5-4-2016

Applications of Advanced Signal Processing Techniques in Pure and Applied Geophysics

Robert W. Mehl
University of Connecticut - Storrs, robert.mehl@uconn.edu

Applications of Advanced Signal Processing Techniques in Pure and Applied Geophysics

Robert Mehl

B.S., University of Connecticut, 1998

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science
at the
University of Connecticut
2016

APPROVAL PAGE

Master of Science Thesis

Applications of Advanced Signal Processing Techniques in Pure and Applied Geophysics

Presented by

Robert Mehl, B.S.

Major Advisor_____
Lanbo Liu

Associate Advisor_____
Vernon Cormier

Associate Advisor_____
Lisa Park Boush

University of Connecticut

2016

Acknowledgments

I would like to express my deep appreciation and gratitude to my advisor, Dr. Lanbo Liu, for the patient guidance and mentoring he provided me over the many evolutions of the thesis topic and the many years it took to complete this degree. Dr. Liu's intellect can only be eclipsed by his compassion.

I would also like to think my advisory committee of Dr. Vernon Cormier and Dr. Lisa Park Boush. Dr. Cormier has (perhaps unknowingly) provided thought provoking ideas and new avenues of attacking problems. Dr. Cormier has been very generous over the years with his time and energy. I would like to thank Dr. Park Boush for the fierce review of my these and holding me to the academic standards of the department. I greatly appreciate her editorial review and comments.

I would like to acknowledge Dr. Stephen Billings, Dr. Lenard Pasion and Dr. Doug Oldenberg of the University of British Columbia for all that they have done to help me grow as a scientist.

I would like to thank my fellow students: Jason Keener, Tra Pham, Sarah Morton, Eric Rohrbach, and Rory Henderson. I would like to thank Seth Travis for thesis edits and being an unbiased sounded (and venting) board.

Lastly, I would like to thank my wife, Jennifer, who provided me with the necessary kick to finish my degree. It is with her love, faith in me, and brains that I owe the completion of this thesis too.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Data analysis and signal processing provide the tools that scientists can use to investigate and understand the physical world. However, these tools contain assumptions and/or limitations that can make the interpretation of the data difficult or lead to false conclusions. Currently, the most widely used analysis tool, the Fourier Transform (FT), decomposes continuous and discrete signals into their rudimentary frequency structure. Fourier-based analysis tools are inadequate to accurately describe physical signals. The Hilbert-Huang Transform (HHT) is a newcomer signal processing tool and this thesis applies this transform to different geophysical data.

The Fourier Transform uses linear combinations of simple harmonic functions as the basis to transform signals between the frequency and time or space domains [*Bracewell*, 2000]. Nevertheless, the fundamental theoretical requirement of a signal or series to be decomposed using Fourier polynomials is that the signal must be linear and stationary [*Huang et al*., 1998]. Very few real-world data are stationary and linear. A stationary signal is one that the mean and variance do not change over time. A linear signal follows the principle of superposition and scaling. While many signals can be accurately described by using the summation of linear functions, many others simply cannot.

Although the FT and its inverse are powerful tools to present the data in time and frequency domains, it falls short when accounting for both temporal and frequency characteristics simultaneously. It is important to understand that when a frequency change relative to time or space occurs, linear functions generate harmonic frequencies to account for these transient signals [*Bracewell*, 2000]. Development of the short time Fourier and Wavelet Transforms offer some improvements in time-frequency analysis

[*Huang et al*., 1998], but these techniques still use the linear combinations of linear functions to transform signals from one domain to the other. These two techniques still require the signal to be linear and stationary.

To address the issue of processing nonlinear and non-stationary signals, Dr. Norden Huang developed the Hilbert-Huang Transform (HHT) [*Huang et al*., 1998]. This technique uses the Empirical Mode Decomposition (EMD) to process an original signal into a small and finite number of intrinsic mode functions (IMFs) as the basis where the instantaneous frequencies are extracted using a Hilbert Transform [*Huang et al*., 1998]. The HHT method is empirically adaptive, it can successfully and effectively process nonlinear and non-stationary signals to extract useful characteristics to assist the understanding of the physical nature of the signals [*Huang et al*., 1998].

To understand the nature of a signal, it is imperative to remove the noise and emphasize the signal. Unlike the Fourier Transform, implementation of EMD with the HHT signal processing technique allows for the decomposition and removal of certain undesirable constituents from the input signals. When the IMFs are combined, the exclusion of the noise improves the signal-to-noise ratio and eases the interpretation. For this thesis, signal processing and data analysis used this technique for both one-dimensional (1D) time series as well as two-dimensional (2D) space domain data.

As part of standard data processing procedures, 1D signal processing filters remove unwanted signals (noise) and constant bias (DC offset). These rudimentary 1D filtering techniques give satisfactory results for most cases. However, these filtering techniques start to introduce artificial peaks in the signal or generate unwanted power in the harmonics when noisy signals are transient and/or nonlinear. Therefore, filtering

techniques should retain the necessary time/frequency relationship information so that the filter applies the correct value at the correct time.

This thesis explores the signal processing of time and space domain signals where the signal and noise components are non-stationary and nonlinear in nature. After an introduction of the development of the signal processing techniques in Chapter 2, I focus on the application of the EMD/HHT technique in a number of geophysical data sets to extract the useful information. These individual studies will be presented in the chapters following Chapter 2.

Chapter 3 addresses the improvement of microtremor horizontal-to-vertical spectral ratio (H/V) analysis using the Hilbert-Huang transform and Hilbert-Huang Spectrum. Chapter 4 highlights improvements of the analysis of a series of magnetotelluric data from Japan by removing transient signals and unwanted noise from the data record for the Schumann Resonance. Chapter 5 processes image textures to demonstrate the technique using bi-dimensional empirical mode decomposition (BEMD). For the geophysical application, BEMD is used to process helicopter-borne total field magnetic data by removing low and high frequency noise from data images for improved analysis. Chapter 6 provides a summary and restatement of the entire thesis study with comments on future developments.

Chapter 2: Signal Processing Methods and Theory

In this chapter, I describe the most widely used signal processing transforms presently used in geophysical data analysis. I compare these techniques with the Hilbert Huang Transform and show how the HHT algorithm addresses the shortcomings of the Fourier based algorithms.

2.1 Fourier Transform and Analysis

Of the many mathematical discoveries and theoretical developments made over the last few centuries, the Fourier series and Fourier Transform (FT) are two of the most important because their use spans many physical, financial and pure mathematical sciences. The basis for the success of the Fourier series and FT is found in a simple yet elegant solution and its ability to describe many different systems. The Fourier series, named in honor of its founder French mathematician and physicist Joseph Fourier (1768 - 1830), was created as a solution to the partial differential equation describing heat flow [*Bracewell*, 2000]. The series is simply a linear combination of weighted trigonometric (basis) functions [*Bracewell*, 2000]. The Fourier series is defined by [*Bracewell*, 2000]

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}\left(a_n \cos(nx) + b_n \sin(nx)\right) \tag{2.1}$$

where the coefficients a_n and b_n are

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx \end{aligned} \tag{2.2}$$

Using Euler's formula, $e^{ix} = \cos x + i \sin x,$ the series equation reduces to

$$f(x) = \sum_{n=\infty}^{\infty} c_n e^{inx}dx \tag{2.3}$$

where $i = \sqrt{-1}$ and the coefficient c_n is

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}dx \qquad (2.4)$$

By this substitution, the terms of the Fourier series are a set of vectors mapped in a complex plane. Since the Fourier series is a linear combination of simple harmonic basis functions, each linearly independent function is a vector on the complex plane with its own amplitude and phase angle [*Bracewell*, 2000]. Harmonic frequencies are linear combinations of existing vectors mapped in the complex plane.

