Investigation of total electron content variability due to seismic and geomagnetic disturbances in the ionosphere

Secil Karatay,[1] Feza Arikan,[2] and Orhan Arikan[3]

Received 19 October 2009; revised 23 March 2010; accepted 27 May 2010; published 20 October 2010.

[1] Variations in solar, geomagnetic, and seismic activity can cause deviations in the ionospheric plasma that can be detected as disturbances in both natural and man-made signals. Total electron content (TEC) is an efficient means for investigating the structure of the ionosphere by making use of GPS receivers. In this study, TEC data obtained for eight GPS stations are compared with each other using the cross-correlation coefficient (CC), symmetric Kullback-Leibler distance (KLD), and L2 norm (L2N) for quiet days of the ionosphere, during severe geomagnetic storms and strong earthquakes. It is observed that only KLD and L2N can differentiate the seismic activity from the geomagnetic disturbance and quiet ionosphere if the stations are in a radius of 340 km. When TEC for each station is compared with an average quiet day TEC for all periods using CC, KLD, and L2N, it is observed that, again, only KLD and L2N can distinguish the approaching seismicity for stations that are within 150 km radius to the epicenter. When the TEC of consecutive days for each station and for all periods are compared, it is observed that CC, KLD, and L2N methods are all successful in distinguishing the geomagnetic disturbances. Using sliding-window statistical analysis, moving averages of daily TEC with estimated variance bounds are also obtained for all stations and for all days of interest. When these bounds are compared with each other for all periods, it is observed that CC, KLD, and L2N are successful tools for detecting ionospheric disturbances.

Citation: Karatay, S., F. Arikan, and O. Arikan (2010), Investigation of total electron content variability due to seismic and geomagnetic disturbances in the ionosphere, *Radio Sci.*, *45*, RS5012, doi:10.1029/2009RS004313.

1. Introduction

[2] The ionosphere is an important layer of Earth's upper atmosphere that extends from 60 to 1000 km and is ionized to a plasma state primarily by radiation from the sun with density N_e varying with altitude up to 10^{12} m^{-3}. The determining parameter of the ionospheric plasma is the electron concentration, which is a complex function of variations and coupling in solar, geomagnetic, and seismic activities such as solar flares, sunspot number, solar wind, geomagnetic storms, and earthquakes. An important measurable quantity of the electron density is the total electron content (TEC), which provides an efficient means to investigate the structure of the ionosphere and upper atmosphere. TEC is defined as the line integral of electron density along a raypath or as a measure of the total number of electrons along a raypath. The unit of TEC is given in TECU, where 1 TECU = 10^{16} el/m^2 [*Nayir et al.*, 2007; *Arikan et al.*, 2003]. The variations and disturbances of the ionosphere can be obtained effectively and efficiently by computing and monitoring TEC. In recent decades, the Global Positioning System (GPS), with its network of worldwide receivers, provides a cost-effective solution in estimating TEC (GPS-TEC) and monitoring ionospheric variability over a significant proportion of global landmass [*Nayir et al.*, 2007; *Arikan et al.*, 2003].

[3] The general trends of temporal and spatial variability of the ionosphere depend on Earth's diurnal and annual rotation and the distribution of magnetic field lines of the geomagnetic dipole. Earth's magnetic field is seldom quiet, even when there are no storms. The underlying trends and standard periodical variations make up the dynamics of quiet ionosphere [*Rishbeth and Garriot*, 1969]. It has long been observed that variations

[1]Department of Physics, Faculty of Arts and Sciences, Firat University, Elazig, Turkey.
[2]Department of Electrical and Electronics Engineering, Hacettepe University, Ankara, Turkey.
[3]Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey.

0048-6604/10/2009RS004313

in the solar and geomagnetic activity and seismicity can cause deviations from the quiet conditions and these changes can be detected as disturbances in both natural and man-made signal parameters. If the magnetic field is severely disturbed, a magnetic storm is said to occur. The geomagnetic storms can be listed as one of the major sources of severe temporal and spatial variability in the ionosphere. Several empirical indices have been developed to describe the amount of the variability at any given time. These disturbances are due to the coupling of solar activity with Earth's magnetic field that involves highly complicated dynamics in the magnetosphere and ionosphere. A large number of studies in the literature [*Rishbeth and Garriot*, 1969; *Biqiang et al.*, 2007; *Vlasov et al.*, 2003; *Zhang and Xiao*, 2000] investigating the ionospheric disturbances suggest that geomagnetic storms can cause strong disturbances in the electron density distribution and TEC. Also, in recent years, the coupling of seismic activity in the lithosphere with the troposphere and ionosphere has been observed through the variations in electromagnetic signals, Earth's electric and magnetic fields, and the chemical composition of the atmosphere. Recently, there have been some theories that try to explain electromagnetic anomalies associated with preseismic activity and their effects in the ionosphere. The forecast methods [e.g., *Ondoh*, 2000; *Pulinets*, 2004; *Pulinets et al.*, 2004, 2005, 2007; *Liu et al.*, 2000, 2004; *Chuo et al.*, 2001; *Plotkin*, 2003; *Trigunait et al.*, 2004] suggest that, before the strong earthquakes, there are several disturbances in the ionospheric parameters, especially in the critical frequency of the F2 layer (foF2), ion temperatures (T_i), and TEC.

[4] In the literature, the basic statistical tools that are used to investigate the effect of ionospheric disturbances on ionospheric parameters include but are not limited to relative deviation [*Kouris and Fotiadis*, 2002; *Kouris et al.*, 2006], time derivative analysis [*Ciraolo and Spalla*, 1999], interquartile range analysis [*Liu et al.*, 2004; *Chuo et al.*, 2001; *Lazo et al.*, 2004; *Zhang et al.*, 2004], correlation analysis [*Pulinets*, 2004; *Pulinets et al.*, 2004, 2005, 2007], TEC difference analysis [*Plotkin*, 2003], and ionospheric correction [*Trigunait et al.*, 2004]. All of these methods are applied to severe geomagnetic storms or major earthquakes with magnitudes $M \geq 6$. Yet the investigated disturbance periods and data sets are still very limited. Also, the reliability and accuracy of the applied statistical tools still need to be reconsidered for a precursor alarm signal before geomagnetic or seismic disturbances.

[5] In statistics and information theory, the Kullback-Leibler divergence is a widely used measure of distance of discrimination between two probability density distributions [*Cover and Thomas*, 2006; *Hall*, 1987; *Inglada*, 2003]. Similarly, the L2 norm is used to define the Euclidian distance between two vectors [*Kreyszig*, 1988]. Sliding-window statistical analysis with moving average and variance bounds is a useful tool in defining the time-varying general trend of the data and characterization of the underlying structure of the disturbances in terms of wide sense stationarity [*Arikan and Erol*, 1998; *Erol and Arikan*, 2005]. In this study, the variability of GPS-TEC is investigated by comparison of the disturbed days with the quiet-day trends of the ionosphere over a large data set using the measures of Kullback-Leibler divergence, L2 norm, and sliding-window statistical analysis for the first time in the literature [*Arikan et al.*, 2009; *Karatay et al.*, 2009a, 2009b]. The correlation coefficients of data sets are also computed in spatial and temporal domains. Six earthquakes with different seismic properties and two very severe geomagnetic disturbances are chosen for investigation in this study. GPS-TEC is computed for 15 days before and after each earthquake (earthquake-days period) for all the GPS stations at various distances from the earthquake epicenter. TEC values are also obtained for the periods when there is no seismic activity but the ionosphere is under the influence of strong geomagnetic disturbances (disturbed-days period) and also for the periods when there are no significant disturbances or seismic activity in the regions of interest (quiet-days period). The results are obtained for three groups of application. In the first group, the statistical tools are applied between neighboring stations for all periods. In the second group, an average quiet-day TEC estimate is obtained for each station and TEC estimates for all periods are compared with this average quiet-day TEC using statistical tools. In the third group, TEC estimates for consecutive days of all periods are compared with each other. The statistical methods used in the study and the results for the data are presented in sections 2 and 3, respectively.

