JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 115, B04406, doi:10.1029/2009JB006421, 2010

Long-term monitoring of ULF electromagnetic fields at Parkfield, California

K. N. Kappler,[1] H. Frank Morrison,[2] and G. D. Egbert[3]

Received 3 March 2009; revised 13 October 2009; accepted 30 October 2009; published 9 April 2010.

[1] Electric and magnetic fields in the (10^{-4}–1.0) Hz band were monitored at two sites adjacent to the San Andreas Fault near Parkfield and Hollister, California, from 1995 to 2007. A data window (2002–2005), enclosing the 28 September 2004 *M*6 Parkfield earthquake, was analyzed to determine if anomalous electric or magnetic fields or changes in ground conductivity occurred before the earthquake. The data were edited, removing intervals of instrument malfunction, leaving 875 days in the 4 year period. Frequent, spikelike disturbances were common but were not more frequent around the time of the earthquake; these were removed before subsequent processing. Signal-to-noise amplitude spectra, estimated via magnetotelluric processing, showed the behavior of the ultralow frequency fields to be remarkably constant over the period of analysis. These first-order plots make clear that most of the recorded energy is coherent over the spatial extent of the array. Three main statistical techniques were employed to separate local anomalous electrical or magnetic fields from the dominant coherent natural fields: transfer function estimates between components at each site were employed to subtract the dominant field, and look deeper at the "residual" fields; the data were decomposed into principal components to identify the dominant coherent array modes; and the technique of canonical coherences was employed to distinguish anomalous fields which are spatially broad from anomalies which occur at a single site only, and furthermore to distinguish anomalies present in both the electric and magnetic fields from those present in only one field type. Standard remote reference apparent resistivity estimates were generated daily at Parkfield. A significant seasonal component of variability was observed, suggesting local distortion due to variations in near-surface resistance. In all cases, high levels of sensitivity to subtle electromagnetic effects were demonstrated, but no effects were found that can be reasonably characterized as precursors to the Parkfield earthquake.

Citation: Kappler, K. N., H. F. Morrison, and G. D. Egbert (2010), Long-term monitoring of ULF electromagnetic fields at Parkfield, California, *J. Geophys. Res.*, *115*, B04406, doi:10.1029/2009JB006421.

1. Introduction

[2] Many investigators have reported observations of unusual geoelectromagnetic behavior associated with seismic activity [e.g., *Corwin and Morrison*, 1977; *Varotsos and Alexopoulos*, 1984; *Molchanov et al.*, 1992; *Fraser-Smith et al.*, 1990; *Uyeda et al.*, 2000; *Johnston*, 1989, 1997, and references therein; *Johnston et al.*, 2006; *Pulinets and Boyarchuk*, 2004; *Park et al.*, 1993]. The reported anomalies take on a variety of forms, including variations in quasi-static electric fields, self-potentials, ultralow frequency (ULF) magnetic fields, alternating electric fields in the ULF, very low frequency (VLF), or extra low frequency (ELF) bands, and variations in ground resistivity. Although *Park et al.* [1993] concluded (albeit with some reservations) that there did seem to be credible observations of some of these precursory phenomena, there is still vigorous debate in the geophysical community on the validity and scientific merit of these reports [e.g., *Geller*, 1996]. In general, efforts to observe these phenomena have suffered from a lack of multiple earthquake observations to demonstrate a consistent relationship between the observed phenomena and the earthquake. Furthermore, observations have typically been of short duration, making it difficult to tell whether an anomalous observation was truly a useful precursor or whether such anomalies also happen at many times when there is no significant seismic activity.

[3] In 1995, Berkeley researchers installed two electromagnetic (EM) monitoring sites on the San Andreas Fault (SAF) near Parkfield (PKD) and Hollister (SAO), California, shown in Figure 1. The Parkfield segment of the SAF was chosen for a focused prediction experiment on the basis

[1]Berkeley Seismological Laboratory, University of California, Berkeley, California, USA.
[2]Department of Earth and Planetary Science, University of California, Berkeley, California, USA.
[3]College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA.

0148-0227/10/2009JB006421

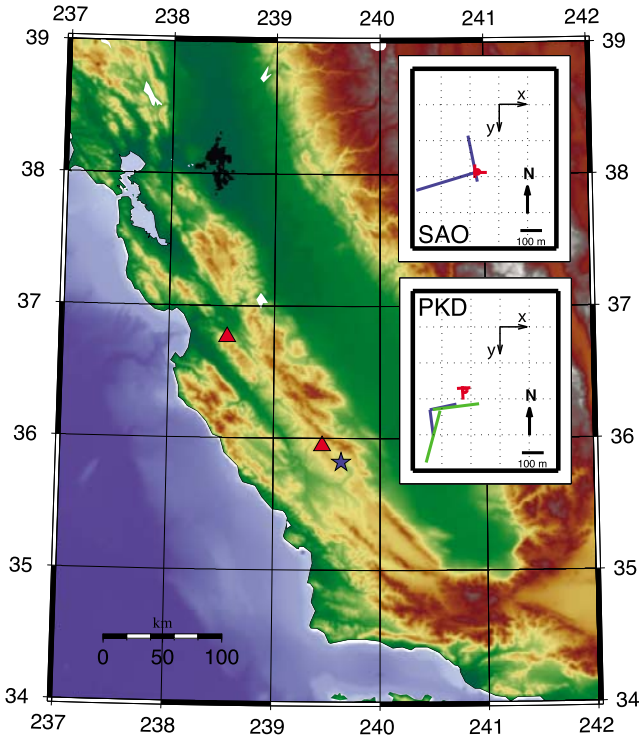

Figure 1. Map illustrating the two site locations. Red triangles denote the magnetotelluric sites. The blue star denotes the epicenter of the 2004 Parkfield earthquake. Inset showing the azimuths of the electrodes and coils. Electrode lengths are shown to scale, whereas the coils are scaled to be visible.

of repeating *M*6 earthquakes, with a recurrence interval of approximately 22 years [*Bakun and McEvilly*, 1984]. By 1995, the Parkfield earthquake was late, but the probability of recurrence was high, making this a good place for an EM monitoring experiment. The site was also chosen for the availability of other geophysical data and experiments relating to the dynamics of the SAF. The expected Parkfield earthquake did not occur until 28 September 2004 (M_w 6), by which time the array had been in operation (although not without interruption) for 9 years. The Parkfield segment has also been monitored with a long dipole array for long period (>300 s) electric fields since 1988 [*Park et al.*, 2003, 2007].

[4] At the Parkfield segment of the strike-slip SAF, a granitic basement is overlain by sedimentary cover to the west and the conductive Franciscan melange to the east. Low-resistivity structures are clearly seen to be coincident with the SAF zone at both PKD and SAO [*Unsworth and Bedrosian*, 2004]. Such variations in conductivity are thought to be controlled by variations in fluid saturation and fracturing. It is unclear whether the presence of fluids is causing the creep and repeated earthquakes along the SAF or whether the fluids are simply indicative of tectonic deformation zones. A discussion of geoelectric structure near Parkfield and its role in fault dynamics is given by *Unsworth and Bedrosian* [2004] and *Becken et al.* [2008] and references therein. Since the primary purpose of this study was to monitor temporal changes in fields and ground resistivity, we have made no attempt to interpret the impedance in the context of local conductivity structure.

[5] The Berkeley EM monitoring array was designed to monitor general fluctuations in the spectra of natural fields in the 10^{-4}–10 Hz band, changes in ground resistivity through the magnetotelluric (MT) impedance tensor, and variations in the amplitude and phase of intersite transfer functions. Each of these properties has been studied using this array by researchers interested in ULF fields in the past. Four prior studies are especially relevant to this paper. *Boyd* [2000] discussed the array signal quality and introduced the idea of looking for subtle electromagnetic fields in the array residual fields, which are calculated by removing the signal which is coherent at both sites from the sensor data. *Egbert et al.* [2000] studied the Parkfield-Hollister intersite transfer function and showed that the array is strongly contaminated by nonplane wave fields in the 10–30 s period. The major sources of this contamination were shown to be the San Francisco Bay Area Rapid Transit (BART) DC electric train and natural continuous geomagnetic micropulsations (Pc3) activity. *Eisel and Egbert* [2002] (hereinafter referred to as EE02) carried out a study of the temporal variation in the robustly calculated impedance tensor, which was estimated daily at PKD for 2 years using data from this array. As for *Egbert et al.* [2000], this study used data from 1996 to 1997, a time window when no earthquakes larger than *M*5 occurred within 50 km of the PKD site. Finally, the array data were briefly examined by *Johnston et al.* [2006], who focused on the possible presence of earthquake precursors immediately before the 28 September 2004 earthquake. Residual plots, computed by comparing the Parkfield data with the remote site SAO for a 3 week section centered on the earthquake, reveal no clear precursors such as the sharp increase in magnetic field amplitude immediately prior to Loma Prieta reported by *Fraser-Smith et al.* [1990].

[6] We build on these efforts in a number of ways. A signal such as the one observed by *Fraser-Smith et al.* [1990] could start and end outside the boundaries of the time window considered by *Johnston et al.* [2006]. We address this problem by plotting the residual time series in daily averages over a much longer window and also provide a time series of residual fields on the day of the earthquake from the band around 0.01 Hz (the frequency which was the main focus of *Fraser-Smith et al.* [1990]). We also report on a number of different analyses of electric and magnetic fields variations in the 5×10^{-4}–1 Hz band, to better characterize the long-term behavior of recorded fields and quantities derived from the recorded fields and to better identify and understand the sources of outliers in the time series. Then we apply the same codes as EE02 but over a longer, 4 year time period, more clearly demonstrating that there is a seasonal variation in the impedance tensor estimates. We also apply a distortion analysis which separates time-dependent variations in the impedance tensor which are frequency independent from those variations which are both time- and frequency-dependent. The frequency-independent changes are likely due to near-surface distortion or changes in calibration when instruments were swapped, whereas frequency-dependent shifts are more likely related to deeper changes in subsurface conductivity.

[7] The array was also originally intended to be used to search for anomalous variations in quasi-DC electric field of

the sort reported by Varotsos and collaborators [e.g., *Varotsos and Alexopoulos*, 1984; *Varotsos and Lazaridou*, 1991]. However, we found many anomalous transients in the electric field channels, including both spikes and box-car-like steps, which were clearly noise. Such spurious transient signals are commonly encountered in ULF-ELF recording and are especially troublesome in MT studies. Common sources are sferics which become distorted through amplifier saturation, and a host of other local effects such as power-line or battery switching transients, nearby electrical equipment, and poorly understood electrochemical reactions in and near the electrodes. Given the prevalence of these sorts of noise in our long-term recordings, we concluded that it would be difficult to verify the presence of so-called Varotsos-Alexopoulos-Nomicos (VAN)-type signals in our data. Instead, we chose to implement a despiking algorithm to remove these high-frequency noise sources (and possibly also some signals) from the 1 Hz data that we focus on here.

[8] The analysis of the data is broken into seven main tasks, as follows.

[9] 1. The first task was removal of spike transients from the raw digitized time series of E and H field observations.

[10] 2. The second task was transformation of despiked records to the spectral domain, with correct scaling for instrument transfer functions and gains.

[11] 3. The third task was application of the robust multivariate errors-in-variables (RMEV) method [*Egbert*, 1997] to separate the spectral time series into coherent signal and incoherent noise, and to estimate the spectral density matrix (SDM) of the array for each day.

[12] 4. The fourth task was subtraction of large spatial scale (mostly natural source) EM fields from Parkfield channels predicted from the EM fields at Hollister, followed by plotting and analysis of residual fields.

[13] 5. The fifth task was the analysis of principal components of the daily SDM. This analysis is based on *Egbert* [1997, 2002] and is intended to examine the dimensionality and signal source characteristics in the array data as a function of time and frequency.

[14] 6. The next task was to use the canonical coherence analysis [*Brillinger*, 1969] to identify signals which are coherent between various collections of array channels.

[15] 7. The final task was unbiased daily calculations of the PKD impedance tensor and an examination of its variation with time.

[16] Unless otherwise noted, analyses are conducted for each day separately and displayed as long-term time series of daily averages. We occasionally use 2 h analysis windows when treating 40 Hz data, to examine residuals, and to present plots of unsmoothed Fourier coefficients over short time windows. Processing is applied to two distinct long time segments. Results for a 4 year section of data provide a view of longer term variations and possible trends in the data. These plots also emphasize the difficulty of obtaining reliable and meaningful results over long time periods which include equipment swaps, malfunctions, and site maintenance. Because of such difficulties, we focus much of our analysis on a 163 day long section from 16 May to 26 October 2004 (Julian days 137–299). This time window is significant because it brackets the 2004 Parkfield *M*6.0 earthquake, with no equipment malfunctions or interruptions for service. Extending for more than 4 months prior to the earthquake, this data segment represents a longer period of uninterrupted ULF monitoring for precursors than has been reported previously.

2. Sites and Instruments

[17] The array consists of two sites separated by 120 km (Figure 1). At each observatory, three orthogonal induction coils (model EMI BF-4) measured the time-varying magnetic field, and two 100 m long electrodes measured the electric field in the surface plane. The horizontal coils were buried in trenches 0.5 m deep, and the vertical coils were in drilled holes approximately 2 m deep. The electrodes were Pb-PbCl nonpolarizing type, placed in 3 m holes with moist bentonite packing to keep contact resistance to a minimum. The entire system was powered by 12 V batteries with constant trickle charge provided by on-site solar cells. The dipole data were preprocessed with an electric field signal conditioner (EFSC), consisting of a preamplifier, optical isolator, and main amplifier in series, with an optional high-pass filter. The preamplifier can be set at 10, 20, 30, or 40 dB of gain. The data were then digitized by 24 bit Quanterra digitizers (Q935 at PKD and Q4120 at SAO), at a sampling rate of 40 Hz, with time synchronization maintained via GPS. The data were then telemetered in packets to the Northern California Earthquake Data Center (NCEDC), where they are archived and are available for public download. A schematic diagram of the site instrumentation is provided in Figure 2. At PKD, an added independent pair of 200 m dipoles collected data alongside the 100 m pair. This is useful for recognizing when an electrode is creating voltage noise, or for checking linearity of measurements, and is required should one want to run VAN method analysis [*Varotsos and Lazaridou*, 1991]. For the analysis reported here we use 8 of the 12 available channels: the horizontal magnetic coils together with the 100 m electrodes at the two sites. The 200 m electrodes and vertical magnetometers were omitted because preliminary analysis of signal-to-noise ratios (SNR) indicated that these channels were frequently severely corrupted.

3. Data Processing

3.1. Data Selection and Cleaning

[18] The time window of analysis for this study spans 4 years, a total of 1461 days. Earlier data were omitted because of extended periods of instrument malfunction at the sites. Although the entire data set was recorded at both 1 Hz and 40 Hz, unless otherwise stated, we focus on the 1 Hz data. The raw data are stored as $8 \times 86{,}400$ point (day-long) arrays, in units of datalogger machine counts. Before the main processing, a data selection filter was applied to omit from further analysis any days on which the array was not functioning adequately.

[19] To implement this filter, each day was divided into 450 equal-width time windows and the variance of each channel was calculated in each window. Ratios of these variances at coincident times were calculated for corresponding fields at the two sites, e.g. the ratio of Ex at PKD to Ex at SAO, or the ratio of Hy at PKD to Hy at SAO. A window was considered contaminated by spikes if the $\log_{10}$ variance ratios

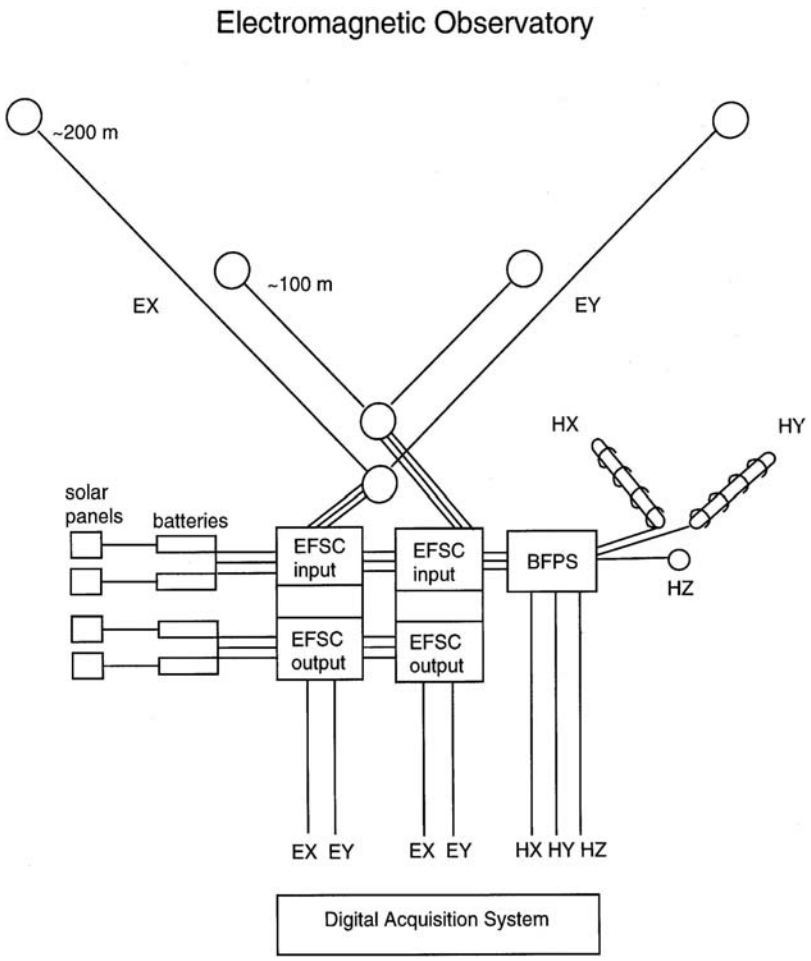

Figure 2. Schematic of the electromagnetic observatory.

