Investigation of ionospheric electron content variations before earthquakes in southern California, 2003–2004

Thomas Dautermann,[1] Eric Calais,[2] Jennifer Haase,[2] and James Garrison[3]

Received 12 April 2006; revised 26 July 2006; accepted 18 September 2006; published 15 February 2007.

[1] It has been proposed that earthquakes are preceded by electromagnetic signals detectable from ground- and space-based measurements. Ionospheric anomalies, such as variations in the electron density a few days before earthquakes, are one of the precursory signals proposed. Since Global Positioning System (GPS) data can be used to measure the ionospheric total electron content (TEC), the technique has received attention as a potential tool to detect ionospheric perturbations related to earthquakes. Here, we analyze 2 years (2003–2004) of data from the Southern California Integrated GPS Network (SCIGN), a dense network of 265 continuous GPS stations centered on the Los Angeles basin, for possible precursors. This time period encompasses the December 2003, M6.6, San Simeon and September 2004, M6.0, Parkfield earthquakes. We produce TEC time series at all SCIGN sites and apply three different statistical tests to detect anomalous TEC signals preceding earthquakes. We find anomalous TEC signals but no statistically significant correlation, in time or in space, between these TEC anomalies and the occurrence of earthquakes in southern California for the 2003–2004 period. This result does not disprove the possibility of precursory phenomena but show the signal-to-noise ratio of a hypothetical TEC precursor signature is too low to be detected by the analysis techniques employed here. Precursors may still be revealed for future large earthquakes in well instrumented areas such as California and Japan, if the tests can be developed into techniques that can better separate external influences from the actual TEC signal.

Citation: Dautermann, T., E. Calais, J. Haase, and J. Garrison (2007), Investigation of ionospheric electron content variations before earthquakes in southern California, 2003–2004, *J. Geophys. Res.*, *112*, B02106, doi:10.1029/2006JB004447.

1. Introduction

[2] It has long been proposed that earthquakes are preceded by electromagnetic signals in the ULF, ELF, and VLF bands, detectable by ground or space-based measurements [e.g., *Gokhberg et al.*, 1982; *Tate and Daily*, 1989; *Fraser-Smith et al.*, 1990; *Parrot and Lefeuvre*, 1985]. Some of these signals are thought to propagate to the ionosphere [e.g., *Pulinets*, 2004] and trigger variations of the ionospheric critical plasma frequency (*foF*2) and of the ionospheric Total Electron Content (TEC). For instance, phase shifts in VLF and ULF signals from Omega transmitters across Japan have been observed a few days before the Kobe earthquake [*Hayakawa et al.*, 1996, 1999, 2000]. *Liu et al.* [2000, 2001] report an *foF*2 and TEC decrease 1 to 6 days prior to the 1999, Mw = 7.7, Chi-Chi earthquake in Taiwan. More recently, *Pulinets* [2004] also reported a TEC decrease the day preceding the December 2003, M6.6, San Simeon earthquake in central California. Measuring TEC can now be done straightforwardly using Global Positioning System (GPS) data [e.g., *Manucci et al.*, 1993; *Calais and Minster*, 1995]. Confirming earlier reports of TEC perturbations precursory to earthquakes using independent statistical tests would have a significant impact on earthquake forecasting and may open the way to a better understanding of the physical processes that lead to earthquakes.

[3] The primary difficulty, however, is the lack of a consistent physical framework to explain electromagnetic earthquake precursors as well as their propagation to the ionosphere. Often quoted processes for ground electromagnetic emissions include stress changes in rocks through piezoelectric and piezomagnetic effects [*Finkelstein et al.*, 1973; *Bishop*, 1981], induction effects due to motion of electric charges in the geomagnetic field (e.g., electrokinetic effects [*Morrison et al.*, 1989; *Lonre et al.*, 1999; *Revil et al.*, 1999]), the emission of radioactive gas or metallic ions, or positive hole defects in minerals [*Freund*, 2000]. The mechanism that may link the electromagnetic signal at ground level with ionospheric anomalies remains poorly understood as well, although some models have been

[1]Department of Physics, Purdue University, West Lafayette, Indiana, USA.
[2]Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana, USA.
[3]School of Aeronautics and Astronautics, Purdue University, West Lafayette, Indiana, USA.

0148-0227/07/2006JB004447$09.00

proposed that involve direct coupling from the Earth's surface to the lower ionosphere through electric currents [*Pulinets et al.*, 2000; *Sorokin et al.*, 2001] or the triggering of atmospheric gravity waves by gas releases or thermal anomalies [*Molchanov and Hayakawa*, 2001; *Shvets et al.*, 2004]. Because the physics of the processes which could lead to a precursor signal is unclear, the theoretical structure of the expected signals remains unknown and hence cannot be used to separate these signals from those generated by processes other than seismic activity. At present, we are left with studying precursor signals empirically.

[4] A second difficulty resides in demonstrating a causal link between a signal identified as anomalous and a given earthquake. Doing so requires the ability to quantify the background variability of the ionosphere (i.e., in the absence of earthquake-related signals) and to show, statistically, a temporal and spatial correlation between anomalous signals and earthquakes. Many studies have included only short time intervals before a known event, which makes it difficult to unambiguously address the statistical significance of their results. Some studies have however used a statistical approach based on long data time series [e.g., *Larkina et al.*, 1989; *Parrot and Lefeuvre*, 1985; *Serebryakova et al.*, 1992]. For instance, using 3.5 years of TEC data from the Topex-Poseidon altimetry satellite, *Zaslavski et al.* [1998] found significant variations compared to background variability within the 48 hours preceding earthquakes in 34% of the cases studied within 300 km of the epicenter, whereas a random distribution would only lead to a 16% correlation. More recently, *Fujiwara et al.* [2004] used 1.5 years of VHF data between two Japanese stations 294 km apart to argue that anomalies in the data are systematically enhanced within 5 days of M > 4.8 earthquakes. *Henderson et al.* [1993], however, using a large data set of ELF/VLF measurements from low-altitude satellites, argue for no distinguishable ELF/VLF signals associated with earthquakes.

[5] An additional obstacle in detecting earthquake precursors is that the chances that an adequate sensor exists and operates at the right time and place, i.e., near the epicenter of a large imminent earthquake, are fairly small. Repeating a successful precursor detection is even more difficult. For instance, the quasi-static electromagnetic energy building up prior to the 1989 Loma Prieta earthquake observed by *Fraser-Smith et al.* [1990] has not yet been repeated, in spite of a number of moderate to large earthquakes in California since that date. This may simply reflect the spatial complexity of precursory signals, making it difficult to detect them with only a few sensors not optimally located. Indeed, if some studies have argued for a statistically significant temporal correlation between earthquakes and preceding ground or space signals, very few had the adequate spatial resolution to unambiguously test for spatial patterns. Given the large number and high spatial density of GPS stations in seismically active regions such as California or Japan, GPS-derived TEC measurements may help address this problem.

[6] This study is based on continuous data from the Southern California Integrated GPS Network (SCIGN), a densely spaced GPS array of 265 stations centered on the Los Angeles basin. We use the SCIGN data to compute TEC time series for the entire network for years 2003 and 2004, that span the San Simeon (22 December 2003, M6.6) and Parkfield (29 September 2004, M6.0) earthquakes. We then search the TEC time series for anomalous TEC signals using three different methods and try to correlate TEC anomalies with the occurrence and location of the earthquakes.

2. From GPS Data to Total Electron Content

[7] The space segment of the GPS system consists of a constellation of a minimum of 24 satellites in six circular orbits 20,200 km above the Earth (see, e.g., *Dixon* [1991] for a complete overview of the Global Positioning System). Each plane is filled with four to six satellites, spaced for an optimum visibility of at least four satellites from any point on Earth. Each satellite continuously broadcasts a pseudo-random noise (PRN) code modulated radio signal in the L-band on two carrier-frequencies, f_1 = 1.57542 GHz and f_2 = 1.2276 GHz. GPS receivers track the coded information and generate two main types of observables: pseudoranges, calculated by multiplying the (uncorrected) travel time from satellite to receiver by the speed of light, and carrier phases ϕ_1 and ϕ_2 on both frequencies.

[8] The propagation delay experienced by the GPS signals in the ionosphere represents one of the most important error sources that affect GPS positioning. However, since the ionosphere is a dispersive medium, the differential delay between f_1 and f_2 can be used to mitigate and/or quantify that effect. Here we take advantage of that property to calculate the integral of the electron density (also called integrated electron content, or IEC) along the ray path from a satellite to a ground receiver from dual frequency GPS pseudorange and phase data. The IEC is derived from the phase data using:

$$IEC = \left(\underbrace{\phi_2 - \frac{f_2}{f_1}\phi_1}_{L_G} + f_2(t_{IFB} - t_{TGD}) + N \right) \times \frac{c f_1^2 f_2}{A\left(f_1^2 - f_2^2\right)} \qquad (1)$$

where A = 40.3 m^3/s^2 and c is the speed of light [e.g., *Klobuchar*, 1985]. IEC measurements from GPS phase observables are ambiguous because of integer cycle ambiguities N, constant for a given observation arc (if no cycle slips occur), and uncalibrated code offsets (transmitter group delay t_{TGD} and interfrequency bias t_{IFB} [e.g., *Manucci et al.*, 1993]). The transmitter group delay is specific to each satellite and is given in the navigation message, but the interfrequency bias must be estimated from the GPS data using inversion techniques.

[9] The IEC is typically expressed in units of 10^{16} m^{-2} = 1 TECU. To estimate N, the IEC can be computed similarly from dual-frequency pseudorange observations ρ_1 and ρ_2 as for instance described by *Calais and Minster* [1995]; we call the unbiased observable $L_{GU} = L_G + N$. The equation above gives the electron content integrated over a line-of-sight from a ground receiver to a satellite. It can be converted to a vertical TEC (sometimes called vertical electron content (VEC)), using an elevation mapping func-

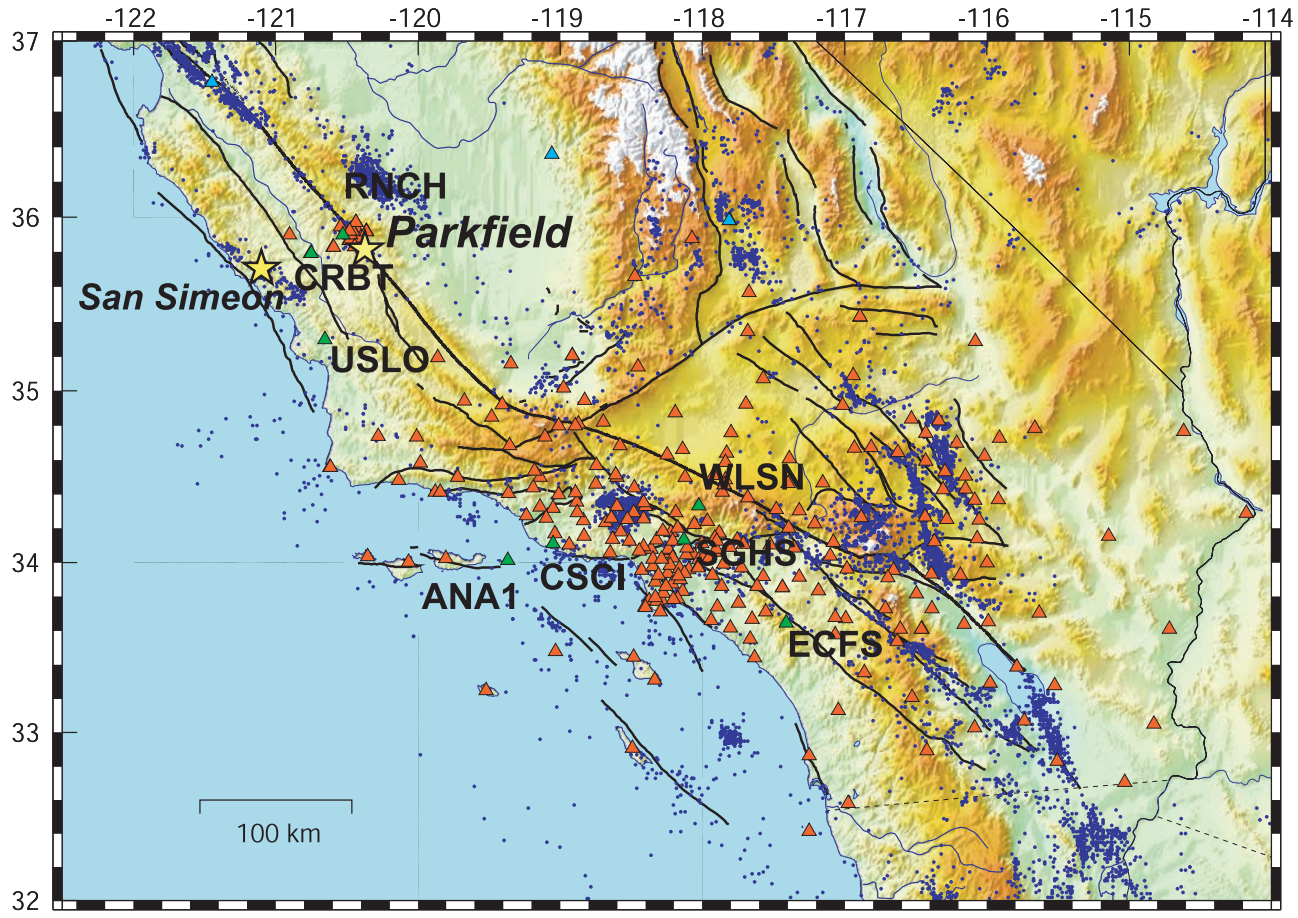

Figure 1. Map of continuous GPS stations in southern California (triangles). Black lines show major active faults. Dots show seismicity (1973–present, NEIC catalog). The epicenters of the 2003 San Simeon and 2004 Parkfield earthquakes are shown with yellow stars, sites mentioned in the text by name are indicated by green triangles.