Fourier Transform

At present, the most widely used technique for calculating the frequency spectrum of a series of digital signals is the Fourier transform (FT). In a general sense, FT is a complex Fourier series [*Bracewell*, 2000]. The transform expresses digital and analog signals as a linear combination of elementary trigonometric functions. These functions have unique amplitudes and frequencies as defined by the terms of the series created [*Bracewell*, 2000]. Plotting these amplitudes and frequencies generate a frequency spectrum unique to the input signal. In the most simplistic description, FT converts a time or one-dimensional (1D) spatial (i.e., distance) domain signal to a series in a frequency or wave number domain.

For a continuous time signal, the FT can be expressed as [*Bracewell*, 2000]

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt \qquad (2.5)$$

and its inverse as

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}dt \qquad (2.6)$$

There are many advantages to working in the frequency domain over the time domain for many signal-processing operations. The mathematical process of convolution of two

signals becomes simple multiplication, and the design of digital filters is greatly simplified.

Discrete Fourier Transform

While the continuous Fourier Transform (CFT) operates on signals of infinite length and generates an infinite number of frequencies, infinite frequencies are not very practical. The Discrete Fourier transform (DFT) offers a way to process signals of finite length and sample size. The definition of DFT is [*Bracewell*, 2000]

$$X_k = \sum_{n=0}^{N-1} x_n e^{-2\pi i k n / N}, \text{ k} = [0, \ldots, \text{N-1}] \qquad (2.7)$$

Like the (CFT), the DFT maps time domain signals to the frequency domain [*Bracewell*, 2000]. In the frequency spectrum, the number of effective frequencies calculated is equal to one-half of the number of samplings in the time domain, and the highest frequency, known as the Nyquist frequency, is equal to the reciprocal of two times the sampling interval [*Bracewell*, 2000].

Implementation of the DFT algorithm requires a large number of CPU cycles to compute the sums of the transform and generate the frequency structure of a signal. For large data sets, solving for many terms (approximately 1000s), this method of computation was inefficient and required a number of CPU cycles that grows as a quadratic function [*Cooley and Tukey*, 1965]. The development of the Fast Fourier Transform (FFT) sped up the process and can solve for the terms without losing accuracy and yet use far fewer CPU cycles for large data sets [*Cooley and Tukey*, 1965].

2.1.1 Signal Processing

Signal processing commonly uses Fourier Transforms as the basis to analyze the frequency structure of signals. Whether the signal is audio, electromagnetic (for example

radar, light, X-rays), or seismic in nature, the Fourier Transform decomposes the wave into its component frequencies. Structural analysis uses Fourier Transforms to analyze the harmonic structure of buildings and bridges under different ground motion and wind conditions. While signal processing uses these analyses for noise reduction or signal amplification. When desired signal amplitudes are close to the noise level, it is helpful to reduce the amount of noise first through environmental controls (such as: removing sources of noise, steadying the sensor and providing clean power) and second through signal processing. It is always best to mitigate the noise at the source rather than processing it away with filters. However, some noise sources cannot be removed, as is the case with helicopter-borne magnetometry. The noise created by the main rotor is a source that cannot be removed since it is needed for flight. The removal of this periodic noise uses band-pass filters. This method works nicely when noise is periodic and continuous throughout the signal, however if the noise is under-sampled, transient (lasting only a short time) or quasi-periodic, then this method does not work as well.

2.1.2 Image Processing

The Fourier Transform works equally well in the space-domain as it does in the time-domain. Images or pictures can be thought of as three-dimensional surfaces with two dimensions representing the X and Y Cartesian plane and the third dimension representing the value at that location. Images can be 1-channel gray scale or 3-channel red, green, blue (RGB) matrices. Some common analyses performed from image processing are image filtering (blur, sharpen, noise reduction), texture analysis, edge detection, pattern recognition, and compression [*Linderhed*, 2004].

2.1.3 Limitation of the Fourier Transform

Despite the versatility of the FFT for signal and image processing, the transform has limitations. One limitation of the Fourier transform is that the transform defines the frequency spectrum with no regard to the time at which the frequency occurred [*Huang et al*., 1998]. Another limitation is that the signal must be linear and stationary [*Huang et al*., 1998]. These limitations are problematic when introducing a transient signal, a signal with a short life span, into an otherwise well behaved signal. In the case of a transient signal, the transform introduces more frequency terms into the series as it tries to resolve the change in frequency. Since the signal has finite energy, the transform smears the signal energy across all frequencies. In addition, the transform does not state when the frequency occurred in time. The transform works on linear data only because the series terms are linear functions, however a clear picture of how the frequency changes in time is not resolved. Another technique has been developed to address these limitations by changing the length of time the transform works on the data.

2.1.4 Short Time Fourier Transform

The development of the Short Time Fourier transform (STFT) attempts to locate when in time or where in space a transient signal occurred. STFT computes the Fourier transform of the signal by moving a windowing function along the signal [*Huang et al*., 1998]. The result is a 3-D plot of time, frequency and amplitude. This method works well for non-stationary signal; however, the data inside the window must be stationary for this technique to work efficiently. If the transient signal has a small time window, the STFT window can be made smaller to locate smaller transient signals but if the signal has a large bandwidth resolution it is lost at the low frequencies. Using a narrow window gives

good time resolution but poor frequency resolution whereas using a wide window gives good frequency resolution but poor time resolution [*Huang et al*., 1998]. This is an unfavorable consequence of the STFT method. The determination of the window size is an *a priori* decision and rooted in Fourier theory, therefore this method is non-adaptable and is insufficient with nonlinear signals [*Huang et al*., 1998]. There needs to be a method that can address these realistic nonlinear and non-stationary signals.

2.2 Wavelet Transform

In the 1980s, the Wavelet Transform was developed and found traction in addressing the limitations of the Fourier Transform. The Wavelet Transform uses a signal multiplied with a wave function to address non-stationary signals and solve the problem of time/frequency resolution. The wave function is a small wave or mother wavelet. The function is considered a wave because it is oscillatory in nature and is made of sine and cosine functions [*Linderhed,* 2004]. The Wavelet Transform has a continuous and a discrete transform.

2.2.1 Continuous Wavelet Transform

The continuous wavelet transform is defined as [*Linderhed,* 2004],

$$W(s,\tau) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-\tau}{s} \right) dt, \tag{2.8}$$

where x is the original signal, ψ is the wavelet basis function or mother wavelet, s is the scale or size of the basis function, τ is the time translation from the basis function, and the factor normalizes the energy for the given scale.

The STFT has uniform resolution across all frequencies and times, and the WT addresses this problem [*Huang et al*., 1998]. However, the WT has its own limitations in the

resolution of frequencies and when in time those signals occur. The Heisenberg Uncertainty Principle for signal processing states that the frequency and the time cannot be known simultaneously [*Huang et al*., 1998]. Unlike the STFT which has a constant resolution for all times and frequencies, the WT has a good time and poor frequency resolution at high frequencies, and good frequency and poor time resolution at low frequencies. This maintains the limitation of Heisenberg and gives a better analysis when compared to the STFT [*Huang et al*., 1998].

The Wavelet transform operates on the original signal by convolving the wavelet with the original signal [*Linderhed*, 2004]. The result is a new signal that resembles a non-normalized correlation coefficient. This correlation coefficient will have high values when the frequency of the wavelet and the original signal are similar [*Linderhed*, 2004]. The continuous Wavelet transform is an iterative process with the original wavelet compressing or dilating to adjust the wavelets frequency. As the scale of the wavelet decreases (compression), higher frequencies in the data correlate with the wavelet. The opposite holds for the dilating wavelet.