2. Methods of the Statistical Analysis

[6] In GPS-TEC computation, TEC on the slant raypath from the satellite to the receiver is called the slant TEC (STEC). When the STEC values are projected to the local zenith at the ionospheric pierce point, the computed TEC value is called the vertical TEC (VTEC) [*Nayir et al.*, 2007; *Arikan et al.*, 2003]. Let

$$\mathbf{x}_{u;d} = \left[x_{u;d}(1) \cdots x_{u;d}(n) \cdots x_{u;d}(N)\right]^{\mathrm{T}} \qquad (1)$$

represent the set of VTEC data of length N estimated for day d. Here, u denotes the receiver, n is the index where $1 \leq n \leq N$, and T is the transpose operator. To compare VTEC values obtained from different seasons and days, data vectors as in equation (1) are normalized. The experimental probability density function (PDF) of TEC

for station u and day d can be approximated using the TEC measurements as

$$\hat{\mathbf{P}}_{u;d} = \mathbf{x}_{u;d}\left[\sum_{n=N_i}^{N_s} x_{u;d}(n)\right]^{-1} \tag{2}$$

where N_i and N_s denote the initial and final indices for the measurement set, respectively. To compare the behavior of TEC for the quiet days with those from the disturbed and earthquake days, an average quiet-day TEC (AQDT) estimate for each GPS station is obtained. For N_d quiet days for station u, the AQDT is defined as

$$\mathbf{x}_{u;d_i-d_s} = \frac{1}{N_d}\sum_{n_d=d_i}^{d_s} \mathbf{x}_{u;n_d}, \tag{3}$$

where n_d is the index for the quiet-day period which extends from d_i to d_s. An approximation for the PDF of the AQDT is defined as follows:

$$\hat{\mathbf{P}}_{u;d_i-d_s} = \mathbf{x}_{u;d_i-d_s}\left[\sum_{n=N_i}^{N_s} x_{u;d_i-d_s}(n)\right]^{-1}. \tag{4}$$

[7] Using the normalized distributions given in equations (2) and (4), the following statistical tools are applied to the VTEC data.

2.1. Cross-Correlation Coefficient

[8] In statistics, the correlation function measures the relationship between two random variables. In this study, daily cross-correlation coefficients are computed between two GPS stations as presented in the literature. First, the station nearest the epicenter, which is denoted as the central station, is chosen. Then, the correlation coefficients are computed between VTEC data of the central station and that of the other stations. Finally, correlation coefficients are computed between the TEC values for all station pairs. For N_T samples from N_i to N_s, the cross-correlation coefficient (CC) for day d between stations u and v is defined as

$$r\left(\mathbf{x}_{u;d};\mathbf{x}_{v;d}\right) = \frac{1}{N_T\sigma_{u;d}\sigma_{v;d}}\sum_{n=N_i}^{N_s}\left(x_{u;d}(n)-\overline{\mathbf{x}}_{u;d}\right)\left(x_{v;d}(n)-\overline{\mathbf{x}}_{v;d}\right), \tag{5}$$

where $\overline{\mathbf{x}}_{u;d}$ and $\sigma_{u;d}$ denote the mean and the standard deviation of $\mathbf{x}_{u;d}$, respectively. Using the normalized average quiet-day TEC given in equation (4), the CC is computed for station u between day d and the normalized average quiet day as $r(\hat{\mathbf{P}}_{u;d};\hat{\mathbf{P}}_{u;di-ds})$. For the consecutive days of station u, the CC is computed as $r(\hat{\mathbf{P}}_{u;d};\hat{\mathbf{P}}_{u;d+1})$.

2.2. Kullback-Leibler Distance

[9] Kullback-Leibler (KL) divergence is used in various disciplines to define similarity and difference between two distributions [*Cover and Thomas*, 2006; *Arikan et al.*, 2009; *Karatay et al.*, 2009a, 2009b]. The KL divergences of normalized experimental PDFs defined in equation (2) for day d between stations u and v can be defined as

$$\mathrm{KL}\left(\hat{\mathbf{P}}_{u;d}\setminus\hat{\mathbf{P}}_{v;d}\right) = \sum_{n=N_i}^{N_s}\hat{P}_{u;d}(n)\ln\left(\frac{\hat{P}_{u;d}(n)}{\hat{P}_{v;d}(n)}\right), \tag{6}$$

$$\mathrm{KL}\left(\hat{\mathbf{P}}_{v;d}\setminus\hat{\mathbf{P}}_{u;d}\right) = \sum_{n=N_i}^{N_s}\hat{P}_{v;d}(n)\ln\left(\frac{\hat{P}_{v;d}(n)}{\hat{P}_{u;d}(n)}\right), \tag{7}$$

where $N_i < n < N_s$. The symmetric Kullback-Leibler distance (KLD) is defined as the sum of the Kullback-Leibler divergences [*Cover and Thomas*, 2006; *Hall*, 1987; *Inglada*, 2003; *Arikan et al.*, 2009; *Karatay et al.*, 2009a, 2009b] as

$$\mathrm{KLD}\left(\hat{\mathbf{P}}_{u;d};\hat{\mathbf{P}}_{v;d}\right) = \mathrm{KL}\left(\hat{\mathbf{P}}_{u;d}\setminus\hat{\mathbf{P}}_{v;d}\right) + \mathrm{KL}\left(\hat{\mathbf{P}}_{v;d}\setminus\hat{\mathbf{P}}_{u;d}\right). \tag{8}$$

[10] Using normalized AQDT, for day d of station u, symmetric KLD is defined as $\mathrm{KLD}(\hat{\mathbf{P}}_{u;d};\hat{\mathbf{P}}_{u;di-ds})$. For consecutive days of station u, symmetric KLD is defined as $\mathrm{KLD}(\hat{\mathbf{P}}_{u;d};\hat{\mathbf{P}}_{u;d+1})$.

2.3. L2 Norm

[11] The L2 norm (L2N) is a metric measure that quantifies the distance between two vectors. Using the normalized experimental PDFs given in equation (2), the L2 norm for day d, between stations u and v, can be defined as [*Kreyszig*, 1988; *Arikan et al.*, 2009; *Karatay et al.*, 2009a, 2009b]

$$\mathrm{L2N}\left(\hat{\mathbf{P}}_{u;d}\setminus\hat{\mathbf{P}}_{v;d}\right) = \sqrt{\sum_{n=N_i}^{N_s}\left(\hat{P}_{u;d}(n)-\hat{P}_{v;d}(n)\right)^2}, \tag{9}$$

where $N_i < n < N_s$. For station u between day d and normalized AQDT, the L2 norm is defined as $\mathrm{L2N}(\hat{\mathbf{P}}_{u;d}\setminus\hat{\mathbf{P}}_{u;di-ds})$. For consecutive days of station u, the L2 norm is computed as $\mathrm{L2N}(\hat{\mathbf{P}}_{u;d}\setminus\hat{\mathbf{P}}_{u;d+1})$.