(corrected for amplifier gain) were found to be unnaturally high (deviated by more than 1 from the corresponding 4 year median ratio). Windows with one or more missing channels were also considered contaminated. If more than 10% of the time windows for a given day were contaminated by missing data or spikes, then that day was removed from the analysis. A total of 586 days were omitted this way. The remaining 875 days were treated with a time domain MT data despiking algorithm (K. N. Kappler, unpublished data, 2009). In this process, around 3500 spikes were identified in the remaining data and replaced by predictions computed from the simultaneous noncorrupt data from other channels using a Wiener filter. This corresponds to around 0.1% of the data windows being contaminated. The coseismic signal which accompanies the ground motion at either site [*Kappler et al.*, 2006] was identified by this process and removed, but because the coseismic energy is present in all channels, the removal is imperfect, and some trace signal remains in the time series (see *Kappler* [2008; see also Kappler, unpublished data, 2009] for further details on the despiking process). An example of application of the despiking algorithm is illustrated in Figure 3. The distribution over time of the spikes and gaps identified is shown as a sum over all channels in Figure 4. Of the remaining 875 days there do not seem to be any periods of time which are especially prone to spikes, nor is there any evident relationship between spike occurrence and the *M*6 earthquake.

3.2. Harmonic Representation

[20] The despiked data were fast-Fourier-transformed using 256 point Hamming windows with a 64 point overlap. Long-period Fourier coefficients (FCs) are obtained by repeatedly decimating the data by a factor of two and reapplying the same windowing scheme. This results in six decimation levels, spanning periods from around 3 to 1500 s. The resulting daily harmonic time series of FCs have lengths 449, 224, 112, 55, 27 and 13, for decimation levels 1–6, respectively. During the FC calculation, the manufacturer's instrument transfer functions are used to correct for amplifiers in EFSCs and the frequency-dependent response of the coils, so that FC time series data are expressed in the standard units of MT data: $\mathrm{mV/km/\sqrt{Hz}}$ for electrodes and $\mathrm{nT/\sqrt{Hz}}$ for coils. Although processing is applied to each harmonic independently, most of the results are further band-averaged before plotting. For this 32 distinct bands were selected, each having a constant Q.

3.3. Spectral Density Matrices and Signal-to-Noise Ratios

[21] When dealing with multiple channel data, an important quantity is a table of the cross power spectra between

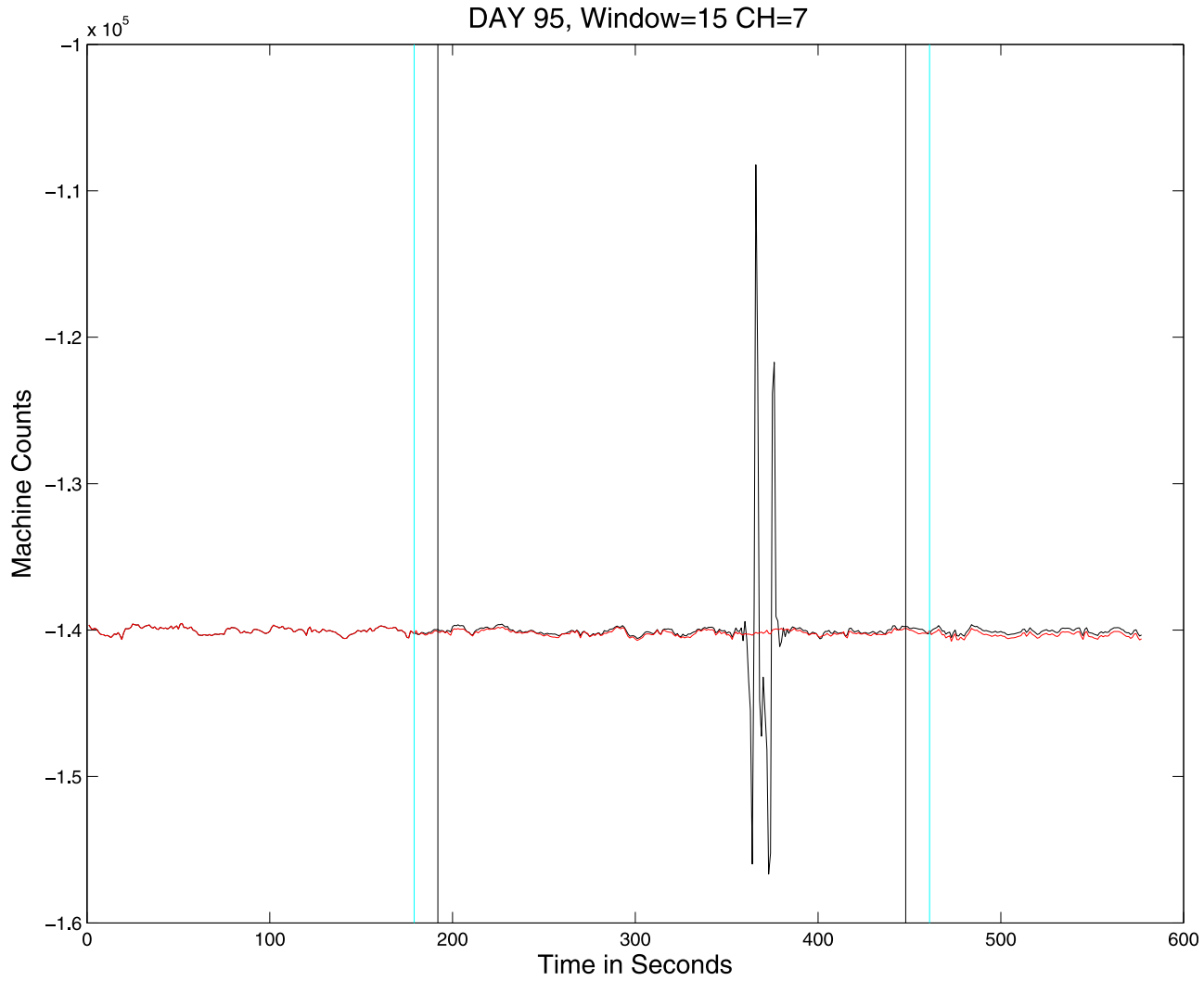

Figure 3. An example of the despiking routine results. The recorded data are in black, and the plausible data calculated by the method outlined in the text are shown in red. Black vertical lines bound the clipped window, and in the regions bounded between the cyan and black vertical lines a weighted average of the predicted and recorded data is taken to "splice" the two time series together.

the different channels, i.e., the SDM. Consider an N channel instrument array in which the time series of complex FCs (harmonics) is calculated as prescribed in section 3.2 for each of the 1, 2,… N channels and stored as $\mathbf{x}_1(t)$, $\mathbf{x}_2(t)$,…, $\mathbf{x}_N(t)$ $t = 1,\ldots, T$. The SDM is simply an $N \times N$ complex-valued matrix $\mathbf{S}$ such that $S_{i,j} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$, where the angle brackets denote averaging over the T segments, i.e., the i^{th} diagonal element of the SDM is simply the time-averaged autopower of the i^{th} sensor, and $S_{i,j}$ is the cross-power between channels i and j. Alternatively, if $\mathbf{X}$ is the $N \times T$ matrix of FCs, then we have

$$\mathbf{S} = \frac{\mathbf{X}\mathbf{X}^*}{T}, \tag{1}$$

where the asterisk denotes the complex conjugate transpose.

[22] Although Equation 1 conveys the idea behind the SDM, the method of calculation described is only appropriate when the data are outlier free. In reality, a more robust calculation procedure is needed. A variant of the RMEV frequency domain cleaning algorithm is thus employed to generate estimates of SDMs separately for each day and for each FC harmonic. This scheme employs an iterative procedure for alternately downweighting multivariate outliers, and cleaning up isolated single-channel outliers in the time series of FCs. Multivariate outliers are detected and downweighted using the affinely invariant approach of *Huber* [1981]. To downweight outliers in individual channels, each channel's FC time series is fit by robust least squares to all other array channels, and FCs which have large residuals are pulled toward predicted values. These residual variances serve as initial, but imperfect estimates of channel noise variances, as they also incorporate noise from other channels. This contamination is reduced by an updating step which approximately decouples the mixed incoherent channel noise estimates from one another [see *Egbert*, 1997]. A byproduct of the RMEV calculation of the SDM is thus an estimate of incoherent noise power (i.e., the fraction of the measured signal in each channel that is incoherent with all

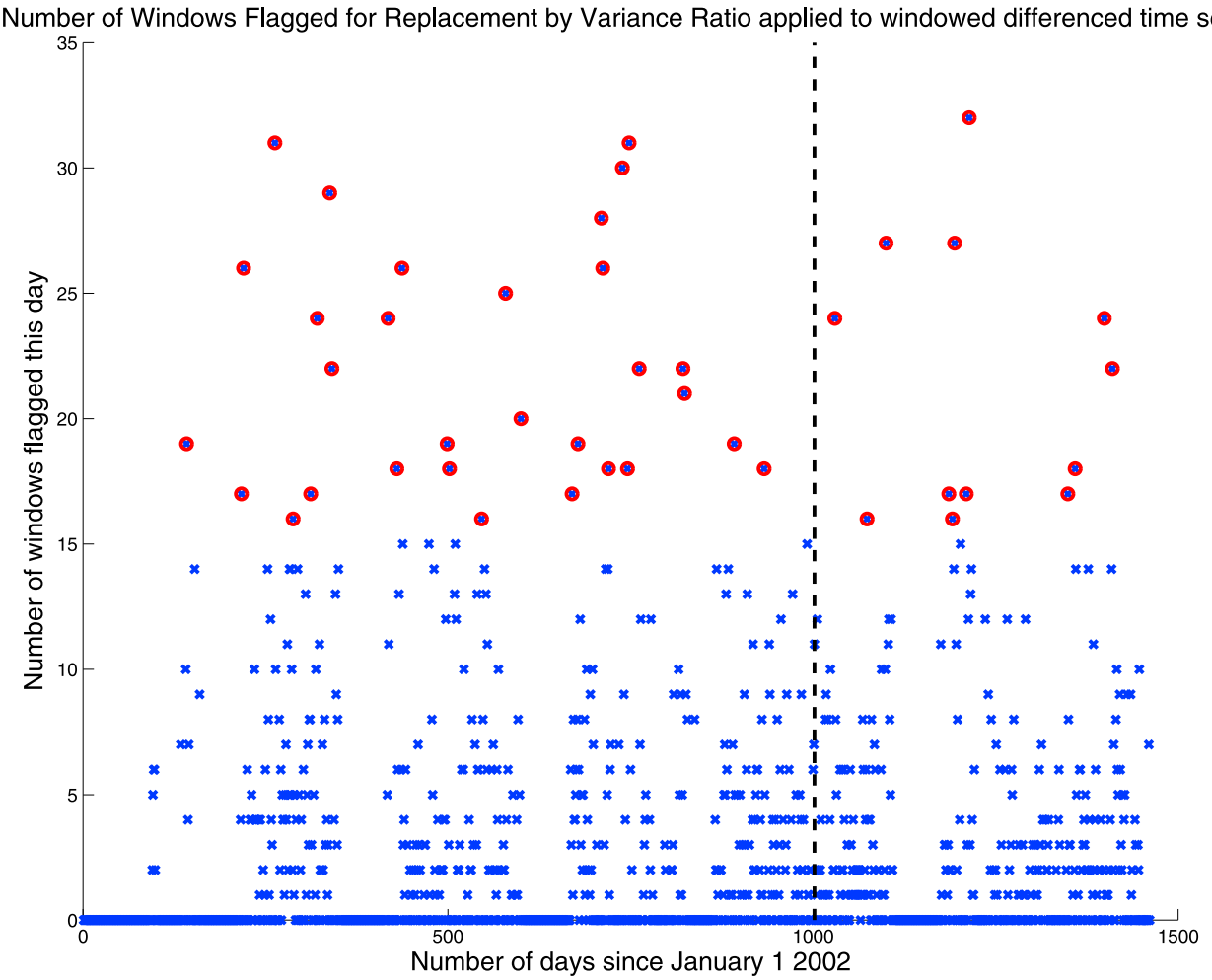

Figure 4. Distribution of flagged windows for all channels versus time. The *y* axis shows the number of spikes on a given day, whereas the *x* axis is an index of the days. The 95th percentile of these points is shown in red. The vertical black dashed line denotes the date of the 28 September 2004 Parkfield *M*6 earthquake.

other channels in the array), thus allowing the recorded fields to be expressed in terms of SNR (see *Egbert* [1997, and references therein] for further details on the RMEV estimation scheme).

[23] The robust estimates of the SDM, together with the estimates of incoherent noise levels, serve as the basis for our initial discussion of signal and incoherent noise. We define the signal power in channel i as the i^{th} diagonal element of the robustly estimated SDM minus the estimate of incoherent noise variance. We define the average daily signal amplitude as the positive square root of the average daily signal power. The effective instrument noise level is taken to be the square root incoherent noise variance returned by the last iteration of the RMEV scheme. We estimate noise for a whole band as the mean of the noise levels at each harmonic in the band.

[24] Once robust estimates of the SDM and individual channel noise levels are obtained, we form a scaled version of the SDM for each day. This SDM $\mathbf{S}^{\sigma}$ is formed as in equation (1), but using the final cleaned matrix of FC time series $\mathbf{X}$, with the i^{th} row (channel) normalized by the daily incoherent noise standard deviation for the i^{th} channel σ_i.

3.4. Residual Fields

[25] Much of the recorded data results from large-scale external sources and is thus coherent between the two sites. To the extent that this is true, the field components at PKD are related to the observed fields at SAO through a time invariant complex-valued intersite transfer function G:

$$\mathbf{D}_{\text{PKD}}(t) = \mathbf{G}\mathbf{M}_{\text{SAO}}(t), \tag{2}$$

where $\mathbf{D}$ is a multivariate time series of FCs for some channels at PKD, and the N-vector valued time series of FCs for some N channels (where N defines the dimensionality of the sources) at SAO makes up the matrix $\mathbf{M}$. If sources are truly spatially uniform (as in the classical MT assumption), it would be sufficient to use only $N = 2$ channels at SAO to capture all of the natural field variability at PKD. In reality, the source fields are more complicated, due to both complications in the natural source and large-scale cultural noise (BART in our case [*Egbert et al.*, 2000]). Because of these

complications we use all ($N = 4$) of the standard SAO channels for prediction and compute the residual times series or local 'site noise' $\mathbf{R}(t)$ as

$$\mathbf{R}(t) = \mathbf{D}_{\mathrm{PKD}}(t) - \mathbf{G}\mathbf{M}_{\mathrm{SAO}}(t). \tag{3}$$

[26] Note that the data $\mathbf{D}$, input to the residual calculation, are not downweighted for outliers, but rather they are the raw time series of FCs calculated from the time series of despiked data. If there were a precursory signal near Parkfield, it should be more clearly visible against a background of residuals calculated with this TF rather than the background of combined site noise plus MT field.

[27] Estimation of the intersite transfer function may be biased by uncorrelated noise at SAO [e.g., *Egbert*, 2002]. In principle, an unbiased estimate could be obtained using a third reference site, which is presumed to see the same naturally occurring fields, but where noise is uncorrelated with the first two. Unfortunately, data were obtained at three sites simultaneously only during the 1 month period of February 1999 and so this approach cannot be used for the long-term analysis reported here.