tion E_Θ, which accounts for the different ray path length through the atmosphere as the satellite elevation angle Θ varies:

$$E_\Theta = \sqrt{1 - \left(\frac{\sin\left(\frac{\pi}{2} - \Theta\right) R_E}{R_E + h_{ion}}\right)^2} \tag{2}$$

where h_{ion} is the mean ionospheric height, taken equal to 350 km, close to the height of the maximum electron density. Rearranging (1) and substituting $TEC = IEC \times E_\Theta$ provides an expression for the (corrected) absolute L_{GU} as a function of TEC and receiver bias (t_{IFB}) as:

$$L_{GU} - f_2 t_{TGD} = \frac{K}{E_\Theta} TEC - f_2 t_{IFB} \tag{3}$$

with:

$$K = \frac{A\left(f_1^2 - f_2^2\right)}{c f_1^2 f_2} \tag{4}$$

Equation (3) provides a linear relationship between the observations L_{GU} and the unknowns TEC and t_{IFB}, for each site. This set of equations is overdetermined at any given measurement epoch, if more than one satellite is visible, and can therefore be solved using a least-squares technique.

3. GPS Data Processing

[10] We used the method described above to compute TEC time series at all continuous GPS sites in southern California for years 2003 and 2004 (Figure 1). In order to lower the computational burden, we decimated the data, originally sampled at 30 s, to 90 s. Figure 2 shows the raw TEC time series averaged over the entire network and the corresponding power spectrum for the period considered in this study. As expected, the TEC is dominated by diurnal (24 hours) and semidiurnal (12 hours) periods reflecting daily solar activity variations and lunar tides [e.g., *Nicholson and Steiger*, 1963]. We also find significant power at a 27 day period, also seen on *Dst* and solar flux time series (see below, Figure 3), that coincides with the equatorial rotation rate of the Sun. Finally, the TEC time series also shows an annual fluctuation reflecting seasonal variations in solar activity.

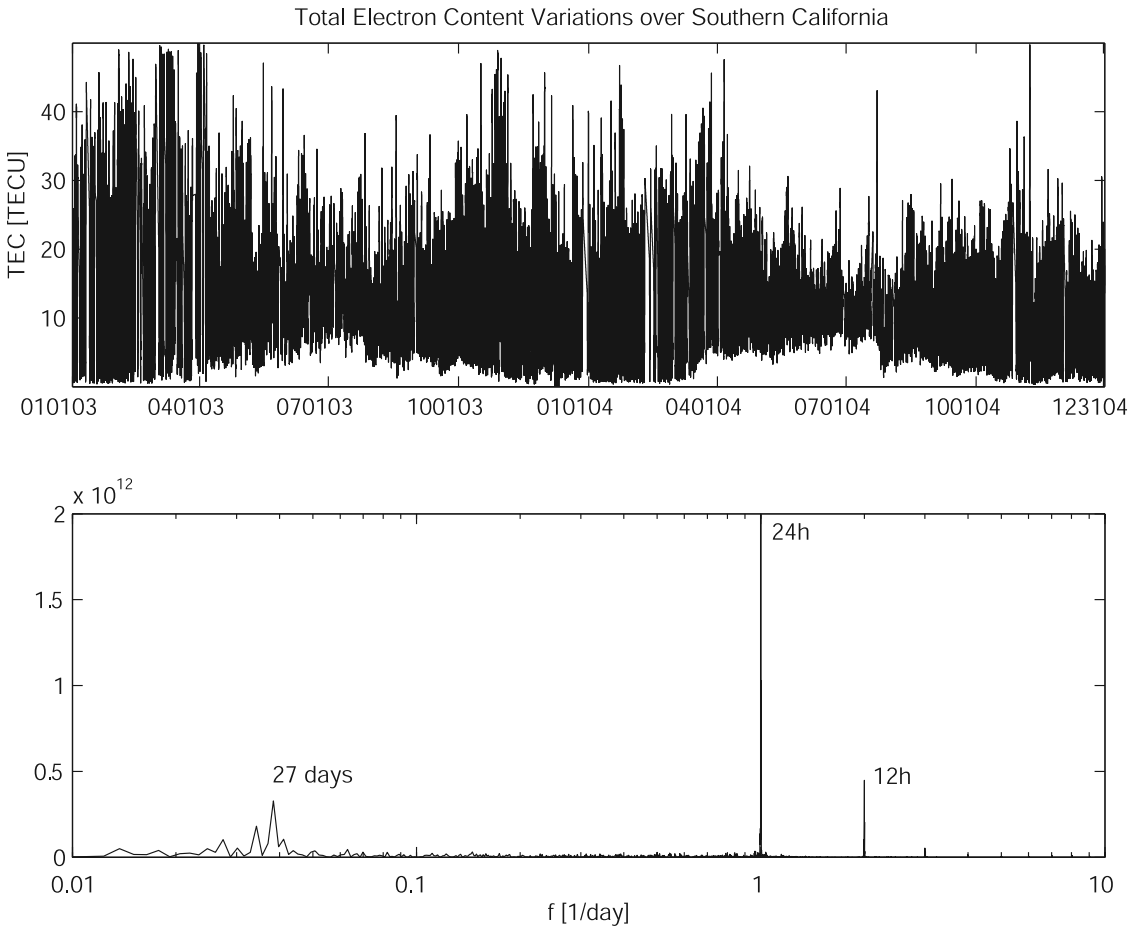

Figure 2. (top) Stacked TEC time series for 2003 and 2004 in Southern California. (bottom) Corresponding spectrum, periodicity of selected peaks is indicated.

[11] Since ionospheric perturbations are correlated with solar activity and magnetic storms, we extracted the magnetic disturbance storm time index (*Dst*) and solar activity index from the U.S. National Geophysical Data Center, Space Physics Interactive Data Resource (http://spidr.ngdc.noaa.gov). The *Dst* measures the magnitude of the horizontal component of the Earth's magnetic field at ground level in nT using hourly readings from midlatitude and equatorial stations. The solar activity is measured in terms of solar flux in the 10.7 cm wavelength in units of 10^{-22} W/m^{2}Hz at a radio telescope near Ottawa, Canada. Figure 3 shows the *Dst* and solar flux times series for 2003 and 2004. Note that both the San Simeon and Parkfield events are preceded by periods of large fluctuations in the solar flux.

[12] Once the TEC time series are obtained, objective procedures and criteria must be established to detect TEC anomalies linked to earthquakes. Since anomalous TEC signals may result from a number of causes besides earthquakes, and since the structure of the expected precursory signals is not established, one must rely on a statistical analysis in order to test for a causal link between anomalous signals and seismic events. In the following, we will base our interpretation on several criteria. First, the possibility that the anomaly may have an extra-terrestrial source must be ruled out by showing a lack of correlation with high *Dst* or solar flux index. Second, the anomaly must be detected at a significant number of GPS stations and be spatially coherent at stations located close to each other. Third, we expect the anomaly to show a spatial pattern correlated with the epicentral area, with, in general, larger amplitudes closer to the epicenter. Fourth, similar signals should be absent in the absence of significant earthquakes. In the following, we use three different methods to detect anomalous TEC signals from the SCIGN GPS stations in southern California for 2003 and 2004 and apply simple tests based on the criteria listed above to assert the significance level of these detections.

4. Envelope Method

4.1. Description

[13] This method was used by *Liu et al.* [2000] and *Pulinets et al.* [2004] to identify precursors to the 1999 Chi Chi earthquake in Taiwan and 2003 San Simeon earthquake in California from ionospheric measurements. The natural variability of the TEC at each station is obtained by averaging several days of TEC measurements at each measurement epoch (e.g., each 90 s) and calculating the associated standard deviation. The averaging window must be long enough to efficiently average out signals that do not

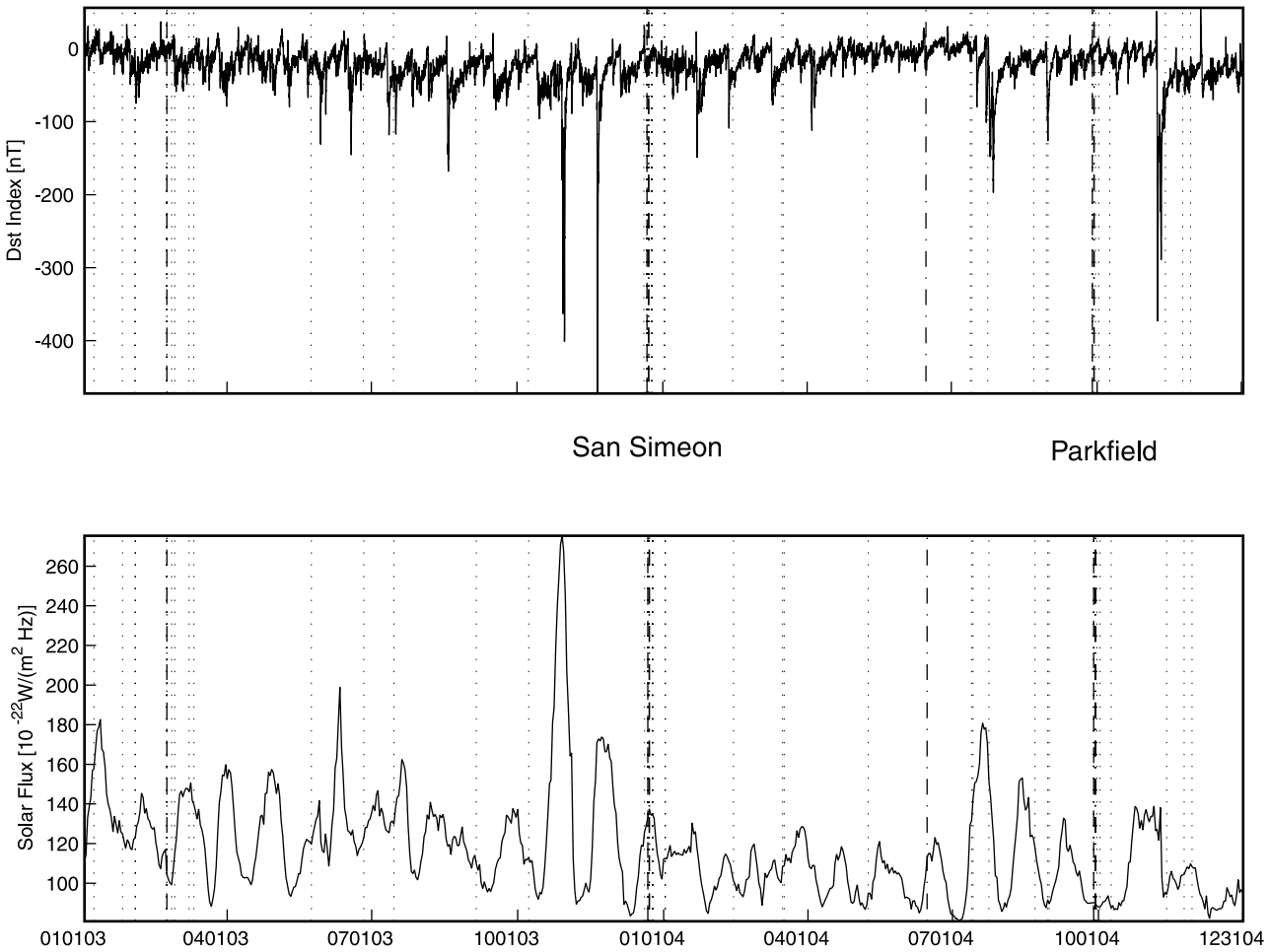

Figure 3. Solar flux and *Dst* index for 2003 and 2004. Earthquakes larger than M = 3 (limited to the geographic area of the SCIGN, from 32°N to 36°N and 123°W to 115°W, http://neic.usgs.gov) are marked by vertical lines: magnitudes 6 to 7 with dashed lines, magnitudes 5 to 6 with dash-dotted lines, magnitudes 4 to 5 with dotted lines.

repeat daily. However, its length is limited by the natural variability of the ionosphere at longer periods, for instance at seasonal timescales. Here we used a sliding window average that includes TEC values measured at the same time of each day during the 10 preceding, the day considered and the 10 following days. The envelope is calculated as:

$$TEC_E(t) = \frac{1}{21}\left(\sum_{k=-10}^{k=10} TEC(t-k)\right) \pm 2\sigma_E \qquad (5)$$

where t is the time (in fractions of day of the year, e.g., 100.001, 100.002, ...) and k runs over the days in the sliding window. A detection is counted each time the observed TEC measurement exceeds the envelope in positive or negative direction.

