Identification and classification of transient pulses observed in magnetometer array data by time-domain principal component analysis filtering

Karl N. Kappler · Daniel D. Schneider · Laura S. MacLean · Thomas E. Bleier

Received: 28 January 2017 / Accepted: 28 August 2017 / Published online: 27 October 2017
© The Author(s) 2017. This article is an open access publication

Abstract A method for identification of pulsations in time series of magnetic field data which are simultaneously present in multiple channels of data at one or more sensor locations is described. Candidate pulsations of interest are first identified in geomagnetic time series by inspection. Time series of these “training events” are represented in matrix form and transpose-multiplied to generate time-domain covariance matrices. The ranked eigenvectors of this matrix are stored as a feature of the pulsation. In the second stage of the algorithm, a sliding window (approximately the width of the training event) is moved across the vector-valued time-series comprising the channels on which the training event was observed. At each window position, the data covariance matrix and associated eigenvectors are calculated. We compare the orientation of the dominant eigenvectors of the training data to those from the windowed data and flag windows where the dominant eigenvectors directions are similar. This was successful in automatically identifying pulses which share polarization and appear to be from the same source process. We apply the method to a case study of continuously sampled (50 Hz) data from six observatories, each equipped with three-component induction coil magnetometers. We examine a 90-day interval of data associated with a cluster of four observatories located within 50 km of Napa, California, together with two remote reference stations—one 100 km to the north of the cluster and the other 350 km south. When the training data contains signals present in the remote reference observatories, we are reliably able to identify and extract global geomagnetic signals such as solar-generated

noise. When training data contains pulsations only observed in the cluster of local observatories, we identify several types of non-plane wave signals having similar polarization.

Keywords Time series · Magnetic fields · Array data · Signal processing · Principal component analysis

1 Introduction

There are many reports of transient pulses observed before earthquakes in the electromagnetic (EM) field (Varotsos and Alexopoulos 1984; Fraser-Smith et al. 1990, 1991; Bleier et al. 2009; Dunson et al. 2010; Han et al. 2014). While some pulses may be sourced from subsurface phenomena (Bleier et al. 2010; Freund 2007a, b; Takeuchi and Nagao 2013), it is well understood that many of these spurious signals are sourced from cultural noise (Szarka 1988; Junge 1996) and natural processes such as lightning, solar storms, or other magnetospheric/ionospheric interactions (Campbell 2003; Vozoff 1991; Egbert et al. 2000). To better understand the sources of these pulses, we have been monitoring the magnetic field via the QuakeFinder (QF) network of observatories for approximately 10 years. This network consists of 124 stations mostly along the San Andreas Fault in California and another 40 stations along fault zones in Greece, Taiwan, Peru, Chile, and Indonesia. Because we are ultimately interested in understanding pulses associated with earthquakes, we are using a subset of the QF data set collected near Napa, California, over a time window that includes the August 24, 2014, *M*6.0 Napa earthquake.

Here we present a method for pulse identification based on a variant of principal component analysis seeded by

K. N. Kappler (✉) · D. D. Schneider · L. S. MacLean · T. E. Bleier
QuakeFinder, Palo Alto, CA 94306, USA
e-mail: kkappler@quakefinder.com

training data. This time-domain principal component analysis (TDPCA) filter readily recognizes pulses that appear to be sourced from the same process as the training data and is particularly well suited to finding multiple occurrences of the same pulse type observed at more than one station.

For preliminary data inspection, we used a visualization tool which shows the multi-channel time series from the selected stations. We found that there are many spurious transients in the observed magnetic field time series which can readily be identified by visual inspection. Traditionally in EM surveys employing a remote reference station (Gamble et al. 1979), the presence of a signal at multiple stations far apart is presumed to indicate natural space-weather-sourced magnetic activity (Kappler 2012). In our study, it was clear that some of the transients were array-wide and hence likely due to space weather sources. Many

other pulses could be seen in two or more of the stations near Napa but not at either reference station. Using the high density of the QF array, we will show that many pulses which occur at multiple stations are in fact cultural noise which extends over tens of kilometers or more.

In section two, we introduce a six-station subset of the QF array focused on Napa and show some examples of pulses observed at multiple stations. In section three, we describe the mathematical approach to the TDPCA filtering method. In section four, we apply the TDPCA filtering method using the example pulses of section one as training data and show the distribution of these pulses in time. In section five, we interpret the results in terms of sources based on the distribution of the example pulses in time-of-day, relative amplitudes, and directions of the fields. Finally, we discuss the significance of these pulses and the role of the TDPCA filtering method in monitoring the EM field.

