Statistical Study of Receiver Operating Characteristic Curve on Seismo-Ionospheric Precursors of the Total Electron Content Associated with Large Earthquakes in Japan

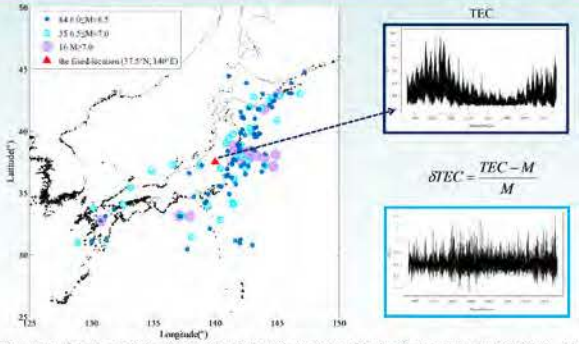
Tiger Jann-Yenq Liu[1]*, Katsumi Hattori[2], Dimitar Ouzounov[3] , Yuh-Ing Chen[4], and Cheng-Yan Liu[5]

[1]Institute of Space Science, National Central University, Taoyuan City, Taiwan
[2]Chiba University, Yayoi 1-33, Inage, Chiba, 263-8522, Japan
[3]CEESMO, Chapman University, Orange, CA, USA
[4]Institute of Statistics, National Central University, Taoyuan City , Taiwan
[5]Institute of Earthquake Prediction, Beijing University of Technology, Beijing, China

Abstract

This paper reports on statistical results of seismo-ionospheric precursors (SIPs) of the total electron content (TEC) in the global ionosphere map associated with M≥6.0 earthquakes in Japan during 1999–2016. A median-based method together with the *z* test is employed to examine the TEC variations 30 days before and after the earthquake. Results show that the SIP is mainly the TEC significant increase few days before the earthquakes. The receiver operating characteristic (ROC) curve is further used to evaluate the efficiency of the TEC for predicting the earthquakes in Japan during the time period with statistical significance detected by the z test. The SIP is further confirmed since the area under the ROC curve is positively associated with the earthquake magnitude.

Data



$$\delta TEC = \frac{TEC - M}{M}$$

Figure 1. Locations of the GIM TEC and the 135 M≥6.0 earthquakes in Japan during 1999-2016. The earthquake catalog is retrieved from USGS. The triangle denote (37.5° N, 140° E), the location of GIM TEC used in this study.

TEC Reference

M, *LQ* and *UQ* are the median, first and third quartiles of TEC based on previous 15-day TECs
Set the upper bound $LB=M-k(M-LQ)$ and the lower bound $UB=M+k(UQ-M)$

The standardized TEC change $\delta TEC = \frac{TEC - M}{M}$

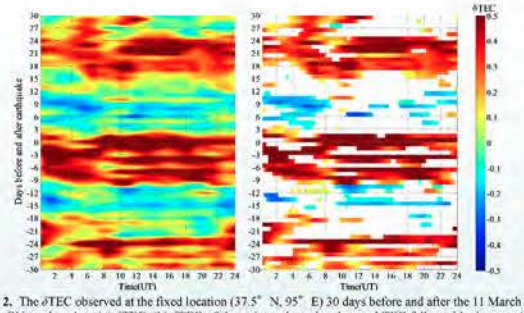


Figure 2. The δTEC observed at the fixed location (37.5° N, 95° E) 30 days before and after the 11 March 2011 M9.0 Tohoku-Oki earthquake. (a) δTEC, (b) δTEC of the points where the observed TEC fall outside the associated LB or UB with k=1.5.

Anomaly Zone

p : the observed proportion of earthquake-related anomalies
p_0 : the background proportion of anomalies n : the number of earthquakes

The *z* value is $z = \frac{p - p_0}{\sqrt{p_0(1-p_0)/n}}$

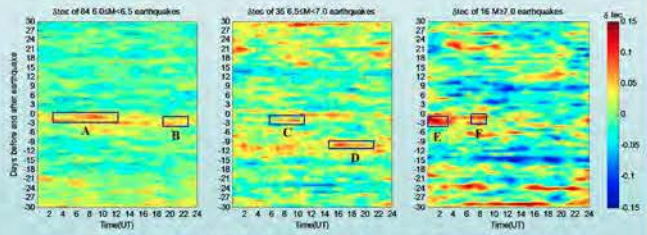


Figure 3. The median values of δTEC in 30 days before and after Japan earthquakes with magnitude (a) 6.0≤ M<6.5 (b) 6.5 ≤M<7.0, and (c) M ≥7.0. The contour denotes significant *z* test at significance level 0.05