[28] The robust SDMs are used to calculate the TFs for residual calculations, rather than estimating these directly from regression on the time series of FCs. This has the effect of stabilizing the TF estimates against outliers. Consider an arbitrary but fixed frequency bin centered at ω. Let one group of channels (for example the channels at PKD) be represented as the time series of FCs $\mathbf{D}(t)$ (in this case $\mathbf{D}(t)$ has dimension $4 \times T$), and similarly the other group as $\mathbf{M}(t)$ (in this case the channels at SAO). The channels can be ordered so that the SDM at the frequency ω is approximately

$$\mathbf{S} \approx \frac{1}{T}\left[\begin{array}{c|c} \mathbf{DD}^* & \mathbf{DM}^* \\ \hline \mathbf{M}^*\mathbf{D} & \mathbf{MM}^* \end{array}\right] = \frac{1}{T}\left[\begin{array}{c|c} \mathbf{S}_{1,1} & \mathbf{S}_{1,2} \\ \hline \mathbf{S}_{2,1} & \mathbf{S}_{2,2} \end{array}\right], \tag{4}$$

where approximate, instead of exact, equality holds due to the robust SDM calculation. Thus, the transfer function between the channels in $\mathbf{D}$ and the channels in $\mathbf{M}$ can be expressed in terms of the partitioned SDM as

$$\mathbf{G} = \mathbf{S}_{1,2}{\mathbf{S}_{2,2}}^{-1}, \tag{5}$$

where each row of $\mathbf{G}$ corresponds to a TF between one Parkfield channel and all remote (Hollister) channels. An SDM is calculated separately on each day for each FC. To compute residuals for the 163 day time window bracketing the earthquake in 2004, we use an average TF computed from the average SDM obtained by stacking the daily SDMs over the first 3 weeks of the 163 day section (days 114–135 or 2004). Thus, the same TF is used for all residual calculations so that any slowly evolving trend in the intersite TF will not be subtracted out, as would be the case if an adaptive TF for each day were used for the residual calculation. We are implicitly assuming that the 3 week time window used for calculation of the TFs is not contaminated by any earthquake precursor signals.

3.5. SDM Eigenmodes

[29] Once expressed in units of SNR, the number of eigenvalues of the SDM significantly greater than 1 is an indicator of the number of uncorrelated sources present in the array data, as discussed by *Egbert and Booker* [1989] and *Egbert et al.* [2000]. If a process with an electromagnetic signature were occurring in association with the earthquake, it would likely be uncorrelated with the ambient MT variations, but still produce a coherent multichannel signal at the local site, and hence should appear in the data as an added significant eigenvalue of the SDM. For an arbitrary but fixed combination of day t and frequency band ω, the daily normalized averaged SDM $\mathbf{S}^{\sigma}_{\omega}(t)$ is decomposed as

$$\mathbf{S}^{\sigma}_{\omega}(t) = \mathbf{U}_{\omega}(t)\mathbf{D}_{\omega}(t)\mathbf{U}^{*}_{\omega}(t), \tag{6}$$

where $\mathbf{U}$ is the unitary matrix whose columns are the normalized eigenvectors of $\mathbf{S}^{\sigma}$, and $\mathbf{D}$ is a diagonal matrix comprising the eigenvalues of $\mathbf{S}^{\sigma}$. With $\mathbf{D}$ ordered so that the diagonal entries λ_i^2 are descending, the n^{th} dominant mode of the data distribution which generated $\mathbf{S}^{\sigma}$ is given by the n^{th} column of $\mathbf{U}$ denoted by $\mathbf{u}_{:,n}$. The coordinate system in which u is described is $\mathbb{C}^8$, such that each axis is associated with one of the eight sensors, and where one graduation on the i^{th} axis equals the standard deviation of the noise of the i^{th} channel. The nth dominant mode M_n of the daily normalized time series $\mathbf{X}^{\sigma}(t)$ is described by the time series

$$M_n(t) = \sum_{i=1}^{8} u_{i,n}\mathbf{X}^{\sigma}_{i,:}(t), \tag{7}$$

where $u_{i,n}$ is the i^{th} element of the n^{th} column of $\mathbf{U}$, and $\mathbf{X}^{\sigma}_{i,:}(t)$ denotes the i^{th} noise normalized instrument channel making up $\mathbf{X}^{\sigma}$.

[30] If the ambient field was truly two-dimensional, there would only be two eigenvalues of the SDM above 0 dB. There are typically four eigenvalues present in the array data with amplitudes significantly above the reference incoherent noise level. At least some of the energy in eigenvalues 3 and 4 is noise from the BART DC electric trains [*Egbert et al.*, 2000].

[31] The principal axes of the SDM for a particular harmonic will be different from day to day. For example, during times of intense magnetic storms we would expect a different set of principal axes than that during quiet times. In order to get a stable estimate of the typical data distribution, we stack together 3 weeks of robustly calculated SDMs from the first 21 days of our 163 day window. These 3 weeks are free of both solar storms and significant earthquake activity. The principle axes of the stacked SDM can be interpreted as the average modes of the data distribution. These can be used as used by *Egbert and Booker* [1989] to project the full time series of FCs onto a fixed set of modes, each of which corresponds to a fixed spatial pattern of EM fields.

3.6. Canonical Coherences

[32] Canonical coherence analysis (CCA), sometimes referred to as the method of canonical correlations [*Brillinger*, 1969; *Hardle and Simar*, 2007], has been applied to long-term geophysical monitoring by *Lyubushin* [1998]. Here the method is used to differentiate between signals which are common to both sites versus signals which

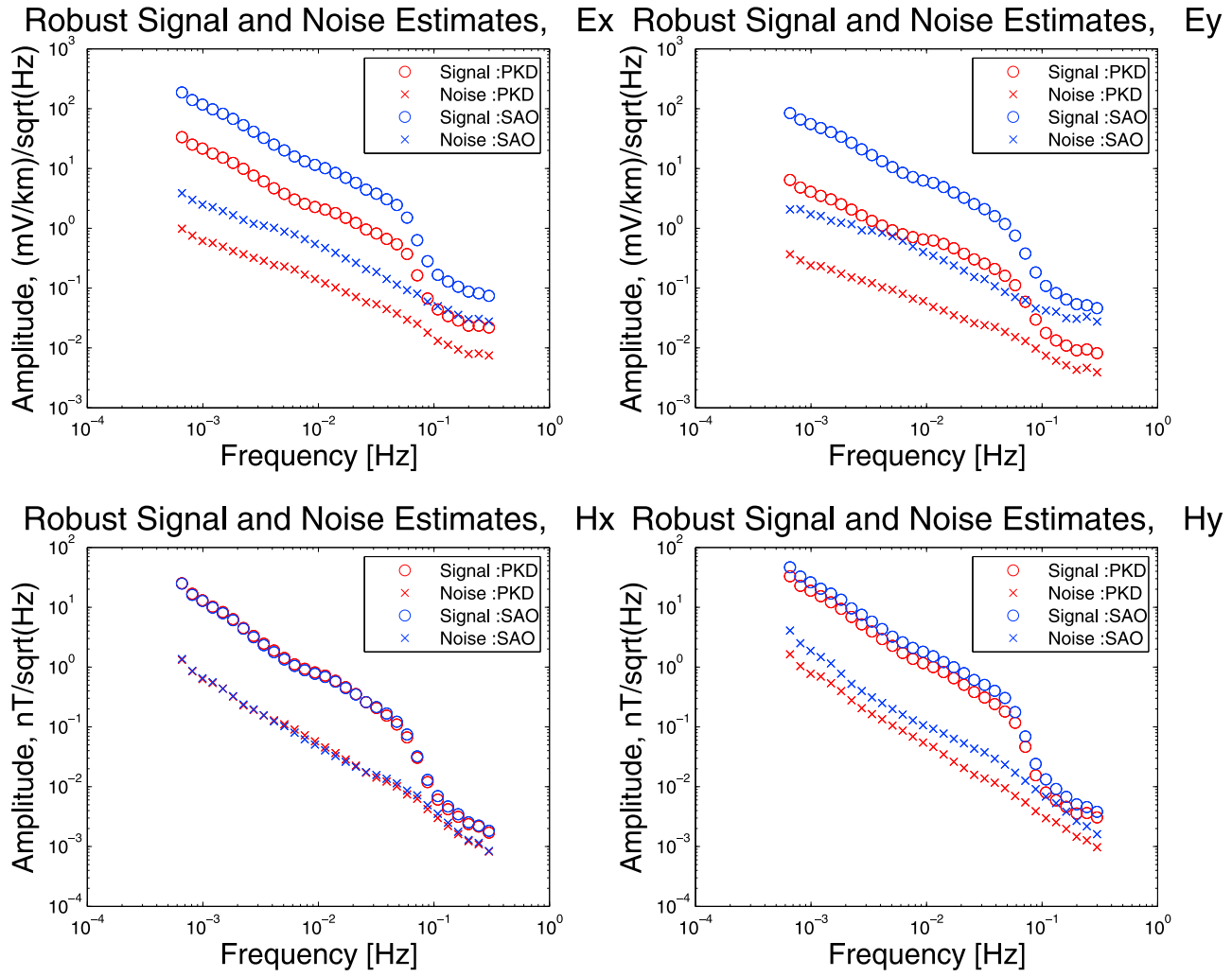

Figure 5. Signal and noise median amplitudes as a function of frequency calculated over the whole 4 year interval.

are present at one site only. This analysis complements the SDM analysis in which all array channels are lumped together, leaving ambiguity over the channels which are the sources of various signals.

[33] The CCA technique identifies linear combinations of array channels which are maximally correlated. A time series of FCs for all array channels $\mathbf{Z}(t)$ is broken into two disjoint channel groups $\mathbf{X}(t)$ and $\mathbf{Y}(t)$. We seek to best approximate the behaviour of a linear combination of channels $\mathbf{X}(t)$ using a linear combination of channels in the other group $\mathbf{Y}(t)$, where x and y denote the number of channels in each group, and without loss of generality, $x \leq y$. Then, there exist x ordered pairs of row vectors $\{\mathbf{p}_1, \mathbf{s}_1\} \ldots \{\mathbf{p}_x, \mathbf{s}_x\}$, where $\mathbf{p}_i$ and $\mathbf{s}_i$ are of dimensions x and y, respectively with the following two properties: (1) the correlation between the canonical variates $\mathbf{p}_i^*\mathbf{X}$ and $\mathbf{s}_i^*\mathbf{Y}$ is maximal in the space defined by the projections of $\mathbf{X}(t)$ and $\mathbf{Y}(t)$ orthogonal to the first $i-1$ canonical variates (for $i = 1$ no projection is required) and (2) the canonical variates within $\mathbf{X}$ or $\mathbf{Y}$ are uncorrelated, i.e., $\langle \mathbf{p}^*_i\mathbf{X}, \mathbf{p}^*_j\mathbf{X} \rangle = 0 = \langle \mathbf{s}^*_i\mathbf{Y}, \mathbf{s}^*_j\mathbf{Y} \rangle$ for $i \neq j$.

[34] Canonical coherences can be obtained from the robust SDMs by computing the ordered eigenvalues of the matrix

$$\mathbf{S}_{XX}^{-1/2}\mathbf{S}_{XY}\mathbf{S}_{YY}\mathbf{S}_{YX}\mathbf{S}_{XX}^{-1/2} \tag{8}$$

using the submatrices of the SDM arising from input time series partitioned into two groupings as in equation (4). For further details, see *Kappler* [2008].

3.7. Estimation of the Magnetotelluric Impedance Tensor and Apparent Resistivity

[35] The MT impedance tensor $\mathbf{Z}$ relates co-located $\mathbf{E}$ and $\mathbf{H}$ fields

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix} \tag{9}$$

under the usual MT assumption that incident fields are plane waves of infinite horizontal extent. This is a reasonable approximation provided EM skin depths are small compared to source length scales, as generally holds for the period

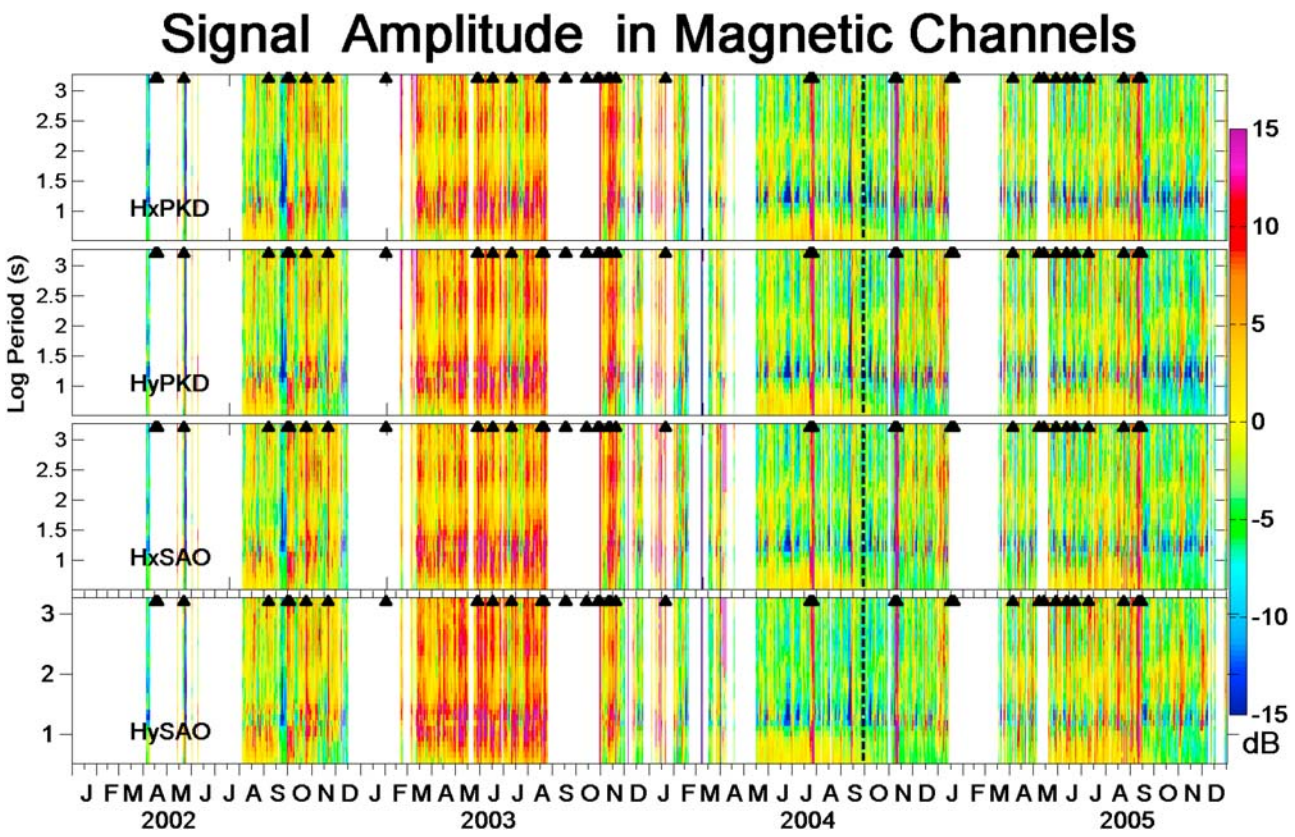

Figure 6. Median normalized signal amplitudes as a function of period calculated over the whole 4 year interval. Black triangles mark days of major geomagnetic storms.

range and site latitudes considered here. The off-diagonal elements of the MT impedance tensor can be converted to an apparent resistivity in the usual way:

$$\rho_{ij} = \frac{1}{\mu\omega}\|Z_{ij}^2\| \qquad \text{for} \quad i \neq j. \tag{10}$$

We used a remote reference processing scheme [*Gamble et al.*, 1979], modified to be robust as outlined by EE02, to compute impedance tensor estimates for each day of data, allowing us to track time- and frequency-dependent variations in the apparent resistivity.

4. Results

4.1. Variation in Signal and Noise Over 4 Years

[36] The median signal and noise amplitude spectra for all channels over the full 4 year period are shown in Figure 5. These exhibit a $1/f^{\alpha}$-type frequency dependence for noise, and signal levels exhibit a similar pattern but show a significant drop in the "dead" band (0.1–1.0 Hz). Note that the amplitude of the electric fields at SAO is significantly larger than that at PKD, suggesting that the ground there is much more resistive than that at PKD. This is consistent with the site geologies. It is not understood why the Hy SAO coil has a higher signal and noise level (roughly a factor of 2) across all bands. This may result from a bad calibration file for the Hy coil, which was never swapped out over all 4 years.

[37] Dynamic behaviour of median-normalized signal amplitude is shown in Figure 6. Figure 6 shows several broadband pulses, narrow in time, where signal levels were suddenly unusually large. For example, these can be seen in July and November of 2004. These pulses correspond to days of anomalously high global geomagnetic activity. NOAA classifies days where the A_p index is greater than 49 as major geomagnetic storms, and $Ap > 29$ as minor geomagnetic storms. Days when major storms occurred are denoted by black triangles in Figure 6.