The Wavelet transform only resolves either the time or the frequency, but not both. As the wavelet dilates, the frequency of the wavelet becomes better known because more of the wavelet data is contained in the oscillations of the wavelet [*Linderhed*, 2004]. However, when convolving the wavelet with the original signal, the wavelet operates over longer periods. This is the resolution problem with the wavelets, when we know more about the frequency; we know less about the time the frequency occurred. When compressing the wavelet, the frequency of the wavelet becomes lesser known because the number of points representing the energy of the wavelet decreases. The Nyquist

frequency limits the compression of the wavelet [*Linderhed*, 2004]. The energy of the wavelet is concentrated in a short time, so the time at which the frequency occurs is known.

The resolution capability of the wavelet transform can be thought of as a box with fixed area. In the Wavelet Transform, the mother wavelet has a predetermined time-scale resolution that is represented by a box with area an, *a*, on the time-scale plot [*Linderhed*, 2004]. As the scale increases (low frequency), the frequency becomes better known which is represented by a shrinking scale dimension. Conversely, the time dimension gets longer because of the amount of uncertainty in when the frequency happened. As much as the scale dimension shrinks, the time dimension increases that same amount which creates a box with different dimensions but same area [*Linderhed*, 2004].

The Wavelet transform provides advancements in signal processing/analysis and information about when certain frequencies occur, however certain limitations do exist. The first of these limitations is in the selection of the mother wavelet. Each mother wavelet has specific properties and how this wavelet interacts with the signal changes the time-frequency-energy spectra [*Huang et al*., 1998]. Different parts of a signal can have different properties and be governed by different equations; since the mother wavelet cannot change during the analysis, the wavelet transform is non-adaptive [*Huang et al*., 1998]. The best current practice is to perform the wavelet analysis using different mother wavelets then draw conclusions from all the results. This practice may be insightful but would be time consuming.

Another limitation is that the wavelet transform does not detect very short period transient signals. The Morlet wavelet is the most commonly used and this wavelet is

Fourier based, therefore it suffers from the many shortcoming of Fourier spectral analysis [*Huang et al*., 1998]. For instance, the Morlet wavelet has limitations when working on frequency modulated signals. When frequency changes are rapid and short lived, for example: a chirp, the wavelet cannot detect the change in frequency. The wavelet only works effectively when the frequency changes are gradual, for example: the analysis of sound from an oncoming train [*Huang et al*., 1998]. The Morlet wavelet has a length of 5.5 waves, that makes it unable to resolve intra-wave frequency modulations.

2.2.2 Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) computes the Wavelet transform using a filter bank of low- and high-pass filters. The filtering operation and the scale determine the resolution of the signal by up-sampling and down-sampling operations. This process is called the Mallat algorithm or Mallat tree decomposition (Figure 2.1). The decomposition uses a number of filters associated with different scaling functions to produce a course approximation of the Wavelet transform. Each of the iterations through the algorithm produces two signals, one from the low-pass filter and the other from the high-pass filter with the filter down sampling the signal by two [*Linderhed*, 2004].

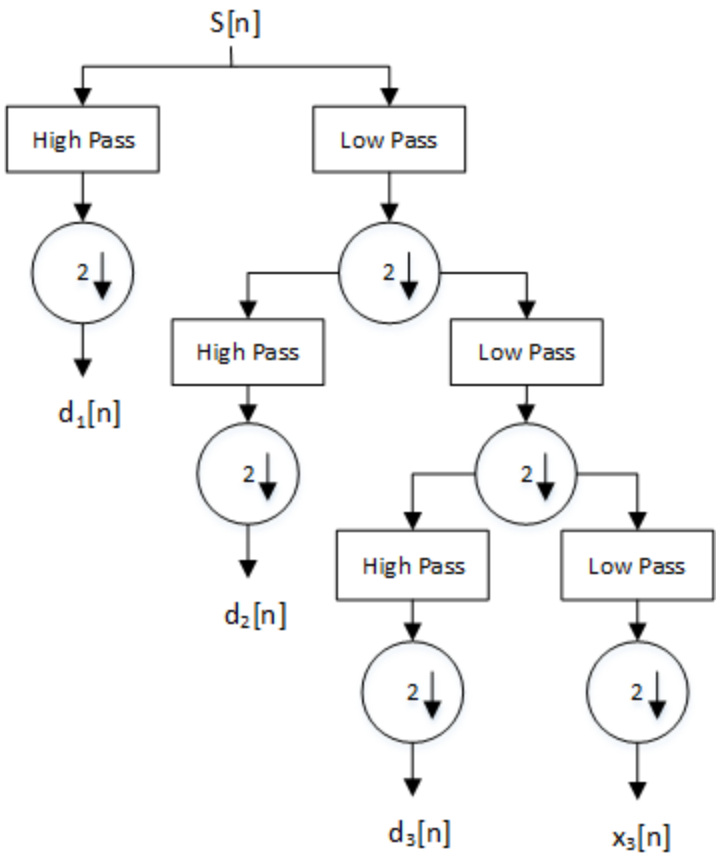

Figure 2.1: The Mallet Decomposition Tree outlining the discrete wavelet transform steps. Original signal X[n] processed with a high and low pass filter, the outputs are down-sampled by 2, the output from the high pass side is done, while the low pass side is processed again with the same method.

This process doubles the frequency resolution but reduces the frequency uncertainty by half. In accordance with Nyquist's rule if the original signal has a highest frequency of ω, which requires a sampling frequency of 2ω radians, then it now has a highest frequency of ω/2 radians. It can now be sampled at a frequency of ω radians thus discarding half the samples with no loss of information. Linderhed [2004] found that, "This decimation by two halves the time resolution as the entire signal is now represented by only half the number of samples. Thus, while the half band low pass filtering removes half of the frequencies and thus halves the resolution, the decimation by 2 doubles the scale."

2.3 Empirical Mode Decomposition

The newest processing technique is Empirical Mode Decomposition (EMD). EMD is an adaptive signal processing method that decomposes a signal into a finite number of intrinsic mode functions (IMF) and a residual function originally proposed by N. E. Huang [*Huang et al*., 1998]. Unlike the Fourier and Wavelet Transforms, the adaptive nature of this method allows the EMD algorithm to process nonlinear and non-stationary signals giving a more intuitive interpretation of the original signal. EMD is able to describe the short time changes in frequency the Fourier analysis is unable to resolve. EMD uses a sifting process to create IMFs that have well behaved Hilbert transforms from which the instantaneous frequencies are easily calculated [*Huang et al*., 1998]. Each IMF describes the local signal characteristics and not the entire signal. The summation of the IMFs and residue equal the original signal without loss or distortion [*Huang et al*., 1998].

Like the Wavelet transform, EMD results in an energy-time-frequency distribution or Hilbert spectrum. However, EMD has a few key differences from the Wavelet transform which improves analysis [*Huang et al*., 1998] and they are:

1. Time-frequency resolution is not constant,
2. EMD works on the scale of the distance between the signals extrema and not on the scale of the wavelet,
3. EMD does not require a priori information,
4. IMF is not represented as a series of trigonometric basis functions.

This empirical decomposition process works on the signal locally and decomposes the signal into basis functions (or IMF) based on the signal. The Heisenberg Uncertainty

Principle limits the signal resolution of STFT and WT [*Huang et al*., 1998]. However, unlike Fourier based methods, the Hilbert Transform of the IMF calculates the instantaneous frequencies, so the EMD method overcomes the time-frequency resolution posed by Heisenberg for signal processing [*Huang et al*., 1998].