Table 1. Date, Time, Geographical Location, Magnitude, and Depth of the Chosen Earthquakes

Earthquakes	Date	Time (UT)	Latitude	Longitude	M	z (km)	Central Station
E1	25 Sep 2003	1950	41°N	143°E	8.3	27	mizu
E2	5 Sep 2004	1457	33°N	137°E	7.4	10	kgni
E3	13 Jun 2008	2343	39°N	140°E	6.9	10	mizu
E4	11 Jun 2006	2001	33°N	131°E	6.3	154	usud
E5	23 Jul 2005	0734	35°N	139°E	5.9	65	mtka
E6	12 May 2008	0628	30°N	103°E	7.9	19	kunm

2.4. Sliding-Window Statistical Analysis

[12] The ionosphere presents a spatial and temporal variability at different scales. To determine the characterization of the variability of the TEC values, the sliding-window statistical analysis method is applied to the VTEC data sets. The sliding-window moving average and estimated standard deviation of normalized TEC distributions of station u and day d can be given as [*Arikan and Erol*, 1998; *Erol and Arikan*, 2005; *Karatay et al.*, 2009b]

$$\hat{\mu}_{u;d}(n) = \frac{1}{N_w} \sum_{i=\frac{-N_w+1}{2}}^{\frac{N_w-1}{2}} \hat{P}_{u;d}(n+i), \tag{10}$$

$$\hat{\sigma}_{\mu;u;d}(n) = \sqrt{\frac{1}{N_w} \sum_{j=-(N_w-1)}^{N_w-1} r_{u;d}(|j|;n)\left(1 - \frac{|j|}{N_w}\right) - \hat{\mu}^2_{u;d}(n)}, \tag{11}$$

where N_w is the window size which is chosen as an odd number and $r_{u;d}(j;n)$ is the local correlation function that can be approximated as [*Arikan and Erol*, 1998; *Erol and Arikan*, 2005; *Karatay et al.*, 2009b]:

$$r_{u;d}(j;n) \simeq \frac{1}{2N+1-j} \sum_{i=-N}^{N} \hat{\mathbf{P}}_{u;d}(n-i)\hat{\mathbf{P}}_{u;d}(n-i+j), \tag{12}$$

where N can be chosen close to N_w. The window size N_w is chosen to be as long as possible to provide better statistical characterization and to be short enough to capture the local variability. The sliding-window moving average and estimated standard deviation are also computed for the normalized average quiet-day TEC given in equation (4). Using the sliding-window moving average, [equation (10)], and estimated standard deviation [equation (11)], a bound b for station u and day d is computed as follows:

$$b_{u;d}(n) = \hat{\mu}_{u;d}(n) + \hat{\sigma}_{\mu;u;d}(n). \tag{13}$$

[13] Using equation (13), for AQDT, and those of each day for all earthquake-, disturbed-, and quiet-day periods of each station, CC, KLD, and L2N are computed for station u between day d and the normalized average quiet day as r_b ($\hat{\mathbf{P}}_{u;d}$; $\hat{\mathbf{P}}_{u;di-ds}$), KLD_b ($\hat{\mathbf{P}}_{u;d}$; $\hat{\mathbf{P}}_{u;di-ds}$), and $\mathrm{L2N}_b$ ($\hat{\mathbf{P}}_{u;d}\backslash\hat{\mathbf{P}}_{u;di-ds}$), respectively. Between the bounds of consecutive days of all periods of station u, CC, KLD, and L2N are computed as r_b ($\hat{\mathbf{P}}_{u;d}$; $\hat{\mathbf{P}}_{u;d+1}$), KLD_b ($\hat{\mathbf{P}}_{u;d}$; $\hat{\mathbf{P}}_{u;d+1}$), and $\mathrm{L2N}_b$ ($\hat{\mathbf{P}}_{u;d}\backslash\hat{\mathbf{P}}_{u;d+1}$), respectively.

3. Results and Discussion

[14] The statistical analysis tools described in section 2 are applied to VTEC data in search of a precursor signal for geomagnetic and seismic disturbances. These methods are used in three major groups of application. In group I, the relationship between the GPS stations is investigated in terms of distance between them. In group II, VTEC data of each station are compared to average quiet-day VTEC data for that station. In group III, VTEC data of consecutive days for each station are compared with each other.

[15] For this study, five earthquakes with magnitudes between 5.9 and 8.3 that occurred in Japan and one earthquake with magnitude 7.9 that occurred on 12 May 2008 in China are chosen. The geographical location of the epicenter, GPS station closest to the epicenter (central station), date, time (universal time), magnitude (M, Richter scale), and depth (z, kilometers) of the chosen earthquakes, E, are presented in Table 1 (data available at http://earthquake.usgs.gov/regional/world). The earthquake day periods (EDPs) for each earthquake are defined as the time period from 15 days prior to the earthquake, the earthquake day, and 15 days after the earthquake (31 days). There are no significant geomagnetic disturbances during the chosen EDPs. The days of the earthquakes in EDPs are indicated by an arrow in all figures.

[16] The geomagnetic disturbance days are chosen such that there is no significant seismic activity in the region of interest. The first period is chosen as DDP_1, from 14 October to 11 November 2003 (29 days), and the second period is chosen as DDP_2, from 23 August to 21 September 2005 (30 days). The quiet days are chosen such that there are no significant geomagnetic or seismic

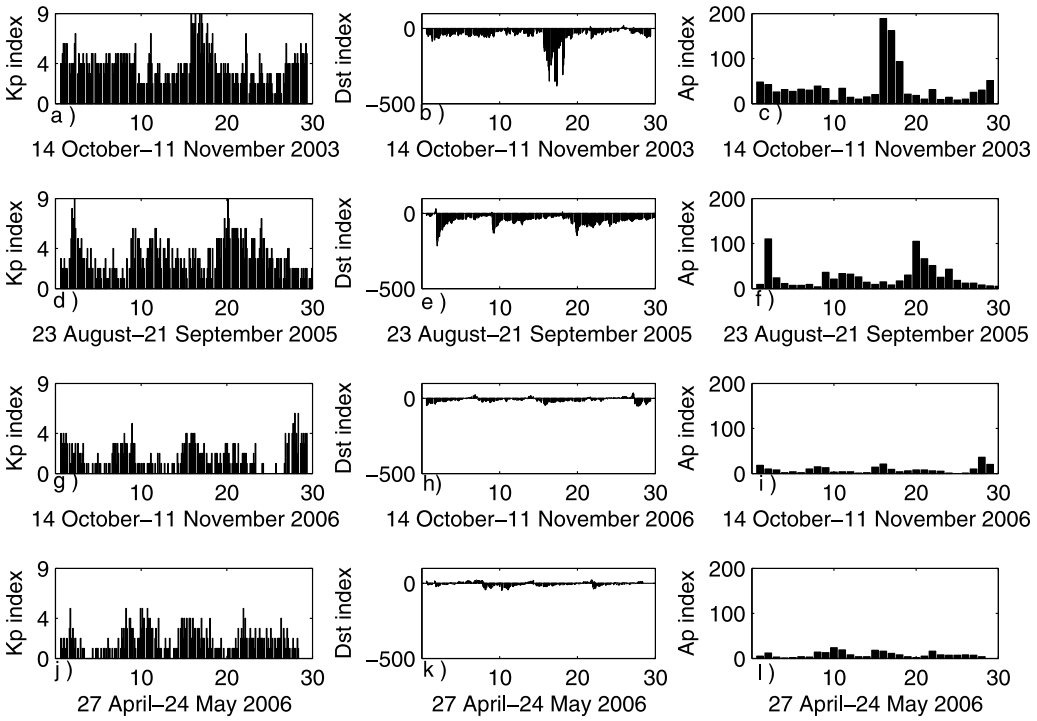

Figure 1. Daily geomagnetic indices *Kp*, *Dst*, and *Ap* for (a–c) DDP_1, (d–f) DDP_2, (g–i) QDP_1, and (j–l) QDP_2.

activities in the region of interest. The first quiet-days period is chosen as QDP_1, from 14 October to 11 November 2006 (29 days), and the second quiet-days period is chosen as QDP_2, from 27 April to 24 May 2006 (28 days). The time interval from 27 April to 24 May 2006 for the AQDT is chosen for the GPS stations placed in Japan and the time interval from 14 October to 11 November 2006 for the AQDT is chosen for the GPS station placed in China. (Quiet- and disturbed-day periods are chosen according to the information provided at http://www.swpc.noaa.gov/ftpmenu/indices/old_indices, http://wdc.kugi.kyoto-u.ac.jp/dstdir/, and http://www.cbk.waw.pl/rwc/idce.html.) The *Kp*, *Ap*, and *Dst* indices corresponding to DDP_1 and DDP_2 are given in Figures 1a–1f. The *Kp*, *Ap*, and *Dst* indices corresponding to QDP_1 and QDP_2 are provided in Figures 1g–1l.