[38] We find it interesting that the signal levels in all channels seem high in 2003. In order to determine if this is an artifact of signal processing, we looked at the magnetic field strengths calculated directly from the time series of FCs with no RMEV processing applied, and found the same result. The fact that the electric fields exhibit a similar increase in signal strength during this interval suggests that this variation in signal strength is real.

4.2. Variation of Signal and Noise Over 6 Months

[39] To further explore signal and noise characteristics around the time of the PKD *M*6.0 earthquake, we zoom in on the 163 day section from 16 May to 26 October 2004. A more complete set of plots for all channels and instruments is available from *Kappler* [2008]. Figures 7 and 8 show the signal amplitude in the magnetic sensors and the SNR in dB for this time window, respectively.

[40] On the scale showing the full dynamic range of signal amplitude over all frequencies (Figure 7), the fields appear

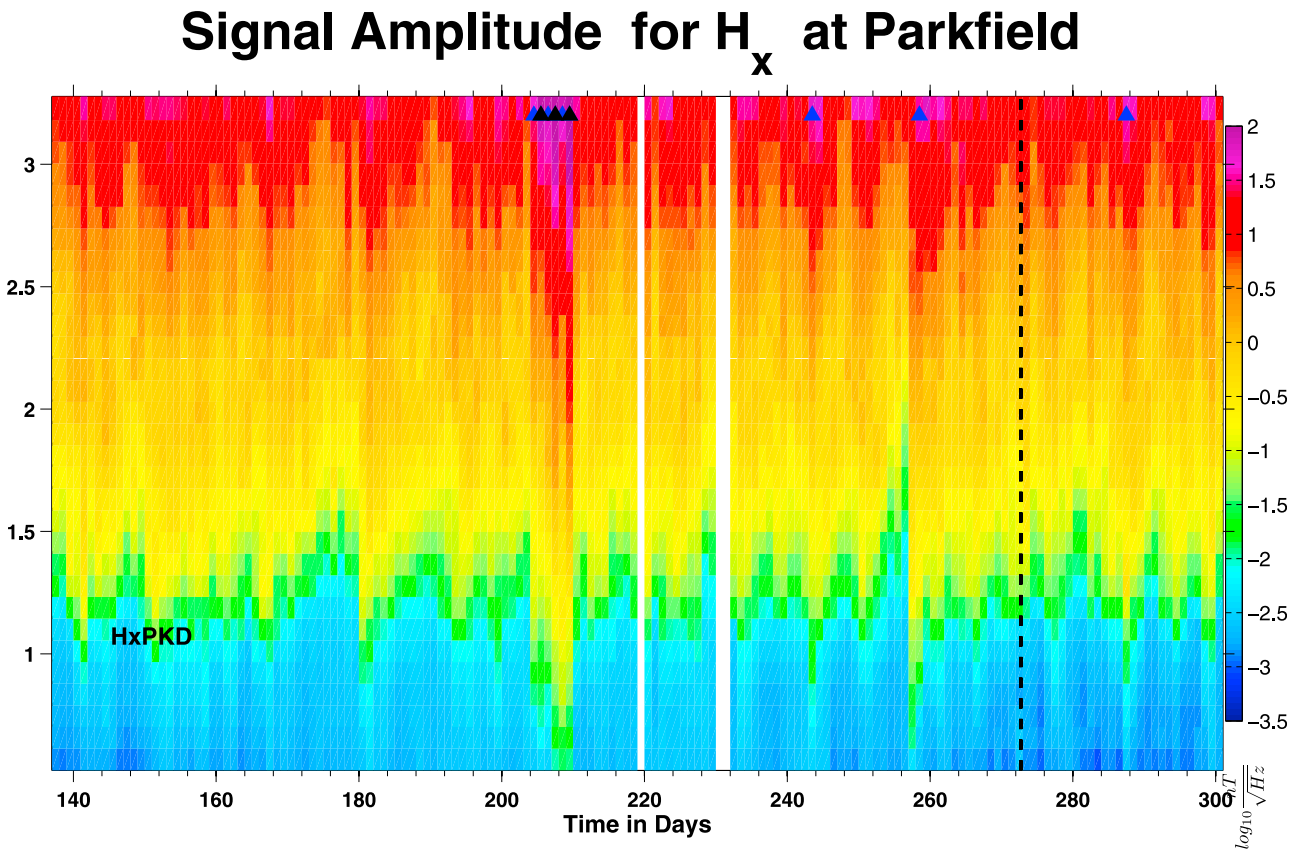

Figure 7. Signal amplitudes in $\log_{10}(\mathrm{nT}/\sqrt{\mathrm{Hz}})$. Black triangles mark days of major geomagnetic storms, and blue triangles mark minor storms.

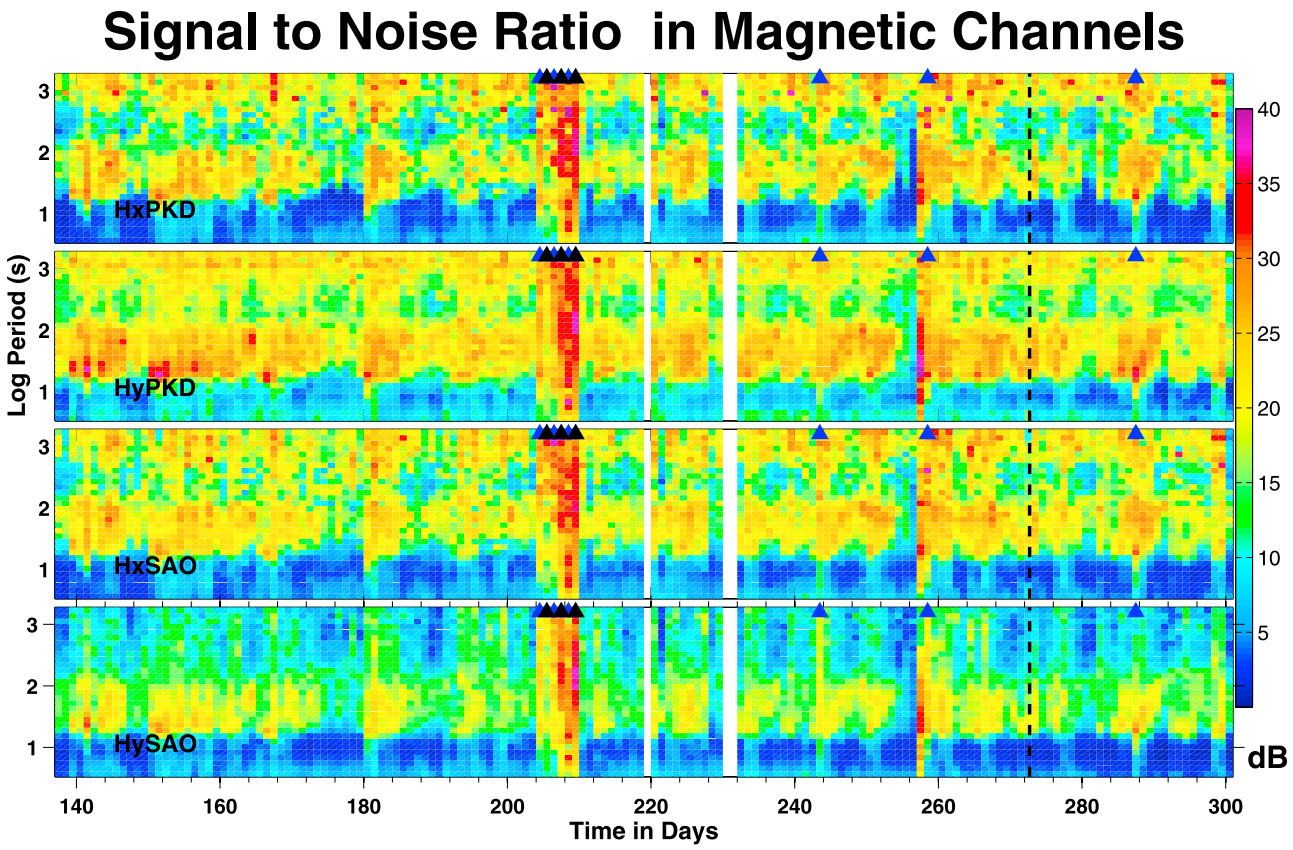

Figure 8. Daily median SNR in decibels for magnetic channels.

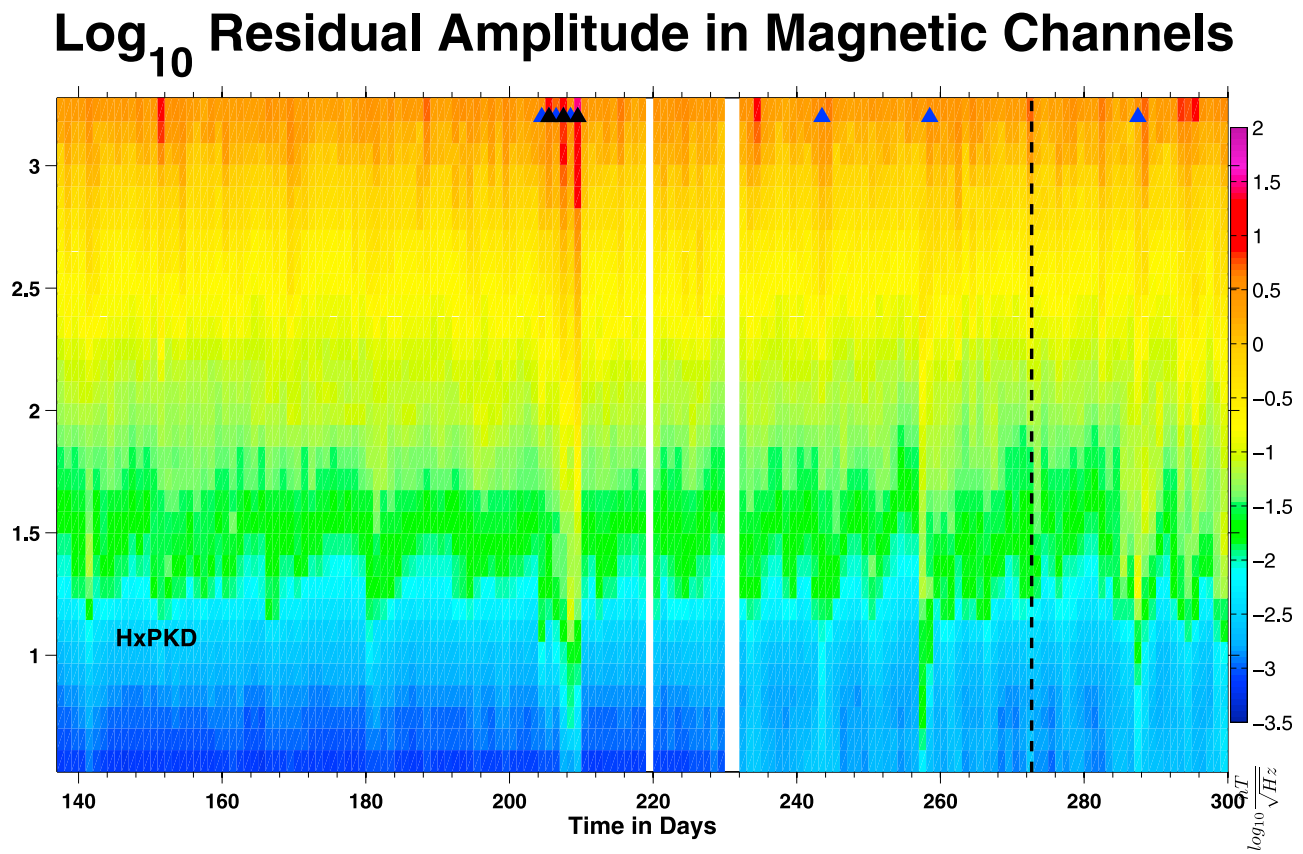

Figure 9. Residual signal amplitudes in $\log_{10}(\mathrm{nT}/\sqrt{\mathrm{Hz}})$. Black triangles mark days of major geomagnetic storms, and blue triangles mark minor storms. Color scale is kept the same as in Figure 7.

remarkably stable. SNR plots (Figure 8) more clearly reveal temporal variations in these spectra. The SNR is clearly linked to geomagnetic storm activity, as shown by the correlation between days of high geomagnetic activity (black and blue triangles) and high SNR. These are ruled out as being earthquake related for two reasons. First, they correlate with days of high A_p geomagnetic activity index, and second, they can be observed at both sites.

[41] These plots may not be ideally suited to the search for EM-earthquake anomalies as they are made from data which downweight outliers. However, they do illustrate that to first order the instrument record shows only the broadband, natural, spatially correlated ULF fields. In order to look for signals originating in the subsurface, the components of the signal which are correlated between sites must first be subtracted from the data.

4.3. Residual Fields

[42] Residual spectra, calculated from residual time series as discussed in Section 3.4, are plotted for magnetic components in Figure 9, on the same scale used for Figure 7. Note the highly stationary appearance of these fields. Broadband amplitude anomalies can be seen to correlate with days of high geomagnetic activity. There appears to be an anomalous change in the residual field amplitudes in the shortest period bands, beginning on day 230, suspiciously immediately following a gap at SAO in available data. These anomalous signals are dangerously close the Nyquist period of 2 s and could possibly result from the instrument/acquisition systems, rather than actual variations in the fields.

[43] Residuals spectra calculated daily over the 2 h window 0000-0200H PST (as opposed to for the full day), with no band averaging over the narrow range of frequencies which encompass the three shortest period bands, are shown in Figure 10. The residual amplitude change is evident in these data, and is relatively broadband. An additional anomaly at 4.1 s period shows itself to be extremely narrow (only 3FCs wide at $df = 1/256$).

[44] If a real signal caused either of these anomalies, it would be visible in the 40 Hz data where the sampling is adequate to clearly resolve signals at these periods. Calculating the residuals (for each FC) from the 40 Hz data over the same 2 h window, the broadband phenomena are not present (Figure 11). We also decimated the 40 Hz data down to 1 Hz and found no rise in the residual spectra of the highest frequency bands. This suggests that the broad anomaly is related to the decimation to 1 Hz done by the Quanterra and is possibly related to a change in anti-alias filter of the data logger after the gap.

[45] Although Figure 9 shows that there was no long-term rise in magnetic field before the earthquake, we also consider whether there may have been a sharp spike such as that reported by *Fraser-Smith et al.* [1990] around 3 h prior to

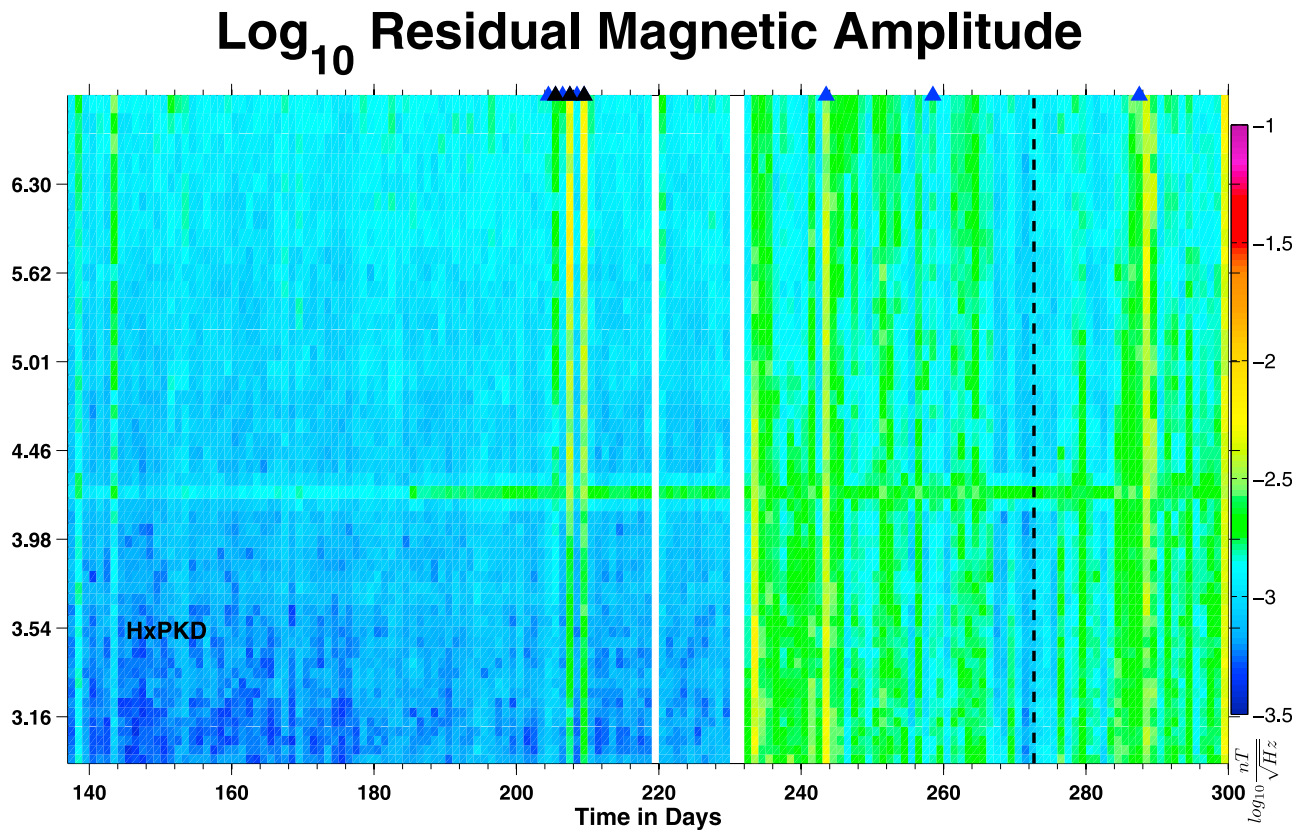

Figure 10. Residual signal amplitudes in $\log_{10}(\mathrm{nT}/\sqrt{\mathrm{Hz}})$ for the lowermost 81 Fourier coefficients.