4.2. Time Series

[14] Figure 4 shows the TEC time series along with its 2 standard deviations upper and lower bounds (as defined above) at four GPS sites for a period of 10 days centered on the time of the 22 December 2003, M6.6, San Simeon earthquake. These four sites are chosen at distances increasing by 250 km intervals from the epicenter. As reported by *Pulinets et al.* [2004], we do find a detection on the day of the San Simeon earthquake and the day before (crosses on Figure 4). Note that we also find detections on the day following the earthquake. We find similar detections at 65% of the SCIGN stations for 21 December, 77% of the stations for 22 December, and 49% of the stations on 23 December.

[15] As shown on Figure 3, the time period of the San Simeon earthquake corresponds to a 10-day period of high solar flux, the third in the last 2 months of 2003. For comparison, we processed GPS data from additional continuous GPS stations across North America at the latitude of southern California (Flagstaff, AZ; Summerfield, TX; French Bayoux, AR; Castle Hayne, NC) for the same time period (Figure 5). As for the SCIGN stations, we find detections on the day preceding the San Simeon earthquake, although these stations are located between 842 and 3906 km from the earthquake epicenter. The fact that the detections observed are correlated with a period of high solar flux and are seen all across North America suggests no causal link with the San Simeon earthquake. A similar analysis for

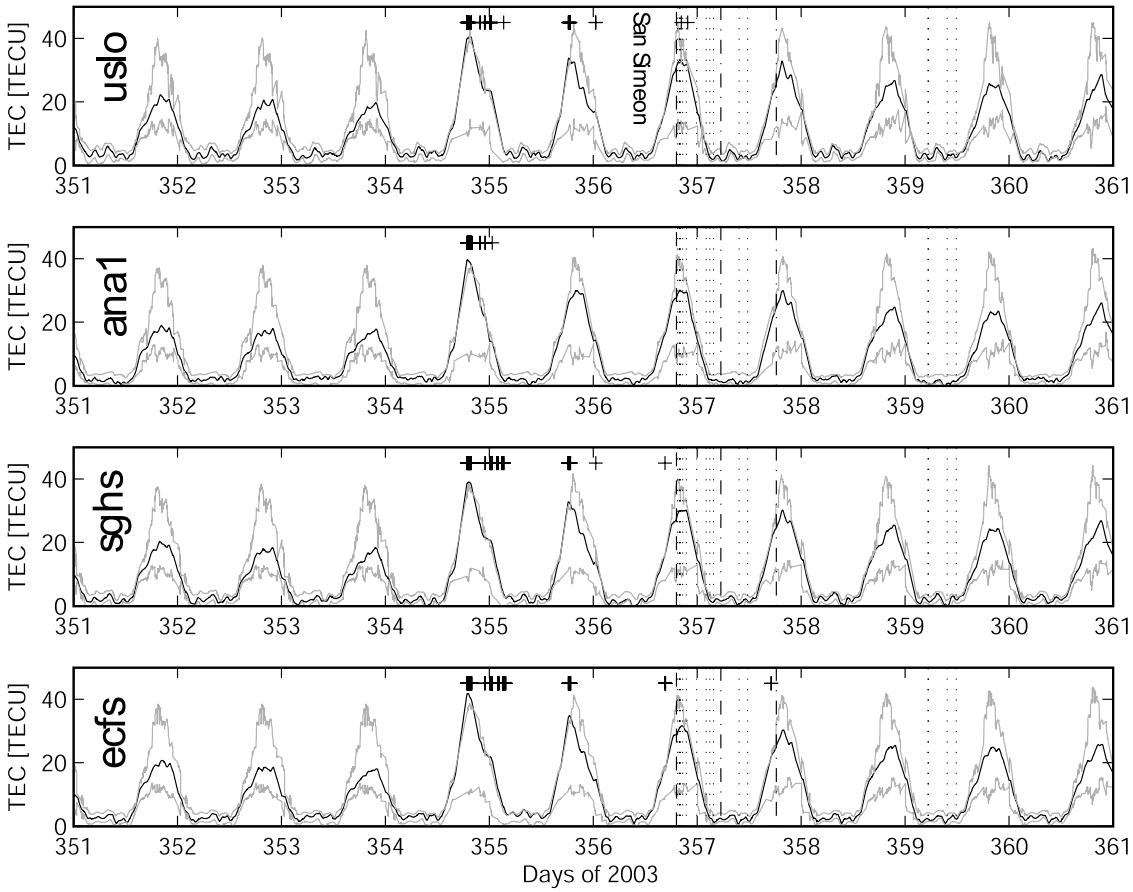

Figure 4. GPS-derived TEC time series at four sites of the SCIGN network for the 5 days preceding and following the 22 December 2003 San Simeon earthquake (black lines). The gray lines show the 2-standard deviation bounds derived from averaging the TEC time series over the 21 day period. Significant earthquakes are marked by vertical lines: magnitudes 6 to 7 with dashed lines, magnitudes 5 to 6 with dash-dotted lines, magnitudes 4 to 5 with dotted lines. The time of the 22 December 2003 San Simeon earthquake is indicated. Crosses show detections.

the 29 September 2004 M6.0 Parkfield earthquake shows no consistent detection preceding the event (Figure 6).

4.3. Search for a Spatial Pattern

[16] *Pulinets and Legen'ka* [2003] showed spatially coherent changes of *foF*2 within 29 hours before and 18 hours after the March 1964, M = 9.2 Alaska earthquake, and 46 hours before and 8 hours after the November 1980, M = 7.2 Campania/Basilicata earthquake in Italy. *Khegai et al.* [2002] also reported spatially coherent ionospheric signals starting 7 hours before the 1964 Alaskan earthquake. We therefore investigated the possibility of a spatial pattern in the geographic distribution TEC before the San Simeon and Parkfield earthquakes. We produced TEC maps by averaging TEC time series at all SCIGN sites over 30 min intervals and linearly interpolating between stations. The resulting maps (Figures 7 and 8) show no obvious spatial pattern around the location and time of the earthquake. However, an overall increase in TEC on the day before the San Simeon earthquake can be seen compared to the maps created 3, 4, and 5 days before at the same local time as the earthquake. This corresponds to the detections seen with the envelope method in the previous section. The maps around the Parkfield earthquake do not show such elevated TEC.

4.4. Search for a Temporal Pattern

[17] We applied the envelope method to the entire SCIGN network for the 2003–2004 time period, using the same criteria as above to define a "detection," weighting the detections linearly according to their duration, then binning these detections by day:

$$N_{weighted}(t) = \sum_{SCIGN} N_{station}(t)\frac{\Delta\tau}{90s} \tag{6}$$

where the fraction on the right side denotes the weighting factor. $\Delta\tau$ is the duration of the detection in seconds, t is the day considered, and $N_{station}$ denotes a detection (N = 1) or no detection (N = 0) at a single GPS site. Figure 9 shows that detections are found all year round, with a maximum in the summer. We then compute the mean and standard deviation of the resulting time series over 2 years to provide an objective threshold to identify periods of abundant and long lasting detections. We find that 33% of the detections

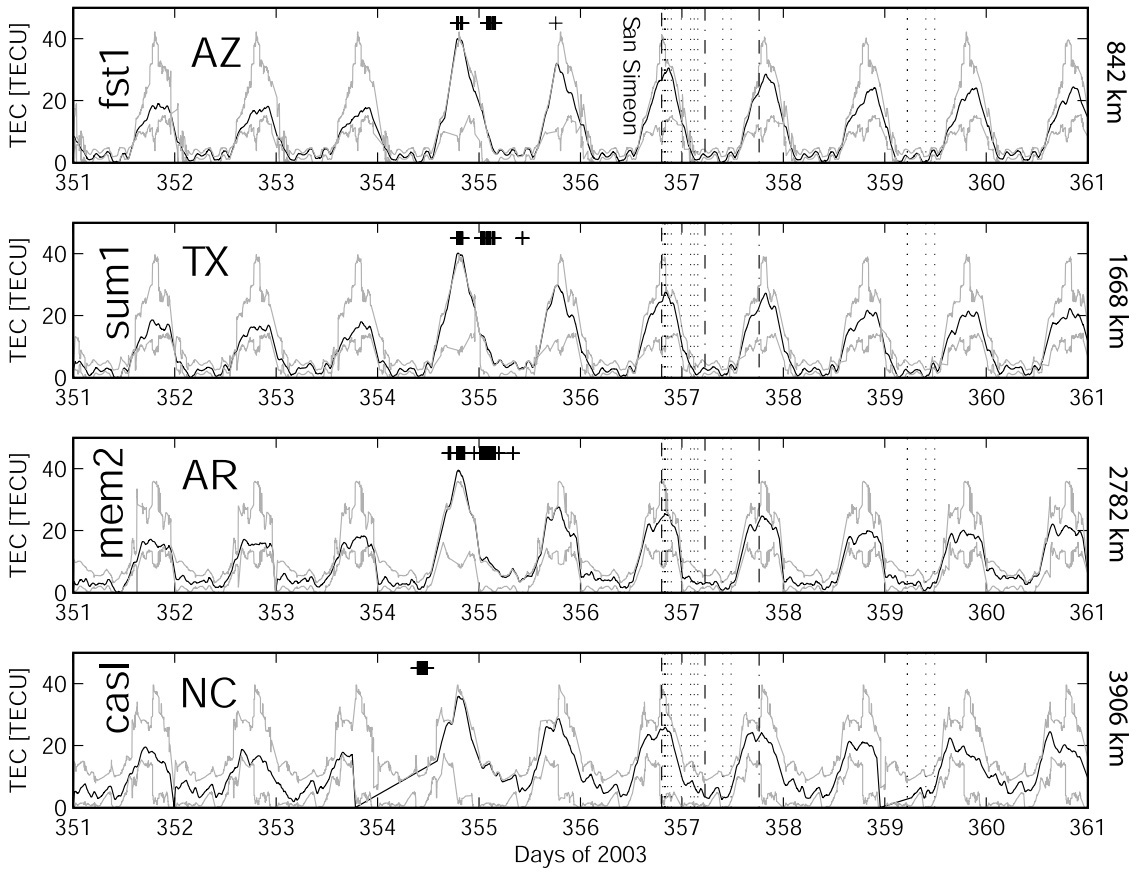

Figure 5. GPS-derived TEC time series at four continuous GPS sites across the U.S., located at a latitude similar to that of the SCIGN network and at distances ranging from 842 to 3906 km from the San Simeon epicenter (distances indicated on the right side of each plot). All other markings are the same as in Figure 4.

that are above the one-standard deviation threshold (black and gray dashed bins on Figure 9) are correlated with high solar flux or *Dst*. The remaining detections do not temporally correlate with earthquakes within the established time frame.

5. Correlation Method

5.1. Description

[18] This detection method builds on the concept of an "earthquake preparation area." The concept was first introduced by *Keylis-Borok and Malinovskaya* [1964], who found that the occurrence rate of moderate-size earthquakes increases during several years prior to large earthquakes over a large area centered on the future event. They proposed that the radius R of the preparation area scales with the magnitude M of the large event as $R = 10^{0.43M}$ km [see also *Dobrovolsky et al.*, 1979]. Since then, several studies have reported earthquake precursors, from earthquake clustering [e.g., *Zaliapin et al.*, 2002; *Kossobokov et al.*, 2000] to geochemical anomalies [e.g., *Fleischer*, 1981; *Toutain and Baubron*, 1998], with a spatial distribution consistent with the definition of an earthquake preparation area.

[19] The method, first proposed by *Pulinets et al.* [2004], is based on the fact that first-order TEC variations due to solar and geomagnetic activity are strongly correlated over distances up to several hundreds of kilometers. However, this correlation will break in the presence of precursory signals, which should be absent outside of the earthquake preparation area. Given the magnitude of the San Simeon and Parkfield earthquakes, the formula above leads to a 624 and 380 km preparation area radius, respectively. Both earthquakes occurred at the northern edge of the SCIGN network so that there is data available close to the epicenter and up to 800 km away.