The EMD signal processing method was developed to analyze both non-stationary and nonlinear signals. In nature, many measured signals are nonlinear and non-stationary but are approximated using linear and stationary signals [*Huang et al*., 1998]. Fourier analysis was used to study these signals but this analysis can provide misleading results because Fourier analysis assumes the input signals are both linear and stationary [*Huang et al*., 1998]. Many natural systems can be approximated with linear systems; however, the measuring of those systems can yield nonlinear results due to sampling rate, instrument response and/or instrument dynamic range [*Huang et al*., 1998]. The same conclusion and measurements argument can be made for the stationary case.

Intrinsic Mode Functions

An intrinsic mode function (IMF) is a simple harmonic function where the number of extrema and zero crossings must either be equal or differ by only one and the mean value of the envelope defined by the local maxima and minima is zero [*Huang et al*., 1998]. The second property (mean value) is necessary to prevent unwanted fluctuations in the computation of the instantaneous frequency cause by asymmetric waveforms. By this definition, the IMF has a well-behaved Hilbert Transform.

Sifting is the name of process used by EMD to calculate the IMF [*Huang et al*., 1998]. The sifting process iteratively decomposes the signal by extracting the highest oscillatory

mode from original signal. The sifting iterations continue until two criteria are met. The criteria are [*Huang et al*., 1998]:

1. The standard deviation, SD, between consecutive sifting results is no larger than a predetermined value.

$$SD = \sum_{t=0}^{T} \left[\frac{\left| \left(h_{1(k-1)}(t) - h_{1k}(t) \right) \right|^2}{h_{1(k-1)}^2(t)} \right], \text{ where k = the sifting iteration;} \qquad (2.9)$$

2. The number of zero-crossings and extrema are equal or at most differing by one.

Once both criteria are satisfied, the resulting signal is classified as an IMF and subtracted from the pre-sifted signal. If the resulting signal is a monotonic function, then the signal is classified as the residue and the sifting process is complete [*Huang et al*., 1998]. As an exception to the rule, the number of IMF extracted can be predetermined where the resulting residue may not be a monotonic function. For a signal to be sifted it must have at least two extrema (1 minimum and 1 maximum) if one does not exist the signal can be differentiated one or more times to reveal the extrema [*Huang et al*., 1998]. As stated in Huang [1998], the sifting algorithm is as follows:

1. Find the positions and amplitudes of local minima and maxima.
2. Create upper envelope from maxima and lower envelope from minima using spline interpolation.
3. Calculate the mean for each time using the upper and lower envelope.
4. Subtract the mean from the input signal.
5. Check the difference signal to see if it meets the criteria for IMF, if not, process this signal from step 1. If so, move to step 6.

6. Find the next IMF by using the residue as the input signal and iterating through again.

In step 2 of the process, the upper and lower envelopes are created using an interpolation method. This step can cause many issues with the sifting process if the interpolation fails to conform to the maxima and minima properly [*Huang and Wu*., 2008]. The interpolation may generate over- and under-shoots that can create artifacts in the resulting iterations and subsequent IMF. Linear interpolation will surely prevent these over- and under-shoots but the resulting envelopes are not smooth. Piecewise cubic spline interpolation creates a smooth envelop but over- and under-shoots can still occur. Piecewise cubic Hermite interpolation creates a smooth envelop, though less smooth than the spline, and does not over- or under-shoot the extrema values. The use of different types of splines for interpolation are one of the areas that need to be further explored by researchers. The use of the Hermite spline was selected as the spline method used for envelop creation and has not been fully vetted as the best interpolation algorithm.

Instantaneous Frequency

The instantaneous frequency is calculated from the IMF by taking the derivative of the instantaneous phase which is an output of the Hilbert Transform. The instantaneous frequency does not need a full oscillation to calculate the frequency, but can identify the partially oscillating non-stationary signals contained within the IMF [*Huang et al*., 1998]. Unlike Fourier analysis, which uses harmonics to decompose the wave to simulate the signal, no spurious harmonics are needed to correctly represent the nonlinear and non-stationary aspects of the signal. The Hilbert Transform of the IMF creates an analytical

signal that greatly simplifies the computation of the instantaneous frequency [*Huang et al*., 1998]. For any time series, X(t), the Hilbert transform, Y(t), is [*Huang et al*., 1998]

$$Y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{X(t')}{t - t'} dt', \tag{2.10}$$

where P is the Cauchy principal value. X(t) and Y(t) form a complex conjugate pair, which yields an analytic function, Z(t), as

$$Z(t) = X(t) + iY(t) = a(t)e^{i\theta(t)}. \tag{2.11}$$

The above equation can be transformed into its polar form,

$$a(t) = \sqrt{X^2(t) + Y^2(t)}, \quad \theta(t) = \arctan\left(\frac{Y(t)}{X(t)}\right). \tag{2.12}$$

Hilbert transforms define the imaginary part in such a way as to allow the result to be an analytic function. The instantaneous frequency is calculated by taking the derivative of the phase of the analytic function, θ(t) [*Huang et al*., 1998]. The instantaneous frequency can have a meaningless result if the function has direct current component greater than one. To overcome this, the function is restricted in order to be symmetric locally with respect to the zero mean level. It is important to define a process to decompose the data into a function or set of functions that have a zero-mean property because these functions will have well-behaved Hilbert Transforms. These functions are the intrinsic mode functions [*Huang et al*., 1998].

Looking back at the Fourier based methods, the uncertainty principle made it impossible to know the frequency at the time it occurred and the best that could be done was to shrink the window over which the Fourier transform operated. When the instantaneous frequency is calculated, we know the frequency of the signal at one time while the

frequency for the rest of the signal is unknown. Each IMF contains only one frequency at any point, which is how the instantaneous frequency is calculated.

The Hilbert Spectrum

The results of performing FT or WT analysis is a frequency-amplitude or frequency-time-amplitude spectrum, respectively; similarly the EMD analysis results in a frequency-time-amplitude spectrum. The difference between the EMD and WT spectra is the EMD spectrum has higher frequency-time resolution. The increase in frequency resolution is due to the Hilbert transform that facilitated the computation of the instantaneous frequency. After the Hilbert transform, the IMF components can be expressed in the following form [*Huang et al*., 1998]:

$$X(t) = \sum_{j=1}^{n} a_j(t) exp\left(i \int \omega_j(t) dt\right). \qquad (2.13)$$

It is important to mention that the residue function is not used in the computation for the spectrum, because the monotonic function that the residue represents provides no useful information and tends to overpower the spectrum [*Huang et al*., 1998].

The above equation shows that both the amplitude and frequency for each IMF are a function of time in the three-dimensional plot, in which the amplitude can be contoured as a function of time and frequency [*Huang et al*., 1998]. This plot is called the Hilbert Spectrum. If needed, an energy distribution plot can be created by squaring the amplitude. These plots are similar to the amplitude or energy distribution plots seen with WT analysis, except the frequency and time resolution is greater with the Hilbert spectrum.

2.4 Bi-dimensional EMD

Building upon the successes of the one-dimensional EMD, the technique eventually expanded to the two-dimensional case for image processing and analysis. Texture analysis, image compression and spatial filtering, to name a few, greatly benefit from the adaptive nature of the bi-dimensional EMD (BEMD). BEMD generates a series of IMF surfaces with the first IMF containing the highest wave numbers present in the data and each subsequent IMF containing lower and lower wave numbers. This makes the BEMD an ideal tool for image processing especially when the desired output is a separation of wavelengths.

Since first developed, the two-dimensional EMD algorithm has seen development and refinement by several people. The first development was by Huang [2001] when the algorithm processed a two-dimensional image as a series of one-dimensional slices. This method, however, did not process the image as a continuous surface so discontinuities appeared between the slices. In 2003, a further development came in an approach which processed the surface as a whole. The approach sifted the data using morphological operators to detect the regional maxima and radial basis functions for surface interpolation [*Nunes et al*., 2003a]. Linderhed [2004] also developed BEMD method for image compression that used the spline for surface interpolation. Damerval et al. [2005] computed the upper and lower envelopes based on Delaunay triangulation and piecewise cubic polynomial interpolation.