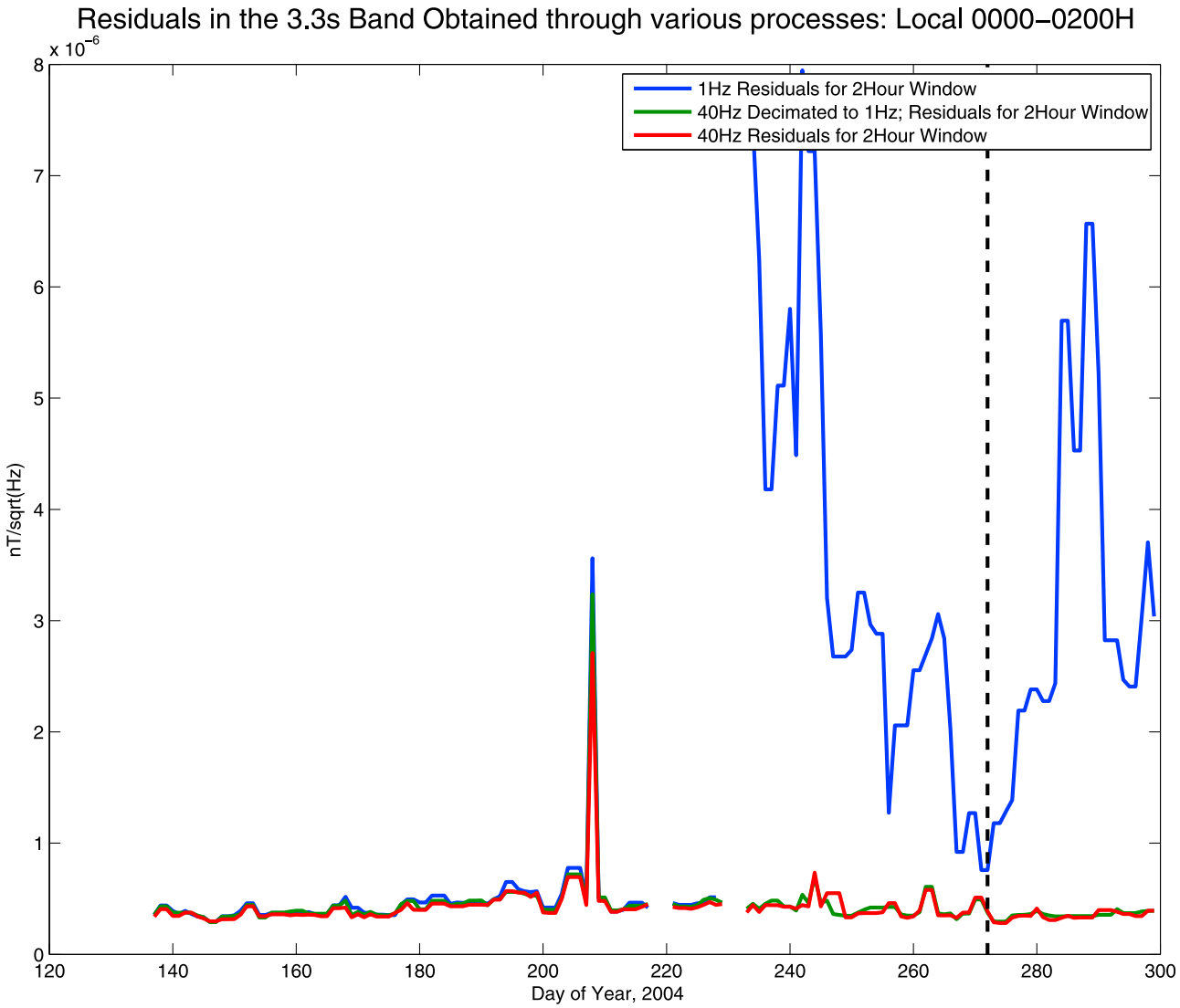

Figure 11. Residuals in cases: 1 Hz, 40 Hz, and 40 Hz decimated to 1 Hz.

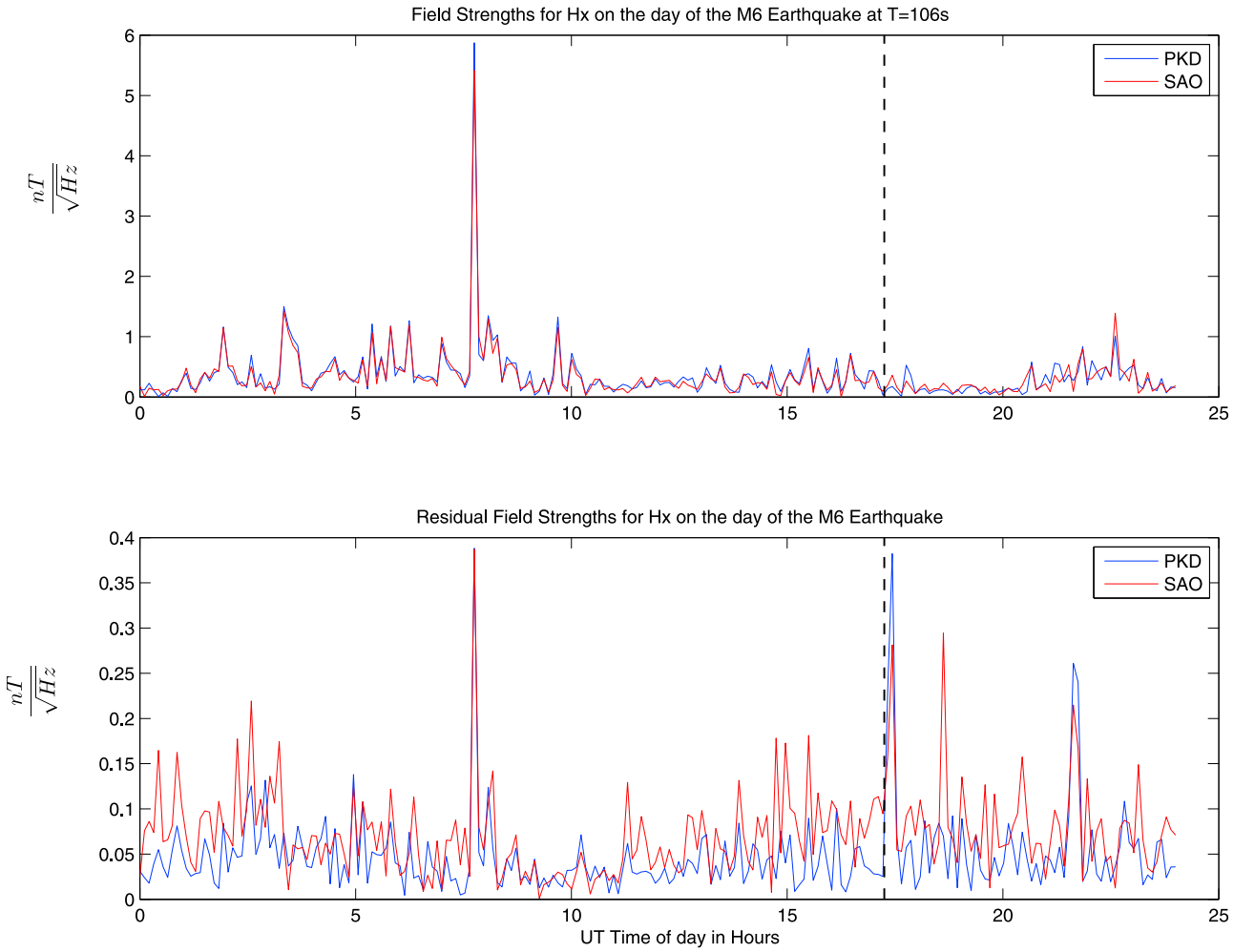

Figure 12. Ambient and residual fields at PKD and SAO for Hx on the day of the earthquake (28 September 2004, day of year 272).

the event in their "MA3" index. Figure 12 shows band-passed ambient and residual field amplitudes for Hx on the day of the earthquake, for a band which is effectively MA3.

[46] There is no activity unique to the PKD site before the earthquake. Note that by plotting the residuals we should be able to detect phenomena an order of magnitude smaller at this frequency. The importance of monitoring at more than one site is again underscored in Figure 12—with data only recorded at PKD the spike in activity at around 0800H UT may have been interpreted as a precursor. We know it is not local because it was seen with equal amplitude at a site 120 km away.

[47] There is an underlying premise in this study that anomalous electric or magnetic fields are generated at or near the earthquake hypocenter. Fields from such sources would be attenuated rapidly and thus should have much lower amplitudes at distances of 12 times the hypocentral depths (i.e., at site SAO). Furthermore, the magnetic-electric relationship in the near field of a source would be different from the plane wave relationship of the dominant fields common to two separated sites. This relationship of fields due to an earthquake source would manifest itself in the eigenmode analysis even if the fields were present at both sites.

4.4. Eigenvalues of the Spectral Density Matrix

[48] If the natural source (MT) fields were truly plane wave, there would only be two eigenvalues of the SDM above 0 dB. The first four ordered eigenvalues of **S** are plotted for the 163 day window around the earthquake in Figure 13. This plot shows that not only are there two dominant sources of energy (mostly the MT field) but also a nontrivial third source is clearly present, especially in the bands around 30–100 s. At least some of the energy in the third eigenvalue is noise from the BART DC electric trains, as shown by *Egbert et al.* [2000]. On days when there are geomagnetic storms (shown by the black and blue triangles), more energy can be seen in the lower-order modes, demonstrating that complications in the natural source fields also contribute to the higher modes, at least under storm-time conditions.

[49] There is an increase in all eigenvalue amplitudes in the dead band on day 272. This appears to be related to the signal produced during ground motion by the earthquake itself, and is discussed by *Kappler et al.* [2006] and *Kappler* [2008]. The increases in residual energy at 4 s period correlate with increased energy in the third eigenmode at this period around the time of the earthquake, suggesting that the

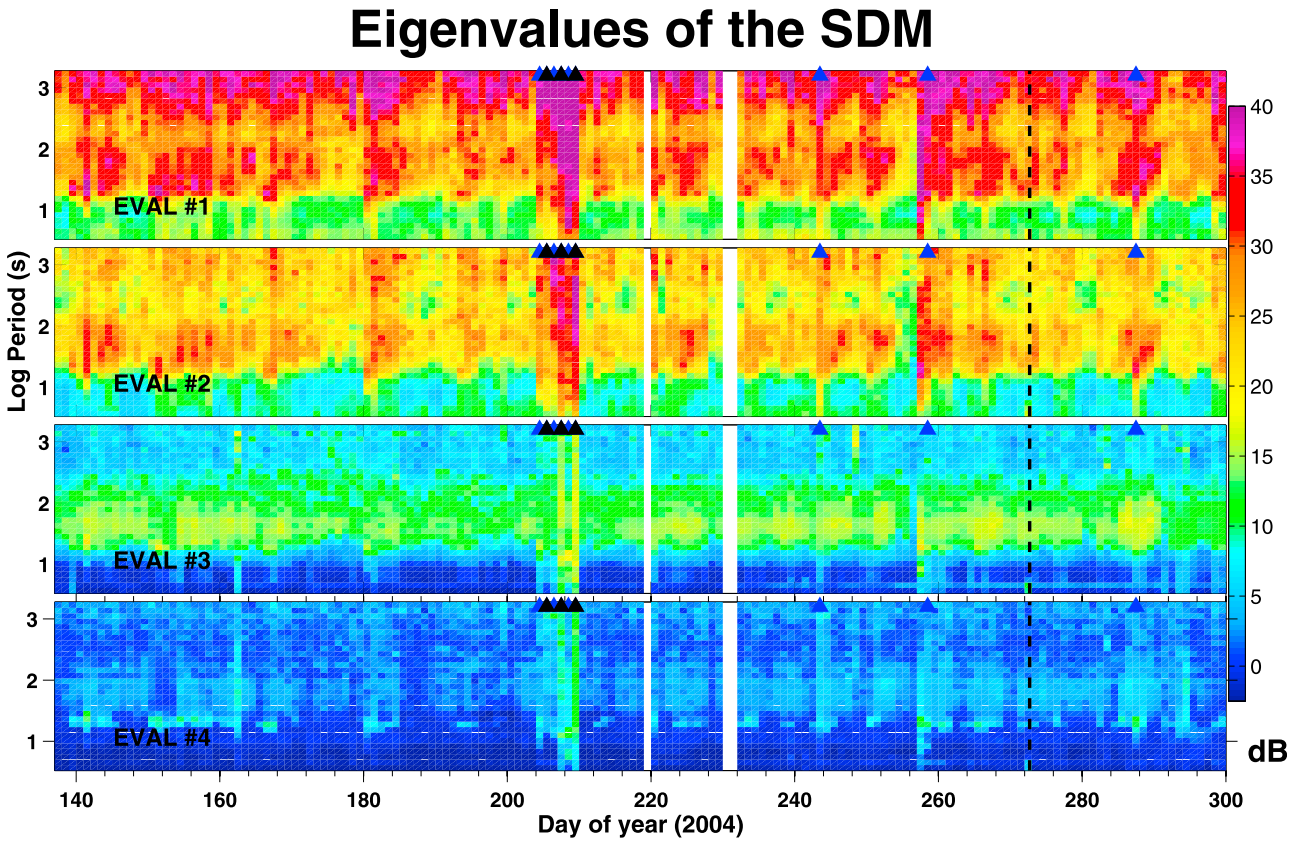

Figure 13. Dominant four eigenvalues of the SDM in decibels plotted for the 163 day interval surrounding the 2004 Parkfield earthquake.