[17] The raw data for the corresponding seven GPS stations in the region of interest are obtained from the International Global Navigation Satellite Systems (GNSS) Service (IGS; data available at http://igscb.jpl.nasa.gov/). The distance between IGS-GPS stations to the earthquake epicenters varies from 33 to 2000 km. The geographical locations of the GPS stations are listed in Table 2. The VTEC values for each station are estimated by IONOLAB-TEC using the Reg-Est algorithm described in the literature [*Nayir et al.*, 2007; *Arikan et al.*, 2003, 2004; see also http://www.ionolab.org] with a time resolution of 2.5 min. The missing values in tables and figures in this section are due to the lack of raw data for those stations and days.

[18] In group I analysis, the GPS stations are ordered in pairs according to the great circle distance between them. For the earthquakes given in Table 1 and GPS stations given in Table 2, the distance k for station pairs are categorized into six groups: $k1$, $k < 20$ km; $k2$, 30 km $< k <$ 70 km; $k3$, 80 km $< k <$ 150 km; $k4$, 150 km $< k <$ 340 km; $k5$, 340 km $< k <$ 450 km; and $k6$, 450 km $< k$. These distance categories are chosen such that there is at least one station pair in each category. The IONOLAB-TEC for each station and for all the days of EDP; DDP_1

Table 2. Selected GPS Receiver Stations

Receiver Station	Station ID	Latitude	Longitude
Koganei, Japan	kgni	35.5°N	139.4°E
Kashima, Japan	ksmv	35.7°N	140.6°E
Mizusawa, Japan	mizu	38.9°N	141.1°E
Mitaka, Japan	mtka	35.4°N	139.5°E
Tsukuba, Japan	tskb	35.9°N	140.0°E
Usuda, Japan	usud	35.9°N	138.3°E
Yuzh-Sakh, Russia	yssk	46.8°N	142.7°E
Kunminimumg, China	kunm	24.8°N	102.8°E

Table 3. Percentage of CC Values Less Than 0.9 for Station Pairs for Quiet-Day Period, Days Before the Earthquakes, and Disturbed-Day Period

Period	*k*1	*k*2	*k*3	*k*4	*k*5	*k*6
Quiet-day period (QDP_1)	0	1.2	5	7.6	2.9	31.6
Disturbed-day period (DDP_1, DDP_2)	0	1.6	5.1		0	0
Days before each earthquake	2.5	3	3	3.3	9.2	35.8

for kgni, ksmv, mtka, and tskb; and DDP_2 for mizu, usud, and QDP_1 are compared with each other using CC, KLD, and L2N given in equations (5), (8), and (9), respectively. The total number of CCs in all distance categories and for all periods is 2171. In the literature [*Pulinets*, 2004; *Pulinets et al.*, 2004, 2005, 2007], two data sets are considered to be highly correlated if the daily CC is higher than 0.9. In this study, the ratio of CC that is under the threshold 0.9 to the total number of CC in each period and distance category is computed and presented in Table 3 as percentage values. It is observed from Table 3 that CC is not a good measure to differentiate either geomagnetically or seismically disturbed days from the quiet days. For distance categories of *k*1 and *k*2, the VTEC is highly correlated for all EDP, DDP_1, DDP_2, and QDP_1. Although VTEC is still highly correlated during DDP_1 and DDP_2 for categories *k*5 and *k*6, the correlation decreases for both QDP_1 and days before the earthquakes. In Table 3, the number of days before the earthquakes is chosen as 15 days before each earthquake. These time intervals are 10–24 September 2003 for E1, 21 August to 4 September 2004 for E2, 29 May to 12 June 2008 for E3, 27 May to 10 June 2006 for E4, and 8–22 July 2005 for E5. A similar result is demonstrated in Figure 2 for two station pairs, one pair in category *k*2, mtka-tskb (67 km), and the other pair in *k*6, mtka-yssk (1436 km). In Figures 2a, 2b, and 2c, the correlation coefficients for the mtka-tskb pair are given for earthquake E5, QDP_1, and DDP_1, respectively. It is observed that the station pair that is in category *k*2 has high correlation coefficients for all EDP, QDP_1, and DDP_1. In Figures 2d, 2e, and 2f, the correlation coefficients for the mtka-yssk pair are given for earthquake E5, QDP_1, and DDP_1, respectively. It is observed that the station pair that is in category *k*6 has low correlation coefficients for all EDP, QDP_1, and DDP_1. Thus, it is impossible to discriminate the earthquake and disturbed-day periods from quiet-day periods.

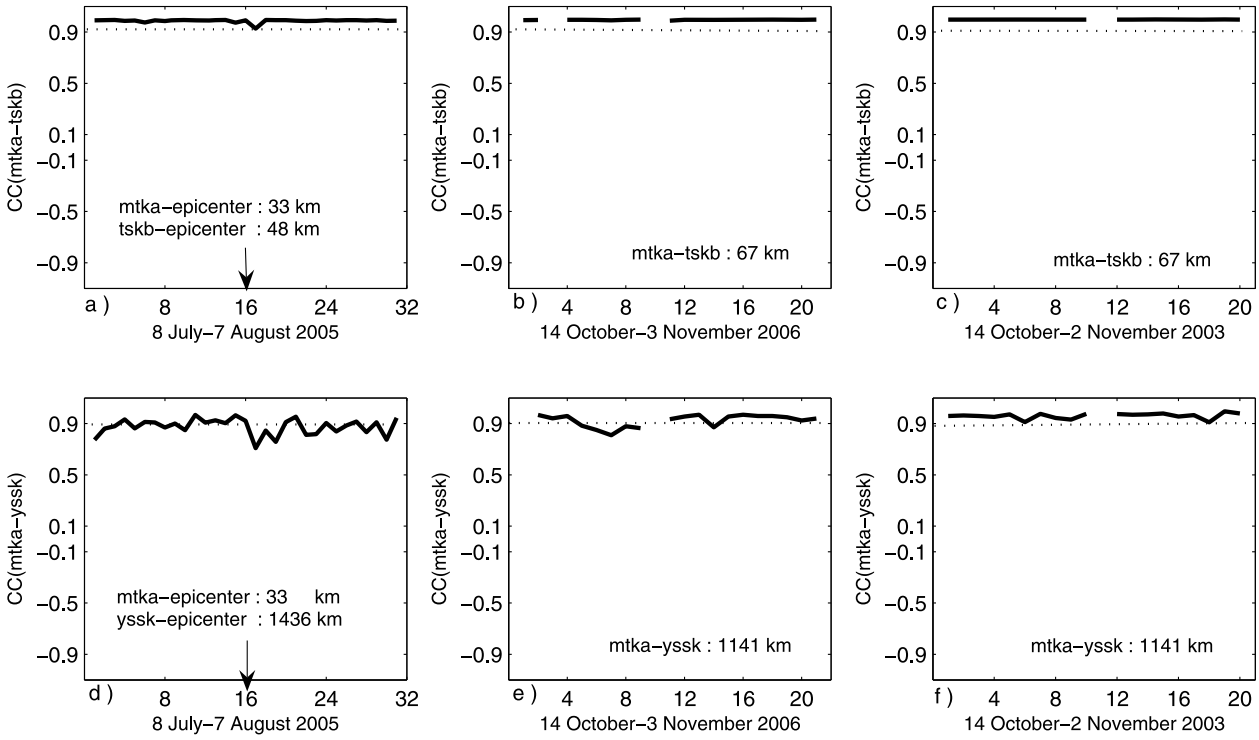

Figure 2. Cross-correlation coefficients between (a) mtka-tskb in EDP for E5, (b) mtka-tskb in QDP_1, (c) mtka-tskb in DDP_1, (d) mtka-yssk in EDP for E5, (e) mtka-yssk in QDP_1, and (f) mtka-yssk in DDP_1. The day of the earthquake is indicated by the arrow in Figures 2a and 2d.