Furthermore, Xu et al. [2006] developed the method used for this thesis and uses finite-elements and characteristic points to compute the mean surface. Characteristic points are defined as the local maxima, local minima and saddle points. Saddle points are

simultaneously local maxima and minima points and therefore should be included in the determination of the mean surface. The corner points are included in the set of characteristic points because they define a closed polygonal domain for the finite-element analysis and this ensures the Delaunay triangular mesh covers the whole domain. This method represents the local mean surface of the data as a linear combination of two-dimensional linear basis functions smoothed with a bi-cubic spline. The coefficients for the basis functions are computed from the local extrema using a general low-pass filter [*Xu et al*., 2006].

The algorithm used for by Xu et al. [2006] for BEMD is very similar to that used in the one-dimensional EMD. With a two-dimensional dataset, *f*, the decomposition algorithm is as follows:

1. Find the local extrema and saddle points of *f*.
2. Form a triangular mesh using the Delaunay method.
3. Smooth the characteristic points set using a generalized low-pass filter.
4. Generate the local mean surface, *m*.
5. Compute $h = f - m$.
6. Determine if the stopping criterion has been met. If the stopping criterion is met, stop the iteration, otherwise treat *h* as the data and iterate on *h* through steps 1 – 6.

There are subtle differences between all the methods mentioned above but the basic algorithm holds true. Like EMD, the summation of the BEMD IMFs will nearly equal the original surface. However, different interpolation methods introduced different errors and the fitting is only an approximation of the surface since the spline could not go through all the actual data points [*Huang and Wu*, 2008].

The definition of the two-dimensional IMF is the same as for the one-dimensional, that is, zero-mean, maxima are positive values, minima are negative values, and the number of maxima and minima points is equal. However, in the algorithm used for this thesis, there is a large difference in how the IMFs are computed. Damerval et al. [2005] calculate the local mean surface from the mean of the upper and lower envelopes. The methods employed by Xu et al. [2006] calculate the local mean surface directly from the characteristic points. This method prevents the possible over-shootings between the upper and lower envelops when fitting the surfaces with a spline function therefore reducing the error associated with splines.

Chapter 3: Application of Empirical Mode Decomposition to Microtremor Data

3.1 Microtremor – HHT Technique

Data background

Microtremor are seismic data produced by local noise and collected using a wide-band seismometer. Typically, seismic data is most useful when there is minimum noise because noise prohibits studying deep Earth structures, however, microtremor studies use noise as the source. Transient noise is one type of noise that contaminates microtremor data and requires removal. Placing a seismometer at a single location and collecting data for hours describes the microtremor data collection process. Recording and analyzing the coherent noise produced by natural and anthropomorphic sources provides the data for understanding the subsurface as applied to engineering such as structural design. These data provide valuable information regarding the local structure beneath and around survey sites. This particular microtremor data collection was part of a larger project to analyze the sediment thickness of Beijing, China's central city.

The city of Beijing sits on the alluvial fan with mountains circling from the west to north and onward to the east. Hundreds of meters of sediment cover bedrock where a series of faults underlie the city and surrounding communities, resulting in an earthquake hazard [*Liu et al*., 2015]. In fact, an historic magnitude 8 earthquake occurred just 70 km east of the city center (the 1679 Sanhe-Pinggu earthquake) [*Liu et al*., 2015]. When sediments exist over bedrock with seismically active faults, there often exists a concern about how structures react to ground motion. For example, a magnitude 8.1 earthquake rattled Mexico City on September 15, 1985 and the city suffered major damage [*Liu et al.*,

2015]. This damage resulted from seismic energy resonating in the sediment-filled valley underneath Mexico City. The sediment layer of the valley had a natural frequency of 2.5 Hertz and buildings with a height of 15 to 16 stories have similar resonance frequencies [*Liu et al.*, 2015]. As the seismic energy traveled through the bedrock and into the sediment, the waves began to resonate between the bedrock and the surface, buildings started to crumble and the soil started to liquefy. Beijing is in a similar situation; even though the sediment composition is different the sediment thickness focuses the seismic energy in much the same way [*Liu et al.*, 2015]. Microtremor data can be used to non-intrusively determine the sediment thickness and is more economical than using a drill rig.

To compute the sediment thickness from microtremor data, a common practice in microtremor data pre-processing is to exclude the transients by using an anti-triggering technique (which is opposite to the triggering approach of traditional seismic records) which uses the short-term average to long-term average (STA/LTA) ratio [*Liu et al*., 2015]. The commonly used values in seismology are 1 second for STA and 25 seconds for LTA and typically generates a ratio between 0.2 and 4.0 [*Liu et al*., 2015]. This is the microtremor data selection criterion used in GEOPSY, a microtremor analysis software package [*Bard et al*., 2004].

While GEOPSY removes data contaminated with transients and local monochromic noise, the enhanced approach accomplishes the same effect but removes signals by eliminating certain IMFs. Unlike the anti-triggering method using the STA/LTA ratio to excluded transient portions, the data after HHT enhancement remains continuous. This

greatly improves the signal analysis by eliminating edge effects of the time series because the data remains continuous.

3.2 Microtremor Data Analysis Enhancement

The enhancement of microtremor data using HHT follows these three processes. The processes are:

1. The elimination of the local transient signals through EMD analysis
2. Identification and elimination of continuous monochromatic noise
3. Application of the Hilbert-Huang spectrum to determine the H/V ratio

Processing the microtremor data collected from the Beijing area demonstrated the aforementioned processing steps [*Chen et al*., 2008].

Elimination of Local Transients

This processing step involves reducing the time series data into their IMFs using EMD then removing the IMF that contains the majority of the transient energy. Figure 3.1 shows the IMFs of 1 minute of East-West component microtremor data. The top row represents the original microtremor data with each successive row showing the IMFs in sequence from 1 to 5. In this particular example, IMF1 and IMF2 are the predominant contributors in the EW components. In general, IMF1 always possesses the highest frequency content and is most likely associated with transient noise. By definition of the IMF, the IMF order is from highest to lowest frequency.

To demonstrate that IMF1 is most likely associated with transient noise, the Fourier Transform was applied to the vertical component to analyze the frequency structure of the microtremor signal. The left column of Figure 3.2 shows the raw data, IMF1 and the

summation of the remaining IMFs for the vertical component from site OL-14. The right column shows the corresponding behavior of the left column in the frequency domain. In a comparison of the left column images to the original data, it is clear that IMF1 contains the dominant signals contained in the raw signal. The frequency analysis or the right column of plots, shows that IMF1 dominates the high frequency band above 10 Hz. Therefore, IMF1 originates locally and contains the transient noise. Usually, the high frequency constituents in the microtremor records are generated locally by transient noise sources that do not contain much useful geological or engineering information. Theoretical analysis has also indicated that noise within one wavelength of the site may be useless for microtremor analysis [*Arai and Tokimatsu*, 2004].