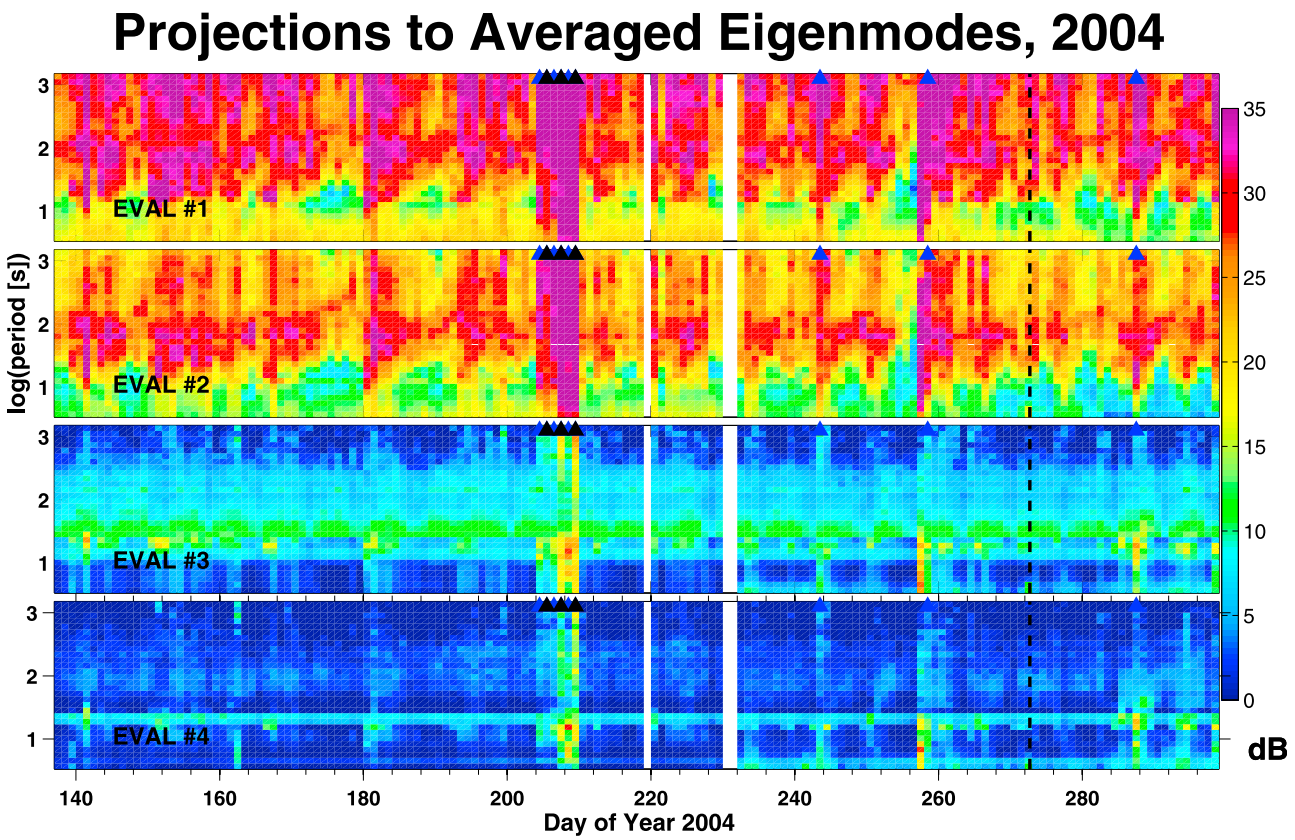

Figure 14. Projections of daily SDM eigenvectors onto the averaged modes of the four dominant eigenvectors.

signals in the two plots are caused by the same phenomena. No other anomalous energy is present around the time of the earthquake in this plot.

[50] Projecting the eigenvectors of each day onto the "averaged modes" (section 3.5) results in Figure 14. Most of the energy which was present in the fourth mode in Figure 13 is no longer present in this plot, implying that this mode is not particularly stable. Figure 14 reveals a weekly periodicity in the bands around 30 s of the third eigenmode. Isolating the band which corresponds to 31 s, the third eigenmode projection is shown in Figure 15. The time series values corresponding to Sundays are plotted in red, showing that there is typically around 2 dB less signal in that mode on Sundays than on other days. This is consistent with the suggestion of *Egbert et al.* [2000] that fields near this period are significantly influenced by the BART system, considering that BART operations are reduced on Sundays. A persistent peak in fourth average mode energy near 10 s period is apparent in Figure 14, which may be due to PC3 field line resonance. In the "averaged modes" plots, the narrowband signal, previously observed in the third eigenvalue at ~4 s in Figure 13, is not visible. This implies that the signal was not typically present on the days used to create the averaged modes, i.e., the first 3 weeks of the section.

4.5. Canonical Coherences

[51] With only two sites and two field types (E and H), there are two natural signal groupings: (1) electrics in one group $\mathbf{X}(t)$, magnetics in the other $\mathbf{Y}(t)$, and (2) Parkfield channels in one group, Hollister channels in the other.

[52] Canonical coherences over the 163 day window surrounding the earthquake are shown for the E-H and PKD-SAO channel groupings in Figures 16 and 17, respectively.

[53] To first order, these show for either grouping that high correlation coefficients exist across all bands, and for all times, for two linear combinations of field components. Note the increase in coherence of the 4 s band in the third electric-magnetic canonical coherence around the time of the earthquake. A corresponding phenomenon is not visible in the PKD-SAO plot. This implies that the signal responsible is present at only one site.

[54] The increase in correlation coefficients around the earthquake appears to be similar to an increase which occurred in 2002 [*Kappler*, 2008]. To reduce the appearance of variations in this time series which are broadband, we normalize the third CC coefficient in the 4 s band by the third CC coefficient in neighboring bands. The time series (normalized by the 3.3 s band) is filtered by a median smoother and shown in Figure 18. A plot of the CC coefficient ratios, normalized by the 5.0 s band, is also shown in Figure 19 for the 163 day section. It is tempting to ascribe significance to the local maxima in the normalized coherence near the earthquake in these plots. However, Figure 18 shows that similar large variation in the CC ratio occurs in spring of 2002. At both times, the signal is narrowband and inhabits the exact same three FCs of the band centered around 4.09 s. Selecting one day near each maxima, a plot

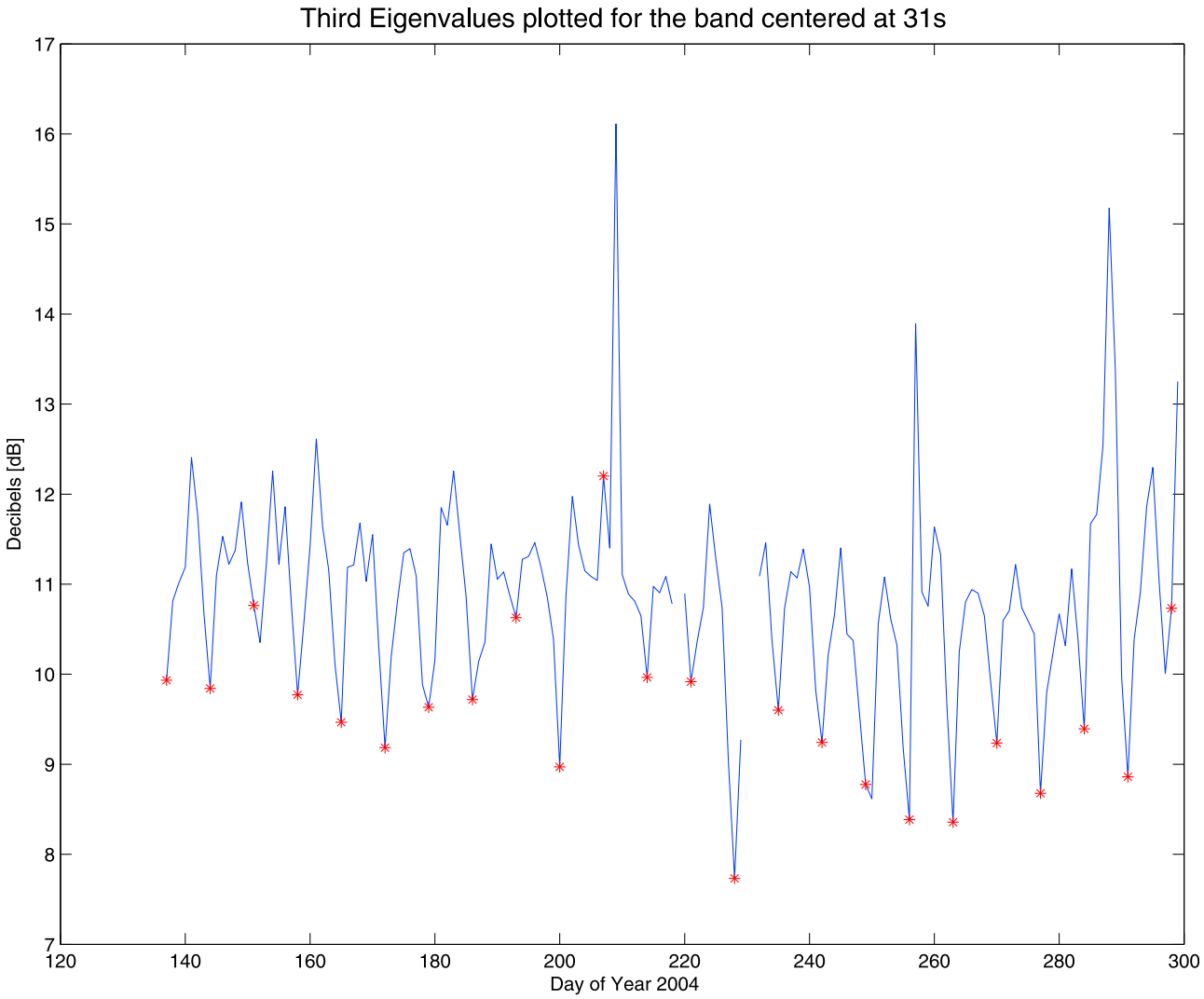

Figure 15. The band of the third eigenmode shown in Figure 14 corresponding to a period of 31 s. Sundays are denoted by a red asterisk.

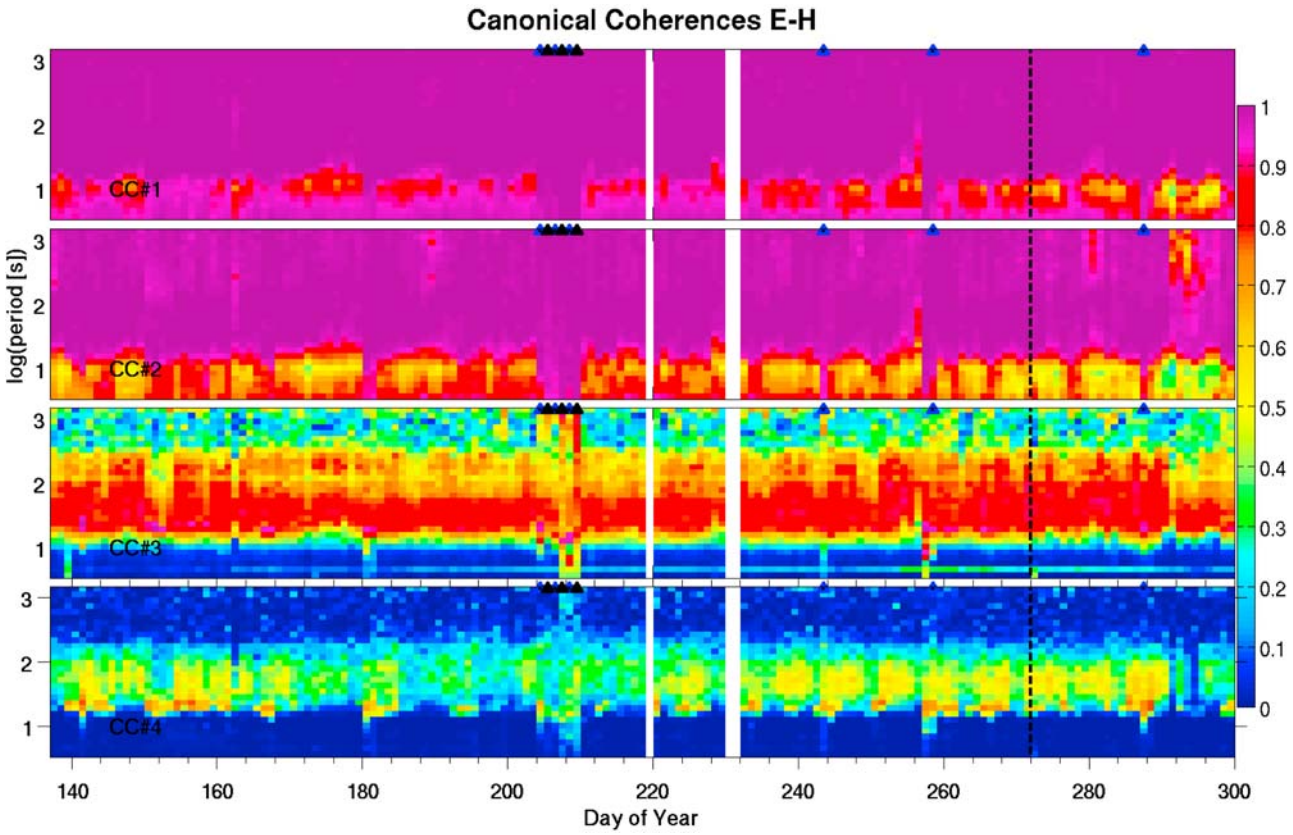

Figure 16. Canonical coherences between electric and magnetic field channels for the 163 day interval surrounding the 2004 Parkfield earthquake.

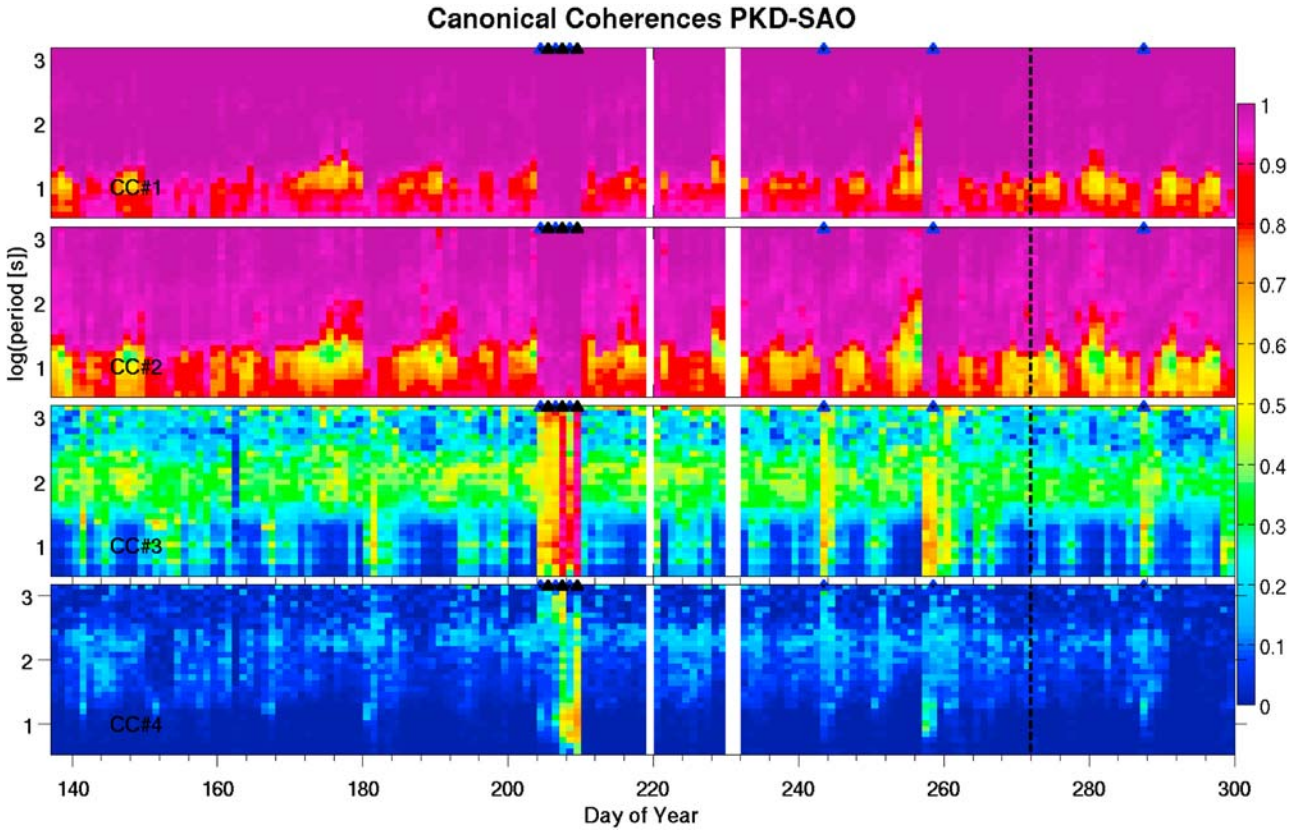

Figure 17. Canonical coherences in 2004 between channel groupings: Parkfield and Hollister for the 163 day interval surrounding the 2004 Parkfield earthquake.