Fig. 1 The six magnetometer stations studied comprised four clustered near Napa, California, and two remote reference stations. All stations are shown as *yellow circles* and the location of the *M6.0 Napa Earthquake* on August 24, 2014 is shown as a *red star*. The two reference stations—Ukiah (914) and Cholame (850)—relative to the four stations of interest. Inset shows zoom of the four primary stations—East Petaluma (854), Napa Valley (871), Black Point (809), and Benicia (870) approximately bounded by the *black rectangle*. This map was created using Google Earth

Table 1 Information about the four stations near Napa, California, and the two remote reference stations

Station name	Station ID	Lat.(°)	Long.(°)	Magnetometer model	Epicentral distance (km)
Ukiah	914	39.19	−123.19	QF 3	132.11
East Petaluma	854	38.29	−122.58	ANT 4	24.88
Napa Valley	871	38.34	−122.26	QF 3	14.02
Black Point	809	38.17	−122.51	QF 3	18.38
Benicia	870	38.12	−122.11	QF 3	20.75
Cholame	850	35.67	−120.20	ANT 4	339.69

2 Instrumentation and data

2.1 Instrumentation

The set of data used for this study was selected to include the *M*6.0 Napa Earthquake on August 24, 2014. Our window of analysis starts two months prior to the earthquake (June 24, 2014) and completes one month after it (September 24, 2014). The data were collected from six QF stations in Northern California with four located around Napa Valley and two reference stations (Fig. 1). Each station is equipped with three feedback induction magnetometers, two ion sensors (one for detection of positive ions and one for the detection of negative ions), a geophone, a temperature sensor, and a humidity sensor. The magnetometer data, which were the focus of this study, include north, east, and vertical sensors buried approximately 15 cm below the surface. Information for each station and the specific model of its magnetometer are listed in Table 1. The data were recorded at 50 samples per second with GPS antennas supplying pulse-per-second reference time stamps, continuously collected, and transmitted daily to the QF data center in Palo Alto, California.

2.2 Data

Described below are four different pulses we refer to as Type-1, Type-2, Type-3, and Type-4. These pulses were first identified visually and were then used as training data for the TDPCA filtering. They were chosen because they had large amplitudes compared to the background noise and were recorded by several stations indicating they were caused by physical phenomena. Using a multiple-channel plotting utility with the time scale zoomed out (to 20,000 s), most of these pulse types appear non-unique relative to each other—the main attribute that differentiates the events is the active channels sensing the pulse (Fig. 2). Specifically, Type-1, Type-2, and Type-3 pulses are not recorded at the reference stations, while Type-4 pulses are coincident at all stations. Unless specially stated, all figures are in UTC time.

On all but the narrowest time scales, a first-order search based on thresholds is not able to differentiate the characteristics of individual energetic excursions of the time series and therefore is not able to separate the sources. However, a close examination of the different pulses reveals differences in their character (Fig. 3). The primary distinguishing traits are the collection of channels that are simultaneously triggered and the relative phases of the pulse in each channel. For each type of pulse shown in Fig. 3, we list the stations, their channels which were being triggered, and their relative phases to each other for the identified pulse used as training data (Table 2). These polarities are specific to the training pulse. In another instance, the pulse may appear differently, e.g., the phases for all of the channels in Type-4 may be $-$ instead of $+$, or a Type-1 may appear as $+++--$ rather than $---+-$. Despite looking different, the relationship, or pattern, between the channels and their relative phases will remain the same. Most pulses are visible only on a subset of the 18 channels that were monitored for this study. The specific subset offers a clue to the spatial location of the source. Other traits include the pulse duration and the relative phases of the channels. Each ‘type’ of training data was used to quantitatively identify similar pulses in the data.

3 TDPCA filtering method

After isolating pulses of interest, we sought to identify similar pulses in the full data set. To compare these training pulses with the entire time series data, we characterized each one by the eigenvectors of its Pearson correlation matrix (Bendat and Piersol 1980). These eigenvectors were then compared to the eigenvectors of the windowed time series data. In the following description we apply a linear detrending operation to each channel before performing any Pearson calculations.