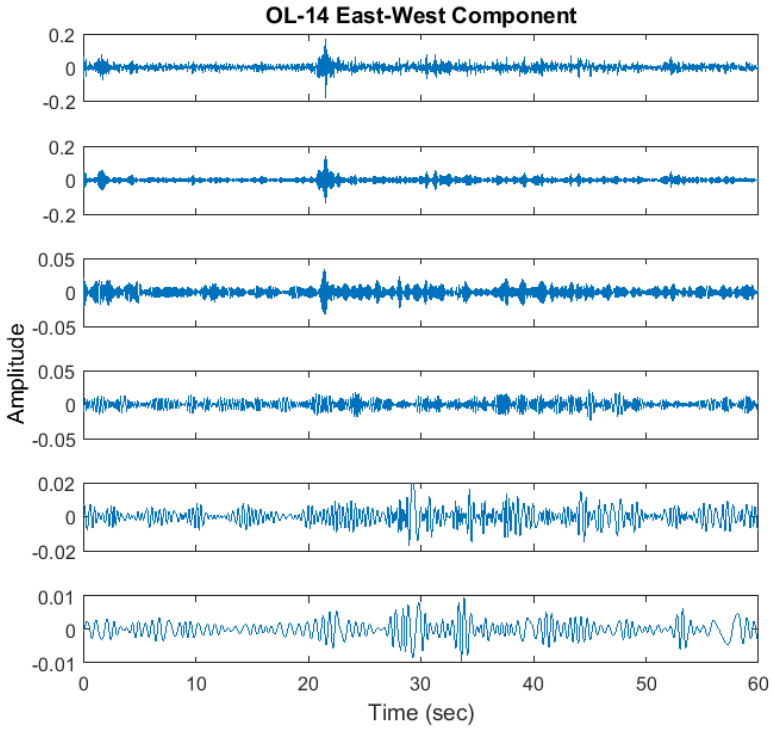

Figure 3.1: An example of the East-West component microtremor records and their IMFs from the 1st to the 5th mode after the empirical mode decomposition.

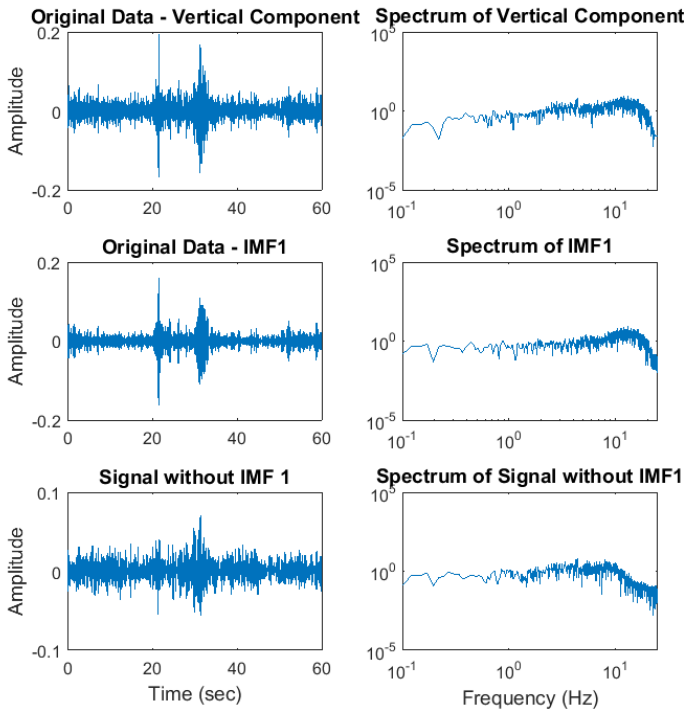

Figure 3.2: An example of the time domain record for the original vertical component, IMF1 and the signal with IMF1 removed (left column) and their corresponding amplitude spectra in frequency domain (right column).

The EMD filtering technique was applied to microtremor data collected at OL-14 (ϕ=40.006°, λ=116.450°, and h=40.5 m), a site near the Olympic Green and north of the center of Beijing. Since the site is located within the bustling city of Beijing, local transients heavily contaminate the signal, as shown in Figure 3.3. Using EMD and eliminating IMF1, the power of local transients is suppressed. Using signals without IMF1, a comparison between the STA/LTA ratio without IMF1 and the raw data shows a significant drop in the ratio at the time of the transients. Figure 3.4 shows the east-west component and this comparison with the overall reduction of the influence of the local transients on the STA/LTA ratio. For the first 30,000 data points, a reduction of the maximum STA/LTA ratio goes from 4.0 for the raw data to 2.94 after eliminating IMF1. Excluding IMF1 eliminated the seven peaks with the STA/LTA ratio exceeding 3.0. The

new data are continuous with transients removed or their influence to the ratio reduced. If using GEOPSY to process, the removal of the data creates a discontinuous dataset. Discontinuities are unfavorable for the remaining processing steps.

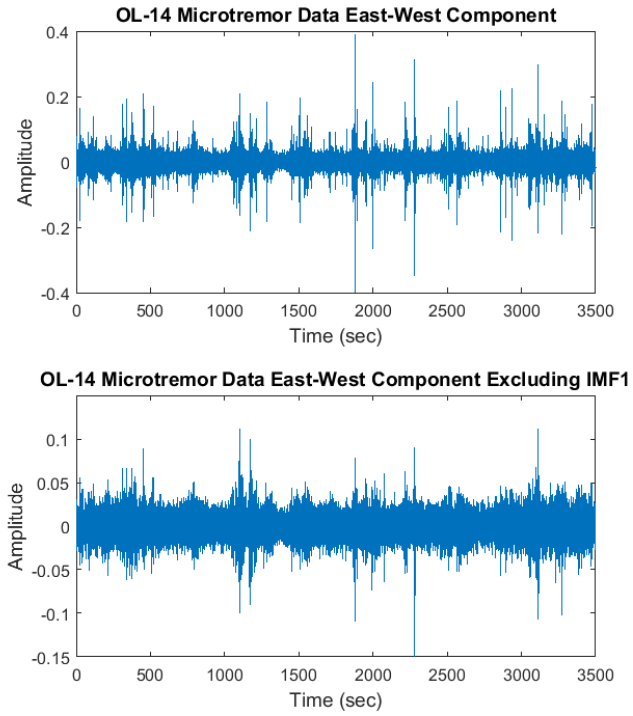

Figure 3.3: The original 3-component microtremor data recorded at site OL-14; and same data with IMF1 removed using the EMD analysis.

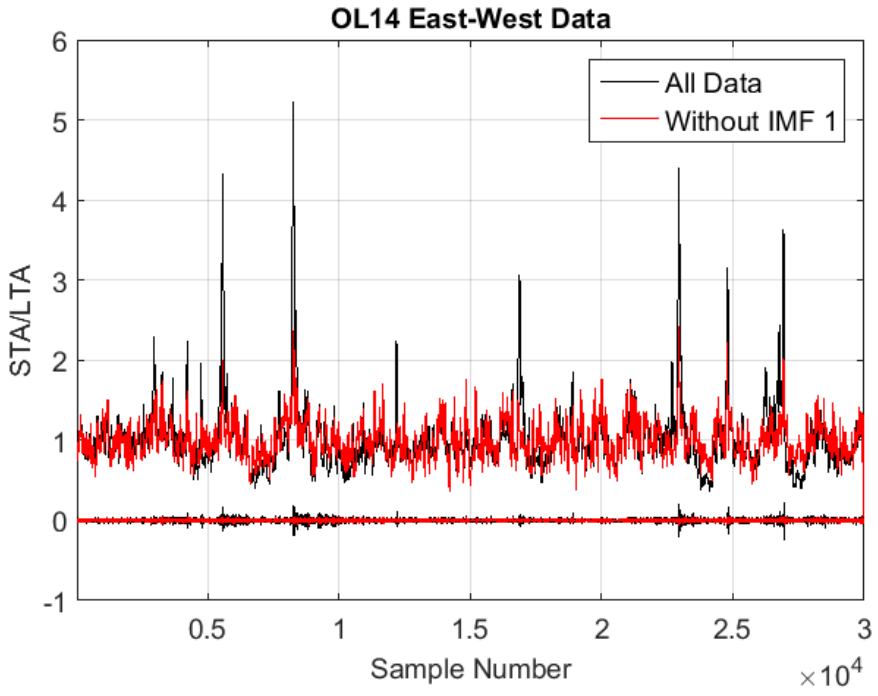

Figure 3.4: The upper lines show the STA/LTA ratio and the lower lines show the signal amplitude with all the data (black line) and with IMF1 removed (red line).

