ULTRA-LOW FREQUENCY MAGNETIC FIELDS IN THE SAN FRANCISCO BAY AREA: MEASUREMENTS, MODELS, AND SIGNAL PROCESSING

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Thomas T. Liu
June 1999

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Antony C. Fraser-Smith
(Principal Adviser)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

G. Leonard Tyler

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert M. Gray

Approved for the University Committee on Graduate Studies:

Abstract

Efforts to detect magnetic precursors to earthquakes are often complicated by the presence of man-made noise. In the San Francisco Bay Area of California, the magnetic fields generated by the Bay Area Rapid Transit (BART) system exhibit power spectral amplitudes that are up to 1000 times greater than those of the natural geomagnetic background fields in the ultra-low frequency (ULF, 0.01 to 10 Hz) band. Measurements from an array of three-axis magnetometers show that the BART interference appears as a series of transients with temporal durations ranging from 4 to 30 seconds. In addition, at each observation site, the polarization vectors of the transients can be bounded by a non-isotropic subset of the three-dimensional polarization space. Electromagnetic modeling results support these experimental findings.

A framework for the identification and removal of the BART transients is proposed. The identification problem is formally posed as the detection, in the presence of $1/f^{\gamma}$ noise, of transients of known shapes, but unknown amplitudes, delays, and scales, where the range of unknown scales corresponds to the observed range of temporal durations. A detection method based upon pattern matching in the undecimated discrete wavelet transform (UDWT) domain is introduced. The method detects the most significant BART transients, while discriminating against non-BART-related transients with a false alarm rate of approximately 8 percent. As an additional constraint, only those transients whose polarization vectors lie within the previously described bounding subsets are considered to be BART-related.

An estimate of each detected transient is formed in the UDWT domain from those transform coefficients that are in the vicinity of the transient location, and whose amplitudes exceed one standard deviation of the transform coefficients computed for the

assumed $1/f^{\gamma}$ background noise. The estimate is converted to the time domain with an inverse translation invariant wavelet transform, and then subtracted from the original time series. Removal of the transients significantly increases the ability to test for the presence of precursor signals.

Acknowledgments

This work was supported in part by U.S. Geological Survey (USGS) Grants No. 1434-HQ-97-GR-03124 and 1434-HQ-96-GR-02715, Electric Power Research Institute (EPRI) Grant No. WO8035-2, and a Gerald J. Lieberman Fellowship. In addition, many people have contributed in one way or another to the success of this research project.

I would like to thank my principal advisor, Professor Antony C. Fraser-Smith, for giving me the opportunity to work on this project. His patience, constant good cheer, and genuine concern for students are an inspiration to me as I continue with my career. Professor G. Leonard Tyler served as my associate advisor, and am I grateful for his support and encouragement throughout my graduate education. I would also like to thank Professor Robert Gray for serving as the chairman of my defense panel and the third reader of this thesis; and Professor Howard Zebker for serving as the fourth member of my panel.

The installation of the ULF magnetic field stations would not have been possible without the generous assistance of Doug Chrissan, Martin Füllekrug, Darcy Karakelian, Judy Liu, and Bill Trabucco. They willingly braved difficult conditions to help me dig trenches, drill holes, and accomplish the myriad tasks involved with a field installation. Paul McGill and Cal Teague provided valuable technical advice in the early stages of the project and were always willing to talk about the latest Macintosh programming secrets. Mike Johnson and Jeff DiCarlo provided software code and circuitry for the GPS timing system of the Mission Peak system. Malcolm Johnston, Nancy King, Jim Lienkaemper, and Doug Myren from the USGS supplied much needed advice on good sites for the systems. John Maunder and Jimmy Parsons from the Chabot Gun Club, Jack Kenny, supervisor of Chabot Regional Park, Dan Reasor, supervisor of Mission Peak Regional Preserve, and Nancy Schley of the East Bay Regional Park District (EBRPD) provided invaluable cooperation and assistance in the placement of our systems on EBRPD lands.

Ed Nichols, Ugo Conti, and Giuseppe Guarino of EMI Inc. made it possible for me to install low-noise induction coils at Mission Peak and were always very helpful in answering my many questions.

Frank Morrison and Sierra Boyd of the U.C. Berkeley Department of Materials Science and Mineral Engineering, and Doug Neuhauser and Barbara Romanowicz of the U.C. Berkeley Seismographic Station kindly provided advice, software support, and data from the SAO and PKD1 geomagnetic observatories. Darcy Karakelian graciously shared her data from the station at Jasper Ridge.

John Evans, Abdulhaque Shaikh, Peter Todd, and Michael Epperson from the Bay Area Rapid Transit (BART) Engineering Department generously shared their knowledge of the BART system. Chester Livingston and Joshua Sheldon of Pacific Gas and Electric kindly provided me with access to the BART traction power data.

I thank Dr. H.T. Tang of EPRI for his assistance and encouragement during the initial phase of this project. I am also grateful for the advice provided by Bilal Ahmad, Alper Erdogan, Dana Porrat, and Cal Teague during a practice run of my dissertation defense. I thank John Baron and Snezana Maslaković for their assistance on the day of the defense and on countless other occasions over the years. Max Lui, Sarah Lundgren, Erik Toomre, and Dave Wilson have been great friends since before my Stanford days, and I owe them all a debt of gratitude for their continued support and inspiration. Frank Doran Sensei and the members of the Stanford Aikido Club have played an integral part in my overall education, and I am grateful for their generous spirits.

I thank my parents for their constant patience and encouragement, my sister Judy for her laughter and love, and my uncle Jung-Chu Lin for inspiring me to study electrical engineering.

My years at Stanford have been blessed with the love and companionship of my wife Shin. I am truly grateful for her compassion, her sense of humor, and her constant delight with the beauty of life. Our son Adrian has already taught me much about wonder and joy. Thanks, Adrian.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Monitoring for Magnetic Precursors to Earthquakes

The ability to predict earthquakes is of far-reaching practical interest. Over the last several decades, significant progress has been made in the field of *long-term* earthquake prediction, which typically yields probabilities of occurrences of earthquakes and estimates of potential damage over periods of years or decades. Such information has been useful to both urban planners and structural engineers involved in hazard mitigation efforts. However, the ability to consistently predict earthquakes on shorter time scales, such as hours or days, continues to elude us. *Short-term* prediction of earthquakes would be useful for the evacuation of at-risk areas and also for hazard mitigation in certain industrial sectors, such as the electrical power industry.

One of the more promising areas for short-term prediction is the monitoring of ultra-low frequency (ULF, 0.01 to 10 Hz) magnetic fields. Previous observations [*Park*, 1996; *Park et al.*, 1993] suggest that major earthquakes may be preceded by large anomalous increases in ULF magnetic activity. For example, significant increases in ULF magnetic activity were observed days and hours prior to the 1989 Loma Prieta earthquake (M 7.1) which impacted much of the San Francisco Bay Area [*Fraser-Smith et al.*, 1990]. There are a number of processes by which earthquakes potentially can generate electromagnetic fields (e.g., electrokinetic, piezomagnetic, and piezoelectric effects [*Park et al.*, 1993]), but additional experimental evidence, including more measurements of anomalies, is required to determine which processes are actually responsible for the reported observations.

Efforts to monitor ULF fields in regions of interests are often hampered by the presence of man-made interference. The problem is of particular significance in urban areas where both the interference levels and the potential damage due to an earthquake are the greatest. This thesis focuses on the situation in the San Francisco Bay Area of California. Similar situations exist in Japan and Greece, among other areas. For example, in Greece a controversy currently exists as to whether signals that were reported as earthquake precursors were in fact due to electric currents from an industrial power plant [*Gruszow et al.*, 1996].

In the San Francisco Bay Area, the Hayward Fault provides a promising and important setting for furthering our understanding of ULF earthquake precursors. There are several reasons for this. First, the probability of an earthquake of magnitude 7 or larger occurring somewhere on the fault in the next 30 years is estimated to be 45 percent [*Working Group on California Earthquake Probabilities*, 1990]. Second, there is a wide array of instrumentation, including seismometers, strainmeters, dilatometers, and creepmeters, in and along the fault. The extent and variety of the instrumentation makes it possible to study in detail the relationships among various phenomena that may precede an earthquake. Finally, the Hayward Fault runs through a heavily populated area, and a major earthquake on the fault has the potential to affect thousands of lives and cause billions of dollars of property damage. Additional knowledge about the behavior of the fault may help to mitigate the effects of an earthquake.

Despite the importance of measuring ULF magnetic fields along the Hayward Fault, efforts to establish a long-term monitoring program have been deterred by the high-levels of man-made electromagnetic noise in the vicinity of the fault. The predominant source of interference is the Bay Area Rapid Transit (BART) System, a 1000 volt direct-current electric railway that serves the San Francisco Bay Region [*Fraser-Smith and Coates*, 1978]. Over most of its 70 km length the fault lies within 5 km of the closest BART tracks. The power spectral amplitudes of the ULF magnetic interference due to BART can be as much as 1000 times greater than typical natural background levels, making it difficult to detect earthquake related anomalies.

The goal of this dissertation is to show how useful measurements can be made despite the severity of the interference. Underlying our approach is the observation that the BART interference has unique spatial and temporal features that distinguish it from the natural background. We make use of these features to develop a combined space-time approach

for the identification and removal of the interference, and to show that the application of the method improves our ability to monitor for precursors.

1.2 Review of Previous Methods

The problem of electromagnetic interference due to electrical railways is not new. It was first reported in 1890 by the Greenwich Observatory in London [*Burbank*, 1905] and has been studied by numerous researchers over the years [*Szarka*, 1988]. A wide variety of techniques for minimizing the effects of the interference have been suggested [*Junge*, 1996], primarily by workers in the field of magnetotellurics. In this section, we review a number of the previously proposed methods and examine their applicability to the task of monitoring for earthquake precursors.

The simplest method for dealing with man-made interference is to move the measurement systems to a region far from the source. For example, researchers have tried to locate geomagnetic observatories at least 30 km from the nearest electrical railway line [*Chaize and Lavergne*, 1970; *DuPouy*, 1950; *Yanagihara*, 1977]. This method is valid when one is interested primarily in measuring the fields due to upper atmospheric or magnetospheric sources, but is clearly not applicable when the interference sources are located within the region of interest, as is the case with the BART system and the Hayward Fault.

In the magnetotelluric method, simultaneous measurements of magnetic and electric fields are used to calculate the local impedance of the earth. The impedance values are then used to estimate the regional electrical conductivity structure [*Vozoff*, 1991]. Since the local structure is of interest, relocation of the systems is not a viable option when faced with man-made interference. As a consequence, other techniques have been developed. A number of the approaches that we describe make use of a so-called *remote reference* site. Such a site is located far from the region of interest and is assumed to be free of interference. Because geomagnetic fields are highly coherent over distances of several hundred kilometers or more [*Egbert and Booker*, 1989], the fields at the remote reference site can be used to form an estimate of the noise-free fields at the local site of interest. We show examples of this estimation process in Chapter 3.

1.2.1 Outlier Replacement

When the interference consists primarily of sparsely distributed spikes, one approach is to detect and then replace the affected data. *Fontes et al.* [1988] deleted data segments whose partial variance exceeded 3 to 4 times the average variance observed at a certain site. The gaps created by deletion were filled in using estimates based upon the remaining data — e.g., maximum entropy extension of the good data segments. *Larsen et al.* [1996] identified outliers in their data set as those points whose absolute deviation from the median was greater than $4.5\sigma_m$ where σ_m is the median absolute deviation. Outliers were replaced by values extrapolated from the remaining data.

The outlier replacement approach fails in instances when the time series are dominated by the interference. In these cases the majority of the data points can be labeled as outliers, and it is not practical to replace the outliers with values extrapolated from a few 'noise-free' points. As we show in Chapter 3, the time series data we acquire are almost completely dominated by BART interference. In addition, when viewed within the context of monitoring for unknown anomalies, the use of an outlier criterion is difficult to justify since the outliers may be due to the very anomalies we wish to detect.

1.2.2 Data Selection or Downweighting

In instances where the interference is especially severe, various researchers have chosen to select only interference-free data segments or to downweight segments with significant interference. Downweighting refers to the process of assigning less weight to certain data segments when constructing an estimate. For example, *Fontes et al.* [1988] used their partial variance criterion to determine which data segments were not affected by interference. Only these segments were used to form estimates of the local conductivity.

Egbert [1997] conducted a magnetotelluric study with a three station array consisting of two stations just to the south of San Jose, CA, area and one station south of Hollister, CA. The San Jose stations showed clear signs of BART interference, while the Hollister station was shown to be relatively free of interference and was used as a remote reference station. Egbert showed that frequency domain, multivariate statistical techniques could be used to form a test diagnostic that indicated the presence of coherent noise sources such as BART. Calculation of the diagnostic required the availability of a noise-free remote

reference site. Egbert used the diagnostic to downweight or eliminate noisy data segments before forming an estimate of the conductivity.

Liu and Fraser-Smith [1996a] used data selection to determine if anomalous increases in ULF activity occurred prior to a M 4.7 earthquake located 50 km to the southeast of station LKC (see Chapter 2 for map). They examined a month's worth of half hourly average spectral indices during periods when no BART trains were scheduled to be in operation. These periods were 02:00 to 03:00 local time on weekdays and 02:00 to 04:00 local time on weekends. After comparing these spectral indices with similar indices computed for a remote reference site located in Southern California, they concluded that no anomalies were observed during these early morning time periods.

Data selection represents a reasonable approach for monitoring for precursors if there are periods during which the interference is absent. The major disadvantage in the case of the BART interference is that the useful period of observation is limited to about 2 hours per day, roughly one tenth of the overall daily monitoring period. If the precursors exhibit sustained activity over many days, as was the case with the 1989 Loma Prieta anomalies, then the activity may be detectable, even with the limitations of data selection. However, precursors that appear as transient phenomena lasting for only a few hours at a time could go undetected.

1.2.3 Direction Dependent Bayesian Estimation

Santarato and Spagnolini [1995] addressed the problem of performing a magnetotelluric survey in Northern Italy, a region with significant interference due to electric railways. After first whitening the time series so that step-like noise components appeared as impulses, they examined the residual signals obtained by subtracting an estimate of the natural background, obtained using a remote reference or by an adaptive filtering technique, from the local time series. They assumed that the residuals e were the sum $e = n + e_p$ of a noise component n modeled by a mixture model with an impulsive portion with variance σ_n^2 and a prediction error e_p modeled by a Gaussian model with variance σ^2, where $\sigma^2 \ll \sigma_n^2$. From the observed residuals e, they calculated the Bayes estimate of the noise $\hat{n} = E[n|e]$ and then subtracted this estimate from the original time series. To take advantage of the fact that the interference due to electric railways often exhibits strong polarization (see, for example, Chapter 3), directional dependence was introduced by allowing the parameters of the mixture model to vary with azimuth.

From the point of view of precursor detection, the Bayesian estimation approach suffers from the same drawback as the outlier replacement approach. Namely, only the amplitude of the residual signal is used to distinguish the man-made noise from the prediction error. The portion of a signal that exceeds a predetermined threshold is considered to a man-made component even if it happens to come from a precursor. The use of directional dependence partially addresses the shortcoming by effectively increasing the threshold in directions where the probability of observing man-made noise spikes is decreased.

1.2.4 Filtering Approaches

In cases where the man-made noise is confined to certain narrow frequency bands (e.g., power line harmonics), frequency domain filtering approaches can be used [*Fontes et al.*, 1988] to remove the interference. Interference due to electric railways, however, tends to be fairly wideband (see Chapter 3). While a filter with a wide bandwidth could be used to reduce the interference, such a filter would also reduce any precursor activity in that band.

Hattingh [1988] proposed an adaptive noise canceling approach for the removal of interference from geophysical time series. The method uses a slight modification of the adaptive noise canceler proposed by *Widrow et al.* [1975]. It assumes that two channels are available, each consisting of the desired signal component plus a noise component that is not correlated with the signal. In addition, it assumes that the noise processes on the two channels are uncorrelated. The method adaptively filters the noisy signal on one channel to provide the best match to the noisy signal on the other channel. Since the noise processes are uncorrelated, the filtering process attempts to remove the noise from the channel. The difficulty in using this approach in our situation is that both the desired signal (e.g., earthquake precursors) and the noise (i.e., BART) can be highly correlated across the channels in the array.

1.3 The Proposed Approach

The majority of the methods described above make the assumption that the amplitude of the man-made interference is on average much larger than the amplitude of either the natural signal of interest or, in the case of *Santarato and Spagnolini* [1995], the prediction error. In general this assumption is applied to the whitened time series. There is an

implicit second assumption that the man-made signals are not as 'smooth' as the natural signals, so that a whitening procedure amplifies the difference in the respective amplitudes. While these assumptions typically are suitable for magnetotelluric studies, where the natural signals of interest are well understood, they are difficult to justify for an earthquake precursor research and monitoring effort. Because very little is known about the temporal features of precursor signals, we must be careful in making assumptions that may result in the mis-identification of a possible precursor signal as a man-made signal. As an example, we consider the Loma Prieta precursor data [*Fraser-Smith et al.*, 1990], which are in the form of spectral indices. These data indicate that the precursor amplitudes can be quite large and that the spectral energy tends to fall off with increasing frequency. As a result, the whitened precursor signals will be broadband and may not necessarily be 'smooth' when compared to the man-made interference. The amplitudes of the whitened precursors may be larger than or comparable to the man-made signal amplitudes. Thus, an amplitude based interference removal procedure carries the risk of removing the precursor signals. The introduction of directional dependence [*Santarato and Spagnolini*, 1995] reduces this risk somewhat by increasing the amplitude threshold in directions where the probability of observing man-made interference is low. However, even with directional dependence, precursors that either lie in the direction of the man-made interference or have sufficiently large amplitudes will be incorrectly identified as interference.

In this thesis we develop a description of the temporal and spatial features of the man-made interference and then to use that description to identify and remove the interference. In the temporal domain, we characterize the interference as the superposition of transients with a range of characteristic time scales. We use a mathematical transient model that approximates the general shape of the observed transients associated with BART. With this description we are able to distinguish the man-made transients from possible precursor signals of similar amplitudes but with different shapes or characteristic time scales.

Our working assumption is that the precursor signals are unlikely to exhibit the same temporal features as the interference. Still, we cannot completely rule out this possibility. To reduce the risk of mis-identification when only the time-domain method is employed, we develop a method in which the signal polarization is used to distinguish further the precursor signals from the man-made interference. In Chapter 3 we show that the polarization of the BART interference has a strong directional component. Then, in Chapter 7,

we construct three-dimensional subsets that describe the distribution of the BART polarization vectors at each observation station and use these subsets to decide if a detected transient is related to BART. Note that in contrast to the method of *Santarato and Spagnolini* [1995], our use of the polarization information should be viewed as a secondary step that is used to constrain the results of the primary step in which the identification of the interference is performed solely in the time domain. In addition, our method makes simultaneous use of all stations in the array, whereas previous methods have tended to treat each station independently. We show in Chapter 7 that simultaneous use of data from all stations enables us to determine if a transient is man-made in situations where data from a single station would be inadequate.

1.4 Outline of the Dissertation

In this dissertation we provide a comprehensive description of our efforts to monitor ULF magnetic fields along the Hayward Fault. We begin in Chapter 2 with a description of the measurement systems. We describe the design of a typical system and discuss methods for obtaining estimates of the magnetic field.

Chapter 3 describes in detail the nature of the interference from BART. To the best of our knowledge, this is the most complete description presently available of the magnetic interference due to a direct current electric railway system. We present example time series, polarization plots, power spectra, and amplitude histograms. We also present the results of applying wavelet transform analysis to the geomagnetic background.

Chapter 4 describes how the magnetic fields are generated by the BART system. We present a simplified circuit model of the BART system and use it to calculate the propulsion and leakage currents due to trains at various points in the system. Using these currents, we calculate the magnetic fields that would be observed at the stations in the array.

Chapter 5 provides a brief overview of our approach for the identification and removal of BART transients. The identification of transients relies upon a novel, undecimated wavelet transform based detection method. We explain the theory and implementation of the detector in Chapter 6.

Chapter 7 synthesizes the results of previous chapters and presents a comprehensive approach for the identification and removal of BART transients. We show that by identi-

fying and removing the most significant transients, we can significantly improve our ability to detect anomalous increases in the geomagnetic background.

Finally, Chapter 8 summarizes the results of the dissertation and suggest areas for future research.

Earlier versions of the work presented in this dissertation have appeared in [*Liu and Fraser-Smith*, 1996a,b, 1998a,b,c].

1.5 Contributions

We briefly list the contributions described in this thesis. A more detailed summary is provided in Chapter 8.

1. Extensive characterization, using measurements from an array of multi-component sensors, of the magnetic fields due to BART.

2. The use of circuit and propagation models to calculate the magnetic fields due to BART.

3. Formulation of a new method for the detection of transients in the presence of $1/f$ noise.

4. Development of a coherent framework for the identification and removal of transients due to BART.

Chapter 2

Measurement Systems

2.1 Introduction

In this chapter we describe the ULF magnetic field measurement systems. After a brief overview of the history and locations of these systems, we outline the design of a typical system. We then introduce expressions for the transfer function associated with each system. For later chapters, we will find it useful to form an estimate of the underlying magnetic field from the acquired time series data. To do this we deconvolve the time series data using the transfer function descriptions. We discuss various options for performing the deconvolution and then describe our specific implementation.

2.2 Overview

The primary experimental data for our study were obtained from three ULF magnetic field measurement systems situated in the San Francisco Bay Area. Two of these systems Lake Chabot (LKC) and Mission Peak (MPK) are permanently installed along the Hayward Fault, while the third Jasper Ridge (JSR) was temporarily operated at a site near Stanford University. We refer to the collection of these three systems as the SFBay array.

The LKC system was installed in August of 1995 at a site within the Chabot Regional Park at Castro Valley, CA. It is located approximately three kilometers from the San Leandro salient section of the fault. The salient is a significant bend in the fault characterized by high seismic activity. It is believed that the salient divides the northern and southern sections of the fault and may serve as a starting point for large earthquakes [*Lienkaemper*,

1989; *Lienkaemper et al.*, 1991]. The MPK system was installed in September of 1996 at a site within the Mission Peak Regional Reserve in Fremont, CA. It is located near the southern end of the Hayward Fault and lies halfway between the Hayward and Calaveras Faults. The Jasper Ridge (JSR) system was operated on a temporary basis from September 1997 through September 1998 by D. Karakelian of the Geophysics Department at Stanford University. All three systems acquire three axes of magnetic data, while the LKC system also acquires two horizontal axes of tilt data. The time series data are stored to disk and are accessible either by phone or by site visits. Spectral estimates are calculated at LKC and MPK and are also available by remote access. Both the time series and spectral data are archived at Stanford.

In addition to the SFBay data, we made use of experimental data from seven systems located in other regions of California. Five of these systems, COR, HAL, VAR, TBM and PFO, are operated by Stanford University (SU), while the remaining two, SAO and PKD1, are operated by the Berkeley Seismographic Station (BKS). Table 2.1 provides a listing of the systems. Maps showing the system locations are presented in Figures 2.1 and 2.2.

System	Location	Coordinates	Axes	Inst.
COR	Corralitos	37.015° N, 121.804° W	NS	SU
HAL	Haliburton Ranch	35.888° N, 120.381° W	NS	SU
JSR	Jasper Ridge	37.404° N, 122.238° W	NS,EW,V	SU
LKC	Lake Chabot	37.738° N, 122.103° W	NS,EW,V	SU
MPK	Mission Peak	37.551° N, 121.876° W	NS,EW,V	SU
PKD1	Haliburton Ranch	35.873° N, 120.425° W	NS,EW,V	BKS
PFO	Pinõn Flat Observatory	33.609° N, 116.455° W	NS	SU
SAO	Hollister	36.764° N, 121.446° W	NS,EW,V	BKS
TBM	Table Mountain	34.400° N, 117.700° W	NS	SU
VAR	Varian Ranch	35.926° N, 120.394° W	NS	SU

Table 2.1: ULF Geomagnetic Systems in California. The symbols SU and BKS are abbreviations for Stanford University and Berkeley Seismographic Station, respectively. The symbols NS and EW refer to North-South and East-West geomagnetic axes for the SU systems and to the corresponding geographic axes for the BKS systems. The V symbol refers to the vertical axis parallel to the local gravity vector.

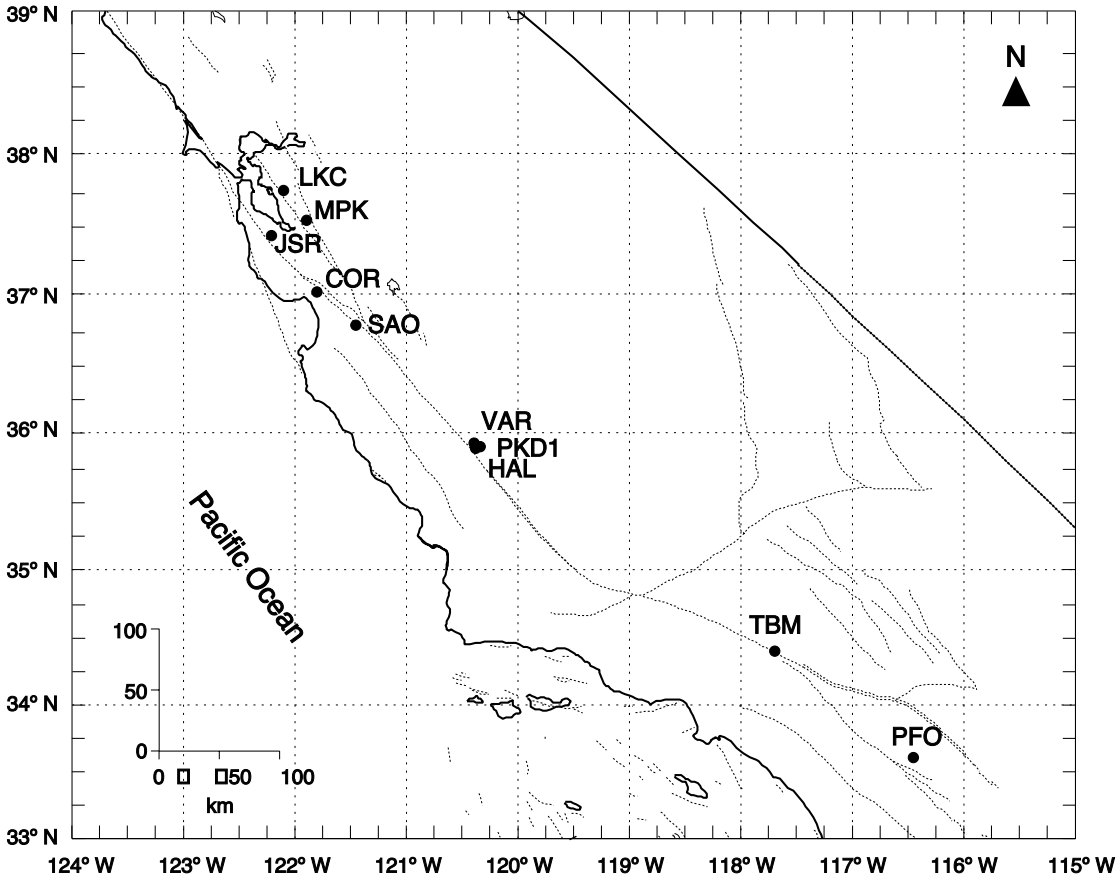


Figure 2.1: Geomagnetic systems in California. Dotted lines represent earthquake faults.

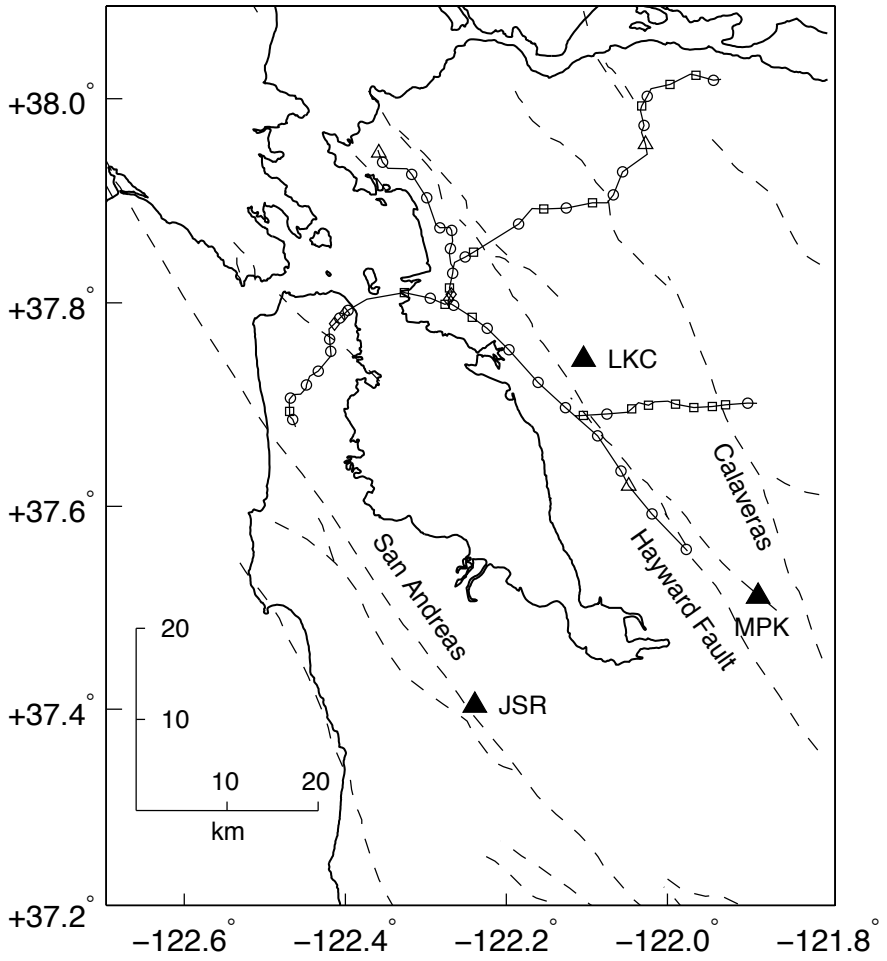


Figure 2.2: The SFBay array. Open symbols and solid lines represent the BART stations and tracks, respectively. The dashed lines represent active earthquake faults.

2.3 The MPK System

To provide some insight into the design of the geomagnetic measurement systems we provide a description of the MPK system. A block diagram of the MPK system is shown in Figure 2.3. The PKD1, SAO, and JSR systems are similar in design to the MPK system and make use of the same type of commercial induction coil sensors, while the LKC system and the remaining SU systems employ induction coil sensors designed and built at Stanford. More detailed descriptions of the LKC and MPK systems are provided in [*Liu and Fraser-Smith*, 1996a].

2.3.1 Coils and Amplifiers

The MPK system employs BF-4 and BF-7 Magnetic Field Induction Sensors manufactured by Electromagnetic Instruments Inc. (EMI) of Richmond, CA. Each sensor consists of a coil and a low-noise pre-amplifier sealed inside a waterproof, impact resistant Nema G-10 fiberglass tube. The length, diameter, and weight of the BF-4 sensors are 142 cm, 6 cm, and 7.9 Kg, respectively. The BF-7 sensor is roughly equivalent in performance to the BF-4, but has a shorter length (104 cm) to facilitate installation in the field. The BF-4 sensors are used to measure the two horizontal magnetic components, while the BF-7 sensor is used to measure the vertical component.

The EMI sensors are widely used by the magnetotelluric research community. They are based upon a feedback design [*Morrison et al.*, 1984] which makes the frequency response relatively flat above a frequency of 0.3 Hz and allows the gain to be determined primarily by the value of a precision feedback resistor. The design also reduces the magnetic field imposed on the μ-metal core and improves the linearity of the measurements. The recommended operating range of the sensor is 0.0003 Hz to 500 Hz. The inductance and natural coil resonance of the sensor are 1250 H and 140 Hz, respectively.

The horizontal sensors are placed in 10.2 cm diameter PVC pipe and buried at a depth of approximately 75 cm. They are aligned in the geomagnetic North-South and East-West directions. The vertical sensor is placed in a 1.2 meter deep hole lined with 15.2 cm diameter PVC pipe. The minimum distance between any two coils is eight meters. The distance from the sensors to the data acquisition system is 270 meters.

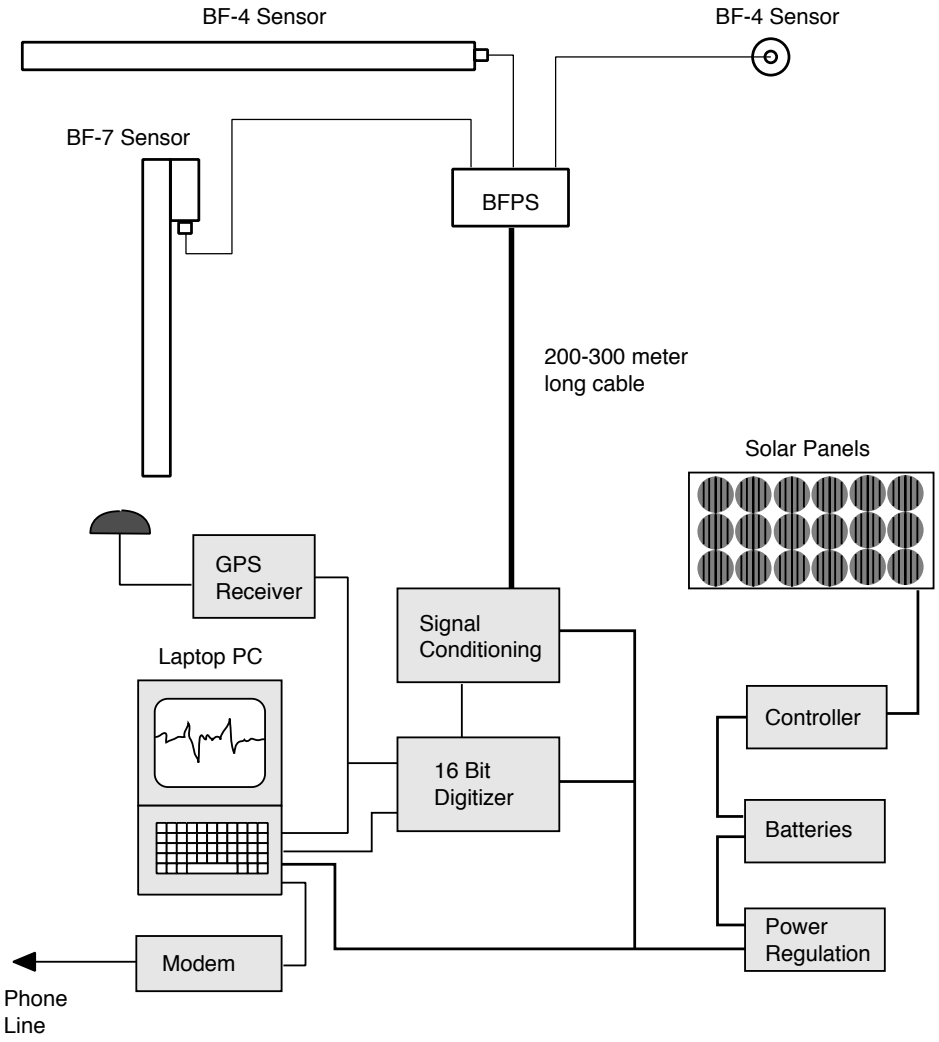


Figure 2.3: Block diagram of the system installed at Mission Peak Regional Preserve. The hard disk drive is contained within the laptop computer.

2.3.2 Data Acquisition, Timing, and Telemetry

A National Instruments DAQPad-MIO-16XE-50 16 bit data acquisition system is used to digitize the data. Each channel is sampled at a rate of 640 samples per second. The anti-aliasing filters are 4 pole Butterworth filters with a cut-off frequency of 100 Hz. An on-site laptop computer (Cyrix 586, 100 MHz processor, Microsoft Windows NT 4.0) receives and processes the data from the digitizer. The 16 bit data are digitally filtered and downsampled by a factor of 16 to a rate of 40 samples per second. The filtering and downsampling operation increase the dynamic performance of the acquisition system to approximately 18 effective bits. The time series data are compressed and stored on a magnetic disk. With the use of Steim compression algorithms (obtained from J. Steim, Quanterra Inc.) and a 810 MB hard disk, about 100 days of data can be stored. These algorithms perform lossless coding of the first and second differences of the data. Power spectral estimates are calculated every 15 minutes on 30 minute blocks of data and are used to compute MA indices covering the frequency range 0.001 to 10 Hz. Details of the spectral estimation procedure can be found in [*Liu*, 1996].

A Garmin 25 GPS receiver and a clock synchronization circuit designed and built at Stanford are used to provide timing signals accurate to several micro-seconds. These signals are used to synchronize the DAQPad data acquisition system. In addition, the timing information is used to time-stamp the stored time series data.

Telemetry is provided by a 28.8 Kbaud PCMCIA modem card operating on a standard phone connection. Commercial communications software allows remote control of the on-site computer and access to data stored on the magnetic disk.

2.3.3 Power System

The power system consists of three 75 Watt solar panels, two deep-cycle batteries, and assorted control and regulation circuitry. The system is designed to operate through at least three cloudy days. The maximum power dissipation of the MPK system is 20 watts.

2.4 System Transfer Functions

In order to use the data from our array, we need to know the transfer function that relates the output of each system to the magnetic field that we are trying to measure. In this

section we begin with a general description of the signal path common to all the systems. We then describe in detail the transfer functions of the various systems.

2.4.1 Signal Path Overview

All the systems consist of an analog subsystem followed by a digital subsystem. We refer to the analog subsystem as the *magnetometer*. It includes the induction coil sensor array and various stages of amplification and filtering. We use the notation $h(t)$ for the transfer function of each channel of the magnetometer, where $h(t)$ incorporates both the $db(t)/dt$ response of the induction coil sensor and the effects of the amplifiers and filters. The output of the magnetometer is $y(t) = b(t) * h(t)$ where $b(t)$ is the relevant component of the magnetic field. In the frequency domain we have $Y(f) = B(f)H(f)$ where $Y(f)$, $B(f)$ and $H(f)$ are the Fourier transforms of $y(t)$, $b(t)$, and $h(t)$, respectively. In describing $h(t)$ we will also make use of its Laplace transform $H(s)$, where $H(s) = H(f)$ for $s = j2\pi f$.

The digital subsystem consists of a digitizer followed by digital filters and downsamplers. The digitizer samples and quantizes $y(t)$ to form a discrete signal $y[n]$ such that

$$y_s[n] = y(nT_s) + e[n] \tag{2.1}$$

where T_s is the sampling period and $e[n]$ represents quantization noise. For the MPK, JSR and BKS systems, $y_s[n]$ is digitally filtered and downsampled to form the sequence $y[n]$. That is,

$$y[n] = (y_s[n] * h_2[n]) \downarrow M \tag{2.2}$$

where $h_2[n]$ is a lowpass filter and M is the downsampling factor. For the LKC system, we have $y[n] = y_s[n]$, i.e., $h_2[n] = \delta[n]$ and $M = 1$.

In the frequency band of interest, 0.01 to 5 Hz, the effects of the sampling and digital filtering operations on the overall response of the system are negligible. As a result, we will focus on the analog transfer function $h(t)$ in the presentation that follows.

2.4.2 LKC Transfer Function

In this section we describe the LKC transfer function, which is also the transfer function of all the other SU systems except for MPK and JSR. The transfer function is

$$H_c(s) = G_0 H_{ci}(s) H_{hi}(s) H_{lo}(s) H_8(s) \tag{2.3}$$

where

$$H_{ci}(s) = \frac{s}{s+\beta} \tag{2.4}$$

and G_0 is a gain factor. $H_{ci}(s)$ models the $db(t)/dt$ response of the induction coil and the effect of passive elements in the coil and pre-amplifier. $H_{hi}(s)$ is a second-order Butterworth high pass filter with a nominal cut-off frequency (f_c) of 0.005 Hz, while $H_{lo}(s)$ is a second order Butterworth low pass filter with a nominal $f_c = 12.5$ Hz. $H_8(s)$ is an eighth-order switched capacitor Butterworth low pass filter with a nominal $f_c = 13.5$ Hz. The magnitude response of the LKC system is shown in Figure 2.4. The sampling rate for the LKC system is 30 samples per second (SPS) and the digitizer resolution is 12 bits.

The parameter β is given by the expression $\beta = \frac{R_s+R_c+R_{in}}{L_s}$, where R_s and L_s are the series resistance and inductance of the coil, R_c is the resistance due to an RFI filter, and R_{in} is the input resistance of the pre-amplifier. Nominal values for R_{in} and R_c are $4000\,\Omega$ and $200\,\Omega$, respectively. The measured value of R_s is between $170\,\Omega$ and $180\,\Omega$, while the value of L_s is estimated to be between 72 H and 97 H (P. McGill, personal communication). A fit of the theoretical magnitude response $|H(f)|$ to the measured LKC response indicates that $\beta \approx 44$ for the vertical channel and $\beta \approx 40$ for the horizontal channels. Assuming that $R_s = 175\,\Omega$, the values for L_s are 109 H and 99 H for β values of 40 and 44, respectively. Because of the uncertainty in the measured values of L_s, we use the values of β obtained from the model fit. The series inductance of the coil results in a roll-off of the $db(t)/dt$ response at the pole frequencies 6.4 Hz and 7 Hz, corresponding to β values of 40 and 44, respectively.

2.4.3 MPK Transfer Function

The MPK system transfer function is

$$H_m(s) = G_0 H_{mi}(s) H_{lo}(s) H_{aa}(s) \tag{2.5}$$

where

$$H_{mi}(s) = \frac{s}{s+\omega_{hi}} \tag{2.6}$$

and

$$H_{lo}(s) = \frac{\omega_{lo}^2}{s^2 + \sqrt{2}s\omega_{lo} + \omega_{lo}^2}. \tag{2.7}$$

The $H_{mi}(s)$ term models the $db(t)/dt$ response of the induction coil and the effect of the feedback design. Typical values for G_0, ω_{hi}, and ω_{lo} are 0.3 volts/nT, $2\pi(0.2)$ radians per second, and $2\pi(450)$ radians per second, respectively. EMI Inc. determines the actual values with a calibration measurement. $H_{aa}(s)$ is the response of the first stage anti-aliasing filter. It is a 4th order Butterworth lowpass filter with $f_c = 100$ Hz and consists of two second order Sallen-Key active filter sections. The system magnitude response $|H_m(s)|$ is shown in Figure 2.4.

In the digital subsystem, $h_2[n]$ is a 1281 point finite impulse response (FIR) filter with a cut-off frequency of 17 Hz. The filtering operation is performed with double precision floating point arithmetic, and the result is truncated to 20 bits of precision. As mentioned previously, the downsampling factor is $M = 16$.

2.4.4 BKS and JSR Transfer Functions

The BKS and JSR systems also employ EMI coils. The BKS systems use a BF-4 coil for the vertical axis, while the JSR system uses a BF-21 coil. The BF-21 is considerably shorter than either the BF-4 or BF-7 coil. Nevertheless the transfer functions for the various types are well-described by Equation (2.5).

The BKS systems employ a 24 bit Quanterra 4120 Digitizer at SAO and a 24 bit Quanterra 935 Digitizer at PKD1. Both digitizers oversample at 32 KSPS and generate output data at a rate of 40 SPS. The JSR system uses a 24 bit RefTek digitizer which oversamples at 256 KSPS with an output data rate of 40 SPS. The RefTek uses a Crystal Semiconductor CS5321/CS5322 chipset to implement the sigma-delta analog-to-digital conversion and FIR filtering operations. A 6-pole anti-aliasing filter is used.

2.5 Noise Performance

Typical noise spectra of the LKC and MPK systems are shown in Figure 2.5. For both systems we have converted the electrical noise amplitudes to equivalent magnetic field amplitudes. Our estimate of the LKC spectrum is based upon time series data acquired with the inputs of the pre-amplifier shorted. We have also added the theoretical value of the thermal noise due to the coil resistance. The MPK noise spectrum is a typical curve obtained from EMI data sheets. Their method for calculating noise spectra is described in [*Morrison et al.*, 1984].

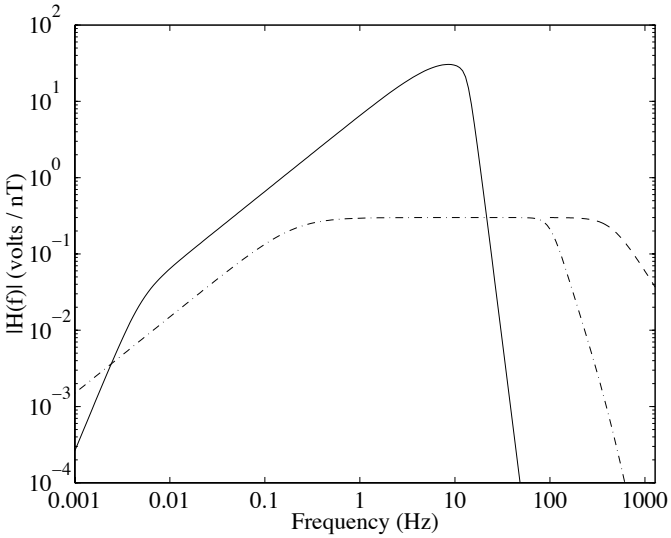


Figure 2.4: Response of the LKC (solid) and the MPK (dash-dot) systems. The dashed line shows the MPK response without the 4th order Butterworth lowpass filter for frequencies above 100 Hz.

The noise spectra reflect the noise of the coils and amplifiers and do not include any effects due to digital processing of the sampled data. In addition the noise spectrum for the MPK system does not include the effects of the anti-aliasing filters. Note that the noise level of the MPK system is 5 to 10 times lower than that of the LKC system. Especially for frequencies between 0.1 and 1 Hz, the noise floor of the LKC system is sometimes greater than the natural geomagnetic background as measured by the MPK system. During periods of enhanced magnetic activity, however, the noise floors of both systems are well below the geomagnetic background.

2.6 Transfer Function Compensation

The measurement systems can be divided into two groups: those with transfer functions similar to that of LKC and those with responses similar to that of MPK. In order to combine the data from the various systems, we apply compensation filters to the sampled time series data so that the compensated transfer functions of the two system types are nearly identical over a specified bandwidth. We have chosen to compensate both system

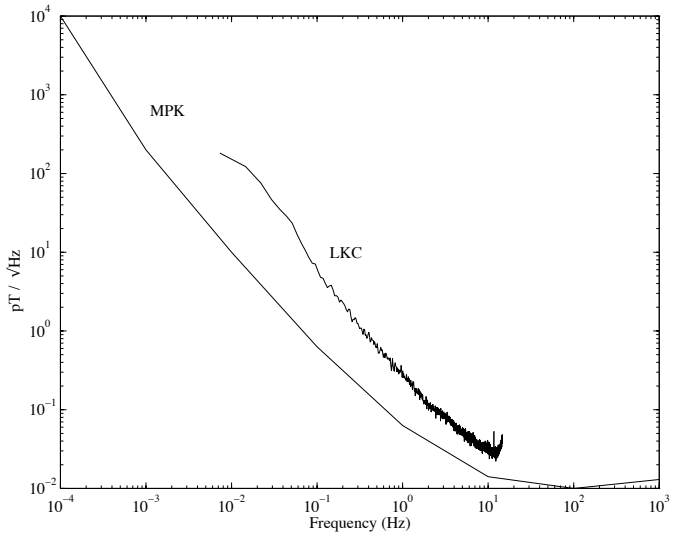


Figure 2.5: Noise characteristics of the LKC and MPK systems.

types so that the overall response yields an output that is proportional to the magnetic field $b(t)$ over the frequency range 0.01 to 5 Hz. In effect, we invert the transfer functions over a limited bandwidth. This process is also referred to as deconvolution or inverse filtering [*Scherbaum*, 1996].

There is an extensive literature that discusses the deconvolution problem [*Harvey and Choy*, 1982; *Hunt*, 1972; *Oldenburg*, 1981; *Treitel and Lines*, 1982]. In its most basic form, the problem may be stated as follows: Given the *a priori* relation $y(t) = h(t) * b(t)$, use the observed data $y(t)$ and knowledge of the transfer function $h(t)$ to form an estimate $\hat{b}(t)$ of $b(t)$. If there exists an inverse filter function $g(t)$ such that $g(t) * h(t) = \delta(t)$, where $\delta(t)$ is the Dirac delta function, then the desired estimate is simply $\hat{b}(t) = g(t) * y(t)$. The Fourier transform of $g(t)$ is $G(f) = \frac{1}{H(f)}$ where $H(f)$ is the Fourier transform of $h(t)$.

When singularities exist in $H(f)$, the deconvolution problem is referred to as an *ill-posed* problem [*Riad*, 1986]. Note, for example, that $G(f)$ is undefined for frequencies where $H(f) = 0$. Furthermore, when the amplitude of $H(f)$ is close to zero, the deconvolution process becomes unstable in the sense that small changes in the measurement $y(t)$ yield large changes in the estimate. To better understand this effect, note that, in practice, measurements always reflect the presence of noise, i.e., $y(t) = h(t) * b(t) + n(t)$ where $n(t)$ represents the noise. Assuming that $b(t)$ and $n(t)$ are stationary processes,

and that $b(t)$ and $n(t)$ are independent, we have the frequency domain relation $S_Y(f) = |H(f)|^2 S_B(f) + S_N(f)$ where $S_Y(f)$, $S_B(f)$, and $S_N(f)$ are the power spectra of $y(t)$, $b(t)$, and $n(t)$, respectively. The power spectrum of the estimate $\hat{b}(t) = g(t) * y(t)$ is $S_{\hat{B}}(f) = S_B(f) + \frac{S_N(f)}{|H(f)|^2}$, where the previous definition of $G(f)$ holds. Note that for frequencies where $|H(f)|$ approaches zero, $S_{\hat{B}}(f)$ is dominated by the noise term.

We now present a brief overview of three commonly used methods for solving the deconvolution problem — *regularization*, *linear estimation*, and *inverse filter design*. Each method takes a slightly different approach to dealing with the ill-posed nature of the problem. After a discussion of the various methods we describe the specific implementation used for the ULF data set. For convenience and ease of understanding we concentrate on continuous-time domain descriptions of the deconvolution methods. Of course, since the time series data are sampled sequences $y[n]$, the actual implementation is in the discrete-time domain. The intuition gained from the continuous-time descriptions carries over well to the discrete-time case, and we defer treatment of the discrete-time solution to the end of our presentation.

2.6.1 Regularization

Regularization methods typically handle singularities in $H(f)$ by placing constraints on the estimate $\hat{b(t)}$ [*Banham and Katsaggelos*, 1997; *Oldenburg*, 1981; *Treitel and Lines*, 1982]. That is, they find the solution $\hat{b}(t)$ that minimizes a cost function $J(\hat{b}(t))$, where a common form for the cost function is

$$J(\hat{b}(t)) = ||y(t) - h(t) * \hat{b}(t)||^2 + \sum_i \lambda_i ||F_i(\hat{b}(t))||^2, \tag{2.8}$$

The $F_i(\cdot)$ terms represent different operators (e.g., $\frac{d}{dt}, \int$), and the λ_i terms are the so-called *regularization parameters* [*Banham and Katsaggelos*, 1997]. The notation $||b(t)||$ denotes the L_2 or energy norm of $b(t)$. The first term in (2.8) represents the error $e(t) = y(t) - h(t) * \hat{b}(t)$ in solving the convolution equation, while the remaining cost terms represent various constraints on the form of the solution. For example, the use of the operator $F(\hat{b(t)}) = \hat{b}(t)$ is equivalent to requiring that the solution $\hat{b}(t)$ have finite energy. The resulting cost function

$$J(\hat{b}(t)) = ||e(t)||^2 + \lambda ||\hat{b}(t)||^2 \tag{2.9}$$

represents a trade-off between the accuracy of the solution and the energy of the solution. To gain more insight into the nature of the cost function, we may use Parseval's theorem to rewrite (2.9) in the frequency domain as

$$J(\hat{B}(f)) = ||Y(f) - H(f)\hat{B}(f)||^2 + \lambda||\hat{B}(f)||^2, \tag{2.10}$$

where, for the sake of argument, we have assumed that the Fourier transform $Y(f)$ of $y(t)$ is well-defined. We can see that the choice $\hat{B}(f) = \frac{Y(f)}{H(f)}$ makes the first term go to zero. But if there are any zeros in $H(f)$ this choice causes the second term $||\hat{B}(f)||^2$ to be unbounded. The form of the cost function thus forces us to choose a solution $\hat{B}(f) \neq \frac{Y(f)}{H(f)}$ in order to minimize the sum of the terms.

In using regularization methods, we make *a priori* assumptions about the regularity or smoothness of the solution. In addition, through our choice of regularization parameters, we choose a specific trade-off between accuracy and regularity. We also note that the operation is data dependent since the minimization procedure in (2.8) depends explicitly on $y(t)$. In other words, if the regularization process is represented by an operator $L(\cdot)$ such that $\hat{b}(t) = L(y(t))$, then the form of $L(\cdot)$ depends on $y(t)$.

Another regularization strategy focuses on forming a regularized inverse filter. Instead of solving for $\hat{b}(t)$, we solve for the inverse filter $g(t)$ that minimizes the cost function

$$J(g(t)) = ||\phi(t) - g(t) * h(t)||^2 + \sum_i \lambda_i ||F_i(g(t))||^2 \tag{2.11}$$

where $\phi(t)$ is the 'data resolution kernel' [*Menke*, 1989]. This form of the regularization problem is closely related to both the 'spiking' filter in seismic deconvolution [*Treitel and Lines*, 1982] and the Backus-Gilbert formalism in geophysics [*Oldenburg*, 1981; *Treitel and Lines*, 1982; *Twomey*, 1997]. Here the choice of $\phi(t)$ and the form of the additional cost terms reflect our *a priori* assumptions about the nature of the desired solution. This deconvolution process is *data independent* since the estimate is $\hat{b}(t) = g(t) * y(t)$, where $g(t)$ does not depend on $y(t)$.

2.6.2 Linear Estimation

In the framework of estimation theory, we model $b(t)$ as a random process and explicitly take into account the additive noise $n(t)$, which is assumed to be stationary and independent of $b(t)$. As a result, $y(t) = h(t) * b(t) + n(t)$ is also a random process. We seek an operator $G(\cdot)$ that minimizes the expected value of an error functional $f\left(G(y(t)) - w(t) * b(t)\right)$ where $w(t)$ is a smoothing filter. It is common to use a norm-based functional

$$f(e(t)) = \left(||e(t)||_p\right)^{\gamma}. \tag{2.12}$$

where $e(t) = G(y(t)) - w(t) * b(t)$ is the error and $||e(t)||_p$ is its L_p norm. For $p = \gamma = 2$, the functional corresponds to the square of the L_2 norm of the error, and the estimate is referred to as the *least squares* (LS) estimate. If $b(t)$ and $y(t)$ are jointly Gaussian random processes, then the operator $G(\cdot)$ which minimizes the LS error is a linear operator of the form $G(y(t)) = g(t) * y(t)$. The resulting estimate is referred to as the *linear least squares estimate* (LLSE) [*Kailath*, 1981]. It is most appropriate when applied to processes that are relatively well modeled by Gaussian distributions [*Menke*, 1989], but, because of its ease of implementation and the existence of closed-form solutions, it is widely used even when the Gaussian condition is not met.

If we assume that $b(t)$ is stationary, then we may obtain a closed-form solution for the inverse filter $g(t)$. In the frequency domain the non-causal solution is

$$G(f) = \frac{W(f)H^*(f)S_B(f)}{|H(f)|^2 S_B(f) + S_N(f)}. \tag{2.13}$$

where $S_B(f)$ and $S_N(f)$ are the power spectra of $b(t)$ and $n(t)$, respectively, $G(f)$, $H(f)$ and $W(f)$ are the Fourier transforms of $g(t)$, $h(t)$ and $w(t)$, respectively, and $H^*(f)$ is the complex conjugate of $H(f)$. This solution is also referred to as the non-causal Wiener filtering solution [*Kailath*, 1981]. We see that the inclusion of the noise term $S_N(f)$ stabilizes the inverse filter when $|H(f)|$ is small. The inverse filter $G(f)$ depends on the second order statistics of both the signal $b(t)$ and the noise $n(t)$. Since $n(t)$ represents primarily the electronic noise of the system, we can safely assume that it is a stationary process. The assumption of stationarity is less applicable to $b(t)$, since geomagnetic activity varies significantly over the course of a day [*Jacobs*, 1970]. As a result, the optimum inverse filter $G(f)$ will vary with changes in $S_B(f)$.