We started by considering the properties of a particular training datum (i.e., pulse), which can be considered as a sub-array of the full time series. The pulse was seen to

Fig. 2 Time series data from different stations and channels. **a** Type-1 pulse, **b** Type-2 pulse, **c** Type-3 pulse, **d** Type-4 pulse for 20,000 s. For simplicity, we only include the stations and channels where the identified pulse is actually seen

Fig. 2 continued

Fig. 3 Different types of identified pulses viewed at higher resolution. **a** Type-1 pulse, **b** Type-2 pulse, **c** Type-3 pulse, **d** Type-4 pulse zoomed into show their distinct characteristics. For simplicity, we only include the stations and channels where the identified pulse is visually available

Table 2 Station, channel, and phase information for each type of pulse

Type	Station 809			Station 854			Station 870			Station 871			Station 914			Station 850		
	N	E	Z	N	E	Z	N	E	Z	N	E	Z	N	E	Z	N	E	Z
1	–	–	–		+	–												
2	–	~	~		~	+												
3	–	–	+	–	–	+	+	+	~									
4	+			+			+			+			+				+	

The symbols +, –, and ~ denote positive, negative, and mixed polarity, respectively. The values entered here are from the training data in Fig. 3a–d

occur across N channels during a particular time interval which has M samples. We represent the training data as an $N \times M$ array denoted as T , where the i th row of

T corresponds to the time series for i th channel in the set of N channels. Multiplying T by its transpose T^* forms the $N \times N$ correlation matrix C following Eq. (1):

$$\mathbf{C} = \mathbf{T}\mathbf{T}^*, \quad (1)$$

$\mathbf{C}$ is then normalized to obtain the Pearson correlation matrix $\mathbf{R}$ where the i th row and j th column of $\mathbf{R}$ are normalized such that

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}. \quad (2)$$

The matrix $\mathbf{R}$ effectively encodes the cross-correlation information in the active channels and stores neatly as a relatively small $N \times N$ array. The eigenvectors of $\mathbf{R}$ are then calculated, $\hat{t}_1, \hat{t}_2, \dots, \hat{t}_N$. These eigenvectors are then ordered with respect to descending eigenvalue amplitude. These normalized eigenvectors are essentially the pattern we aimed to match in the data.

To search the general data time series for the training pattern, we began by windowing the data via a striding window of length M with some overlap. If the overlap is too small, it is possible to miss a pulse that straddles an adjacent window. For this reason, we used a rather generous overlap of 75%.

For a given window of data $\mathbf{D}$ (an $N \times M$ array—the same as the training data $\mathbf{T}$), we calculated $\mathbf{R}$ where $\mathbf{C} = \mathbf{D}\mathbf{D}^*$. The result was a collection of ordered eigenvectors, $\hat{d}_1, \hat{d}_2, \dots, \hat{d}_N$. Matching was then done by comparing the colinearity of the dominant eigenvectors ($\hat{d}_i$) from the windowed data against those from the training data ($\hat{t}_i$). This condition was imposed by insisting the inner products of the respective eigenvectors were large and positive. Here, we only compared the first two dominant eigenvectors and chose a threshold of 1.95, i.e., a window was flagged as having the same behavior as the training data if $\hat{d}_1\hat{t}_1^* + \hat{d}_2\hat{t}_2^* \geq 1.95$.

Due to the overlapping windows, a single event may be flagged in more than one consecutive time window. Thus, we performed a second level of filtering where we merged adjacent flagged time intervals. Thus, if two consecutive intervals $\mathbf{D}_i$ and $\mathbf{D}_{i+1}$ were flagged, we counted this as only one interval with the starting time of $\mathbf{D}_i$ and the ending time of $\mathbf{D}_{i+1}$.

Initially we found many false-positive matches. After investigating this, we found most of the false positives were likely a result of randomly oriented eigenvectors due to noise. To mitigate this, we applied a threshold of 0.01 V to the sum of the variances of the individual channels in $\mathbf{D}$ to exclude low energy windows (e.g., noise).