Table 4. Values of Δ_{KLD} for All Station Pairs in Distance Categories $k1$ to $k6$ for QDP_1, DDP_1, and DDP_2 and E1, E2, E3, E4, and E5 Earthquakes

Period	$k1$	$k2$	$k3$	$k4$	$k5$	$k6$
QDP_1	0.0005	0.0008	0.0005	0.0006	0.008	0.06
DDP_1, DDP_2	0.0006	0.0014	0.0014		0.009	0.064
E1	0.0196	0.0094	0.0053	0.0009	0.023	0.028
E2		0.0013	0.00069	0.0008	0.017	0.021
E3			0.002	0.023	0.006	
E4	0.0156	0.0096	0.013	0.001	0.0087	0.02
E5	0.0898	0.049	0.048	0.0047		0.054

[19] The symmetric Kullback-Leibler distance (KLD) and L2N given in equations (8) and (9), respectively, are computed using IONOLAB-TEC for each station pair in categories $k1$ to $k6$ and for all periods in group I. It is observed that when the distance between the stations is less than 150 km (categories $k1$ to $k3$), KLD and L2N values of earthquake days are significantly greater than those for quiet days. For the distance between stations greater than 340 km (category $k4$), KLD and L2N values of earthquake days and quiet days are similar. In addition, if the station pairs are close to the earthquake epicenters as in E3 and E5, the KLD and L2N values of those station pairs in earthquake days are significantly greater than those of quiet days. To demonstrate this observation, the difference between the maximum and minimum of KLD and L2N values in a given period are computed as $\Delta_{KLD} = \max(KLD) - \min(KLD)$ and $\Delta_{L2N} = \max(L2N) - \min(L2N)$. Δ_{KLD} and Δ_{L2N} for all station pairs in a distance category k, and for all earthquake-, quiet-, and disturbed-day periods according to the GPS stations of concern in Japan are given in Tables 4 and 5, respectively. The missing values in Tables 4 and 5 are due to the lack of raw data for the GPS stations in those station pairs. It is observed from Tables 4 and 5 that for distances larger than 340 km between the station pairs corresponding to $k5$ and $k6$, the Δ_{KLD} and Δ_{L2N} values are similar for both quiet and disturbed days. Even for distances less than 150 km between the stations (categories $k1$ to $k3$), the Δ_{KLD} and Δ_{L2N} are very similar for quiet and disturbed days. Yet, for earthquake-day periods, Δ_{KLD} and Δ_{L2N} are significantly larger from those of quiet- and disturbed-day periods for all distance categories. Thus, KLD and L2N are possible candidates to be used as indicators of earthquakes if these measures are constantly monitored for the stations that are close to the earthquake zones.

[20] To demonstrate the performance of CC, KLD, and L2N methods for earthquake and disturbed days, an example is given in Figure 3 for E2 and DDP_1 for the station pair kgni-ksmv in category $k3$. Station kgni is the central station for E2 and the distance between the stations is 109 km. In Figures 3a and 3b, the CC for the station pair is always very high all through the earthquake and disturbed days. In Figures 3c and 3e, it is observed that KLD and L2N have high values for 6 days (day 10) to 2 days (day 14) before the earthquake. KLD and L2N values of disturbed days are smaller than those of earthquake days as shown in Figures 3d and 3f, respectively. Thus, KLD and L2N are better indicators of approaching seismic disturbance than CC.

[21] In the group II analysis, an average distribution of TEC is obtained from the days in the chosen quiet-day period for each station as in equation (3). The VTEC data for earthquake-, disturbed-, and quiet-day periods are compared with AQDT using the cross-correlation coefficient $r(\hat{\mathbf{P}}_{u;d}; \hat{\mathbf{P}}_{u;di-ds})$, symmetric Kullback-Leibler distance $KLD(\hat{\mathbf{P}}_{u;d}; \hat{\mathbf{P}}_{u;di-ds})$, and L2 norm $L2N(\hat{\mathbf{P}}_{u;d} \setminus \hat{\mathbf{P}}_{u;di-ds})$. It is observed that if the distance of the station to the epicenter is less than 150 km, KLD and L2N values of this station for earthquake days are greater than those of quiet and disturbed days. For stations that are more than 150 km away from the epicenter, KLD and L2N values of earthquake and quiet days are similar to each other. It is observed that KLD and L2N values computed between the disturbed-day TEC and AQDT cannot be differentiated between disturbed- and quiet-day periods. The correlation coefficients between AQDT and earthquake-, disturbed-, and quiet-day periods vary between +0.2 and +0.7 for any station regardless of the distance to the epicenter. An exception to this observation is provided in Figure 4 for E6 and station kunm for the 12 May 2008 earthquake. Although the distance of kunm is more than 600 km to the epicenter, the CC, KLD, and L2N values for 2 days before the earthquake (day 14) are significantly different than those of the other days in E6, as shown in Figures 4a, 4d, and 4g, respectively. The comparison of AQDT values with those of DDP_1 and QDP_2 using CC, KLD, and L2N are provided in Figures 4b, 4c, 4e, 4f, 4h, and 4i, respectively. The correlation coefficients are close to +1 in quiet- and disturbed-day periods. The averages of KLD and L2N values for DDP_1 and QDP_2 are similar to each other.

[22] In group II, the bounds defined in equation (13) are computed for all days in earthquake-, disturbed-,

Table 5. Values of Δ_{L2N} for All Station Pairs in Distance Categories $k1$ to $k6$ for QDP_1, DDP_1, and DDP_2 and E1, E2, E3, E4, and E5 Earthquakes

Periods	$k1$	$k2$	$k3$	$k4$	$k5$	$k6$
QDP_1	0.0013	0.0013	0.0012	0.00082	0.0023	0.0048
DDP_1, DDP_2	0.0011	0.001	0.001		0.0026	0.0057
E1	0.0083	0.0049	0.0036	0.00082	0.0053	0.0052
E2		0.0035	0.0024	0.00054	0.0029	0.0037
E3			0.009	0.0063	0.0031	
E4	0.0066	0.0026	0.0035	0.00087	0.0026	0.0036
E5	0.011	0.0055	0.0065	0.0018		0.0073

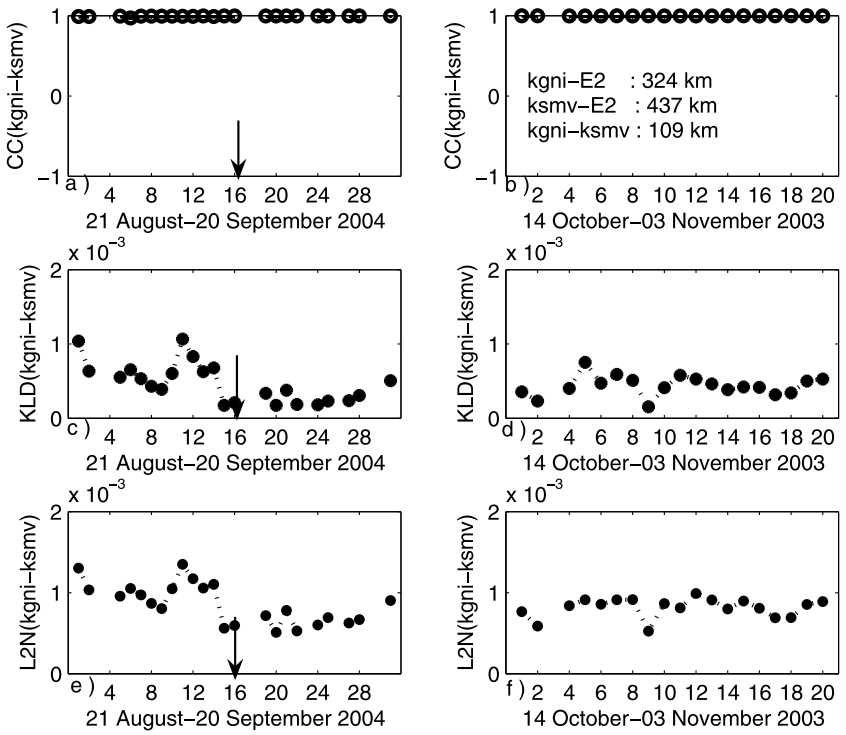

Figure 3. For stations kgni and ksmv, the values of (a) CC for E2, (b) CC for DDP_1, (c) KLD for E2, (d) KLD for DDP_1, (e) L2N for E2, and (f) L2N for DDP_1. The day of the earthquake is indicated by the arrow.