2.6.3 Inverse Filter Design

The inverse filter design approach makes use of the intuitive notion that we should not try to reconstruct $b(t)$ for frequencies where $H(f)$ approaches zero. This approach has been employed by various researchers [*Choy and Cormier*, 1986; *Harvey and Choy*, 1982; *Kitzinger et al.*, 1988] for the deconvolution of seismic waveform data. Instead of forming an estimate of $b(t)$, we attempt to estimate a bandlimited version $b_w(t) = b(t) * w(t)$ where $w(t)$ is a bandpass filter. The estimate is

$$\hat{b}_w(t) = g(t) * w(t) * y(t) = b_w(t) + g(t) * w(t) * n(t). \tag{2.14}$$

Define the bandlimited inverse filter $g_w(t) = g(t) * w(t)$ with corresponding Fourier transform $G_w(f) = G(f)W(f) = \frac{W(f)}{H(f)}$, where $W(f)$ is the Fourier transform of $w(t)$. The transform $G_w(f)$ is well-defined for all frequencies if the zeros of $W(f)$ are chosen to cancel the zeros of $H(f)$. Aside from this constraint on the placement of zeros, the design of $W(f)$ depends primarily on the desired frequency characteristics of the estimate.

Note that $g_w(t)$ is equal to the linear estimation filter defined in (2.13) when there is no noise, i.e., $S_N(f) = 0$. In addition, $g_w(t)$ is equal to the regularized inverse filter associated with (2.11) when $\phi(t) = w(t)$ and there are no additional cost terms, i.e., $\lambda_i = 0$.

The results achieved with the filter design technique have been shown to be fairly accurate when applied to seismic data sets [*Choy and Cormier*, 1986]. Still, the use of a filter $w(t)$ has been criticized by *Sipkin and Lerner-Lam* [1992], who claim that it introduces distortion into pulse shapes and argue that regularization approaches perform better. It is important to realize, however, that the parameters of the regularization approach need to be chosen correctly in order to obtain the improved performance. While the correct parameters can easily be chosen when deconvolving simulated data as was done in [*Sipkin and Lerner-Lam*, 1992], the choice of parameters for a real data set necessarily reflects the assumptions of the practitioner under conditions of uncertainty and may not provide better performance.

2.6.4 Choosing an Approach

After reviewing the approaches described above, we selected the inverse filter design approach for deconvolution of the ULF data. The regularization approach of (2.8) is most appropriate when there are strong *a priori* beliefs about the smoothness of the solution.

In the case of the SFBay data set, however, it is difficult to justify any *a priori* constraints on either the smoothness or roughness of the solution. Examination of the time series data reveals both periods when the waveforms are quite smooth (e.g., narrowband micropulsations) and periods when the waveforms are quite rough (e.g., strong man-made and natural transient phenomena). In addition, there are many periods in which the two types of waveforms are superposed.

The linear estimation approach makes the assumption that the second-order statistics of the underlying process $b(t)$ are known and stationary. In practice, the statistics are not known and we must estimate $S_B(f)$ from the data. Because $S_B(f)$ varies greatly over the course of a day, and also from system to system, the procedure is computationally intensive. Furthermore, our ability to obtain a good estimate of $S_B(f)$ decreases as the signal power $|H(f)|^2 S_B(f)$ decreases with respect to the noise power $S_N(f)$. Examination of (2.13) reveals that the effect of the term $S_N(f)$ on the form of $G(f)$ is greatest when the $|H(f)|^2|S_B(f)|$ is equal to or less than $|S_N(f)|$ — but this is precisely the operating region where our estimate of $S_B(f)$ is poor.

For the MPK system the low-noise design of the sensors typically results in the condition $|H(f)|^2|S_B(f)| \gg |S_N(f)|$. With this assumption we find that $G(f) \approx \frac{W(f)}{H(f)}$ so that the linear estimation solution approaches the inverse filter design approach solution. For the LKC system, because of the high noise floor of the sensors we find that during periods of low levels of geomagnetic activity $|S_N(f)| > |H(f)|^2|S_B(f)|$, so that the linear estimation solution can differ significantly from the inverse filter design solution. In the presence of BART interference, however, we almost always find that $|H(f)|^2|S_B(f)| \gg |S_N(f)|$ for the LKC system. Since our work primarily addresses the detection and removal of the BART interference, we may assume that $|H(f)|^2|S_B(f)| \gg |S_N(f)|$ for all systems in the SFBay array.

We conclude that the linear estimation and the inverse filter design approaches are nearly equivalent for all systems in the SFBay array and for all systems outside of the SFBay that employ low-noise EMI sensors. We selected the inverse filter design approach since it does not require estimation of $S_B(f)$ and thus requires less computation.

2.6.5 Specific Implementation

In implementing the inverse filter design approach, we do not attempt to find an inverse filter for the complete transfer function, but instead concentrate on finding an inverse

filter for that part of the transfer function which has the greatest impact on the frequency response in the range 0.01 to 5 Hz. As an example, for the LKC system, a straightforward application of the inverse filter design approach would use a filter $G(s) = \frac{W(s)}{H_c(s)}$, which yields an estimate of the bandlimited magnetic field $B_w(f) = W(f)B(f)$. We modify the approach to define an inverse filter $G(s) = \frac{W(s)}{H_{ci}(s)}$ which only inverts the $H_{ci}(s)$ term of $H_c(s)$. This filter is used to form an estimate of $B_w(f) = W'(f)B(f)$ where $W'(f) = W(f)H_{hi}(f)H_{lo}(f)H_8(f)$. In essence, we do not invert the bandlimiting terms $H_{hi}(s)$, $H_{lo}(s)$, $H_8(s)$, but instead use them as part of our effective bandpass filter $W'(f)$. The advantages of this approach are increased simplicity and numerical stability of the inverse filter $G(s)$. For the MPK system we invert only the $H_{mi}(s)$ term since the effects of the $H_{lo}(s)$ and $H_{aa}(s)$ terms and the digital interpolation filter $h_2[n]$ (see Equation 2.2) are negligible in the range of interest.

We now describe in detail the construction of the inverse filter for the LKC system. The non-bandlimited inverse filter is $G_{ci}(s) = \frac{1}{H_{ci}(s)} = \frac{s+\beta}{s}$ which has a singularity at $s = 0$. Since we carry out the deconvolution in the discrete-time domain, we need to construct a discrete-time filter $\tilde{G}_{ci}(z)$ which has the approximately the same frequency response as $G_{ci}(s)$. We use the bilinear transform $s = 2F_s\frac{1-z^{-1}}{1+z^{-1}}$ [*Oppenheim and Schafer*, 1989] to construct

$$\begin{aligned} \tilde{G}_{ci}(z) &= \left.\frac{s+\beta}{s}\right|_{s=2F_s\frac{1-z^{-1}}{1+z^{-1}}} \\ &= \frac{(\beta+2F_s)+(\beta-2F_s)z^{-1}}{2F_s(1-z^{-1})} \end{aligned} \tag{2.15}$$

where F_s is the sampling frequency. The bilinear transform introduces a frequency warping such that a frequency f in the continuous-time domain is warped into a frequency

$$\tilde{f} = \frac{F_s}{\pi}\arctan\left(\frac{\pi f}{F_s}\right) \tag{2.16}$$

in the discrete-time domain [*Oppenheim and Schafer*, 1989]. To justify the use of the bilinear transform, we verify that the warping is negligible for the frequencies of interest. For $F_s = 30$ samples per second, continuous-time frequency values of 0.01, 0.1, 1.0 and 5.0 Hz are warped into discrete-time frequency values of 0.01, 0.10, 0.996 and 4.61 Hz. The warping is clearly negligible over the range 0.01 to 1.0 Hz. Over the range 1.0 to 5.0

the degree of warping has increased but is still small with a maximum relative error of 8 percent at 5.0 Hz.

Our next step is to construct a filter $\tilde{W}(z)$ to cancel the pole at $z = 1$ in $\tilde{G}_{ci}(z)$. We first multiply $\tilde{G}_{ci}(z)$ by a one-pole high-pass filter

$$\tilde{W}_1(z) = \frac{b_0(1 - z^{-1})}{1 - a_1 z^{-1}} \tag{2.17}$$

with $f_c = 0.005$ Hz. The combined response

$$\begin{aligned} \tilde{G}_1(z) &= \tilde{G}_{ci}(z)\tilde{W}_1(z) & (2.18) \\ &= \frac{b_0}{2F_s}\frac{(\beta + 2F_s) + (\beta - 2F_s)z^{-1}}{1 - a_1 z^{-1}} & (2.19) \end{aligned}$$

no longer has a singularity on the unit circle in the z-domain. In practice, because $\tilde{G}_1(z)$ lacks a zero at $z = 1$, we find that the inverse filtering process exhibits problems with baseline drift. To reduce this effect, we place a zero at $z = 1$ by multiplying $\tilde{G}_1(z)$ by an additional first order high-pass filter $\tilde{W}_2(z)$ with $f_c = 0.005$ Hz.

We also multiply by an non-causal low-pass FIR filter $\tilde{W}_3(z)$ with $f_c = 10$ Hz and a duration of 4 seconds. The primary purpose of $\tilde{W}_3(z)$ is to eliminate a 13.5 Hz calibration tone that is used to monitor the operation of the LKC system. A FIR design was chosen in order to obtain a steep roll-off while maintaining linear phase. Since $\tilde{W}_3(z)$ is non-causal, the output of the filter depends on inputs extending up to 2 seconds into the future. One of the effects of using a non-causal filter is that a spike in the input data will appear in the output as a spike preceded by ripples beginning 2 seconds prior to the spike. As we are interested in possible earthquake precursors, this phenomenon is a potential cause for concern. In fact, *Scherbaum and Bouin* [1997] addressed the problem of misinterpreting FIR filter artifacts as seismic nucleation phases. With regard to possible ULF magnetic field precursors, previous observations [*Fraser-Smith et al.*, 1990] indicate that if a precursor exists it is likely to persist for hours or days before the main shock. As a result, we assume that the 2 second precursor artifact due to $\tilde{W}_3(z)$ can be neglected in our application. In effect we accept the presence of a minimal precursor artifact in order to obtain a linear phase response. A causal solution, i.e., one in which the filter output depends only on past and present inputs, would necessarily introduce some phase distortion. Finally, we note that both $\tilde{W}_1(z)$ and $\tilde{W}_2(z)$ are causal by design.

Putting all the pieces together, we have constructed an inverse filter

$$\tilde{G}(z) = \tilde{W}(z)\tilde{G}_{ci}(z)$$

where $\tilde{W}(z) = \tilde{W}_1(z)\tilde{W}_2(z)\tilde{W}_3(z)$. Using the substitution $z = e^{j2\pi f/F_s}$ on the unit circle in the z-domain [*Oppenheim and Schafer*, 1989], we may write the overall response of the compensated system as

$$\begin{aligned} W_c'(f) &= H_c(f)\tilde{G}\left(e^{j2\pi f/F_s}\right) \\ &= H_{ci}(f)\tilde{G}_{ci}\left(e^{j2\pi f/F_s}\right)\tilde{W}\left(e^{j2\pi f/F_s}\right)H_{lo}(f)H_{hi}(f)H_8(f) \\ &\approx \tilde{W}\left(e^{j2\pi f/F_s}\right)H_{lo}(f)H_{hi}(f)H_8(f) \end{aligned}$$

where we have used the fact that $G_{ci}\left(e^{j2\pi f/F_s}\right)$ is the approximate inverse of $H_{ci}(s)$ in the frequency range of interest. The magnitude response, phase response, and group delay of $W_c'(f)$ are shown in Figure 2.6. Note that the magnitude response is fairly flat over the range 0.01 to 5 Hz. As shown by the group delay plots, there is noticeable phase distortion or dispersion at the lower frequencies, e.g., about 20 percent distortion at 0.01 Hz. Among other effects, dispersion can lead to the distortion of the shapes of transient signals. Because the detection approach that we will develop in this dissertation is not highly dependent on the exact shape of the transients, this degree of distortion is acceptable. We are also aided by the fact that the characteristic time scales of the transients of interest are less than 10 seconds — there is negligible phase distortion at 0.1 Hz. While it is possible to modify the design to decrease the phase distortion, the resulting inverse filter exhibits decreased numerical stability.

For the MPK system we construct an inverse filter of the form

$$\tilde{G}(z) = \tilde{G}_{mi}(z)\tilde{W}(z)$$

where $\tilde{G}_{mi}(z)$ is given by the bilinear transform of $\frac{1}{H_{mi}(s)}$ and

$$\tilde{W}(z) = (\tilde{W}_1(z))^2\tilde{W}_2(z),$$

where $\tilde{W}_1(z)$ is a first order high-pass filter with $f_c = 0.005$ Hz and $\tilde{W}_2(z)$ is a second order Butterworth high-pass filter with $f_c = 0.005$ Hz. The form of $\tilde{W}(z)$ was chosen so

that the response $W'_m(f) = H_m(f)\tilde{G}\left(e^{j2\pi f/F_s}\right)$ of the compensated MPK system matches that derived above for the LKC system. The magnitude response, phase response, and group delay of $W'_m(f)$ are shown in Figure 2.6.

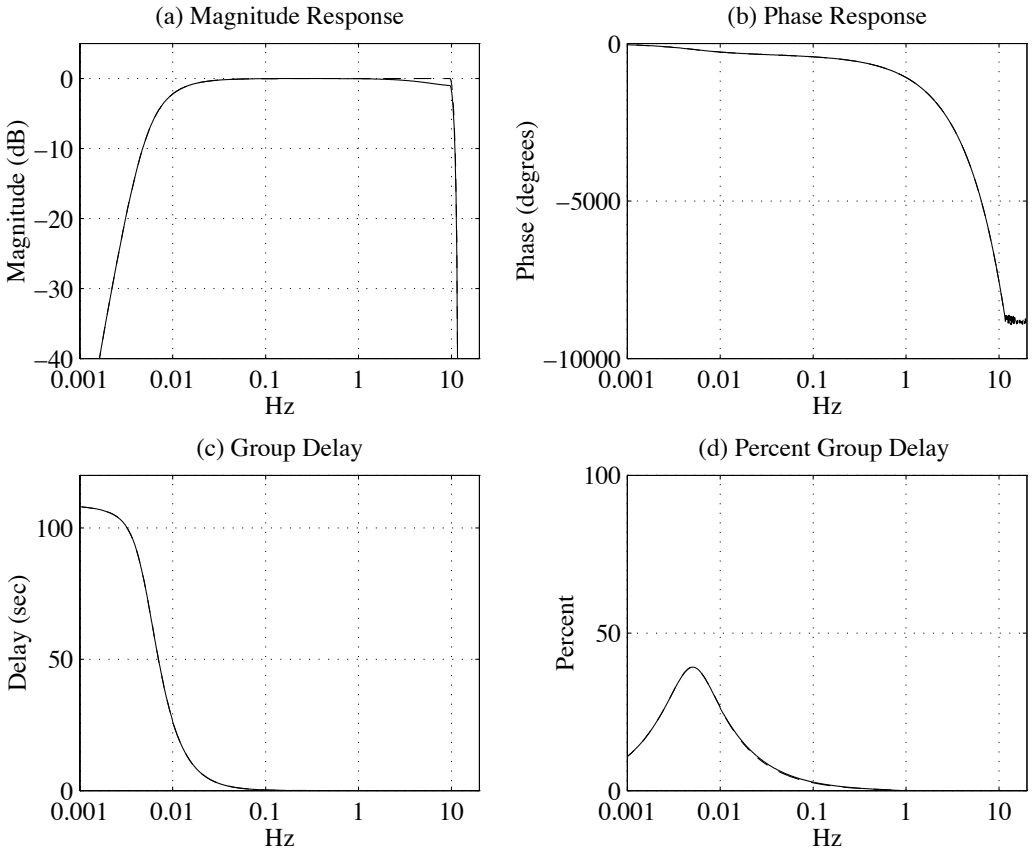


Figure 2.6: Transfer functions after system response compensation for LKC (solid) and MPK (dashed) systems. Panel (a) Magnitude response; (b) Phase Response; (c) Group Delay; (d) Percent Group Delay, i.e., group delay as a percentage of period. For the most part, the curves for the two systems coincide. In panel (a), the dashed line is slightly above the solid line for frequencies between 5 and 10 Hz.

Chapter 3

ULF Magnetic Field Measurements

3.1 Introduction

In this chapter we describe the ULF fields measured by the systems in the SFBay array. We focus on the fields generated by the BART system. As shown by *Fraser-Smith and Coates* [1978], the BART system is the dominant source of man-made ULF magnetic fields in the SFBay. Later work [*Bernardi et al.*, 1989; *Egbert*, 1997; *Ho et al.*, 1979] provided additional observations of the fields due to BART. Still, a detailed description of the temporal and spatial features of the BART fields has been lacking. With the SFBay array we are able to simultaneously observe the BART fields at multiple sites, each equipped with a three-axis magnetometer. As a result, we can present a fairly exhaustive and unprecedented characterization of the fields.

After a brief recap of the previous work in Section 3.2, we examine in Section 3.3 the use of the fields at a remote reference site, SAO, to estimate the natural background at the SFBay sites. By subtracting this estimate from the observed data, we can obtain a better description of the BART fields. In Section 3.4 we present a detailed treatment of the interference in the time domain. In addition to providing plots of example time series data, we discuss the characteristic time scales and amplitudes of the interference. We also examine the spatial features of the fields by constructing polarization vectors for the interference at each site. In Section 3.5 we present both power spectra and half hourly spectral indices to demonstrate the frequency domain behavior of the interference.

Furthermore, in order to provide some perspective on the field amplitudes, we compare the spectra of the BART interference to the precursor levels observed at Corralitos prior to the 1989 Loma Prieta earthquake.

While our primary focus is on the characterization of the BART fields, we also consider in Section 3.6 the approximate $1/f$ nature of the natural geomagnetic background. Using both power spectra and recently developed wavelet based analysis methods, we examine the extent to which the geomagnetic background can be considered an approximate $1/f$ process. The results are relevant to the signal processing methods developed in later chapters.

Times are cited in both local time (LT) and universal time (UT). UT is used for geomagnetic data, since some of the data (SAO and PKD1) are obtained from public databases which follow a UT convention. LT is a more appropriate convention with respect to BART operating hours. LT is also used in instances where geomagnetic data and BART operational data are superimposed, e.g., plots of magnetic activity and power consumption. The majority of the geomagnetic data shown are from August 1998. Unless otherwise stated, subtract 7 hours from the UT value to convert to LT. For example, on August 11, 1998, 10:46 UT corresponds to 3:46 LT.

When discussing the interference due to BART, we typically show data from all channels in the array, since the observed interference varies greatly from channel to channel. By contrast, the natural geomagnetic fields do not vary greatly from system to system. The fields are strongest in the horizontal plane, and the features of the N-S and E-W fields tend to be similar. In our treatment of the background fields, we have typically chosen, therefore, to present data from a single channel that best illustrates the point at hand.

3.2 Summary of Previous Observations

The measurements reported by *Fraser-Smith and Coates* [1978] were performed using a single coil magnetometer aligned with the North-South geomagnetic axis and located at Stanford, CA, a site about 20 km from the nearest BART tracks. BART-related spectral amplitudes were shown to be about 10 times greater than the the average natural background level, with a typical spectral amplitude of 500 pT/$\sqrt{\text{Hz}}$ at 0.1 Hz. Later, *Ho et al.* [1979] used a three-axis fluxgate magnetometer to measure BART fields at Laney College in Oakland, a site only 100 meters from the closest tracks. Spectral amplitudes

for the vertical and horizontal components were approximately 100 times and 10 times, respectively, greater than corresponding amplitudes observed during hours when BART was not in operation. Peak increases were observed at frequencies near 0.1 Hz with values of approximately 2.5×10^3 pT/$\sqrt{\text{Hz}}$, 1400 pT/$\sqrt{\text{Hz}}$ and 600 pT/$\sqrt{\text{Hz}}$ for the vertical, North-South (N-S), and East-West (E-W) fields, respectively. Time series data showed that the vertical component peak amplitudes associated with BART were as high as 2000 to 3000 nT.

The time of occurrence of the interference was found to be well correlated with the hours of BART operation — nominally 04:00 to 01:30 LT on weekdays, 06:00 to 01:30 LT on Saturdays, and 08:00 to 01:30 LT on Sundays [*Fraser-Smith and Coates*, 1978]. In fact, half-hourly spectral averages of the data showed a striking resemblance to the number of BART cars in use over the course of a day [*Bernardi et al.*, 1989].

Using a three-site magnetotelluric array located approximately 50 km to the south of the SFBay, *Egbert* [1997] showed that the predominant source of coherent noise in his array data was the BART system. The array consisted of two sites that exhibited BART interference and one southernmost 'quiet' site that was shown to be free of observable BART signals. Egbert found that the coherent noise was significant over the frequency range 0.02 to 0.25 Hz with peak interference levels at around 0.1 Hz. Confirming the observations of *Fraser-Smith and Coates* [1978], Egbert reported the absence of coherent noise during the period 01:30 to 04:00 LT, when BART is not in operation. From examination of the electric field components of the coherent noise, Egbert concluded that the observations were consistent with a model of the BART noise source as a series of grounded electric dipoles varying with time.

3.3 The Use of SAO as a Remote Reference

Throughout this thesis we will use the data from SAO as an estimate of the natural background signals. The presence of such an estimate helps us to determine if an observed transient is man-made or natural. Natural ULF geomagnetic fields are highly coherent over distances of several hundred kilometers or more because the external sources are "well approximated by uniform plane waves of infinite horizontal extent [*Egbert and Booker*, 1989]." The *remote reference* method in magnetotellurics takes advantage of this fact and uses the horizontal magnetic fields at a remote site to estimate the magnetic fields at a

local site [*Gamble et al.*, 1979]. In our case, the SAO site is located about 100 km to the south of the SFBay array and can be expected to provide good estimates of the horizontal fields measured at the SFBay sites. Minor differences do appear in the fields between the sites, however. These are due to deviations from the uniform plane wave model and lateral variations in local conductivity that distort the magnetic field differently at each site [*Egbert and Booker*, 1989]. The higher noise levels of the LKC coils as compared to those in use at SAO further contribute to differences in the observed fields.

In Figure 3.1 we demonstrate the use of the SAO data to predict the E-W magnetic field at JSR. Denoting the E-W time series data at JSR as $b_0[n]$ and the N-S and E-W channels at SAO as $b_1[n]$ and $b_2[n]$, respectively, we assume a relation of the form

$$b_0[n] = h_1[n] * b_1[n] + h_2[n] * b_2[n] + e[n] \tag{3.1}$$

where $e[n]$ is the estimation error and $h_1[n]$ and $h_2[n]$ are causal finite impulse response filters, which are referred to as inter-station transfer functions [*Egbert and Booker*, 1989]. Similar relations are used for the other SFBay channels. In geomagnetic work (e.g., geomagnetic sounding and magnetotelluric studies), transfer functions are typically computed in the frequency domain because the quantities of interest (e.g., resistivity) are best interpreted as functions of frequency. Time-domain methods have been proposed [*McMechan and Barrodale*, 1985; *Spagnolini*, 1994], but are not widely used. Since our study focuses on the time-domain behavior of the BART interference, we adopt a time-domain method and compute the filters $h_1[n]$ and $h_2[n]$ that minimize the mean of the squared error $e^2[n]$. Both filters have 21 coefficients.

In panel (a) of Figure 3.1 we show the result of the estimation process for data from an eight minute period beginning at 10:46 UT on August 11, 1998. The SAO data predict the JSR E-W data very well — the standard deviation of the estimation error is 11.3 pT. The optimum filters vary slightly over time, but in general, the filters computed for one time period can be applied to another time period with minimal degradation in performance. In panel (b) we show the result of applying the filters optimized for the period 9:30 to 9:40 UT on August 11, 1998 to the data for the 8 minute period beginning at 10:46 UT. The standard deviation of the estimation error has increased to 15 pT, but the estimation is still quite good.

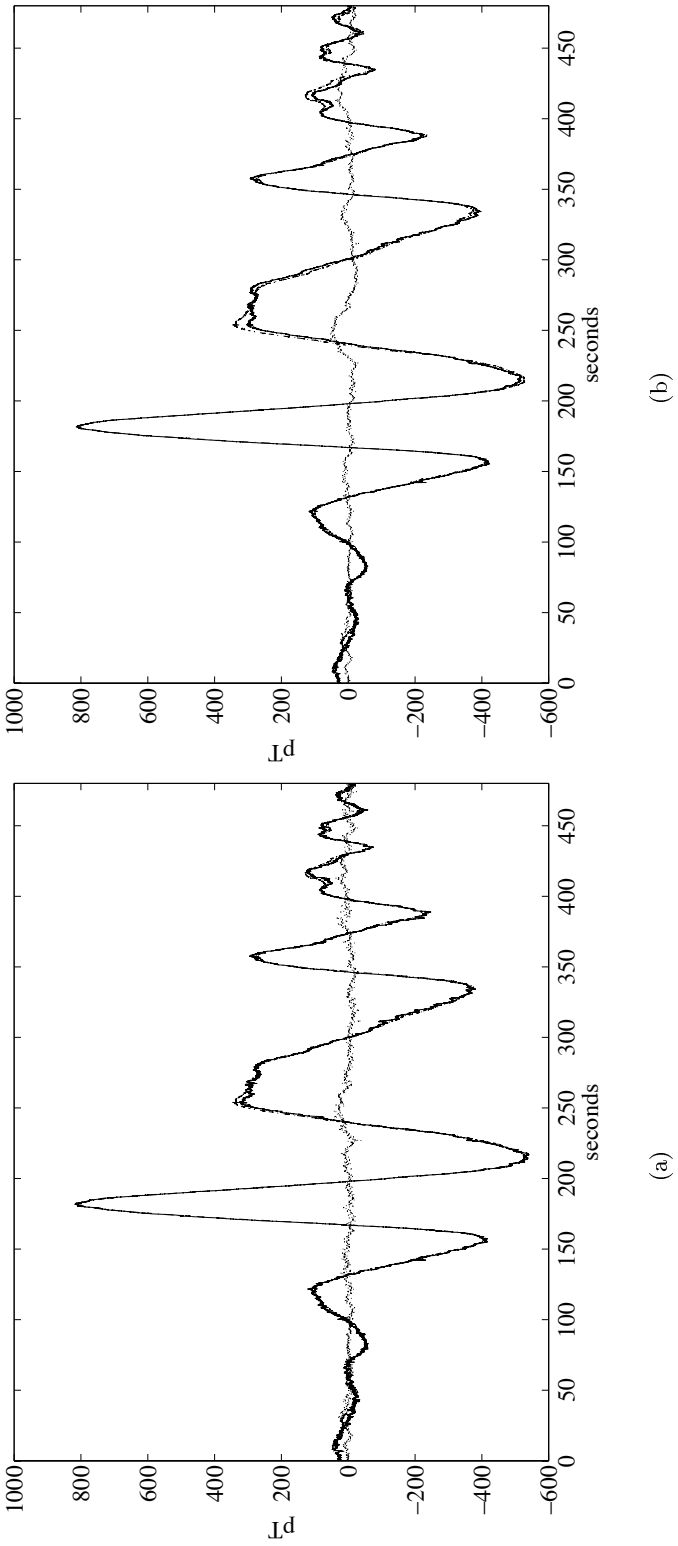


Figure 3.1: Example use of SAO data to estimate the fields at JSR. Panel (a): Estimation of JSR E-W channel for the period 10:46 to 10:54 UT. Panel (b): Estimation of JSR E-W channel for the period 10:46 to 10:54 UT using filters computed for the period 9:30 to 9:40 UT. For both panels, the dash-dot line represents the data at JSR, the superimposed solid line is the estimate, and the dotted line is the estimation error. For the most part, the solid and dash-dot lines coincide. A slight separation can be observed near the 250 second mark, where the dash-dot line is above the solid line.

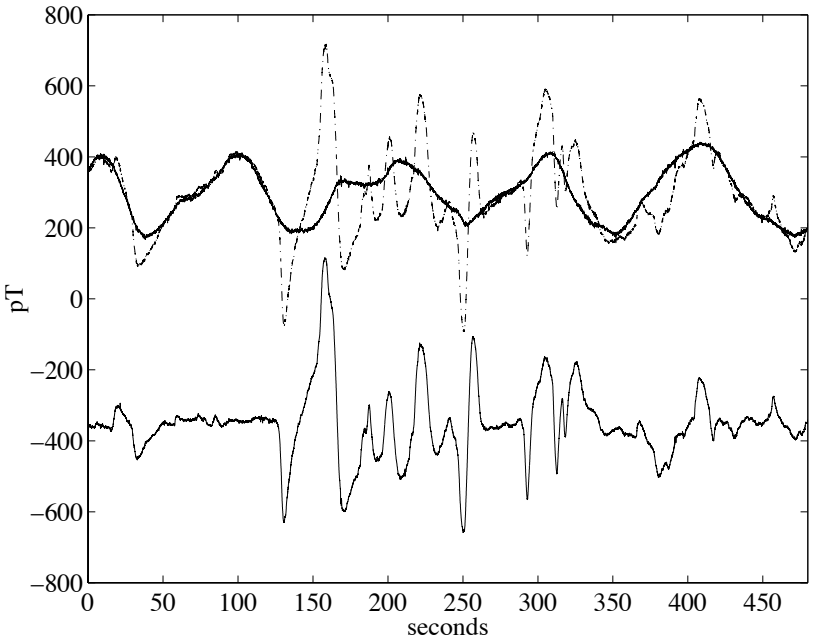


Figure 3.2: Estimation of JSR N-S channel for the period 11:05 to 11:13, when BART is in operation. Field at JSR (dash-dot), estimated field (upper solid), and estimation error (lower solid) are shown. Note that the estimation error shows the BART transients after the contribution of the natural fields has been removed.

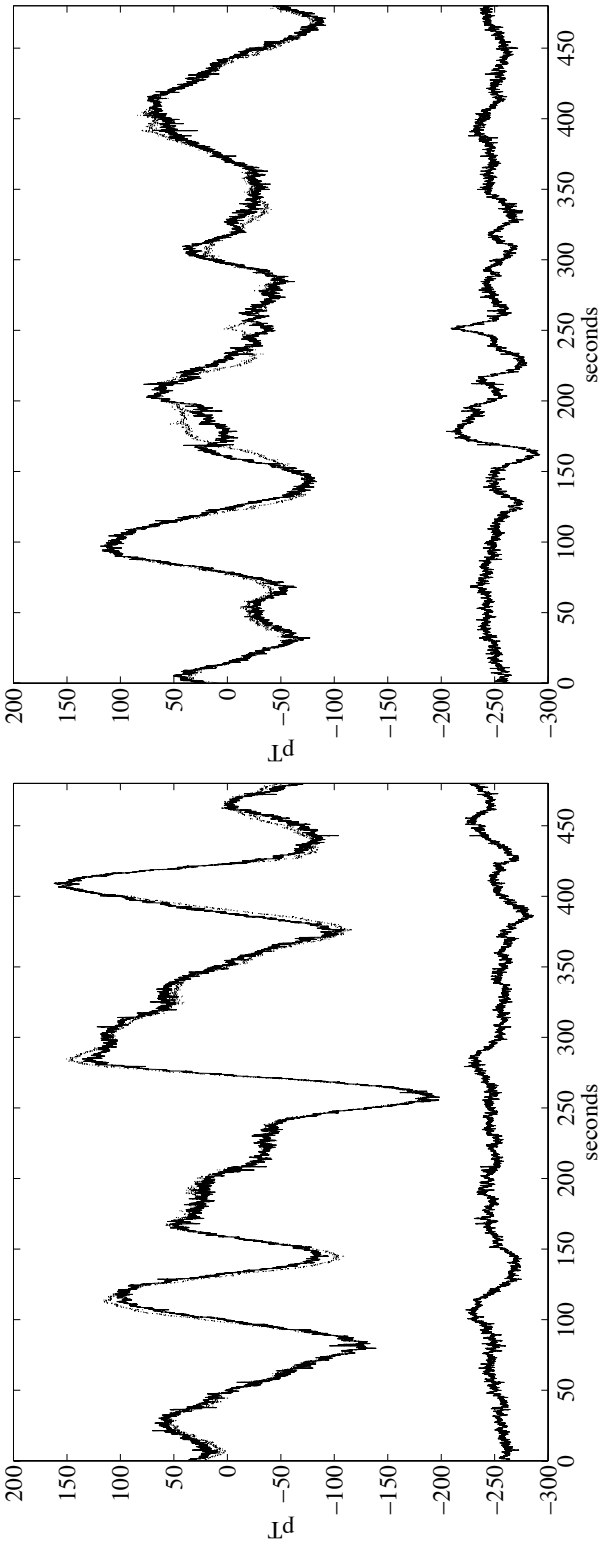


Figure 3.3: Use of SAO data to predict E-W channel at PKD1. Panel (a): estimation for the period 9:32 to 9:40 UT. Panel (b): estimation for the period 11:06 to 11:14 UT when there is significant BART activity. Dotted line is PKD1 data, upper solid line is the estimate and lower solid line is the estimation error. For the most part, the dotted and upper solid lines coincide. To separate the two lines, note that the dotted line is above the upper solid line at the 110 and 275 second marks in panel (a) and at the 175 and 250 second marks in panel (b).

In Figure 3.2 we apply the filters computed for the period 10:45 to 10:55 UT to estimate the natural background on the N-S channel at JSR for the 8 minute period beginning at 11:05 UT on August 11, 1998. Because the contribution of the natural background has been removed, the residual signal shows clearly the behavior of the BART induced transients at JSR.

Up to this point we assumed implicitly that the SAO data are unaffected by BART activity. While the levels of interference are much lower at SAO, as compared to the SFBay sites, we can still observe the effects of BART, especially during periods of decreased geomagnetic activity. By examining the spectral indices computed for SAO (see Section 3.5.3), we found that the effect is most noticeable in the 0.05 to 0.10 Hz spectral band. As an aside, we note that *Egbert* [1997] observed BART interference at a site located 50 km to the southeast of SFBay, but did not detect interference at a 'quiet' site located 130 km to the southeast of the SFBay. The SAO site lies between the two sites — it is about 50 km to the southeast of the 'noisy' site and 30 km to the northwest of the 'quiet' site.

In order to determine whether the SAO time series data provide an adequate estimate of the natural background, we used the SAO horizontal fields to predict the PKD1 E-W field. PKD1 is located approximately 200 km to the south of the SFBay and does not exhibit signs of BART interference. In Figure 3.3 panel (a), we show the result of the estimation procedure for the time period 9:32 to 9:40 UT, when BART is not in operation. The standard deviation of the estimation error is 10.6 pT. In panel (b) we show the result for the time period 11:06 to 11:14 UT when there is significant BART activity (see Figures 3.1c and 3.5). Here the standard deviation is 12.2 pT, and there is a possibility that the excursions in the estimation error near the 175 and 260 second marks are BART-related. The standard deviation of the estimation error, however, is well within the bounds of the typical values, 10 to 20 pT, observed for estimates performed when there is no BART activity. In addition, the standard deviation is considerably smaller than the median values of the BART transient amplitudes (see Section 3.4.2). As a result, we conclude that the SAO data are suitable for use as a remote reference.

Finally, it would be preferable to use data from PKD1, the only remaining 3-axis system with low-noise sensors, as a remote reference. Unfortunately, due to instrument difficulties, we were unable to obtain data for a time period during which all SFBay and PKD1 channels were operational.

3.4 Time Series Examples

In Figures 3.4, 3.5, and 3.6 we present examples of time series data from the SFBay array and from SAO. The first example (Figure 3.4) shows an isolated set of BART transients occurring around 10:40 UT, just before the BART system begins its regularly scheduled activity. Each row of the figure shows one channel of time series data. For example, the rows labeled LKC0, LKC1, LKC2 are data from the vertical, N-S and E-W channels at LKC, respectively. The scale unit for the ordinate axis is pico-Tesla (pT) and the ordinate axis scale of the LKC channels is doubled to accommodate the larger amplitude of the transients seen at LKC. In Figure 3.5 we see that by 11:07 UT both the number and amplitudes of the BART transients have increased significantly. Figure 3.6 shows the activity during the morning rush hour period. Using the SAO data as a reference, we observe that the time series are completely dominated by the BART interference. In both Figures 3.5 and 3.6 the ordinate axis scale of the LKC data channels is four times as large as that of the other systems to accommodate the larger signal levels.

3.4.1 Characteristic Time Scales

As shown in Figures 3.1(c) and 3.4 through 3.6, the BART interference can be considered to be the superposition of a series of transients. In Chapter 4 we show that these transients are associated primarily with changes in the propulsion current during train accelerations and decelerations. Here we examine the characteristic time scales of the observed transients.

In Figure 3.7 we provide an example of the BART interference on the vertical channel at MPK during a period when the natural background signals are negligible. We have extracted and plotted the most significant transients below the main time series. For each of these transients we compute the square root of the variance of the squared modulus, as defined in [*Bracewell*, 1986], as a measure of the temporal width σ_t of a transient. In this specific example, the temporal widths range from 0.9 to 3.6 seconds. In general, we have found that the temporal widths tend to lie in the range 0.8 to 6 seconds. The time between the start and stop of a transient is approximately $5\sigma_t$, so that the duration of the transients range from 4 to 30 seconds. As we will show in Chapter 4, the observed temporal widths are consistent with the temporal widths of propulsion current transients.

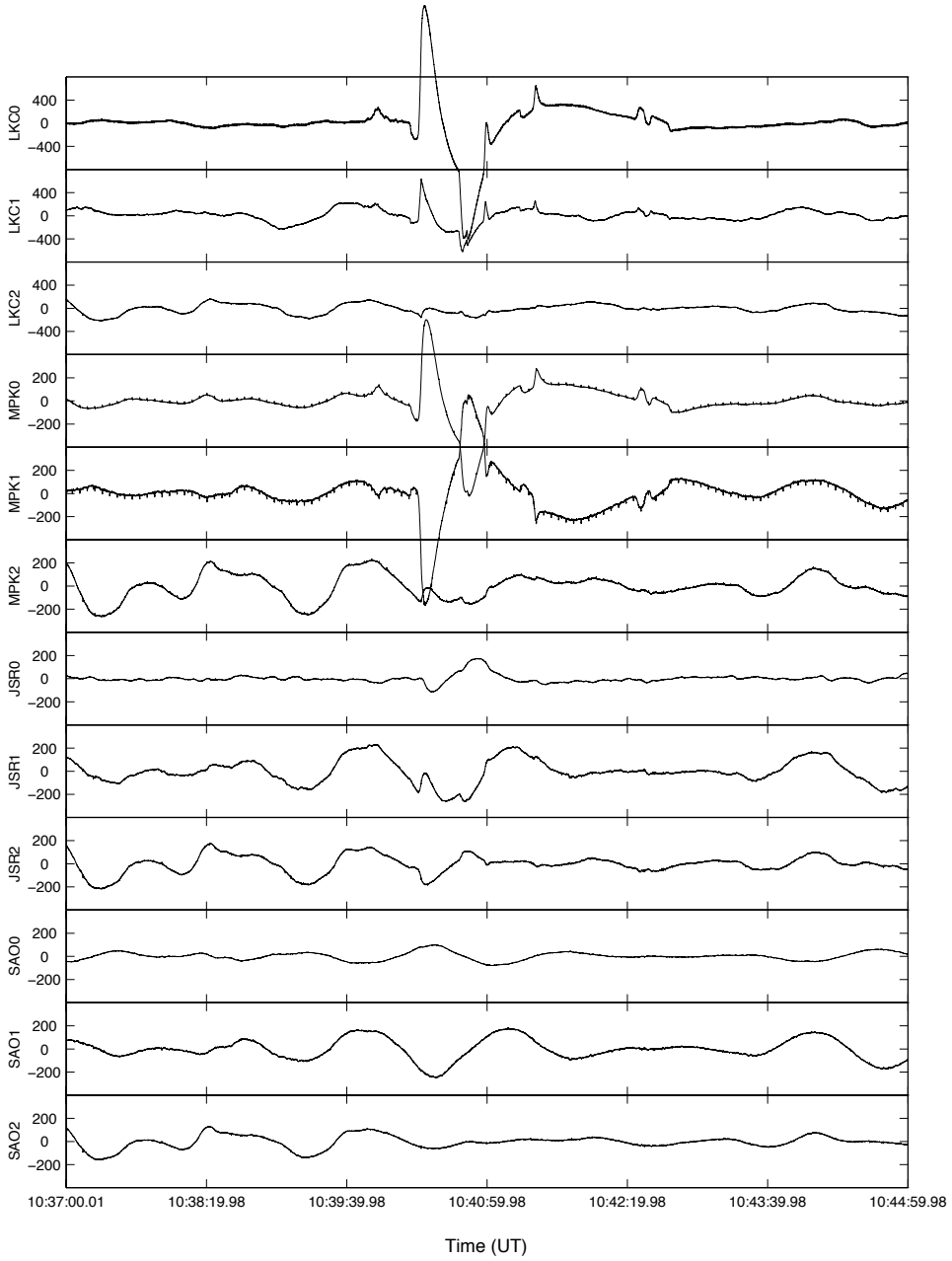


Figure 3.4: Time series data acquired at 10:37 UT on August 11, 1998 from the SFBay array and remote reference site SAO. Channels are labeled 0: vertical, 1: North-South and 2: East-West. Ordinate axis scale units are pT. Note that the ordinate axis scale for site LKC is twice as large as the scale for the other sites. Small spikes, occurring approximately every four seconds in the MPK0 and MPK1 data and, to a lesser extent, in the MPK2 data are due to an unidentified local source.

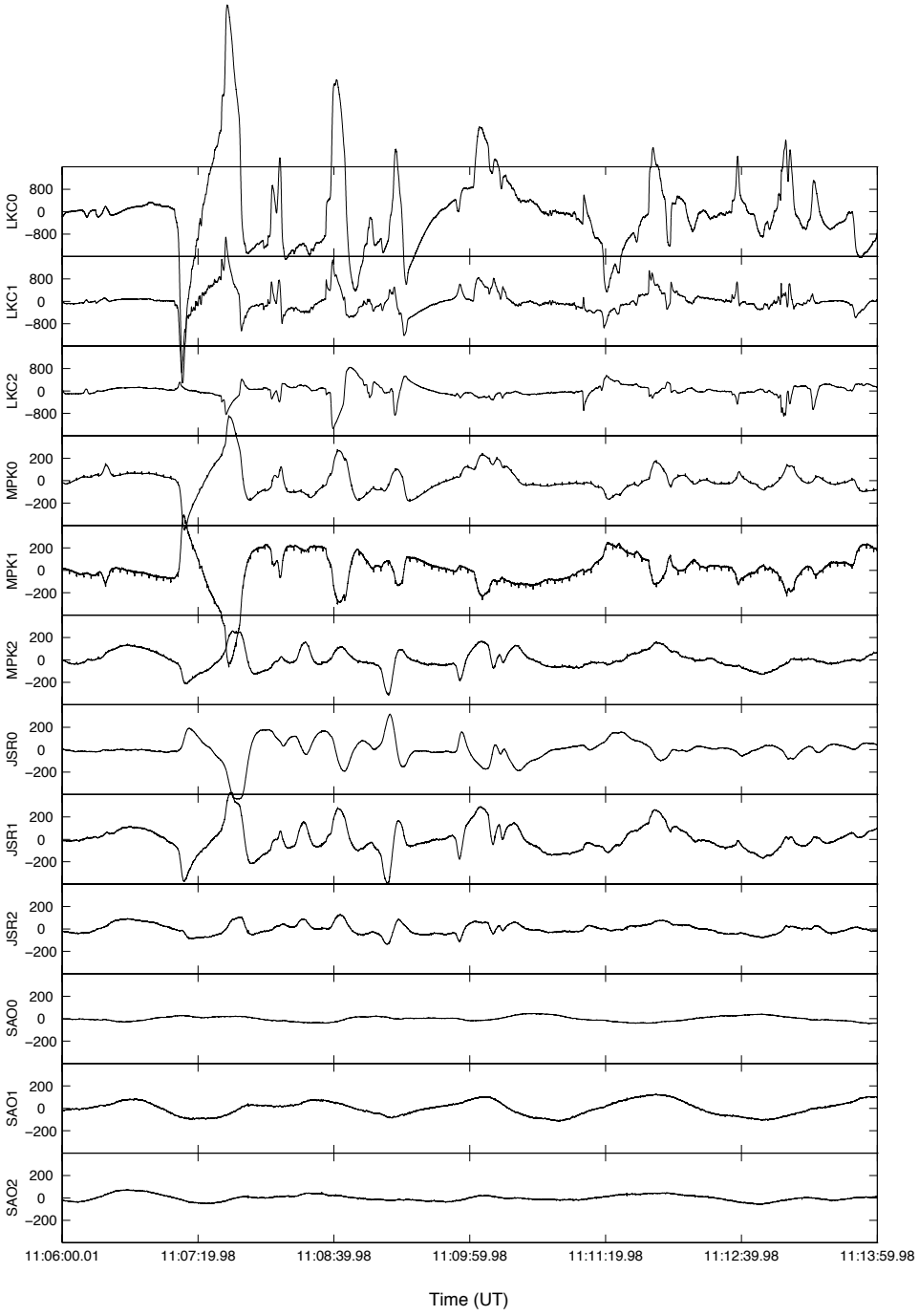


Figure 3.5: Time series data acquired at 11:06 UT on August 11, 1998 from the SFBay array and remote reference site SAO. Channels are labeled 0: vertical, 1: North-South and 2: East-West. Ordinate axis scale units are pT. Note that the ordinate axis scale for site LKC is four times as large as the scale for the other sites.

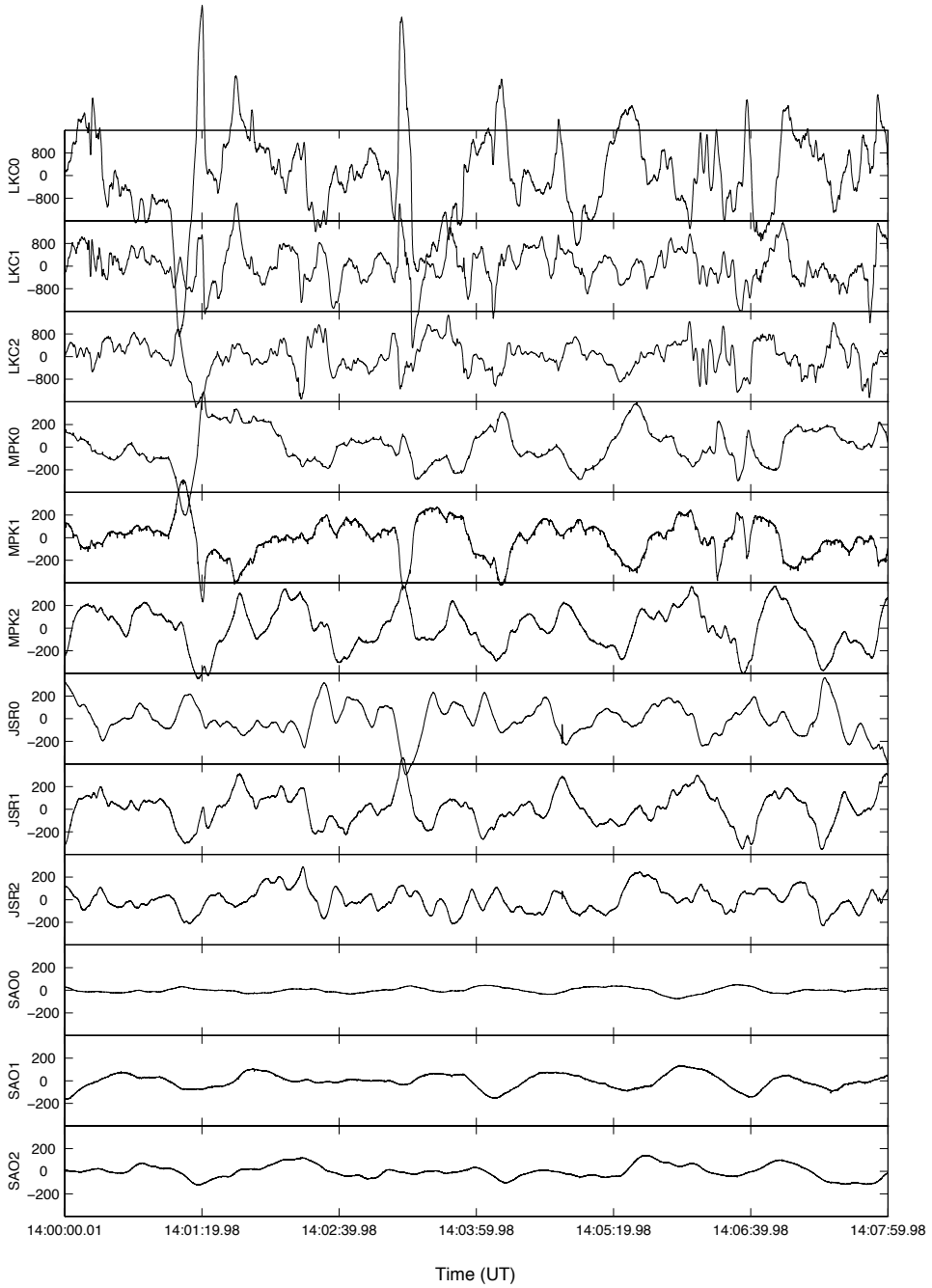


Figure 3.6: Time series data acquired at 14:00 UT on August 11, 1998 from the SFBay array and remote reference site SAO. Channels are labeled 0: vertical, 1: North-South and 2: East-West. Ordinate axis scale units are pT. Note that the ordinate axis scale for LKC is four times as large as the scale for the other sites.

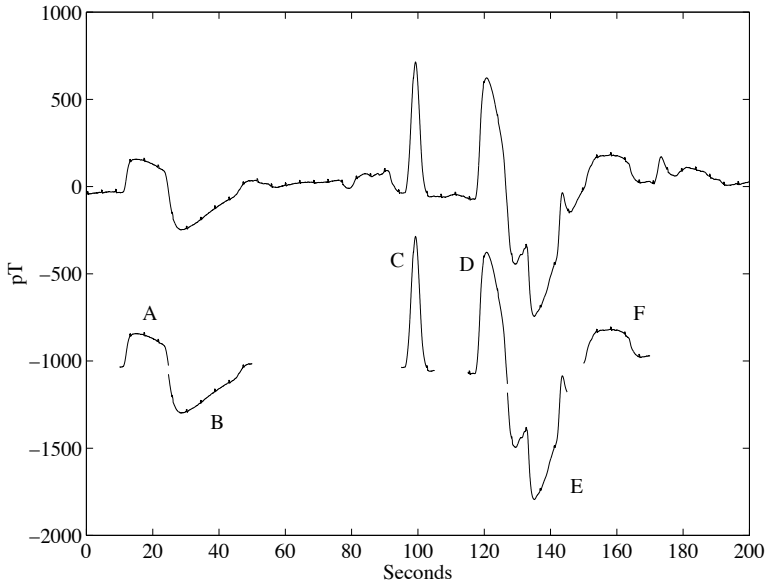


Figure 3.7: Example of BART interference acquired along the vertical axis at MPK. Significant transients are re-plotted below the original time series. The temporal widths for transients A through F are 3.1, 4.4, 0.9, 1.8, 3.6 and 3.5 seconds, respectively.

3.4.2 Amplitude Distribution

In this subsection we examine the amplitude distribution of the BART transients. To obtain the distributions, we first compute an estimate of the natural background for each channel using the remote reference data from SAO. We subtract the estimate from the original time series and then divide the resulting time series into overlapping sections. Each section is 20 seconds long and the overlap between sections is 10 seconds. For each 20 second section we determine the maximum of the absolute value of the time series. We then bin the data to obtain a histogram of the computed maxima for each channel. Since the standard deviation of the remote reference estimation process is between 10 and 20 pT, we use only those maxima with amplitudes greater than 50 pT to compute the histogram. In addition, if the same maximum occurs in two different overlapping sections, it is only counted once.

In Figure 3.8 we show the amplitude histograms computed for a half hour of data during the morning rush hour period on August 11, 1998. The largest amplitudes are observed on the vertical channel at LKC, where the range is approximately 1000 to 9000 pT with a median value of 2700 pT. The range is considerably smaller for the other

channels, with median values of approximately 1000 pT for the LKC horizontal channels, 300 pT for the MPK vertical and N-S channel, 200 pT for the MPK E-W channel and the JSR vertical and N-S channel, and 150 pT for the JSR E-W channel. By way of comparison, some typical amplitudes for natural mid-latitude geomagnetic micropulsations in this frequency range are 100 pT in the Pc1 frequency range (0.2 to 5 Hz) and 500 pT in the Pc2 to Pc3 frequency range (0.02 to 0.2 Hz) [*Campbell*, 1997]. It is important to note, however, that micropulsations are intermittent phenomena and that the geomagnetic background amplitudes in the absence of micropulsations are considerably lower than the cited micropulsation amplitudes. As a result, we find that even the smaller observed BART amplitudes for the JSR E-W channel are significant when compared to the natural background — see, for example, Figure 3.1(c).

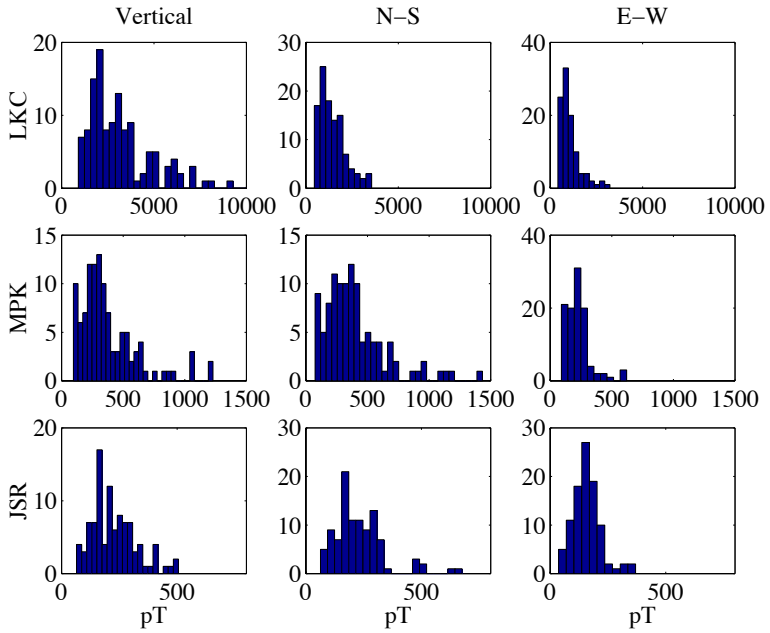


Figure 3.8: Histograms of transient amplitudes observed over a half hour period starting at 14:00 UT on August 11, 1998. The y-axis unit is number of occurrences. The x-axis limits vary with the site.

3.4.3 Polarization

Because the BART interference consists primarily of transients, we examine the polarization in the time domain. Our goal is to generate polarization vectors of the form

$$\mathbf{P}_j = \left[\begin{array}{cccc} A_{1,j} & A_{2,j} & \cdots & A_{9,j} \end{array} \right]^T \tag{3.2}$$

where $A_{i,j}$ is the amplitude of the jth transient as observed on the ith channel of the array. Note that $\mathbf{P}_j$ is a 9-dimensional vector because there are 9 channels in the SFBay array. As with the estimation of amplitude distributions, we first compute an estimate of the natural background for each channel and then subtract the estimate from the original time series. We divide the resulting time series into overlapping sections, where each section is 20 seconds long and the overlap between sections is 10 seconds, and then find the maximum of the absolute value of the data in each 20 second section. To form the polarization vectors, we associate these maxima across channels, with the constraint that only maxima with amplitude greater than 50 pT are used.