4 Results of TDPCA from Napa data

To identify pulses in the data with characteristics similar to those used as training data, we used the attribute of

similarly oriented dominant eigenvectors. This method revealed pulses with the same overall pattern as the training data, yet the identified pulses also displayed some unique and complex traits. Many of the Type-1 and Type-2 pulses that were found were shifted in phase relative to the training data (Fig. 4). For example, the majority of the signals have negative polarity in Fig. 4, while those at station 854 have positive polarity. During a 180° phase shift, the majority of the signals then have a positive polarity and those at station 854 then have a negative polarity. Although the relative phases of the signals were consistent for each pulse instance, the absolute phase change was apparently random. We tabulated the distributions of both Type-1 and Type-2 pulses with respect to phase by visual inspection (Table 3). We did this by using a crude set of attribute bins: 0°, 90°, and 180°. Despite this phase shift, the pulse still contained the same characteristics as the training data.

By using normalized eigenvectors, TDPCA filtering is also unaffected by the amplitude of the training data. Therefore, Type-2 pulses revealed signals in the data with the same pattern but with very different amplitude ratios (Fig. 5). Despite the obvious difference in the amplitude ratio between stations 809 and 854, these pulses were identified by using the same training datum.

Similarly, TDPCA filtering is also indifferent to the complexity and the duration of the pulse. Type-1 and Type-2 signals were relatively simple with uniform duration and retrieved similar pulses from the data. When using an admittedly more complex training data such as Type-3, we identified complicated pulses (Fig. 6). Although the training pulse for Type-3 was only 30 s long, these pulses vary in duration up to 190 s yet possess the same pattern as the training data. A similar phenomenon is apparent with the Type-4 pulses (Fig. 7). Although these training data were not as chaotic as Type-3, they retrieved a range of pulses varying in complexity and duration.

To determine more information about the source of these various pulses, we analyzed the number of pulses per day (Fig. 8). The number of detected pulses per day for all pulse types varied significantly with no obvious patterns. To gain further insight into the sources, we then evaluated the number of occurrences versus day-of-week for each type of pulse (Fig. 9). Both Type-1 and Type-2 pulses were distributed fairly evenly amongst the weekdays. We cannot ascribe statistical significance to the modest variations in the histogram bar heights in Fig. 9a or Fig. 9b. Type-3 pulses were less common during the weekends with the fewest pulses by far on Sunday (Fig. 9c). Type-4 pulses appear to be bimodally distributed, with more pulses occurring on Mondays and Tuesdays, but all other days being approximately equal (Fig. 9d).

Fig. 4 Different phases found for Type-1 pulse. Examples of pulses found in the time series data with **a** 0° , **b** 90° , **c** 180° phase shifts, **d** a false positive using the Type-1 pulse as training data

Table 3 Information about the identified pulses

	Total flagged	0°	180°	90°	False positive
Type-1	867	288	391	157	31
Type-2	760	283	275	78	124
Type-2a	652	237	247	44	NA
Type-2b	108	46	28	34	NA
Type-3	933	NA	NA	NA	NA
Type-4	407	NA	NA	NA	NA

We also evaluated the number of occurrences versus time-of-day for each ‘type’ of pulse (Fig. 10). Both Type-1 and Type-2 pulses had obvious peaks during the day, but were otherwise fairly quiet. Type-1 pulses had a single

large peak around 07:00 h in local time with a total of 680 pulses in this data set (Fig. 10a). Similarly, Type-2 pulses had narrow peaks at approximately 07:00 and 21:30 h in local time with a total of 615 pulses found in the data set (Fig. 10b). In contrast, Type-3 pulses had a clear diurnal signature with broad peaks in the mid-morning around 09:00 and 20:30 h in local time with an obvious quiet period between 01:00 and 04:30 h in local time (Fig. 10c). There were 790 Type-3 pulses in the data set. Lastly, Type-4 pulses had a lower number of occurrences but were very bimodal (Fig. 10d). The first peak was in the early morning between 01:00 and 04:00 h. The second peak was in the afternoon between 14:00 and 18:00 h local time with about twice as many pulses as in the first peak. There were a total of 367 Type-4 pulses in the data set.

Fig. 5 Sub-types found for Type-2 pulse. Examples of the two sub-types, **a** Type-2a, **b** Type-2b, found in the time series data using the Type-2 pulse as training data

5 Discussion

In an attempt to identify pulses in the data from the same source, we developed a filtering method called TDPCA. We visually identified four different spurious transients in the data to use as training data. We then compared the eigenvectors from these four types of training data to the eigenvectors of the windowed time series data allowing us to isolate pulses with similar characteristics and those occurring simultaneously at more than one observation station. While the algorithm found similar pulses when using a particular training datum, visual inspection revealed considerable variabilities for these pulse types (e.g., shape, polarity, time-of-day).