After removal of IMF1, the data processing continued as normal. As a comparison, the data was processed with the transients and the data with IMF1 removed. As shown in Figure 3.5, there is no clear H/V peak when using all of the data. In contrast, there is a clear peak, albeit broad, in the data with IMF1 removed. This is a clear improvement in the ability to analyze the data effectively.

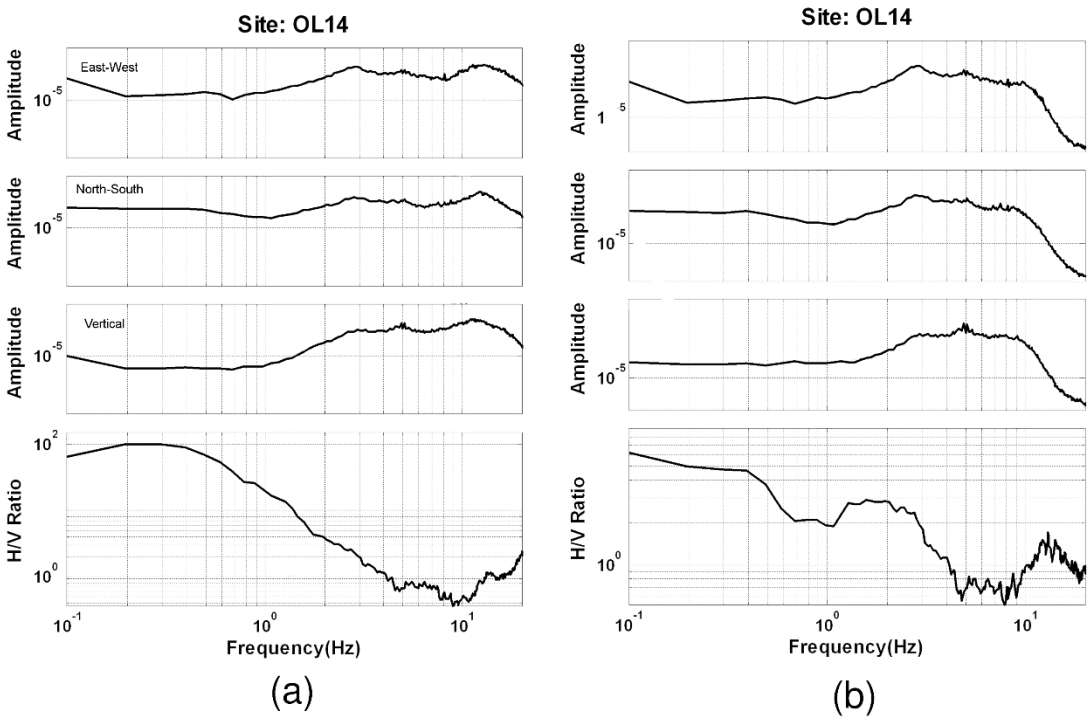

Figure 3.5: The Fourier spectra and H/V results from the original 3-component microtremor data recorded at site OL-14; and same analysis with IMF1 removed using the EMD analysis.

Elimination of Continuous Monochromatic Noise

Like Fourier based analysis, EMD can be used to remove continuous frequencies such as monochromatic noise. Monochromatic noise has a limited frequency band and can generate false peaks in the H/V ratio analysis especially when the noise appears in the horizontal or vertical components. To eliminate these false peaks, the data is processed into IMFs and the IMFs that are dominated by the noise are removed. Some of the survey data collected in the Beijing survey contained this type of noise. An analysis of the noise shows that it has a frequency around 1.5 Hz in the horizontal component. An industrial park under construction was the known source of the noise [*Liu et al*., 2015].

Data collected at NW282 had a 1.5 Hz resonance in the EW component and this noise would adversely affect any H/V analysis (Figure 3.6). When breaking the data down into

their IMFs, the third IMF contains most of the 1.5 Hz noise. Therefore, this IMF was excluded, the remaining IMFs were summed together and processed using the H/V analysis steps. The H/V analysis done on the original data shows the monochromatic noise and a resonance peak at 1.5 Hz. Figure 3.7 shows the H/V analysis on the original data on the left and the H/V analysis on the data with the third IMF removed has improved and a resonance frequency at 1.2 Hz.

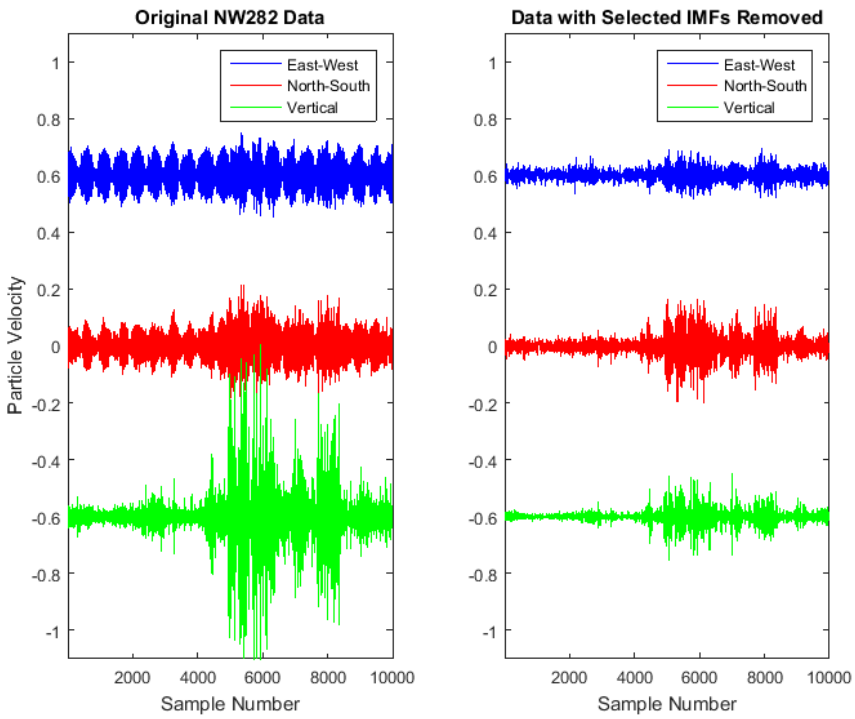

Figure 3.6: The original 3-component data at site NW282 and after the elimination of the 1.5Hz continuous wave (IMF3) in the horizontal components and IMF3 with additional transient in the vertical component.