The association process proceeds as follows. We begin by considering all maxima on channel 0 as 'seed' maxima for potential polarization vectors. For each seed maximum, we determine which maxima on the other channels lie closest in the time domain to the given seed maximum. The closest maxima that are within a 10 second window of the seed maximum are then used to form a polarization vector. If there are no maxima within the 10 second window, then no polarization vector is generated from the given seed maximum. This step is taken to reduce the number of 'one-dimensional' polarization vectors — a one-dimensional polarization vector has only one non-zero component, and is formed when the maxima closest to the seed maxima lies just outside the 10 second window. The drawback is that a few transients whose polarization vectors lie exactly along one of the nine coordinate axes are omitted. Once a maxima is used as part of a polarization vector it cannot be used again for another vector. After we have exhausted all seed maxima from channel 0, we repeat the process using all remaining maxima at channel 1 as seed maxima. Once all seed maxima from channel 1 have been used, we then repeat the process for channel 2, then channel 3, and so on.

In panels (a)-(c) of Figure 3.9 we have plotted the polarization vectors for a half hour from the morning rush hour period on August 11, 1998. For display purposes, we have split each 9-dimensional polarization vector into three smaller vectors, each of which is

the 3-dimensional polarization vector for one of the three sites in the array. We see that although the source of the interference is complex and distributed, each set of polarization vectors exhibits a strong directional component. In addition, the space occupied by the vectors is fairly well constrained. That is, if we denote the 3-dimensional vector for the jth transient at the ith site as $\mathbf{P}_{i,j}$, then we find that $\mathbf{P}_{i,j} \in S_i$ where $S_i \subset \mathcal{R}^3$ is a subset of Euclidean 3-space. In Chapter 7, we present a method for constructing these subsets and show how they can be used to distinguish BART-related transients from transients due to natural sources.

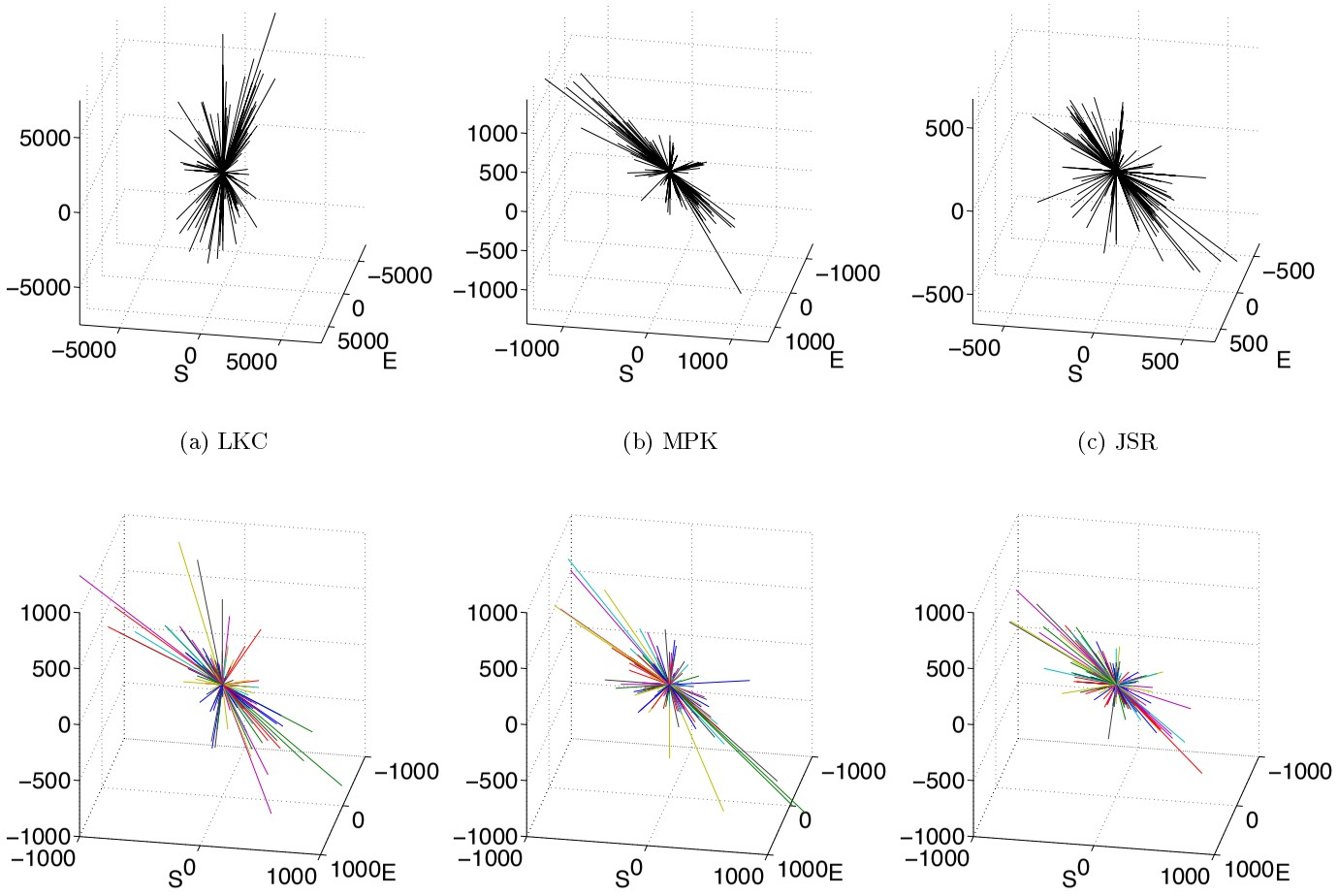
In panels (d)-(f), we show the 3-dimensional polarization vectors at MPK for the same half hour period (11:30 to 12:00 UT) over the course of three days. While the exact features of the vectors vary slightly from day to day, the overall shape of the subset occupied by the vectors is roughly the same. As a result, for each half hour period of the day, we can construct subsets that approximately describe the distribution of the polarization vectors during that period.

3.5 Power Spectra

3.5.1 Observations

Representative power spectra are shown in Figures 3.10 and 3.5.1. Spectra were computed with the averaged periodogram method [*Oppenheim and Schafer*, 1989] and the following parameters: 8192 point sections, 50% overlap between adjacent sections, and a Hanning window. Figure 3.10 shows the spectra for the half hour period 09:30 to 10:00 UT. In each subfigure we plot the spectra for the SFBay and SAO sites along a specified orientation (e.g., North-South). Although there is no scheduled BART activity during this period, trains are still moved about for maintenance and testing purposes. This activity is reflected in the increased levels at around 0.1 Hz for the SFBay sites. The rise in the LKC vertical channel spectra for frequencies above 2.0 Hz is due to a problem with the instrumentation[1]. Spikes in the MPK spectra are due to the periodic interference discussed in the caption to Figure 3.4. The spectral peaks near 8 Hz in the N-S and E-W components are due to the Schumann resonance phenomenon [*Vozoff*, 1991]. Aside from the differences just described, we can see that the spectra from the SFBay agree fairly well with the spectra

[1]The cause of this problem is still under investigation. It has minimal impact on the results of this thesis, since BART interference is most pronounced for frequencies below 1.0 Hz



(a) LKC (b) MPK (c) JSR

(d) MPK August 10 (e) MPK August 11 (f) MPK August 12

Figure 3.9: BART interference polarization plots. Panels (a)–(c) show the polarization at each site for the period 14:00 to 14:30 UT on August 11, 1998. Panels (d)–(f) show the polarization at MPK for the period 11:30 to 12:00 UT on August 10 through 12, 1998. Units are pT. Horizontal axes are labeled south (S) and east (E).

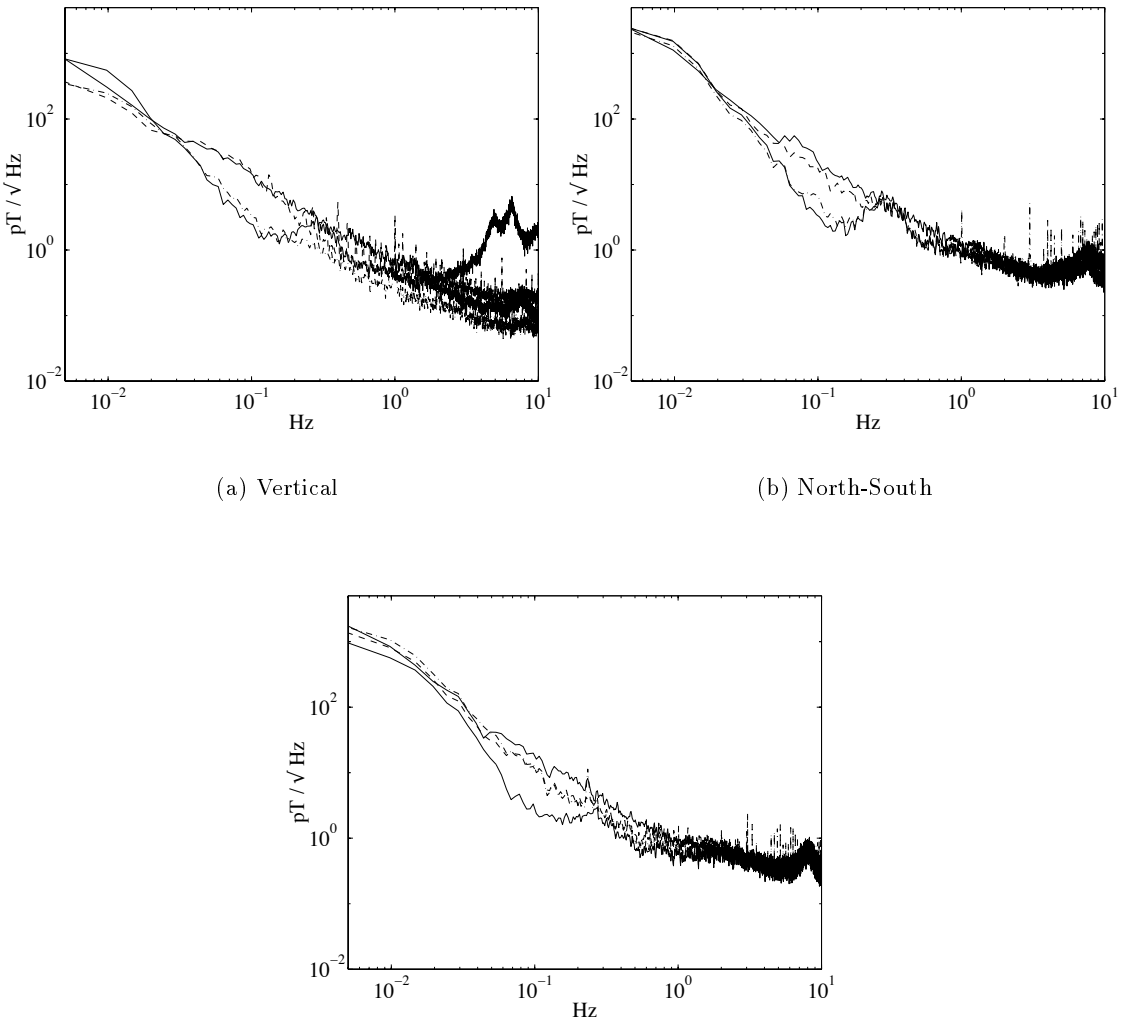
from SAO when there is limited BART activity. The increase in spectral amplitudes for frequencies near 0.3 Hz is due to natural sources — a similar increase is observed on the N-S component at TBM, which is approximately 500 km to the south of the SFBay and thus far outside the range of the BART interference.

The distances from the JSR and Stanford sites to the BART tracks are essentially the same. In Figure 3.5.1, panels (a)–(c), we show the spectra for a half hour span (14:00 to 14:30 UT) during the morning rush hour period. The ratios of the SFBay spectra to the SAO spectra are shown in panels (d)–(f). Consistent with previous observations [*Egbert*, 1997; *Fraser-Smith and Coates*, 1978; *Ho et al.*, 1979], we see that the largest increase occurs between 0.1 and 0.2 Hz with BART spectral amplitudes ranging from 10 to 1000 times greater than the natural background levels. In Figure 3.12, we group the spectra by site and observe that at LKC the interference is strongest on the vertical channel, while at MPK and JSR the interference is roughly the same level for all channels. The spectral amplitudes at JSR are of the same order as those previously observed at the Stanford, CA site [*Fraser-Smith and Coates*, 1978], which is located 5.3 km to the east of JSR.

3.5.2 Comparison with Loma Prieta Data

To provide some perspective on the magnitude of the observed increases, we compare the rush hour spectra with spectral averages recorded at Corralitos (COR) prior to the October 17, 1989 Loma Prieta M_s 7.1 earthquake. At the time of the earthquake the COR system stored only half hourly spectral averages computed for the N-S channel — no raw time series data were saved. The anomalous ULF signals observed by *Fraser-Smith et al.* [1990] prior to the earthquake can be divided into roughly three periods. An anomalous narrowband signal that alternated between the two bands, 0.05-0.10 Hz and 0.1-0.2 Hz, appeared on September 12, 1989 and lasted until October 5, 1989. On October 5, 1989, a significant increase was seen in all frequency bands. The elevated signals gradually decreased in strength until the day of the earthquake. Three hours prior to the earthquake, at approximately 21:00 UT on October 17, 1989, an exceptionally large increase was observed in the frequency band 0.01 to 0.5 Hz, with the observed amplitudes at 0.01 Hz exceeding the typical natural background levels by a factor of over 100.

In Figure 3.13(a) we show the rush hour (14:00 to 14:30 UT on August 11, 1998) spectra for the N-S channels at LKC, MPK, JSR, and SAO along with example COR spectral averages from each of the three time periods described above. We can see that



(a) Vertical

(b) North-South

(c) East-West

Figure 3.10: Power spectra from SAO (lower solid), LKC (upper solid), JSR (dash), and MPK (dash-dot) for the period 09:30 to 10:00 UT on August 11, 1998. Distinction between upper and lower solid lines is based on spectral amplitudes at 0.1 Hz. In panels (a) and (b), the upper solid and lower solid lines cross at about 0.02 Hz.

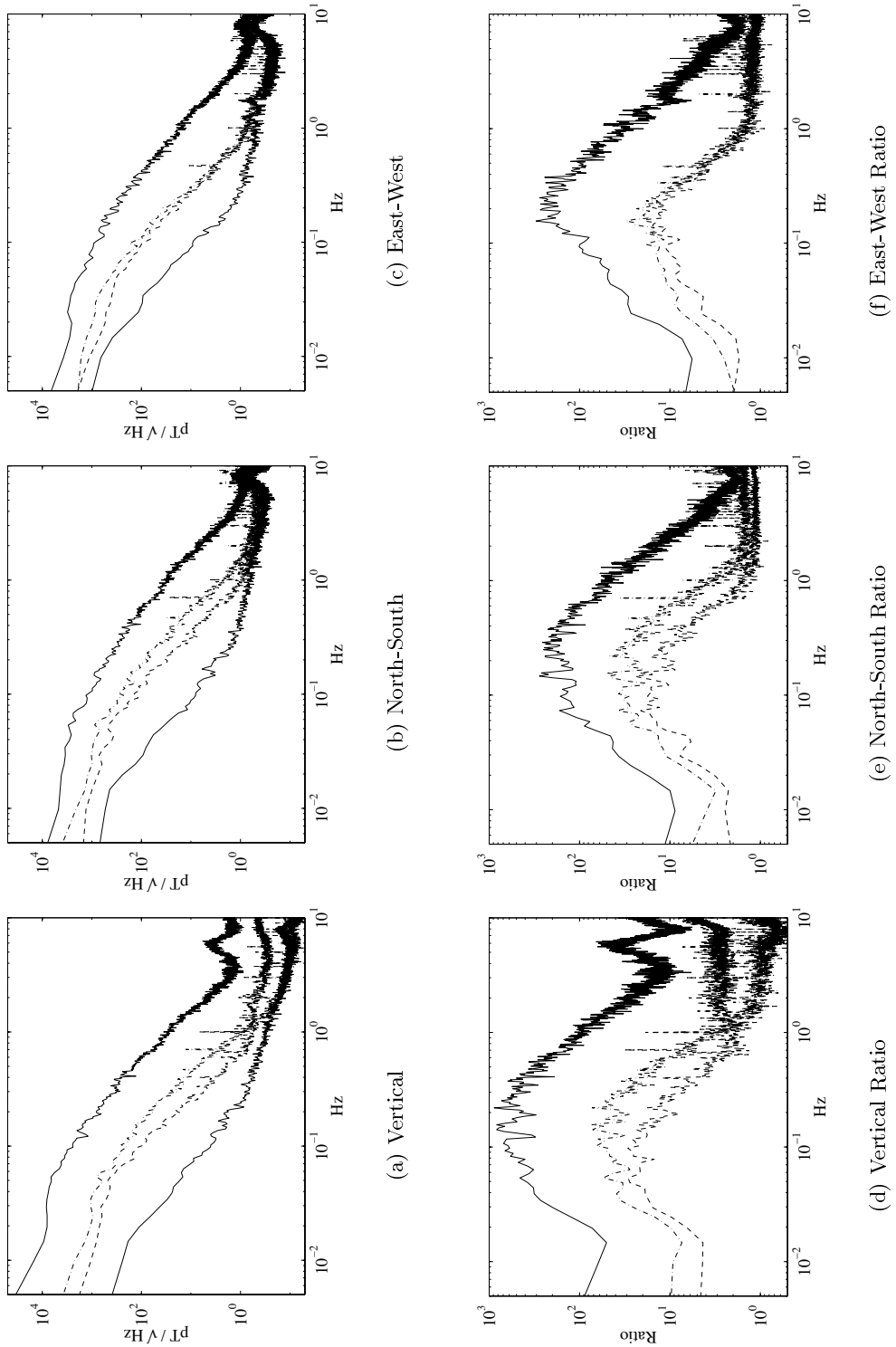
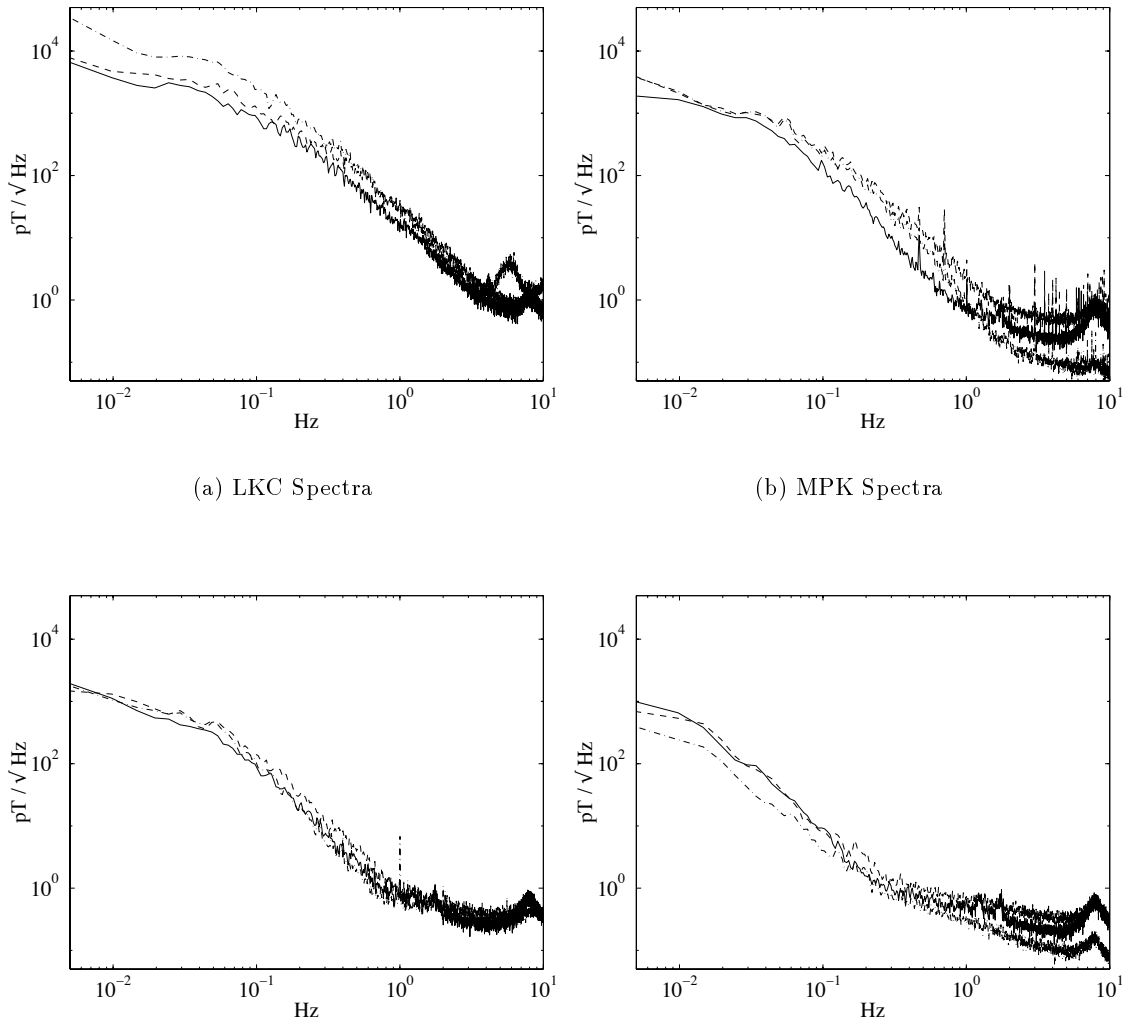


Figure 3.11: Power spectra and spectral ratios. Panels (a)–(c), power spectra from SAO (lower solid), LKC (upper solid), JSR (dash), and MPK (dash-dot) for the period 14:00 to 14:30 UT on August 11, 1998. Panels (d)–(e), LKC (solid), MPK (dash-dot), and JSR (dash) power spectra divided by the SAO power spectrum.



(a) LKC Spectra

(b) MPK Spectra

(c) JSR Spectra

(d) SAO Spectra

Figure 3.12: Power spectra for the period 14:00 to 14:30 UT on August 11, 1998 grouped by site. Spectra are shown for vertical (dash-dot), North-South (dash), and East-West (solid) channels.

COR spectral averages from the second and third time periods (i.e., averages from October 9, 1989 and October 17, 1989) are greater than the observed amplitudes at MPK and JSR for all frequency bands and of the same order or greater than the observed amplitudes at LKC for the lower frequency bands. The narrowband anomaly observed near 0.1 Hz on September 25, 1989 is roughly the same amplitude as the LKC spectra.

Although the spectra shown in Figure 3.13(a) seem to imply that an earthquake precursor of the type observed at COR would be easily observable across the SFBay array, even in the presence of BART, we must be careful not draw such a conclusion. The COR system was located 7 km from the epicenter of the Loma Prieta earthquake. The depth of the quake hypocenter was estimated to be 17.6 km [*Fraser-Smith et al.*, 1993]. As a first order approximation, we may assume that the source is a horizontal electric dipole (HED) located at the hypocenter [*Fraser-Smith et al.*, 1993]. Due to the non-zero conductivity of the earth, the magnetic fields produced by the dipole decay rapidly as a function of distance from the source, and the rate of decay increases as function of frequency. As a result, we would not expect the fields due to a precursor to be of the same magnitude at all sites in the SFBay array.

To expand on this point, we consider the fields due to a HED located at a depth of 5 km on the section of the Hayward Fault just to the west of LKC. This section of the fault is referred to as the San Leandro salient and was the site of a M_L 4.7 earthquake that occurred in 1984 at a depth of 5 km [*Lienkaemper et al.*, 1991]. The assumed HED points to the southeast and is aligned with the strike of the fault. As this is a first-order analysis, we assume a homogeneous earth with a conductivity of 0.03 S/m, which is the value in the SFBay conductivity model, presented in Appendix A, for the depth range 0.5 to 15 km. For comparison, we note that *Fraser-Smith et al.* [1993] assumed a homogeneous earth with a conductivity of 0.1 S/m. We use the formulas in [*Fraser-Smith and Bubenik*, 1980] to compute the fields at each of the sites in the array, assume that the N-S fields at LKC are of the same magnitude as the COR measurements, and then calculate the ratios between the magnitudes of N-S fields at MPK and JSR to the N-S field at LKC. We then multiply the COR measurements by these ratios to obtain the expected field magnitudes at MPK and JSR. In panels (b) and (c) of Figure 3.13, we present spectra for MPK, JSR, and SAO (time period is the same as for panel (a)), along with the downweighted COR spectral averages. Due to propagation effects, the expected field magnitudes at MPK and JSR are significantly reduced. For frequencies above 0.3 Hz, the expected magnitudes are

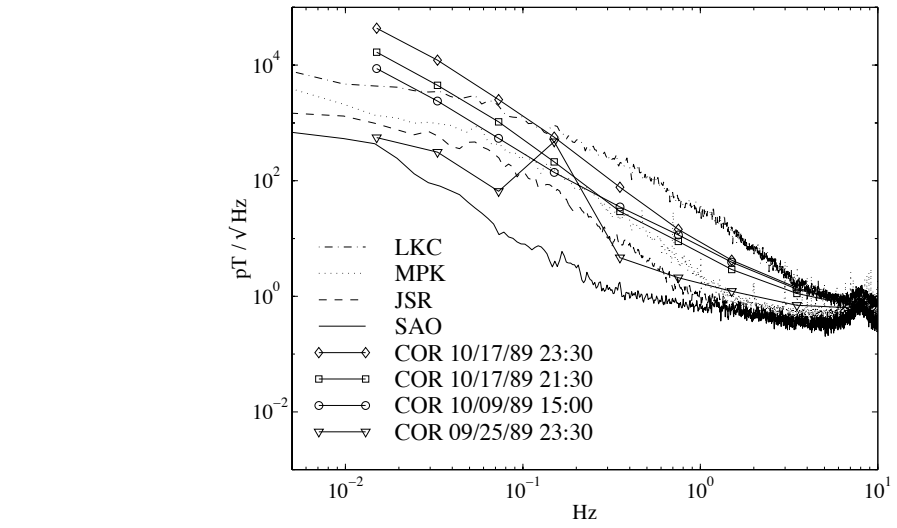
less than the natural background levels at SAO. For frequencies below 0.3 Hz, most of the expected magnitudes lie between the SAO background levels and the MPK or JSR levels. This example highlights the difficulty in observing precursor signals at all sites in the SFBay without a means of removing the effects of the BART interference. In our example, the source is located close to one site, LKC. The challenge of observing precursor signals increases if the source were to be located at a significant distance from all sites in the array — for example, near the northern end of the Hayward Fault

3.5.3 Spectral Indices

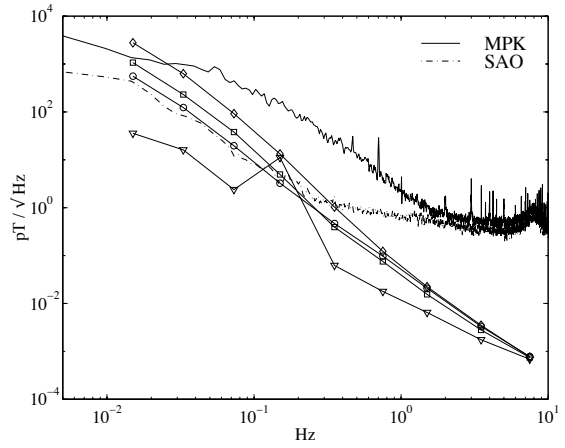
Spectral or magnetic activity (MA) indices are useful for examining the daily variation in the BART generated ULF fields. For the LKC system, each MA index is the logarithm to the base 2 of the average spectral power, computed for a half-hour interval, in one of 9 distinct frequency bands covering the range 0.01 to 10 Hz. The indices are generated every half hour and can be converted to units of $\mathrm{pT}/\sqrt{\mathrm{Hz}}$ through the use of a calibration table. See [*Bernardi et al.*, 1989] for a more detailed description of the MA indices and their relation to other geomagnetic indices. For the MPK and BKS systems, the MA indices cover the range 0.001 to 10 Hz [*Liu*, 1996]. The indices are generated every 15 minutes from overlapping half-hour sections. At present MA indices are not automatically generated for the JSR system.

In Figure 3.14 we have plotted the MA indices of the N-S component at LKC for the week April 22, 1996 through April 28, 1996. The data depict very clearly the levels of the BART interference as a function of the time of day and day of the week. The levels are well correlated with the hours of BART interference, nominally 04:00 to 01:30 LT on weekdays, 06:00 to 01:30 LT on Saturdays, and 08:00 to 01:30 LT on Sundays. For example, we note that the periods without BART interference are longer in duration for the weekend days (April 27 and 28).

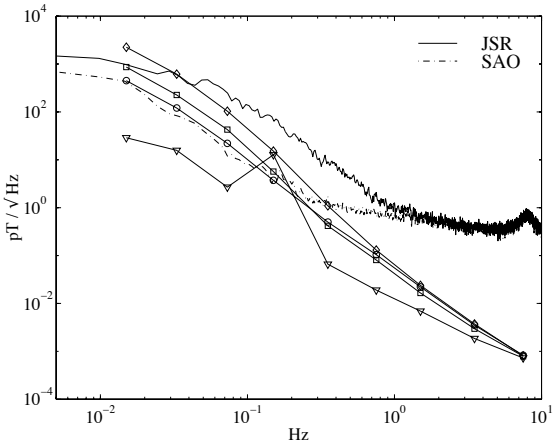
As shown in the next chapter, the BART interference is primarily due to propulsion and regenerative braking currents flowing along the tracks and through the earth. Given this relation, we would expect there to be a correlation between the interference levels and the amount of power used by the BART system. We obtained power data from Pacific Gas and Electric (PG&E), which monitors and records, on a half-hourly basis, the power supplied to the BART traction system. For each half hour of the day, we computed the median MA index (LKC, N-S component) for the 0.02-0.04 Hz frequency band and the



(a)



(b)



(c)

Figure 3.13: Comparison of power spectra with Loma Prieta Data. Panel (a): Spectra from LKC, MPK, JSR, and SAO (N-S channel) for the period 14:00 to 14:30 UT August 11, 1998 compared to spectral averages from COR (N-S channel) for dates prior to the 1989 Loma Prieta earthquake. The notation COR 9/25/89 23:30 corresponds to the spectral average for the half hour period ending at 23:30 UT on September 25, 1989. Panels (b) and (c): Spectra with COR magnitudes multiplied by ratios of expected N-S fields at MPK and JSR, respectively, to expected N-S field at LKC.

median BART total traction power for all weekdays in the period March 1, 1996 through April 30, 1996. The results are shown in Figure 3.15, where we have used the logarithm to the base 2 of the power (in units of Kilo-watt hours). The striking resemblance between the two curves indicates that the median traction power is well correlated with the median BART interference. In contrast, attempts to use regression analysis to predict the half-hourly sample values of the MA indices from the corresponding traction power values have met with limited success, thus indicating that the good correlation observed for median values does not extend to the sample values.

3.6 Characteristics of the Geomagnetic Background

In this section we examine the $1/f$ behavior observed in the power spectra of the natural geomagnetic background. After a review of previous work, we present example power spectra using data from SAO. Although processes which exhibit $1/f$ power spectra have been widely observed in nature, a useful model for signal processing applications has been lacking. Recently, *Wornell* [1996] proposed a wavelet-transform-based model for what he terms *nearly*-$1/f$ *processes*. The model allows one to both synthesize and analyze processes with spectra that are approximately $1/f$. In addition, it provides a convenient framework for the detection of transients in the presence of $1/f$ noise, which we discuss in detail in a later chapter.

3.6.1 Previous Observations

We begin with a working definition for $1/f$ power spectra. Denoting the magnetic field as $b(t)$ and its power spectrum as $S_b(f)$, we say that the spectrum is an approximate $1/f$ spectrum if

$$S_b(f) \approx \frac{\sigma_b^2}{f^\gamma} \tag{3.3}$$

over some specified range $f_{lo} < f < f_{hi}$. We assume that the parameter $\gamma > 0$. If $S_b(f)$ satisfies (3.3), then $b(t)$ is referred to as an *approximate* $1/f$ process. We also use the term '$1/f^\gamma$ process' in situations where we wish to make clear that γ need not be equal to one, e.g., in the abstract of this dissertation. Spectra that are approximately $1/f$ appear as straight lines on log-log plots of power spectral density versus frequency, where $-\gamma$ is the slope of the line. We will also introduce the term *piecewise* $1/f$ to refer to power

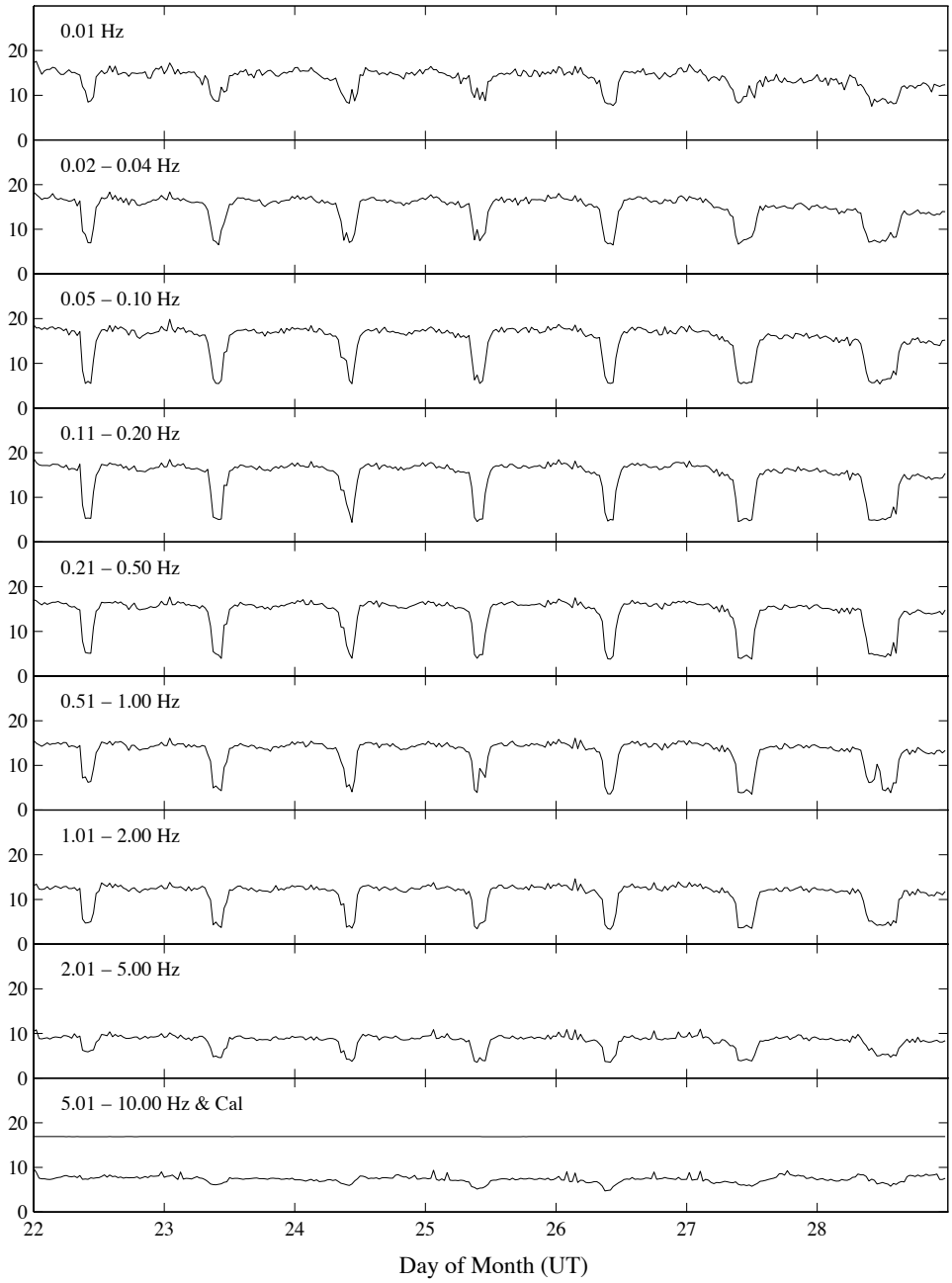


Figure 3.14: MA Indices (LKC, N-S component) for the period April 22, 1996 through April 28, 1996. The straight line in the bottom row is the calibration tone. Subtract 8 hours to obtain time in LT.

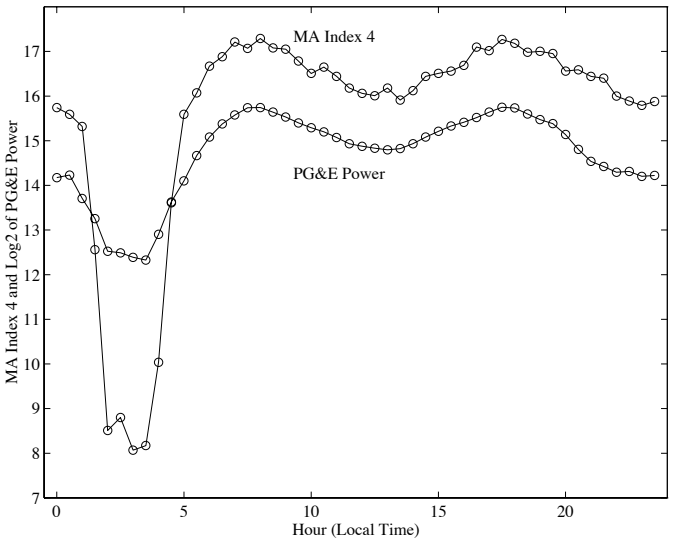


Figure 3.15: Median values for MA Index 4 (0.02 to 0.04 Hz; LKC, N-S component) and logarithm to the base 2 of PG&E total power (KWH) for the period March 1, 1996 through April 30, 1996.

spectra which are $1/f$ on connected intervals of the frequency axis. For example, we say that $S_b(f)$ is piecewise $1/f$ over the interval $[f_1, f_3]$ if

$$S_b(f) \approx \begin{cases} \frac{\sigma_1^2}{f^{\gamma_1}} & \text{for } f \in [f_1, f_2] \\ \frac{\sigma_2^2}{f^{\gamma_2}} & \text{for } f \in [f_2, f_3] \end{cases} \tag{3.4}$$

where $f_1 < f_2 < f_3$, $\gamma_1 \neq \gamma_2$, and $\frac{\sigma_1^2}{f_2^{\gamma_1}} = \frac{\sigma_2^2}{f_2^{\gamma_2}}$.

Various researchers have observed that geomagnetic background spectra exhibit approximate $1/f$ type behavior over a wide range of frequencies. *Lanzerotti et al.* [1990] considered both daily and monthly averaged power spectra from Arrival Heights, Antarctica for the frequency range 10^{-5} to 10^5 Hz (frequencies from 0.1 to 10 Hz were not included in the study). They found that up to frequencies of 3 KHz, the averaged spectra were approximately $1/f$ with an estimated γ in the range 2.0 to 3.0.

Campbell [1976] computed the average spectra over the range 7×10^{-5} to 3.3×10^{-3} Hz for a month of data collected from over ten sites located in the American hemisphere. He found that the average spectra was approximately $1/f$ with $\gamma = 2.0$. He studied in more

detail the spectra at a pair of high latitude and at a pair of low latitude sites, dividing the spectra into three ranges, 7×10^{-5} to 2.8×10^{-4} Hz, 2.8×10^{-4} to 1.7×10^{-3} Hz, and 1.7×10^{-3} to 3.3×10^{-3} Hz and computing the γ values for each range. For the high latitude sites, he found γ values ranging from 1.0 to 4.0 with the highest values associated with the high frequency range and the lowest values associated with the low frequency range — i.e., the $1/f$ fall-off was steeper for higher frequencies. For the low latitude sites, he found γ values ranging from 0.8 to 2.6 with the highest values associated with the low frequency range and the lowest values associated with the high frequency range — i.e., the $1/f$ fall-off became less steep for higher frequencies. In both cases, the variation in γ values was determined by looking at the daily averages over a four month period. Systematic variations in γ were observed also as a function of time of day and season.

For the range 0.083 to 1.3 Hz, *Wertz and Campbell* [1976] found γ values ranging from 3.0 for low-latitude and mid-latitude sites to 4.0 for high-latitude sites. *Davidson* [1964] found that spectra over the range 1×10^{-3} to 0.4 Hz could be broken into two ranges: an interval from 0.1 to 0.4 Hz characterized by $\gamma = 1.3$, and a second interval from 1×10^{-3} to 0.1 Hz characterized by $\gamma = 2.6$. The data were from a mid-latitude (geographic 39.63° N) site.

3.6.2 New Observations

In Figure 3.16(a) we present the mean spectral amplitudes during the month of August 1998 for the E-W channel at SAO. The means are calculated using the spectral indices, which are calculated every 15 minutes for overlapping half-hour periods. The mean values are grouped into six 4-hour intervals. For example, the mean value for 08:15 - 12:00 UT represents the mean of all values acquired for that interval over the course of a month. Over the frequency range 0.001 to 0.35 Hz, the spectra are approximately $1/f$ with $\gamma \approx 3.2$. From 0.35 to 8 Hz the spectra are approximately $1/f$ with $\gamma \approx 1.3$. These values are consistent with those reported in [*Davidson*, 1964] and [*Wertz and Campbell*, 1976]. The discontinuity in γ value near 0.35 Hz coincides roughly with the transition between magnetospheric and lower atmospheric sources. Typically, fields at frequencies above 1.0 Hz are attributed to electric storms in the lower atmosphere, while fields at frequencies below 1.0 Hz are associated with hydromagnetic waves in the magnetosphere [*Vozoff*, 1991]. The division is not absolute — magnetospheric sources can contribute to fields above 1.0 Hz and lower atmospheric sources can give rise to spectral energy below 1.0 Hz.

For example, Pc1 hydromagnetic narrowband emissions have been observed at frequencies ranging from 0.2 to 5 Hz [*Jacobs*, 1970].

The range of spectral amplitudes observed over a month can be quite wide, and a few large amplitude events can have a large effect on the mean. In addition, intermittent interference sources or instrumentation problems also can give rise to outlier values that distort the mean. To address these problems, the median is often chosen as a more robust estimate of the 'center' of a population [*Press et al.*, 1988]. In Figure 3.16(b) we show the mean, median, maximum, and minimum values for the one of the 4-hour intervals, as observed during the month of August 1998. For most frequency bands, the median value is significantly smaller than the mean value — in some cases more than an order of magnitude smaller. The disparity between the mean and median values indicates that the median values may provide a better description of a typical spectrum. As an example, in Figure 3.16(c) we plot the median spectra for two different 4-hour intervals along with sample spectra for half hour periods that lie within each of the intervals for the date August 11, 1998. Here we see that the shapes of the median and sample spectra show good agreement. There is a steep $1/f$ transition zone with $\gamma = 4.5$ from about 0.025 Hz to 0.2 Hz. On either side of the transition zone the spectra are $1/f$ with γ values of 2.5 and 1.3. In Figure 3.16(d) we plot the median values for the month of August 1998 for the six 4 hour intervals. We note that the steep transition zone appears for all intervals. The significance of this zone is not known. In summary, both the mean and median values indicate that geomagnetic background is piecewise $1/f$ in the ultra-low frequency band.

3.6.3 Wavelet Based Models for $1/f$ Processes

A variety of models have been proposed for the synthesis of processes that exhibit $1/f$ power spectra. Among these are the fractional Brownian motion model [*Flandrin*, 1992], the autoregressive moving average (ARMA) model [*Kershner*, 1982], the clustering Poisson process model [*Gruneis and Musha*, 1986], and the wavelet transform based model [*Wornell*, 1996]. *Chrissan* [1998] has shown that the statistics of lightning flashes are well-described by a clustering Poisson process model. Such a model may explain the $1/f$ nature of the observed spectra for frequencies above 0.35 Hz. The physical significance of the $1/f$ behavior observed at lower frequencies is less clear. Over the frequency range 1.7×10^{-3} to 5 Hz, the typical amplitudes of continuous (Pc) micropulsations from the magnetosphere decrease as a function of increasing frequency [*Campbell*, 1967]. However,

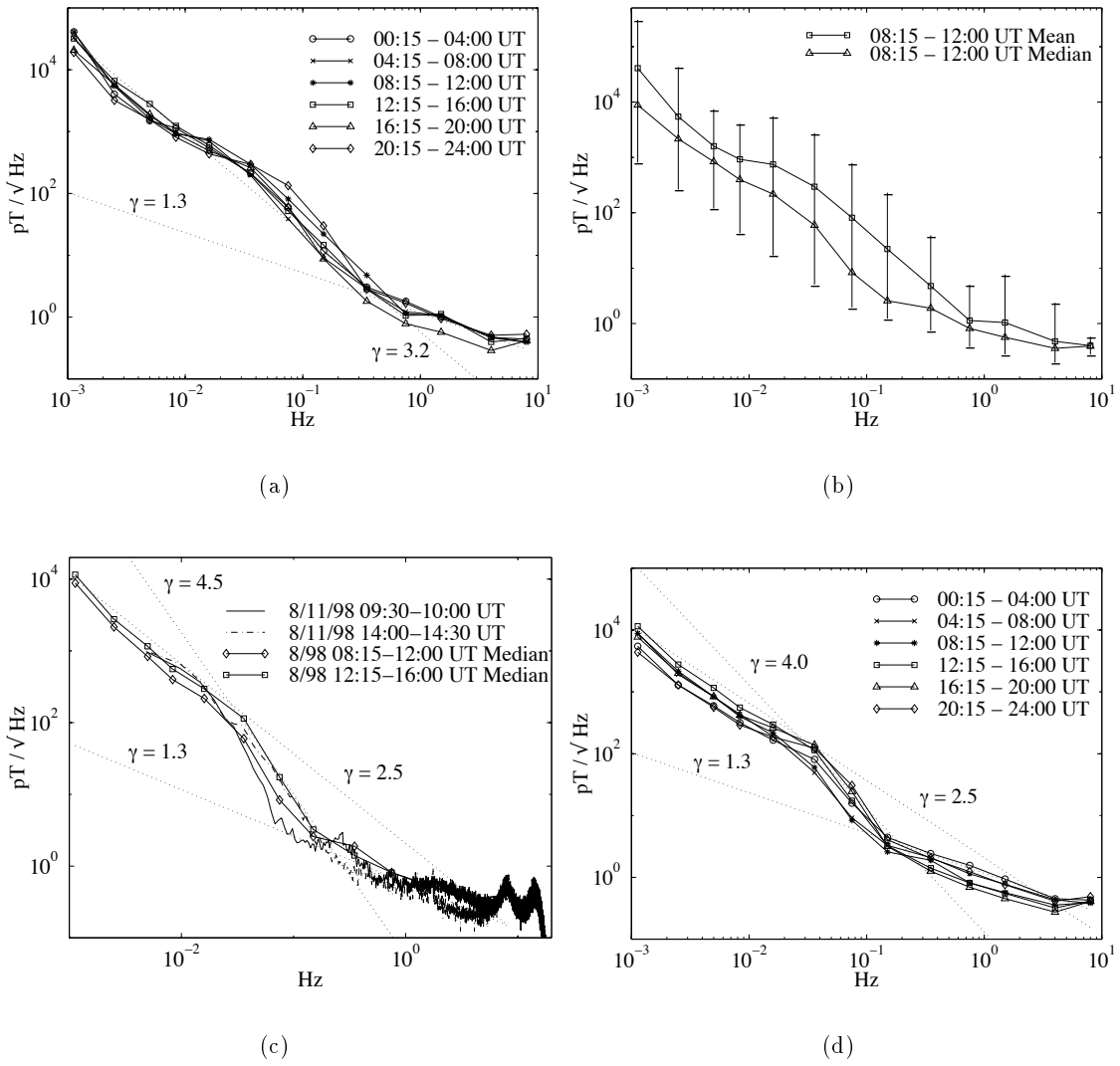


Figure 3.16: Mean and median half hourly spectral averages observed during August 1998 on E-W channel at SAO. (a) Mean spectral averages. (b) Mean, median, maximum, and minimum spectral averages. (c) Median spectral averages with two sample spectra. (d) Median spectral averages. Dotted lines represent slopes of ideal $1/f$ processes.

Pc micropulsations are narrowband phenomena which give rise to spectral peaks as, for example, the peak near 0.3 Hz shown in Figure 3.10. *Wertz and Campbell* [1976] suggested that the absence of peaks in monthly averages of the spectra were due to the fact that (1) broadband irregular (Pi) type micropulsations are more significant and (2) variations in the frequencies of the Pc micropulsations cause the spectral peaks to be averaged out. *Davidson* [1964] hypothesized that the $1/f$ nature of the spectra may bear some relation to the theoretical and experimentally observed power spectra ($\gamma = 5/3$) of turbulent processes.

Synthesis models are important for understanding the physical mechanisms which give rise to $1/f$ spectra and for simulating $1/f$ processes, but they are generally not useful for signal processing tasks such as detection and estimation. An exception is the discrete wavelet transform (DWT) model which provides an effective framework for both synthesis and analysis [*Wornell*, 1996]. As an analysis tool, the DWT acts as an approximate whitening transform for $1/f$ processes ([*Wornell*, 1996] and references therein). In addition, *Abry et al.* [1995] have shown that wavelet based methods provide an unbiased estimate of the parameter γ in contrast to periodogram based power spectrum estimation methods which yield biased estimates of γ.

We focus on the whitening transform property of the DWT since it will be used in subsequent chapters to develop a detection scheme for transients in the presence of the geomagnetic background. From [*Wornell*, 1996] we have the following theorem: If $v(t)$ is a $1/f$ process with power spectrum $\frac{\sigma_v^2}{f^\gamma}$ over a range of frequencies $[f_{lo}, f_{hi}]$ then the DWT coefficients

$$V_{m,n} = \int v(t) 2^{-m/2} \psi(2^{-m}t - n) dt$$

are approximately uncorrelated, i.e.,

$$E\left[V_{m,n} V_{m',n'}\right] \approx \sigma_m^2 \delta[m - m'] \delta[n - n'] \tag{3.5}$$

where $E[V]$ denotes the expected value of V. The variance term is defined as $\sigma_m^2 = \sigma_w^2 2^{\gamma m}$ where $\sigma_w^2 = (2\pi)^\gamma \eta \sigma_v^2$,

$$\eta = \int_{-\infty}^{\infty} \frac{|\Psi(f)|^2}{|2\pi f|^\gamma} df,$$

and $\Psi(f) = \int_{-\infty}^{\infty} \psi(t) e^{-j2\pi f t} dt$ is the Fourier Transform of $\psi(t)$. The wavelet $\psi(t)$ is defined using the framework of an orthogonal multiresolution analysis [*Daubechies*, 1992]. From

the theory of frames [*Daubechies*, 1992], we find that $\eta = \frac{ln2}{\pi}$ for the case of orthonormal wavelets and $\gamma = 1$. Closed form expressions for the special case of ideal bandpass wavelets have been derived for $\gamma \neq 1$ [*Wornell*, 1996].

The approximate decorrelation property can be made slightly more precise with the following result [*Wornell*, 1996],

$$E\left[V_{m,n}V_{m',n'}\right] \sim O\left(|2^{-m}n - 2^{-m'}n'|^{-\lceil 2R-\gamma \rceil}\right) \tag{3.6}$$

as

$$|2^{-m}n - 2^{-m'}n'| \to \infty$$

where R is the number of vanishing moments of the wavelet $\psi(t)$ and $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. Expression (3.6) states that the correlation between DWT coefficients falls off rapidly with the distance $|2^{-m}n - 2^{-m'}n'|$ between coefficients. A similar, but slightly stronger, statement was made in [*Tewfik and Kim*, 1992] for the special case of fractional Brownian motion.

Variance of Wavelet Coefficients

We now present the results of applying the DWT analysis to time series data from SAO. Our goal is to determine the extent to which the DWT acts as a whitening transform for the geomagnetic signals. In Figure 3.17 we plot the logarithm (base 2) of the sample variance of the DWT coefficients $V_{m,n}$ as a function of scale m for three half-hour periods on August 11, 1998. For each half-hour period, we computed the DWT of the first 2^{16} points (27.3 minutes). Coefficients subject to end effects were not included in the computation of the sample variances. We used a Coiflet parameter 2 wavelet (C_2), which has $R = 4$ vanishing moments [*Daubechies*, 1992]. The decorrelating property of the DWT relies on the condition $2R > \gamma$ [*Wornell*, 1996]. Based upon the spectral plots shown previously, we can assume that $\gamma < 5$, so that any wavelet with $R > 2.5$ satisfies the condition.

For a $1/f$ process, a plot of the logarithm (base 2) of the sample variance versus scale m yields a straight line with slope $-\gamma$. From Figure 3.17, we see that for the range of scales shown, the geomagnetic sample variances lie approximately upon two straight lines, with $\gamma = 1.3$ for scales 1 to 7 and $\gamma = 3.6$ for scales 7 to 11. In order to relate the wavelet variance plot to the spectral plots, note that the wavelet coefficients at a given scale m may be interpreted as samples of a bandpass filtered version of the input signal $v(t)$, where

$2^{-m/2}\psi(-2^{-m}t)$ is the bandpass filter. That is,

$$V_{m,n} = 2^{-m/2}\ \left(v(t) * \psi(-2^{-m}t)\right)\big|_{t=2^m n}\,.$$

For each scale m, the variance of the coefficients $V_{m,n}$ reflects the amount of energy in the signal $v(t)$ that lies within the bandwidth of the bandpass filter. In Figure 3.18 we plot the frequency response $|\Psi(f)|$ of the C_2 wavelet for scales $m = 2$ to 11. The responses are equally spaced along the logarithmic frequency axis. As a result, a plot of the logarithm of the wavelet sample variance versus scale has the same form as a logarithmic plot of power spectrum versus frequency. In both cases, the analysis indicates that the geomagnetic background is piecewise $1/f$.

Referring to Figure 3.17, we note that the transition in slope occurs at $m \approx 7$ which corresponds roughly to a frequency of 0.2 Hz — the same value as obtained for the spectra shown in Figure 3.16(c). The variance lacks a steep transition zone as was observed in Figure 3.16(c). The value of $\gamma = 3.6$, for scales 7 to 11, however, is very close to the arithmetic mean of the gamma values 2.5 and 4.5 observed in Figure 3.16(c) for frequencies below 0.2 Hz. This close agreement may indicate that the finite bandwidth of the wavelet frequency responses tends to smooth out the steep transition zone. Overall, the wavelet variance plots are in good agreement with the power spectral plots and further indicate the presence of two primary sources in the ULF band.

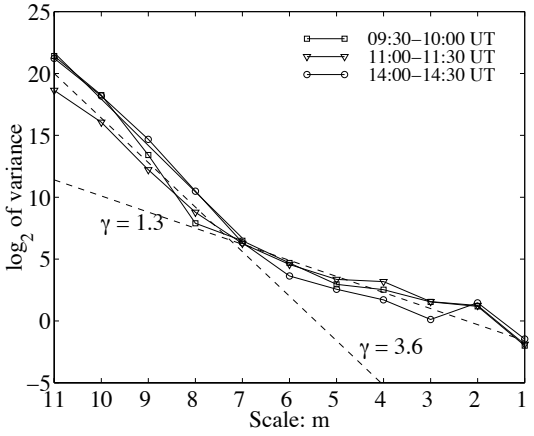


Figure 3.17: Variance of DWT coefficients for data from E-W channel at SAO.