Type-1 and Type-2 were both relatively simple pulses with an irregular time-of-day distribution and yet a fairly uniform day-of-week distribution. They stimulated the same station-channel groups and had similar pulse durations. Despite lacking a strong preference toward a particular day of the week, the large discontinuities in the time-of-day histograms of Type-1 and Type-2 pulses suggest anthropogenic sources. Another interesting characteristic of these two types of pulses was the relative phase changes. The channels at each station experienced simultaneous phase changes, e.g., each experienced a 90° or 180° phase change in some of the identified pulses. These phase changes could be due to the field generating the current pulsing in the opposite direction. However, Type-1 and Type-2 pulses were clearly from different sources as can be seen by the variations in their relative polarities. Specifically, the vertical channels for Type-2 pulses had opposite polarities at the two stations, while the vertical channels in Type-1 pulses had coherent polarities. The

opposite was true for the east channels. Therefore, there is no way that the same current system can be responsible for both types.

An interesting attribute of this method is that it is insensitive to the relative amplitudes of the pulses; it is only sensitive to the correlations and phases of the relative channels. Using the Type-2 training data, we actually identified two sub-types (Type-2A and Type-2B) of pulses which had similar characteristics yet had very different amplitude ratios. These must be from different sources due to the nonlinear relationship in their amplitude ratios, which is an area for further exploration into source signatures.

Although Type-3 pulses were admittedly more complex and highly variable, we believe their source was more obvious. They had a distinct quiet period between 24:00 and 04:00 h in local time with a bimodal daily distribution during the morning and evening. This corresponds to the local ‘rush hours’ of daily commuters strongly suggesting we have isolated a signal from the Bay Area Rapid Transit DC electric railway. However, the variability in the pulses per day suggests these signals were either (a) not generated by routine train operations or (b) we do not have a large enough sample size. Since there were not any false positives evident for this type of pulse in this study, we believe the latter to be true. More of these pulses would be found if we relaxed the thresholds.

Because Type-4 pulse amplitudes are highly similar at stations several hundred kilometers apart, it is reasonable to suppose that the source of these pulses is approximately plane wave (McPherron 2005). Type-4 pulses exhibited a variety of waveforms with frequency content primarily in the 2–5 Hz band. These pulses would be classed as ‘P11’ in

Fig. 6 Examples of complex Type-3 pulses found using training data. **a** Type-3 pulse on 28 Jun 2014, 07:53:55. **b** Type-3 pulse on 30 Jun 2014, 16:38:20. **c** Type-3 pulse on 02 Jul 2014 19:00:30. **d** Type-3 pulse on 12 Jul 2014, 03:11:50. Note pulses are of varying durations but there are several common traits, for example the significant energy and correlation in 870-N and 809-Z

the space physics nomenclature (Jacobs et al. 1964). Type-4 pulses could be found at any time-of-day; however, they did display two daily peaks in the intervals 23:00–02:00 h and 15:00–21:00 h in local time, with the evening peak significantly more active than the late night peak. It is well understood that various natural magnetic processes exhibit a dependence on local time of day (McPherron 2005), but we could not identify a specific process consistent with the frequency content and time-of-day for these pulses.

There are several areas that warrant further investigation. In this analysis, we employed only the two dominant eigenvectors for signal pattern matching. It is possible the filter could be made less prone to false positives by adding a third eigenvector. Also we did not explore the effect on relaxing the number of eigenvectors down to one.

Furthermore, the only channels used for matching the cross-correlation pattern were those actively stimulated by the training pulse. In addition to requiring a fixed (high amplitude) correlation between these channels exhibiting the signal, we could also require poor correlation between those channels not exhibiting the signal with those that do. These additional constraints would require the windowed data to better match the training data, presumably leading to fewer false positives at the risk of identifying fewer pulses overall.

False positives were only observed for Type-1 (4%) and Type-2 (16%)—the application of this study to Type-3 and Type-4 pulses resulted in zero false positives. This contrast may be due to the number of training data channels. That is, the training data for Type-1 and Type-2 both used 5

Fig. 7 Examples of Type-4 pulses found using training data. **a** Type-4 pulse on 27 Jun 2014, 10:20:46. **b** Type-4 pulse on 28 Jun 2014, 08:01:29. **c** Type-4 pulse on 28 Jun 2014, 08:50:18. **d** Type-4 pulse on 29 Jul 2014, 12:34:20. Note pulses are of varying durations and shape but arraywide correlation is present in all cases

channels, while Type-3 and Type-4 had 9 and 6 channels, respectively.

