hisada, T., Karakama, G., & Kuwano, O. (2020). Stochastic evaluation of pre-earthquake TEC enhancements. *Journal of Geophysical Research: Space Physics*, *125*, e2020JA027899.

https://doi.org/10.1029/2020JA027899

Ikuta, R., Oba, R., Kiguchi, D., & Hisada, T. (2021). Reanalysis of the ionospheric total electron content anomalies around the 2011 Tohoku-Oki and 2016 Kumamoto earthquakes: Lack of a clear precursor of large earthquakes. *Journal of Geophysical Research: Space Physics*, *126*, e2021JA029376. https://doi.org/10.1029/2021JA029376

Iwata, T., & Umeno, K. (2016). Correlation analysis for preseismic total electron content anomalies around the 2011 Tohoku-Oki earthquake. *Journal of Geophysical Research: Space Physics*, *121*, 8969–8984. https://doi.org/10.1002/2016JA023036

Iwata, T., & Umeno, K. (2017). Preseismic ionospheric anomalies detected before the 2016 Kumamoto earthquake. *Journal of Geophysical Research: Space Physics*, *122*, 3602–3616. https://doi.org/10.1002/2017JA023921

Kamogawa, M., & Kakinami, Y. (2013). Is an ionospheric electron enhancement preceding the 2011 Tohoku-oki earthquake a precursor? *Journal of Geophysical Research: Space Physics*, *118*, 1–1754. https://doi.org/10.1002/jgra.50118

Le, H., Liu, J. Y., & Liu, L. (2011). A statistical analysis of ionospheric anomalies before 736 M6.0+ earthquakes during 2002–2010, Journal of Geophysical Research, 116, A02303, https://doi.org/10.1029/2010JA015781.

Liu, J. Y., Chen, Y. I., Pulinets, S. A., Tsai, Y. B., & Chuo, Y. J. (2001). Seismo-ionospheric signatures prior to M≥6.0 Taiwan earthquakes. *Geophysical Research Letters*, *27*, 19, 3113-3116, https://doi.org/10.1029/2000GL011395

Liu, J. Y., Chen, Y. I., Chuo, Y. J., & Tsai H. F. (2001). Variations of ionospheric total electron content during the Chi-Chi earthquake. *Geophysical Research Letters*, *28*, 7, 1383-1386, https://doi.org/10.1029/2000GL012511

Liu, J. Y., Chen, Y. I., Jhuang, H. K., & Lin, Y. H. (2004). Ionospheric foF2 and TEC Anomalous Days Associated with M≥5.0 Earthquakes in Taiwan during 1997-1999. Terrestrial Atmospheric and Oceanic Sciences, 15, 3, 371-383, https://doi.org/10.3319/TAO.2004.15.3.371

Liu, J. Y., Chen, Y. I., Chen, C. H., Liu, C. Y., Chen, C. Y., Nishihashi, M., Li, J. Z., Xia, Y. Q., Oyama, K. I., Hattori, K., Lin, C. H. (2009). Seismoionospheric GPS total electron content anomalies observed before the 12 May 2008 Mw7.9 Wenchuan earthquake, *Journal of Geophysical Research*: *Space physics*, 114, A04320, https://doi.org/10.1029/2008JA013698

Liu, J. Y., Chen, Y. I., Chen, C. H., & Hattori, K. (2010a), Temporal and spatial precursors in the ionospheric global positioning system (GPS) total electron content observed before the 26 December 2004 M9.3 Sumatra Andaman Earthquake, *Journal of Geophysical Research: Space Physics*, 115, A09312, https://doi.org/10.1029/2010JA015313

Liu., J.Y., Chen, C. H., Chen, Y. I., Yang, W.H., Oyama, K.I., & Kuo, K.W. (2010b). A statistical study of ionospheric earthquake precursors monitored by using equatorial ionization anomaly of GPS TEC in Taiwan during 2001–2007. *Journal of Asian Earth Sciences*, *39*, 76-80, https://doi.org/10.1016/j.jseaes.2010.02.012

Liu, J. Y., Le, H., Chen, Y. I., Chen, C. H., Liu, L., Wan, W., Su, Y. Z., Sun, Y. Y., Lin, C. H., & Chen, M. Q. (2011). Observations and simulations of seismoionospheric GPS total electron content anomalies before the 12 January 2010 M7 Haiti earthquake, *Journal of Geophysical Research: Space physics*, *116*, A04302, https://doi.org/10.1029/2010JA015704.

Liu, J. Y., Chen, C. H., & Tsai, H. F. (2013a) *A statistical study on seismo-ionospheric precursors of the total electron content associated with 146 M ≥ 6.0 earthquakes in Japan during 1998–2011*. *Earthquake Prediction Studies: Seismo Electromagnetics*, edited by M. Hayakawa, 1-13. TERRAPUB

Liu J Y, Chen C H, Tsai H F, Le H. (2013b). A statistical study on seismo-ionospheric anomalies of the total electron content for the period of 56 M ≥ 6.0 earthquakes occurring in China during 1998−2012. *Chinese Journal of Space Science*, *33*, 3, 258-269, https://doi.org/10.11728/cjss2013.03.258.

Liu, C.-Y., Liu, J.-Y., Chen, Y.-I., Qin, F., Chen, W. S., Xia, Y. Q., & Bai, Z.-Q. (2018). Statistical analyses on the ionospheric total electron content related to M ≥ 6.0 earthquakes in China during 1998-2015. *Terrestrial Atmospheric and Oceanic Science*, *29*, 485-498, https://doi.org/10.3319/TAO.2018.03.11.01

Masci, F., Thomas, J. N., Villani, F., Secan, J. A., & Rivera, N. (2015). On the onset of ionospheric precursors 40 min before strong earthquakes. *Journal of Geophysical Research: Space Physics*, *120*, 1383–1393. https://doi.org/10.1002/2014JA020822

Tozzi, R., Masci, F., & Pezzopane, M. (2020). A stress test to evaluate the usefulness of Akaike information criterion in short-term earthquake prediction. *Scientific Reports*, *10*(1), 21153. https://doi.org/10.1038/s41598-020-77834-0

Utada, H., & Shimizu, H. (2014). Comment on “Preseismic ionospheric electron enhancements revisited” by K. Heki and Y. Enomoto. *Journal of Geophysical Research: Space Physics*, *119*, 6011–6015. https://doi.org/10.1002/2014JA020044