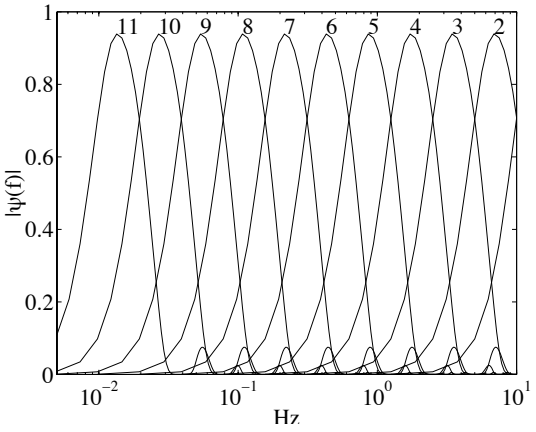


Figure 3.18: Frequency response of the C_2 wavelet for scales $m = 2$ to 11

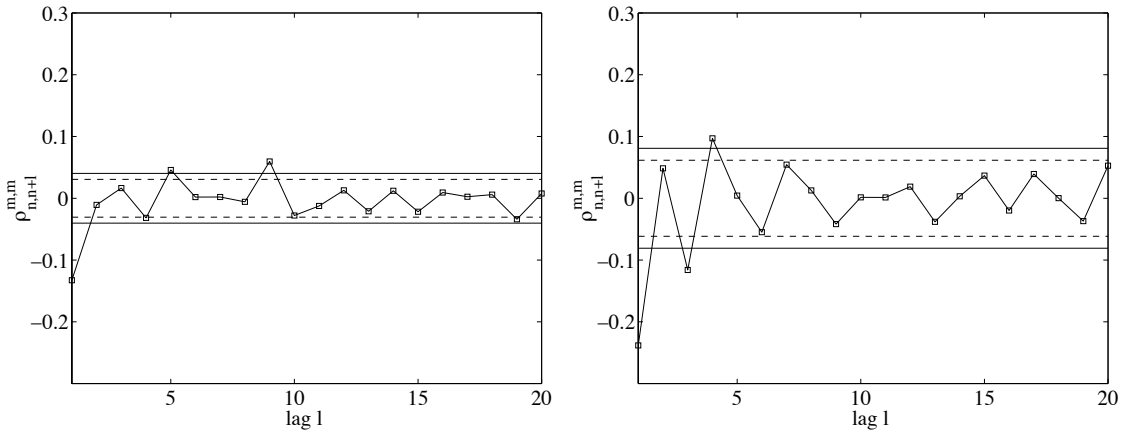
Correlation of Wavelet Coefficients

To determine the extent to which the DWT acts as an approximate whitening transform for geomagnetic noise, we computed the correlation of the resultant DWT coefficients. The general expression for the correlation coefficient is

$$\rho_{n,n'}^{m,m'} = \frac{E\left[V_{m,n}V_{m',n'}\right]}{\sqrt{\mathrm{var}(V_{m,n})\mathrm{var}(V_{m',n'})}} \tag{3.7}$$

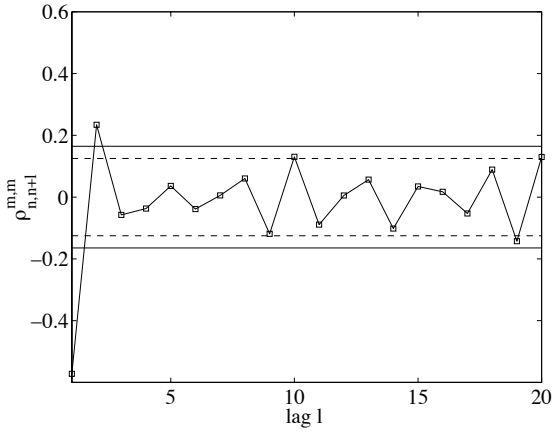
where $\mathrm{var}(V)$ denotes the variance of V. For the case of autocorrelation, we have $m = m'$ and $n' = n + l$, where l denotes the lag. In Figure 3.19 we present sample auto-correlation plots for scales $m = 4, 6, 8,$ and 10. To assess the significance of the correlation values, we have superimposed lines corresponding to 1% and 5% significance levels [*Press et al.*, 1988]. The probability of a correlation value exceeding the 1% or 5% significance levels when the underlying waveforms are uncorrelated is less than 0.01 or 0.05, respectively. Note that due to the dyadic sampling grid inherent to the DWT, the number of coefficients used to compute the correlations decreases with increasing scale value m. As a result, the significance levels increase with increasing m. At all scales, there is significant correlation at lag $l = 1$, but correlations at greater lag values are, for the most part, not significant.

There are many ways to examine the correlation between wavelets across analysis scales. *Wornell* [1996], for example, derived a theoretical expression for the case where $2^{-m}n = 2^{-m'}n'$ and showed that the result was a function only of $m - m'$. *Johnstone and Silverman* [1996] considered the cross-correlation between the coefficients at scale m with the upsampled (by a factor of 2) coefficients at scale $m + 1$. The upsampling of the scale $m + 1$ coefficients is required to compensate for the dyadic sampling of the DWT. This definition of cross-correlation corresponds to the choice $m' = m - 1$ and $n' = 2n + l$, i.e., $\rho_{n,2n+l}^{m,m-1}$. In Figure 3.20 we present sample cross-correlation ($\rho_{n,2n+l}^{m,m-1}$) plots for scales $m = 4, 6, 8,$ and 10. The 1% and 5% significance levels are also shown [*Press et al.*, 1988]. For all the scales shown, there are significant correlation values around lag $l = 5$. These high correlation values are due primarily to the presence of wideband signals, such as impulses due to lightning, which contribute energy to more than one DWT analysis scale. Aside from this region of high correlation, the values at other lag values are, for the most part, not significant.

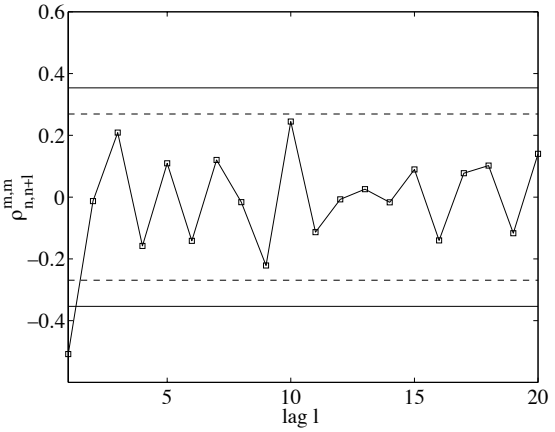


(a) $m = 4$

(b) $m = 6$

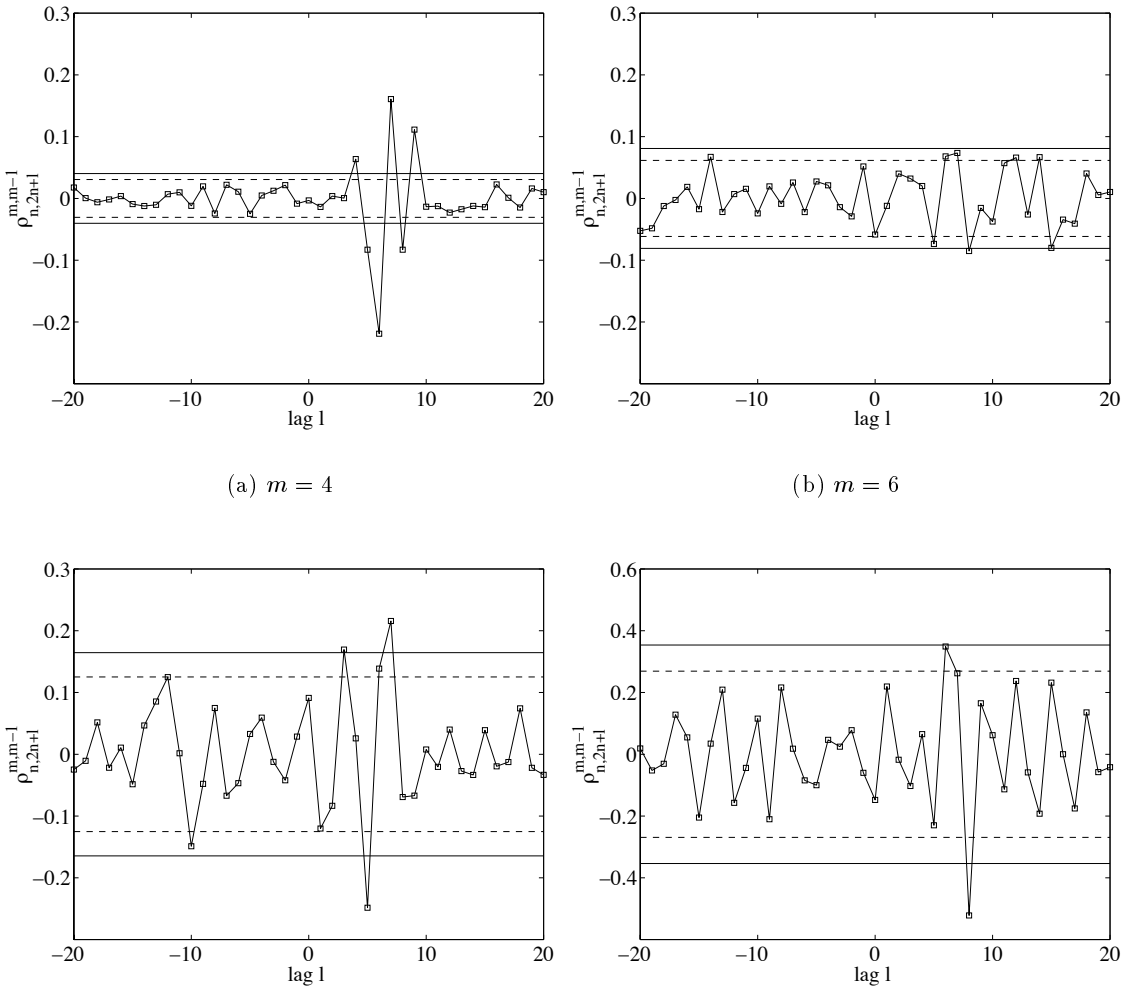


(c) $m = 8$



(d) $m = 10$

Figure 3.19: Autocorrelation of DWT coefficients for data from E-W channel at SAO for the period 11:00 to 11:30 UT on August 11, 1998. Solid and dashed lines represent 1% and 5% significance levels, respectively.



(a) $m = 4$

(b) $m = 6$

(c) $m = 8$

(d) $m = 10$

Figure 3.20: Cross-correlation $\rho^{m,m-1}_{n,2n+l}$ of DWT coefficients for data from E-W channel at SAO for the period 11:00 to 11:30 UT on August 11, 1998. Solid and dashed lines represent 1% and 5% significance levels, respectively.

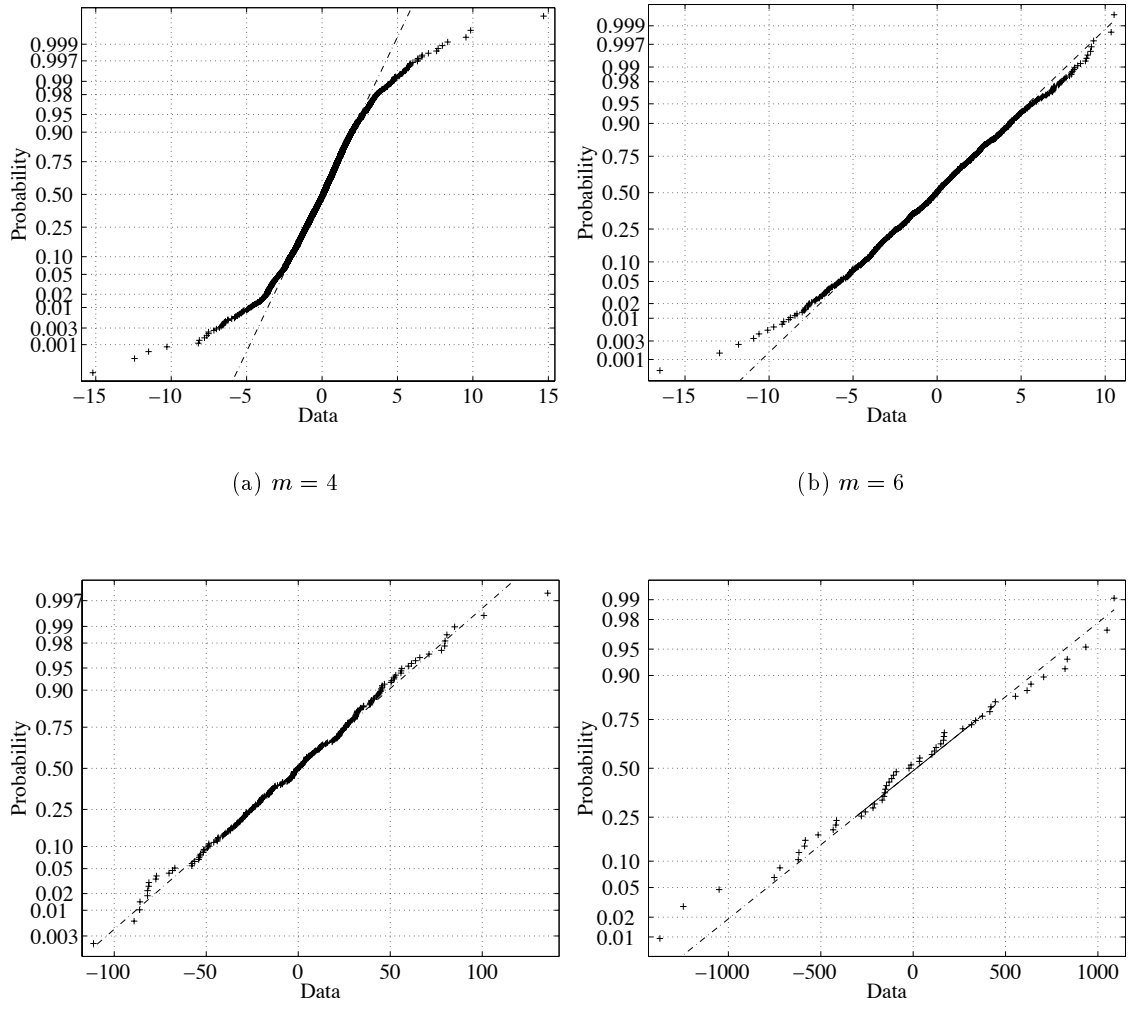
Distribution of Wavelet Coefficients

In this section we examine the probability distribution of the geomagnetic background DWT coefficients. We are particularly interested in determining if the DWT coefficients are normally distributed, since the assumption of a normal distribution simplifies the development of a wide array of signal processing tasks. In specific, as we will show in Chapter 6, the assumption of a normal distribution, in conjunction with the whitening transform property of the DWT, permits us to develop a matched filter detection scheme for transients in the presence of the geomagnetic background.

To assess the degree to which the data follow a normal distribution, we computed normal probability plots [*Rice*, 1995] for the DWT coefficients at analysis scales $m = 4, 6, 8,$ and 10. These are shown in Figure 3.21. At scale $m = 4$ the coefficients clearly do not come from a normal distribution, while at scales $m = 8$ and $m = 10$ the assumption of normality is more reasonable. The normal probability plot at $m = 4$ indicates that distribution has long tails. These are due to the presence of lightning induced sferics which are known to give rise to long-tailed distributions [*Chrissan*, 1998]. At the coarser analysis scales, the energy due to the sferics is small compared to energy due to magnetospheric sources. In the section on the wavelet coefficient variances, we noted that the scale $m = 7$ serves as the transition point between lightning induced and magnetospheric sources. Here we note that this scale also serves as a rough dividing point between non-normal and normal behavior of the coefficients. That is, for $m \geq 7$ the data appear to be normally distributed, while for $m < 6$ the data are not normally distributed and the deviation from normality increases with decreasing scale. It is important to point out that we have not shown that the DWT coefficients for $m \geq 7$ are characterized by a joint normal distribution, but only that the distribution of coefficients at each of these scales is consistent with a normal distribution.

3.7 Summary

In this chapter we have examined in detail the properties of the BART interference as observed by an array of multi-component sensors in the SFBay. At each channel, the interference can be approximated as the superposition of transients with typical durations of 4 to 30 seconds. In addition, the polarization at each site exhibits a strong directional



(a) $m = 4$

(b) $m = 6$

(c) $m = 8$

(d) $m = 10$

Figure 3.21: Normal probability plots of DWT coefficients for data from E-W channel at SAO during the period 11:00 to 11:30 UT on August 11, 1998.

component. In Chapter 7, we take advantage of these characteristics to develop a method to distinguish the interference from the natural background activity.

We also considered the $1/f$ behavior of the geomagnetic background. In agreement with previous observations, we showed that the geomagnetic background power spectra are piecewise $1/f$. In addition, we demonstrated that the DWT acts as an approximate whitening transform for the geomagnetic background, and that the DWT coefficients appear to be normally distributed for scales $m \geq 7$. These properties will be used in Chapters 6 and 7 to develop a method for the detection of transients in the presence of the geomagnetic background.

Chapter 4

BART Magnetic Field Models

4.1 Introduction

In this chapter we model the magnetic fields produced by the BART system. We begin in Section 4.2 with a review of the design and operating characteristics of the BART system. We find that significant transients in the train propulsion currents have time scales similar to those observed for the magnetic field transients. In Section 4.3 we present a simplified electric circuit model of the BART system and use that model to calculate the distribution of currents along the rails and in the earth. In Section 4.4 we derive formulas for the static and quasistatic magnetic fields generated by these currents. For the quasistatic case we show that the conducting earth acts as a lowpass filter of the magnetic fields. In Section 4.5 we compute polarization plots for both the static and quasistatic cases. The polarization exhibits a strong directional component, in agreement with the experimental findings.

4.2 Description of the BART System

4.2.1 Overview

Since its inception in 1972, the BART system has served over one billion riders in the San Francisco Bay Area. At present the BART system has approximately 150 km of double track with 39 passenger substations. In addition, there are four maintenance yards. A map of the BART system is shown in Figure 4.1. For reference purposes, we divide the system into five physical lines: A-line, C-line, L-line, M-line, and R-line. The A-line, the M-line and the R-line join together at a location just west of the Lake Merritt station.

From there the A-line extends to the southeast and ends at the Fremont Station. The M-line extends to the southwest and ends at the Colma Station. The R-line extends to the north and ends at the Richmond Yard. The C-line branches off the R-line just north of the MacArthur station and extends to the northeast, ending at the Pittsburg station. The L-line branches off the A-line just to the south of the Bay Fair station and extends to the east, ending at the Dublin station.

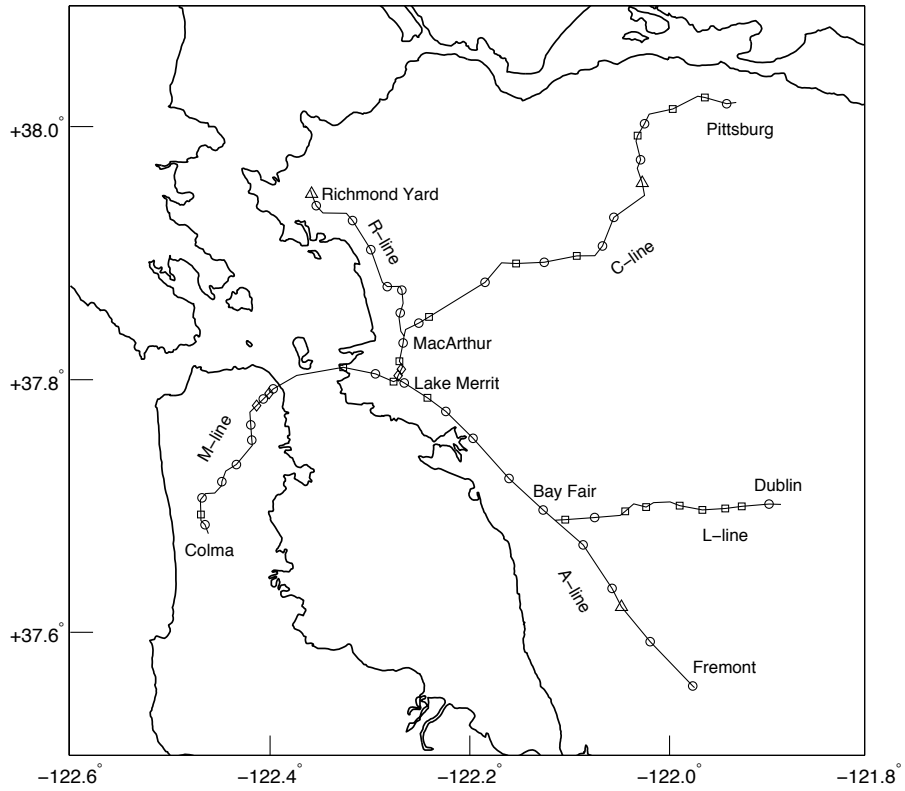


Figure 4.1: Map of the BART system. Circles: Passenger stations with a nearby substation. Diamonds: Passenger stations without a nearby substation. Squares: Substations not located near a passenger station. Triangles: Maintenance yard with a substation.

The BART system uses a 1000 volt DC third rail, located adjacent to the running rails, to provide traction power to its trains. The third rail is, for the most part, electrically continuous over the span of the system. It is fed by 60 power substations. Each substation contains either one or two transformer-rectifier units, with the majority of the stations containing two units operating in parallel [*BART Engineering Dept.*, 1996]. Depending

on the load requirements at the substation, the capacity of the transformer-rectifier units is 3, 4 or 5 MW. As shown in Figure 4.1, the majority of these substations are located in the vicinity of a passenger station or maintenance yard.

A BART train may draw up to 10,000 amperes of current under conditions of maximum acceleration. A train operating at constant velocity draws about 2000 amperes. By design, the current returns to the substations primarily through the running rails. However, due to the finite resistance between the rails and the earth, some current does return through the earth and conducting objects within the earth, e.g., metal cables and pipes, and is referred to as *leakage current*. Figure 4.2 shows a schematic view of the current flowing to the train along the third rail and returning to the substations along the running rails and through the earth.

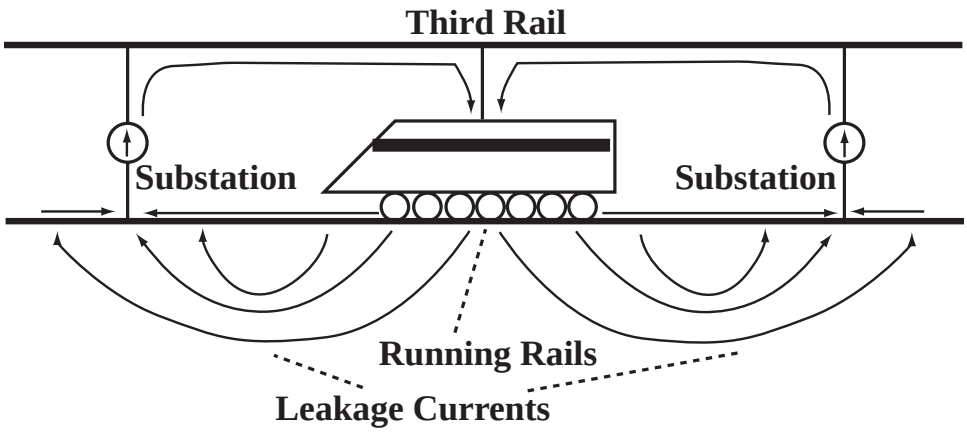


Figure 4.2: A simplified view of the flow of propulsion current along the third rail and return currents along the running rails and through the earth. For display purposes, the third rail is shown above the train. In fact, it is located adjacent to the running rails, as shown in Figure 4.7.

Train deceleration is accomplished primarily through a combination of regenerative and dynamic braking. In both cases the propulsion motors are re-configured to act as generators. In the regenerative braking mode, the train returns power to the third rail. With the present control system, about 20 percent of the traction power can be recovered through regenerative braking [*Evans*, 1993]. This mode can be used only if another train is present to accept the power. Field tests have shown that power may be supplied to another train as far as 10 km away [*Evans*, 1996]. It is important to note that power

does not return to the substations during the deceleration period. Power that cannot be returned to the third rail is dissipated by onboard dynamic braking resistor grids. For speeds below four miles per hour, friction braking augments the other braking modes.

4.2.2 Train acceleration and deceleration patterns

To first order, train operation may be divided into the following four operating modes:

Maximum Acceleration

For speeds less than a base speed of 36 MPH, the train undergoes a constant maximum acceleration of $a = 3$ MPH/s. This period typically lasts for about 11 seconds. The tractive effort F required to move the train is $F = M(a + g\sin\alpha) + R$ where M is the mass of the train, R is the rolling resistance and air resistance, g is the gravitational constant and α is the angle of the grade. To first order, the required power is then $P(t) = F \cdot v(t) = M(a + g\sin\alpha)at$, where we have assumed that the train velocity is $v(t) = at$ and that $R \ll Ma$. Assuming that the third rail voltage V_3 is relatively constant during this period, the third rail current is

$$I_3 = \frac{P(t)}{V_3} = \left[\frac{aM(a + g\sin\alpha)}{V_3}\right] t$$

Thus, the third rail current increases linearly with time in this mode.

Constant Power

To achieve speeds higher than the base speed of 36 MPH, the train switches into a constant power mode as it continues to accelerate up to a maximum speed of 80 MPH. In this mode, the tractive effort and acceleration decrease with time. The third rail current I_3 is fairly constant in this mode.

Constant Velocity

Once the train has reached its maximum speed, it maintains a constant velocity. The required tractive effort and third rail current are nominally constant in this mode. They will increase or decrease to accommodate changes in track grade, train rolling resistance, and air resistance.

Deceleration

As described previously, the primary means of deceleration are regenerative braking and dynamic braking. Simulations show that a train may inject up to 7500 amperes of current onto the third rail when regenerative braking is in effect [*Evans*, 1996].

Figure 4.3a presents the results of a simulation [*Evans*, 1996] in which a train accelerates to maximum speed and then decelerates. The train velocity and third rail current are plotted as a function of time. The discontinuous behavior of the third rail current in the deceleration mode is due to the switching in of dynamic braking resistors.

The test case shown in Figure 4.3a is useful for observing the various train operating modes but does not reflect typical train operations. Throughout its course of operation, each train must follow maximum speed limits that are set by the centralized control system. During acceleration, for example, a train may accelerate up to a speed of 48 MPH, before maintaining that speed until it reaches a zone with a higher speed limit, at which point it accelerates to its maximum speed. A train velocity and third rail current profile for a train traveling between two stations is shown in Figure 4.3b. Note the plateaus in the velocity profile due to imposed speed limits.

In the profile shown in Figure 4.3b, we observe that the largest currents are associated with transients that have durations ranging from approximately 5 to 30 seconds. These transients occur primarily when the train is accelerating from a passenger station or decelerating before arriving at a station. Between stations transients can occur when the train encounters an increase in grade. The range of durations for the current transients is approximately the same as the range observed for the magnetic field transients presented in Chapter 3. As a result, the current profile lends further support to our characterization of the BART interference as the superposition of transients with a range of characteristic time scales.

Finally, we should note that the velocity and current profiles in Figure 4.3b are only typical profiles that reflect a train following a nominal schedule. Actual profiles are a complicated function of the run-time schedule imposed by the central control system. Deviations from the ideal schedule, caused by signaling system errors, equipment malfunction, or safety concerns, will determine the effective velocity-time profile of each train [*Stoilov et al.*, 1994]. For example, a problem occurring in one train will affect the profiles of all trains that follow.

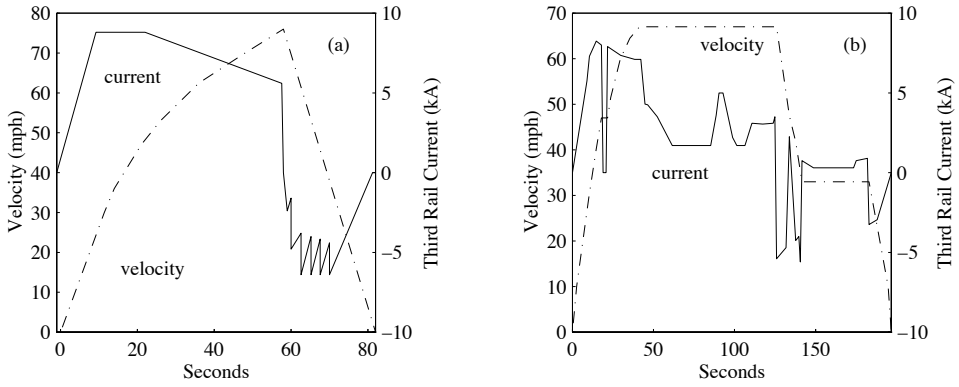


Figure 4.3: BART current and velocity profiles. Panel (a) shows the profiles for a train accelerating and then decelerating. Panel (b) shows the profiles for a train moving from one station to the next.

4.3 Electrical Model of the BART system

4.3.1 Distributed Parameter Model

The effect of direct current railways on electromagnetic measurements was noted as early as 1890 [*Burbank*, 1905] and has been studied by various researchers over the course of the past century (see [*Szarka*, 1988] for a comprehensive review). Several researchers have demonstrated that leakage current effects must be taken into account when modeling the magnetic disturbances due to electric railways [*La Cour and Hoge*, 1937; *DuPouy*, 1950; *Yanagihara*, 1977]. As mentioned above, leakage currents are those currents which do not return to the substation via the running rails.

We can model the effect of the leakage currents on the rail current distribution with a distributed resistance to ground [*Riordan*, 1932], as shown in Figure 4.4. The rail voltage $V(x)$ and current $I(x)$ satisfy the equations:

$$\eta \frac{dI(x)}{dx} = -V(x) \tag{4.1}$$

$$\frac{dV(x)}{dx} = -RI(x) \tag{4.2}$$

where R is the resistance per unit length (with units of Ω/km) of the running rail and η

is the leakage resistivity (with units of Ω-km). Combining these two equations yields

$$\frac{d^2I(x)}{dx^2} = \gamma^2 I(x) \tag{4.3}$$

where $\gamma = \sqrt{\frac{R}{\eta}}$. Solutions to (4.3) are of the form:

$$I(x) = Ae^{-\gamma x} + Be^{\gamma x} \tag{4.4}$$

where A and B are constants that are determined by the boundary conditions. Initially, the use of this model was justified by its good fit to experimental measurements [*Riordan*, 1932]. *Sunde* [1936] later demonstrated that it was also a good theoretical approximation for the currents of a long, straight, partially insulated wire lying on the earth's surface.

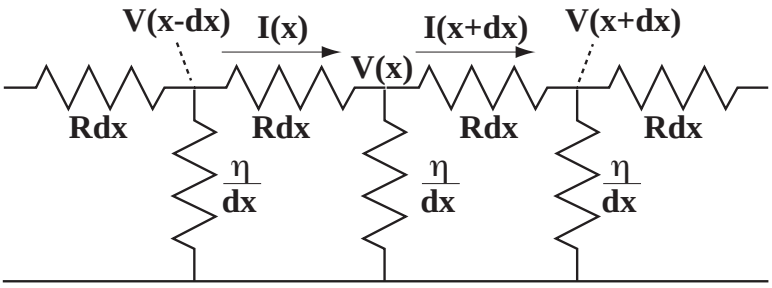


Figure 4.4: Distributed parameter model for the running rail. Effect of leakage resistance to earth is modeled by shunt resistances with value η/dx

The voltages and currents in a railway system may be determined by using solutions of the form given by (4.4) with the appropriate boundary conditions [*DuPouy*, 1950; *Yanagihara*, 1977]. For large railway circuits the unknown coefficients are readily obtained by solving a system of linear equations with a program such as MATLAB. This approach becomes cumbersome, however, when we wish to model a modern railway circuit containing multiple branches and an arbitrary number of loads. In addition, the approach is not well-suited to calculating the effect of non-linear circuit elements, such as grounding diodes. Many modern railway systems, such as BART, use grounding diodes at the substations to both limit the negative voltage excursion on the running rail and reduce the amount of stray current that flows between substations [*Bomar et al.*, 1974].

4.3.2 Lumped-Element Circuit Model

We have adopted a more flexible approach to the determination of the system currents. Instead of a distributed parameter model, we form a lumped-element circuit model of the BART system and use a circuit simulation program (HSPICE) to solve for the voltages and currents. Lumped-element circuit models have been used for stray current studies [*Shaffer*, 1982] and for general purpose DC railway traction power simulations [*Cai et al.*, 1995]. The approach entails dividing the running rail into a series of finite-length sections, where each section has a series resistor to represent the running rail resistance and a shunt resistor to represent the leakage resistance. *Cai et al.* [1995] state that for most practical purposes a section length of 0.1 km provides sufficient accuracy. We have verified their conclusion by comparing results obtained with the distributed parameter model to those obtained with the lumped-element model. Unless otherwise stated, we use 0.1 km for the section length in our calculations. A schematic showing a typical portion of the model is provided in Figure 4.5.

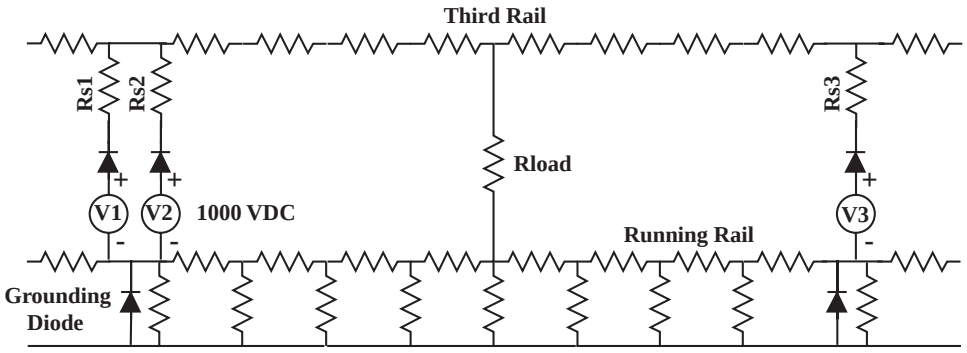


Figure 4.5: Schematic showing a portion of the lumped-element model for the BART electrical system.

At each substation location, the transformer-rectifier units are modeled by voltage sources in series with a source resistance and a diode. We use a 1000 volt source since this is the nominal open circuit voltage of the rectifier units. The presence of the diode reflects the fact that the rectifier units are not receptive, i.e., they cannot draw current from the third rail. The source resistance models the load characteristic of the rectifier unit in its normal operating mode [*Flowers*, 1995]. The rectifier units are nominally rated at 6% voltage regulation [*Evans*, 1993], which corresponds to a source resistance of 0.06 Ω. In

practice, we have found that a value of 0.04 Ω better models the current sharing between substations. The BART system is designed so that about 30% of an accelerating train's power is supplied by the two substations on either side of the substation closest to the train [*Bomar et al.*, 1974; *Todd*, 1997]. The train is modeled as a resistive load with a value of 0.1 Ω representing the approximate load for peak power consumption. Higher resistance values can be used to model the train for operating modes with lower power levels. Finally, grounding diodes are placed at each substation except for those on the L-line [*Todd*, 1997].

Except for complicated junctions, the BART system is a double track system with four running rails and two third rails. To minimize the return resistance, the four running rails are electrically connected with crossbonds. The parallel resistance of the rails is about 8.2 mΩ/km [*Bomar et al.*, 1974]. For the majority of the BART system, the resistance of the third rail is typically 11.8 mΩ/km. There are some sections with third rail resistances of 6.6 mΩ/km. These are: the entire L-line, the C-line from Rockridge to Lafayette and also from Concord to Pittsburg, and the M-line from West Oakland to Powell and also south of Daly City [*Todd*, 1997].

Leakage resistivity values vary greatly with both the weather and track location. The BART system contains tracks that are at-grade (i.e., directly on the earth's surface), tracks on aerial structures, and tracks in subways and tunnels. At-grade tracks consist of steel rails mounted on concrete ties with bolts and insulating pads. Timber ties are used at switches and through seismic zones. Aerial and subway tracks are mounted to supporting concrete structures with special insulating fasteners [*Bomar et al.*, 1974]. The leakage resistivity depends highly on the condition of the insulating pads and fasteners, with moisture and corrosion leading to decreased resistivity values. For the BART system, *Galler and Todd* [1992] report measured resistivity values ranging from 4 to 4000 Ω-km with most values in the middle of that range. By comparison, *Shaffer and Fitzgerald* [1981] report values for the Washington Metro of 1.5 to 54 Ω-km (mean value of 7.5 Ω-km) for timber ties and 7.2 to 3000 Ω-km (mean value of 200 Ω-km) for aerial and subway structures. *Shaffer* [1987] states that most modern fasteners should provide resistivity values of about 150 Ω-km under relatively dry and clean conditions, with degraded performance for fasteners covered by moisture or dirt films. *Hill* [1996] notes that established practical values range from 2.5 to 500 Ω-km. Finally, *Yu and Goodman* [1992] used values of 5 to 10 Ω-km for timber ties and 5 to 100 Ω-km for concrete ties.

We calculated current distributions for system-wide leakage resistivity values of 0.5, 5 and 50 Ω-km. While our model can accommodate resistivity values that vary across the system, we do not incorporate this degree of freedom in the examples presented below. Figure 4.6 shows an example of the results obtained by the model with a resistivity value of 5 Ω-km.

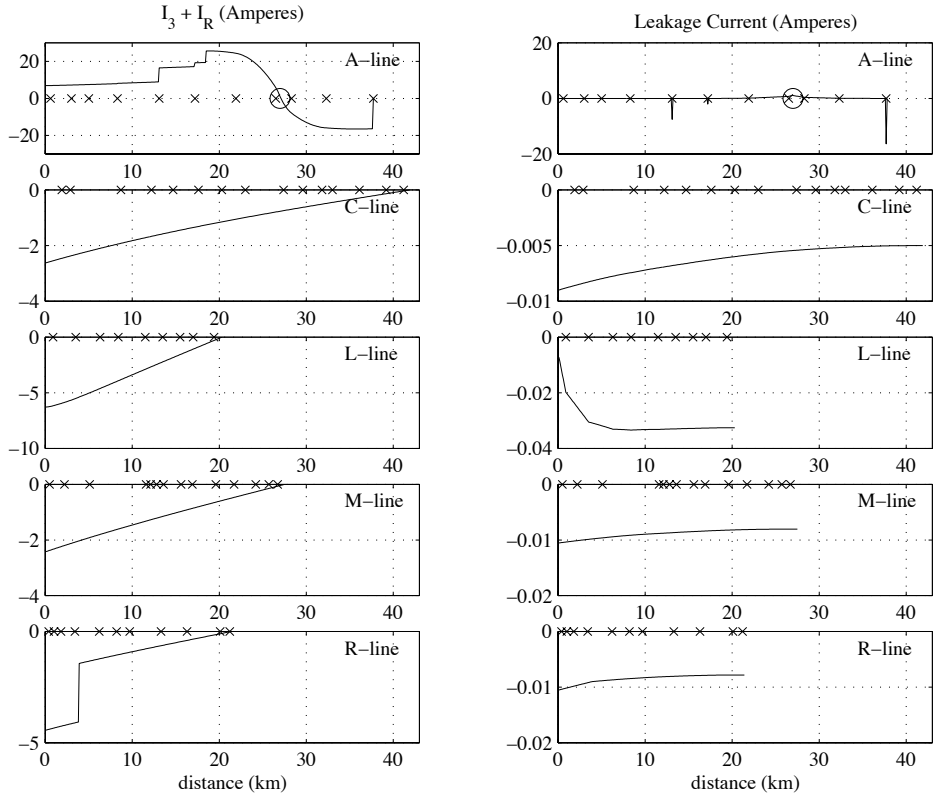


Figure 4.6: Rail current distributions obtained with the lumped-element circuit model. The left column shows the sum of the third rail and running rail currents along the 5 lines. The right column shows the leakage current. The station locations are marked with 'x' symbols. The train is on the A-line and its location is marked with a circle. The train load and leakage resistivity are assumed to be 0.1 Ω and 5 Ω-km, respectively.

4.4 Magnetic Field Expressions

The lumped-element circuit model generates values for leakage currents and rail currents. The effect of the leakage currents may be modeled by point electrodes injecting current into the ground, while the rail current for each track section is simply a current flowing along a finite length wire. In this section we derive formulas for the magnetic fields due to point electrodes and finite length wires. The total magnetic field due to any specified current distribution may then be found by superposing the solutions from all the sections of the model.

4.4.1 The Magnetostatic Case

DuPouy [1950] and *Yanagihara* [1977] calculated magnetostatic field amplitudes using analytical expressions for the current distributions and found that the predicted fields agreed well with observed amplitudes. In addition DuPouy calculated the slowly varying fields due to a train moving from one substation to the next and was able to obtain a good match to the measured fields as a train moved between two substations. We begin our treatment of the BART fields by following DuPouy and Yanagihara's approach.

Fields due to Rail Currents

In the magnetostatic case, the magnetic fields are due solely to primary currents flowing in the rails and in the earth. There are no induction or displacement currents. The fields due to currents in the third and running rails are calculated using the Biot-Savart Law

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{V'} \frac{\mathbf{J}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} dV' \tag{4.5}$$

where $\mathbf{J}(r')$ is the source current density and $\mathbf{r}$ and $\mathbf{r}'$ are the observation and source vectors, respectively. Recall that the lumped-element circuit model divides the system into a finite number of rail sections. Each section may be considered to be a *current stick*, i.e. a section of current I extending from a start point $\mathbf{r}'_0$ to an end point $\mathbf{r}'_1$. Application of the Biot-Savart law to a current stick [*Haus and Melcher*, 1989] yields

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0 I}{4\pi} \frac{\vec{c} \times \vec{a}}{|\vec{c} \times \vec{a}|^2} \left(\frac{\vec{a} \cdot \vec{c}}{|\vec{c}|} - \frac{\vec{a} \cdot \vec{b}}{|\vec{b}|} \right), \tag{4.6}$$

where $\vec{a} = \vec{c} - \vec{b}$, $\vec{b} = \mathbf{r}'_0 - \mathbf{r}$, and $\vec{c} = \mathbf{r}'_1 - \mathbf{r}$.

We estimate the position vectors for the rails in the following manner. For each station in the system we obtain its geographic coordinates from a digital topographic mapping program (TOPO! version 1.2.1, Wildflower Productions, San Francisco, CA). We then use the mapping program to approximate the tracks with piecewise linear sections. For example, if there is a 90 degree bend in the tracks between two stations, we approximate the track as two linear sections that join near the location of the bend. The approximation is *ad hoc*, i.e., we use the mapping program to choose linear sections which provide a good visual fit to the actual track layout. We record the coordinates of the start and stop points of each linear section for use in the current stick calculations.

When calculating the magnetic field, we also need to take into account the relative spacing of the various running rails and third rails. Recall that the BART track system consists of four running rails and two third rails, except at complicated junctions. The BART system employs a track gauge of 1.72 meters between running rails. The distance between pairs of running rails is about 2.55 meters [*Todd*, 1997]. Exceptions to this case occur in the Transbay and Berkeley Hills tunnels. For example, the distance between running rail pairs in the Transbay tunnel is about 7 meters [*Friedlander*, 1972]. The two third rails lie between the running rail pairs for at-grade tracks and to the outside of the running rail pairs for aerial tracks [*Todd*, 1997]. A diagram of the typical at-grade rail spacing is shown in Figure 4.7. In our calculations we use the at-grade rail spacing for the entire system. We also assume that the running rail current is equally divided among the four running rails.

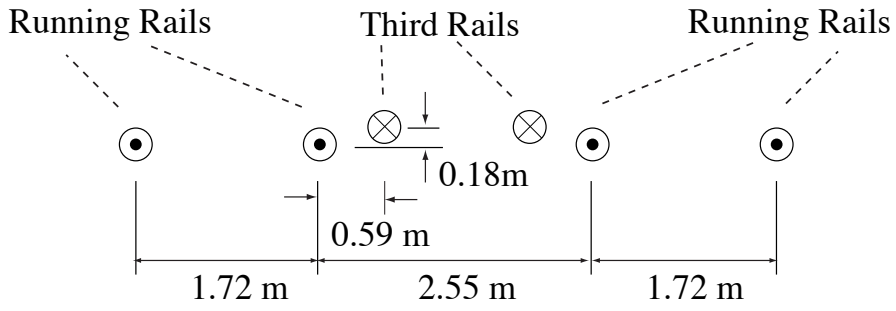


Figure 4.7: Rail spacing for at-grade track sections.

In the absence of leakage currents the rail spacing determines the magnitude of the observed magnetic fields. For typical leakage current magnitudes, however, the rail spacing becomes less important, and the magnetic field is well-approximated by the field due to a current I_{sum} flowing along just one of the rails (e.g., the third rail), where I_{sum} is the sum of the third rail and running rail currents. In our calculations we found that the difference between the two methods was negligible for system-wide leakage resistivity values of 0.5, 5 and 50 Ω-km.

Fields due to Leakage Currents

To calculate the fields due to leakage currents we first model each contact between the rail and the earth with a point electrode injecting a current $I_{i,L}$ into the earth, where $I_{i,L}$ is the current flowing in the ith shunt resistor of our lumped-element circuit model. Consistent with the circuit model, we use a finite number of contact points to approximate the actual situation where there are infinitely many contacts between the rails and the earth. While a continuous analysis is possible [*DuPouy*, 1950], it is impractical for a complex system such as BART.

Next we assume a horizontally stratified conducting earth. That is, we approximate the true three-dimensional conductivity profile $\sigma(x, y, z)$ with a one-dimensional conductivity profile $\sigma(z)$. By assuming a one-dimensional profile, we can obtain convenient analytical expressions for the magnetic fields in both the static and quasistatic cases.

With the use of numerical modeling techniques (see, for example, [*Hohmann*, 1987]) it is possible to calculate the fields assuming a three-dimensional profile, but there are several factors that place such an effort beyond the scope of this thesis. First, a map of the three-dimensional conductivity of the San Francisco Bay region currently does not exist. The generation of an accurate map represents a significant effort in its own right. In Appendix A we present an initial attempt at characterizing the conductivity of the region. We argue there that the significant features of the region can be described with a two-dimensional profile $\sigma(x, z)$. Even with a two-dimensional profile, however, the required numerical model will need to be three-dimensional because the currents in the BART system are functions of both horizontal coordinates (x and y).

It is unlikely that the advantages to be gained with a more detailed conductivity model, either two or three dimensional, would be significant given the large uncertainties already present in modeling the distribution of currents in the BART system. For example, at

present, it is only possible to make first-order estimates of the leakage resistivities in the system. Finally, if we assume a three-dimensional model of the San Francisco Bay region with a minimum grid size of 0.1 km per dimension and an overall extent of 100 km per dimension, the resulting model would contain one billion grid points. At this point we do not believe that the potential gains of such a model warrant the significant computational complexity and effort.

With the assumption of a stratified earth, it can be shown that the above ground static magnetic fields due to a point electrode injecting a current I into the earth are the same as those due to a semi-infinite line current I extending from the surface of the earth to positive infinity [*Kaufman and Keller*, 1983]. From the Biot-Savart law applied to a semi-infinite line current we obtain

$$\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{\hat{z} \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^2} \left(1 + \frac{\hat{z} \cdot (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\right) \tag{4.7}$$

where $\hat{z}$ points into the earth and the $\hat{z}$ components of $\mathbf{r}$ and $\mathbf{r}'$ are assumed to be negative and zero, respectively.

4.4.2 Fields above a Stratified Earth: Quasistatic Case

In the static case, the magnetic fields are independent of the one-dimensional conductivity profile, as can be easily verified by examining Equation 4.7. The conductivity of the earth comes into play, however, as we consider time-varying fields. In this section we present expressions for the propagation of ULF magnetic fields over a stratified earth.

We begin by presenting the formulas for a horizontal electric dipole (HED) located at the the surface of a horizontally stratified conducting earth. We then use the HED formulas to obtain expressions for the fields due to a point current electrode and a finite length wire. The basic theory and notation used in this section are reviewed in Appendix B. We use the one-dimensional conductivity model discussed in Appendix A and shown in Figure 4.8. The use of this model allows us to make a number of approximations.

For a HED at the surface of the earth, we consider the fields in region 0+ and evaluate for d_0 and z equal to zero. Since there are no boundaries for $z < d_0$, both $R^{TM}_{0,-1}$ and $R^{TE}_{0,-1}$ are zero. From eqs. (B.46) through (B.49), we obtain

$$A_{0+} = E_{hed}, \quad B_{0+} = R^{TM}_{0,1} E_{hed}, \quad C_{0+} = H_{hed}, \quad D_{0+} = R^{TE}_{0,1} H_{hed}. \tag{4.8}$$

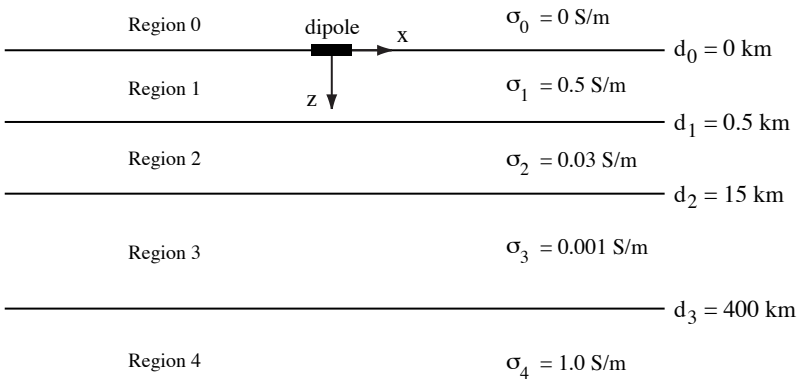


Figure 4.8: Conductivity profile for calculation of quasistatic fields.

Substitution of these values into eqs. (B.25), (B.28), and (B.29) yields the desired expressions for H_z, H_r, and H_ϕ (with $z = 0$, $d_0 = 0$)

$$\begin{aligned} H_{0+,r} &= \frac{Ids\sin\phi}{4\pi r}\left[\int_0^\infty \left[\left(1 - R_{0,1}^{TE}\right) - \left(1 + R_{0,1}^{TM}\right)\right] J_1(k_\rho r)dk_\rho \right. \\ &\quad \left. - \int_0^\infty k_\rho r\left(1 - R_{0,1}^{TE}\right) J_0(k_\rho r)dk_\rho\right] \qquad (4.9) \\ H_{0+,\phi} &= \frac{Ids\cos\phi}{4\pi r}\left[\int_0^\infty \left[\left(1 + R_{0,1}^{TM}\right) - \left(1 - R_{0,1}^{TE}\right)\right] J_1(k_\rho r)dk_\rho \right. \\ &\quad \left. - \int_0^\infty \left(1 + R_{0,1}^{TM}\right) k_\rho r J_0(k_\rho r)dk_\rho\right] \qquad (4.10) \\ H_{0+,z} &= \frac{Ids\sin\phi}{4\pi}\int_0^\infty \frac{k_\rho^2}{j\,k_{0,z}}\left(1 + R_{0,1}^{TE}\right) J_1(k_\rho r)dk_\rho \qquad (4.11) \end{aligned}$$

Quasistatic Approximation

In the quasistatic approximation we assume that $k_\rho \gg k_0$ where $k_0 = \omega\sqrt{\mu_0\epsilon_0}$ is the free space wave number [*Wait*, 1982]. The quasistatic approximation is equivalent to ignoring the effects of displacement currents in the region above the earth. To understand the approximation we note that the fields due to an HED are computed using integrals that include Bessel functions with the argument $k_\rho r$. The quasistatic approximation is valid when the value of an integral with limits of integration 0 to ∞ is approximately equal to the same integral with limits of $5k_0$ to ∞. The lower limit $5k_0$ is chosen to satisfy the quasistatic approximation. The approximation is valid if

$$\begin{aligned} H_{0+,z} &= \frac{Ids\sin\phi}{4\pi}\int_0^\infty \frac{k_\rho^2}{jk_{0,z}}\left(1+R_{0,1}^{TE}\right)J_1(k_\rho r)dk_\rho \\ &\approx \frac{Ids\sin\phi}{4\pi}\int_{5k_0}^\infty \frac{k_\rho^2}{jk_{0,z}}\left(1+R_{0,1}^{TE}\right)J_1(k_\rho r)dk_\rho \end{aligned} \tag{4.12}$$

Since Bessel functions decay slowly as a function of their argument, the approximation holds if $5k_0 r \ll 1$ or equivalently $r \ll \lambda_0$ where $\lambda_0 = 2\pi/k_0$ is the free space wavelength. In our case the conditions for the quasistatic approximation are easily met — λ_0 is 3×10^4 km for $f = 10$ Hz, while the largest distance of interest is less than 100 km. With the quasistatic approximation, we may assume that $k_{0,z} \approx -jk_\rho$, where $k_{0,z}$ is the free space longitudinal wave number (see Equation B.22).

We also make use of the observation that, for the conductivities and frequencies of interest, the condition $\sigma_i \gg \epsilon_i\omega$ is satisfied where we assume that $\epsilon_i = \epsilon_0$. In other words, the conduction current is much greater than the displacement current in the conducting earth.

Using the quasistatic approximation we may write

$$R_{0,1}^{TE} \approx \frac{Z_1^{TE}(0) - \frac{j\omega\mu_0}{k_\rho}}{Z_1^{TE}(0) + \frac{j\omega\mu_0}{k_\rho}} \tag{4.13}$$

$$R_{0,1}^{TM} \approx \frac{Y_1^{TM}(0) - \frac{j\omega\epsilon_0}{k_\rho}}{Y_1^{TM}(0) + \frac{j\omega\epsilon_0}{k_\rho}} \tag{4.14}$$

We first show that $R_{0,1}^{TM} \approx 1$. Let us consider $R_{0,1}^{TM}$ as a function of k_ρ. With the quasistatic assumption, we assume that $k_\rho \gg k_0$ where $k_0 = 2\times10^{-7}$ rad/m for $f = 10$ Hz, the highest

frequency of interest. It is sufficient, therefore, to consider the behavior of $R_{0,1}^{TM}$ for k_ρ over the range 1×10^{-6} (i.e., $5k_0$) to infinity. From (B.66) we have

$$Y_1^{TM}(0) = \hat{Y}_1^{TM} \frac{Y_2^{TM}(d_1) + \hat{Y}_1^{TM} \tanh(j k_{1,z} h_1)}{Y_2^{TM}(d_1) \tanh(j k_{1,z} h_1) + \hat{Y}_1^{TM}} \tag{4.15}$$

We observe that $\tanh(x) \approx 1$ for $x \geq 3$ and conclude that $Y_1^{TM}(0) \approx \hat{Y}_1^{TM}$ when the condition

$$h_1 \sqrt{k_\rho^2 - k_1^2} = h_1 \sqrt{k_\rho^2 + (\sigma_1 \mu \omega)^2} \geq 3 \tag{4.16}$$

holds. We have used the approximation $k_1 \approx -j\sigma_1 \mu\omega$, which is valid with our assumption that $\sigma_1 \gg \epsilon_0 \omega$. Condition (4.16) is always satisfied for $k_p \geq \frac{3}{h_1} = 0.006$ where $h_1 = 0.5$ km in our model. In fact, for sufficiently high frequencies, $|k_1| \geq 0.006$ and (4.16) holds for all values of k_ρ. This behavior reflects the fact the layered medium appears as a half space of conductivity σ_1 when h_1 is much greater than the skin depth of the top layer. With the assumption that (4.16) holds and $\sigma_1 \gg \epsilon_0 \omega$, we obtain

$$R_{0,1}^{TM} = \frac{\hat{Y}_1^{TM} - \frac{j\omega\epsilon_0}{k_\rho}}{\hat{Y}_1^{TM} + \frac{j\omega\epsilon_0}{k_\rho}} \approx 1 \tag{4.17}$$

for $k_\rho \geq 0.006$. For the range $10^{-6} \leq k_\rho \leq 0.006$ we numerically evaluated $Y_1^{TM}(0)$ for $10^{-3} \leq f \leq 10$ and found that the condition $|Y_1^{TM}(0)| \gg \frac{\omega\epsilon_0}{k_\rho}$ holds. As a result, we may assume that $R_{0,1}^{TM} \approx 1$ for our application.

For the development that follows, we find it convenient to rewrite $R_{0,1}^{TE}$ as

$$R_{0,1}^{TE} = \frac{k_\rho - \gamma_1^*}{k_\rho + \gamma_1^*} \tag{4.18}$$

where $\gamma_1^* = \gamma_1/\Lambda$, $\gamma_1 = j k_{1,z} = \sqrt{k_\rho^2 - k_1^2}$, and Λ is defined as

$$\Lambda = \frac{Z_2^{TE}(d_1) + \hat{Z}_1^{TE} \tanh(-\gamma_1 h_1)}{Z_2^{TE}(d_1) \tanh(-\gamma_1 h1) + \hat{Z}_1^{TE}} \tag{4.19}$$

Note that $Z_1^{TE}(0) = \hat{Z}_1^{TE} \Lambda$.