It is plausible there are many more Type-3 and Type-4 events not admitted by an overly conservative threshold. If true, the thresholds on variance and/or eigenvector colinearity could be iteratively relaxed until some percentage (5% for example) of false positives appears. Another way to obtain more pulses would be to employ a cross-validation method. One could randomly draw from the pool of found pulses for a particular type, and rerun the TDPCA filter using these randomly selected pulses as the updated training data. While this could find more pulses, the selectivity may become unstable resulting in the filter returning many more false positives than pulses similar to the training type.

Lastly, in order to better understand these types of pulses and their origins, we could explore the ability of simple source models (e.g., electric and/or magnetic dipoles in different locations and orientations) to explain the physical phenomenon. Modeling the pulses would help identify cultural, solar, and other phenomena of interest in our data.

6 Conclusions

We visually identified several types of pulses within the electromagnetic time series data. When viewed at lower resolution, the pulses appeared quite similar, yet were very different upon inspection at higher resolution. Using

Fig. 8 The number of pulses per day. The number of **a** Type-1 pulses, **b** Type-2 pulses, **c** Type-3 pulses, **d** Type-4 pulses for each day of the 100 days of data (day 175 to day 267, which is June 24, 2014, to September 24, 2014). The vertical line is the day of the Napa earthquake. The red line is a 7-day median filter applied to the daily pulse counts

these manually identified pulses as training data and comparing the eigenvectors of each training datum to the eigenvectors from windowed time intervals, we built a process to automatically find and isolate pulses with similar characteristics. We have demonstrated this technique, TDPCA, is useful for finding pulses that are likely emitted from the same source as each relative training datum or ‘type.’ Analysis of the TDPCA outputs enabled

us to place further constraints on the possible sources of each pulse type, allowing for the identification of several sources such as the DC train noise and natural field events. It has also identified several sources of unknown (though likely cultural) origin. We are continuing research into this classification method and aim to follow up with a study modeling the sources as dipoles. At a minimum, being able to positively identify cultural noise will allow

Fig. 9 The number of occurrences versus day-of-week. **a** Type-1 pulses, **b** Type-2 pulses, **c** Type-3 pulses, **d** Type-4 pulses for each day of the week (daylight savings time, UTC-7) from Monday to Sunday

Fig. 10 The number of occurrences versus time-of-day. **a** Type-1 pulses, **b** Type-2 pulses, **c** Type-3 pulses, **d** Type-4 pulses in local time (daylight savings time, UTC-7) from June 24, 2014, to September 24, 2014