[20] We calculate the correlation coefficient between two TEC time series using:

$$c_{1,2} = \frac{Cov_{1,2}}{\sqrt{Cov_{1,1}Cov_{2,2}}} \tag{7}$$

$Cov_{i,j}$, $i, j = 1, 2$ is the covariance matrix at zero lag. The covariance is defined as:

$$Cov_{i,j} = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle \tag{8}$$

$\langle X \rangle$ is the expectation value of X (i.e., its mean) and indices 1 and 2 correspond to the two GPS stations considered.

[21] We calculated one correlation coefficient per day between all pairs of sites that include one site within

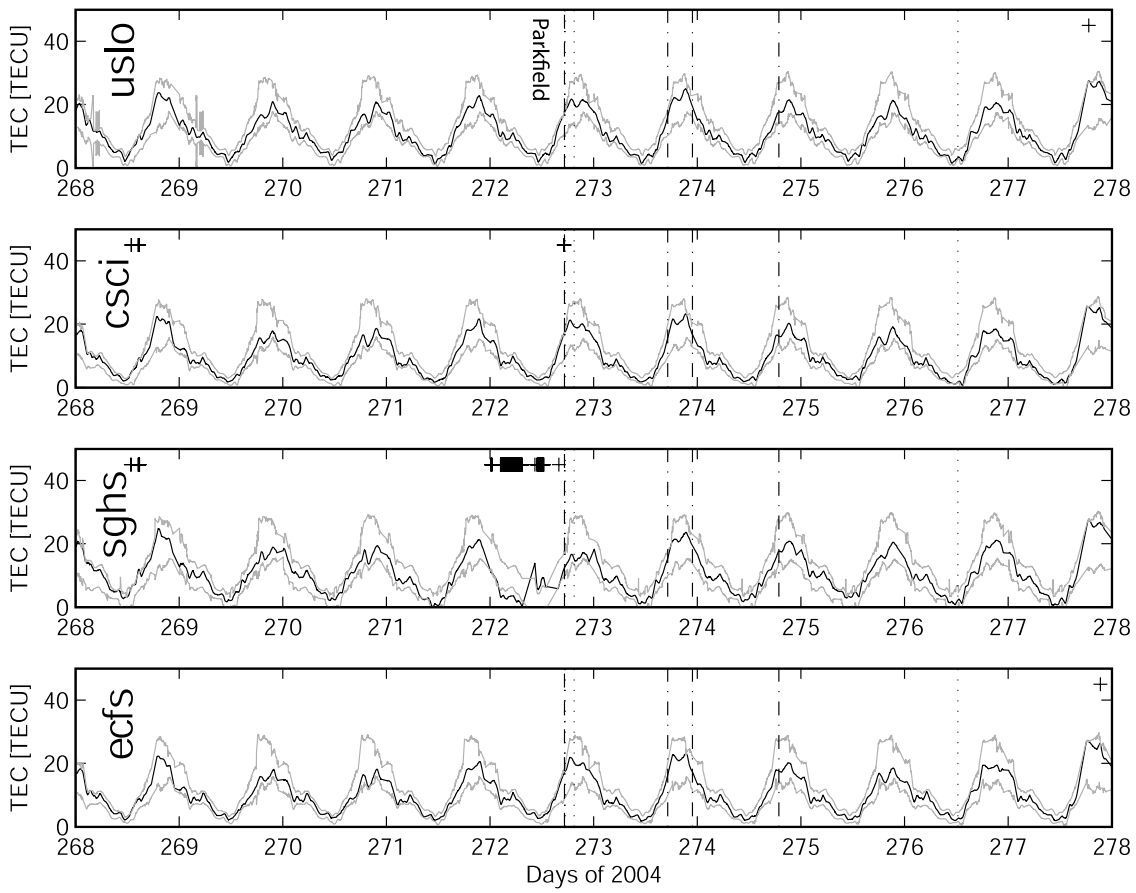

Figure 6. GPS-derived TEC time series at four sites of the SCIGN network for the 10 days surrounding the 29 September 2004 Parkfield earthquake (black lines). The time of the Parkfield earthquake is indicated. All other markings are the same as in Figure 4.

100 km from the epicenter (''earthquake area''), the other one between $r_{n-1} = (n - 1) \times 100$ km to $r_n = n \times 100$ km away (''outer ring area''). This typically amounts to about 3750 correlation coefficients per day for the 265 GPS stations of the SCIGN network with 15 (San Simeon) and 16 (Parkfield) stations within a 100 km radius of the epicenter. In order to increase the signal to noise ratio, we then stacked and normalized all correlation coefficients for each day and for each outer ring area. We repeated this for a range of outer ring areas using successive threshold distances r_n for $n = 2, \ldots, 7$. If a precursory signal is present in the earthquake preparation zone, then one should observe, before the earthquake, larger correlation coefficients for small threshold distances (i.e., within the preparation zone), while they should decrease for larger distances.

[22] However, as we increase the threshold distance in our tests, the average distance between sites also increases. This itself, in the absence of any precursory signal, will introduce a natural decorrelation of the TEC time series because of the natural spatial variability of the TEC [e.g., *Araujo-Pradere et al.*, 2005]. To quantify this relationship between correlation and distance in the absence of precursory signals, we computed the correlation coefficient for every possible pair of sites over two years (Figure 10). We separated this analysis for days of low (below one standard deviation), high (above one standard deviation), and regular (within ±1 standard deviation) solar flux and *Dst* index, calculated the averaged correlation coefficient and standard deviations for baseline lengths between 0–50 km, 50–100 km, …, and 800–850 km, and performed linear least square fits.

[23] We find that the correlation coefficient decreases linearly with distance at a slow rate, ranging from 0.115/1000 km to 0.061/1000 km (Figure 10). The decrease rate is lower for days of high *Dst* or solar flux, although the difference may not be significant given the scatter in the distribution of correlation coefficients (error bars on Figure 10). In all cases, the standard deviation about the mean is on the order of 0.2. We hence use the decrease found from the fit to the mean correlation coefficient minus one standard deviation as a minimum threshold below which the correlation coefficient has to drop in order to be considered significantly different from the natural spatial scatter of the TEC for the period considered.

5.2. Results

[24] Figure 11 shows the resulting correlation coefficient time series for 2003 (San Simeon earthquake) and 2004 (Parkfield earthquake). Overall, we find only very few days when the correlation coefficient drops below the one-sigma threshold: from 5 in 2003 and 4 in 2004 for the 100–200 km range to 28 in 2003 and 15 in 2004 for the 600–700 km

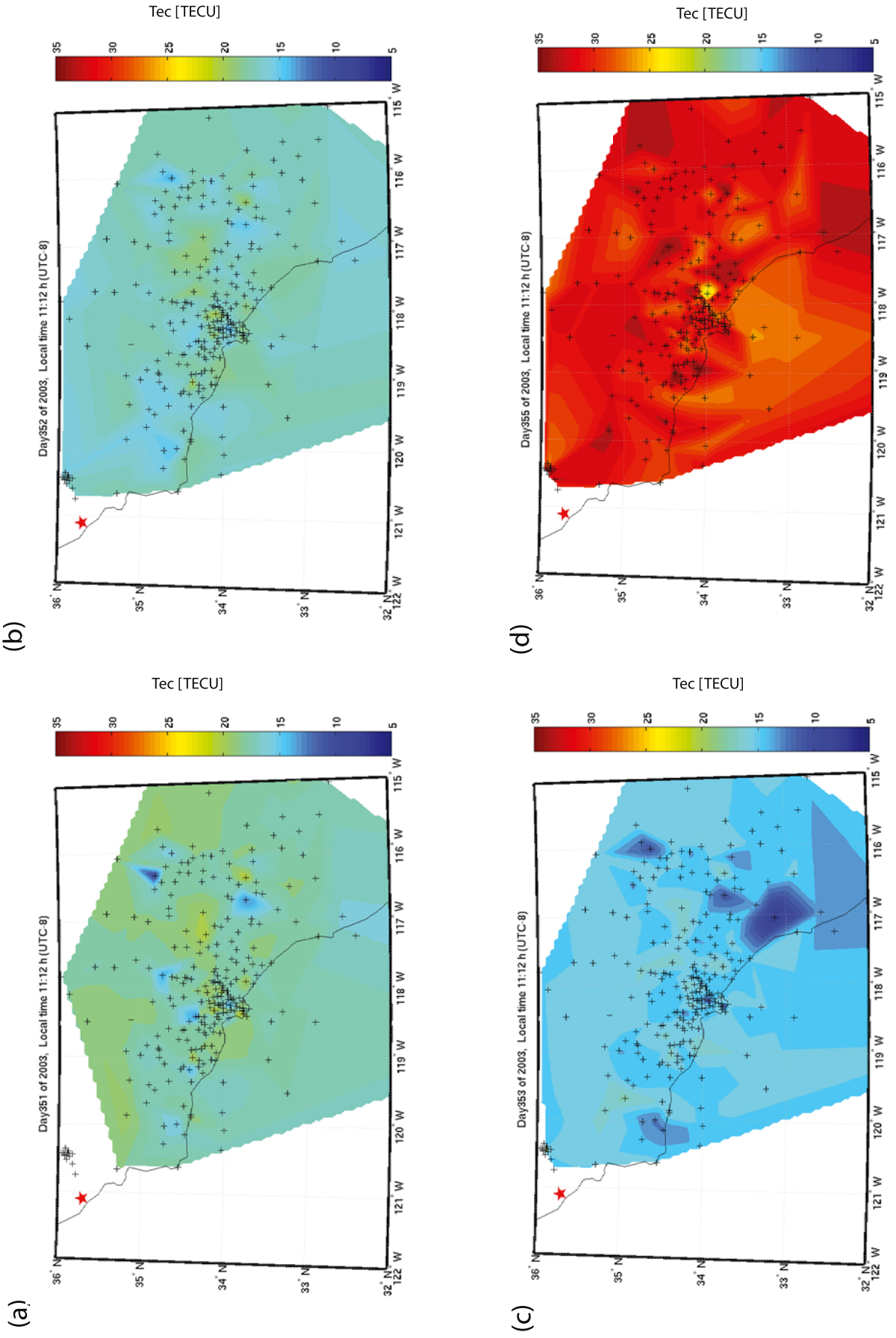

Figure 7. Interpolated maps of TEC around the time of the 22 December 2003 San Simeon earthquake (1115 local). Crosses show GPS stations, red star shows the earthquake epicenter, units are TECU. The indicated time is the beginning of the 30 min interval.

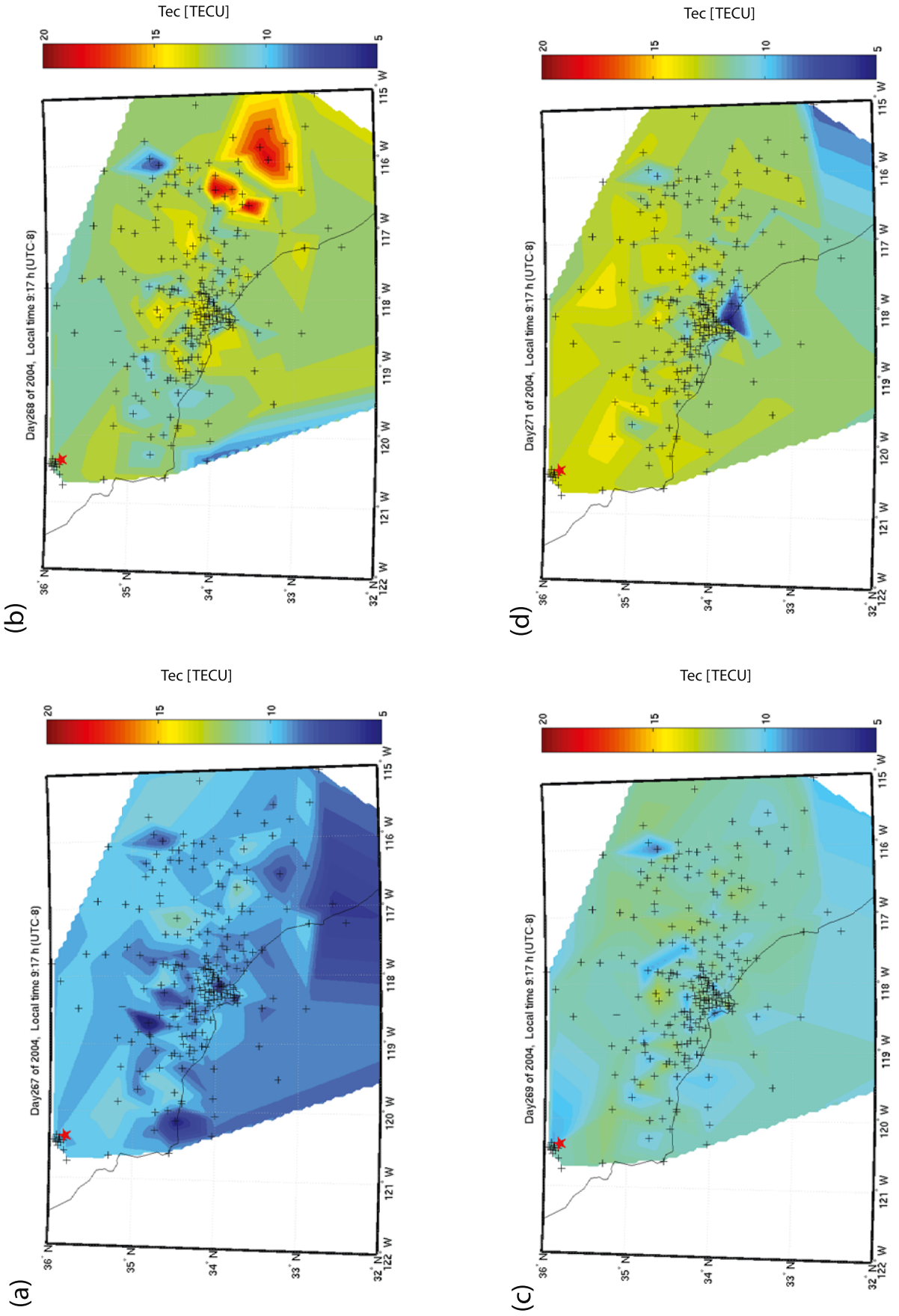

Figure 8. Interpolated maps of TEC around the time of the 28 September 2004 Parkfield earthquake (1015 local). Crosses show GPS stations, red star shows the earthquake epicenter, units are TECU. The indicated time is the beginning of the 30 min interval.