HED Expressions

Using the approximations and notation from the previous section, we rewrite equations (4.9)-(4.11) as

$$\begin{aligned} H_{0+,r} &= -\frac{Ids}{2\pi}\sin\phi\left[\int_0^\infty \frac{k_\rho\gamma_1^*}{k_\rho+\gamma_1^*}J_0(k_\rho r)dk_\rho + \frac{1}{r}\int_0^\infty \frac{k_\rho}{k_\rho+\gamma_1^*}J_1(k_\rho r)dk_\rho\right] \\ H_{0+,\phi} &= \frac{Ids}{2\pi r}\cos\phi\int_0^\infty \frac{k_\rho}{k_\rho+\gamma_1^*}J_1(k_\rho r)dk_\rho \\ H_{0+,z} &= \frac{Ids}{2\pi}\sin\phi\int_0^\infty \frac{k_\rho^2}{k_\rho+\gamma_1^*}J_1(k_\rho r)dk_\rho \end{aligned}$$

To obtain the $\hat{x}$ and $\hat{y}$ components of the magnetic field we apply the relations

$$\begin{aligned} H_x &= H_r\cos\phi - H_\phi\sin\phi \\ H_y &= H_r\sin\phi + H_\phi\cos\phi \end{aligned}$$

and find that

$$H_x = \frac{\partial}{\partial x}\left(\frac{y}{r}\frac{Ids}{2\pi}\int_0^\infty \frac{k_\rho}{k_\rho+\gamma_1^*}J_1(k_\rho r)dk_\rho\right) \tag{4.20}$$

$$\begin{aligned} H_y &= -\frac{Ids}{2\pi}\int_0^\infty \frac{k_\rho\gamma_1^*}{k_\rho+\gamma_1^*}J_0(k_\rho r)dk_\rho \\ &\quad -\frac{\partial}{\partial x}\left(\frac{x}{r}\frac{Ids}{2\pi}\int_0^\infty \frac{k_\rho}{k_\rho+\gamma_1^*}J_1(k_\rho r)dk_\rho\right) \end{aligned} \tag{4.21}$$

Point Current Electrode

As in the static case, we model the effects of the leakage currents with an isolated point electrode injecting current into the earth. The isolated point electrode is a convenient abstraction [*Ward and Hohmann*, 1987], and it is equivalent to a *time-varying electric monopole*. While it is possible to calculate the BART fields without appeal to the fiction of a point electrode, we have found that the use of the point electrode concept streamlines the computational process and makes clear the contributions of the various currents involved. An isolated point electrode cannot exist in reality because charge conservation requires that charges flow in and out to feed the electrode ([*Jackson*, 1975], p. 394). Any attempt to calculate the fields due solely to the point electrode will be inconsistent with Maxwell's

equations. Instead, the point electrode must be considered as part of a closed current system. For example, *Kauahikaua* [1978] and *Ward and Hohmann* [1987] have pointed out that the fields due to a finite length horizontal wire may be split into solutions due to point electrodes at the grounded ends and to current flow in the wire. We extend this reasoning to derive the fields due to a point electrode located at the origin from the solution for a semi-infinite horizontal wire that extends from the origin to infinity.

We find the fields due to a semi-infinite horizontal wire by integrating the HED expressions derived in the previous section along the x' axis. To make explicit the dependence of the dipole expressions on x' we replace $r = \sqrt{x^2 + y^2}$ by $r' = \sqrt{(x - x')^2 + y^2}$. In addition we note that $\frac{\partial r'}{\partial x} = -\frac{\partial r'}{\partial x'}$ and we replace ds by dx'. For H_x we obtain

$$\begin{aligned} H_x &= -\int_0^\infty \frac{\partial}{\partial x'} \frac{Iy}{2\pi r'} \int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r') dk_\rho dx' \\ &= -\frac{Iy}{2\pi r'} \int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r') dk_\rho \Bigg|_{x'=0}^{\infty} \\ &= \frac{I}{2\pi} \sin\phi \int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho. \end{aligned}$$

Note that the term at $x' = \infty$ vanishes since the fields must drop off at least as fast as $\frac{1}{r'}$. This claim can be made more rigorous by observing that the inner integral can be approximated (see Section 4.5.1) as a finite sum of terms of the form $\frac{1}{r'} I_1\left(\frac{jkr'}{2}\right) K_1\left(\frac{jkr'}{2}\right)$. Using asymptotic expressions for $I_1(x)$ and $K_1(x)$, we find that each of these terms falls off as $\frac{1}{(r')^2}$ as $r' \to \infty$.

From the derivation above we can see that H_x is due to the unconnected end, i.e. the end at $x' = 0^-$, of an incremental dipole at the origin and thus can be considered to be the field of a point electrode. The other end of this incremental dipole (at $x' = 0^+$) is connected to the next incremental dipole that makes up the semi-infinite wire and does not contribute to H_x. In the limit as $\omega \to 0$ we note that $\gamma_1^* \to \gamma_1$ and $H_x \to \frac{I}{4\pi r} \sin\phi$, in agreement with the results for the static case.

In a similar fashion, we obtain the following expression for H_y

$$\begin{aligned} H_y &= -\frac{I}{2\pi} \cos\phi \int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho \\ &\quad -\frac{I}{2\pi} \int_0^\infty \int_0^\infty \frac{k_\rho \gamma_1^*}{k_\rho + \gamma_1^*} J_0(k_\rho r') dk_\rho dx' \end{aligned}$$

As with H_x, the first term is the contribution due to the grounded end at $x' = 0^-$. The second term represents the fields due to currents induced by the primary magnetic field of the wire. This term disappears as $\omega \to 0$. The H_z component is given by

$$H_z = \frac{I}{2\pi} \int_0^\infty \frac{y}{r'} \int_0^\infty \frac{k_\rho^2}{k_\rho + \gamma_1^*} J_1(k_\rho r') dk_\rho dx', \tag{4.22}$$

and represents the fields due to currents flowing along the wire and associated induction currents. Note that at $\omega = 0$ and $x = 0$, (4.22) yields $H_z = \frac{I}{4\pi r}$ which is the static field of a semi-infinite line current.

To summarize, the fields due to a point electrode (i.e., the grounded end) at the origin with current I are given by

$$H_x = -\frac{I}{2\pi} \sin\phi \int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho \tag{4.23}$$

$$H_y = \frac{I}{2\pi} \cos\phi \int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho \tag{4.24}$$

$$H_z = 0 \tag{4.25}$$

To be consistent with the static case, we have multiplied the expressions for H_x and H_y by -1 in order to obtain the fields for a point electrode injecting current into the earth.

Finite Length Wire

In this section we consider the fields due to currents flowing along the rails. Each rail section is modeled as a wire of length L, where typically $L = 100$ meters. We use the HED expressions derived previously and ignore the fields due to the grounded ends, since these fields are included in our calculation of the fields due to leakage currents. We find that

$$\begin{aligned} H_x &= 0 \\ H_y &= -\frac{I}{2\pi} \int_{-L/2}^{L/2} \int_0^\infty \frac{k_\rho \gamma_1^*}{k_\rho + \gamma_1^*} J_0(k_\rho r') dk_\rho dx' \\ H_z &= \frac{I}{2\pi} \int_{-L/2}^{L/2} \frac{y}{r'} \int_0^\infty \frac{k_\rho^2}{k_\rho + \gamma_1^*} J_1(k_\rho r') dk_\rho dx. \end{aligned}$$

In our array the minimum distance from a rail section to an observation site is about 5 km so that $r \gg x'$. We take advantage of this fact to make the approximation $r' \approx r$ and write

$$H_y \approx -\frac{IL}{2\pi}\int_0^\infty \frac{k_\rho \gamma_1^*}{k_\rho + \gamma_1^*} J_0(k_\rho r) dk_\rho \tag{4.26}$$

$$H_z \approx \frac{IL}{2\pi}\frac{y}{r}\int_0^\infty \frac{k_\rho^2}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho. \tag{4.27}$$

In geophysical applications, this approximation is commonly made when $r \geq 5L$ [*Ward and Hohmann*, 1987].

4.5 Results

4.5.1 Numerical Evaluation Methods

The expressions for the magnetic fields consist of three types of integrals

$$\text{Type 1} \quad \int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho$$

$$\text{Type 2} \quad \int_0^\infty \frac{k_\rho \gamma_1^*}{k_\rho + \gamma_1^*} J_0(k_\rho r) dk_\rho$$

$$\text{Type 3} \quad \int_0^\infty \frac{k_\rho^2}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho$$

which belong to the class of integrals commonly referred to as Sommerfeld integrals. The slow convergence of the Sommerfeld integrals has led to the development of a variety of techniques for their numerical evaluation [*Goldman*, 1990; *Mohsen*, 1982]. In our case, the integrands for H_z and H_y diverge, and it is necessary to express the integral as the sum of a term for a homogeneous half-space with conductivity σ_1 and a residual integral term which converges [*Frischknecht*, 1967]. We rewrite the integrals as

$$\begin{aligned}
\int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho &= \int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1} J_1(k_\rho r) dk_\rho \\
&\quad + \int_0^\infty k_\rho \left(\frac{1}{k_\rho + \gamma_1^*} - \frac{1}{k_\rho + \gamma_1} \right) J_1(k_\rho r) dk_\rho \\
\int_0^\infty \frac{k_\rho \gamma_1^*}{k_\rho + \gamma_1^*} J_0(k_\rho r) dk_\rho &= \int_0^\infty \frac{k_\rho \gamma_1}{k_\rho + \gamma_1} J_0(k_\rho r) dk_\rho
\end{aligned}$$

$$
\begin{aligned}
& & & + \int_0^\infty k_\rho \left(\frac{\gamma_1^*}{k_\rho + \gamma_1^*} - \frac{\gamma_1}{k_\rho + \gamma_1} \right) J_0(k_\rho r) dk_\rho \\
\int_0^\infty \frac{k_\rho^2}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho & = & \int_0^\infty \frac{k_\rho^2}{k_\rho + \gamma_1} J_1(k_\rho r) dk_\rho \\
& & + \int_0^\infty k_\rho^2 \left(\frac{1}{k_\rho + \gamma_1^*} - \frac{1}{k_\rho + \gamma_1} \right) J_1(k_\rho r) dk_\rho .
\end{aligned}
$$

We can use various identities [*Kaufman and Keller*, 1983; *Ward and Hohmann*, 1987] to obtain the following analytical solutions for the homogeneous half-space integrals:

$$
\begin{aligned}
\int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1} J_1(k_\rho r) dk_\rho &= \frac{1}{r} I_1(\frac{jk_1 r}{2}) K_1(\frac{jk_1 r}{2}) \\
\int_0^\infty \frac{k_\rho \gamma_1}{k_\rho + \gamma_1} J_0(k_\rho r) dk_\rho &= -\frac{k_1^2}{8} \left(I_0(\frac{jk_1 r}{2}) K_0(\frac{jk_1 r}{2}) - I_2(\frac{jk_1 r}{2}) K_2(\frac{jk_1 r}{2}) \right) \\
\int_0^\infty \frac{k_\rho^2}{k_\rho + \gamma_1} J_1(k_\rho r) dk_\rho &= -\frac{1}{k_1^2 r^4} \left(3 - e^{-jk_1 r} (3 + 3jk_1 r - k_1^2 r^2) \right)
\end{aligned}
$$

where $I_1(x)$ and $K_1(x)$ are modified Bessel functions that are evaluated numerically with the *bessel.m* routine in MATLAB 5.1. The residual integral terms are evaluated over the interval $0 \leq k_\rho \leq 0.1$ using the *quad8.m* (adaptive recursive Newton-Cotes 8-panel rule) routine with an error tolerance of 10^{-3} in MATLAB 4.2C. The results are verified with the *d01akf.m* (adaptive Gauss 30-point and Konrod 61-point rules) routine in MATLAB 5.1 with absolute error tolerance of 10^{-13} and relative error tolerance of 10^{-3}.

As an additional check on the accuracy of our answers, we developed a method to approximate the integrals as the sum of homogeneous half-space solutions. Numerical evaluation of $R_{0,1}^{TE} = \frac{k_\rho - \gamma_1^*}{k_\rho + \gamma_1^*}$ showed that the reflection coefficient could be fairly well approximated at each frequency as

$$
R_{0,1}^{TE} \approx \sum_{l=1}^{L} \alpha_l \mathbf{Re} \left(R^{TE}(\sigma_l^{re}) \right) + j \sum_{m=1}^{M} \beta_m \mathbf{Im} \left(R^{TE}(\sigma_m^{im}) \right) \tag{4.28}
$$

where $R^{TE}(\sigma)$ is the reflection coefficient for a homogeneous half-space with conductivity σ, and α_l and β_m are coefficients which minimize the mean squared error of the approximation for a given set of σ_l^{re} and σ_m^{re}. We found that 4 terms in each sum ($L = M = 4$) provided a good fit. The value for σ_1^{im} is chosen so that the peaks of $\mathbf{Im} \left(R^{TE}(\sigma_1^{im}) \right)$ and $R_{0,1}^{TE}$ occur at the same value of k_ρ. We also use the heuristic values $\sigma_1^{re} = \sigma_1^{im}$, $\sigma_L^{re} = 1$,

and $\sigma_M^{im} = 30$. The intermediate values of σ_l^{re} and σ_m^{im} are chosen to span the the interval from the first to the Lth or Mth value with logarithmic spacing. Note that the parameter values must be recomputed for each frequency of interest.

As an example, if we use the approximation for a Type 1 integral, we obtain

$$\begin{aligned}
\int_0^\infty \frac{k_\rho}{k_\rho + \gamma_1^*} J_1(k_\rho r) dk_\rho &= \frac{1}{2}\int_0^\infty (1 + R_{0,1}^{TE}) J_1(k_\rho r) dk_\rho \\
&\approx \frac{1}{2}\int_0^\infty \left(1 + \sum_{l=1}^{L} \alpha_l \mathbf{Re}\left(R^{TE}(\sigma_l^{re})\right)\right. \\
&\quad \left. + j\sum_{m=1}^{M} \beta_m \mathbf{Im}\left(R^{TE}(\sigma_m^{im})\right)\right) J_1(k_\rho r) dk_\rho \\
&= \sum_{l=1}^{L} \alpha_l \mathbf{Re}\left(\int_0^\infty \frac{k_\rho}{k_\rho + \gamma_l} J_1(k_\rho r) dk_\rho\right) \\
&\quad + j\sum_{m=1}^{M} \beta_m \mathbf{Im}\left(\int_0^\infty \frac{k_\rho}{k_\rho + \gamma_m} J_1(k_\rho r) dk_\rho\right) \\
&= \frac{1}{r}\sum_{l=1}^{L} \alpha_l \mathbf{Re}\left(I_1\left(\frac{jk_l r}{2}\right) K_1\left(\frac{jk_l r}{2}\right)\right) \\
&\quad + \frac{j}{r}\sum_{m=1}^{M} \beta_m \mathbf{Im}\left(I_1\left(\frac{jk_m r}{2}\right) K_1\left(\frac{jk_m r}{2}\right)\right)
\end{aligned}$$

where γ_l, k_l, γ_m, and k_m are derived from the values for σ_l^{re} and σ_m^{re}, and we have used the fact that $\mathbf{Im}\left(R^{TE}(\sigma_m^{im})\right) = \mathbf{Im}\left(1 + R^{TE}(\sigma_m^{im})\right)$. The effectiveness of the approximation is shown in Figure 4.9, where we have plotted the magnitude of the Type 1 integral versus radial distance for a range of frequency values.

4.5.2 Frequency Response of Integrals

In Figure 4.10 we show the frequency responses of the various integrals for a range of radial distances. The magnitude responses for the Type 1 and Type 3 integrals are normalized by the magnitudes of the static case, $\frac{1}{2r}$ and $\frac{1}{2r^2}$, respectively. Since the Type 2 integral vanishes in the static case, we normalize by the magnitude of the Type 3 integral (static case).

The frequency response plots show that the conducting earth acts a low-pass filter of the horizontal fields of a point electrode (Type 1 integral) and the vertical field of a finite length wire (Type 3 integral), with the cut-off frequency decreasing as the distance to

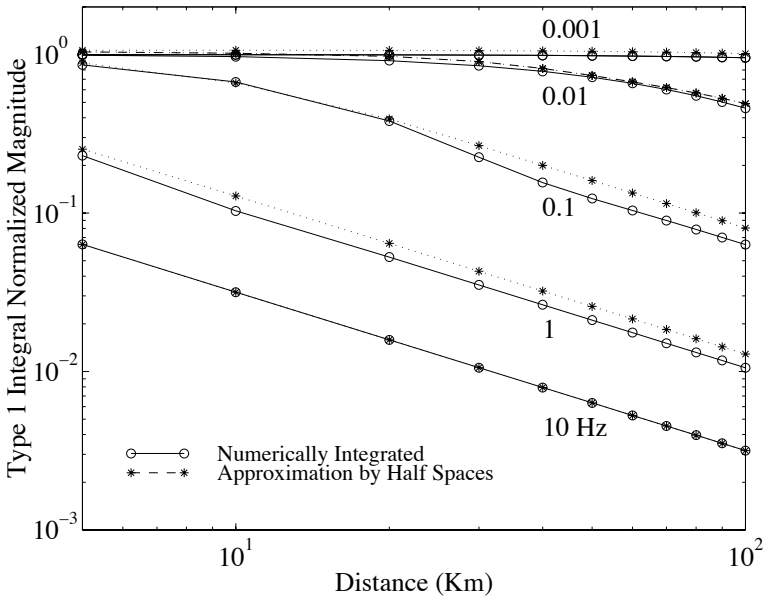
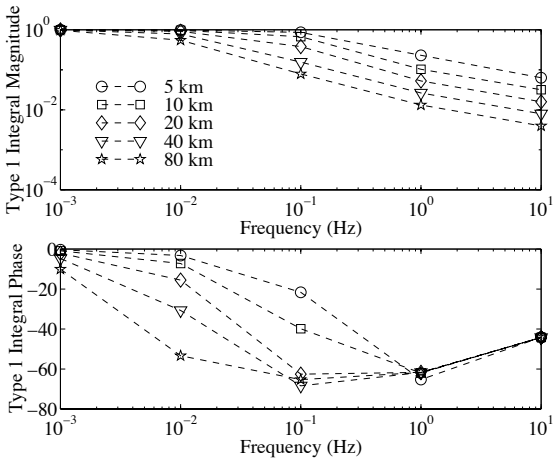


Figure 4.9: Magnitude of Type 1 integral normalized by the magnitude of the integral for the zero frequency case. The magnitude is plotted versus radial distance with frequency as a parameter. Both numerically integrated (circles) and approximation by half space (stars) values are shown. The lines are labeled by frequency in units of Hz.

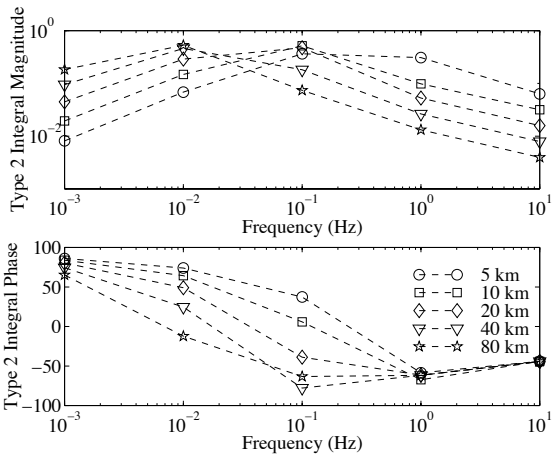
the observation point increases. By contrast, the horizontal field of the finite length wire (Type 2 integral) increases with frequency before dropping off. The peak frequency of the response increases as the distance to the observation point decreases.

4.5.3 Polarization Plots

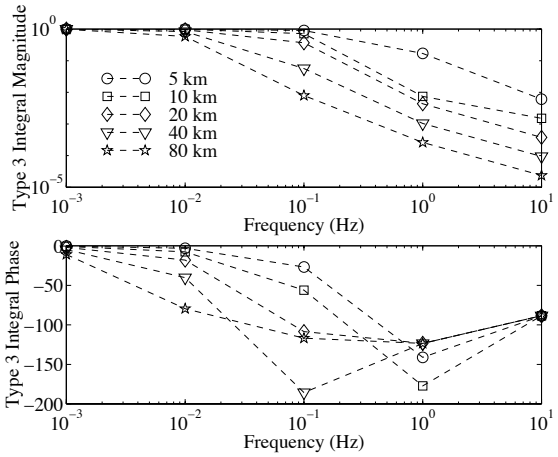
In this section we use the electric circuit model presented in Section 4.3 and the formulas presented in Section 4.4 to calculate the magnetic fields due to BART. Because of the complexity of the BART system, we must necessarily make some simplifying assumptions. First, we assume that only one train is present in the system. We then position the train at various locations and compute the fields for each location. In the examples shown below, we computed the fields for 63 different train locations. In each case the train was modeled as a 0.1 Ω load resistance positioned at distance of 0.5 km from a passenger station or substation. The 0.1 Ω load resistance value models a train drawing 10 kilo-amperes of current at a rail voltage of 1000 VDC. The 0.5 km distance value was chosen to reflect



(a) Type 1 Integral



(b) Type 2 Integral



(c) Type 3 Integral

Figure 4.10: Magnitude and phase response of dipole integrals for a range of radial distances. For Type 1 and Type 3 integrals the magnitudes are normalized by the magnitude of the zero frequency case. For the Type 2 integral the magnitude is normalized by the value of the Type 3 zero frequency case.

the fact that the maximum current draw typically occurs about 10 seconds after the train has departed from a station. For example, in Figure 4.3(a), the peak current draw occurs when the train is between 0.3 and 0.8 km from the station of departure.

For each train location, we use the SPICE circuit model to calculate first the rail and leakage currents. We then use the formulas from Section 4.4 to compute both the static and quasistatic fields generated by these currents. Figures 4.11 and 4.12 present the results in the form of polarization plots, where for each observation site we have plotted the computed fields associated with each of the train locations. Because the train locations are distributed across the BART system, the resultant fields provide a first-order estimate of the expected polarization at each observation site.

In Figure 4.11, we show the static field polarization plots for system-wide leakage resistivity values of $\eta = 0.5$ Ω-km and $\eta = 5$ Ω-km. At each station, we find that the polarization vectors exhibit a strong directional component. This finding agrees with the results presented in Figure 3.9 of Chapter 3. In addition the magnitudes of the computed fields for $\eta = 0.5$ Ω-km agree fairly well with the measured magnitudes. There are, however, clear differences between the detailed features of the computed and measured polarization vectors. For example, the measured polarization at LKC has more of a vertical orientation than does the computed polarization. We attribute these differences to uncertainties in both the circuit and propagation models and the use of simplifying assumptions.

The primary simplifications we have made are (1) the assumption of a single leakage resistivity value for the entire system, (2) the use of a single train to calculate currents and fields, and (3) the use of a one-dimensional conductivity profile. The limitations of simplification (3) were discussed previously. With regards to simplification (1), we noted in Section 4.3.2 that the leakage resistivity can vary greatly across the system since it depends on local factors such as the moisture of the soil, the presence of underground pipes, and the build-up of dirt films on the rails and ties. Our use of a single value for leakage resistivity reflects the difficulty in obtaining information on these local, and potentially time-varying, conditions. We have found that the use of a low system-wide value $\eta = 0.5$ Ω-km leads to the best agreement with the experimental data. This value is much lower than the values (4 to 4000 Ω-km) reported by *Galler and Todd* [1992] for the BART system. But as shown in panels (d)-(f) of Figure 4.11, the use of a more reasonable value of $\eta = 5$ Ω-km results in predicted field magnitudes that are significantly less than the

observed magnitudes shown in Figure 3.9. For both values of leakage resistivity, however, we do observe the strong directional component in the polarization.

Our use of a single train also represents a significant simplification, since there are typically more than 40 trains in operation at any given time [*Nishinaga et al.*, 1994]. The current distribution due to a single train will inevitably vary from the distribution for multiple trains. For example, the single train assumption does not take into account the possibility of one train supplying propulsion current to another train during a period of regenerative braking. Also note that superposition does not apply because the trains are modeled as resistances and not as independent sources. Since our primary goal is to develop a first-order understanding of the polarization of the fields, we may justify the simplification in the following manner. As shown in Chapter 3, we can consider the BART interference to be the superposition of transients. The time scales of the transients indicate that the most significant transients are associated with train accelerations and decelerations. As a first-order approximation, we attribute the polarization vectors of the more significant transients to fields associated with a single train leaving from or arriving at a station. In support of this conclusion, we observe that the nominal train departure times (as obtained from published BART schedules) are, for the most part, staggered so that trains drawing peak currents at the same point in time are generally spaced at least several stations apart. The spacing is greater during off-peak hours when the number of departures per hour decreases. As shown in Figure 4.6 the most significant net propulsion and leakage currents arise within a 15 km distance of the train. Thus, to first order, trains accelerating at the same time but spaced sufficiently apart can be treated independently.

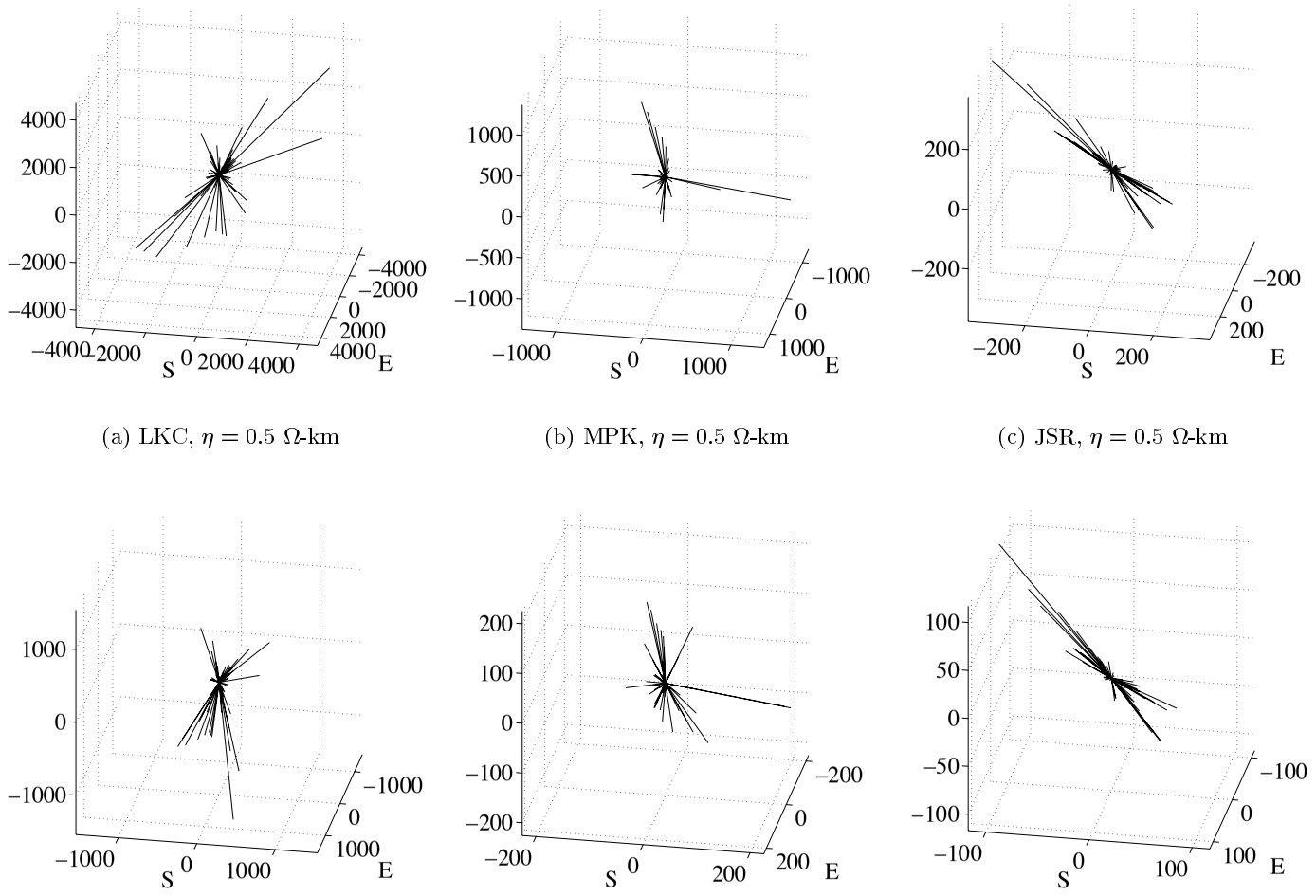
Up to this point, we have concentrated on the behavior of the static fields. Turning to Figure 4.12, we consider the polarization of the quasistatic fields at 0.01 and 0.1 Hz with an assumed system-wide leakage resistivity of 0.5 Ω-km. At 0.01 Hz, the polarization of the quasistatic fields is essentially the same as that for the static fields. This is not a surprising result, since the frequency response curves in Figure 4.10 show that the field magnitudes are not greatly affected at 0.01 Hz. At 0.1 Hz, the effect of the conducting earth becomes more pronounced, especially at the larger radial distances. Thus, we see that the fields at MPK and JSR are significantly affected while the fields at LKC are still quite similar to the static fields. On average, LKC is much closer to the BART tracks than either MPK or JSR. At 0.1 Hz, the effect of the upper conducting layer is much more pronounced because the 2.2 km skin depth is of the same order as the 0.5 km thickness

of the layer. At 0.01 Hz, the skin depth of this upper layer is 7.1 km. To verify the effect of the upper layer at 0.1 Hz, we also computed BART magnetic fields assuming a homogeneous half space with $\sigma = 0.03$ S/m, which is the conductivity of the second layer in our conductivity profile. Under these conditions, the polarizations of the fields at MPK and JSR are not significantly different from the static case.

It is important to note that because the BART interference is broadband, the quasistatic field computations that we have presented serve only as a rough indicator of the effect of the conducting earth on the BART fields. Essentially, they indicate that our general observation of a strong directional polarization still holds in the presence of a conducting earth.

4.6 Summary

In this chapter we presented a first-order model for the generation and propagation of the magnetic fields due to BART. We developed a lumped-element circuit model for the BART system and used the model to calculate the distribution of currents in the system. We then developed expressions for the static and quasistatic magnetic fields due to currents flowing in point electrodes and finite length wires. We computed the fields due to trains at various locations in the BART system, and found that the polarization of the fields exhibited a strong directional component. This observation is in agreement with the experimental data presented in Chapter 3 and supports our use of polarization vectors in Chapters 5 and 7 to differentiate the BART fields from the natural background. We also showed that the characteristic time scales of propulsion current transients were similar to those of the observed magnetic field transients. This observation supports our use of a fixed range of unknown scales in the detection method discussed in Chapters 5 through 7.



(a) LKC, $\eta = 0.5$ Ω-km

(b) MPK, $\eta = 0.5$ Ω-km

(c) JSR, $\eta = 0.5$ Ω-km

(d) LKC, $\eta = 5$ Ω-km

(e) MPK, $\eta = 5$ Ω-km

(f) JSR $\eta = 5$ Ω-km

Figure 4.11: Polarization plots for static fields. The top row shows the plots for a system-wide leakage resistivity of $\eta = 0.5$ Ω-km, while the bottom row show plots for $\eta = 5$ Ω-km. Units are pT.

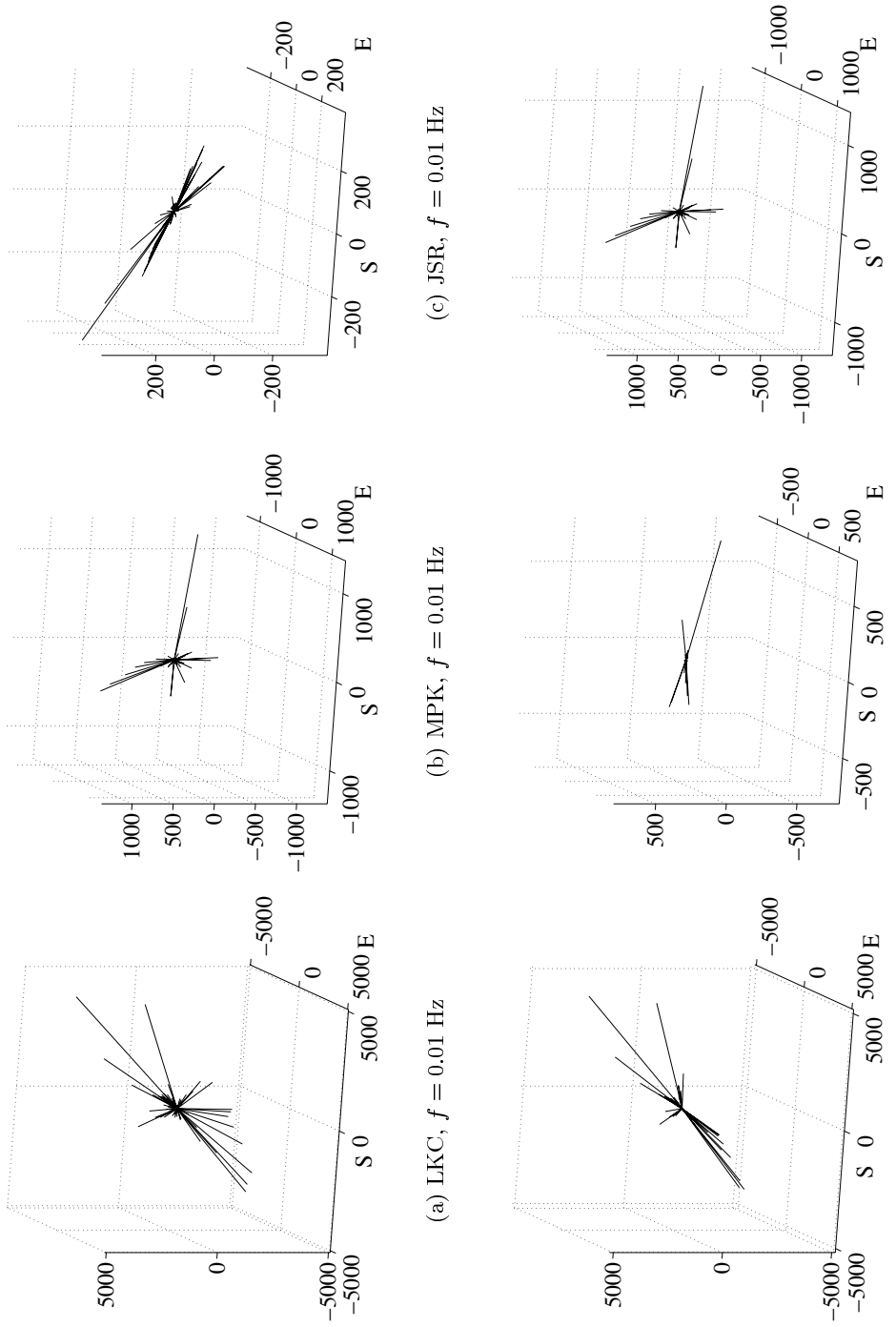


Figure 4.12: Polarization plots for quasistatic fields. The top row show the polarization at $f = 0.01$ Hz, while the bottom row shows the polarization at $f = 0.1$ Hz. The system-wide leakage resistivity was $\eta = 0.5$ Ω-km. Units are pT.

Chapter 5

Signal Processing: Overview of the Approach

5.1 Introduction

In this very short chapter we provide an overview of the signal processing approach that will be explained in detail in the following two chapters. As discussed in Chapter 1, our goal is to monitor for anomalous changes in the geomagnetic background in the presence of the BART interference. As a first step towards that goal, we have developed a method for identifying and then removing the transients generated by BART. This chapter presents the signal model that is used in developing our method. It also provides an outline of the method.

5.2 Signal Model

As shown in the previous chapter, the generation and propagation of the magnetic fields due to BART are complex processes. We do not attempt to model the myriad complexities but instead propose a signal model that aims to represent the most significant features of the interference.

We begin by recalling that the SFBAY array consists of three measurement sites. There is a three-axis (vertical, geomagnetic North-South, and geomagnetic East-West)

magnetometer at each site, resulting in a total of nine data channels. We model the signal on the ith channel as

$$x_i(t) = \sum_j A_{i,j} s_{k_{i,j},l_{i,j}}(t) + v_i(t) + o_i(t) \tag{5.1}$$

where $i = 1$ to 9, and $A_{i,j}$, $k_{i,j}$, and $l_{i,j}$ are unknown amplitudes, scales, and delays, respectively. We assume that $A_{i,j} \neq 0$, and that $k_{i,j}$ and $l_{i,j}$ are integers. We define $s_{k,l}(t) = 2^{\gamma k/2}\xi_k(t-l)$ where $\xi_k(t) = 2^{-k/2}\xi(2^{-k}t)$ and $\xi(t)$ is the transient signal model. The term $v_i(t)$ represents the natural geomagnetic background and $o_i(t)$ represents other natural signals of interest, such as possible earthquake precursors. We assume that both $v_i(t)$ and $o_i(t)$ are approximate $1/f$ processes. We showed in Chapter 3 that the geomagnetic background can be treated as a *piecewise* approximate $1/f$ process. The Loma Prieta precursor spectra presented in that chapter also exhibit a $1/f$ type behavior.

In addition to the nine data channels in the SFBAY, we also use the three channels acquired at SAO. For these channels we assume that $x_i(t) = v_i(t), i = 10$ to 12, i.e., BART interference and earthquake precursor signals are negligible. Because SAO is located in a region along the San Andreas Fault that is prone to frequent earthquakes, with magnitudes typically ranging from M 4.0 to M 5.0, it is entirely possible that there are precursor signals due to earthquakes local to SAO. Our assumption is that SAO data will not exhibit precursor signals due to a major earthquake in the SFBAY area, e.g., an event along the Hayward Fault. This assumption is consistent with the modeling in [*Fraser-Smith et al.*, 1993], which placed a 100 km upper limit on the distance over which electromagnetic precursors may be detected.

The signal model in (5.1) approximates the BART interference as the superposition of scaled and delayed versions of a single transient $\xi(t)$. Although the transients comprising the observed interference vary in shape, they can be characterized roughly as non-zero mean transients with signal energies concentrated in the main lobes. Our transient signal model serves as a first order approximation if we choose a $\xi(t)$ with these general characteristics. In Figure 5.1 we show that the C_2 scaling functions and Gaussian functions serve as good models for the transient waveforms. Other models are possible and the choice of a model appears somewhat arbitrary. For the remainder of this thesis we choose $\xi(t)$ to be a C_2 scaling function.

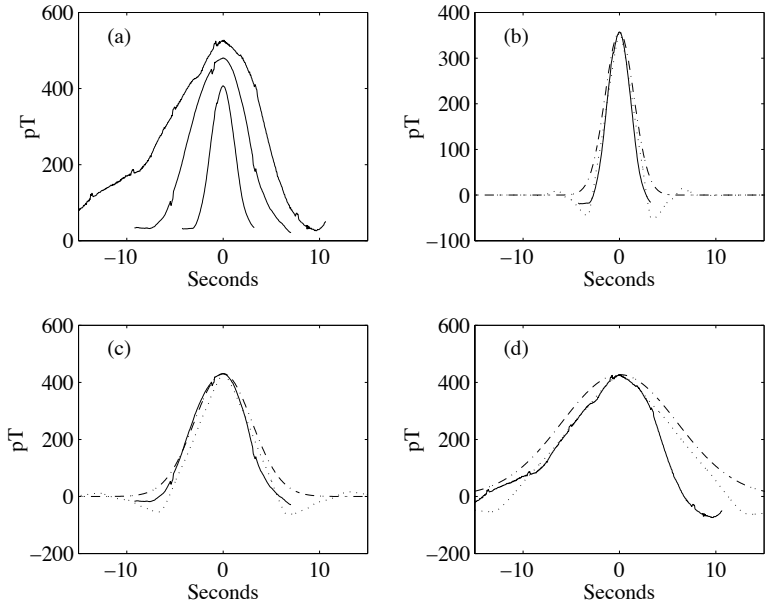


Figure 5.1: Panel (a): BART transients with temporal widths of 5.8, 2.9 and 0.8 seconds. Waveforms of C_2 scaling functions (dotted lines) and Gaussian functions (dash-dotted lines) are superimposed in panels (b)–(d).

In Chapter 3 we showed that the temporal widths of the observed transients lie within a limited range. To be consistent with these observations, we limit the scale parameter k to lie within a range $\mathcal{K} = \{6, 7, 8, 9\}$ in order to obtain C_2 scaling functions with the desired temporal widths.

5.3 Overview of the Approach

Given the signal model defined in (5.1), our goal is to detect and remove the transients $A_{i,j}\, s_{k_{i,j},l_{i,j}}(t)$. We develop our approach to this problem in two steps. First, we consider the detection and removal of transients when we have only a single channel. We then extend the single channel approach to the case of multiple channels.

5.3.1 Single Channel

In the single channel case, we process each channel in the array without making use of the information from other channels. For each channel, the detection problem is one of detecting multiple transients in the presence of $1/f$ noise. In Chapter 6 we develop

an undecimated discrete wavelet transform (UDWT) based method for the detection of a single transient in the presence of $1/f$ noise. We extend the method to the case of multiple transients in Chapter 7. To remove detected transients, we form an estimate of each transient in the UDWT domain. We then use an inverse translation invariant wavelet transform (TIWT) to form the time domain estimate of the transient and subtract this estimate from the original waveform. Details of the estimation process are presented in Chapter 7.

5.3.2 Multiple Channels

The use of multiple channels allows us to take advantage of the spatial characteristics of the BART interference to refine our detection process. For each detected transient, we form a polarization vector consisting of the peak amplitudes of the transient as observed at all channels in the array. We then determine whether or not the polarization vector lies within a multidimensional subset constructed from the observed BART polarization vectors. If the vector lies within the BART subset, we use the method described above to remove the associated transients from each channel in the array. Details of this process are provided in Chapter 7.

5.4 Summary

We have briefly described our approach to the identification and removal of transients due to BART. In Chapters 6 and 7 we develop this approach in detail and present examples of its performance.

Chapter 6

Detection of Transients in $1/f$ Noise

6.1 Introduction

In the previous chapter we presented an outline of our overall approach for the detection and removal of transients due to BART. A key component of this approach is the detection of transients in $1/f$ noise. We now develop in detail the theory for a generalized likelihood ratio test (GLRT) detector based upon the undecimated discrete wavelet transform (UDWT). We consider the case of a single transient with unknown amplitude, scale, and delay parameters in the presence of $1/f$ Gaussian noise. The extension to multiple transients is presented in the next chapter.

Wavelet transforms have been widely applied to the problem of transient detection and processing, primarily because the transform basis functions provide good time localization [*Friedlander and Porat*, 1992; *Frisch and Messer*, 1994; *Mallat and Hwang*, 1992]. Applications have appeared in biomedical signal processing [*Chin et al.*, 1994] and power system transient detection [*Robertson et al.*, 1994], among other areas.

A number of detection methods have been proposed. *Frisch and Messer* [1994] observed that the wavelet transform acts as bank of matched filters, and can therefore be used as a generalized likelihood ratio test (GLRT) detector for transients modeled by wavelets in the presence of white noise. For transients of unknown shape, GLRT detectors that form the detection statistic in the wavelet domain have also been presented [*Friedlander and Porat*, 1992; *Frisch and Messer*, 1992]. For a transient with known parameters in $1/f$

Gaussian noise, *Wornell* [1996] described a matched filter detector in the discrete wavelet transform (DWT) domain. This detector relies on the observation that the DWT acts as an approximate whitening transform for $1/f$ noise. As shown in Section 6.3.3, the performance of the detector is shift variant.

Several authors have proposed detectors based on maxima tracking in either the UDWT domain [*Chin et al.*, 1994; *Mallat and Hwang*, 1992] or the analytic wavelet transform domain [*Abry and Flandrin*, 1994]. These techniques rely on the observation that the evolution of the transform maxima across scales provides a measure of the local regularity of the signal [*Mallat and Hwang*, 1992]. Maxima tracking makes intuitive sense, but the connection to standard detection theory has not been clear.

The method that we introduce in this chapter utilizes pattern matching in the UDWT domain and extends the work of *Wornell* [1996] to the unknown parameter case. We explain the theory underlying our method in Section 6.3 and then discuss its implementation in Section 6.4. In particular, we argue that the local maxima in the UDWT domain provide estimates of the scales and delays that are most likely to maximize the proposed detection statistic, and therefore a maxima tracking approach may be used to reduce the order of complexity of the implementation. In Section 6.5 we use Monte Carlo simulations to compute the receiver operating characteristics (ROC) for transients modeled by either scaling functions, Gaussian functions, or two-sided exponential functions. The simulation results indicate that the reduction in computational complexity using maxima tracking methods can be achieved with negligible impact on detector performance.

6.2 Notation

The symbols $\phi(t)$ and $\psi(t)$ represent scaling functions and wavelets which are defined using the framework of an orthogonal multiresolution analysis (MRA) [*Daubechies*, 1992]. For wavelets at other scales, we use the subscript notation $\psi_m(t) = 2^{-m/2}\psi(2^{-m}t)$ and $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t-n)$. The inner product of two functions is defined as $\langle f(t), g(t)\rangle = \int_{-\infty}^{\infty} f(t)g^*(t)dt$ where g^* is the complex conjugate of g.

The DWT of a signal $x(t)$ is $X_{m,n} = \langle x(t), \psi_{m,n}(t)\rangle$. The choice of time origin for the basis functions $\psi_{m,n}(t)$ is arbitrary, and we define other DWT's with basis functions

$\psi_{m,n}(t-J)$ and the notation

$$X_{m,n}^{[J]} = \langle x(t), \psi_{m,n}(t-J) \rangle \tag{6.1}$$

where J is an arbitrary integer shift. If M denotes the largest analysis scale of interest, then the $X_{m,n}^{[J]}$ are invariant to shifts by integer multiples of 2^M, i.e. $X_{m,n}^{[J]} = X_{m,n+2^{(M-m)}}^{[J-2^M]}$. As a result there are 2^M unique DWT shifts, with each shift giving rise to a different decomposition of the signal $x(t)$. The UDWT is defined as

$$\tilde{X}_{m,n} = \langle x(t), \psi_m(t-n) \rangle. \tag{6.2}$$

6.3 Theory

6.3.1 Detection Model

We consider the detection of a transient signal with unknown amplitude, scale, and delay. The standard hypothesis test is

$$\begin{array}{lrcl} H_0: & x(t) & = & v(t) \\ H_1: & x(t) & = & A s_{k,l}(t) + v(t) \end{array} \tag{6.3}$$

where A, k, and l are the unknown amplitude, scale, and delay, respectively, of the transient, and we assume that $A \neq 0$ and k and l are integers. We define

$$s_{k,l}(t) = 2^{\gamma k/2} \xi_k(t-l) \tag{6.4}$$

where $\xi_k(t) = 2^{-k/2}\xi(2^{-k}t)$ and $\xi(t)$ is the signal model. The reason for the leading $2^{\gamma k/2}$ factor is given in Section 6.3.3. The additive noise $v(t)$ is assumed to be a $1/f$ Gaussian random process.

A standard scheme for detection with unknown parameters is the generalized likelihood ratio test (GLRT) [*Frisch and Messer*, 1994; *Kay*, 1998]. It has the following form: choose H_1 if the likelihood ratio $r(x(t))$ satisfies $r(x(t)) > r_1$ where

$$r(x(t)) = \frac{\max_{\{A,k,l\}} f\left(x(t)|A,k,l,H_1\right)}{f\left(x(t)|H_0\right)} \tag{6.5}$$

and r_1 is a threshold value chosen to achieve a desired probability of false alarm (PFA); and

choose H_0 otherwise. In most cases the likelihood ratio can be replaced with a sufficient test statistic that has a simpler expression [*Kay*, 1998].

6.3.2 Detection in white Gaussian noise

In the special case where $\gamma = 0$, $v(t)$ is a stationary white Gaussian noise process, and the GLRT detector is: choose H_1 if the sufficient test statistic

$$\lambda(x(t)) = \max_{k,l} \left| \int x(t) s_{k,l}(t) dt \right| \; > \; \lambda_1. \tag{6.6}$$

Implementation of the GLRT detector requires calculation of the inner products $\langle x(t), \xi_k(t-l) \rangle$. In the special cases where $\xi(t)$ is either a scaling function or a wavelet, the inner products are simply the lowpass or highpass filter outputs, respectively, of an iterated filter bank implementation of the UDWT (see, for example, [*Vetterli and Kovačević*, 1995]). As a result, they can be computed in $O(N \log_2 N)$ time. The performance of this detector when $\xi(t)$ is a wavelet was discussed in [*Frisch and Messer*, 1994].

6.3.3 Detection in $1/f$ Gaussian noise

Shift Variant Detection with Known Parameters

Detection in the presence of a $1/f$ noise process is based on the observation by Wornell and others ([*Wornell*, 1996] and references therein) that the DWT acts as an approximate whitening transform for such a process (see Chapter 3 for more details). We consider first the detection problem with known signal parameters in order to develop concepts that will be useful in understanding the unknown parameter case. Following *Wornell* [1996], we note that the equivalent hypothesis test in the DWT domain is

$$\begin{array}{lrcl} H_0: & X_{m,n}^{[0]} & = & V_{m,n}^{[0]} \\ H_1: & X_{m,n}^{[0]} & = & A S_{m,n}^{[0],\{k,l\}} + V_{m,n}^{[0]} \end{array} \tag{6.7}$$

where A, k and l are known parameters. Since A is known, we may assume $A > 0$ without loss of generality. $X_{m,n}^{[J]}$ was defined in (6.1) and

$$S_{m,n}^{[J],\{k,l\}} = \langle s_{k,l}(t), \psi_{m,n}(t - J) \rangle \tag{6.8}$$

is the DWT of $s_{k,l}(t)$ with shift J. We have used the value $J = 0$ since this detector uses the standard, unshifted DWT. The likelihood ratio $r(\mathbf{X}^{[0]})$ is

$$\frac{f\left(\mathbf{X}^{[0]}|H_1\right)}{f\left(\mathbf{X}^{[0]}|H_0\right)} = \frac{\prod_{m,n} \exp\left(-\frac{\left(X^{[0]}_{m,n} - AS^{[0],\{k,l\}}_{m,n}\right)^2}{2\sigma_w^2 2^{\gamma m}}\right)}{\prod_{m,n} \exp\left(-\frac{\left(X^{[0]}_{m,n}\right)^2}{2\sigma_w^2 2^{\gamma m}}\right)} \tag{6.9}$$

where $\mathbf{X}^{[J]} = \{X^{[J]}_{m,n}, m, n \in \mathcal{Z}\}$ is the vector of DWT observations. We can simplify (6.9) to obtain the sufficient test statistic $\lambda(\mathbf{X}) = \sum_{m,n} 2^{-\gamma m} X^{[0]}_{m,n} S^{[0],\{k,l\}}_{m,n}$, where we have dropped the superscript on $\mathbf{X}$ for notational convenience.

Because the DWT is shift variant, the performance of the detector is also shift variant. To see this, we consider the performance index $d = \frac{A\sqrt{\mathcal{E}_k(l)}}{\sigma_w}$, which is the normalized distance between the distributions of $\lambda(\mathbf{X})$ under the two hypotheses stated in (6.7). For each scale k, the term $\mathcal{E}_k(l)$ is a function of l and is defined as

$$\mathcal{E}_k(l) = \frac{1}{A} E\left[\lambda(\mathbf{X})|H_1\right] = \sum_{m,n} 2^{-\gamma m} \left(S^{[0],\{k,l\}}_{m,n}\right)^2 \tag{6.10}$$

where $E[\lambda(\mathbf{X})|H_1]$ is the expected value of $\lambda(\mathbf{X})$ given hypothesis H_1. Figure 6.1 shows an example of the variation of $\mathcal{E}_k(l)$ with input shift l when $\xi(t)$ is a C_2 scaling function and $\gamma = 1$. We can define a ratio $\rho = \frac{\max_l \mathcal{E}_k(l)}{\min_l \mathcal{E}_k(l)}$ as a measure of the shift variance of the detector. In Figure 6.2 we plot ρ as a function of γ for C_2 scaling functions at various scales and show that the degree of shift variance is an increasing function of γ. Note that by Parseval's identity [*Vetterli and Kovačević*, 1995] we should obtain $\rho = 1$ when $\gamma = 0$. In practice, however, we obtain $\rho \approx 1$ because we use a finite range of analysis scales in the DWT.

Shift Invariant Detection with Known Parameters

We construct a shift invariant detector by first noting that the noise statistics are independent of DWT shift, since all shifted transforms also act as approximate whitening transforms. As a result, for a given known signal $s_{k,l}(t)$, we are free to choose the DWT

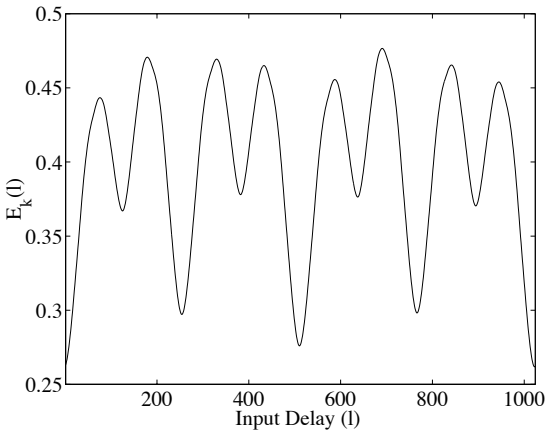


Figure 6.1: Variation of $\mathcal{E}_k(l)$ with input delay l for a C_2 scaling function ($k = 7$) analyzed with a C_2 wavelet over the range $m = 1$ to 10 with $\gamma = 1$. The dependence on l shows that the performance of the detector defined in Section 6.3.3 is shift variant.

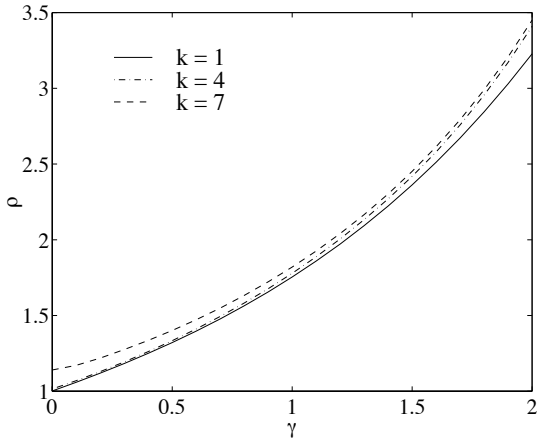


Figure 6.2: The ratio $\rho = \frac{\max_l(\mathcal{E}_k(l))}{\min_l(\mathcal{E}_k(l))}$ as a function of γ for a C_2 scaling function analyzed with a C_2 wavelet over the range $m = 1$ to 10 for k values 1, 4, and 7. The ratio ρ is a measure of the shift variance of the detector in Section 6.3.3.

shift that maximizes the performance index. For the signal $s_{k,0}(t)$ with delay $l = 0$, we define a detector with a test statistic of the form

$$\lambda(\mathbf{X}) = \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[J_k]} S_{m,n}^{[J_k],\{k,0\}}, \tag{6.11}$$

where

$$\begin{aligned} J_k &= \arg\max_J \frac{1}{A} E[\lambda(\mathbf{X})|H_1] \\ &= \arg\max_J \sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[J],\{k,0\}} \right)^2 . \end{aligned} \tag{6.12}$$

The subscript k on J_k indicates that the optimum DWT shift is a function of the scale k of the input signal. The performance index of this detector is equal to the maximum performance index of the shift variant detector, since

$$\begin{aligned} \max_J \sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[J],\{k,0\}} \right)^2 &= \max_{-l} \sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[0],\{k,l\}} \right)^2 \\ &= \max_l \mathcal{E}_k(l), \end{aligned} \tag{6.13}$$

where we have used the identity

$$S_{m,n}^{[J],\{k,0\}} = S_{m,n}^{[0],\{k,-J\}}. \tag{6.14}$$

We introduce the notation

$$\mathcal{E}_k = \max_l \mathcal{E}_k(l) \tag{6.15}$$

for use in the remainder of this dissertation.