TEC increment Zone A: 0600-1000 UT 1-2 days Zone B: 1900-2200 UT 2-3days before EQKs
Zone C: 0700-1000 UT 2-3 days Zone D: 1500-2100 UT 9-10 days before EQKs
Zone E: 0100-0300 UT 2-3 days Zone F: 0700-0900 UT 1-3 days before EQKs

Anomaly Occurrence Frequency

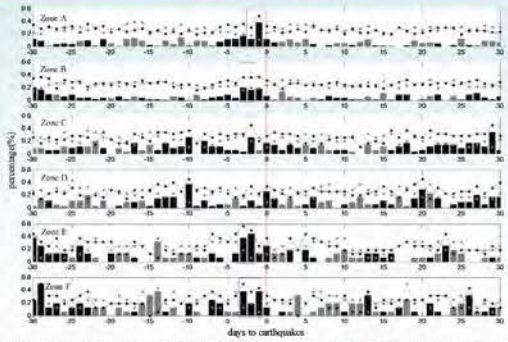


Figure 4. Percentages of the earthquakes with positive (black dot) and negative (gray dot) precursory days. The black bar represents the amount of percentage in which positive anomaly is over negative anomaly, while the gray bar denotes the amount of percentage in which negative anomaly is over positive anomaly.

Receiver Operating Characteristic (ROC) Curve

- For a cutoff value *k*, We compute the true positive rate (TPR) and false positive rate (FPR) as
  TPR=(No. of alarmed earthquake days)/(No. of earthquake days)
  and FPR=(No. of alarmed non-earthquake days)/(No. of non-earthquake days).
  The ROC curve with FPR as the x-axis and TPR as the y-axis is obtained for a variety of possible *k* value.
- The area under the ROC curve (AUC) is used for assessing the effectiveness of the TEC precursor.
- For testing H_0 (SIPs are randomly observed) versus H_A (SIPs are not randomly observed), We randomly generate anomaly days based on Gamma distribution during the study period, and then calculate the ROC curve and max-AUC for each simulation.

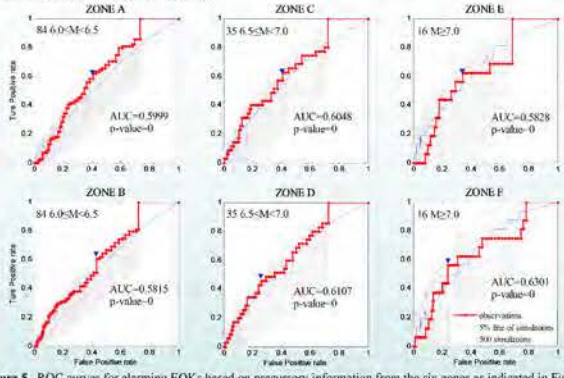


Figure 5. ROC curves for alarming EQKs based on precursory information from the six zones as indicated in Figure 3. The red, gray and blue curves denote the ROC curves of the observations, simulations, and the 5% line of simulations respectively. The red inverted triangle denotes the *k* value yielding the maximum R score (= TPR-FPR), which is called the Youden index (Youden, 1950)

Table 1. AUCs of Different Earthquak Groups

EQ	Alarming day	Anomaly period (LT)	AUC	k-value	TPR	FPR	p-value(AUC*)	Alarms
84 6.0≤M<6.5	Zone A	1-2	0600-1000	0.5999	1.4	0.6190	0.4279	0(0.5247)	2751
	Zone B	2-3	1900-2200	0.5815	1.2	0.6071	0.4538	0(0.5265)	2913
35 6.5≤M<7.0	Zone C	2-3	0700-1000	0.6048	1.3	0.6286	0.4268	0(0.5482)	2735
	Zone D	9-10	1500-2100	0.6107	2.1	0.4571	0.2774	0(0.5398)	1779
16 M≥7.0	Zone E	2-3	0100-0300	0.5825	1.6	0.6250	0.3600	0(0.5641)	2305
	Zone F	1-3	0700-0900	0.6301	2.7	0.5625	0.2505	0(0.5521)	1606

Conclusion

1. The *z* test at significance level 0.05 shows the SIPs appears in different time/day zones prior to three group of earthquakes (84 6.0≤ M<6.5, 6.5 ≤ M<7.0, and M ≥ 7.0) in Japan during 1999-2016
2. The ROC curves and the associated AUC*s recommend that a significant TEC increment in six zone provides with a reliable SIP for three group earthquakes in Japan, respectively.
3. The average of median δTECs is increasing with the associated earthquake magnitude and is closely related to the associated AUC. It implies that larger earthquakes have the greater preparation, and in turn release a stronger SIP of GPS TEC.