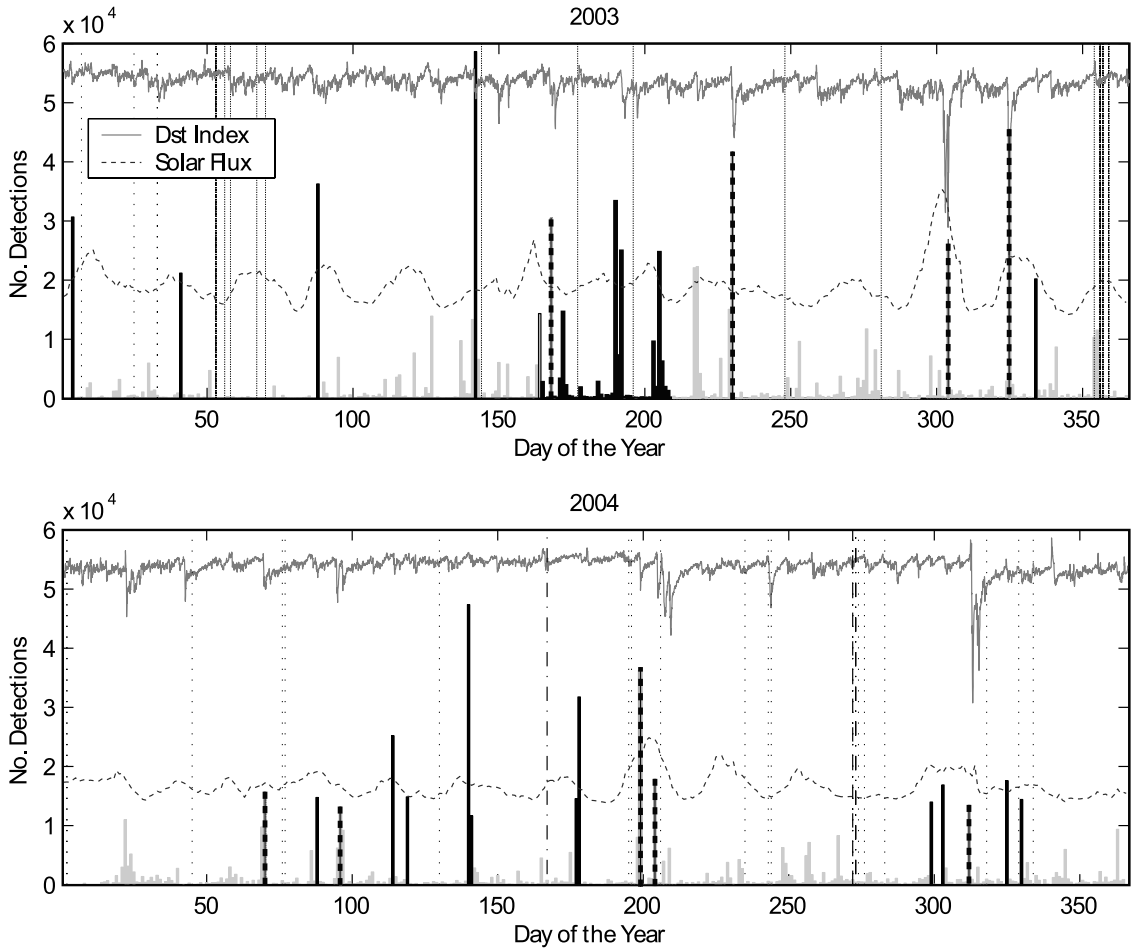

Figure 9. Total number of weighted daily detections for the entire SCIGN network using the envelope method. Earthquakes are marked by vertical lines: magnitude 6 to 7 with dashed lines, magnitude 5 to 6 with dash-dotted lines, magnitude 4 to 5 with dotted lines. Solar flux and *Dst* time series are shown with dashed and solid gray lines, respectively. Peaks in the detections (upper 33%) are shown as black-gray dashed bins if they coincide with a day of high solar flux or low *Dst*, in black otherwise. High solar flux or *Dst* are defined when the data deviate from the mean of the entire flux or *Dst* time series by more than one standard deviation.

range. These drops do not necessarily coincide with a preparation period before a large earthquake.

[25] In the last 2 months of 2003, before the San Simeon earthquake, we do observe a systematic increase in the number of days with a significant decrease in correlation for all threshold distances over 400 km. However this time period also coincides with two months of increased solar flux and magnetic activity (Figure 11). The onset of the decorrelation (around day 300), coincident with the yearly maximum in solar flux and a large spike in the *Dst* index is particularly striking. The TEC anomaly found here before the San Simeon earthquake is therefore more likely to be due to solar or geomagnetic activity. In 2004 we do not observe a significant decrease in correlation coefficients during the month preceding the Parkfield earthquake.

6. Signal Detection Approach

6.1. Description

[26] In order to further quantify the statistics associated with the detection of a signal, we adapted the Neyman-Pearson (NP) test, a standard method of detection theory used to quantify the probability of detecting a signal buried in noise [e.g., *Kay*, 1998]. The method is based on a quantitative comparison of the conditional probability density functions (PDFs) of a measurement, or "test statistic," in the presence of noise, under two hypotheses: the signal is present or the signal is not present (noise only). The two conditional PDFs are illustrated schematically in Figure 12. In the NP test, a probability of false alert P_{fa} is set, which will define (using the PDF conditioned on the "no signal" hypothesis) a detection threshold for the measurement. The probability of detection can then be computed from the area to the right of this threshold, under the PDF conditioned on the "signal present" hypothesis. Other probabilities, such as the probability of missed detection, can be similarly computed once the two conditional PDFs are known, and the threshold is defined.

[27] To apply this technique, we first remove the diurnal TEC variations from the observed TEC time series. To do so, we average the TEC time series at all stations to obtain a "virtual TEC time series" for California for 2003 and 2004

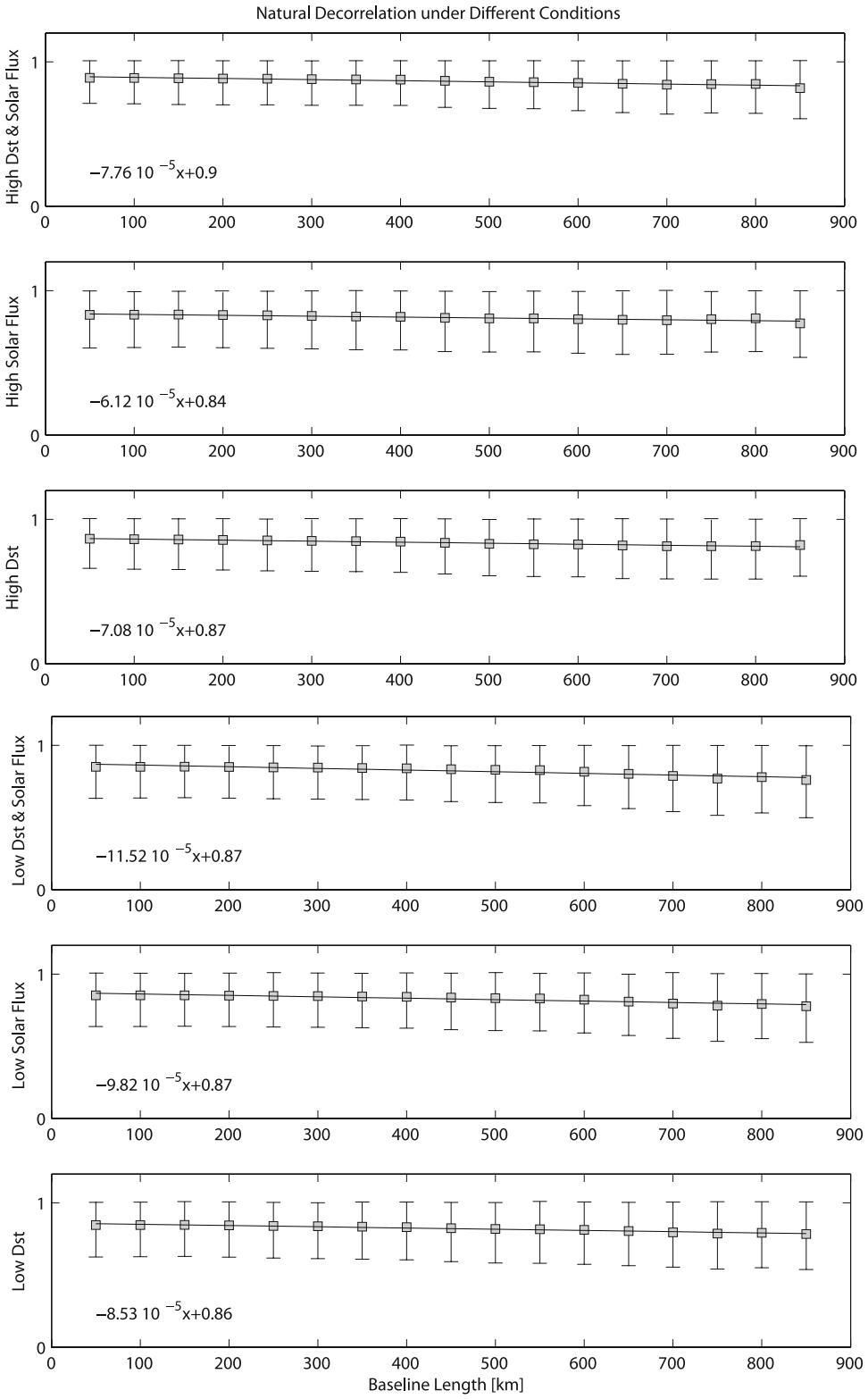

Figure 10. Mean correlation coefficient between TEC time series at all site pairs of the SCIGN network as a function of the distance between sites. The error bars show the associated 1-sigma standard deviation up to the maximum value of 1.