For the case of arbitrary delay l, we define a test statistic

$$\lambda(\mathbf{X}, l) = \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[J_k(l)]} S_{m,n}^{[J_k(l)],\{k,l\}}, \tag{6.16}$$

where

$$J_k(l) = \arg\max_J \sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[J],\{k,l\}} \right)^2. \tag{6.17}$$

Using the identity stated in (6.14), we note that

$$\sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[J_k],\{k,0\}} \right)^2 = \sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[J_k+l],\{k,l\}} \right)^2 \tag{6.18}$$

so that $J_k(l) = J_k + l$. We may therefore rewrite (6.16) as

$$\begin{aligned} \lambda(\mathbf{X}, l) &= \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[J_k+l]} S_{m,n}^{[J_k+l],\{k,l\}} \\ &= \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[J_k+l]} S_{m,n}^{[J_k],\{k,0\}}. \end{aligned} \tag{6.19}$$

In the UDWT domain the statistic is written as

$$\lambda(\mathbf{X}, l) = \sum_{m,n} 2^{-\gamma m} \tilde{X}_{m,2^m n+J_k+l} \tilde{S}_{m,2^m n+J_k}^{\{k,0\}}, \tag{6.20}$$

where we have introduced the notation

$$\tilde{S}_{m,n}^{\{k,l\}} = \langle s_{k,l}(t), \psi_m(t-n) \rangle \tag{6.21}$$

for the UDWT of $s_{k,l}(t)$ and used the relation $\tilde{S}_{m,2^m n+J}^{\{k,l\}} = S_{m,n}^{[J],\{k,l\}}$.

The test statistic $\lambda(\mathbf{X}, l)$ has conditional means

$$E\left[\lambda(\mathbf{X}, l)|H_0\right] = 0 \text{ and } E\left[\lambda(\mathbf{X}, l)|H_1\right] = \mathcal{E}_k A,$$

and conditional variances

$$var\left[\lambda(\mathbf{X}, l)|H_0\right] = var\left[\lambda(\mathbf{X}, l)|H_1\right] = \sigma_w^2 \mathcal{E}_k.$$

As a result, the performance index is $d = \frac{A\sqrt{\mathcal{E}_k}}{\sigma_w}$, and the detector is shift invariant with respect to the delay parameter l. The likelihood ratio that corresponds to this test statistic is

$$r\left(\mathbf{X}\right) = \frac{f\left(\mathbf{X}^{[J_k+l]}|H_1\right)}{f\left(\mathbf{X}^{[J_k+l]}|H_0\right)}, \tag{6.22}$$

where the form of the ratio emphasizes the fact that the detector uses the known parameter values k and l to choose the DWT shift $J_k + l$ that maximizes its performance.

Shift Variant GLRT Detector

We now extend the shift variant detector of Section 6.3.3 to the case of unknown parameters $\{A, k, l\}$. From the definition of the GLRT detector presented in (6.5) the likelihood ratio is $r\left(\mathbf{X}\right) = \max_{\{A,k,l\}} \frac{f\left(\mathbf{X}^{[0]}|A,k,l,H_1\right)}{f\left(\mathbf{X}^{[0]}|H_0\right)}$ which yields a sufficient test statistic

$$\lambda(\mathbf{X}) = \max_{\{A,k,l\}} \left[A \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[0]} S_{m,n}^{[0],\{k,l\}} - \frac{A^2}{2} \sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[0],\{k,l\}}\right)^2 \right]. \tag{6.23}$$

Upon substituting the maximum likelihood estimate

$$\hat{A} = \frac{\sum_{m,n} 2^{-\gamma m} X_{m,n}^{[0]} S_{m,n}^{[0],\{k,l\}}}{\sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[0],\{k,l\}}\right)^2}, \tag{6.24}$$

we can rewrite (6.23) in the more useful form

$$\lambda(\mathbf{X}) = \max_{\{k,l\}} \frac{\left(\sum_{m,n} 2^{-\gamma m} X_{m,n}^{[0]} S_{m,n}^{[0],\{k,l\}}\right)^2}{\mathcal{E}_k(l)}. \tag{6.25}$$

To demonstrate the shift variance of the detector, we examine the normalized distance between the expected values of $\lambda(\mathbf{X})$ under the two hypotheses H_0 and H_1. If this distance is shift variant, we say that the detector is also shift variant. We first note that $E[\lambda(\mathbf{X})|H_0]$ is independent of the signal and therefore we need only determine if $E[\lambda(\mathbf{X})|H_1]$ is shift variant. We define

$$g_{k,l} = \frac{1}{\sqrt{\mathcal{E}_k(l)}} \sum_{m,n} 2^{-\gamma m} X^{[0]}_{m,n} S^{[0],\{k,l\}}_{m,n} \tag{6.26}$$

so that $\lambda(\mathbf{X}) = \max_{\{k,l\}} g^2_{k,l}$. For hypothesis H_1, we assume that $X^{[0]}_{m,n} = A_0 S^{[0],\{k_0,l_0\}}_{m,n} + V^{[0]}_{m,n}$, where A_0, k_0 and l_0 are the true but unknown parameter values. For each index pair $\{k,l\}$, the random variable $g_{k,l}$ is Gaussian with mean

$$m_{k,l} = E[g_{k,l}] = \frac{A_0}{\sqrt{\mathcal{E}_k(l)}} \sum_{m,n} 2^{-\gamma m} S^{[0],\{k_0,l_0\}}_{m,n} S^{[0],\{k,l\}}_{m,n} \tag{6.27}$$

and variance σ^2_w. Furthermore, the random variables are dependent since there exist index pairs $\{k,l\}$ and $\{k',l'\}$ for which the covariance

$$cov(g_{k,l} g_{k',l'}) = \frac{\sigma^2_w}{\sqrt{\mathcal{E}_k(l)\mathcal{E}_{k'}(l')}} \sum_{m,n} 2^{-\gamma m} S^{[0],\{k,l\}}_{m,n} S^{[0],\{k',l'\}}_{m,n} \tag{6.28}$$

is not equal to zero. The random variables $g^2_{k,l}$ are also correlated, and each variable $g^2_{k,l}$ follows a noncentral chi-squared distribution with one degree of freedom (χ^2_1) and has mean $m^2_{k,l} + \sigma^2_w$ and variance $4m^2_{k,l}\sigma^2_w + 2\sigma^4_w$. As a result, $\lambda(\mathbf{X})$ is the maximum of a collection of correlated non-central χ^2_1 random variables with varying means and variances.

Obtaining an estimate of $E[\lambda(\mathbf{X})|H_1]$, through either straightforward calculations or Monte Carlo simulations, is at best a cumbersome task. As an approximation, we claim that the distribution of $\lambda(\mathbf{X})$ will be dominated by the χ^2_1 variable with the largest mean (and hence the largest variance as well). The accuracy of this approximation increases as the noise variance decreases, since as $\sigma^2_w \to 0$, we find that $E[\lambda(\mathbf{X})|H_1] \to \max_{\{k,l\}} m^2_{k,l}$. In addition we observe that $\max_{\{k,l\}} m^2_{k,l}$ serves as a lower bound on $E[\lambda(\mathbf{X})|H_1]$. We may therefore use $\max_{\{k,l\}} m^2_{k,l}$ to estimate the dependence of $E[\lambda(\mathbf{X})|H_1]$ on the signal delay.

Applying the Cauchy-Schwartz inequality, we find that

$$\begin{aligned}\left(\sum_{m,n} 2^{-\gamma m} S_{m,n}^{[0],\{k_0,l_0\}} S_{m,n}^{[0],\{k,l\}}\right)^2 &\leq \sum_{m,n} 2^{-\gamma m}\left(S_{m,n}^{[0],\{k_0,l_0\}}\right)^2 \sum_{m,n} 2^{-\gamma m}\left(S_{m,n}^{[0],\{k,l\}}\right)^2 \\ &= \mathcal{E}_{k_0}(l_0)\mathcal{E}_k(l), \end{aligned} \tag{6.29}$$

which enables us to write

$$m_{k,l}^2 \leq A_0^2 \mathcal{E}_{k_0}(l_0). \tag{6.30}$$

The upper bound is attained when $k = k_0$ and $l = l_0$, so that in fact $\max_{\{k,l\}} m_{k,l}^2 = A_0^2\mathcal{E}_{k_0}(l_0)$ and is a function of the assumed signal delay l_0. Thus, our estimate is

$$E[\lambda(\mathbf{X})|H_1] \approx \sigma_w^2 + \max_{k,l} m_{k,l}^2 = \sigma_w^2 + A_0^2\mathcal{E}_{k_0}(l_0). \tag{6.31}$$

To normalize this estimate we divide by the square root of the variance of $\lambda(\mathbf{X})$. We approximate the variance of $\lambda(\mathbf{X})$ by the the variance of the χ_1^2 variable with the largest mean and variance. The resulting performance index is

$$d \approx \frac{\sigma_w^2 + A_0^2\mathcal{E}_{k_0}(l_0)}{\sqrt{2\sigma_w^4 + 4A_0^2\mathcal{E}_{k_0}(l_0)\sigma_w^2}} \tag{6.32}$$

This performance index is shift variant. To see this, note that as σ_w^2 decreases in value, the performance index becomes $d = \frac{A_0\sqrt{\mathcal{E}_{k_0}(l_0)}}{2\sigma_w}$, which is clearly shift variant.

Shift Invariant GLRT Detector

To obtain a shift invariant detector we extend the concepts developed in Section 6.3.3 (e.g., Equation 6.22) to define a likelihood ratio

$$\begin{aligned} r\left(\tilde{\mathbf{X}}\right) &= \max_{\{A,k,l\}} \frac{f\left(\mathbf{X}^{[J_k+l]}|A,k,l,H_1\right)}{f\left(\mathbf{X}^{[J_k+l]}|H_0\right)} \\ &= \max_{\{A,k,l\}} \frac{\prod_{m,n}\exp\left(-\frac{\left(X_{m,n}^{[J_k+l]} - AS_{m,n}^{[J_k+l],\{k,l\}}\right)^2}{2\sigma_w^2 2^{\gamma m}}\right)}{\prod_{m,n}\exp\left(-\frac{\left(X_{m,n}^{[J_k+l]}\right)^2}{2\sigma_w^2 2^{\gamma m}}\right)} \end{aligned} \tag{6.33}$$

where $\tilde{\mathbf{X}} = \{\tilde{X}_{m,n},\ m,n \in \mathcal{Z}\}$ is the vector of UDWT observations and contains all possible DWT coefficients $X^{[J]}_{m,n},\ J \in \mathcal{Z}$. Note that the DWT shift $J_k + l$ is chosen to maximize the detector performance conditioned on the unknown parameters k and l. The likelihood ratio in (6.33) simplifies to yield a sufficient test statistic

$$\lambda(\tilde{\mathbf{X}}) = \max_{\{A,k,l\}} \left[A \sum_{m,n} 2^{-\gamma m} X^{[J_k+l]}_{m,n} S^{[J_k],\{k,0\}}_{m,n} - \frac{A^2}{2} \sum_{m,n} 2^{-\gamma m} \left(S^{[J_k],\{k,0\}}_{m,n} \right)^2 \right]. \tag{6.34}$$

Upon inserting the maximum likelihood estimate $\hat{A}$, we obtain the equivalent test statistic

$$\lambda(\tilde{\mathbf{X}}) = \max_{\{k,l\}} \frac{1}{\mathcal{E}_k} \left(\sum_{m,n} 2^{-\gamma m} X^{[J_k+l]}_{m,n} S^{[J_k],\{k,0\}}_{m,n} \right)^2. \tag{6.35}$$

Let us now demonstrate the shift invariance of the detector. We define

$$g_{k,l} = \frac{1}{\sqrt{\mathcal{E}_k}} \sum_{m,n} 2^{-\gamma m} X^{[J_k+l]}_{m,n} S^{[J_k],\{k,0\}}_{m,n} \tag{6.36}$$

such that $\lambda(\tilde{\mathbf{X}}) = \max_{\{k,l\}} g^2_{k,l}$. The test statistic $\lambda(\tilde{\mathbf{X}})$ is shift invariant if the set of random variables $G = \{g_{k,l}\ , k,l \in \mathcal{Z}\}$ is invariant with respect to the assumed signal delay for hypothesis H_1. We consider two signals $A_0 s_{k_0,l_0}(t)$ and $A_0 s_{k_0,l_1}(t)$ with signal delays of l_0 and l_1, respectively. For signal delay l_0 we have

$$g_{k,l} = \frac{1}{\sqrt{\mathcal{E}_k}} \sum_{m,n} 2^{-\gamma m} \left(A_0 S^{[J_k+l],\{k_0,l_0\}}_{m,n} + V^{[J_k+l]}_{m,n} \right) S^{[J_k+l],\{k,l\}}_{m,n}, \tag{6.37}$$

while for delay l_1 we define a corresponding variable

$$g_{k,l'} = \frac{1}{\sqrt{\mathcal{E}_k}} \sum_{m,n} 2^{-\gamma m} \left(A_0 S^{[J_k+l'],\{k_0,l_1\}}_{m,n} + V^{[J_k+l']}_{m,n} \right) S^{[J_k+l'],\{k,l'\}}_{m,n}. \tag{6.38}$$

We want to show that the sets $G = \{g_{k,l}\}$ and $G' = \{g_{k,l'}\}$ are identical. It is sufficient to show that for each l there is a unique l' such that $g_{k,l} = g_{k,l'}$. Choosing $l' = l + l_1 - l_0$, we find that the two sets are indeed identical. The detector is, therefore, shift invariant.

With the definition $s_{k,l}(t) = 2^{\gamma k/2} \xi_k(t)$, we can simplify the test statistic by noting that this definition results in $\mathcal{E}_k = \mathcal{E}_0$ for all k, where $\mathcal{E}_0$ is a constant (see Appendix A).

Using (6.12) and (6.15) to rewrite (6.34) as

$$\lambda(\tilde{\mathbf{X}}) = \max_{\{A,k,l\}} \left[A \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[J_k+l]} S_{m,n}^{[J_k],\{k,0\}} - \frac{A^2}{2} \mathcal{E}_0 \right], \tag{6.39}$$

we see that the maximization over k and l may be performed independently of the maximization over A. As a result, we may define a test statistic

$$\begin{aligned} \lambda(\tilde{\mathbf{X}}) &= \max_{\{k,l\}} \lambda(\tilde{\mathbf{X}}, k, l) \\ &= \max_{\{k,l\}} \left| \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[J_k+l]} S_{m,n}^{[J_k],\{k,0\}} \right| \qquad (6.40) \\ &= \max_{\{k,l\}} \left| \sum_{m,n} 2^{-\gamma m} \tilde{X}_{m,2^m n+J_k+l} \tilde{S}_{m,2^m n+J_k}^{\{k,0\}} \right|. \qquad (6.41) \end{aligned}$$

It also follows from (6.39) that if the sign of A is known to be non-negative, then the absolute value operation in the above equations may be omitted.

6.4 Implementation

6.4.1 Proposed Methods

Equations (6.40) and (6.41) represent two equivalent ways of computing the detection statistic. Equation (6.40) requires the calculation of 2^M different DWT shifts, and a straightforward implementation takes $O(N2^M)$ operations, where N is the length of the signal and M is the number of DWT analysis scales. It is well known, however, that the task can be performed in $O(NM)$ operations using, for example, a translation invariant wavelet transform (TIWT) [*Coifman and Donoho*, 1995]. Similarly, (6.41) requires the computation of a UDWT, also an $O(NM)$ operation [*Mallat*, 1991]. As shown in [*Coifman and Donoho*, 1995], a mapping exists between the TIWT and UDWT representations. In practice, we implement (6.41) which can be viewed as a pattern matching operation in the UDWT domain, where for each unknown scale the pattern is $\tilde{S}_{m,2^m n+J_k}^{\{k,0\}}$.

The pattern matching procedure in (6.41) requires $O\left(2N \sum_{k \in \mathcal{K}} P_k\right)$ operations, where $\mathcal{K}$ is the set of unknown scales and P_k is the number of non-zero coefficients in $\tilde{S}_{m,2^m n+J_k}^{\{k,0\}}$. As an example, when the signal model $\xi(t)$ is chosen to be a C_2 scaling function with scale $k = 7$ and the UDWT is performed with a C_2 wavelet over the range $m = 1$ to 10, we find

that $P_7 = 1499$. The contribution of most of these coefficients to the detection process is negligible. As a result, we can reduce the number of required coefficients by ranking them according to their contribution to $\mathcal{E}_k$ and selecting only the largest T_k coefficients. In the case where $\xi(t)$ is a C_2 scaling function and $\gamma = 1$, we find that the largest $T_k = 20$ coefficients account for 99.4 percent of the value of $\mathcal{E}_k$.

We can further reduce the number of computations required by observing that the most significant coefficients of the pattern $\tilde{S}^{\{k,0\}}_{m,2^m n+J_k}$ tend to be located near peaks in the UDWT domain. Figure 6.3 shows an example for both the C_2 scaling function and the Gaussian function. This is not a surprising result, since the shift J_k was chosen to maximize $\mathcal{E}_k$. In particular, we note that a number of the coefficients lie somewhere on the maximum peaks at each analysis scale m. We expect, therefore, that the locations in the UDWT domain where the local maxima propagate across scales are likely to correspond to scales and delays for which the detection statistic is maximized. Adapting the nomenclature of [*Alexandrescu et al.*, 1995], we refer to these locations as *transform ridges*. Instead of computing $\lambda(\tilde{\mathbf{X}}, k, l)$ for all possible values of k and l, we constrain the computation to a subset of values that correspond to maxima ridges. If we constrain the search for transform ridges to analysis scales that are much coarser than the finest analysis scales (e.g. $m > 5$), then the number of maxima that need to be considered is typically 1 to 2 orders of magnitude less than the length N of the original signal. As a result, this preselection of locations can typically reduce the computational complexity of the pattern matching procedure by at least an order-of-magnitude.

To place these observations in context, we note that the importance of transform ridges for detecting and characterizing the local regularity of singularities in signals is well known [*Alexandrescu et al.*, 1995; *Mallat and Hwang*, 1992]. In addition, it has been shown that the wavelet transform maxima represent the important features of a signal and can be used for signal reconstruction [*Cvetković and Vetterli*, 1995; *Mallat and Hwang*, 1992]. Our work shows that the transform maxima can also provide information on the parameter values that maximize a GLRT detection statistic.

The preceding discussion leads us to propose the detection methods described below. To simplify the development, we have assumed, without loss of generality, that $A > 0$, which enables us to omit the absolute value operation in the test statistics.

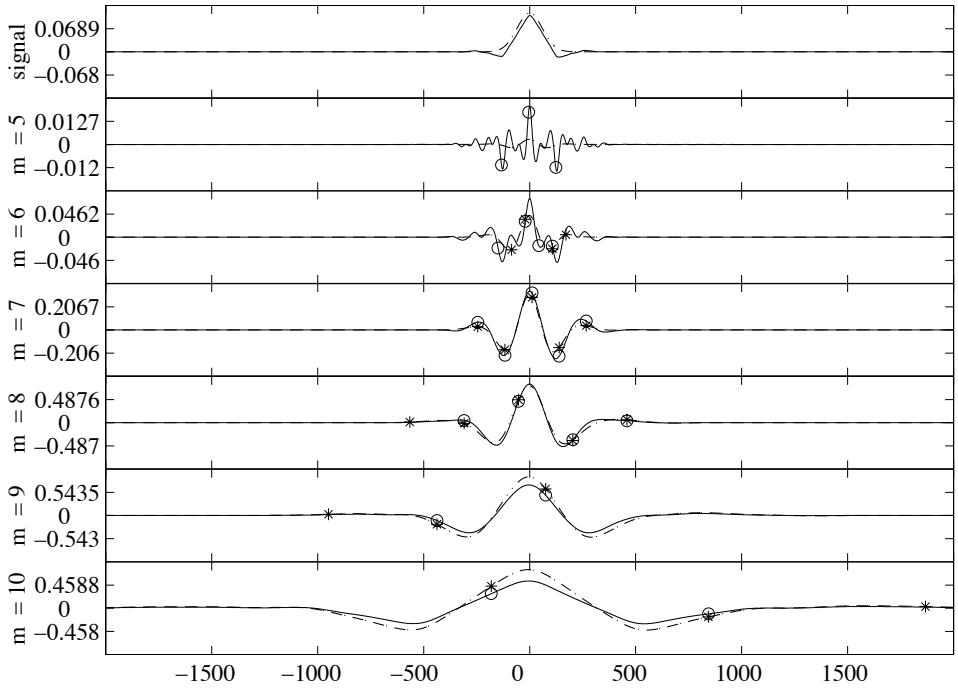


Figure 6.3: The UDWT of a C_2 scaling function (solid line) and the UDWT of a Gaussian function (dashed-dotted line) is shown for scales $m = 5$ to 10. The signals are shown in the top subfigure ($k = 7$ for both cases). The largest 20 coefficients of the pattern $\tilde{S}^{\{k,0\}}_{m,2^m n+J_k}$, as ranked by contribution to $\mathcal{E}_k$, with $\gamma = 1$, are marked with 'o' and '$*$' symbols for the scaling function and Gaussian function, respectively. Note that these coefficients lie near the peaks in the UDWT domain.

Method A. *Baseline GLRT method.* Compute

$$\lambda(\tilde{\mathbf{X}}, k, l) = \sum_{m,n} 2^{-\gamma m} \check{X}_{m,2^m n + J_k + l} \tilde{S}^{\{k,0\}}_{m,2^m n + J_k} \tag{6.42}$$

for all values of k and l and find the maximum.

Method B. *Preselection of Scales and Delays using Transform Ridges.* In this method we use a 'ridge-finding' algorithm to generate estimates of both the scale and delay of the transient. The qth estimate is referred to as an ordered pair $\{\hat{k}_q, \hat{l}_q\}$. We compute the detection statistic $\lambda(\tilde{\mathbf{X}}) = \max_q \lambda(\tilde{\mathbf{X}}, \hat{k}_q, \hat{l}_q)$.

Method C. *Preselection of Delays using Transform Ridges.* In this method we use a ridge-finding algorithm to estimate the delays but not the scales. The detection statistic is $\lambda(\tilde{\mathbf{X}}) = \max_{k,q} \lambda(\tilde{\mathbf{X}}, k, \hat{l}_q)$.

Method D. *Preselection of Delays using Local Maxima at Selected Scales.* In this method we use the positions of the local maxima over a range of analysis scales as delay estimates. The assumptions underlying this method are: (1) most local maxima belong to a maxima ridge and (2) inclusion of local maxima not associated with ridges does not affect the detector performance. The form of the detection statistic is the same as for Method C.

6.4.2 Ridge-Finding Algorithm

There are a number of possible implementations of the ridge-finding algorithm. We describe the implementation that is used in the simulations, while mentioning alternative solutions along the way. Our algorithm is based in part on a ridge finding program contained in the WAVELAB software package [*Buckheit and Donoho*, 1995].

We perform the ridge-finding using a *peak aligned* version of the UDWT, which we obtain by applying a scale dependent time shift D_m to the coefficients of the UDWT. That is, $\hat{X}_{m,n} = \tilde{X}_{m,n+D_m}$, where $\hat{X}_{m,n}$ represents the peak aligned UDWT coefficients. The shifts D_m are chosen such that the peaks of $\hat{X}_{m,n}$ line up across scales when the projection of $x(t)$ onto V_0, the finest MRA subspace of interest, is a Kronecker delta function. This choice of shifts has the effect of approximately aligning the peaks of $\hat{X}_{m,n}$ when other transient models are used, and it facilitates the ridge-finding procedure.

The ridge-finding procedure is limited to a range of R analysis scales

$$\mathcal{M} = \{m_1, \ldots, m_i, \ldots, m_R\},$$

where the range is based on *a priori* knowledge of the unknown transient scales and $m_i = m_1 + i - 1$. As an example, for a C_2 scaling function with unknown scale $k \in \{k_1, \ldots, k_i \ldots, k_{N_k}\}$, where N_k is the number of unknown scales and $k_i = k_1 + i - 1$, we use the range limits $m_1 = k_1$ and $m_R = k_{N_k} + 1$. As can be verified numerically, the UDWT of a scaling function at scale k has its maximum value at analysis scale $m = k+1$. We are thus limiting the ridge-finding to analysis scales where we expect significant peaks in the UDWT domain.

For each $m \in \mathcal{M}$ we find the local maxima of $\hat{X}_{m,n}$. We refer to each local maxima by the ordered pair $L_{m,i} = \{n_{m,i}, p_{m,i}\}$, where $n_{m,i}$ and $p_{m,i}$ are the location and amplitude, respectively, and i ranges from 1 to N_m, the number of local maxima at scale m. Each local maxima satisfies the condition $\hat{X}_{m,n_{m,i}} > \hat{X}_{m,n_{m,i}+1}$ and $\hat{X}_{m,n_{m,i}} > \hat{X}_{m,n_{m,i}-1}$. We use the local maxima of $\hat{X}_{m,n}$ because we have assumed that $A > 0$. In the case where the sign of A is not known, we would use the local maxima of $|\hat{X}_{m,n}|$ since the quantity we would need to maximize is then $|\lambda(\tilde{\mathbf{X}}, k, l)|$.

We begin the ridge-finding algorithm at the coarsest analysis scale m_R in the range $\mathcal{M}$. For each local maxima $L_{m_R,i}$, we look for the maxima at the next finest scale m_{R-1} whose location is closest to the current ridge location $n_{m_R,i}$. Denote the location of this maxima as $n_{m_{R-1},i_{\min}}$. If the absolute difference $|n_{m_R,i} - n_{m_{R-1},i_{\min}}|$ between the locations is less than a threshold parameter ϵ, we add the maxima $L_{m_{R-1},i_{\min}}$ to the current ridge and also delete this maxima from the list of maxima at scale m_{R-1}. We then set the current ridge location to $n_{m_{R-1},i_{\min}}$ and repeat the process at the next finest scale m_{R-2}. The ridge-finding procedure stops when we reach the finest scale of interest m_1 or when we cannot find a maxima location that is within ϵ of the current ridge location. The length of the ridge is equal to the number of included maxima scales. If the length is greater than a minimum ridge length η, we add the ridge to our collection of valid ridges.

Once we have found all ridges that begin at scale m_R, we repeat the process and look for all ridges that begin with local maxima at scale m_{R-1}. Note that, at this point in the procedure, the population of maxima at scale m_{R-1} will no longer include those maxima that belong to ridges with roots at scale m_R. We then repeat the operation at increasingly

finer scales. We stop when we have found all ridges that begin at scale m_η, since no ridges of length η can start at any scale $m < m_\eta$.

At the completion of the ridge-finding algorithm, we have a collection of ridges, where each ridge consists of a set of local maxima. For method B, we use the scales and locations of these local maxima to form the estimates $\hat{k}_q$ and $\hat{l}_q$. For method C, we use only the locations of these local maxima to form the estimates $\hat{l}_q$. As an example of an alternative implementation, we may consider reducing the number of estimates produced by using only the normalized peaks of each ridge. That is, for each ridge we pick the local maxima with the greatest normalized amplitude $2^{-\gamma m/2} p_{m,i}$, and use the location and scale of that maxima to form an estimate. While this implementation reduces the number of pattern matching operations required, we have found that in some cases it has an adverse effect on detector performance.

6.4.3 Computational Complexity

We now examine the computational complexity of the proposed methods. Initially, we consider only the computations required for the pattern matching procedure, since the computational cost of the UDWT is the same for all methods. We assume that $P_k = P$ and $P \gg 1$, i.e., the number of coefficients in the matching patterns is independent of scale and is much greater than 1. Method A takes order $2N_k NP$ operations where N_k is the number of unknown scales. The computational costs for methods B and C are the sum of (1) a local maxima finding cost term, (2) a ridge-finding cost term, and (3) a pattern matching cost term. The local maxima finding cost is $O(NR)$ where R is the number of analysis scales used for ridge-finding. To derive an approximate ridge-finding cost term, we make the following assumptions : (1) there are roughly $2^{-m}N$ maxima at each analysis scale m; (2) half of the maxima at each scale serve as the root of a ridge that continues on to the finest scale in $\mathcal{M}$; (3) the deletion of ridge maxima from the list of maxima at each scale can be ignored; and (4) since the list of maxima at each scale is ordered, the search for the closest maxima is an $O(\log_2 n)$ operation, where n is the length of the list. With these assumptions, which are borne out by simulations, we write the ridge-finding cost C_r as

$$C_r = \frac{N_{m_R}}{2} \sum_{j=1}^{R-1} 2^{(j-1)} \sum_{i=j}^{R-1} \log_2(2^i N_{m_R}), \tag{6.43}$$

where $N_{m_R} = 2^{-m_R} N$. When $R \gg 1$, we can approximate the cost as

$$C_r \approx \frac{N}{2} \left[2^{-(m_R - R)} \left(\log_2 N - m_R \right) + 2^{-m_R} \sum_{i=1}^{R-1} i 2^i \right] \tag{6.44}$$

For $(m_R - R) \gg 1$ and $(\log_2(N) - m_R)$ small compared to $2^{(m_R - R)}$, we find that $C_r \ll N$. As a result, the ridge-finding cost is much less than the local maxima finding cost and can generally be neglected. For example, with $N = 16384$, $m_R = 10$, and $R = 5$, we have $C_r \approx 1800$.

The pattern matching costs depend on the number of maxima associated with ridges. As an upper bound, we may assume that all maxima belong to a ridge. The matching costs for methods B and C are then $O(2^{(R-m_R)} NP)$ and $O(2^{(R-m_R)} NPN_k)$, respectively. Adding the maxima finding costs and neglecting the ridge-finding costs, we find that the total costs for methods B and C are $O\left(N(R + 2^{(R-m_R)} P)\right)$ and $O\left(N(R + 2^{(R-m_R)} PN_k)\right)$, respectively. Since ridge-finding costs are negligible, the cost of method D is the same order as that of method C.

The costs of methods B, C, and D will be much less than that of method A when the following condition is satisfied: $(R + 2^{(R-m_R)} PN_k) \ll 2N_k P$. For example, with the values $N = 16384$, $P = 20$, $N_k = 3$, $R = 5$, and $m_R = 10$, we find that ratios of the cost of method A to the costs of methods B and C are 21.3 and 17.5, respectively. Thus, preselection of delays and scales can reduce by an order-of-magnitude the complexity of the pattern matching. In Section 6.5 we show that this reduction can usually be achieved with no loss of performance.

To gain some perspective on this cost reduction, we must also consider the computational complexity of the UDWT. This cost is $O(4NML)$ where L is the length of the filter used in the iterated filter bank implementation [*Vetterli and Kovačević*, 1995]. For a C_2 wavelet analysis $L = 12$. The total cost of method A is of order $4NML + 2NN_kP$. With $M = 10$ and the parameter values used above, we find that $2NN_kP$ represents 20% of the total cost. Thus, for this example, decreasing the pattern matching cost by an order-of-magnitude reduces the total cost of the detection process by about 20%. We can readily choose parameter values in which the reduction is more significant, e.g. by increasing N_k or decreasing L.

6.4.4 Variations on a Theme

We note that there are many variations of the proposed methods. For example, we can further reduce the pattern matching operations in method D by sorting the maxima by amplitude and using only the delay estimates that correspond to the largest Q maxima. Assuming a total of n maxima and a $n \log_2 n$ sort algorithm, the number of operations is $n \log_2 n + 2QN_kP$ as compared to $2nN_kP$ when all the maxima are used. For the simulation parameters described below, the number of maxima n used in method D is approximately 500, and a value of $Q = 50$ is sufficient to avoid any degradation of the detector performance. The number of pattern matching operations is decreased by a factor of about 6 with respect to method D. We can consider also just using the maxima of the untransformed signal to form estimates of the delays. For $\gamma = 1$, the number of maxima is on the order of $N/3$. With the simulation parameters used in Section 6.5, we have found that using the largest 5% of the maxima provides performance equivalent to that of method A. For values of N such that $\log_2 N$ is roughly the same order as R, the cost of such an algorithm is comparable with methods B, C and D. As the signal-to-noise ratio increases and the signal peaks rise above the noise floor, the largest local maxima of the signal will provide increasingly better estimates of the transient delay.

6.5 Simulation Results

We used Monte Carlo simulations to obtain the receiver operating characteristics (ROC) of the detectors described above. Each simulation consisted of 200 independent trials. We generated $1/f$ noise using the method described in [*Corsini and Saletti*, 1987]. Each $1/f$ noise sample path contained $N = 16384$ samples simulated with the following noise model parameters: $\gamma = 1$, pole frequencies logarithmically spaced over the normalized frequency range 2.5×10^{-5} to 0.125 with 1.5 poles per decade, and $\sigma_v^2 = 1.6 \times 10^{-5}$.

We employed three different transient signal models: a C_2 scaling function, a Gaussian function $g(t) = \alpha e^{-at^2}$, and a two-sided exponential function $e(t) = \beta e^{-b|t|}$. The exponent factors a and b were chosen such that temporal widths (as measured by the variance of the squared modulus [*Bracewell*, 1986]) were the same for all signal models. In addition, the leading scale factors α and β were chosen so that $\mathcal{E}_k$ was equivalent for all signal models. We also defined the functions $g_s(t) = (\sqrt{2})^{(\gamma-1)} g(\sqrt{2}t)$ and $e_s(t) = (\sqrt{2})^{(\gamma-1)} e(\sqrt{2}t)$ for use in examining the effects of scale mismatch. For each signal type, we used the top 20

coefficients of the pattern $\tilde{S}^{\{k,0\}}_{m,2^m n+J_k}$, i.e., $P_k = P = 20$. Examples of the signal models and the noise sample path are provided in Figures 6.4 and 6.5.

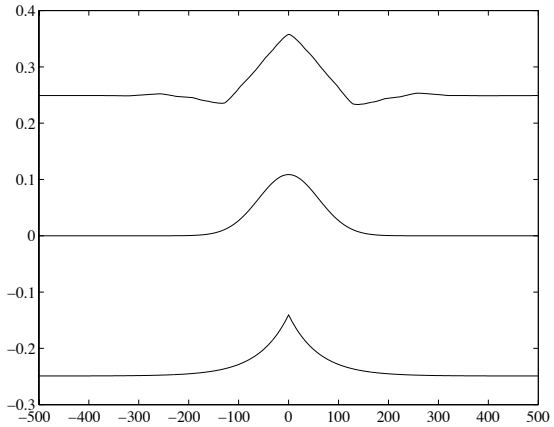


Figure 6.4: Transient signal models. From top to bottom: C_2 scaling function, Gaussian, Two-sided Exponential.

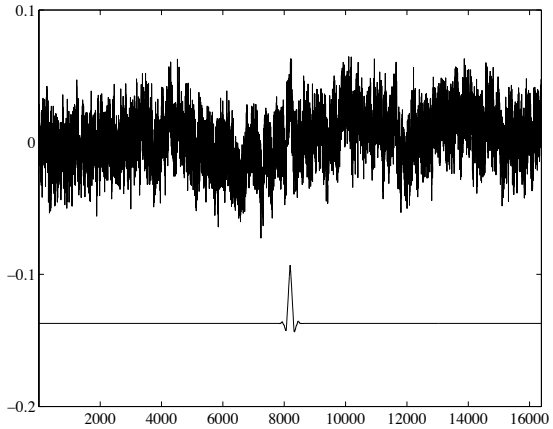


Figure 6.5: Example of transient signal embedded in $1/f$ noise for $\gamma = 1$ and $d = 5$. The noise-free transient is a C_2 scaling function at scale $k = 7$ and is shown in the lower waveform.

We used the performance index $d = \frac{A\sqrt{\mathcal{E}_k}}{\sigma_w}$ as a convenient measure of the signal-to-noise ratio of each simulation. It should be noted that, for a GLRT detector, d no longer represents the normalized distance between the distributions under the two detection hypotheses.

In the examples discussed below, we computed the UDWT using a C_2 wavelet for analysis scales $m = 1$ to 10. The ridge-finding algorithms tracked peaks over analysis scales 6 to 10, with a search width $\epsilon = 25$ and minimum ridge length $\eta = 2$. For methods A, C and D, the maximization over k was performed over the set of unknown scales $\mathcal{K} = \{6, 7, 8\}$. For method A the maximization over l was performed over the range $[1, \ldots, N]$.

Example 1. *Comparison of Methods.* In Figure 6.6 we show the ROC curves for detector methods A through D and the three transient models described above. We used input scales ranging from $k = 6$ to $k = 8$ and performance indices of $d = 4, 5$. The performance of methods C and D is equivalent to that of A for all signals models and input scales. The performance of method B is slightly worse for some cases, indicating

errors in the estimation of the scale of the underlying transient from the scale of the transform maxima.

Example 2. *Effects of Signal Model Mismatch.* In panel (a) of Figure 6.7 we show the ROC curves for a C_2 signal model detected with method A using C_2 UDWT signal patterns, and for a $g(t)$ signal model detected with methods A and C, also with C_2 patterns. Curves are shown for $k = 7$ and $d = 4, 5$. Panel (b) shows the corresponding ROC curves for the $e(t)$ signal model. The plots show that the detector is relatively insensitive to small errors in the assumed signal model. This result reflects the fact that $\mathcal{E}_k$ exhibits some invariance with respect to the assumed signal model. As we can see from Figure 6.3, the UDWT patterns for similar signal models exhibit a great deal of overlap.

Example 3. *Effects of Signal Scale Mismatch.* In panel (c) of Figure 6.7 we show the ROC curve for a $g(t)$ signal model detected with method A, and the ROC curves for a $g_s(t)$ signal model detected with methods A and C. In all three cases, the matching pattern was based on the UDWT of $g(t)$. Curves are shown for $k = 7$ and $d = 4, 5$. Panel (d) shows the corresponding ROC curves for the $e(t)$ and $e_s(t)$ signal models. Note that both $g_s(t)$ and $e_s(t)$ with scale $k = 7$ have an effective scale of 6.5. We are therefore examining the performance of the detector when the actual scale 6.5 lies between the assumed signal scales 6 and 7. The plots show that the detector is fairly robust with respect to the scale mismatch.

Example 4. *Detector Scale Discrimination.* Here we examine the ability of the detector to discriminate against signals with scales outside our assumed scale range $\mathcal{K} = \{6, 7, 8\}$. In panel (a) of Figure 6.8 we show ROC curves for a Gaussian input signal with scales of 4, 5, 5.5 ($g_s(t)$ with $k = 6$), and 6. In panel (b) we show ROC curves for a C_2 input signal with scales of 4, 5, and 6. For both cases, we find that by scale $k = 4$ the detector performance has approached that of an energy detector [*Kay*, 1998], which assumes no *a priori* information about the signal.

All simulations were performed using the MATLAB (version 5.1, The Mathworks Inc., Natick, MA) mathematical software package. We also made use of MATLAB routines from the WAVELAB (version .701) software system available from D. Donoho's group at Stanford University [*Coifman and Donoho*, 1995; *Buckheit and Donoho*, 1995].

6.6 Conclusion

We have proposed a GLRT detector for transients in $1/f$ noise by making use of the approximate whitening properties of the DWT. The detector is shift invariant and is implemented with a pattern matching operation in the UDWT domain. The complexity of the pattern matching procedure can be reduced by at least an order-of-magnitude by using the UDWT local maxima to identify scales and delays that are most likely to maximize the detection statistic. The reduction can be achieved with little or no performance loss. Finally, the proposed detector is robust with respect to signal model and scale mismatch and discriminates against signals with scales outside the desired detection range.

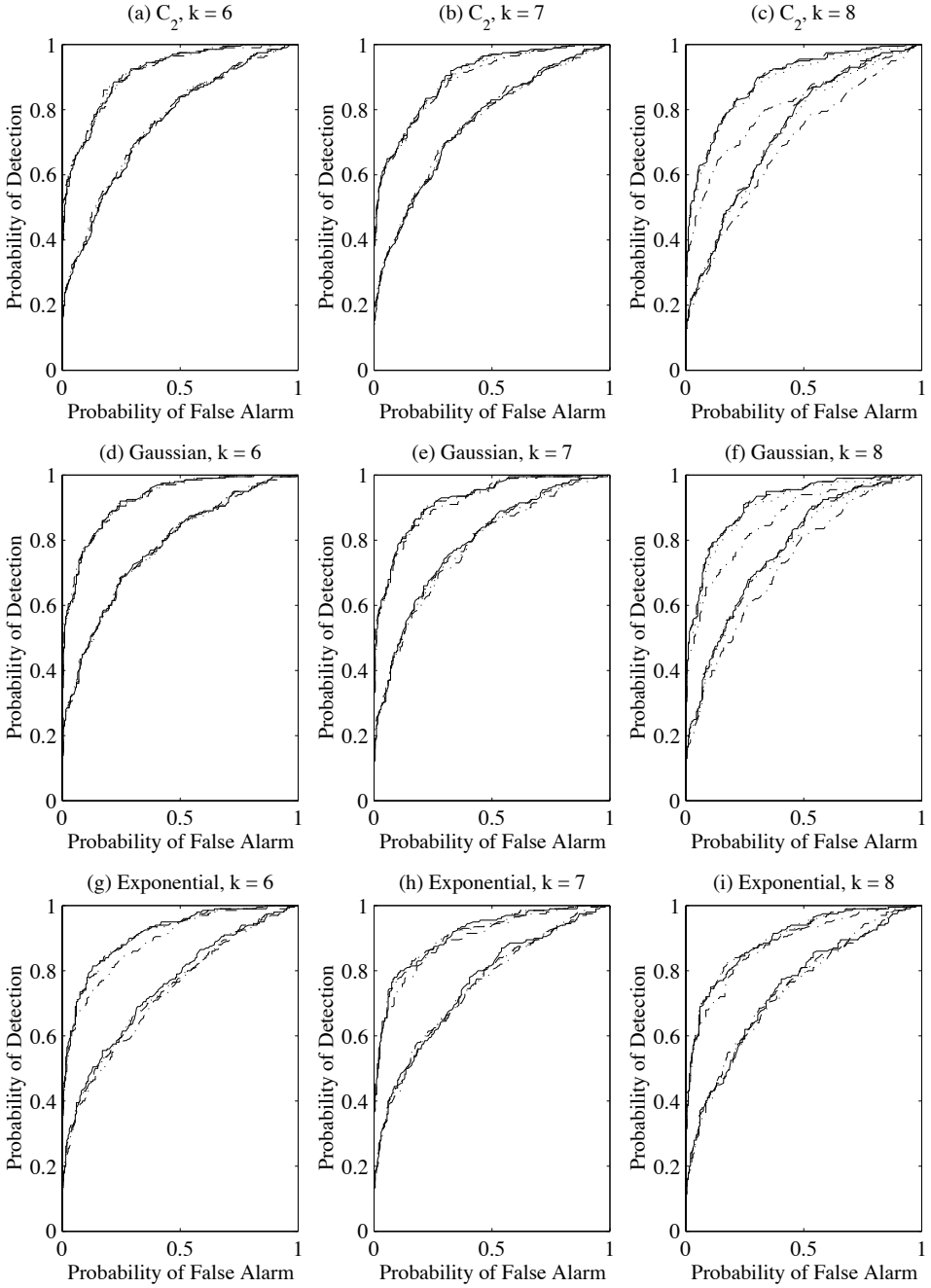


Figure 6.6: ROC curves for methods A (solid line), B (dashed-dotted line), C (dashed line), and D (dotted line). Panels (a)-(c) show ROC curves for the C_2 scaling function with scales k ranging from 6 to 8. ROC curves for the Gaussian function and the two-sided exponential are shown in panels (d)-(f) and (g)-(i), respectively. In each case the upper set and lower set of curves correspond to $d = 5$ and $d = 4$, respectively.

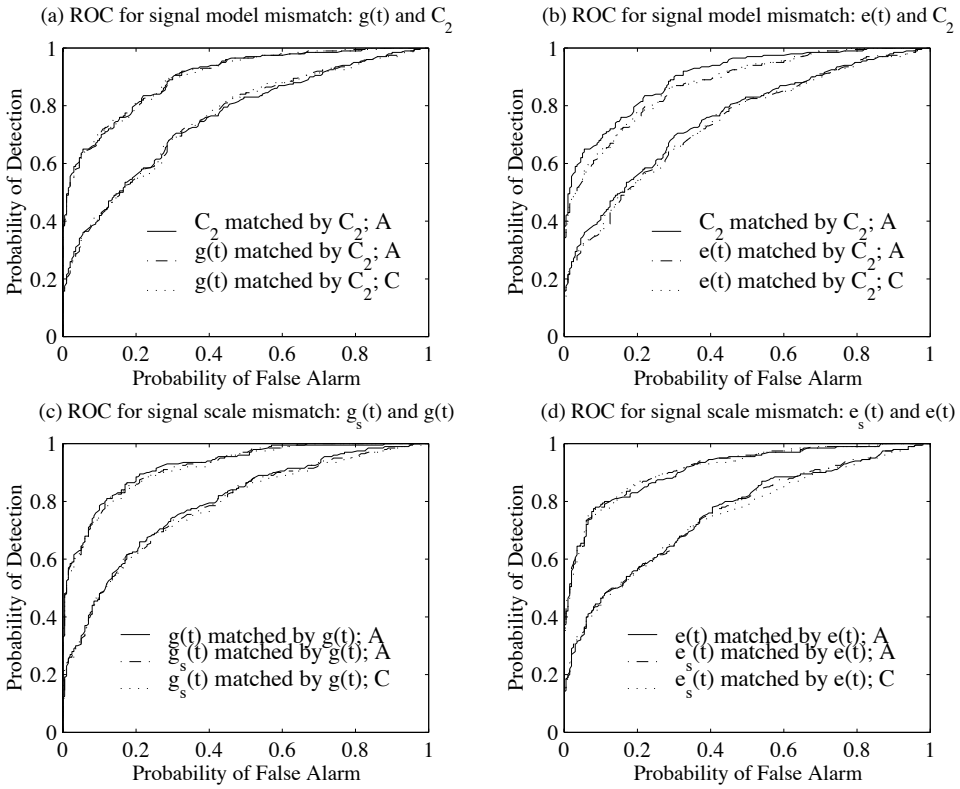


Figure 6.7: ROC curves showing the robustness of the detector with respect to signal model mismatch in panels (a) and (b) and scale mismatch in panels (c) and (d). For each case, the upper set and lower set of curves correspond to $d = 5$ and $d = 4$, respectively. See text for a more detailed explanation.

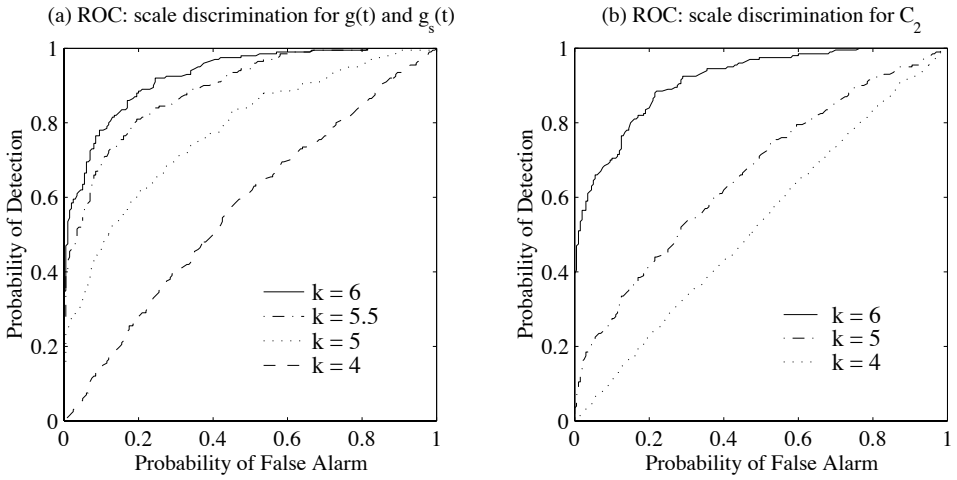


Figure 6.8: ROC curves showing the scale discrimination of the detector for (a) Gaussian and (b) C_2 scaling function at various scales.

Chapter 7

Identification and Removal of BART Transients

7.1 Introduction

In this chapter we develop in detail the framework outlined in Chapter 5 for the detection and removal of BART related transients. We begin in Section 7.2 by considering the detection and removal of transients from a single channel of the array. In this case, the detection criteria depend solely on the temporal characteristics of the transients, and in some instances the method may falsely detect natural transients as BART transients. In order to address this shortcoming, we make use of the polarization of the transients to constrain the detection process. We describe the resultant multi-channel approach in Section 7.3 and show that it improves our ability to distinguish BART transients from the natural background and can reduce the number of false detections.

7.2 Single Channel Case

7.2.1 Detection of Multiple Transients

In Chapter 6 we presented a method for the detection of a single transient with known transient shape and unknown amplitude, scale, and delay. In the BART signal model proposed in Chapter 5 there are multiple transients on each channel. As a result, we need to extend our detection method to the case of multiple transients. Taking a straightforward

approach, we partition the observed time series $x(t)$ into a sequence of overlapping sections where T is the length of each section and P is the percent of overlap between adjacent sections. We then assume that there is at most one transient per section. In the examples that follow, we use $T = 20$ seconds and $P = 50$ percent. Since the transient durations are roughly 4 to 30 seconds, this is a reasonable choice for the spacing of subsequent detection operations.

With the above partitioning in mind, we begin the detection process by first computing $\tilde{X}_{m,n}$, the UDWT of the entire time series, and then computing

$$\lambda(\tilde{\mathbf{X}}, k, l) = \left| \sum_{m,n} 2^{-\gamma m} \tilde{X}_{m,2^m n + J_k + l} \tilde{S}^{\{k,0\}}_{m,2^m n + J_k} \right| \tag{7.1}$$

for all possible values of scale k and delays l. This expression is identical to (6.41) in Chapter 6, where the meaning of the various terms is defined. For the qth section, the detection statistic is

$$\lambda_q(\tilde{\mathbf{X}}) = \max_{k \in \mathcal{K}, l \in L_q} \lambda(\tilde{\mathbf{X}}, k, l) \tag{7.2}$$

where L_q is the set of delay values within the qth section and $\mathcal{K} = \{6, 7, 8, 9\}$, as was first defined in Chapter 5.

For each section, maximum likelihood estimates of k and l are

$$\{\hat{k}_q, \hat{l}_q\} = \arg \max_{\{k \in \mathcal{K}, l \in L_q\}} \lambda(\tilde{\mathbf{X}}, k, l). \tag{7.3}$$

The estimate $\hat{l}$ of the transient delay is used in the transient removal process described in the next section. In practice, the quality of the estimate can be quite poor due to the presence of signal model mismatch errors and deviations of the actual noise process from an ideal $1/f$ noise process. In Chapter 6, it was shown that the local maxima of the UDWT provide good estimates of the delays l which maximize $\lambda(\tilde{\mathbf{X}}, k, l)$. Applying this observation to the geomagnetic data set, we find that delay estimates based on the local maxima of UDWT are more robust than the maximum likelihood estimates. As a result, it is useful to define a modified detection statistic

$$\lambda_q(\tilde{\mathbf{X}}) = \max_{k \in \mathcal{K}} \lambda(\tilde{\mathbf{X}}, k, \hat{l}_q), \tag{7.4}$$

where

$$\hat{l}_q = \arg\max_{l \in L_q} \left(\max_{(m-1) \in \mathcal{K}} |\hat{X}_{m,l}| \right). \tag{7.5}$$

Recall that $\hat{X}_{m,l}$ is the peak-aligned UDWT defined in Section 6.4.2. The modified statistic is used in the examples that follow.

The next step is to compare $\lambda_q(\tilde{\mathbf{X}})$ to a threshold value that depends on the noise exponent γ, the wavelet coefficient variance σ_w^2 (defined in Chapter 3 after Equation 3.5), the maximum likelihood scale estimate $\hat{k}_q$, and $A_{\min}$, which is the minimum amplitude that we wish to detect. In addition, the threshold is chosen to obtain a desired trade-off between the probability of false alarm (PFA) and the probability of detection (PD). Monte Carlo simulations (see Chapter 6) are used to calculate ROC curves and determine PFA and PD as a function of threshold level. In general the threshold is set to obtain a PFA of 0.10 or less and a PD of 0.90 or greater.

Before proceeding, we describe briefly the process for estimating γ and σ_w^2 from the discrete wavelet transforms of the remote reference data. As was shown in Chapter 3, the geomagnetic background exhibits the characteristics of a piecewise $1/f$ process, and two distinct γ values are needed to describe the wavelet coefficient variances. The detection method developed in Chapter 6 assumes, however, that a single value of γ describes the variances. To adapt the method to the observed data, we determine the parameters σ_w^2 and γ that best fit (in the least squares sense) the observed variances over a specified range of analysis scales. That is,

$$\{\sigma_w^2, \gamma\} = \arg\min_{\{\sigma^2, \gamma'\}} \sum_m \left(\log_2 \sigma_m^2 - \log_2 \sigma^2 - \gamma' m \right)^2 \tag{7.6}$$

where σ_m^2 is the estimated variance at scale m. The least squares fit is performed for scales $m = 4$ to 11, since these are the only scales used in the detection process. Recall that only the largest 20 coefficients as ranked by contribution to $\mathcal{E}_k$ are used for pattern matching (see, for example, Figure 6.3). Next, the smallest constant β such that $\beta \sigma_w^2 2^{\gamma m} > \sigma_m^2$ is found, such that the wavelet coefficient variances for the $1/f$ process with parameters $\beta\sigma_w^2$ and γ tightly upper bound the estimated variances. Figure 7.1 illustrates an example of this procedure. The least squares fit captures the gross features of the variance progression for the scales of interest. The parameters $\beta\sigma_w^2$ and γ are used for the Monte Carlo simulations and threshold selection. The resultant detector design is conservative since

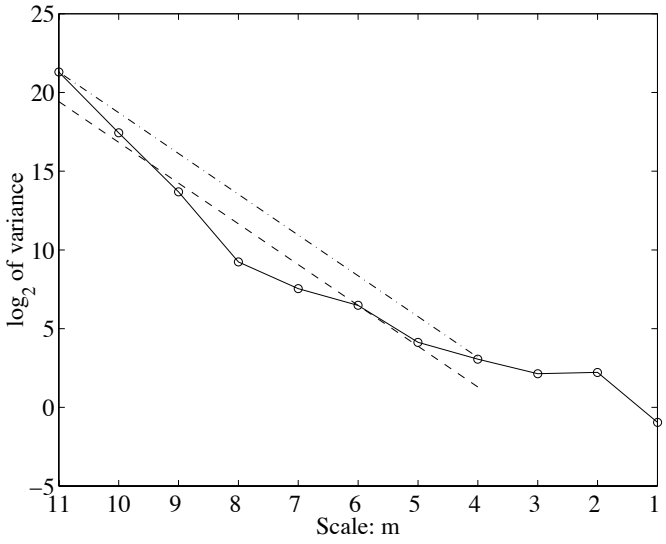


Figure 7.1: Plot of wavelet coefficient variances for 30 minutes of data acquired on SAO N-S channel at 11:00 UT, August 11, 1998. Dashed line represents least squares fit to variances for scales $m = 4$ to 11. Dash-dot line is the least squares fit multiplied by a constant chosen so that the line tightly upper bounds the variances.

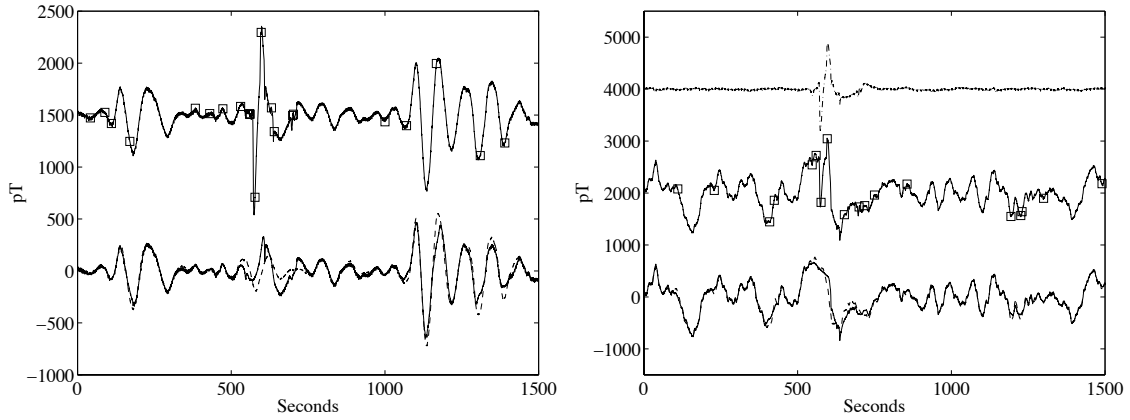
the assumed variances are higher than the estimated variances.