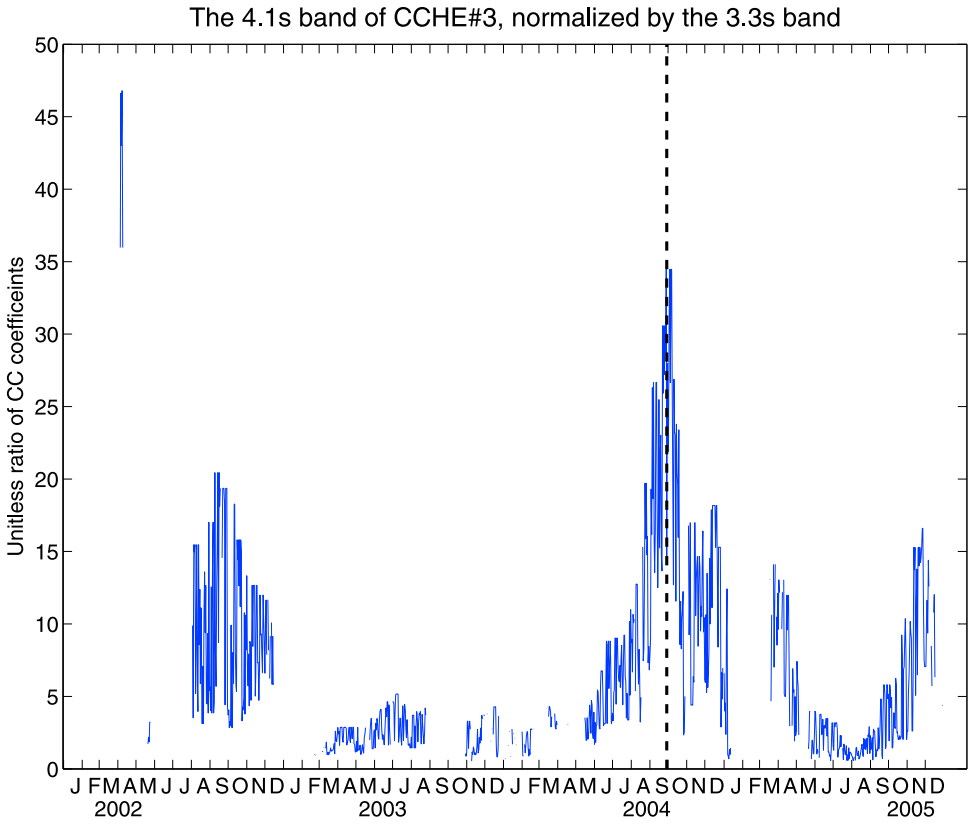

Figure 18. The median smoothed ratio of the 4.1 s CC to the 3.3 s band. The importance of long-term monitoring is highlighted by this plot. Considering only the 2 years centered around the earthquake, the time series might be interpreted as being related to the earthquake. A wider look at the phenomena, however, shows anomalous variations at other times.

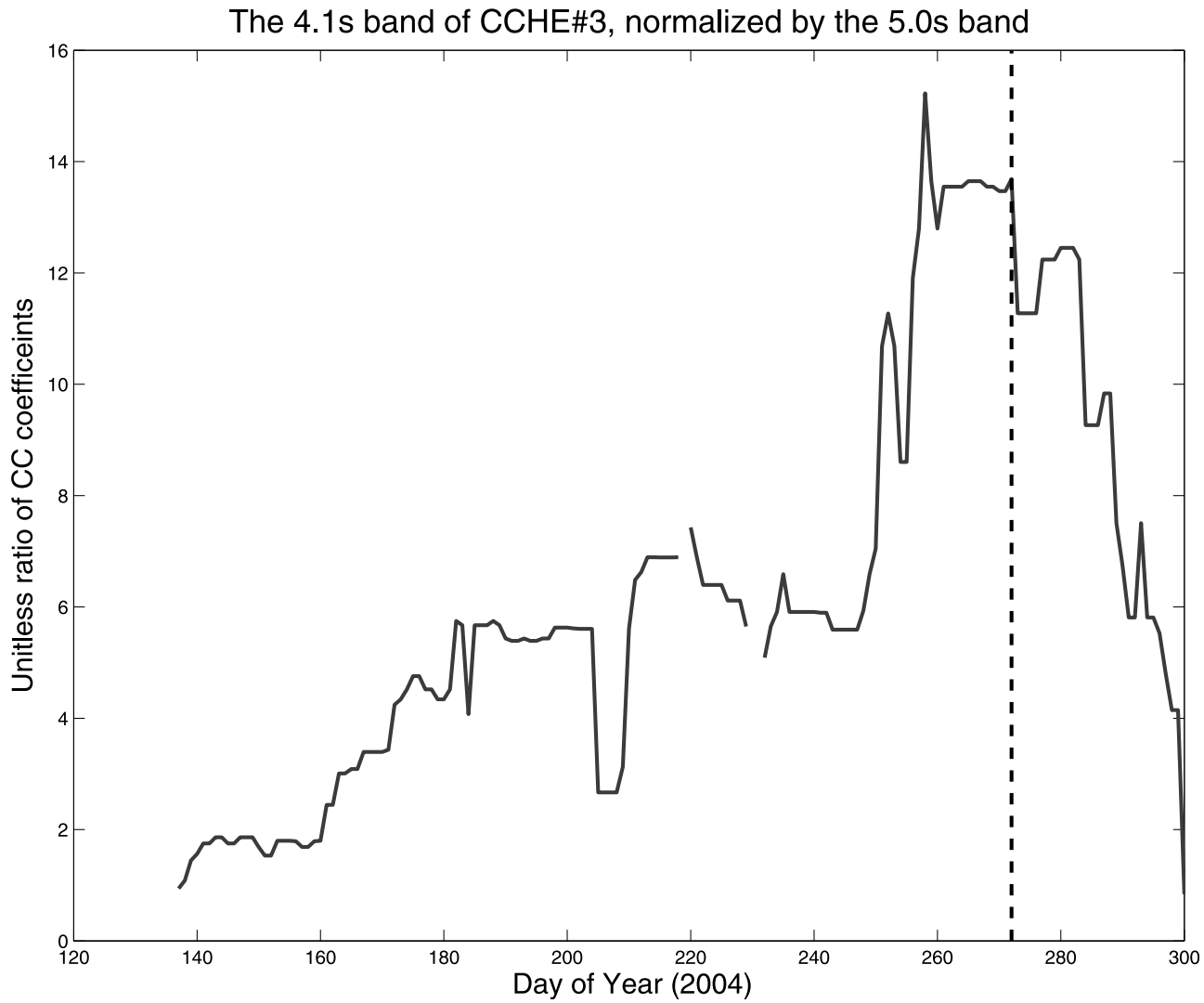

Figure 19. The 9 day median smoothed ratio of the 4.1 s CC to the 5.0 s band in the 163 day window around the earthquake.

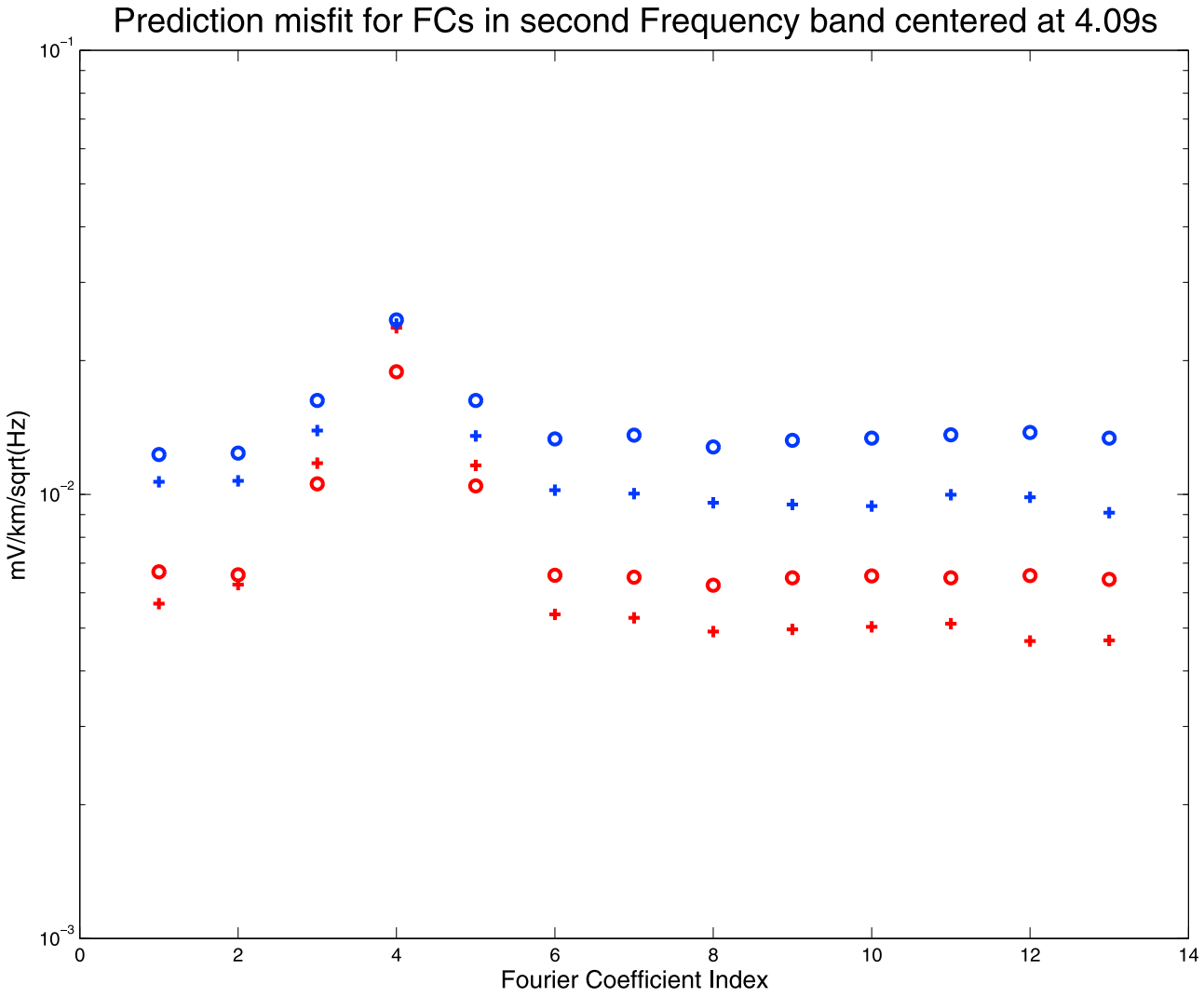

Figure 20. FC prediction misfit in a narrow frequency band centered at 4 s period. Observed field FCs are in blue, and residuals are in red. Crosses are from day 95, 2002, and circles are from day 270, 2004 (2 days prior to the PKD earthquake).

of the FC and residual amplitudes, comparing 5 April 2002 against 26 September 2004 is shown in Figure 20. Note that although the observed field amplitudes are shifted from one time of observation to the next, as can be expected due to natural variability in field strength, the anomalous signal is present at the same amplitude. No significant seismicity occurred at or near PKD 2 months to either side of day 95 in 2002. Although $M1$ and $M2$ earthquakes occurred within less than 10 km of the array frequently, more of these small nearby events occurred in spring 2003 than in spring 2002. Clearly, there is a real phenomenon responsible for these signals. On the basis of Figure 20, however, it is reasonable to conclude that the source of the 2002 anomaly is the same as the source of the 2004 anomaly, and that it is not directly related to the 28 September earthquake.

4.6. Apparent Resistivity Variation and Distortion Correction

[55] Following *EE02*, we quantify the changes in apparent resistivity in percent deviation from the 4 year median value in each band for the 163 day section in Figure 21. Day-to-day variations are mostly random, but there is some indication of a broadband increase in apparent resistivity near the end of the 163 day section. This shift is similar across all frequency bands suggestive of near-surface galvanic distortion effects, instrument swaps, or gain settings. Changes with time in deeper subsurface conductivity should be frequency dependent.

[56] The black circles at the top of each plot around day 290 indicate days of significant rainfall. Thus, it appears that the broadband variation in apparent resistivity after the earthquake is related to the rainfall event. Interestingly, there does appear to be an anomaly in the high-frequency apparent resistivity after the earthquake in the *YX* mode. It is not clear why a corresponding change is not seen in the *XY* mode.

[57] In an effort to decouple the frequency-dependent and frequency-independent changes in apparent resistivity, we express the daily measured impedance tensors as a product of two tensors, where one factor is the complex 4 year

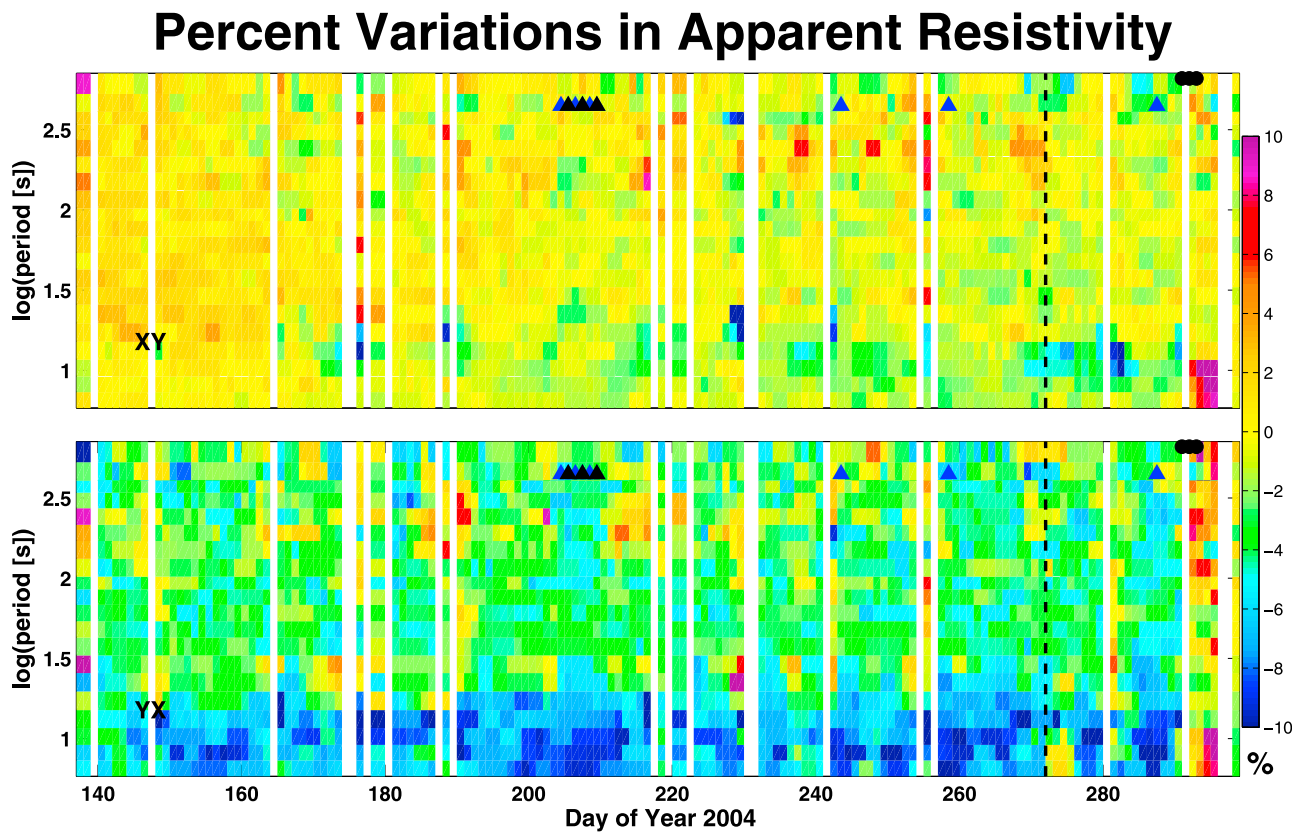

Figure 21. Raw apparent resistivity data in percentage deviation from the 4 year median. *XY* mode is shown on top, and *YX* mode is shown on the bottom. Days of significant (>0.1 inch) rainfall are marked by black circles.

median impedance tensor, and the other factor is a frequency-independent (Real) perturbation tensor we label **D** [see *Smith*, 1995, and references therein].

[58] Formally, for each day t, and each frequency bin ω, we have an impedance tensor estimate $\mathbf{Z}_{t,\omega} \in M_{2,2}(\mathbb{C})$. These 2 × 2 matrices can be concatenated over all frequencies to create $\mathbf{Z}_t \in M_{2B,2}(\mathbb{C})$, where B is the number of frequency bands under consideration. Here we use $B = 23$ log linear spaced bands. Similarly, a 4 year median impedance tensor at each frequency, which we denote by $\overline{\mathbf{Z}}_{t,\omega}$, can be concatenated over the B frequency bins to make $\overline{\mathbf{Z}} \in M_{2B,2}(\mathbb{C})$.

[59] We then define the daily distortion tensor $\mathbf{D}_t$ as the least squares solution which minimizes the expression

$$||\mathbf{Z}_t - \mathbf{D}_t\overline{\mathbf{Z}}||, \tag{11}$$

where we constrain **D** to be real-valued. By multiplying $\mathbf{Z}_t$ on the left by $\mathbf{D}_t^{-1}$, one obtains a "distortion-corrected" impedance tensor, which is nearest the long-term median impedance tensor in a least squares sense.

[60] Raw and distortion-corrected apparent resistivity deviations for both *XY* and *YX* modes are shown smoothed over time and frequency in Figure 22. A 9 day median filter is applied in time to the time series of percent variations, before the bands 1–5 and 6–10 are averaged together. A complete suite of raw and distortion-corrected apparent resistivities in the presentation style of EE02 is given by *Kappler* [2008]. Figure 22 shows that most of the variability in the deviation from the median is removed by the distortion corrections.