and quiet-day periods. These bounds are compared with the bound of AQDT using the cross-correlation coefficient r_b ($\hat{\mathbf{P}}_{u;d}$; $\hat{\mathbf{P}}_{u;di-ds}$), symmetric Kullback-Leibler distance KLD_b ($\hat{\mathbf{P}}_{u;d}$; $\hat{\mathbf{P}}_{u;di-ds}$), and L2 norm $L2N_b$ ($\hat{\mathbf{P}}_{u;d}\backslash\hat{\mathbf{P}}_{u;di-ds}$). It is observed that when bounds in equation (13) are used in comparisons, all three tools are successful in discriminating the earthquake days from AQDT. Yet disturbed days cannot be differentiated from AQDT. An example is provided in Figure 5 for E6 and station kunm. The CC, KLD, and L2N values between the bounds of AQDT and earthquake, quiet, and disturbed days are given in Figures 5a–5c, respectively. It is observed that all of the CC, KLD, and L2N values 2 days before the earthquake are significantly different from those in DDP_1 and QDP_2.

[23] In group III, the daily VTEC data of each station are compared with data of the consecutive day for all earthquake-, disturbed-, and quiet-day periods. For this purpose, the cross-correlation coefficient $r(\hat{\mathbf{P}}_{u;d}; \hat{\mathbf{P}}_{u;d+1})$, symmetric Kullback-Leibler distance $KLD(\hat{\mathbf{P}}_{u;d}; \hat{\mathbf{P}}_{u;d+1})$, and L2 norm $L2N(\hat{\mathbf{P}}_{u;d}\backslash\hat{\mathbf{P}}_{u;d+1})$ are computed. It is observed that when CC, KLD, and L2N methods are applied for disturbed days, all three methods can differentiate the geomagnetically disturbed days, especially for the storm days when $Kp > 6$. No significant variation between the consecutive days of earthquake- and quiet-day periods is observed for CC, KLD, and L2N values. An example is provided in Figure 6 for E3 and station mizu. Station mizu is the central station for E3 and its distance to the epicenter is 43 km. From Figure 6, it is observed that KLD and L2N values for consecutive earthquake and quiet days are very similar to each other. Correlation coefficients for those cases are very high. Yet, for DDP_2, especially in the storm days (2nd and 18th days with $Kp \geq 6$), KLD and L2N values increase significantly and CC values are less than 0.9.

[24] In group III, the bounds defined in equation (13) are computed for all days in earthquake-, disturbed-, and quiet-day periods. These bounds are compared with the bounds of consecutive days using the cross-correlation coefficient r_b ($\hat{\mathbf{P}}_{u;d}$; $\hat{\mathbf{P}}_{u;d+1}$), symmetric Kullback-Leibler distance KLD_b ($\hat{\mathbf{P}}_{u;d}$; $\hat{\mathbf{P}}_{u;d+1}$), and L2 norm $L2N_b$ ($\hat{\mathbf{P}}_{u;d}\backslash\hat{\mathbf{P}}_{u;d+1}$). It is observed that KLD and L2N methods can measure ionospheric disturbance, but they cannot differentiate the seismic disturbances from the geomagnetic ones. When the correlation coefficients of the bounds for consecutive days are computed, it is observed that the CC for disturbed-day periods are consistently lower from those of earthquake- and quiet-day periods. Thus, CC analysis between the bounds of

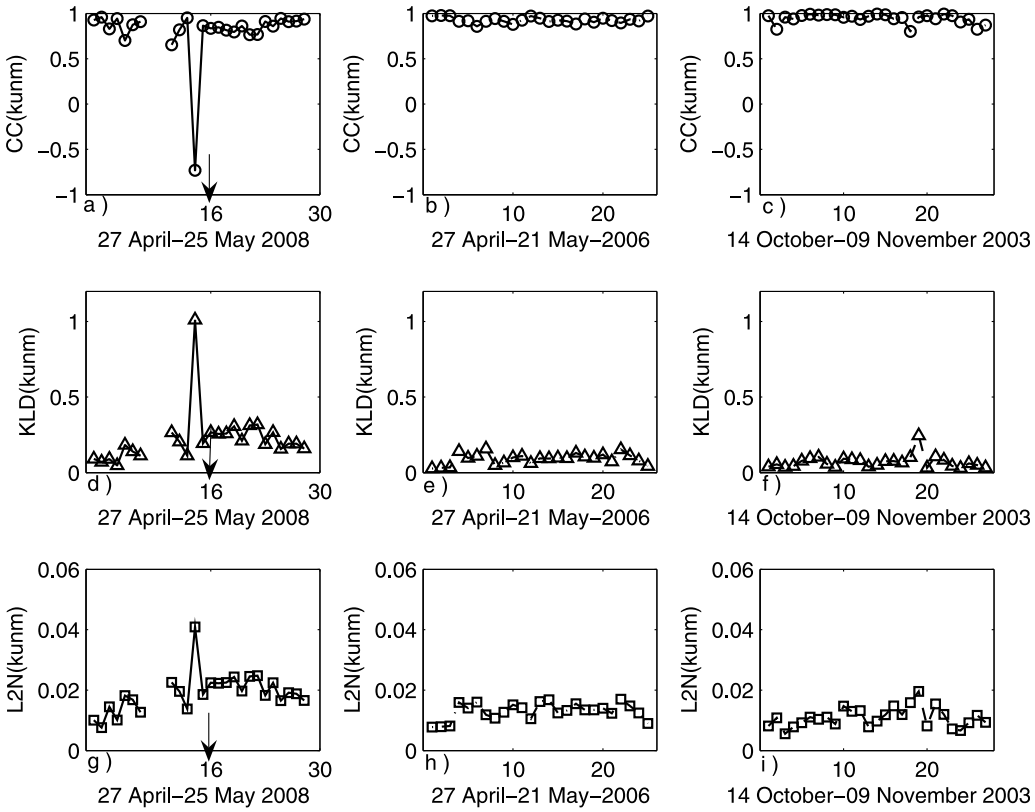

Figure 4. For station kunm, CC values between (a) EDP for E6 and AQDT, (b) QDP_2 and AQDT, and (c) DDP_1 and AQDT; KLD values between (d) EDP for E6 and AQDT, (e) QDP_2 and AQDT, and (f) DDP_1 and AQDT; and L2N values between (g) EDP for E6 and AQDT, (h) QDP_2 and AQDT, and (i) DDP_1 and AQDT. The earthquake day is indicated by the arrow in Figures 4a, 4d, and 4g.

consecutive days can be considered a precursor for geomagnetic disturbance in the ionosphere. An example case is provided in Figure 7 for E3 and station mizu. The CC, KLD, and L2N values between bounds of the consecutive days for E3, DDP_2, and QDP_1 are presented in Figures 7a–7c, respectively. KLD and L2N values in Figures 7b and 7c for the disturbed and earthquake days are greater than those for the quiet days. Yet, in Figure 7a, CC values for DDP_2 are significantly less than those of earthquake days and QDP_1.

[25] In this section, for all six earthquakes, eight GPS stations, two geomagnetic storms, and two quiet-day periods, the comparisons are computed using the cross-correlation coefficient, symmetric Kullback-Leibler distance, and L2 norm between daily VTEC values and sliding-window estimated bounds *b*. To differentiate the ionospheric variability, the statistical tools are applied to both quiet days and an average quiet day for any chosen GPS station. It is observed that CC, KLD, and L2N methods can be used separately or in combination in discriminating disturbed days from quiet days. For certain cases, these methods can even distinguish the type of disturbance. As a result, with further investigation, these methods can be developed into precursors for ionospheric disturbance.