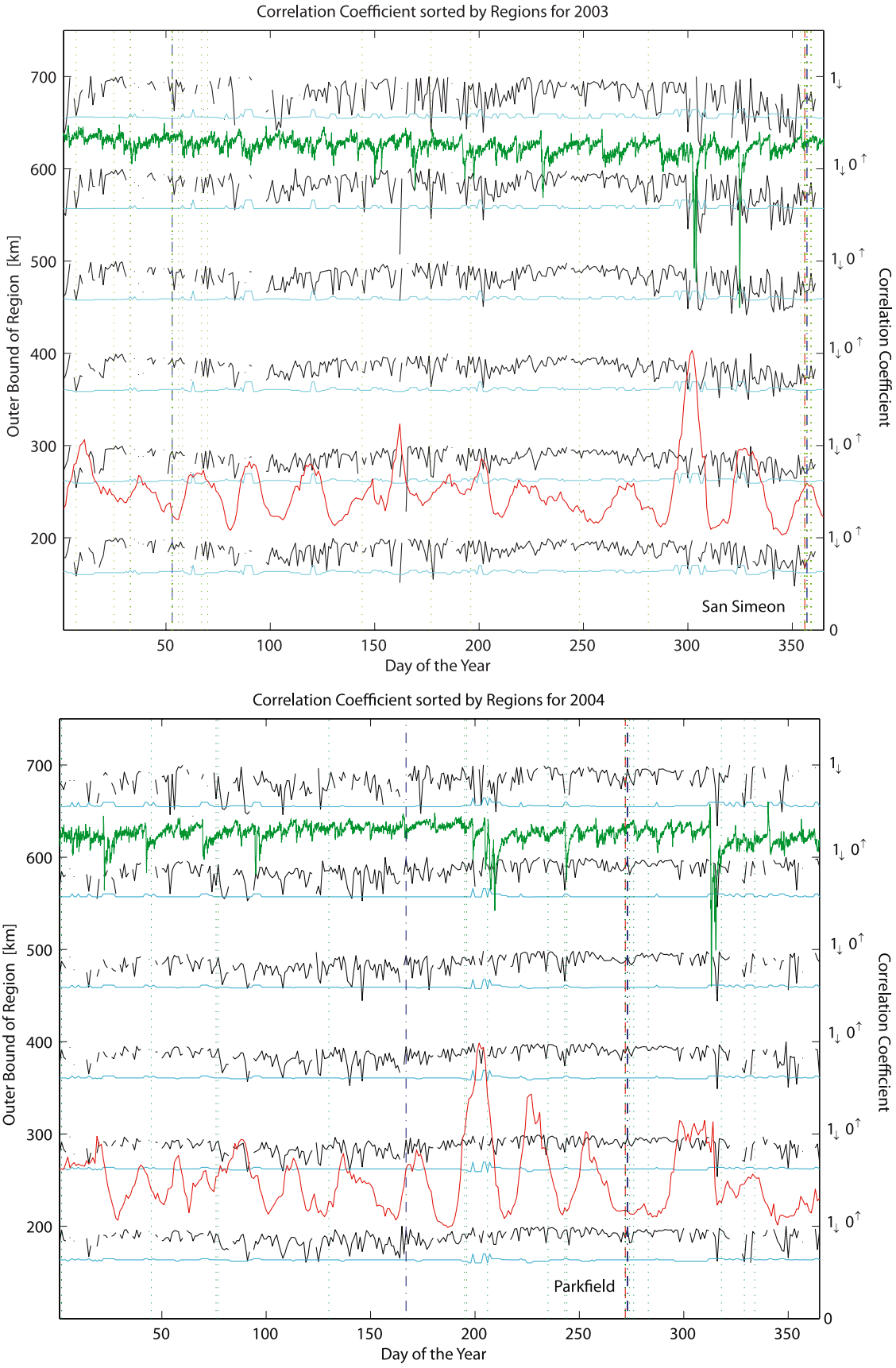
Correlation Coefficient sorted by Regions for 2003
Outer Bound of Region [km]
Day of the Year
Correlation Coefficient
San Simeon
Correlation Coefficient sorted by Regions for 2004
Outer Bound of Region [km]
Day of the Year
Correlation Coefficient
Parkfield

Figure 11

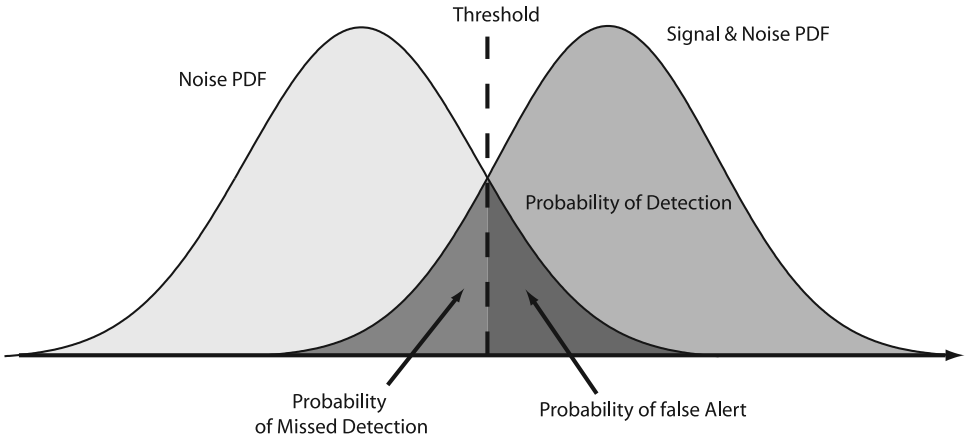

Figure 12. Schematic explanation of the method used to calculate the probability of detection of an anomalous TEC signal. A detection threshold is computed from the noise probability density function for a given probability of false alert P_{fa}, the probability of detection P_D is then computed from the signal + noise PDF.

(Figure 2). We then subtract that time series from the TEC time series at each individual station to obtain residual TEC time series. We define the "signal" as the residual TEC for each day and the "noise" as the residual TEC for all of the preceding and following 10 days. We chose 10 days because it encompasses the time delays that have been claimed between precursory signals and earthquakes in previous studies.

[28] We assume a Gaussian distribution for both the noise and the signal, compute their mean μ and standard deviation σ over the above defined time interval, thus obtaining their probability density functions. Indices n and s below stand for noise and signal, respectively. The probability of false alarm P_{fa} is preset and related to the detection threshold T through [*Kay*, 1998]:

$$P_{fa} = \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_T^\infty e^{-\frac{1}{2}\left(\frac{t-\mu_n}{\sigma_n}\right)^2} dt \tag{9}$$

Substituting

$$q = \frac{t-\mu_n}{\sigma_n\sqrt{2}} \tag{10}$$

with the new boundary

$$Q = \frac{T-\mu_n}{\sqrt{2}\sigma_n} \tag{11}$$

yields

$$P_{fa} = \frac{1}{\sqrt{\pi}} \int_Q^\infty e^{-q^2} dq \tag{12}$$

Using the complementary error function, the threshold T can be calculated iteratively. Once the threshold is found, the probability of detection P_D is given by:

$$P_D = \frac{1}{\sqrt{2\pi\sigma_s^2}} \int_T^\infty e^{-\frac{1}{2}\left(\frac{t-\mu_s}{\sigma_s}\right)^2} dt \tag{13}$$

[29] Note that the probability of detection will decrease as the probability of false detection decreases (i.e., as we accept fewer mistakes, Figure 12). Hence if the probability density functions for signal and noise are the same, which we expect in the case of no signal is present, we expect P_D to be the same as P_{fa}. If a precursor is present, then the probability of detection will indicate the likelihood that it is detected. For the following analysis we chose a probability of false alert $P_{fa} = 0.01$.

6.2. Resulting Time Series

[30] At site RNCH, located 56 km from the December 2003 San Simeon epicenter, we find a probability of detection of 0.012 on the day preceding the event. In other words, there is a 1.2% probability that we detect an anomaly in the TEC on the day preceding the earthquake given that we accept, if no signal was present, that one out of 100 trials would also result in a detection (i.e., false alarm). This preset probability of false detection is rather large, lowering

Figure 11. Mean correlation coefficient of daily TEC time series between stations located within a 100 km radius of the San Simeon (top) or Parkfield (bottom) epicenters and stations located more than 100, 200, 300, 400, 500, and 600 km away. The horizontal blue line shows the 1-sigma threshold of the natural variability of TEC correlations for active and quiet days for each distance considered (see Figure 10). Earthquakes are marked by vertical lines: magnitudes 6 to 7 with red dashed lines, magnitudes 5 to 6 with blue dash-dotted lines, magnitudes 4 to 5 with green dotted lines. Solar flux and *Dst* time series are shown with red and green lines, respectively.

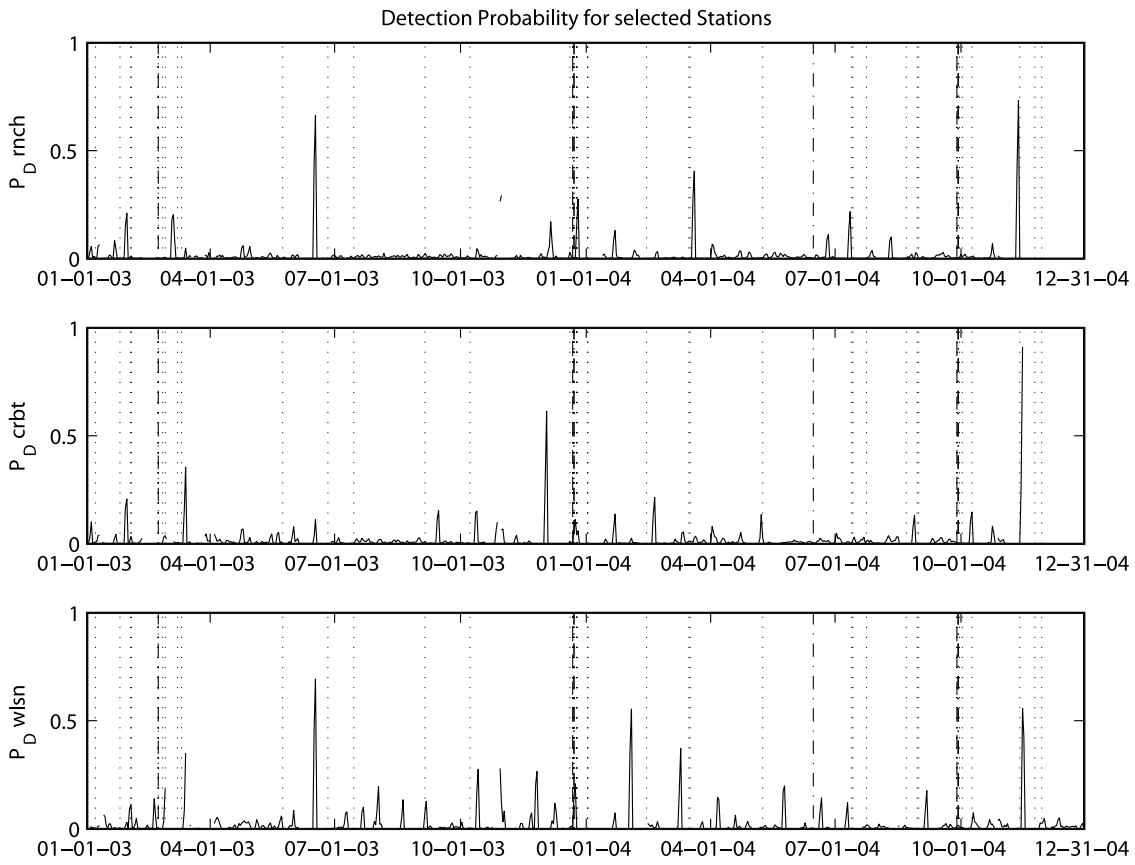

Figure 13. Probability of detection (P_D) time series at sites RNCH (56 km from San Simeon epicenter, top), CRBT (33 km from epicenter, middle), and WSLN (323 km from epicenter, bottom), for a probability of false alarm (P_{fa}) of 0.01. Earthquakes are marked by vertical lines: magnitudes 6 to 7 with dashed lines, magnitudes 5 to 6 with dash-dotted lines, magnitudes 4 to 5 with dotted lines.

it to more conservative values would result in decreasing the probability of detection. Site CRBT, located only 33 km from the epicenter, has a detection probability of 0.005 for the day preceding the San Simeon earthquake, site WLSN, located 323 km from the epicenter, shows a detection probability P_D of 0.013. Overall, the P_D time series for 2003–2004 at sites RNCH, CRBT, and WLSN (Figure 13) is close to zero for most of the time span considered with only 2 to 3 days where it spikes above 0.5. The same holds for the rest of the SCIGN network. There is no apparent temporal and spatial correlation of the increase in P_D with the occurrence of earthquakes and the distance to the epicenter.

[31] If there was a precursory signal in the TEC data, then the probability of detection should be (1) large before the event, (2) systematically larger at sites located inside the earthquake preparation area (see discussion above), and (3) correlated between nearby sites. As mentioned above, we find that the P_D time series show very different patterns from site to site, even for nearby sites. In order to increase the signal-to-noise ratio of the P_D time series, we stacked them for ensembles of sites located within successive distance intervals of 100 km to the epicenter (i.e., we average all stations located within 0–100 km, then 100–200 km, etc.). As for the correlation method described above, this stacking procedure should enhance precursory signals because they are expected to be coherent in time at neighboring sites. On the other hand, signals not related to precursory activity (i.e., "noise") are temporally incoherent between sites and should be attenuated. Note that, since we have already removed the first-order TEC variations by subtracting an average 2003–2004 TEC time series from all SCIGN sites, a large fraction of the correlated noise has already been removed. The resulting P_D time series are plotted on Figure 14 (for a probability of false alarm $P_{fa} = 0.01$) for a series distances to the epicenter. We find that the P_D value rarely rises above 0.1. This is true, in

Figure 14. Averaged time series of the probability of detection (P_D) given a present probability of false alarm (P_{fa}) of 0.01, shown for groups of stations located within successive distance intervals of 100 km from the epicenter (i.e., 0–100 km, 100–200 km, etc.). Solar flux and *Dst* index time series are shown in the background in dashed and solid gray, respectively. Each 100 km interval along the y-axis also corresponds to a 0 to 1 interval of P_D. Earthquakes are marked by vertical lines: magnitudes 6 to 7 with dashed lines, magnitudes 5 to 6 with dash-dotted lines, magnitudes 4 to 5 with dotted lines.