us to remove these events from our search for anomalous signals associated with the earthquake process. Ultimately, we hope to incorporate this method in our algorithmic framework at QuakeFinder to find and isolate pre-earthquake signals.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bendat JS, Piersol AG (1980) Engineering applications of correlation and spectral analysis. Wiley, New York. ISBN 0471058874
- Bleier T, Dunson C, Maniscalco M, Bryant N, Bamberg R, Freund F (2009) Investigation of ULF magnetic pulsations, air conductivity changes, and infra red signatures associated with the 30 October Alum Rock $M_{5.4}$ earthquake. *Nat Hazards Earth Syst Sci* 9:585–603. doi:[10.5194/nhess-9-585-2009](https://doi.org/10.5194/nhess-9-585-2009)
- Bleier T, Dunson C, Alvarez C, Freund F, Dahlgren R (2010) Correlation of pre-earthquake electromagnetic signals with laboratory and field rock experiments. *Nat Hazards Earth Syst Sci* 10:1965–1975. doi:[10.5194/nhess-10-1965-2010](https://doi.org/10.5194/nhess-10-1965-2010)
- Campbell WH (2003) Introduction to geomagnetic fields. Cambridge University Press, Cambridge. ISBN 0521571936
- Dunson JC, Bleier TE, Roth S, Heraud J, Alvarez CH, Lira A (2010) The Pulse Azimuth effect as seen in induction coil magnetometers located in California and Peru 2007–2010, and its possible association with earthquakes. *Nat Hazards Earth Syst Sci* 11:85–105. doi:[10.5194/nhess-11-2085-2011](https://doi.org/10.5194/nhess-11-2085-2011)
- Egbert GD, Eisel M, Boyd OS, Morrison HF (2000) DC trains and Pc3 s: source effects in mid-latitude geomagnetic transfer functions. *Geophys Res Lett* 27:25–28. doi:[10.1029/1999GL008369](https://doi.org/10.1029/1999GL008369)
- Fraser-Smith AC, Bernardi A, McGill PR, Ladd ME, Helliwell RA, Villard OG Jr (1990) Low-frequency magnetic field measurements near the epicenter of the M_s 7.1 Loma Prieta earthquake. *Geophys Res Lett* 17:1465–1468. doi:[10.1029/GL017i009p01465](https://doi.org/10.1029/GL017i009p01465)
- Fraser-Smith AC, Bernardi A, McGill PR, Villard OG Jr (1991) ULF magnetic field measurements near the epicenter of the M_s 7.1 Loma Prieta Earthquake. *Phys Earth Planet Inter* 68:45–63. doi:[10.1016/0031-9201\(91\)90006-4](https://doi.org/10.1016/0031-9201(91)90006-4)
- Freund FT (2007a) Pre-earthquake signals. Part I: Deviatoric stresses turn rocks into a source of electric currents. *Nat Hazards Earth Syst Sci* 7:535–541. doi:[10.5194/nhess-7-535-2007](https://doi.org/10.5194/nhess-7-535-2007)
- Freund FT (2007b) Pre-earthquake signals. Part II: Flow of battery currents in the crust. *Nat Hazards Earth Syst Sci* 7:543–548. doi:[10.5194/nhess-7-543-2007](https://doi.org/10.5194/nhess-7-543-2007)
- Gamble T, Goubau WM, Clarke J (1979) Magnetotellurics with a remote magnetic reference. *Geophysics* 44:53–68. doi:[10.1190/1.1440923](https://doi.org/10.1190/1.1440923)
- Han P, Hattori K, Hirokawa M, Zhuang J, Chen C-H, Febriani F, Yamaguchi H, Yoshino C, Liu J-Y, Yoshida S (2014) Statistical analysis of ULF seismomagnetic phenomena at Kakioka, Japan, during 2001–2010. *J Geophys Res Space Phys* 119:4998–5011. doi:[10.1002/2014JA019789](https://doi.org/10.1002/2014JA019789)
- Jacobs JA, Kato Y, Matsushita S, Troitskaya VA (1964) Classification of geomagnetic micropulsations. *J Geophys Res* 69:180–181
- Junge A (1996) Characterization of and correction for cultural noise. *Surv Geophys* 17:361–391. doi:[10.1007/BF01901639](https://doi.org/10.1007/BF01901639)
- Kappler K (2012) A data variance technique for automated despiking of magnetotelluric data with a remote reference. *Geophys Prospect* 60:179–191. doi:[10.1111/j.1365-2478.2011.00965.x](https://doi.org/10.1111/j.1365-2478.2011.00965.x)
- McPherron RL (2005) Magnetic pulsations: their sources and relation to solar wind and geomagnetic activity. *Surv Geophys* 26:545–592. doi:[10.1007/S10712-005-1758-7](https://doi.org/10.1007/S10712-005-1758-7)
- Szarka L (1988) Geophysical aspects of man-made electromagnetic noise in the Earth: a review. *Surv Geophys* 9:287–318. doi:[10.1007/BF01901627](https://doi.org/10.1007/BF01901627)
- Takeuchi A, Nagao T (2013) Activation of hole charge carriers and generation of electromotive force in gabbro blocks subjected to nonuniform loading. *J Geophys Res Solid Earth* 118:915–925. doi:[10.1002/jgrb.50111](https://doi.org/10.1002/jgrb.50111)
- Varotsos P, Alexopoulos K (1984) Physical properties of the variations of the electric field of the earth preceding earthquakes, I. *Tectonophysics* 110:73–98. doi:[10.1016/0040-1951\(84\)90059-3](https://doi.org/10.1016/0040-1951(84)90059-3)
- Vozoff K (1991) The magnetotelluric method. *Electromagn Methods Applied Geophysics* 2:641–672. doi:[10.1190/1.9781560802686.ch8](https://doi.org/10.1190/1.9781560802686.ch8)