4. Conclusion

[26] In this study, the coupling of seismic and geomagnetic activity to the ionosphere is investigated through the variability of GPS-TEC by using four statistical tools, namely, the cross-correlation coefficient, the symmetric Kullback-Leibler distance, the L2 norm, and sliding-window statistical analysis. Six earthquakes with different seismic properties and two severe geomagnetic disturbances are chosen for investigation in this study. IONOLAB-TEC is computed for each of 15 days before and after each earthquake, geomagnetically disturbed days, and quiet days for eight GPS stations in the regions of interest. For all distance categories (group I),

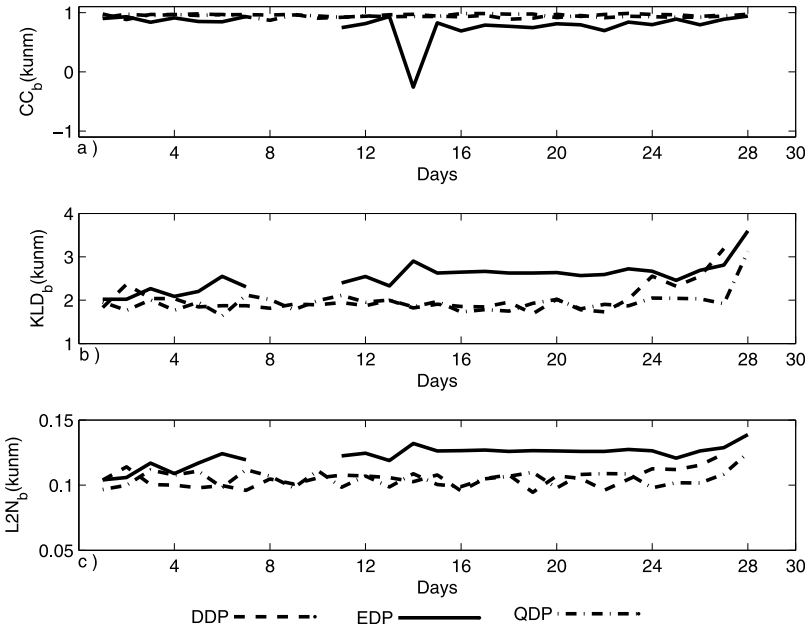

Figure 5. For station kunm, (a) CC, (b) KLD, and (c) L2N values between the bound of AQDT and bounds of EDP for E6, QDP_2, and DDP_1.

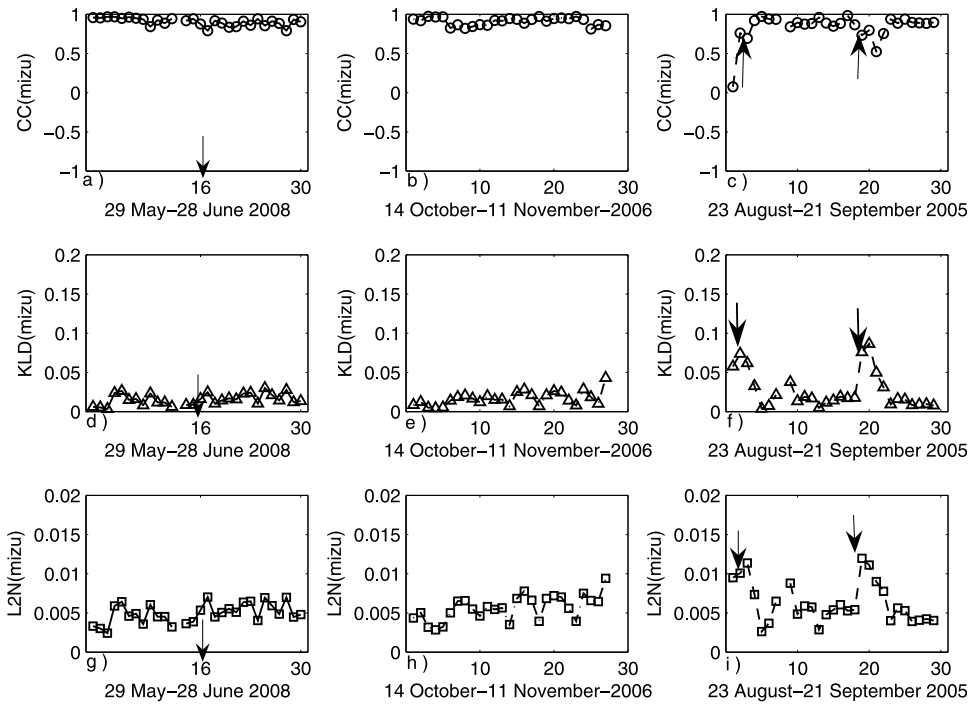

Figure 6. Comparison for consecutive days of station mizu: CC values in (a) EDP for E3, (b) QDP_1, and (c) DDP_2; KLD values in (d) EDP for E3, (e) QDP_1, and (f) DDP_2; and L2N values in (g) EDP for E3, (h) QDP_1, and (i) DDP_2. The earthquake day is indicated by the arrow in Figures 6a, 6d, and 6g, and the days of the geomagnetic storm are indicated by the arrows in Figures 6c, 6f, and 6i.

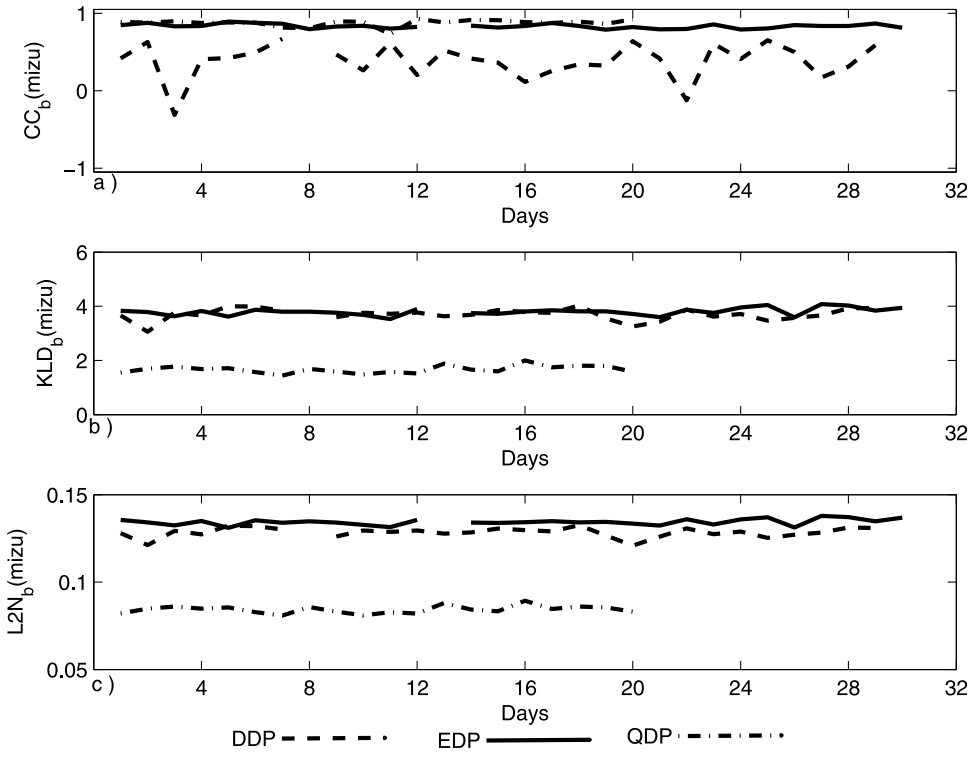

Figure 7. For station mizu, (a) CC, (b) KLD, and (c) L2N values between the bounds of consecutive days of EDP for E3, QDP_1, and DDP_2.