The detector developed in Chapter 6 assumes that the DWT coefficients are uncorrelated and characterized by a joint normal distribution. As shown in Chapter 3, these assumptions are only partially justified — the DWT coefficients are normally distributed for $m \geq 7$, but that there are noticeable departures from normality for $m < 7$. Because the pattern matching operation primarily uses coefficients at scales $m \geq 7$ (see Figure 6.3), we can as a first order approximation ignore the deviation from normality. For the most part the DWT coefficients are uncorrelated, although a few significant correlation values do exist, and these lead to differences between the observed and theoretical performance of the detector.

Panels (a) and (b) of Figure 7.2 provide examples of the detection method applied to time series from the MPK N-S channel at 10:30 UT on August 11, 1998. Panel (a) shows an isolated set of transients in the presence of elevated geomagnetic background signals. An estimate of the geomagnetic background using the horizontal channels from SAO is also shown. In panel (b), we have subtracted out the natural background and added in

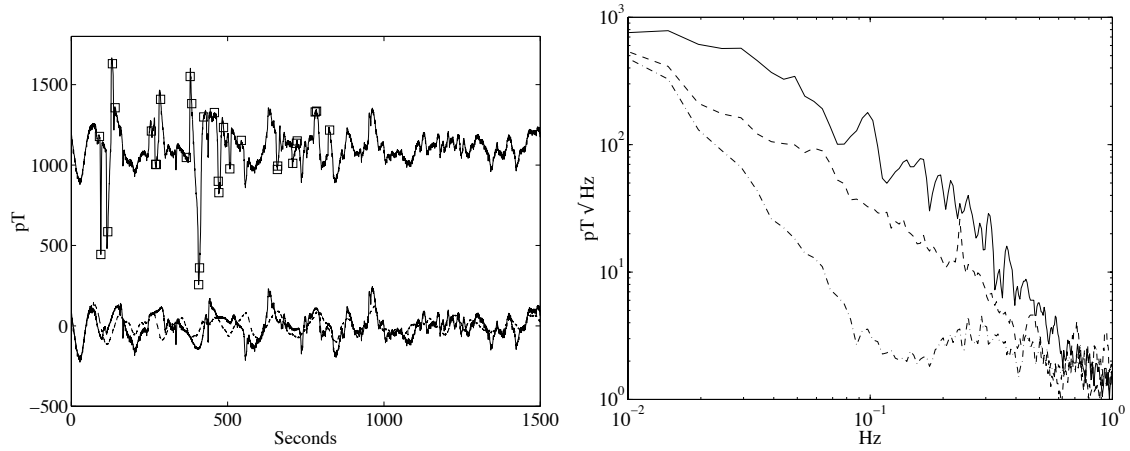
synthetic $1/f$ noise with parameters $\beta\sigma_w^2$ and γ determined using the procedure described above. To determine the detector thresholds, a value of $A_{\min} = 150$ pT is assumed, which is consistent with the amplitude histograms presented in Chapter 3. The detector performs reasonably well at identifying the significant BART transients, while ignoring waveforms with similar time scales and amplitudes. The false alarm rates are approximately 8.5% and 7.1% for panels (a) and (b), respectively.

Figures 7.3(a) and 7.4(a) show data from the MPK N-S channel at 11:00 UT on August 11, 1998. These time series represent a more typical situation in which there are many BART transients. In Figure 7.3(a) we show the raw time series, while in Figure 7.4(a) we have subtracted the background estimate and added in synthetic $1/f$ noise. For both cases, we detect the significant transients while ignoring features in the background (either natural or synthetic) that might otherwise be interpreted as BART transients.



(a) (b)

Figure 7.2: Example of transient detection and removal for data acquired on MPK N-S channel at 10:30 UT, August 11, 1998. Panel (a): MPK time series (upper solid) with estimated transient locations marked with squares; MPK time series (lower solid) after transient removal; and background estimate (dash) using data from SAO. Panel (b): MPK time series (upper dash-dot) with the background estimate removed; modified time series (middle solid) consisting of MPK time series with estimate removed and synthetic $1/f$ noise added — detection locations are marked with squares; modified time series (lower solid) after transient removal; and synthetic $1/f$ noise (dash).



(a) (b)

Figure 7.3: Transient detection and removal for data acquired on MPK N-S channel at 11:00 UT, August 11, 1998. Panel (a): MPK time series (upper solid) with estimated transient locations marked with squares; MPK time series after transient removal (lower solid); and background estimate (dash) using data from SAO. Panel (b): Power spectra of the MPK series (solid), the MPK time series after transient removal (dash), and the background estimate (dash-dot).

7.2.2 Removal of Transients

To remove the detected transients, we first form an estimate of the transient waveforms. The estimation process consists of two steps: neighborhood selection and hard thresholding of the UDWT coefficients. In the neighborhood selection step, only those UDWT coefficients in the neighborhoods of the estimated delays are used to form the estimate. To make this notion precise, we make use again of the peak-aligned UDWT $\hat{X}_{m,n}$ defined in Section 6.4.2. Neighborhood selection obtains, at each analysis scale m, the expression

$$\hat{X}'_{m,n} = \hat{X}_{m,n} \times \sum_j \sqcap \left(\frac{n - \hat{l}_j}{W} \right), \tag{7.7}$$

where

$$\sqcap(n) = \begin{cases} 1 & \text{for } |n| < 1/2 \\ 0 & \text{otherwise} \end{cases}, \tag{7.8}$$

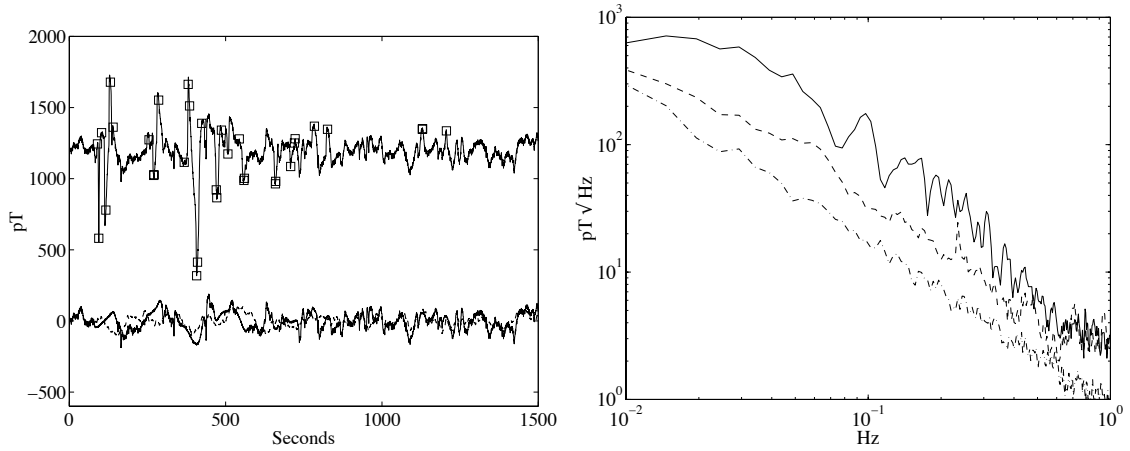


Figure 7.4: Transient detection and removal for data with synthetic $1/f$ noise. Original time series data are from MPK N-S channel at 11:00 UT, August 11, 1998. We have subtracted the estimate of the geomagnetic background from the original time series data and added synthetic $1/f$ noise to generate a modified time series. Panel (a): Modified time series (upper solid) with estimated transient locations marked with squares; modified time series after transient removal (lower solid); and synthetic $1/f$ noise time series (dash). Panel (b): Power spectra of the modified time series (solid), the modified time series after transient removal (dash), and the synthetic $1/f$ time series (dash-dot).

W defines the width of the neighborhood, and the $\hat{l}_j$ are the delay estimates. A value of W that corresponds to a width of 20 seconds is used. We then hard threshold $\hat{X}'_{m,n}$ to obtain

$$\hat{X}''_{m,n} = H(\hat{X}'_{m,n}), \tag{7.9}$$

where the thresholding operation $H(\cdot)$ is defined as

$$H\left(\hat{X}_{m,n}\right) = \begin{cases} \hat{X}_{m,n}, & |\hat{X}_{m,n}| > \sigma_m \\ 0, & |\hat{X}_{m,n}| \leq \sigma_m \end{cases} \tag{7.10}$$

and σ_m is the standard deviation of the background noise at scale m. An inverse translation invariant wavelet transform (TIWT) [*Coifman and Donoho*, 1995] is used to generate an

estimate

$$y(t) = \mathrm{TIWT}^{-1}\left(\hat{X}''_{m,n}\right) \tag{7.11}$$

of the transient signals from the thresholded coefficients. The transient-free time series is defined as

$$z(t) = x(t) - y(t), \tag{7.12}$$

where $x(t)$ is the original time series. Examples of applying this method are shown in Figures 7.2, 7.3, 7.4 and 7.9. We see that we have effectively removed the transients, while preserving some of the features of the desired background. Figures 7.3 and 7.4 show the power spectra of the time series before and after the transient detection and removal process. The removal process has reduced the interference over a wide frequency range.

Thresholding in the wavelet domain (either DWT or UDWT domain) is a commonly used technique [*Coifman and Donoho*, 1995; *Johnstone and Silverman*, 1996] for the denoising of signals, and we make no claims for originality in its use. The original contribution described here is the introduction of neighborhood selection (based upon estimated transient locations) to constrain the thresholding to regions in the UDWT domain where a transient might be located. In Figure 7.5 we compare the results of the neighborhood thresholding method with the standard method of thresholding all the coefficients (global thresholding). We applied both neighborhood thresholding and global thresholding to the signal presented in Figure 7.4 (MPK time series with background estimated removed and synthetic $1/f$ noise added), and compared the results to the desired signal, which in this case is the synthetic $1/f$ noise. Global thresholding suffers the drawback of eliminating many of the features of the desired signal, while these features are preserved, for the most part, by neighborhood thresholding.

7.2.3 Monitoring for Precursors

In this section, we examine the effect of the transient detection and removal process on our ability to monitor for precursors. We consider the following hypothetical scenario. A major earthquake has occurred somewhere along the Hayward Fault and there are no obvious magnetic field anomalies in the hours preceding the event. We are interested in determining, however, if a temporary anomalous increase might have occurred in the days or weeks prior to the earthquake. An effective way of searching for such anomalies is to examine the half-hourly spectral indices [*Fraser-Smith et al.*, 1990]. If the anomalous

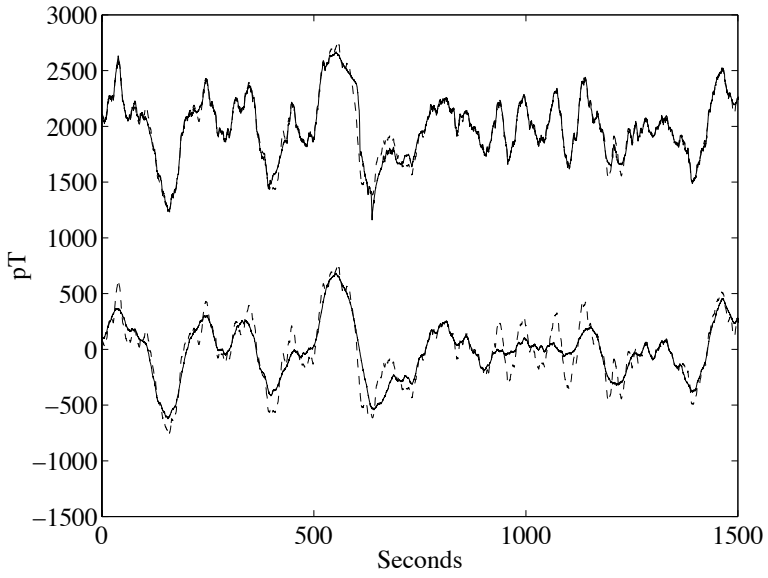


Figure 7.5: Comparison between neighborhood and global thresholding. The upper and lower dashed lines are identical and represent the synthetic $1/f$ process. The upper and lower solid lines are the estimates of the $1/f$ process using transient removal with neighborhood and global thresholding, respectively. In both cases, the detection and transient removal method was applied to the time series represented by the middle solid line in Figure 7.2(b).

increases are small compared with the typical BART spectral amplitudes, it will be difficult to detect such increases. We would like to determine if the transient removal process improves our ability to detect these small increases.

We consider the following two monitoring situations:

Case I. All the stations in the SFBAY array are in operation, but no remote reference stations are available to provide an accurate estimate of the natural geomagnetic background. In this case, we apply the transient detection and removal directly to the observed time series. In the absence of remote reference stations, we may use global magnetic indices, such as the A_p or K_p indices [*Campbell*, 1997], to form rough estimates of the parameters $\beta\sigma_w^2$ and γ used to set the detector thresholds.

Case II. All the stations in the SFBAY array are in operation, and in addition at least one remote reference stations is available to provide an accurate estimate of the natural

geomagnetic background. In this case, we first subtract the background estimate prior to applying the transient removal process.

For both Case I and Case II, we compute the power spectra for the following data sets:

1. Base time series.
2. Base time series with synthetic $1/f$ noise added.
3. Base time series after transient removal.
4. Base time series with synthetic $1/f$ noise added and after transient removal.

The term 'base time series' refers to the original time series data for Case I and the original time series data with the natural background removed for Case II. The synthetic $1/f$ noise is used to represent the precursor signal that we are trying to detect. We perform pairwise comparisons of the spectra – i.e. set 1 compared to set 2, and set 3 compared to set 4. This process allows us to determine the extent to which the spectra with the precursor differs from the spectra without the precursor. In Figure 7.6(a), we present a schematic diagram of the process. In panels (b) and (c) we show the results of applying the procedure to data acquired on the MPK N-S channel at 11:00 UT on August 11, 1998 for cases I and II, respectively. In both cases, the transient removal process significantly improves our ability to determine if a precursor signal is present.

7.3 Extension to Multiple Channels

In the single channel case, we relied solely on temporal features to differentiate the man-made interference from the background. Although unlikely, a precursor signal may have the same temporal features as the interference, in which case the precursor signal would be identified as a BART transient and removed. To reduce the risk of an incorrect identification, we can use the polarization vectors of the transients to further distinguish BART related signals from naturally occurring signals.

7.3.1 Defining Polarization Subsets

As discussed in Chapter 3, the 3-dimensional polarization vectors $\mathbf{P}_{i,j}$ at the ith station are confined to 3-dimensional subsets S_i of Euclidean 3-space. We can extend this observation

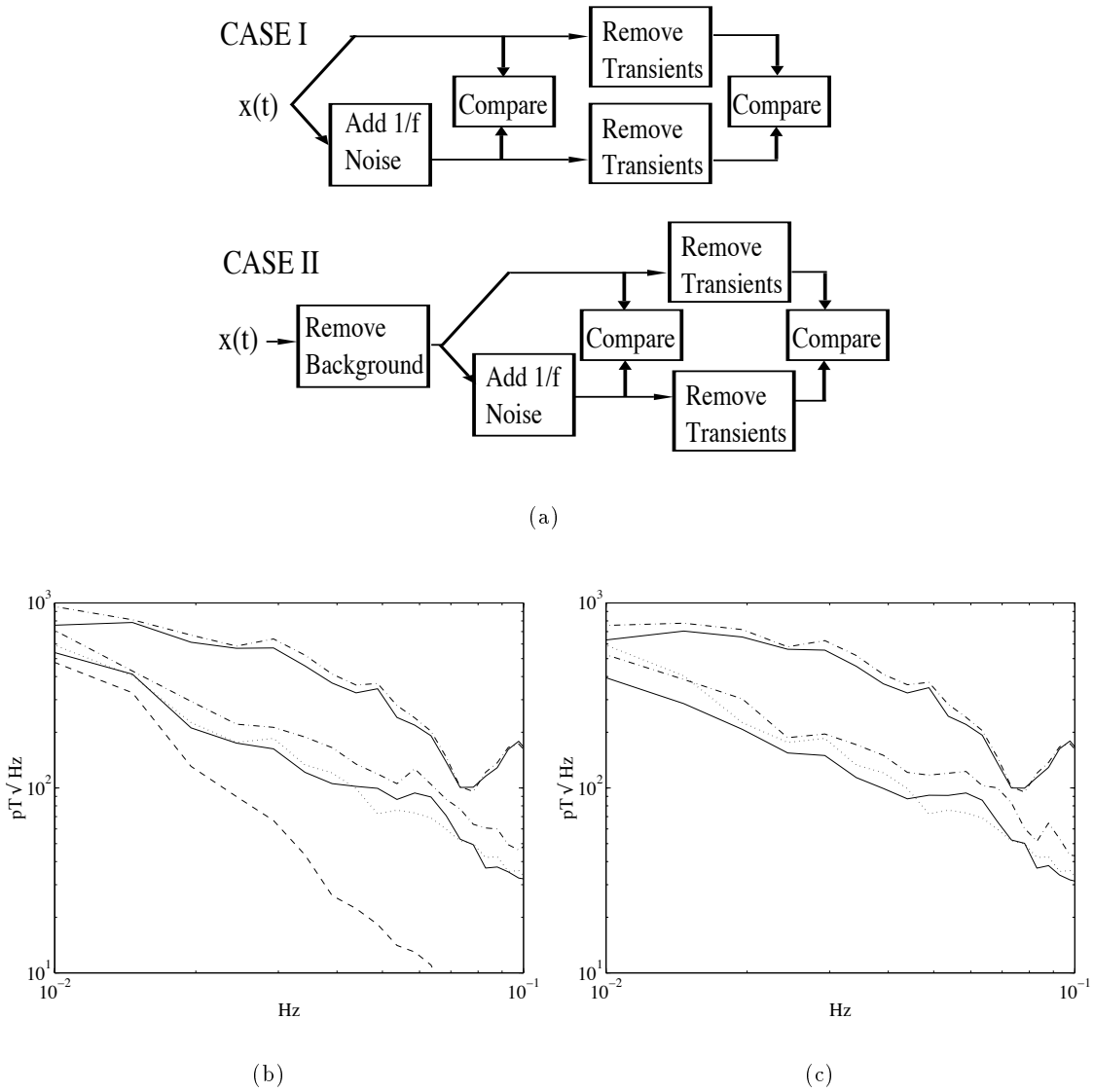


Figure 7.6: Monitoring for precursors. (a) Schematic diagram of procedure. (b) Case I. (b) Case II. For both cases I and II we show power spectra of base time series (upper solid) and base time series with $1/f$ noise added (upper dash-dot); power spectra of base time series (lower solid) and base time series with added $1/f$ noise (lower dash-dot), both after transient removal; and power spectra of $1/f$ noise (dots). For case I we also show the background estimate (dash).

to conclude that the 9-dimensional polarization vectors

$$\mathbf{P}_j = \begin{bmatrix} \mathbf{P}_{1,j} \\ \mathbf{P}_{2,j} \\ \mathbf{P}_{3,j} \end{bmatrix} \tag{7.13}$$

satisfy the condition $\mathbf{P} \in S$, where $S \subset \mathcal{R}^9$. A convenient choice is $S = S_T$ where

$$S_T = \hat{S}_1 \cup \hat{S}_2 \cup \hat{S}_3 \tag{7.14}$$

and each $\hat{S}_i$ is the 9-dimensional continuation of S_i. For example, $\hat{S}_1$ is the subset that contains all vectors of the form $\begin{bmatrix} \mathbf{P}_{1,j}^T & 0 & \cdots & 0 \end{bmatrix}^T$. While in general it is possible to define a subset S that is smaller than S_T, the subset S_T has the advantage of being easy to visualize and work with.

We construct each subset S_i as the union of overlapping ellipsoids which are chosen to cover a set of observed polarization vectors. To do this, we first define a central ellipsoid using a technique adapted from the field of robust statistics [*Rousseuw and Leroy*, 1987]. As the initial step, we compute a sample covariance matrix

$$\mathbf{C} = \frac{1}{N-1} \sum_{j=1}^{N} (\mathbf{P}_{i,j} - \mathbf{P}_c)(\mathbf{P}_{i,j} - \mathbf{P}_c)^T \tag{7.15}$$

where N is the number of polarization vectors and

$$\mathbf{P}_c = \frac{1}{N} \sum_{j=1}^{N} \mathbf{P}_{i,j} \tag{7.16}$$

is the mean polarization vector. We then compute the weighted lengths of the vectors

$$d_{i,j} = (\mathbf{P}_{i,j} - \mathbf{P}_c)^T \mathbf{C}_i^{-1} (\mathbf{P}_{i,j} - \mathbf{P}_c) \tag{7.17}$$

and sort the lengths. Vectors associated with the M smallest lengths are used to define the central ellipsoid, where M/N is typically chosen to be 0.85 or greater. Denoting the set of indices for the chosen vectors as $\mathcal{M}$, we compute the covariance matrix $\mathbf{C}_0$ and

mean vector $\mathbf{X}_0$ for all vectors with indices $j \in \mathcal{M}$. The central ellipsoid is then defined by the equation

$$(\mathbf{X} - \mathbf{X}_0)^T \mathbf{A}_0^{-1} (\mathbf{X} - \mathbf{X}_0) \leq 1 \tag{7.18}$$

where $\mathbf{A}_0 = d_{\max} \mathbf{C}_0$ and

$$d_{\max} = \max_{j \in \mathcal{M}} (\mathbf{P}_{i,j} - \mathbf{X}_0)^T \mathbf{C}_0^{-1} (\mathbf{P}_{i,j} - \mathbf{X}_0). \tag{7.19}$$

Ellipsoids are 'grown' to cover the polarization vectors with indices $j \notin \mathcal{M}$, i.e., those vectors that are not included in the central ellipsoid. For each of these vectors, we generate an ellipsoid with its major axis aligned with the vector and with a major axis length equal to B times the length of the vector, where $B > 1$. Typically, $B = 1.1$. The width of the ellipsoid is determined by the lengths of the minor axes. Each of these ellipsoids can be described by a matrix $\mathbf{A}_j$, a mean vector $\mathbf{X}_j$, and the expression

$$(\mathbf{X} - \mathbf{X}_j)^T \mathbf{A}_j^{-1} (\mathbf{X} - \mathbf{X}_j) \leq 1. \tag{7.20}$$

Figure 7.7 shows the set of ellipsoids generated to cover the polarization vectors observed at station MPK between 11:30 and 12:00 UT over the course of 3 days, August 10, 1998 through August 12, 1998. In order to demonstrate how this subset can be used to distinguish the BART interference from natural signals of interests, we also present the polarization vectors for some possible earthquake precursors. We obtained the vectors by calculating the fields due to horizontal electric dipoles situated at the locations of two of the largest earthquakes observed along the Hayward fault in the last 15 years. These were a March 27, 1984, M 4.40 event near LKC (37.74° N, −122.11° E, hypocentral depth of 7.0 km) and a December 4, 1998, M 4.1 event at the northern end of the fault in the vicinity of Richmond, California (37.92° N, −122.29° E, hypocentral depth of 4.1 km) (Data source: Northern California Earthquake Data Center, http://quake.geo.berkeley.edu). We used a dipole moment of 2.7×10^7 A-m, a value obtained from [*Fraser-Smith et al.*, 1993], and assumed that the dipoles were aligned with the strike of the fault. In addition we assumed a homogeneous earth with a conductivity of 0.03 S/m, the value for the Franciscan basement in the conductivity model outlined in Appendix A. The fields were computed for frequencies of 0.01 and 0.1 Hz. At 0.01 Hz, the fields from both the LKC and Richmond sources are clearly distinguished from the MPK BART subset, while at 0.1 Hz only the

fields from LKC stand out. Note that at 0.01 Hz the fields from the Richmond source are smaller in amplitude than those for the LKC source due to the fact that the Richmond to MPK distance is approximately twice the LKC to MPK distance.

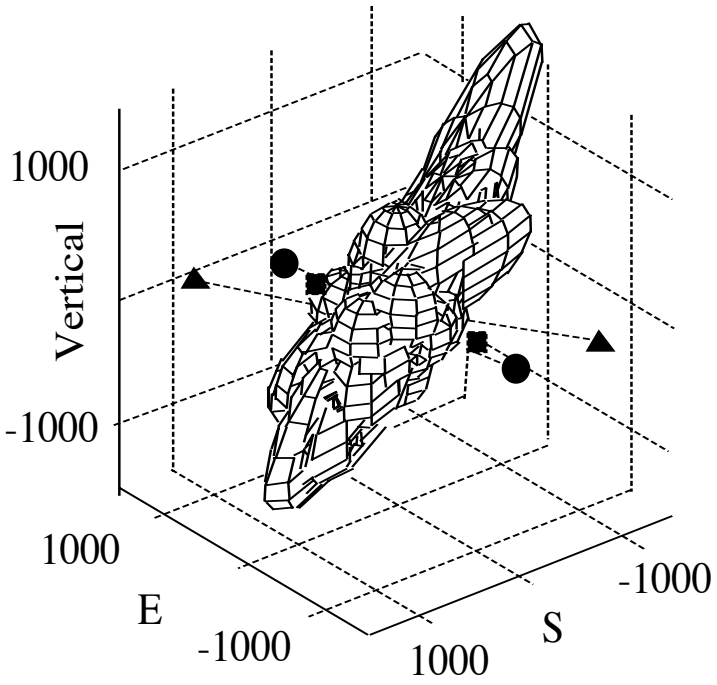


Figure 7.7: BART transient subset at MPK. The subset is the union of overlapping ellipsoids which cover polarization vectors observed between 11:30 and 12:00 UT over the course of August 10, 1998 through August 12, 1998. Hypothetical precursor amplitudes for 0.01 Hz (triangle) and 0.1 Hz (circle) sources near LKC, and 0.01 Hz (square) source near Richmond, CA

7.3.2 Invoking Spatial Constraints

In Section 7.2.1 we described the procedure for the detection of transients on a single channel. The process relied solely on the temporal characteristics of the transients. In this section we make use of the BART polarization subsets to further constrain the detection process. The steps of the procedure are as follows:

1. We apply the detection process described in Section 7.2.1 to each channel in the array. We record the estimated delays of the detected transients.

2. We consider detected transients from different channels to be generated by the same physical BART transient when the absolute difference in the estimated delays are less than a specified width. We then construct a 9-dimensional polarization vector $\mathbf{P}_j$ from the values of the waveforms at the estimated delays. Except for the fact that we use estimated delays instead of maxima locations, the procedure for constructing the polarization vector is identical to that described in Section 3.4.3.

3. For each polarization vector, we wish to determine if $\mathbf{P}_j \in S$. With the definition, $S = S_T$, we may test for the equivalent condition

$$(\mathbf{P}_{1,j} \in S_1) \bullet (\mathbf{P}_{2,j} \in S_2) \bullet (\mathbf{P}_{3,j} \in S_3), \tag{7.21}$$

where $\bullet$ denotes a logical AND operation, and $\mathbf{P}_{i,j}$ are the 3-dimensional polarization vectors that make up $\mathbf{P}_j$. Since each S_i is the union of ellipsoids, and each ellipsoid can be described by a matrix $\mathbf{A}_j$ and a mean vector $\mathbf{X}_j$, it is straightforward to compute membership in S_i. If $\mathbf{P}_j$ satisfies the test condition, then we conclude that $\mathbf{P}_j$ is a BART related polarization vector, and we consider all associated transients to be BART related. If $\mathbf{P}_j$ does not satisfy the test condition, then we consider all the associated detections to be unrelated to BART, and we do not attempt to remove the transients at these locations. A schematic drawing of this procedure is provided in Figure 7.8.

To demonstrate the above approach, we added hypothetical earthquake precursor signals to the data set first presented in Figure 7.2. The 0.1 Hz LKC source previously described was used to calculate the polarization vectors at each of the stations. We reduced the dipole moment to a value of 8×10^6 A-m in order to obtain polarization vectors at MPK and JSR which are fully contained within the BART subsets at the respective stations. The polarization vector at LKC extends far beyond the boundary of the BART subset, due to the proximity of the source to the observation station. This situation is represented schematically in Example 2 of Figure 7.8.

In Figure 7.9 we have added two transients to the MPK N-S time series from 10:30 UT on August 11, 1998. The transients are located at the 450 and 1050 second time

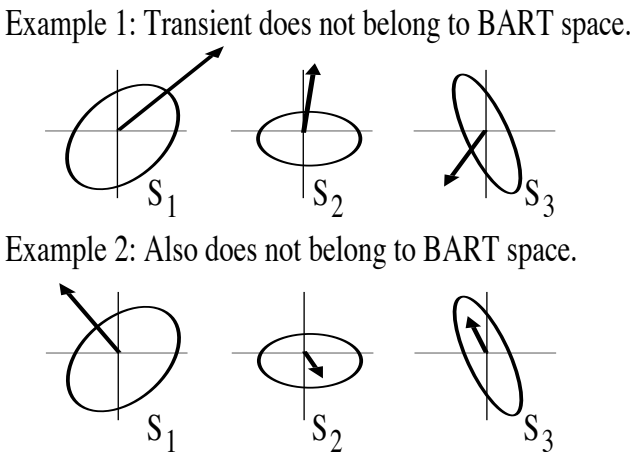


Figure 7.8: Testing for membership in BART subsets. Example 1: The polarization vector lies outside of the subsets at all stations. Example 2: The components of the polarization vector lie within the subsets at stations 2 and 3, but outside the subset at station 1. In both examples, we conclude that the polarization vector is not due to BART.

marks and both have a temporal width of $\sigma_t = 1.8$ seconds, corresponding to a duration of approximately 9 seconds. Transients with the same shapes and locations have also been added to the remaining channels in the array. The amplitudes of the transients on each channel are set equal to the amplitudes of the calculated polarization vectors. The circles and squares indicate the locations of the detections obtained with the single channel method, while the squares indicate the locations of the detections obtained with the multiple channel method that incorporates the BART subset information. Note that the multiple channel method correctly determines that the two added transients are not due to BART. In addition, some of the geomagnetic background transients that led to false alarms in the single channel method are now ignored. We conclude that the multiple channel method increases our ability to distinguish BART transients from precursors and other natural signals of interest.

7.4 Discussion

In this chapter we presented two complementary approaches to the identification and removal of transients generated by the BART system. The first approach addressed the

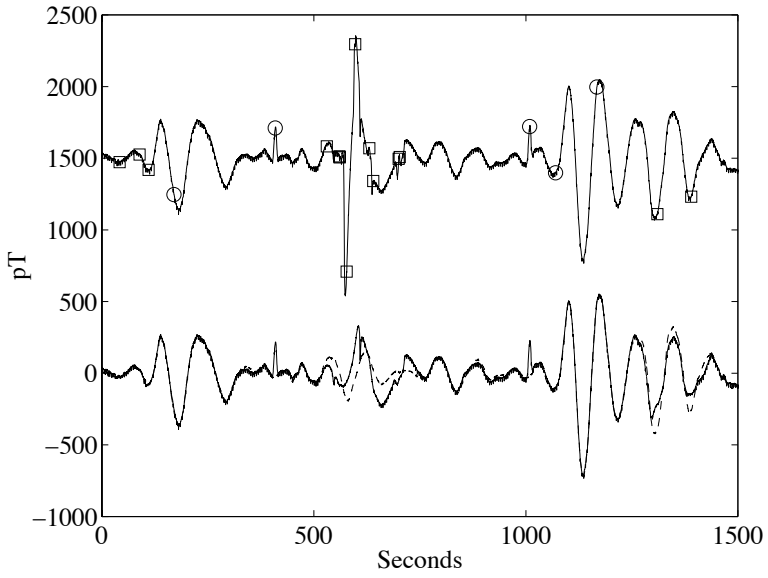


Figure 7.9: Detection with spatial constraints imposed. Upper solid line shows original time series (MPK N-S channel, 10:30 UT, August 11, 1998) with hypothetical precursor transient signals added at 400 and 1000 seconds. Squares indicate detected transient locations, while circles represent transients that would have been detected if spatial constraints had not been imposed. The lower solid and dashed lines are the time series after transient removal and the estimate of the geomagnetic background, respectively.

single channel case, while the second was an extension of the first to the multiple channel case. Both approaches make use of the signal model presented in Chapter 5.

The advantage of the signal model is that it makes explicit our assumptions about both the interference and the assumed background. In the case where there are no precursor signals, the background represents either natural geomagnetic noise or the prediction error associated with the use of a remote reference — it is straightforward to estimate the properties of either background signal. A difficulty arises if we suspect that the background may also contain precursor signals. What assumptions can we make then? In this chapter, we have assumed a $1/f$ model for the precursor signals, and chosen, for the sake of argument, reasonable values for the parameters, variance σ and exponent γ, of the model. In practice, of course, the parameter values are not known, and the $1/f$ model may not be applicable. This difficulty is not a drawback of the signal model, since any identification method must make assumptions about the nature of the signals under

consideration. In fact, all the methods described in Chapter 1 make implicit assumptions about the interference and the background.

One way of addressing our ignorance with regard to the precursor signals is to consider a wide range of signal models, each with a different background signal. For example, we may assume that there are N different signal models, where the ith signal model has a $1/f$ noise background with parameters σ_i and γ_i. The transient identification and removal process is then applied using the assumptions of each signal model, resulting in N different 'transient-free' data sets. The ensemble of 'transient-free' data sets can then be examined for anomalous increases.

Chapter 8

Conclusions and Future Work

This dissertation addresses one of the outstanding problems in the study of magnetic precursors to earthquakes, namely the measurement of natural magnetic fields in the presence of severe man-made interference. An understanding of this problem is crucial to the development of monitoring efforts in urban areas with high earthquake risk.

To the best of our knowledge, this work represents the first attempt to develop a comprehensive approach for the monitoring of ULF magnetic fields in the presence of severe man-made interference. Our goal has been to point out the salient features of the interference and to develop a framework for using these features to distinguish the interference from natural signals of interest. While the work focuses on the development of a monitoring effort in the San Francisco Bay Area, the methods described should also be useful for other situations. In the following sections, we review the main contributions of this thesis and also indicate directions for future work.

8.1 Characterization of Magnetic Fields

8.1.1 Contributions

The description presented in Chapter 3 of the magnetic fields due to BART is unprecedented in its scope and level of detail. The characterization of the characteristic time scales and polarization of the interference provides a basis for the signal processing framework that is developed in later chapters. Finally, the analysis of ULF geomagnetic data with the discrete wavelet transform is new.

8.1.2 Suggestions for Future Research

The mechanisms which give rise to the observed $1/f$ spectra of the geomagnetic background fields are not well understood and deserve further investigation.

8.2 Models of the BART Interference

8.2.1 Contributions

In Chapter 4, the magnetic fields due to BART are calculated using a lumped-element circuit model and formulas for the propagation of magnetic fields over a layered, conducting earth. The use of realistic models to calculate the fields generated by a modern railway system has not been reported previously in the literature. In addition, the two-dimensional conductivity model of the San Francisco Bay Region presented in Appendix A is new.

8.2.2 Suggestions for Future Research

Numerical electromagnetic modeling techniques, coupled with a three-dimensional conductivity model of the San Francisco Bay Region, could be useful for examining the impact of local conductivity structures, such as earthquake faults, on the propagation of the magnetic fields due to BART. In addition, these models could also provide insight into the propagation of the magnetic fields due to sources within the earth, and thus aid in determining the source location of magnetic precursors to earthquakes.

8.3 Detection of Transients in $1/f$ Noise

8.3.1 Contributions

Chapter 6 describes a novel method for the detection, in the presence of $1/f$ noise, of transients with unknown parameters. The method is computationally efficient and robust with respect to transient model mismatches. The observation that the transform local maxima can be used to estimate the parameter values that maximize the detection statistic is also new.

8.3.2 Suggestions for Future Research

The detection method assumes the presence of ideal $1/f$ noise. Reasonable approximations were made in Chapter 7 in order to apply the method to the piecewise $1/f$ geomagnetic background. We have recently determined that it is possible to modify the method to directly handle piecewise $1/f$ noise. Monte Carlo simulations can be used to determine if the modified method yields significantly better performance.

8.4 Identification and Removal of Transients

8.4.1 Contributions

Building upon the signal model proposed in Chapter 5 and the detection method of Chapter 6, Chapter 7 presents methods for the identification and removal of the BART transients. The use of a detection theory framework to identify the transients is new, as is the use of neighborhood thresholding and the inverse translation invariant wavelet transform to estimate the transients. While the use of two-dimensional polarization to distinguish the interference from the background has been reported previously [*Santarato and Spagnolini*, 1995], the use of N-dimensional polarization vectors, where N is the number of channels in the array, is new.

8.4.2 Suggestions for Future Research

The removal of BART transients increases the separation between the spectra of signals with and without added precursor signals. The statistical significance of this separation can be quantified using, for example, a t-test [*Press et al.*, 1988]. A characterization of the variance in the spectral amplitudes of both the BART interference and the remote-reference prediction error is required.

Our method of testing the membership of a polarization vector in pre-defined BART subsets represents an *ad hoc* form of classification. Other methods, such as tree-structured classifiers [*Ripley*, 1996], may yield better performance, and should be investigated.

Efforts to detect electromagnetic precursors are currently underway in Japan [*Uyeda*, 1997], a country with many urban areas that are subject to high earthquake risk. It would be beneficial to determine if the methods developed for the identification and removal of

man-made transients in the San Francisco Bay Area would be useful for monitoring efforts in major urban areas, such as Kyoto or Tokyo.

Appendix A

San Francisco Bay Region Conductivity Model

A.1 Introduction

In this section we present a simplified model for the conductivity structure of the San Francisco Bay (SFBay) Region. We constructed this model to aid in the calculation of the quasistatic magnetic fields produced by currents associated with BART.

A.2 A 2-Dimensional Model

We use a Cartesian coordinate system where the z coordinate corresponds to depth with $z = 0$ at sea level and $z > 0$ for positions below sea level. The y-axis is roughly parallel to the strike of the San Andreas Fault. A map showing the region of interest is presented in Figure A.1.

From the varied geography shown in Figure A.1, it is clear that the conductivity $\sigma(x, y, z)$ will exhibit a dependence on the y coordinate. For example a surface conductivity profile across the bay will reflect the transition from land to seawater. Below the surface, however, the conductivity will be a weak function of y because the predominant geological structures run parallel to the y axis. We simplify our model to reflect this regional feature by restricting the conductivity to be a 2-dimensional function of the x and z coordinates. That is

$$\sigma(x, y, z) = \sigma(x, z). \tag{A.1}$$

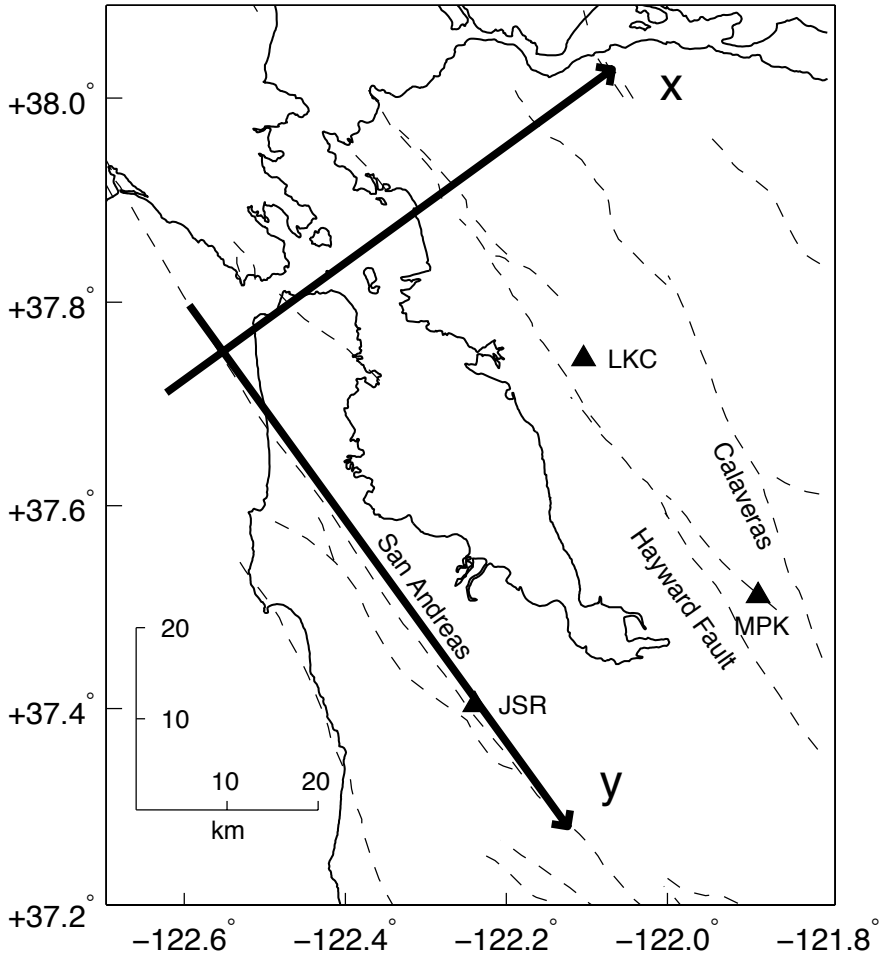


Figure A.1: Map showing SFBay region with site locations (triangles), known surface faults (dash), and x and y coordinate axes.

A schematic of the two-dimensional conductivity model is shown in Figure A.2. In approximating the geology of the SFBay region we have adapted the results of the Bay Area Seismic Imaging Experiment (BASIX) [*Brocher et al.*, 1994; *Holbrook et al.*, 1996]. Our simplified model assumes that a thin sedimentary layer rests upon a granitic basement to the west of the San Andreas Fault and a Franciscan basement to the east of the fault. The basement sections lie on top of subducted oceanic crust.

A.3 Assumptions and Approximations

We now discuss the assumptions and approximations that were used in the construction of our model.

Flat Earth. We ignore changes in surface elevation and assume a flat earth. This assumption is equivalent to a z coordinate uncertainty of about 1 km, since the highest peaks in the region are about 0.8 km tall.

Piecewise Stratified Earth. We assume that the boundaries between various geological structures are parallel to either the xy or yz planes. This represents a simplification since the oceanic lithosphere dips eastward in the SFBay region [*Brocher et al.*, 1994; *Holbrook et al.*, 1996].

Upper Sedimentary Layer. For the purpose of electromagnetic analysis of the SFBay region, we ignore the detailed geology of the region and assume that the key feature of the top layer is that its average conductivity is higher than that of the underlying basement. We assign the top layer a uniform conductivity of 0.5 S/m and a thickness of 0.5 km. As we discuss below, our model simplifies greatly the actual situation.

The geology of the SFBay region is complicated, and so in fact the top layer consists not only of sediments but also of outcroppings of igneous and metamorphic rocks and, of course, the seawater of the bay proper. The bay proper is quite shallow with a mean depth of 6 meters [*Wahrhaftig and Sloan*, 1989] and maximum depths of approximately 100 meters at the Golden Gate Bridge and 20 to 30 meters for various cross-sections across the bay [*Goldman*, 1967]. Typical conductivity values for seawater range from 1 to 5 S/m [*Palacky*, 1987].

The actual thickness of the sedimentary layer varies greatly with location. Below the bay proper, the Franciscan Basement may be encountered at depths as shallow as 0.1 to 0.3 km [*Goldman*, 1967]. *Rogers and Figuers* [1992] estimate the depth to bedrock at approximately 0.5 km along the eastern margin of the bay. A cross-section just to the south of the bay proper reveals sediment deposits that are 2 to 3 km thick [*Wagner et al.*, 1990]. These consist primarily of alluvium and undivided Tertiary sedimentary rocks to the east of the San Andreas Fault, and Tertiary sediment (e.g., Butano sandstone) formations to the west of the San Andreas Fault.

Conductivity also varies greatly with location depending on the the type of sediment present. In the SFBay region we encounter the entire range of sediments, from young, unconsolidated sand and mud deposits located around the shores of the bay to older consolidated marine sediments found both further from the bay and at deeper depths.

Conductivity depends primarily on the presence of electrolytes in the pores and cracks of the rock [*Palacky*, 1987]. The presence of hydrous materials such as clays and serpentine can also increase conductivity [*Grant and West*, 1965]. Typical conductivity values are 0.05 to 0.5 S/m for sedimentary basins [*Hermance*, 1995] and 0.1 to 1 S/m for unconsolidated clastic sediments [*Schmucker*, 1970]. In their magnetotelluric survey of the Loma Prieta epicentral region, *Eberhart-Phillips et al.* [1990] attributed conductivity values of 0.1 to 0.3 S/m to the presence of marine sediments in a section to the west of the San Andreas Fault. *Kurtz et al.* [1990] used an average conductivity value of 0.3 S/m for sediments, but noted that seawater saturated sediments can have conductivities as high as 2.5 S/m. *Unsworth et al.* [1997] found conductivities of about 0.2 S/m to depths of 1.5 km in briny sedimentary rocks to the east of the San Andreas Fault in Parkfield, CA.

Franciscan Basement. We assume that the Franciscan basement has a conductivity of 0.03 S/m and a thickness of 15 km [*Brocher et al.*, 1994; *Holbrook et al.*, 1996]. Conductivity values obtained from magnetotelluric surveys and resistivity soundings of Franciscan formations range from 0.02 to 0.05 S/m [*Eberhart-Phillips et al.*, 1990; *Morrison et al.*, 1977; *Phillips and Kuckes*, 1983; *Unsworth et al.*, 1997].

Granitic Basement. We assume that the granitic basement has a conductivity of 0.002 S/m and a thickness of 15 km [*Brocher et al.*, 1994; *Holbrook et al.*, 1996]. In addition, we assume that it extends all the way to the Farallon Escarpment which is roughly 40 km to

the west of the San Andreas Fault [*Greene and Kennedy*, 1989]. The depth of the ocean at the escarpment is about 0.2 km. Typical *in situ* conductivity values of granite range from 0.001 to 0.003 S/m [*Morrison et al.*, 1977; *Phillips and Kuckes*, 1983; *Unsworth et al.*, 1997].

Pacific Ocean. To the west of the granitic basement we assume an ocean with a depth of 4 km and a conductivity of 4 S/m. Note that this assumption ignores the clastic sediments and San Simeon subduction complex that extend out into the ocean [*Holbrook et al.*, 1996].

Lower Crust and Upper Mantle. We assume that the basement formations are underlain by a lower crust of mafic rocks with a conductivity of 0.001 S/m and a thickness of 10 km. Below the lower crust, we assume a uniform conductivity of 0.001 S/m for the remainder of the lithosphere down to a depth of 85 km. For the upper mantle we use a conductivity of 0.001 S/m for the depths from 85 to 400 km, and a value of 1 S/m for regions below 400 km. We have assumed that the effect of partial melts in the athenosphere is negligible.

As the following discussion shows, our choice of resistivity values for the lower crust and upper mantle is a rough guess based on the published data. Following *Holbrook et al.* [1996], we have assumed that the lower crust is primarily subducted oceanic crust. Oceanic crust consists primarily of gabbro and basalt with a typical conductivity of 0.004 S/m [*Francheteau*, 1983]. *Eberhart-Phillips et al.* [1990] associated a value of 0.001 S/m with mafic rock intrusions in the Loma Prieta area.

In a cross-section across the Vancouver Island region, *Kurtz et al.* [1990] assumed a value of 2×10^{-4} S/m to depths of 30 km and 100 km for oceanic and continental lithospheres, respectively. They assigned a value of 0.005 S/m to the athenosphere and 1 S/m for the halfspace below a 400 km depth. Similarly, *Wannamaker et al.* [1989] employed values of 3×10^{-4} S/m for oceanic lithosphere down to a depth of 50 km, and 0.001 S/m for continental lithosphere and mantle down to a depth of 215 km. *Hermance* [1995] presented typical values for lower crust and mantle of 0.001 S/m, but stated that intracrustal anomalies in active tectonic regions could have conductivities as high as 0.2 S/m at depths of 15 to 50 km. A similar range of values was given for the athenosphere.

In contrast, *Mackie et al.* [1988] assumed lower crust conductivity values of 3×10^{-5} S/m for depths of 9 to 39 km. Mantle conductivity values were assumed to increase

steadily from 1×10^{-4} S/m at 40 km to 1 S/m for depths below 640 km. *Madden et al.* [1993] used similar values.

Fault Zones. Following *Unsworth et al.* [1997] we model the San Andreas, Hayward, and Calaveras fault zones as 0.5 km wide strips with conductivities of 0.2 S/m. We assume that the fault zones extend to the bottoms of the granitic and Franciscan Basements [*Brocher et al.*, 1994].

A.4 A 1-Dimensional Model

In order to facilitate the derivation of analytical expressions for the quasistatic magnetic fields, we have found it useful to construct a 1-dimensional model based upon the parameters of the 2-dimensional model. As discussed in Chapter 4, the assumption of a 1-dimensional model makes it possible to derive convenient formulas for the quasistatic field quantities, whereas a 2-dimensional model requires more computationally intensive numerical modeling techniques. The 1-dimensional model ignores the presence of the fault zones and assumes that the conductivity profile to the east of the San Andreas Fault applies everywhere. A diagram of this profile is given in Figure 4.8.

A.5 Summary

We have presented a simplified 2-dimensional model for the conductivity of the SFBay region. We believe that the model captures the essential features required for electromagnetic analysis of the SFBay region. We have also presented a 1-dimensional model that is more convenient for the computation of quasistatic magnetic fields.

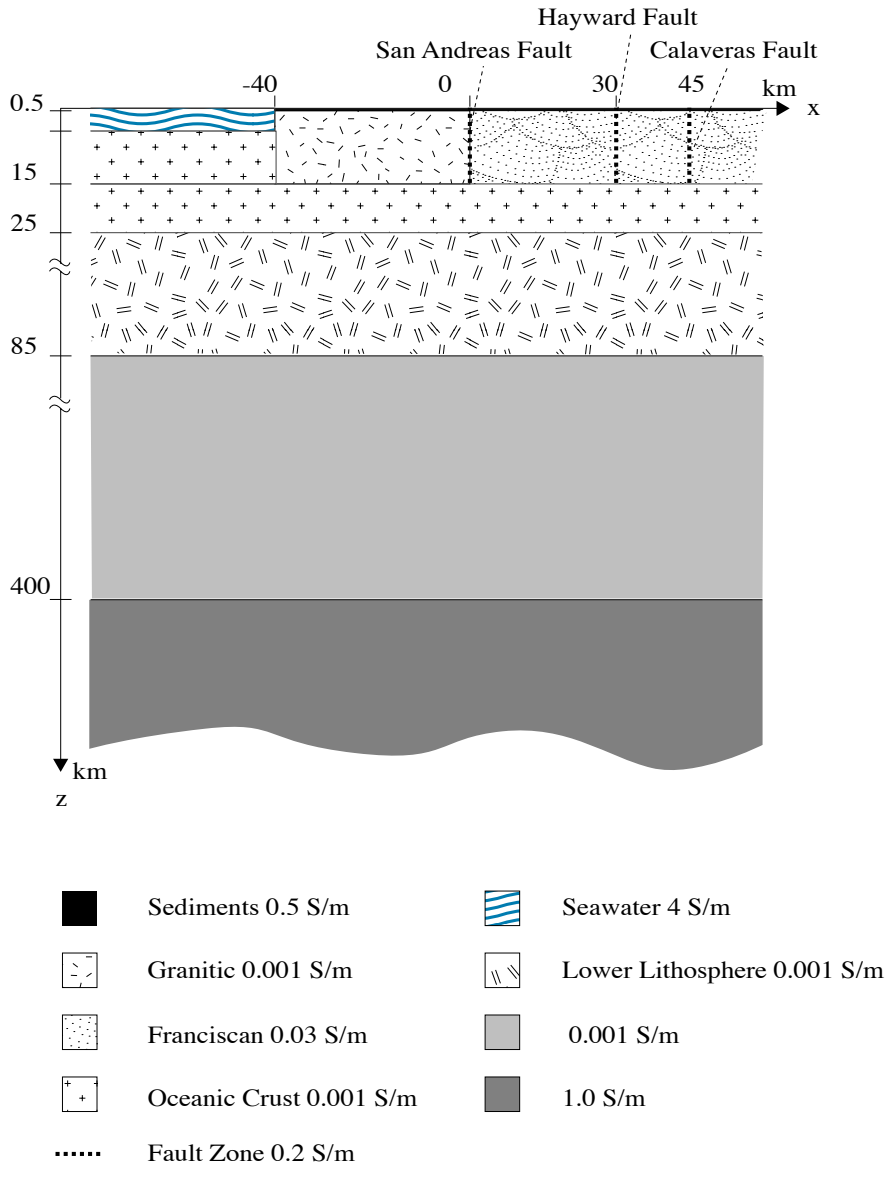


Figure A.2: 2-Dimensional conductivity model for the SFBay region

Appendix B

Dipole Fields

The calculation of the electromagnetic fields due to time-varying electric and magnetic dipoles in stratified media is a well-studied problem with a variety of proposed solutions. In this appendix we present the solution for fields due to a horizontal electric dipole (HED). We adopt the methodology of *Kong* [1981] which considers the propagation of the $\hat{z}$ components of the electric and magnetic fields. Other approaches [*Goldman*, 1990; *Kaufman and Keller*, 1983] utilize the $\hat{x}$ and $\hat{z}$ components of the magnetic vector potential $\mathbf{A}$ or, alternatively, the $\hat{z}$ components of the Schelkunoff electric and magnetic vector potentials [*Ward and Hohmann*, 1987]. We have found Kong's method to be more appealing from an intuitive standpoint.

Our development closely follows *Kong* [1981] with the following differences: (1) Using the widely accepted definition of the Fourier Transform [*Bracewell*, 1986], we assume that the fields vary as $e^{j\omega t}$ instead of as $e^{-j\omega t}$; (2) as is common in the geophysical literature on the subject, we assume that the z axis is oriented downwards into the earth; (3) we use Bessel functions instead of Hankel functions in our expressions; (4) we use a transmission line analogy to derive our formulas for the reflection coefficients. The intent of this appendix is to provide a self-contained explanation of Kong's method. The coordinate system and notation are shown in Figure B.1.

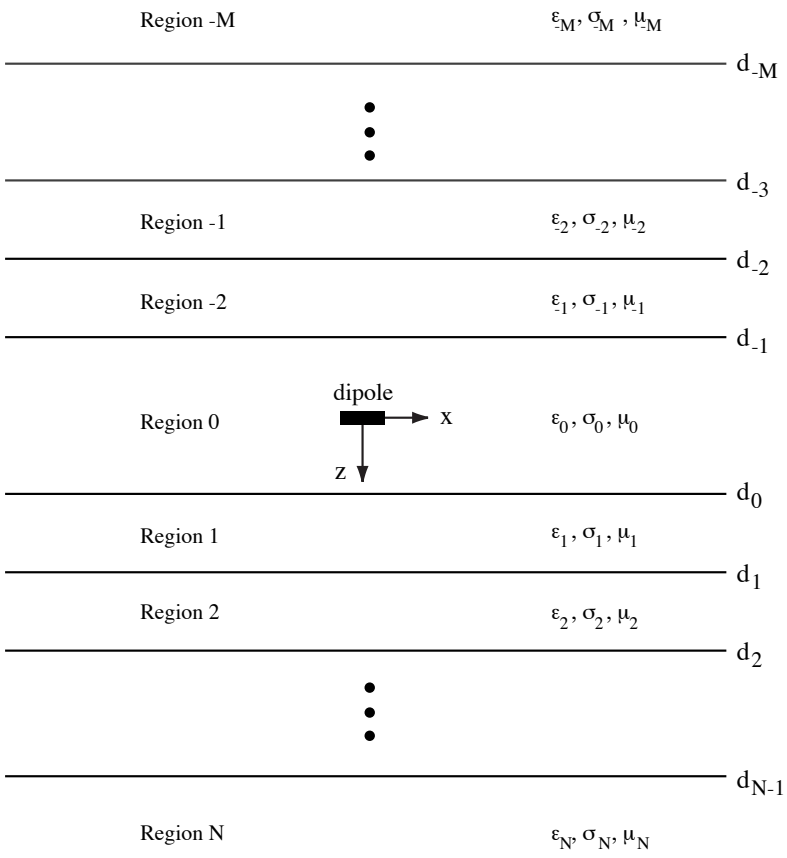


Figure B.1: Notation for calculation of dipole fields in a stratified medium.