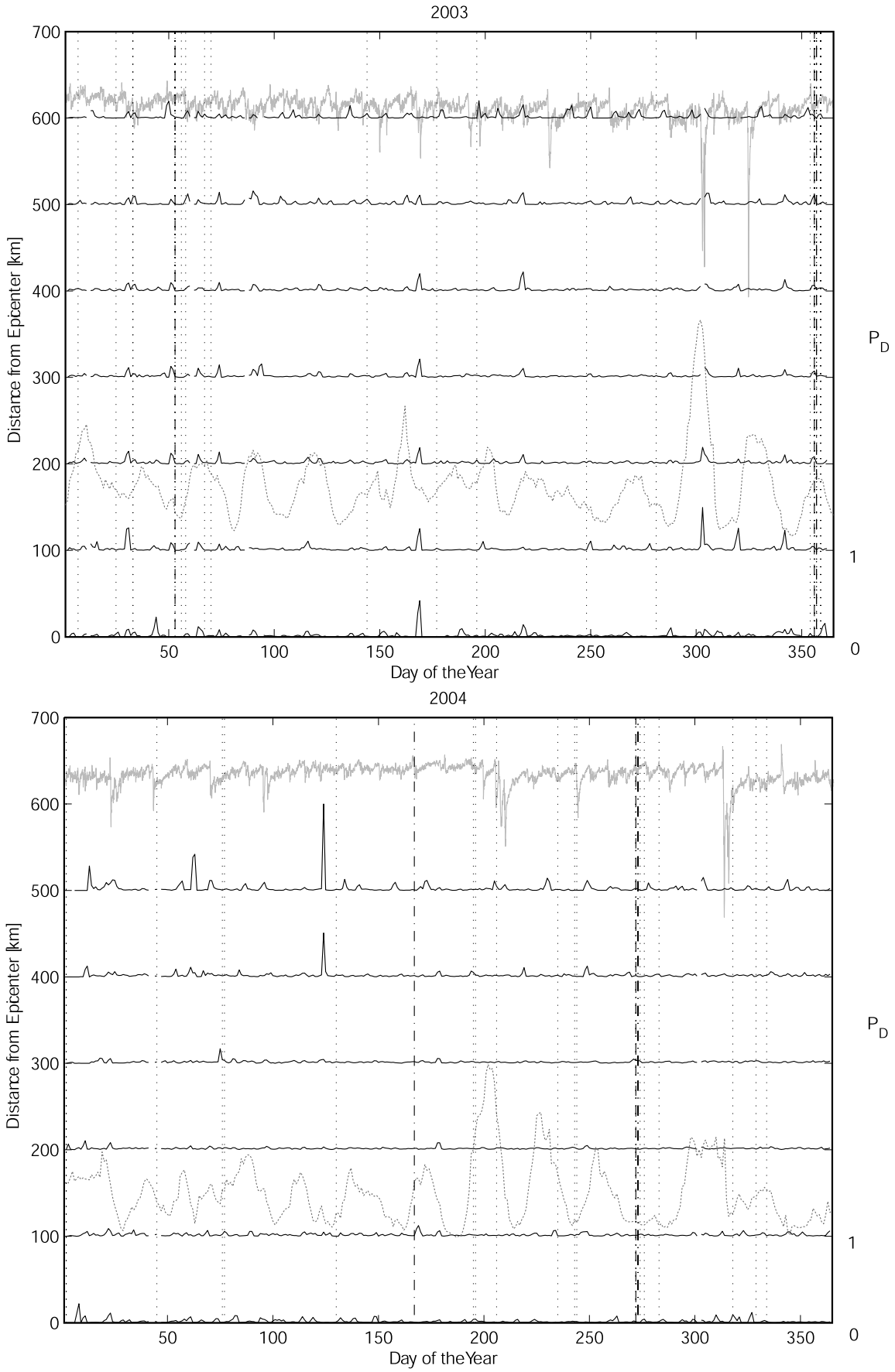
2003
Distance from Epicenter [km]
Day of the Year
P_D
2004
Distance from Epicenter [km]
Day of the Year
P_D

Figure 14

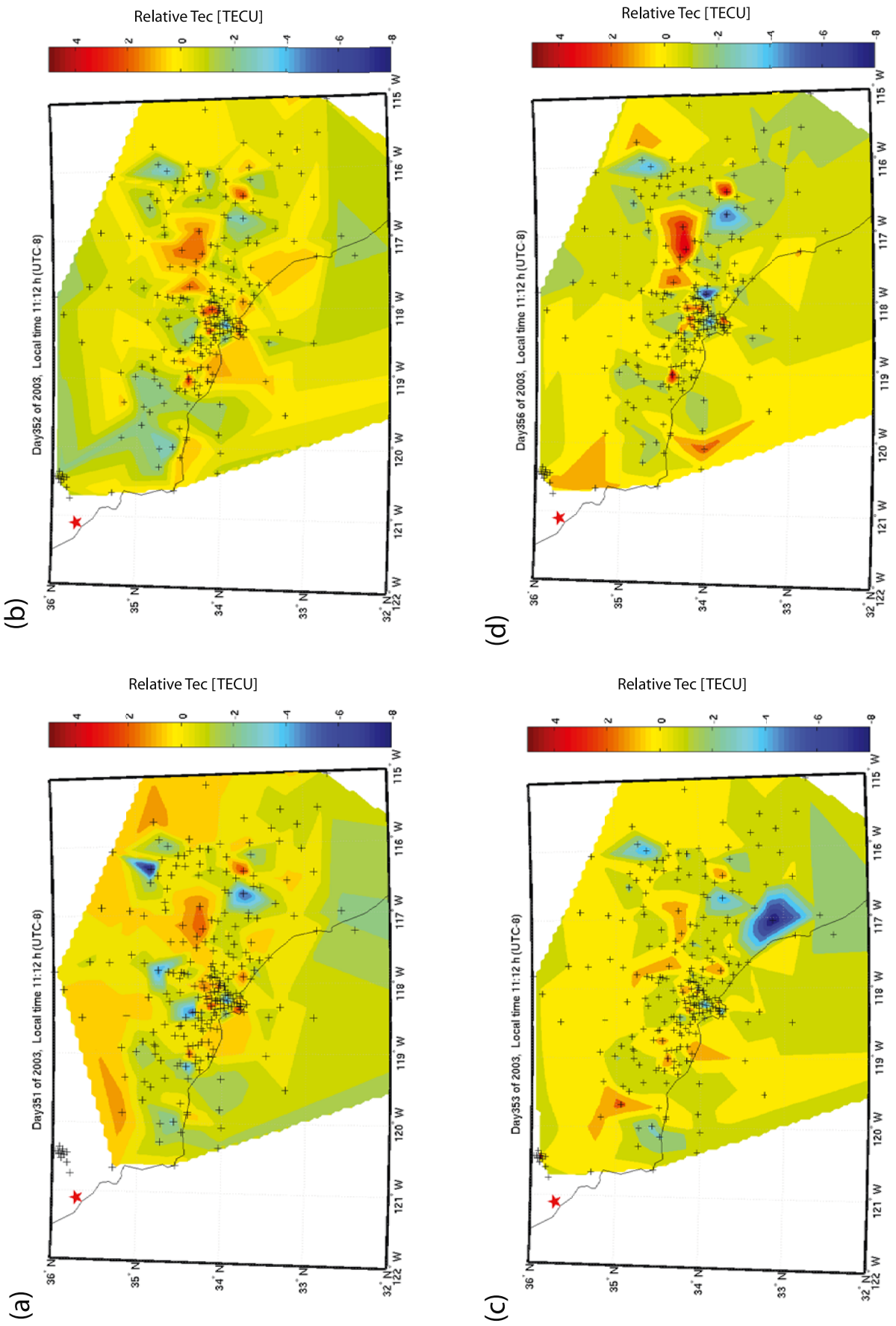

Figure 15. Interpolated maps of residual TEC around the time of the 22 December 2003 San Simeon earthquake (1915 UTC). Crosses show GPS stations, red star shows the earthquake epicenter, units are TECU. The indicated time is the beginning of the 30 min interval.

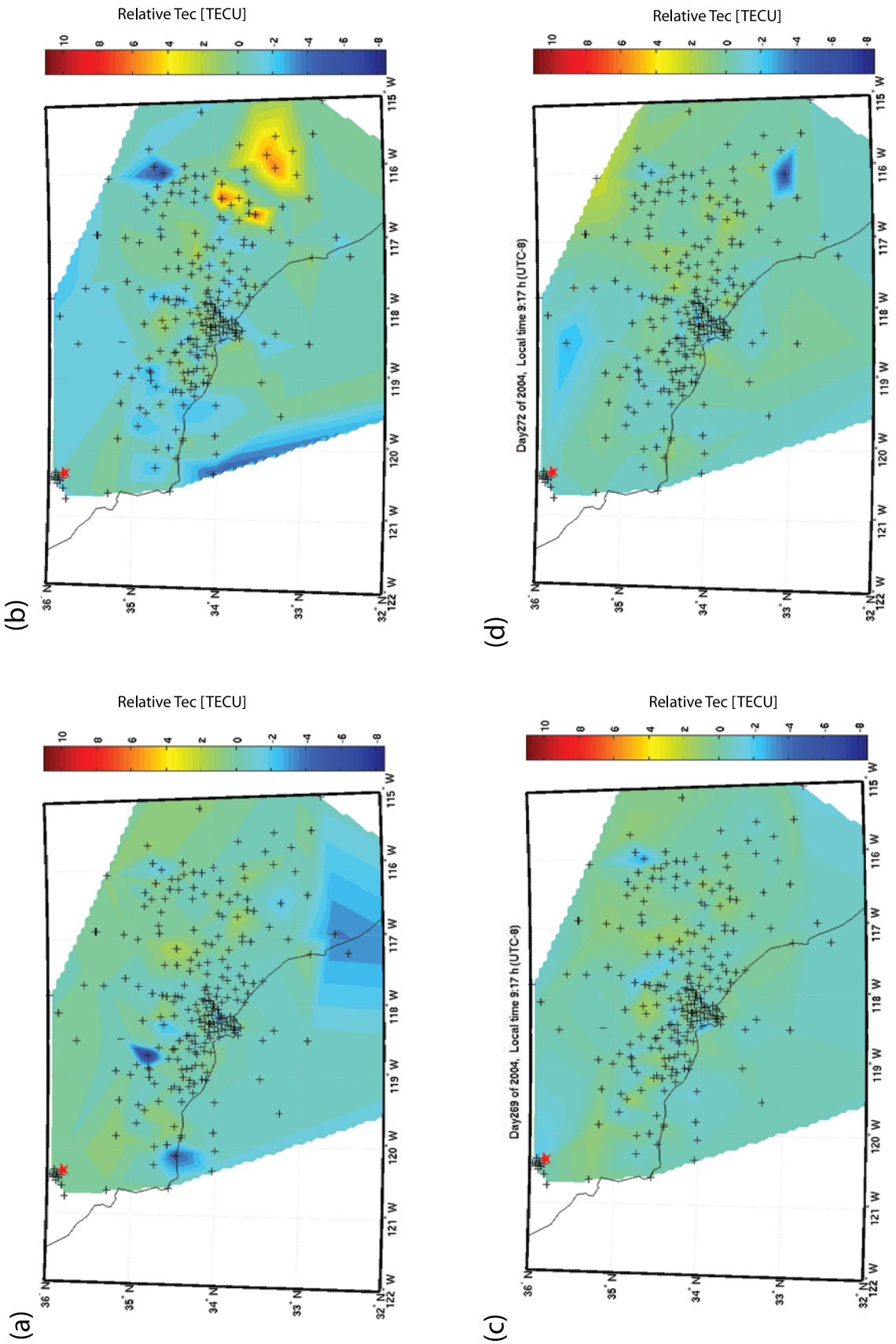

Figure 16. Interpolated maps of residual TEC around the time of the 28 September 2004 Parkfield earthquake (1715 UTC). Crosses show GPS stations, red star shows the earthquake epicenter, units are TECU. The indicated time is the beginning of the 30 min interval.

particular, for the days preceding the San Simeon and Parkfield earthquakes.