AQDT comparisons (group II), and consecutive-day comparisons (group III), more than 9500 values are computed and sorted according to the magnitude of the earthquake, the distance between the stations, the distance between the stations and epicenters, and the depth of earthquakes and periods of quiet and disturbed days. It is observed that the CC between the neighboring GPS stations cannot be used as a definitive earthquake precursor. KLD and L2N between the neighboring stations can be used to distinguish the earthquake days when the distance between the stations is less than 340 km and also when selected station pairs are close to the earthquake epicenter. For the comparison between AQDT and EDP, it is observed that KLD and L2N are strong candidates for developing an earthquake precursor tool for the stations that are located less than 150 km from the earthquake zones. In the comparison of the consecutive days for each station, it is observed that CC, KLD, and L2N can all distinguish geomagnetic disturbance from the seismic disturbance. This study demonstrates that CC, KLD, and L2N can be developed into precursor tools for distinguishing geomagnetic and seismic activity. Yet further long-term analysis is necessary for these tools to produce definitive precursor signals for those GPS stations that are on the earthquake zones. Also, for more reliable estimates of preseismic activity, joint space-time analysis of TEC is necessary over a denser GPS network in the earthquake zones.

[27] **Acknowledgment.** This study is supported by TUBITAK EEEAG grants 105E171 and 109E055.

References

Arikan, F., and C. B. Erol (1998), Statistical characterization of time variability in midlatitude single tone HF channel response, *Radio Sci.*, *33*(5), 1429–1444.

Arikan, F., C. B. Erol, and O. Arikan (2003), Regularized estimation of vertical total electron content from Global Positioning System data, *J. Geophys. Res.*, *108*(A12), 1469, doi:10.1029/2002JA009605.

Arikan, F., C. B. Erol, and O. Arikan (2004), Regularized estimation of vertical total electron content from GPS data for a desired time period, *Radio Sci.*, *39*, RS6012, doi:10.1029/2004RS003061.

Arikan, F., S. Karatay, and O. Arikan (2009), Investigation of ionospheric disturbance due to strong earthquakes using total electron content, *Geophys. Res. Abstr.*, *11*, Abstract EGU2009-8440.

Biqiang, Z., W. Weixing, L. Libo, and M. Tian (2007), Morphology in the total electron content under geomagnetic disturbed conditions: Results from global ionosphere naps, *Ann. Geophys.*, *25*(7), 1555–1568.

Chuo, Y. J., Y. I. Chen, J. Y. Liu, and S. A. Pulinets (2001), Ionospheric f0F2 variations prior to strong earthquakes in Taiwan area, *Adv. Space Res.*, *27*(6), 1305–1310.

Ciraolo, L., and P. Spalla (1999), A statistics of time and space variability of ionospheric electron content at middle latitudes, paper presented at 4th COST 251 Workshop, Madeira, Portugal, 22–25 March.

Cover, T. M., and A. J. Thomas (2006), *Elements of Information Theory*, Wiley-Interscience, New York.

Erol, C. B., and F. Arikan (2005), Statistical characterization of the ionosphere using GPS signals, *J. Electromagn. Waves Appl.*, *19*(3), 373–387.

Hall, P. (1987), On Kullback-Leibler loss and estimation, *Ann. Stat.*, *15*(4), 1491–1519.

Inglada, J. (2003), Change detection on SAR images by using a parametric estimation of the Kullback-Leibler divergence, in *IGARSS 2003: Learning from Earth's Shapes and Sizes*, vol. 6, pp. 4104–4106, Int. of Electr. and Electr. Eng., Piscataway, N. J.

Karatay, S., F. Arikan, and O. Arikan (2009a), Investigation of lithosphere-ionosphere coupling using total electron content, paper presented at SIU-2009 Symposium of Signal Processing and Applications, Int. of Electr. and Electr. Eng., Antalya, Turkey, 9–11 April.

Karatay, S., F. Arikan, and O. Arikan (2009b), Investigation of hourly and daily patterns for lithosphere-ionosphere coupling before strong earthquakes, in *4th International Conference on Recent Advances in Space Technologies, 2009*, pp. 670–674, Inst. of Electr. and Electr. Eng., New York.

Kouris, S. S., and D. N. Fotiadis (2002), Ionospheric variability: A comparative statistical study, *Adv. Space Res.*, *29*(6), 977–985.

Kouris, S. S., K. V. Polimeris, and L. R. Cander (2006), Specifications of TEC variability, *Adv. Space Res.*, *37*(5), 983–1004.

Kreyszig, E. (1988), *Advanced Engineering Mathematics*, John Wiley, New York.

Lazo, B., K. Alazo, M. Rodriguez, and A. Calzadilla (2004), TEC variability over Havana for different solar activity conditions, *Adv. Space Res.*, *34*(9), 2044–2048.

Liu, J. Y., Y. I. Chen, S. A. Pulinets, Y. B. Tsai, and Y. J. Chuo (2000), Seismo-ionospheric signatures prior to $M \geq 6.0$ Taiwan earthquakes, *Geophys. Res. Lett.*, *27*(19), 3113–3116.

Liu, J. Y., Y. J. Chuo, S. J. Shan, Y. B. Tsai, Y. I. Chen, S. A. Pulinets, and S. B. Yu (2004), Pre-earthquake ionospheric anomalies registered by continuous GPS TEC measurements, *Ann. Geophys.*, *22*(5), 1585–1593.

Nayir, H., F. Arikan, O. Arikan, and C. B. Erol (2007), Total electron content estimation with Reg-Est, *J. Geophys. Res.*, *112*, A11313, doi:10.1029/2007JA012459.

Ondoh, T. (2000), Seismo-ionospheric phenomena, *Adv. Space Res.*, *26*(8), 1267–1272.

Plotkin, V. V. (2003), GPS detection of ionospheric perturbations before the 13 February 2001 El Salvador earthquake, *Nat. Hazards Earth Syst. Sci.*, *3*, 249–253.

Pulinets, S. A. (2004), Ionospheric precursors of earthquakes: Recent advances in theory and practical applications, *Terr. Atmos. Ocean Sci.*, *15*(3), 413–435.

Pulinets, S. A., T. B. Gaivoronska, L. A. Contreras, and L. Ciraolo (2004), Correlation analysis technique revealing ionospheric precursors of earthquake, *Nat. Hazards Earth Syst. Sci.*, *4*, 697–702.

Pulinets, S. A., A. L. Contreas, G. Bisiacchi-Giraldi, and L. Ciraolo (2005), Total electron content variations in the ionosphere before the Colima, Mexico, earthquake of 21 January 2003, *Geofis. Int.*, *44*(4), 369–377.

Pulinets, S. A., A. N. Kotsarenko, L. Ciraolo, and I. A. Pulinets (2007), Special case of ionospheric day-to-day variability associated with earthquake preparation, *Adv. Space Res.*, *39*(5), 970–977.

Rishbeth, H., and O. K. Garriot (1969), *Introduction to Ionospheric Physics*, Academic, New York.

Trigunait, A., M. Parrot, S. A. Pulinets, and F. Li (2004), Variations of the ionospheric electron density during the Bhuj seismic event, *Ann. Geophys.*, *22*(12), 4123–4131.

Vlasov, M., M. C. Kelley, and H. Kil (2003), Analysis of ground-based and satellite observations of F-region behavior during the great magnetic storm of July 15, 2000, *J. Atmos. Sol. Terr. Phys.*, *65*, 1223–1234.

Zhang, D., and Z. Xiao (2000), Study of ionospheric TEC using GPS during the large solar flare burst on November 6, 1997, *Chin. Sci. Bull.*, *45*(19), 1749–1752.

Zhang, M. L., J. K. Shi, X. Wang, and S. M. Radicella (2004), Ionospheric variability at low latitude station: Hainan, China, *Adv. Space Res.*, *34*(9), 1860–1868.

F. Arikan, Department of Electrical and Electronics Engineering, Hacettepe University, Beytepe, 06800 Ankara, Turkey. (arikan@hacettepe.edu.tr)

O. Arikan, Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, 06800 Ankara, Turkey. (oarikan@ee.bilkent.edu.tr)

S. Karatay, Department of Physics, Faculty of Arts and Sciences, Firat University, 23100 Elazig, Turkey. (karatays@gmail.com)