B.1 Overview of the Method

Maxwell's equations in isotropic and conducting media yield inhomogeneous wave equations

$$\nabla^2\mathbf{E} + k^2\mathbf{E} = \mu\frac{\partial \mathbf{J}_s}{\partial t} + \nabla\left(\frac{\rho}{\epsilon}\right) \tag{B.1}$$

$$\nabla^2\mathbf{H} + k^2\mathbf{H} = -\nabla \times \mathbf{J}_s \tag{B.2}$$

where

$$k^2 = \mu\epsilon\omega^2 - j\mu\sigma\omega. \tag{B.3}$$

and $\mathbf{J}_s$ is the source current density. In source-free regions the right-hand sides vanish and we obtain the homogeneous wave equations

$$\nabla^2\mathbf{E} + k^2\mathbf{E} = 0 \tag{B.4}$$

$$\nabla^2\mathbf{H} + k^2\mathbf{H} = 0. \tag{B.5}$$

As with all boundary value problems we seek solutions of the form $\mathbf{E} = \mathbf{E}^p + \mathbf{E}^h$ where $\mathbf{E}^p$ is the particular solution to (B.1) and $\mathbf{E}^h$ is the homogeneous solution to (B.4). Values for multiplicative and additive constants in the homogeneous solution are determined by the boundary conditions.

The wave equations are vector equations, and in both Cartesian and cylindrical coordinate systems the $\hat{z}$ component may be found independently, i.e., $\nabla^2 E_z + k^2 E_z = 0$. Kong's method consists of solving the boundary value problems for E_z and H_z. The transverse components are derived then from the $\hat{z}$ components as follows. We express $\mathbf{E}$ and $\mathbf{H}$ as the sum of transverse and longitudinal components

$$\mathbf{E} = \mathbf{E}_t + \hat{z}E_z$$

$$\mathbf{H} = \mathbf{H}_t + \hat{z}H_z.$$

In a similar fashion we may write the gradient operator as $\nabla = \nabla_t + \hat{z}\frac{\partial}{\partial z}$ where $\nabla_t = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y}$ in Cartesian coordinates and $\nabla_t = \frac{\partial}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial}{\partial \phi}\hat{\phi}$ in cylindrical coordinates.

Following *Chew* [1990] (p. 75), we write Maxwell's equation $\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H}$ as

$$\left(\nabla_t + \hat{z}\frac{\partial}{\partial z}\right) \times (\mathbf{E}_t + \hat{z}E_z) = -j\omega\mu\,(\mathbf{H}_t + \hat{z}H_z)\,. \tag{B.6}$$

Equating transverse terms yields

$$\nabla_t \times \hat{z}E_z + \hat{z}\frac{\partial}{\partial z} \times \mathbf{E}_t = -j\omega\mu\mathbf{H}_t. \tag{B.7}$$

Taking the cross product of both sides with $\hat{z}\frac{\partial}{\partial z}$, we obtain

$$\nabla_t \left(\frac{\partial E_z}{\partial z}\right) - \frac{\partial^2 \mathbf{E}_t}{\partial z^2} = -j\omega\mu\hat{z}\frac{\partial}{\partial z} \times \mathbf{H}_t. \tag{B.8}$$

In a similar fashion, from the Maxwell equation $\nabla \times \mathbf{H} = j\omega\bar{\epsilon}\mathbf{E}$, we may derive

$$\nabla_t \times \hat{z}H_z + \hat{z}\frac{\partial}{\partial z} \times \mathbf{H}_t = j\omega\bar{\epsilon}\mathbf{E}_t, \tag{B.9}$$

where

$$\bar{\epsilon} = \frac{\sigma}{j\omega} + \epsilon. \tag{B.10}$$

Substituting the expression for $\hat{z}\frac{\partial}{\partial z} \times \mathbf{H}_t$ into (B.8) yields

$$\mathbf{E}_t = \frac{\nabla_t \left(\frac{\partial E_z}{\partial z}\right) - j\omega\mu\nabla_t \times \hat{z}H_z}{\omega^2\mu\bar{\epsilon} - k_z^2}, \tag{B.11}$$

where we have assumed that $\mathbf{E}_t$ varies as $e^{\pm jk_z}$ so that $\frac{\partial^2 \mathbf{E}_t}{\partial z^2} = -k_z^2\mathbf{E}_t$. The expression $\omega^2\mu\bar{\epsilon} - k_z^2$ may be rewritten as $k_\rho^2 = k^2 - k_z^2$, where k_p and k_z are to be interpreted as the transverse and longitudinal wave numbers, respectively. Note that $k_\rho^2 = k_x^2 + k_y^2$. Making these substitutions leads to the desired expression for $\mathbf{E}_t$

$$\mathbf{E}_t = \frac{1}{k_\rho^2}\left(\nabla_t \left(\frac{\partial E_z}{\partial z}\right) - j\omega\mu\,(\nabla_t \times \hat{z}H_z)\right). \tag{B.12}$$

Taking the cross product of $\hat{z}\frac{\partial}{\partial z}$ with both sides of (B.9) and using the identity (B.7), we can obtain the corresponding expression for $\mathbf{H}_t$

$$\mathbf{H}_t = \frac{1}{k_\rho^2}\left(\nabla_t \left(\frac{\partial H_z}{\partial z}\right) + j\omega\bar{\epsilon}\,(\nabla_t \times \hat{z}E_z)\right). \tag{B.13}$$

The portion of the solution that depends on E_z is known as the transverse magnetic (TM) mode, while the portion of the solution that depends on H_z is known as the transverse electric (TE) mode.

B.2 General Form of the Solutions

Since the boundaries are along planes of constant z, it is convenient to use cylindrical coordinates. The particular solution is most readily determined with the use of a magnetic vector potential $\mathbf{A}$ defined by $\mathbf{H} = \nabla \times \mathbf{A}$. The electric field may be written as $\mathbf{E} = -\nabla\Phi - j\omega\mu A$, where Φ is a scalar potential. Use of the Lorentz Gauge $\nabla \cdot \mathbf{A} = -(\sigma + j\omega\epsilon)\Phi$ leads to the following inhomogeneous equation for $\mathbf{A}$

$$\nabla^2 \mathbf{A} + k^2 \mathbf{A} = -\mathbf{J}_s, \tag{B.14}$$

where $\mathbf{J}_s$ is the source current. The general form of the solution [*Ward and Hohmann*, 1987] is

$$\mathbf{A}(\mathbf{r}) = \int_V \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}\mathbf{J}_s(\mathbf{r}')dV'. \tag{B.15}$$

For a $\hat{x}$ directed electric dipole at the origin $\mathbf{J}_s = Ids\delta(\mathbf{r}')\hat{x}$ where Ids is the dipole moment. This yields

$$\mathbf{A}(\mathbf{r}) = \frac{Ids}{4\pi\mathbf{r}} e^{-jk\mathbf{r}}\hat{x}. \tag{B.16}$$

As we are interested in solutions in cylindrical coordinates we use the Sommerfeld Integral representation ([*Sommerfeld*, 1949], p. 240ff)

$$\begin{aligned} \frac{e^{-jk\mathbf{r}}}{\mathbf{r}} &= \int_0^\infty \frac{k_\rho}{\sqrt{k_\rho^2 - k^2}} e^{-\sqrt{k_\rho^2 - k^2}|z|} J_0(k_\rho r) dk_\rho \\ &= \int_0^\infty \frac{k_\rho}{jk_z} e^{-jk_z|z|} J_0(k_\rho r) dk_\rho \end{aligned} \tag{B.17}$$

to derive formulas for the particular solutions H_z^p and E_z^p in region 0. Note that as before we have made use of the relation $k^2 = k_\rho^2 + k_z^2$. The square root $\sqrt{k_\rho^2 - k^2}$ is understood to have a positive real part so that the integral converges as $z \to \pm\infty$. To be consistent with this choice of sign, we must have $k_z = -j\sqrt{k_\rho^2 - k^2}$.

From the relation $\mathbf{H} = \nabla \times \mathbf{A}$, we find that

$$\begin{aligned} H_z^p &= \frac{Ids}{4\pi} \sin\phi \int_0^\infty \frac{k_\rho^2}{j k_{0,z}} e^{-jk_{0,z}|z|} J_1(k_\rho r) dk_\rho \\ &= \sin\phi \int_0^\infty H_{hed} e^{-jk_{0,z}|z|} J_1(k_\rho r) dk_\rho, \end{aligned} \tag{B.18}$$

where

$$H_{hed} = \frac{Ids}{4\pi} \frac{k_\rho^2}{j k_{0,z}}. \tag{B.19}$$

With the Lorentz Gauge, we may write $\mathbf{E} = \frac{\nabla(\nabla \cdot \mathbf{A})}{\sigma + j\omega\epsilon} - j\omega\mu\mathbf{A}$, so that

$$\begin{aligned} E_z^p &= \frac{sgn(z)}{\sigma_0 + j\omega\epsilon_0} \frac{Ids}{4\pi} \cos\phi \int_0^\infty k_\rho^2 e^{-jk_{0,z}|z|} J_1(k_\rho r) dk_\rho \\ &= sgn(z) \cos\phi \int_0^\infty E_{hed} e^{-jk_{0,z}|z|} J_1(k_\rho r) dk_\rho, \end{aligned} \tag{B.20}$$

where

$$E_{hed} = \frac{Ids}{4\pi} \frac{k_\rho^2}{\sigma_0 + j\omega\epsilon_0} \tag{B.21}$$

and sgn denotes the sign function [*Bracewell*, 1986]. Note that we have adopted the notation $k_{i,z}$ to denote the value of k_z in the ith region, where

$$k_{i,z} = -j\sqrt{k_\rho^2 - k_i^2}. \tag{B.22}$$

General solutions to the homogeneous equations (B.4) and (B.5) in cylindrical coordinates are of the form [*Ramo et al.*, 1984]:

$$\begin{aligned} \left.\begin{matrix} E_z \\ H_z \end{matrix}\right\} &= \left(Ae^{-jk_z z} + Be^{jk_z z}\right)\left(C_1 J_\nu(k_\rho r) + C_2 N_\nu(k_\rho r)\right) \\ &\quad \times \left(F_1 \sin\nu\phi + F_s \cos\nu\phi\right), \end{aligned} \tag{B.23}$$

where J_ν and N_ν are Bessel functions of the first and second kind, respectively, with order ν. From inspection of the particular solution, we can see that for the ith region the homogeneous solutions for a dipole source will be of the form

$$E_{i,z} = \int_0^\infty \left(A_i e^{-jk_{i,z} z} + B_i e^{jk_{i,z} z}\right) J_n(k_\rho r) C_n(\phi) dk_\rho \tag{B.24}$$

$$H_{i,z} = \int_0^\infty \left(C_i e^{-jk_{i,z}z} + D_i e^{jk_{i,z}z}\right) J_n(k_\rho r) S_n(\phi) dk_\rho, \tag{B.25}$$

where, in the case of a HED, $n = 1$, $C_1(\phi) = \cos(\phi)$ and $S_1(\phi) = \sin(\phi)$. If we consider the sum of the particular and homogeneous solutions, we can see that (B.24) and (B.25) are also the general forms for the total solution in region 0.

The transverse components may be found with the use of (B.12) and (B.13). Since the relations hold for each value of k_ρ the indicated operations are performed inside the integral sign. We obtain

$$\begin{aligned}
E_{i,r} &= \int_0^\infty \frac{-jk_{i,z}}{k_\rho}\left(A_i e^{-jk_{i,z}z} - B_i e^{jk_{i,z}z}\right) J_n'(k_\rho r) C_n(\phi) dk_\rho \\
&\quad - \int_0^\infty \frac{j\omega\mu_i}{rk_\rho^2}\left(C_i e^{-jk_{i,z}z} + D_i e^{jk_{i,z}z}\right) J_n(k_\rho r) S_n'(\phi) dk_\rho \qquad \text{(B.26)} \\
E_{i,\phi} &= \int_0^\infty \frac{-jk_{i,z}}{k_\rho^2 r}\left(A_i e^{-jk_{i,z}z} - B_i e^{jk_{i,z}z}\right) J_n(k_\rho r) C_n'(\phi) dk_\rho \\
&\quad + \int_0^\infty \frac{j\omega\mu_i}{k_\rho}\left(C_i e^{-jk_{i,z}z} + D_i e^{jk_{i,z}z}\right) J_n'(k_\rho r) S_n(\phi) dk_\rho \qquad \text{(B.27)} \\
H_{i,r} &= \int_0^\infty \frac{-jk_{i,z}}{k_\rho}\left(C_i e^{-jk_{i,z}z} - D_i e^{jk_{i,z}z}\right) J_n'(k_\rho r) S_n(\phi) dk_\rho \\
&\quad + \int_0^\infty \frac{j\omega\bar{\epsilon}_i}{rk_\rho^2}\left(A_i e^{-jk_{i,z}z} + B_i e^{jk_{i,z}z}\right) J_n(k_\rho r) C_n'(\phi) dk_\rho \qquad \text{(B.28)} \\
H_{i,\phi} &= \int_0^\infty \frac{-jk_{i,z}}{k_\rho^2 r}\left(C_i e^{-jk_{i,z}z} - D_i e^{jk_{i,z}z}\right) J_n(k_\rho r) S_n'(\phi) dk_\rho \\
&\quad - \int_0^\infty \frac{j\omega\bar{\epsilon}_i}{k_\rho}\left(A_i e^{-jk_{i,z}z} + B_i e^{jk_{i,z}z}\right) J_n'(k_\rho r) C_n(\phi) dk_\rho, \qquad \text{(B.29)}
\end{aligned}$$

where the the $f'(u)$ notation is used to denote the derivative of a function $f(u)$ with respect to its argument u. Note that solutions consist of a TM mode (terms involving A_i and B_i) and a TE mode (terms involving C_i and D_i).

B.3 Boundary Conditions

At each boundary $z = d_i$, we require continuity of the tangential components of $\mathbf{E}$ and $\mathbf{H}$. Since the wave equations for E_z and H_z are not coupled, we obtain independent boundary conditions for TM and TE modes. For the TM mode, continuity of E_r and E_ϕ yields

$$jk_{i,z}\left(A_i e^{-jk_{i,z}d_i} - B_i e^{jk_{i,z}d_i}\right) = jk_{i+1,z}\left(A_{i+1} e^{-jk_{i+1,z}d_i} - B_{i+1} e^{jk_{i+1,z}d_i}\right), \tag{B.30}$$

while continuity of H_r and H_ϕ leads to

$$j\omega\bar{\epsilon}_i\left(A_i e^{-jk_{i,z}d_i} + B_i e^{jk_{i,z}d_i}\right) = j\omega\bar{\epsilon}_{i+1}\left(A_{i+1}e^{-jk_{i+1,z}d_i} + B_{i+1}e^{jk_{i+1,z}d_i}\right). \tag{B.31}$$

Similarly, for the TE mode, continuity of tangential **E** leads to

$$j\omega\mu_i\left(C_i e^{-jk_{i,z}d_i} + D_i e^{jk_{i,z}d_i}\right) = j\omega\mu_{i+1}\left(C_{i+1}e^{-jk_{i+1,z}d_i} + D_{i+1}e^{jk_{i+1,z}d_i}\right), \tag{B.32}$$

while continuity of tangential **H** leads to

$$jk_{i,z}\left(C_i e^{-jk_{i,z}d_i} - D_i e^{jk_{i,z}d_i}\right) = jk_{i+1,z}\left(C_{i+1}e^{-jk_{i+1,z}d_i} - D_{i+1}e^{jk_{i+1,z}d_i}\right). \tag{B.33}$$

B.4 Wave amplitudes in Region 0

Kong's method initially solves for the wave amplitudes in region 0 (e.g., A_0, B_0, C_0, and D_0). Once these amplitudes are known, the amplitudes in any other region can be obtained through the use of propagation matrices (see Section B.6). Each region 0 amplitude is the sum of amplitudes due to the particular and homogeneous solutions. Examination of (B.18) and (B.20) shows that the particular solution consists of downward ($e^{-jk_z z}$) going waves for $z \geq 0$ and upward ($e^{jk_z z}$) going waves for $z \leq 0$. As a result, it is necessary to divide region 0 into two regions, a 0+ region for $z \geq 0$ and a 0− region for $z \leq 0$, and to define separate coefficients for each region. For example, instead of one coefficient A_0, we require two coefficients A_{0-} and A_{0+}. In region 0+ we have

$$A_{0+} = E_{hed} + A_0^h \tag{B.34}$$

$$B_{0+} = B_0^h \tag{B.35}$$

$$C_{0+} = H_{hed} + C_0^h \tag{B.36}$$

$$D_{0+} = D_0^h, \tag{B.37}$$

where A_0^h, B_0^h, C_0^h, and D_0^h are the homogeneous solutions for region 0. Similarly, in region 0− we have

$$A_{0-} = A_0^h \tag{B.38}$$

$$B_{0-} = -E_{hed} + B_0^h \tag{B.39}$$

$$C_{0-} = C_0^h \tag{B.40}$$
$$D_{0-} = H_{hed} + D_0^h. \tag{B.41}$$

We assume that upward going wave amplitudes in region 0+ can be expressed as the reflections of the downward going wave amplitudes at the $z = d_1$ interface. That is,

$$B_{0+} = R_{0,1}^{TM} A_{0+} \tag{B.42}$$
$$D_{0+} = R_{0,1}^{TE} C_{0+}, \tag{B.43}$$

where the notation $R_{0,1}^{TM}$ denotes the TM mode reflection coefficient for waves traveling from region 0 to region 1. In region 0−, the downward going wave amplitudes are the reflected quantities so that

$$A_{0-} = R_{0,-1}^{TM} B_{0-} \tag{B.44}$$
$$C_{0-} = R_{0,-1}^{TE} D_{0-}. \tag{B.45}$$

Making use of the above expressions, we can solve for the wave amplitudes. In region 0+ we obtain

$$A_{0+} = \frac{\left(1 - R_{0,-1}^{TM}\right)}{1 - R_{0,-1}^{TM} R_{0,1}^{TM}} E_{hed} \tag{B.46}$$
$$B_{0+} = \frac{R_{0,1}^{TM}\left(1 - R_{0,-1}^{TM}\right)}{1 - R_{0,-1}^{TM} R_{0,1}^{TM}} E_{hed} \tag{B.47}$$
$$C_{0+} = \frac{\left(1 + R_{0,-1}^{TE}\right)}{1 - R_{0,-1}^{TE} R_{0,1}^{TE}} H_{hed} \tag{B.48}$$
$$D_{0+} = \frac{R_{0,1}^{TE}\left(1 + R_{0,-1}^{TE}\right)}{1 - R_{0,-1}^{TE} R_{0,1}^{TE}} H_{hed}, \tag{B.49}$$

while in region 0− we have

$$A_{0-} = -\frac{R_{0,-1}^{TM}\left(1 - R_{0,1}^{TM}\right)}{1 - R_{0,-1}^{TM} R_{0,1}^{TM}} E_{hed} \tag{B.50}$$
$$B_{0-} = -\frac{\left(1 - R_{0,1}^{TM}\right)}{1 - R_{0,-1}^{TM} R_{0,1}^{TM}} E_{hed} \tag{B.51}$$

$$C_{0-} = \frac{R^{TE}_{0,-1}\left(1+R^{TE}_{0,1}\right)}{1-R^{TE}_{0,-1}R^{TE}_{0,1}}H_{hed} \tag{B.52}$$

$$D_{0-} = \frac{\left(1+R^{TE}_{0,1}\right)}{1-R^{TE}_{0,-1}R^{TE}_{0,1}}H_{hed}. \tag{B.53}$$

B.5 Reflection Coefficients

We have found that the calculation of the reflection coefficients is most readily understood with a transmission line analogy [*Wait*, 1970]. For the TM mode, we define a generalized voltage and current in the ith region as

$$V^{TM}_i(z) = jk_{i,z}\left(A_i e^{-jk_{i,z}z} - B_i e^{+jk_{i,z}z}\right) \tag{B.54}$$

$$I^{TM}_i(z) = j\omega\bar{\epsilon}_i\left(A_i e^{-jk_{i,z}z} + B_i e^{+jk_{i,z}z}\right), \tag{B.55}$$

where V^{TM}_i and I^{TM}_i are proportional to tangential $\mathbf{E}^{TM}_i$ and $\mathbf{H}^{TM}_i$, respectively. Then a generalized admittance for the TM mode is

$$Y^{TM}_i(z) = \frac{I^{TM}(z)}{V^{TM}(z)} \tag{B.56}$$

$$= \hat{Y}^{TM}_i \frac{A_i e^{-jk_{i,z}z} + B_i e^{jk_{i,z}z}}{A_i e^{-jk_{i,z}z} - B_i e^{jk_{i,z}z}}, \tag{B.57}$$

where

$$\hat{Y}^{TM}_i = \frac{\omega\bar{\epsilon}_i}{k_{i,z}} \tag{B.58}$$

is the characteristic admittance of region i. We first consider the reflection coefficient for downward-going waves $R^{TM}_{i,i+1} = \frac{B_i}{A_i}$. We assume that the admittance at the lower boundary $z = d_i$ is known and can be considered to be a load admittance Y^L_i. We have the relation

$$Y^L_i = \hat{Y}^{TM}_i \frac{1+R^{TM}_{i,i+1}e^{j2k_{i,z}d_i}}{1-R^{TM}_{i,i+1}e^{j2k_{i,z}d_i}}. \tag{B.59}$$

Solving for $R^{TM}_{i,i+1}$, we obtain

$$R^{TM}_{i,i+1} = \frac{Y^L_i - \hat{Y}^{TM}_i}{Y^L_i + \hat{Y}^{TM}_i} e^{-j2k_{i,z}d_i}. \tag{B.60}$$

Because of the continuity of tangential $\mathbf{E}_i^{TM}$ and $\mathbf{H}_i^{TM}$ at the boundary $z = d_i$, the load admittance may be written $Y_i^L = Y_{i+1}^{TM}(d_i)$ where

$$Y_{i+1}^{TM}(d_i) = \hat{Y}_{i+1}^{TM} \frac{1 + R_{i+1,i+2}^{TM} e^{j2k_{i+1,z}d_i}}{1 - R_{i+1,i+2}^{TM} e^{j2k_{i_1,z}d_i}}. \tag{B.61}$$

By comparison with (B.60), we may write

$$R_{i+1,i+2}^{TM} = \frac{Y_{i+2}^{TM}(d_{i+1}) - \hat{Y}_{i+1}^{TM}}{Y_{i+2}^{TM}(d_{i+1}) + \hat{Y}_{i+1}^{TM}} e^{-j2k_{i+1,z}d_{i+1}}. \tag{B.62}$$

It is clear that the process can be repeated until we reach the final boundary $z = d_N$, where

$$\begin{aligned} Y_{N+1}^{TM}(d_N) &= \hat{Y}_{N+1}^{TM} \frac{A_{N+1} e^{-jk_{N+1,z}d_N}}{A_{N+1} e^{-jk_{N+1,z}d_N}} && \text{(B.63)} \\ &= \hat{Y}_{N+1}^{TM} && \text{(B.64)} \end{aligned}$$

since there are no upward going waves in the $(N+1)$th region. By substituting (B.62) into (B.61), we can make the recursion process explicit and obtain the expression

$$\begin{aligned} Y_{i+1}^{TM}(d_i) &= \hat{Y}_{i+1}^{TM} \times && \text{(B.65)} \\ &\quad \frac{\left(Y_{i+2}^{TM}(d_{i+1}) + \hat{Y}_{i+1}^{TM}\right) + \left(Y_{i+2}^{TM}(d_{i+1}) - \hat{Y}_{i+1}^{TM}\right) e^{-j2k_{i+1,z}h_{i+1}}}{\left(Y_{i+2}^{TM}(d_{i+1}) + \hat{Y}_{i+1}^{TM}\right) - \left(Y_{i+2}^{TM}(d_{i+1}) - \hat{Y}_{i+1}^{TM}\right) e^{-j2k_{i+1,z}h_{i+1}}} \\ &= \hat{Y}_{i+1}^{TM} \frac{Y_{i+2}^{TM}(d_{i+1}) + \hat{Y}_{i+1}^{TM} \tanh(jk_{i+1,z}h_{i+1})}{Y_{i+2}^{TM}(d_{i+1}) \tanh(jk_{i+1,z}h_{i+1}) + \hat{Y}_{i+1}^{TM}}, && \text{(B.66)} \end{aligned}$$

where $h_{i+1} = d_{i+1} - d_i$. *Ward and Hohmann* [1987] note that the use of negative exponentials (Equation B.65), may provide greater numerical stability than the use of hyperbolic tangent functions (Equation B.66).

Using similar arguments, we can show that the TM mode reflection coefficient for upward going waves is given by

$$R_{i,i-1}^{TM} = \frac{Y_{i-1}^{TM}(d_{i-1}) - \hat{Y}_i^{TM}}{Y_{i-1}^{TM}(d_{i-1}) + \hat{Y}_i^{TM}} e^{j2k_{i,z}d_{i-1}}, \tag{B.67}$$

where

$$Y_{i-1}^{TM}(d_{i-1}) = \hat{Y}_{i-1}^{TM} \frac{1 + R_{i-1,i-2}^{TM} e^{-j2k_{i-1,z}d_{i-1}}}{1 - R_{i-1,i-2}^{TM} e^{-j2k_{i-1,z}d_{i-1}}}. \tag{B.68}$$

The expressions for the TE mode reflection coefficients can be obtained from the TM mode equations by replacing $\hat{Y}_i^{TM}$ and $Y_i^{TM}(z)$ by $\hat{Z}_i^{TE}$ and $Z_i^{TE}(z)$, respectively, where

$$\hat{Z}_i^{TE} = \frac{\omega\mu_i}{k_{i,z}} \tag{B.69}$$

is the characteristic impedance of region i and

$$Z_i^{TE}(z) = \hat{Z}_i^{TE} \frac{C_i e^{-jk_{i,z}z} + D_i e^{jk_{i,z}z}}{C_i e^{-jk_{i,z}z} - D_i e^{jk_{i,z}z}} \tag{B.70}$$

is the impedance as a function of z. For example, we may write

$$R_{i,i+1}^{TE} = \frac{Z_{i+1}^{TE}(d_i) - \hat{Z}_i^{TE}}{Z_{i+1}^{TE}(d_i) + \hat{Z}_i^{TE}} e^{-j2k_{i,z}d_i}. \tag{B.71}$$

B.6 Propagation Matrices

Once we have determined the wave amplitudes in region 0 we may use the boundary conditions to calculates the amplitudes in all other regions. Conditions (B.30) and (B.31) may be used to express A_{i+1} and B_{i+1} in terms of A_i and B_i or vice versa. Similarly, conditions (B.32) and (B.33) can be used to express C_{i+1} and D_{i+1} in terms of C_i and D_i or vice versa. These relations can be written in matrix format as follows

$$\begin{aligned}
\begin{bmatrix} A_{i+1} \\ B_{i+1} \end{bmatrix} &= \frac{1}{2} \begin{bmatrix} (\alpha_i + \beta_i)\, e^{j\kappa_i^- d_i} & (\alpha_i - \beta_i)\, e^{j\kappa_i^+ d_i} \\ (\alpha_i - \beta_i)\, e^{-j\kappa_i^+ d_i} & (\alpha_i + \beta_i)\, e^{-j\kappa_i^- d_i} \end{bmatrix} \begin{bmatrix} A_i \\ B_i \end{bmatrix} \\
\begin{bmatrix} A_i \\ B_i \end{bmatrix} &= \frac{1}{2} \begin{bmatrix} \left(\alpha_i^{-1} + \beta_i^{-1}\right) e^{-j\kappa_i^- d_i} & \left(\alpha_i^{-1} - \beta_i^{-1}\right) e^{j\kappa_i^+ d_i} \\ \left(\alpha_i^{-1} - \beta_i^{-1}\right) e^{-j\kappa_i^+ d_i} & \left(\alpha_i^{-1} + \beta_i^{-1}\right) e^{j\kappa_i^- d_i} \end{bmatrix} \begin{bmatrix} A_{i+1} \\ B_{i+1} \end{bmatrix} \\
\begin{bmatrix} C_{i+1} \\ D_{i+1} \end{bmatrix} &= \frac{1}{2} \begin{bmatrix} (\xi_i + \beta_i)\, e^{j\kappa_i^- d_i} & (\xi_i - \beta_i)\, e^{j\kappa_i^+ d_i} \\ (\xi_i - \beta_i)\, e^{-j\kappa_i^+ d_i} & (\xi_i + \beta_i)\, e^{-j\kappa_i^- d_i} \end{bmatrix} \begin{bmatrix} C_i \\ D_i \end{bmatrix} \\
\begin{bmatrix} C_i \\ D_i \end{bmatrix} &= \frac{1}{2} \begin{bmatrix} \left(\xi_i^{-1} + \beta_i^{-1}\right) e^{-j\kappa_i^- d_i} & \left(\xi_i^{-1} - \beta_i^{-1}\right) e^{j\kappa_i^+ d_i} \\ \left(\xi_i^{-1} - \beta_i^{-1}\right) e^{-j\kappa_i^+ d_i} & \left(\xi_i^{-1} + \beta_i^{-1}\right) e^{j\kappa_i^- d_i} \end{bmatrix} \begin{bmatrix} C_{i+1} \\ D_{i+1} \end{bmatrix},
\end{aligned}$$

where

$$\begin{aligned}
\alpha_i &= \frac{\bar{\epsilon}_i}{\bar{\epsilon}_{i+1}} \\
\beta_i &= \frac{k_{i,z}}{k_{i+1,z}} \\
\xi_i &= \frac{\mu_i}{\mu_{i+1}} \\
\kappa_i^+ &= k_{i+1,z} + k_{i,z} \\
\kappa_i^- &= k_{i+1,z} - k_{i,z}.
\end{aligned}$$

Appendix C

Derivation of $\mathcal{E}_k = \mathcal{E}_0$ for all k

We show that $\mathcal{E}_k = \sum_{m,n} 2^{-\gamma m} \left(S_{m,n}^{[J_k],\{k,0\}} \right)^2$ is equal to a constant $\mathcal{E}_0$ for all k if $s_{k,l}(t) = 2^{\gamma k/2} \xi_k(t)$. First we write $\mathcal{E}_k$ in terms of the UDWT $\tilde{S}$,

$$\mathcal{E}_k = \sum_{m,n} 2^{-\gamma m} \left(\tilde{S}_{m,2^m n + J_k}^{\{k,0\}} \right)^2. \tag{C.1}$$

It is straightforward to derive the identity

$$\langle \xi_{k+1}(t), \psi_{m+1}(t-\tau) \rangle = \langle \xi_k(t), \psi_m(t-\tau/2) \rangle. \tag{C.2}$$

Using this identity in conjunction with (6.4) and (6.21), we obtain

$$\tilde{S}_{m+1,2^{m+1}n+2j}^{\{k+1,0\}} = 2^{\gamma/2} \tilde{S}_{m,2^m n+j}^{\{k,0\}}. \tag{C.3}$$

We now show that $\mathcal{E}_k = \mathcal{E}_{k+1}$.

$$\begin{aligned}
\mathcal{E}_k &= \sum_{m,n} 2^{-\gamma m} \left(\tilde{S}_{m,2^m n + J_k}^{\{k,0\}} \right)^2 && \text{(C.4)} \\
&= 2^{\gamma} \sum_{m,n} 2^{-\gamma(m+1)} \left(2^{-\gamma/2} \tilde{S}_{m+1,2^{m+1} n + 2J_k}^{\{k+1,0\}} \right)^2 && \text{(C.5)} \\
&= \sum_{m',n} 2^{-\gamma m'} \left(\tilde{S}_{m',2^{m'} n + 2J_k}^{\{k+1,0\}} \right)^2 && \text{(C.6)} \\
&= \mathcal{E}_{k+1}. && \text{(C.7)}
\end{aligned}$$

Note that we assume infinite limits on the summation index m in order to justify the equality after the substitution $m' = m + 1$. In practice, m lies within a finite range, and the equality becomes an approximation. The quality of the approximation can be tested numerically, and the limits on m can usually be extended to achieved a desired level of accuracy. We have also used the fact that the shift $J_{k+1} = 2J_k$. To prove this fact let us assume that the shift J_{k+1} that corresponds to $\mathcal{E}_{k+1}$ satisfies the relation $J_{k+1} \neq 2J_k$. From the definition of $\mathcal{E}_{k+1}$ in (6.15), we have

$$\sum_{m,n} 2^{-\gamma m} \left(\tilde{S}^{\{k+1,0\}}_{m,2^m n + J_{k+1}} \right)^2 > \sum_{m,n} 2^{-\gamma m} \left(\tilde{S}^{\{k+1,0\}}_{m,2^m n + 2J_k} \right)^2 . \tag{C.8}$$

Using (C.4), (C.5), and (C.6) in reverse order, we obtain

$$\sum_{m,n} 2^{-\gamma m} \left(\tilde{S}^{\{k,0\}}_{m,2^m n + (J_{k+1})/2} \right)^2 > \sum_{m,n} 2^{-\gamma m} \left(\tilde{S}^{\{k,0\}}_{m,2^m n + J_k} \right)^2 . \tag{C.9}$$

By definition, however, J_k is the shift that maximizes the sum $\sum_{m,n} 2^{-\gamma m} \left(\tilde{S}^{\{k,0\}}_{m,2^m n + J} \right)^2$. We conclude, therefore, that our assumption is incorrect and that $J_{k+1} = 2J_k$.

Bibliography

P. Abry and P. Flandrin. Multiresolution transient detection. In *Proc. of IEEE-SP Intl. Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA*, pages 225–8, 1994.

P. Abry, P. Goncalves, and P. Flandrin. Wavelets, spectrum analysis, and $1/f$ processes. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*, pages 15–29. Springer-Verlag, 1995.

M. Alexandrescu, D. Gibert, G. Hulot, J. Le Mouël, and G. Saracco. Detection of geomagnetic jerks using wavelet analysis. *J. Geophys. Res.*, 100(B7):12557–12572, July 10, 1995.

M.R. Banham and A.K. Katsaggelos. Digital image restoration. *IEEE Signal Processing Magazine*, 14(2):24–41, March 1997.

BART Engineering Dept. Technical schematic: Traction power electrification system wide, 1996.

A. Bernardi, A.C. Fraser-Smith, and O.G. Villard, Jr. Measurements of the BART magnetic fields with an automatic geomagnetic pulsation index generator. *IEEE Transactions on Electromagnetic Compatibility*, 31(4):413–417, November 1989.

H.E. Bomar, R.O. Dean, J.A. Hanck, M.D. Orton, and P.L. Todd. Bay area rapid transit system (BART). *Materials Performance*, pages 9–17, December 1974.

R.N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, 1986.

T.M. Brocher, J. McCarthy, P.E. Hart, W.S. Holbrook, K.P. Furlong, T.V. McEvilly, J.A. Hole, and S.L. Klemperer. Seismic evidence for a lower-crustal detachment beneath San Francisco Bay, California. *Science*, 265:1436–1439, September 2, 1994.

J. Buckheit and D.L. Donoho. Wavelab and reproducible research. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*. Springer-Verlag, 1995.

J.E. Burbank. Earth-currents and a proposed method for their investigation. *Terrestrial Magnetism and Atmospheric Electricity*, 10:23–49, 1905.

Y. Cai, M.R. Irving, and S.H. Case. Modeling and numerical solution of multibranched DC rail traction power systems. *IEE Proc. Electr. Power Appl.*, 142(5):323–328, September 1995.

W.H. Campbell. Geomagnetic pulsations. In S. Matsushita and W.H. Campbell, editors, *Physics of Geomagnetic Phenomena*, pages 822–912. Academic Press, 1967.

W.H. Campbell. An analysis of the spectra of geomagnetic variations having periods from 5 min to 4 hours. *J. Geophys. Res.*, 81(7):1369–1390, March 1, 1976.

W.H. Campbell. *Introduction to Geomagnetic Fields*. Cambridge University Press, 1997.

L. Chaize and M. Lavergne. Signal et bruit en magnetotellurique. *Geophysical Prospecting*, 18:64–87, 1970.

W.C. Chew. *Waves and Fields in Inhomogeneous Media*. Van Nostrand Reinhold, 1990.

N. Chin, A. Barreto, J. Riley, and J. Andrian. Temporal multiresolution detection of epileptic spikes based on wavelet transformation. In *Proceedings of SPIE*, volume 2491, pages 819–828, 1994.

G.L. Choy and V.F. Cormier. Direct measurements of the mantle attenuation operator from broadband P and S waveforms. *J. Geophys. Res.*, 91(B7):7326–7342, December 10, 1986.

D.A. Chrissan. *Statistical Analysis and Modeling of Low-Frequency Radio Noise and Improvement of Low-Frequency Communications*. PhD thesis, Department of Electrical Engineering, Stanford University, 1998.

R.R. Coifman and D.L. Donoho. Translation-invariant de-noising. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*. Springer-Verlag, 1995.

G. Corsini and R. Saletti. Design of a digital $1/f^{\gamma}$ noise simulator. In *Noise in Physical Systems and* $1/f$ *Noise*, pages 82–86. World Scientific, 1987.

Z. Cvetković and M. Vetterli. Discrete-time wavelet extrema representation: Design and consistent reconstruction. *IEEE Trans. Signal Processing*, 43(3):681–693, March 1995.

I. Daubechies. *Ten Lectures on Wavelets.* SIAM, 1992.

M.J. Davidson. Average diurnal characteristics of geomagnetic power spectrum in the period range 4.5 to 1000 seconds. *J. Geophys. Res.*, 69(23):5116–5119, December 1, 1964.

G. DuPouy. Perturbation du champ magnétique terrestre et des courants telluriques par les chemins de fer électrifiés. *Annales de Géophysique*, 6(1):18–50, 1950.

D. Eberhart-Phillips, V.F. Labson, W.D. Stanley, A.J. Michael, and B.D. Rodriguez. Preliminary velocity and resistivity models of the Loma Prieta earthquake region. *Geophys. Res. Letters*, 17(8):1235–1238, July 1990.

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

G.D. Egbert and J.R. Booker. Multivariate analysis of geomagnetic array data 1. The response space. *J. Geophys. Res.*, 94(B10):14227–14247, October 10, 1989.

J.D. Evans. Preliminary simulation studies to assess energy savings possible with AATC, reduced braking algorithm. Internal BART Memo, October 14, 1993.

J.D. Evans. BART Engineering Department, Personal Communication, 1996.

P. Flandrin. Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Theory*, 38(2):910–917, March 1992.

J.B. Flowers. Load sharing with thyristor controlled rectifier substations. In *Proc. of the 1995 IEEE/ASME Joint Railroad Conference*, pages 69–73, 1995.

S.L. Fontes, T. Harinarayana, G.J.K. Dawes, and V.R.S. Hutton. Processing of noisy magnetotelluric data using digital filters and additional data selection criteria. *Physics of the Earth and Planetary Interiors*, 52:30–40, 1988.

J. Francheteau. Oceanic crust. *Scientific American*, 249(3):114–129, September 1983.

A.C. Fraser-Smith, A. Bernardi, R.A. Helliwell, P.R. McGill, and O.G. Villard, Jr. Analysis of low-frequency-electromagnetic-field measurements near the epicenter. In *The Loma Prieta, California Earthquake of October 17, 1989 — Preseismic Observations*, USGS Professional Paper 1550-C, pages 17–26. U.S. Gov. Printing Office, 1993.

A.C. Fraser-Smith, A. Bernardi, P.R. McGill, M.E. Ladd, R.A. Helliwell, and O.G. Villard, Jr. Low-frequency magnetic field measurements near the epicenter of the M_S 7.1 Loma Prieta earthquake. *Geophys. Res. Letters*, 17(9):1465–1468, Aug 1990.

A.C. Fraser-Smith and D.M. Bubenik. Compendium of the ULF/ELF electromagnetic fields generated above a sea of finite depth by submerged harmonic dipoles. Technical Report E715-1, Stanford Electronics Laboratories, January 1980.

A.C. Fraser-Smith and D.B. Coates. Large amplitude ULF electromagnetic fields from BART. *Radio Science*, 13(4):661–668, July-August 1978.

B. Friedlander and B. Porat. Performance analysis of transient detectors based on a class of linear data transforms. *IEEE Trans. Inform. Theory*, 38(2):665–673, March 1992.

G.D. Friedlander. The BART chronicle. *IEEE Spectrum*, pages 34–46, September 1972.

M. Frisch and H. Messer. The use of the wavelet transform in the detection of an unknown transient signal. *IEEE Trans. Inform. Theory*, 38(2):892–897, March 1992.

M. Frisch and H. Messer. Detection of a known transient signal of unknown scaling and arrival time. *IEEE Trans. Signal Processing*, 42(7):1859–1863, July 1994.

F.C. Frischknecht. Fields about an oscillating magnetic dipole over a two-layer earth. *Colorado School of Mines Quart.*, 62(1), 1967.

D. Galler and P.L. Todd. Improved rail-fastener insulation for stray-current control. In *Corrosion Forms and Control For Infrastructure*, pages 170–182. ASTM, 1992.

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

H.B. Goldman. *Geology of San Francisco Bay.* San Francisco Bay Conservation and Development Commission, 1967.

M.M. Goldman. *Non-Conventional Methods in Geoelectrical Prospecting.* Ellis Horwood Ltd., 1990.

F.S. Grant and G.F. West. *Interpretation Theory in Applied Geophysics.* McGraw-Hill, 1965.

H.G. Greene and M.P. Kennedy. *Geologic Map of the Central California Continental Margin.* California Dept. of Conservation, Div. of Mines and Geology, 1989.

F. Gruneis and T. Musha. Clustering Poisson process and $1/f$ noise. *Japanese Journal of Applied Physics*, 25(10):1510–1519, October 1986.

S. Gruszow, J.C. Rossignol, A. Tzanis, and J.L. Le Mouel. Identification and analysis of electromagnetic signals in Greece: the case of the Kozani earthquake VAN prediction. *Geophys. Res. Letters*, 23(16):2025–2028, August 1, 1996.

D. Harvey and G.L. Choy. Broad-band deconvolution of GSDN data. *Geophysical Journal of the Royal Astronomical Society*, 69:659–668, 1982.

M. Hattingh. A new data adaptive filtering program to remove noise from geophysical time-or-space-series data. *Computers and Geosciences*, 14(4):467–480, 1988.

H.A. Haus and J.R. Melcher. *Electromagnetic Fields and Energy.* Prentice Hall, 1989.

J.F. Hermance. Electrical conductivity models of the crust and mantle. In *Global Earth Physics: A Handbook of Physical Constants*, pages 190–205. American Geophysical Union, 1995.

R.J. Hill. Iterative techniques for the solution of complex DC-rail traction systems including regenerative braking. *IEE Proc. Gener. Transm. Distrib.*, 143(6):613–614, November 1996.

A.M.-H Ho, A.C. Fraser-Smith, and O.G. Villard, Jr. Large-amplitude ULF magnetic fields produced by a rapid transit system: Close-range measurements. *Radio Science*, 14(6):1011–1015, Nov-Dec 1979.

G.W. Hohmann. Numerical modeling for electromagnetic methods of geophysics. In *Electromagnetic Methods in Applied Geophysics*, volume 1, pages 313–363. Soc. of Exploration Geophysicists, 1987.

W.S. Holbrook, T.M. Brocher, U.S. ten Brink, and J.A. Hole. Crustal structure of a transform plate boundary: San Francisco Bay and the central California continental margin. *J. Geophys. Res.*, 101(B10):22311–22334, October 10, 1996.

B.R. Hunt. Deconvolution of linear systems by constrained regression and its relationship to the Wiener theory. *IEEE Transactions on Automatic Control*, 17(5):703–705, October 1972.

J.D. Jackson. *Classical Electrodynamics*. John Wiley, 1975.

J.A. Jacobs. *Geomagnetic Pulsations*. Springer-Verlag, 1970.

I.M. Johnstone and B.W. Silverman. Wavelet threshold estimators for data with correlated noise. Technical report, Dept. of Statistics, Stanford University, 1996.

A. Junge. Characterization and correction for cultural noise. *Surveys in Geophysics*, 17: 361–391, 1996.

T. Kailath. *Lectures on Wiener and Kalman Filtering*. Springer-Verlag, 1981.

J. Kauahikaua. Electromagnetic fields about a horizontal electric wire source of arbitrary length. *Geophysics*, 43(5):1019–1022, August 1978.

A.A. Kaufman and G.V. Keller. *Frequency and Transient Soundings*. Elsevier, 1983.

S.M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall PTR, 1998.

M.S. Kershner. $1/f$ Noise. *Proc. IEEE*, 70(3):212–218, March 1982.

D.P. Kitzinger, D.R. Breding, and D.M. Etter. Removing seismometer effects using inverse filtering techniques. In *Conference Record of the Twenty-Second Asilomar Conference on Signals, Systems, and Computers*, volume 2, pages 863–866, 1988.

J.A. Kong. Antenna radiation in stratified media. In *Research Topics in Electromagnetic Wave Theory*, pages 211–234. Wiley-Interscience, 1981.

R.D. Kurtz, J.M. DeLaurier, and J.C. Gupta. The electrical conductivity distribution beneath Vancouver Island: a region of active plate subduction. *J. Geophys. Res.*, 95 (B7):10929–10946, July 10, 1990.

D. La Cour and E. Hoge. Note sur les effets de l'électrification du chemin de fer passant dans le voisinage de l'observatoire magnétique de copenhague. In *Transactions of the Edinburgh Meeting, Union géodesique et géophysique internationale*, volume 10, pages 302–306, 1937.

L.J. Lanzerotti, C.G. Maclennan, and A.C. Fraser-Smith. Background magnetic spectra: 10^{-5} to 10^{5} Hz. *Geophys. Res. Letters*, 17(10):1593–1596, September 1990.

J.C. Larsen, R.L. Mackie, A. Manzella, A. Fiordelisi, and S. Rieven. Robust smooth magnetotelluric transfer functions. *Geophys. J. Intl.*, 124:801–819, 1996.

J.J. Lienkaemper. Field trip guide to the Hayward Fault. Open-File Report 89-500, USGS, 1989.

J.J. Lienkaemper, G. Borchadt, and M. Lisowski. Historic creep rate and potential for seismic slip along the Hayward Fault, California. *J. Geophys. Res.*, 96(B11):18261–18283, October 10, 1991.

T.T. Liu. Filter-bank calculations for the Berkeley data set. Technical Note, available at *http://www-star.stanford.edu/~ttliu/*, 1996.

T.T. Liu and A.C. Fraser-Smith. Hayward fault earthquake prediction project: ULF magnetic field measurements. Technical Report D180-1, STARLAB, Stanford University, December 1996a.

T.T. Liu and A.C. Fraser-Smith. Techniques for monitoring ULF geomagnetic fields in the presence of interference due to a mass transit system (BART). *EOS, Transactions, American Geophysical Union*, 77(47, Suppl.):457, November 12, 1996b.

T.T. Liu and A.C. Fraser-Smith. Detection of transients in $1/f$ noise with the undecimated discrete wavelet transform. Submitted to IEEE Trans. Signal Processing, 1998a.

T.T. Liu and A.C. Fraser-Smith. Identification and removal of man-made transients from geomagnetic array time series: A wavelet transform based approach. In *Proc. of the 32nd Asilomar Conf. on Signals, Systems, and Computers*, volume 3, pages 1363–1367, 1998b.

T.T. Liu and A.C. Fraser-Smith. An undecimated wavelet transform based detector for transients in $1/f$ noise. In *Proc. of the 1999 IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, volume 3, 1998c.

R.L. Mackie, B.R. Bennet, and T.R. Madden. Long-period magnetotelluric measurements near the central California coast: a land-locked view of the conductivity structure under the Pacific ocean. *Geophys. J.*, 95:181–194, 1988.

T.R. Madden, G.A. LaTorraca, and S.K. Park. Electrical conductivity variations around the Palmdale section of the San Andreas fault. *J. Geophys. Res.*, 98(B1):795–808, January 10, 1993.

S. Mallat. Zero-crossings of a wavelet transform. *IEEE Trans. Inform. Theory*, 37(4): 1019–1033, July 1991.

S. Mallat and W.L. Hwang. Singularity detection and processing with wavelets. *IEEE Trans. Inform. Theory*, 38(2):617–643, Mar 1992.

G.A. McMechan and I. Barrodale. Processing electromagnetic data in the time domain. *Geophys. J. R. Astron. Soc.*, 81:277–293, 1985.

W. Menke. *Geophysical Data Analysis: Discrete Inverse Theory.* Academic Press, 1989.

A. Mohsen. On the evaluation of Sommerfeld integrals. *IEE Proc.*, 129(4), August 1982.

H.F. Morrison, U. Conti, V.F. Labson, E. Nichols, and E. Goldstein. Field tests of noise in SQUID and induction coil magnetometers. Report 901, Lawrence Berkeley Laboratory, U.C. Berkeley, 1984.

H.F. Morrison, R.F. Corwin, and M. Chang. High-accuracy determination of temporal variations of crustal resistivity. In *The Earth's Crust*, pages 593–614. American Geophysical Union, 1977.

E. Nishinaga, J.A. Evans, and G.L. Mayhew. Wireless advanced automatic train control. In *Proceedings of the 1994 ASME/IEEE Joint Railroad Conference*, pages 31–46, 1994.

D.W. Oldenburg. A comprehensive solution to the linear deconvolution problem. *Geophys. J. R. Astron. Soc.*, 65:331–357, 1981.

A.V. Oppenheim and R.W. Schafer. *Discrete-time Signal Processing.* Prentice Hall, 1989.

G.J. Palacky. Resistivity characteristics of geologic targets. In *Electromagnetic Methods in Applied Geophysics*, volume 1, pages 53–129. Society of Exploration Geophysicists, 1987.

S.K. Park. Precursors to earthquakes: seismoelectromagnetic signals. *Surveys in Geophysics*, 17:493–516, 1996.

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

W.J. Phillips and A.F. Kuckes. Electrical conductivity structure of the San Andreas fault in central California. *J. Geophys. Res.*, 88(B9):7467–7474, September 10, 1983.

W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C.* Cambridge University Press, 1988.

S. Ramo, J.R. Whinnery, and T.Van Duzer. *Fields and Waves in Communication Electronics.* John Wiley and Sons, 1984.

S.M. Riad. The deconvolution problem: An overview. *Proceedings of the IEEE*, 74(1): 82–85, January 1986.

J.A. Rice. *Mathematical Statistics and Data Analysis.* Duxbury Press, 1995.

J. Riordan. Current propagation in electric railway propulsion systems. *Trans. Amer. Inst. of Elec. Engineers*, 51(4):1011–1019, December 1932.

B.D. Ripley. *Pattern Recognition and Neural Networks.* Cambridge University Press, 1996.

D. Robertson, O.I. Camps, and J. Mayer. Wavelets and power system transients: feature detection and classification. In *Proceedings of SPIE*, volume 2242, pages 474–487, 1994.

J.D. Rogers and S.H. Figuers. Late Quaternary stratigraphy of the East Bay plain. In *Proc. of the Second Conf. on Earthquake Hazards in Eastern San Francisco Bay Area*, volume 113, pages 19–27. California Dept. of Conservation, Div. of Mines and Geology Special Publication, 1992.

P.J. Rousseuw and A.M. Leroy. *Robust Regression and Outlier Detection.* Wiley, 1987.

G. Santarato and U. Spagnolini. Cancelling directional EM noise in magnetotellurics. *Geophysical Prospecting*, 43:605–621, 1995.

F. Scherbaum. *Of Poles and Zeros: Fundamentals of Digital Seismology.* Kluwer Academic Publishers, 1996.

F. Scherbaum and M. Bouin. FIR filter effects and nucleation phases. *Geophys. J. Intl.*, 130:661–668, 1997.

U. Schmucker. *Anomalies of Geomagnetic Variations in the Southwestern United States*, volume 13 of *Bull. of the Scripps Inst. of Oceanography.* Univ. of Calif. Press, 1970.

R.E. Shaffer. Design of DC powered rail transit systems to minimize stray currents. *Materials Performance*, pages 17–22, September 1982.

R.E. Shaffer. Stray current control within DC-powered transit systems. In *Rapid Rail Transit and Planning Tools*, volume 1152 of *Transportation Research Record*, pages 42–48. Transportation Research Board, 1987.

R.E. Shaffer and J.H. Fitzgerald. Stray earth current control Washington, D.C. metro system. *Materials Performance*, pages 9–15, April 1981.

S.P. Sipkin and A.L. Lerner-Lam. Pulse-shape distortion introduced by broadband deconvolution. *Bulletin of the Seismological Society of America*, 82(1):238–258, February 1992.

A. Sommerfeld. *Partial Differential Equations in Physics.* Academic Press, 1949.

U. Spagnolini. Time-domain estimation of MT impedance tensor. *Geophysics*, 59(5): 712–721, May 1994.

S.D. Stoilov, J.S. Burns, and P.L. Todd. Traction power system analysis at interfered headways. In *APTA 1994 Rapid Transit Conference*, 1994.

E.D. Sunde. Currents and potential along leaky ground-return conductors. *Electrical Engineering*, pages 1338–1346, December 1936.

Laszlo Szarka. Geophysical aspects of man-made electromagnetic noise in the earth — a review. *Surveys in Geophysics*, 9:287–318, 1988.

A.H. Tewfik and M. Kim. Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. *IEEE Trans. Inform. Theory*, 38(2), March 1992.

P.L. Todd. BART Engineering Department, Personal Communication, 1997.

S. Treitel and L.R. Lines. Linear inverse theory and deconvolution. *Geophysics*, 47(8): 1153–1159, August 1982.

S. Twomey. *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*. Elsevier, 1997.

M.J. Unsworth, P.E. Malin, G.D. Egbert, and J.R. Booker. Internal structure of the San Andreas fault at Parkfield, California. *Geology*, 25(4):359–356, April 1997.

S. Uyeda. Japanese research on electro-magnetic short-term earthquake prediction. *Seismological Research Letters*, 68(2):27, April 1997.

M. Vetterli and J. Kovačević. *Wavelets and subband coding*. Prentice Hall PTR, 1995.

K. Vozoff. The magnetotelluric method. In *Electromagnetic Methods in Applied Geophysics*, volume 2, pages 641–711. Society of Exploration Geophysicists, 1991.

D.L. Wagner, E.J. Bortugno, and R.D. McJunkin. *Geologic Map of the San Francisco–San Jose Quadrangle*. California Dept. of Conservation, Div. of Mines and Geology, 1990.

C. Wahrhaftig and D. Sloan, editors. *Geology of San Francisco And Vicinity*. American Geophysical Union, 1989.

J.R. Wait. *Electromagnetic Waves in Stratified Media*. Pergamon Press, 1970.

J.R. Wait. *Geo-Electromagnetism*. Academic Press, 1982.

P.E. Wannamaker, J.R. Booker, A.G. Jones, A.D. Chave, J.H. Filloux, H.S. Waff, and L.K. Law. Resistivity cross section through the Juan de Fuca subduction system and its tectonic implication. *J. Geophys. Res.*, 94(B10):14127–14144, October 10, 1989.

S.H. Ward and G.W. Hohmann. Electromagnetic theory for geophysical applications. In *Electromagnetic Methods in Applied Geophysics*, volume 1, pages 131–312. Soc. of Exploration Geophysicists, 1987.

R. Wertz and W.H. Campbell. Integrated power spectra of geomagnetic field variations with periods of 0.3-300 s. *J. Geophys. Res.*, 81(28):5131–5140, October 1, 1976.

B. Widrow et al. Adaptive noise cancelling: Principles and applications. *Proc. IEEE*, 63 (12):1692–1716, December 1975.

Working Group on California Earthquake Probabilities. Probabilities of large earthquakes in the San Francisco Bay region, California. *U.S. Geol. Surv. Circ.*, 1053:1–51, 1990.

G. Wornell. *Signal Processing with Fractals: A Wavelet-Based Approach.* Prentice Hall, 1996.

Kazuo Yanagihara. Magnetic field disturbance produced by electric railway. *Geophysical Magazine*, 38(1):17–35, 1977.

J.G. Yu and C.J. Goodman. Stray current design parameters for DC railways. In *Proc. of the 1992 IEEE/ASME Joint Railroad Conference*, pages 19–28, 1992.