6.3. Search for Spatial Patterns

[32] As for the envelope method (see above), we also investigate the possibility of a spatial pattern in the geographic distribution of the residual TEC (i.e., observed TEC at each site minus stacked TEC time series of Figure 2). We averaged the residual TEC values over 30 min intervals and used a simple linear interpolation of the TEC between stations. Figures 15 and 16 show the resulting interpolated TEC maps around the time of the San Simeon and Parkfield earthquakes, spanning similar time periods as previously reported precursors [e.g., *Pulinets and Legen'ka*, 2003; *Zaslavski et al.*, 1998]. As for the TEC maps (Figures 7 and 8), we do not find any convincing spatial pattern that could be correlated with the earthquake epicenters.

7. Conclusions

[33] We have used 2 years (2003–2004) of continuous GPS data at 265 stations in southern California spanning the December 2003, M6.6, San Simeon and September 2004, M6.0, Parkfield earthquakes to search for precursory signals in GPS-derived TEC measurements. The density of continuous GPS stations in southern California allows us to test for the significance of the temporal and spatial correlation of anomalous TEC signals with earthquakes. We produce TEC time series at all sites and use three different methods (two of which were used by other authors to show the existence of earthquake precursors) to test for the presence of anomalous TEC signals related to earthquakes. We found anomalous TEC signals using all available data, but were not able to find any statistically significant correlation, in time or in space, between these TEC anomalies and the occurrence of earthquakes in Southern California for the 2003–2004 period.

[34] The detailed analysis of this very large data set shows, however, that it is almost always possible to find isolated evidence for an anomalous TEC signal in the days preceding an earthquake. By "isolated," we mean detection at either a single station, or detection at multiple stations, but only using a single method, with no detection confirmed by the other methods. We however argue that signals must not only be temporally correlated with earthquakes, but also spatially coherent and detected at a significant level by several techniques, to qualify as precursors. The quality and spatial density of the SCIGN data set allows us to test for such situations.

[35] We find that the period surrounding the San Simeon earthquake suffers from significant space weather effects. This makes any attempt to detect ionospheric effects of earthquakes particularly difficult. It is likely that the precursory event to the San Simeon earthquake reported by *Pulinets* [2004] was actually an artifact resulting from this enhanced space weather activity. The period surrounding the Parkfield earthquake is quieter, but none of the three methods used here provides significant evidence for a TEC perturbation preceding that event.

[36] Although our results do not corroborate previous reports of ionospheric precursors to earthquakes, they do not disprove the possibility of precursory phenomena. Our work illustrates some of the difficulties associated with the search for precursory activity, but also points to directions for future work. Clearly, in the absence of a priori knowledge of the temporal and spatial structure of precursory signals, separating these signals from the background TEC variability is a challenge. Given the spatial and temporal variability of the TEC in the absence of earthquakes, we argue that long time series, spatially dense measurements, and the use of several statistical detection techniques are essential to demonstrate the presence of earthquake precursors.

[37] The methods used give simple robust estimates of the background TEC variability and detection statistics. Filtering out days of large solar and/or magnetic activity is key and could be done in a more quantitative manner. The possibility remains that ionospheric precursors are present at periods shorter than a day, even though, again, discriminating them from nonseismic perturbations such as caused by acoustic-gravity waves might also be problematic. Finally, systematically combining measurements from independent sensors such as GPS (for the ionospheric TEC) and ground or space-based electromagnetic measurements will help to understand the origin of the observed signals. Combining those in a future analysis of large earthquakes in well-instrumented areas such as the western U.S. or Japan may still reveal convincing evidence for precursors.

[38] **Acknowledgments.** This work was supported by Quakefinder LLC and Stellar Solutions (under contract from NASA), subcontract QF204PU1 to Purdue. We acknowledge the Southern California Integrated GPS Network and its sponsors, the W.M. Keck Foundation, NASA, NSF, USGS, SCEC, for providing the data used in this study.

References

Araujo-Pradere, E. A., T. J. Fuller-Rowell, M. V. Codrescu, and D. Bilitza (2005), Characteristics of the ionospheric variability as a function of season, latitude, local time, and geomagnetic activity, *Radio Sci.*, *40*, RS5009, doi:10.1029/2004RS003179.

Bishop, J. (1981), Piezoelectric effects in quartz-rich rocks, *Tectonophysics*, *77*, 297–321.

Calais, E., and J. B. Minster (1995), GPS detection of ionospheric perturbations following the January 17, 1994, Northridge earthquake, *Geophys. Res. Lett.*, *22*, 1045–1048.

Dixon, T. H. (1991), An introduction to the global positioning system and some geological applications, *Rev. Geophys.*, *29*, 249–276.

Dobrovolsky, I. P., S. I. Zubkov, and V. I. Miachkin (1979), Estimation of the size of earthquake preparation zones, *Pure Appl. Geophys.*, *117*, 1025–1044.

Finkelstein, D., U. Hill, and J. Powell (1973), The piezoelectric theory of earthquake lightning, *J. Geophys. Res.*, *78*, 992–993.

Fleischer, R. L. (1981), Dislocation model for radon response to distant earthquakes, *Geophys. Res. Lett.*, *8*, 477–480.

Fraser-Smith, A. C., A. Bernardi, P. R. McGill, M. E. Ladd, R. A. Helliwell, and O. G. Villard (1990), Low-frequency magnetic field measurements near the epicenter of the M_s 7.1 Loma Prieta earthquake, *Geophys. Res. Lett.*, *17*, 1465–1468.

Freund, F. (2000), Time-resolved study of charge generation and propagation in igneous rocks, *J. Geophys. Res.*, *105*, 11,001–11,020.

Fujiwara, H., M. Kamogawa, M. Ikeda, J. Y. Liu, H. Sakata, Y. I. Chen, H. Ofuruton, S. Muramatsu, Y. J. Chuo, and Y. H. Ohtsuki (2004), Atmospheric anomalies observed during earthquake occurrences, *Geophys. Res. Lett.*, *31*, L17110, doi:10.1029/2004GL019865.

Gokhberg, M. B., V. A. Morgounov, T. Yoshino, and I. Tomizawa (1982), Experimental measurement of electromagnetic emissions possibly related to earthquakes in Japan, *J. Geophys. Res.*, *87*, 7824–7828.

Hayakawa, M., R. Kawate, O. A. Molchanov, and K. Yumoto (1996), Results of ultra-low-frequency magnetic field measurements during the Guam earthquake of 8 August 1993, *Geophys. Res. Lett.*, *23*, 241–244.

Hayakawa, M., T. Ito, and N. Smirnova (1999), Fractal analysis of ULF geomagnetic data associated with the Guam earthquake on August 8, 1993, *Geophys. Res. Lett.*, *26*, 2797–2800.

Hayakawa, M., T. Itoh, K. Hattori, and K. Yumoto (2000), ULF electromagnetic precursors for an earthquake at Biak, Indonesia on February 17, 1996, *Geophys. Res. Lett.*, *27*, 1531–1534.

Henderson, T. R., V. S. Sonwalkar, R. A. Helliwell, U. S. Inan, and A. C. Fraser-Smith (1993), A Search for ELF/VLF emissions induced by earthquakes as observed in the ionosphere by the DE 2 satellite, *J. Geophys. Res.*, *98*, 9503–9514.

Kay, S. M. (1998), *Fundamentals of Statistical Signal Processing II Detection Theory*, Prentice-Hall, Upper Saddle River, N. J.

Keylis-Borok, V. I., and L. N. Malinovskaya (1964), One regularity in the occurrence of strong earthquakes, *J. Geophys. Res.*, *69*, 3019–3024.

Khegai, V., A. Legen'ka, S. Pulinets, and V. Kim (2002), Variations in the ionospheric F2 region prior to the catastrophic earthquake in Alaska on March 28, 1964, according to the data of the ground-based stations of the ionospheric vertical sounding, *Geomagn. Aeron.*, *42*, 344–349.

Klobuchar, J. A. (1985), Ionospheric time delay effects on earth space propagation, in *Handbook of Geophysics and the Space Environment*, edited by A. S. Jursa, chap. 10.8, pp. 1084–1088, U.S. Air Force, Washington, D. C.

Kossobokov, V. G., V. I. Keilis-Borok, D. L. Turcotte, and B. D. Malamud (2000), Implications of a statistical physics approach for earthquake hazard assessment and forecasting, *Pure Appl. Geophys.*, *157*, 2323–2349.

Larkina, V. I., V. V. Migulin, O. A. Molchanov, I. P. Kharkov, A. S. Inchin, and V. B. Schvetcova (1989), Some statistical results on very low frequency radiowave emissions in the upper ionosphere over earthquake zones, *Phys. Earth Planet. Inter.*, *57*, 100–109, doi:10.1016/0031-9201(89)90219-7.

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

Liu, J. Y., Y. I. Chen, Y. J. Chuo, and H. F. Tsai (2001), Variations of ionospheric total electron content during the Chi-Chi earthquake, *Geophys. Res. Lett.*, *28*, 1383–1386.

Lonre, B., F. Perrier, and J. P. Avouac (1999), Streaming potential measurements: 2. Relationship between electrical and hydraulic flow patterns from rock samples during deformation, *J. Geophys. Res.*, *104*, 17,879–17,896.

Manucci, A., B. Wilson, and C. Edwards (1993), A new method for monitoring the Earth's ionospheric total electron content using the GPS global network, in *ION GPS-93*, pp. 22–24, Inst. of Navig., Fairfax, Va.

Molchanov, O. A., and M. Hayakawa (2001), VLF monitoring of atmosphere-ionosphere boundary as a tool to study planetary waves evolution and seismic influence, *Phys. Chem. Earth C*, *26*, 453–458.

Morrison, F., W. E. R., and T. Madden (1989), Streaming potentials of westerly granite with applications, *J. Geophys. Res.*, *94*, 12,449–12,461.

Nicholson, J. R., and W. R. Steiger (1963), On lunar semidiurnal tidal variations in the F_2 layer of the ionosphere, *J. Geophys. Res.*, *68*, 3577–3580.

Parrot, M., and F. Lefeuvre (1985), Correlation between GOES VLF emissions and earthquakes, *Ann. Geophys.*, *3*, 733–748.

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

Pulinets, S. A., and A. D. Legen'ka (2003), Spatial-temporal characteristics of large scale disturbances of electron density observed in the ionospheric F-region before strong earthquakes, *Cosmological Res.*, *41*, 221–230.

Pulinets, S. A., K. A. Boyarchuk, V. V. Hegai, V. P. Kim, and A. M. Lomonosov (2000), Quasielectrostatic model of atmosphere-thermosphere-ionosphere coupling, *Adv. Space Res.*, *26*, 1209–1218.

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

Revil, A., P. A. Pezard, and P. W. J. Glover (1999), Streaming potential in porous media: 1. Theory of the zeta potential, *J. Geophys. Res.*, *104*, 20,021–20,032.

Serebryakova, O. N., S. V. Bilichenko, V. M. Chmyrev, M. Parrot, J. L. Rauch, F. Lefeuvre, and O. A. Pokhotelov (1992), Electromagnetic ELF radiation from earthquake regions as observed by low-altitude satellites, *Geophys. Res. Lett.*, *19*, 91–94.

Shvets, A., M. Hayakawa, and S. Maekawa (2004), Results of subionospheric radio LF monitoring prior to the Tokachi (M = 8, Hokkaido, 25 September 2003) earthquake, *Nat. Hazards*, *4*, 647–653.

Sorokin, V. M., V. M. Chmyrev, and A. K. Yaschenko (2001), Electrodynamic model of the lower atmosphere and the ionosphere coupling, *J. Atmos. Terr. Phys.*, *63*, 1681–1691.

Tate, J., and W. Daily (1989), Evidence of electro-seismic phenomena, *Phys. Earth Planet. Inter.*, *57*, 1–2, doi:10.1016/0031-9201(89)90207-0.

Toutain, J.-P., and J.-C. Baubron (1998), Gas geochemistry and seismotectonics: A review, *Tectonophysics*, *304*, 1–27.

Zaliapin, I., Z. Liu, G. Zoller, V. Keilis-Borok, and D. Turcotte (2002), On increase of earthquake correlation length prior to large earthquakes in California, *Comput. Seismol.*, *33*, 141–161.

Zaslavski, Y., M. Parrot, and E. Blanc (1998), Analysis of TEC measurements above active seismic regions, *Phys. Earth Planet. Inter.*, *105*, 219–227.

E. Calais and J. Haase, Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN 47907, USA. (ecalais@purdue.edu; jhaase@purdue.edu)

T. Dautermann, Department of Physics, Purdue University, West Lafayette, IN 47907, USA. (dautermann@purdue.edu)

J. Garrison, School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47907, USA.