[61] Figure 23 shows that the distortion tensor absorbs most of the seasonal variability present in the apparent resistivity time series. The distortion tensor time series for this plot is smoothed in time by using all days in a 9 day window centered at t to generate $\mathbf{Z}_t$. Days where rainfall was greater than 0.25 cm are marked by black circles, and days where rainfall occurred, but was less than 0.25 cm, are marked by green circles. Vertical lines at year changes emphasize the seasonal variations in $D_{1,1}$. There is a tendency for the apparent resistivity estimates to be less stable during the rainy season, as well as to step upward with the onset of the rain. The months-long trends are toward higher apparent resistivity in the winter and to lower apparent resistivity in the summer.

5. Conclusions and Future Work

[62] Observed ULF fields typically comprise natural MT fields, with cultural and instrument noise superposed. There were no anomalies in the measured magnetic fields or the impedances derived from the EM fields prior to the Parkfield earthquake. Several anomalous features in the data that we did identify were eventually shown to be artifacts of signal processing, or to be not uniquely associated with the earthquake when a long time window was considered. The

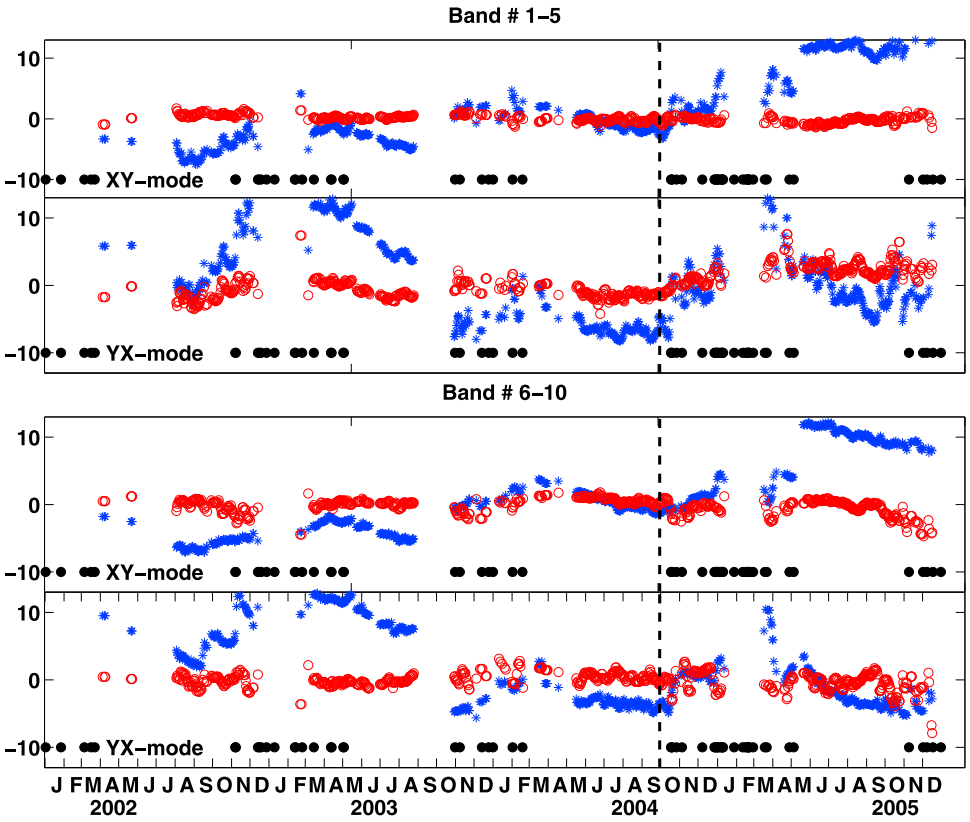

Figure 22. Time series of apparent resistivity deviation from median and distortion-corrected time series deviation from the median. High- and low-frequency band averages are centered at 7 s and 25 s periods, respectively.

array's ability to see clearly the weekly period in a commuter train schedule at a distance of over 100 km is a tribute to its sensitivity. In contrast, no precursor to a *M*6 earthquake at 20 km was observed, leaving some doubt about the practicality of the array spacing that would be required to detect these sorts of anomalies even if they did exist. The *M*6 was located at a depth of 8.6 km, for a hypocentral distance of 21.7 km. The anomalous fields reported by *Fraser-Smith et al.* [1990] were observed at a similar hypocentral distance of 18.7 km (epicentral distance 7 km, depth 17.4 km). The Loma Prieta earthquake was an order of magnitude larger than the Parkfield earthquake but the ULF magnetic field anomaly reported was sufficiently large as to stand out clearly above the natural fields in several frequency bands, most notably, the MA3 band at 0.01 Hz. No such signal was seen at Parkfield, even relative to residual fields, which are an order of magnitude smaller than the natural source fields. It is particularly noteworthy that we found no evidence of the narrowband quasi-harmonic anomalous field amplitudes inferred by *Fraser-Smith et al.* [1990] from the "wandering" of the anomalous fields between the spectral windows. We also did not see any burst of the type reported by *Fraser-Smith et al.* [1990] in the few hours prior to the earthquake. We had a median residual (site-noise) level of 0.07 nT/$\sqrt{\text{Hz}}$ at 0.01 Hz during the 163 day section around the earthquake on which most analysis was focused. Both the ability to identify cultural noise with a weekly period in such a remote area and the stability of the impedances seen in the last section increase our confidence in the validity of the long-term data.

[63] Our results show that long-term monitoring is required in order to establish a baseline of what is anomalous and what is normal. We emphasize that an anomalous behaviour was shown to have occurred around the time of the earthquake, but on further inspection these sorts of phenomena were also found to occur at other times. Specifically, broadband increases in all fields measured were observed prior to the earthquake both on the day of as well as several weeks before. In each case, these are better explained by enhanced geomagnetic activity due to documented solar storms, or natural MT noise, in the case of the same-day anomaly. Curious variations in the amplitude of a signal having period around 4.1 s occurred around the time of the earthquake, apparently in both the electric and magnetic fields at Parkfield. On further inspection, the exact same signature was found in the data years earlier, when no significant seismic activity occurred near Parkfield. Another anomalous signal, this time in the 3–7 s band of the residuals, which occurred several weeks before the earthquake was shown to be an artifact of data processing (or acquisition) rather than an actual signal. Finally, there were multiple spikes in the data around the time of the earthquake, but no more at that time than at any other time which we considered.

[64] Some of the noise and anomalous signals observed in the array data are not completely understood. These severely

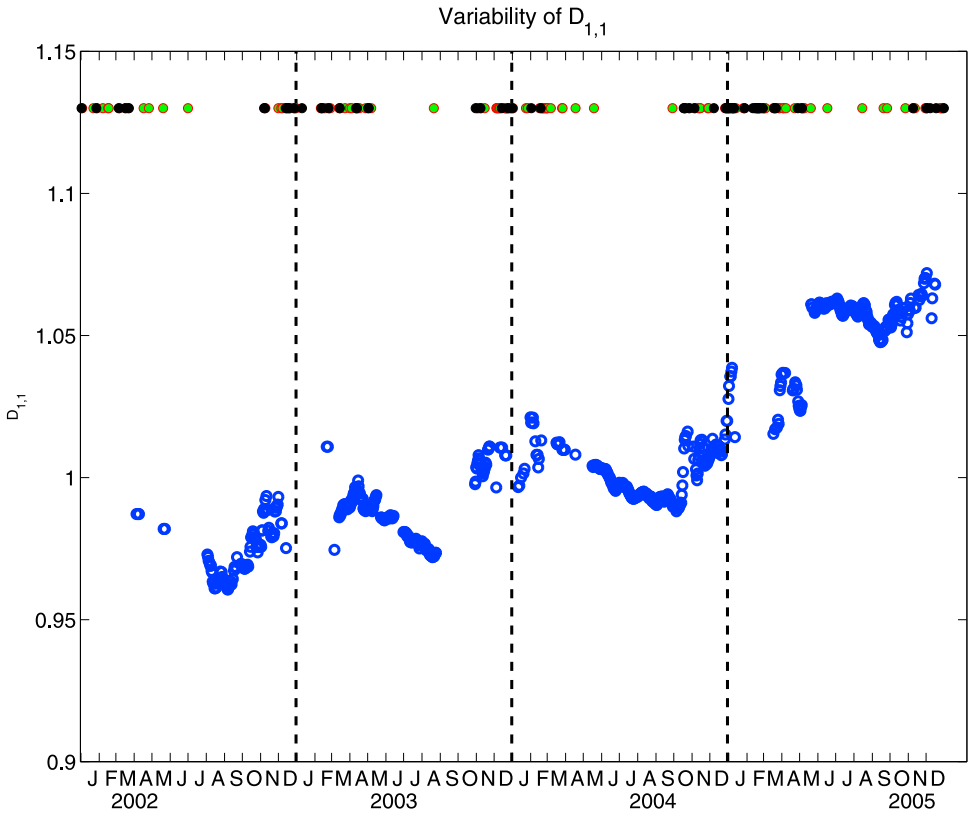

Figure 23. Variability of the 1,1 element of the distortion tensor over the 2002–2005 time interval. Vertical lines are placed at year changes.

complicate the search for anomalies which might be precursory. Our results serve as a caution to researchers who plan to undertake long-term monitoring efforts. Obviously, ongoing maintenance, and equally important, keeping strict records of specific maintenance operations are essential. We found it very difficult to keep instruments running for a long period of time with absolute calibration, particularly with the limited resources that were available.

[65] The application of multivariate signal characterization (using the SDM, principal components, and canonical coherences) was helpful in the search for anomalous processes and signals. Coherent noise and nonplane wave signals were found to be common—due both to cultural noise and source complications. In applying these techniques, however, we still found no clear evidence for anomalous EM signals that could be clearly associated with any earthquake.

[66] Monitoring of apparent resistivity revealed a seasonal pattern. This appears to be at least partly associated with the wetting and drying of the near surface. There does appear to be a slight trend toward an increase in apparent resistivity in the high-frequency bands after the earthquake, which is not present in the lower frequencies. This needs to be examined in further detail before we can say whether or not the earthquake actually caused a change in the near-surface conductivity. Perhaps the most important point to take from this study is that careful long-term monitoring with stable well-calibrated instruments is necessary to characterize and understand "normal background" variations before any truly "anomalous" tectonically related signals could be reliably identified.

[67] **Acknowledgments.** This research was supported by the USGS National Earthquake Hazards Reduction Program (NHERP) grants ending with 05HQGR0077. The authors would also like to thank the Berkeley Seismological Laboratory for the use of their computing facilities.

References

Bakun, W. H., and T. V. McEvilly (1984), Recurrence models and Parkfield, California earthquakes, *J. Geophys. Res.*, *89*, 3051–3058.

Becken, M., O. Ritter, S. K. Park, P. A. Bedrosian, U. Weckmann, and M. Weber (2008), A deep crustal fluid channel into the San Andreas fault system near Parkfield, California, *Geophys. J. Int.*, *173*, 718–732.

Boyd, O. S. (2000), Parkfield-Hollister ULF monitoring array, MS thesis, Univ. of Calif., Berkeley.

Brillinger, D. R. (1969), The canonical analysis of stationary time series, in *Multivariate Analysis II*, edited by P. R. Krishnaiah, pp. 331–350, Academic Press, New York.

Corwin, R. F., and H. F. Morrison (1977), Self-potential variations preceding earthquakes in central California, *Geophys. Res. Lett.*, *4*, 171–174.

Egbert, G. D. (1997), Robust multiple-station magnetotelluric data processing, *Geophys. J. Int.*, *130*, 475–496.

Egbert, G. D. (2002), Processing and interpretation of electromagnetic induction array data, *Surv. Geophys.*, *23*, 207–249.

Egbert, G. D., and J. R. Booker (1986), Robust estimation of geomagnetic transfer functions, *Geophys. J. R. Astron. Soc.*, *87*, 173–194.

Egbert, G. D., and J. R. Booker (1989), Multivariate analysis of geomagnetic array data: 1. The response space, *J. Geophys. Res.*, *94*, 14,227–14,247.

Egbert, G. D., M. Eisel, O. S. Boyd, and H. F. Morrison (2000), DC trains and Pc3s: Source effects in mid-latitude geomagnetic transfer functions, *Geophys. Res. Lett.*, *27*, 25–28.

Eisel, M., and G. D. Egbert (2002), On the stability of magnetotelluric transfer function estimates and the reliability of their variances, *Geophys. J. Int.*, *144*, 65–82.

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 Ms 7.1 Loma Prieta earthquake, *Geophys. Res. Lett.*, *17*, 1465–1468.

Gamble, T. D., W. M. Goubau, and J. Clarke (1979), Magnetotellurics with a remote reference, *Geophysics*, *44*, 53–68.

Geller, R. J. (1996), Debate on evaluation of the VAN method: Editor's introduction, *Geophys. Res. Lett.*, *23*, 1291–1293.

Hardle, W., and L. Simar (2007), *Applied Multivariate Statistical Analysis*, Springer, Berlin.

Huber, P. J. (1981), *Robust Statistics*, John Wiley, Hoboken, N. J.

Johnston, M. J. S. (1989), Review of magnetic and electric field effects near active faults and volcanoes in the U.S.A., *Phys. Earth Planet. Int.*, *57*, 47–63.

Johnston, M. J. S. (1997), Review of electric and magnetic fields accompanying seismic and volcanic activity, *Surv. Geophys.*, *18*, 441–476.

Johnston, M. J. S., Y. Sasai, G. D. Egbert, and S. K. Park (2006), Seismomagnetic effects from the long-awaited 28 September 2004 *M*6.0 Parkfield earthquake, *Bull. Seismol. Soc. Am.*, *96*, 206–220.

Kappler, K. N. (2008), Long-term electromagnetic monitoring at Parkfield, CA, PhD thesis, Univ. of Calif., Los Angeles.

Kappler, K. N., N. H. Cuevas, and J. W. Rector (2006), Response of induction coil magnetometers to perturbations in orientation, paper presented at the 76th Annual Meeting, Soc. of Explor. Geophys., New Orleans, La, 1–6 October.

Lyubushin, A. A., Jr. (1998), Analysis of canonical coherences in the problems of geophysical monitoring, *Izv. Phys. Solid Earth*, *34*, 52–58.

Molchanov, O. A., Y. A. Kopytenko, P. M. Voronov, E. A. Kopytenko, T. G. Matiashvili, A. C. Fraser-Smith, and A. Bernardi (1992), Results of ULF magnetic field measurements near the epicenters of the spitak (*Ms* = 6.9) and Loma Prieta (*Ms* = 7.1) earthquakes: Comparative analysis, *Geophys. Res. Lett.*, *19*, 1495–1498.

Park, S. K., M. J. S. Johnston, T. R. Madden, F. D. Morgan, and H. F. Morrison (1993), Electromagnetic precursors to earthquakes in the ULF band: A review of observations and mechanisms, *Rev. Geophys.*, *31*(2), 117–132.

Park, S. K., W. Dalrymple, and J. C. Larsen (2007), The 2004 Parkfield earthquake: Test of the electromagnetic precursor hypothesis, *J. Geophys. Res.*, *112*, B05302, doi:10.1029/2005JB004196.

Pulinets, S., and K. Boyarchuk (2004), *Ionospheric Precursors of Earthquakes*, Springer, Berlin.

Smith, J. T. (1995), Understanding telluric distortion matrices, *Geophys. J. Int.*, *122*, 219–226.

Unsworth, M., and P. Bedrosian (2004), On the geoelectric structure of major strike-slip faults and shear zones, *Earth Planets Space*, *56*, 1177–1184.

Uyeda, S., T. Nagao, Y. Orihara, T. Yamaguchi, and I. Takahashi (2000), Geoelectric potential changes: Possible precursors to earthquakes in Japan, *Proc. Natl. Acad. Sci. U. S. A.*, *97*, 4561–4566.

Varotsos, P., and K. Alexopoulos (1984), Physical properties of the variations of the electric field of the Earth preceding earthquakes, part I, *Tectonophysics*, *110*, 73–98.

Varotsos, P., and M. Lazaridou (1991), Latest aspects of earthquake prediction in Greece based on seismoelectric signals, *Tectonophysics*, *188*, 321–347.

G. D. Egbert, College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331, USA.

K. N. Kappler, Berkeley Seismological Laboratory, University of California, 215 McCone Hall, Berkeley, CA 94720, USA. (kappler@newton.berkeley.edu)

H. F. Morrison, Department of Earth and Planetary Science, University of California, 215 McCone Hall, Berkeley, CA 94720, USA.