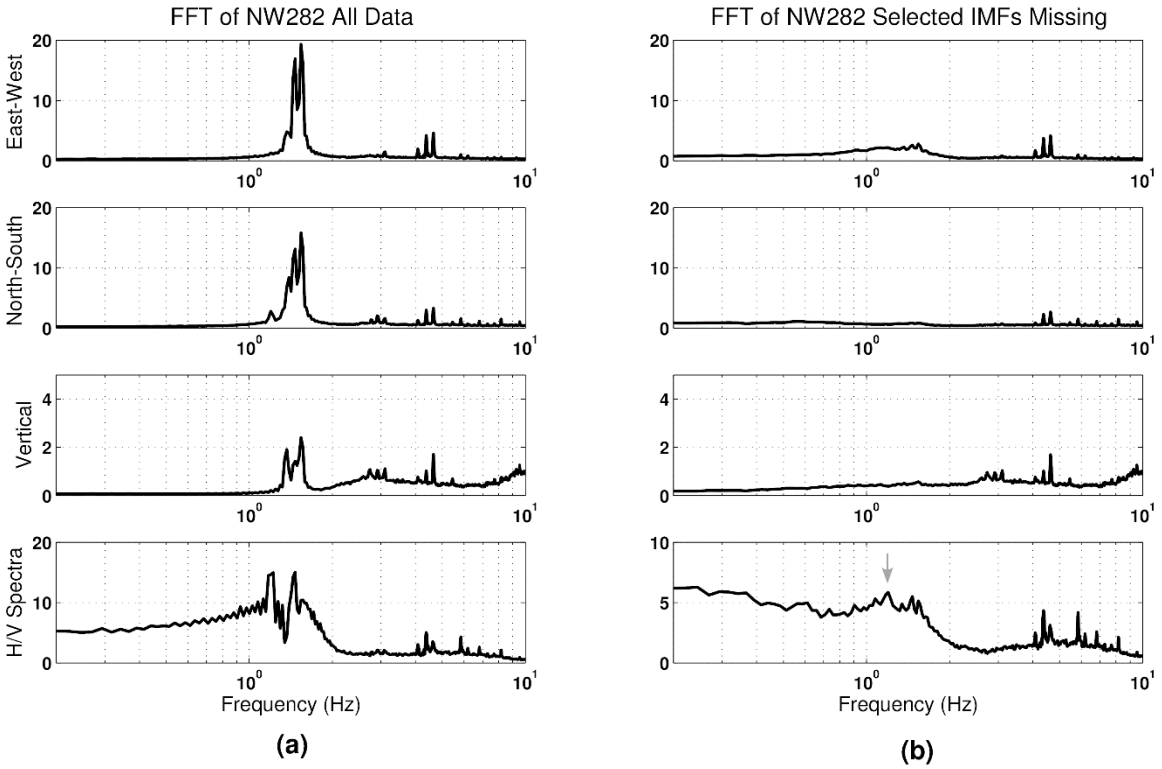

Figure 3.7: The left side (a) shows the H/V result from the FFT algorithm using the original data from the NW282 site. The right side (b) shows the same analysis for data with IMF3 excluded.

Improvement in the H/V Analysis

As was seen in an earlier analysis of data using H/V analysis, the frequency band was broad which leads to the conclusion that there is low frequency resolution. This is typical with frequency analysis using Fourier transform analysis. The goal is to use the Hilbert-Huang transform to increase the frequency resolution during H/V analysis and get sharper resonance peaks.

For the data collected at OL-14, the Hilbert-Huang Spectrum (HHS) was computed for each component. Figures 3.8 to 3.10 show HHS for the East-West, North-South and Vertical components for all the data and with IMF1 removed, respectively. The time series (the top plot), the time-frequency-energy distribution image (the middle plot) and the summation of the energy by frequency (the left plot)when combined form the HHS.

The plot on the left can be interpreted as the Fourier spectrum. The sampling rate of the data is 50 Hz, however, the figures show only the first 10 Hz. The data above the 10 Hz range is inconsequential to the analysis so the upper range of the axis was set to 10 Hz. A notable feature between the original data and the data with IMF1 removed is the decrease in energy above 6 Hz. The time-frequency-energy distribution plots shows the amount of energy in each frequency for each block of time. For these plots, the original data show no coherence. After removing IMF1, the data is no longer saturated with noise and the plots start to show coherence. This is most noticeable in the vertical component where lines start to appear in the image plot. These lines show the connection between the changes in frequency and energy as a function of time.

An objective of this processing is the removal of energy caused by transients. All of the original data plots show a large transient signal at about 1100 seconds. This transient has a lot of energy and is easily seen in the image as the values turn red as the frequency increases. Removing IMF1 removes nearly all the energy from this and other transients.

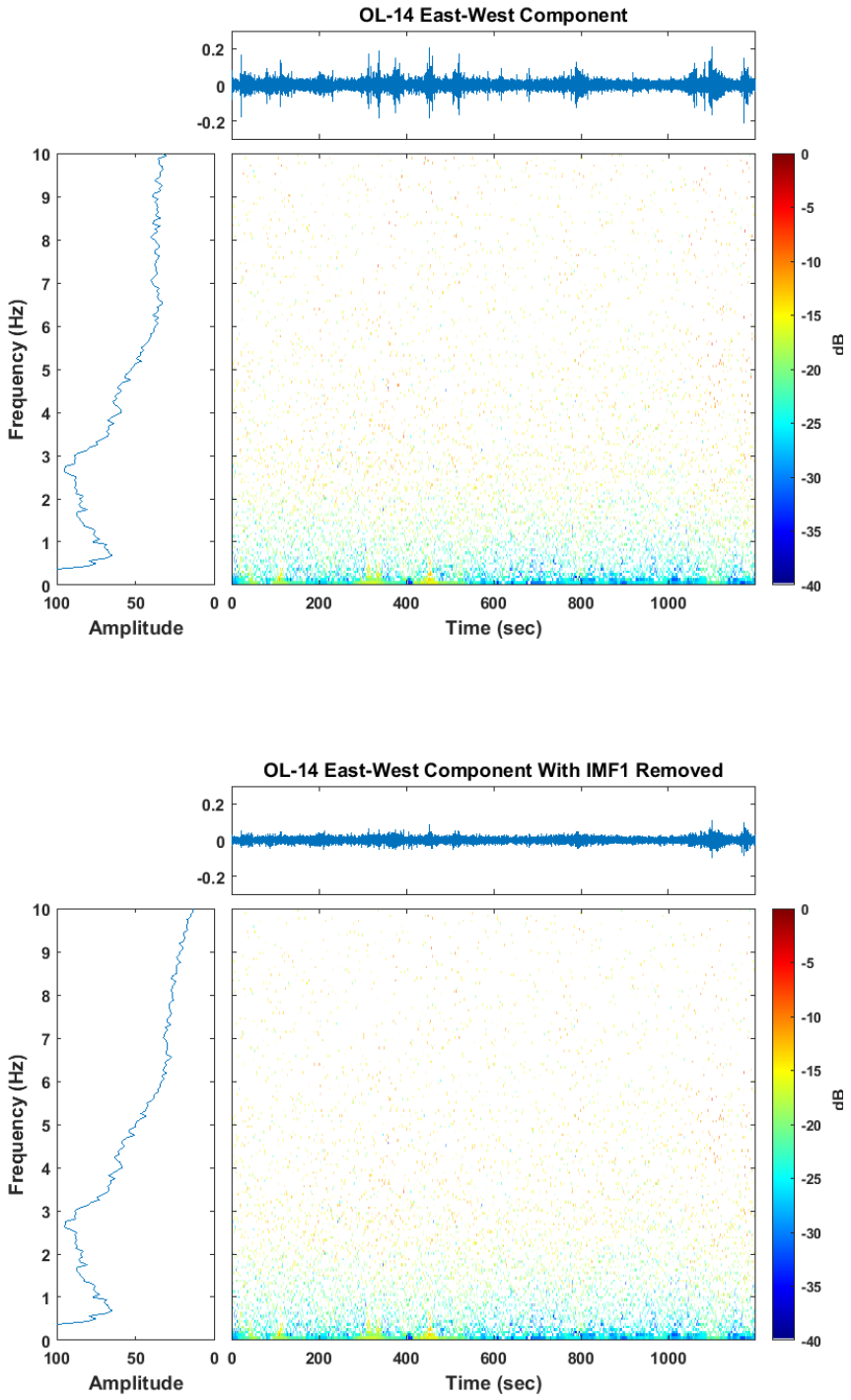

Figure 3.8: HHS of the original East-West component data at OL14 and the HHS of the same component with IMF1 removed.

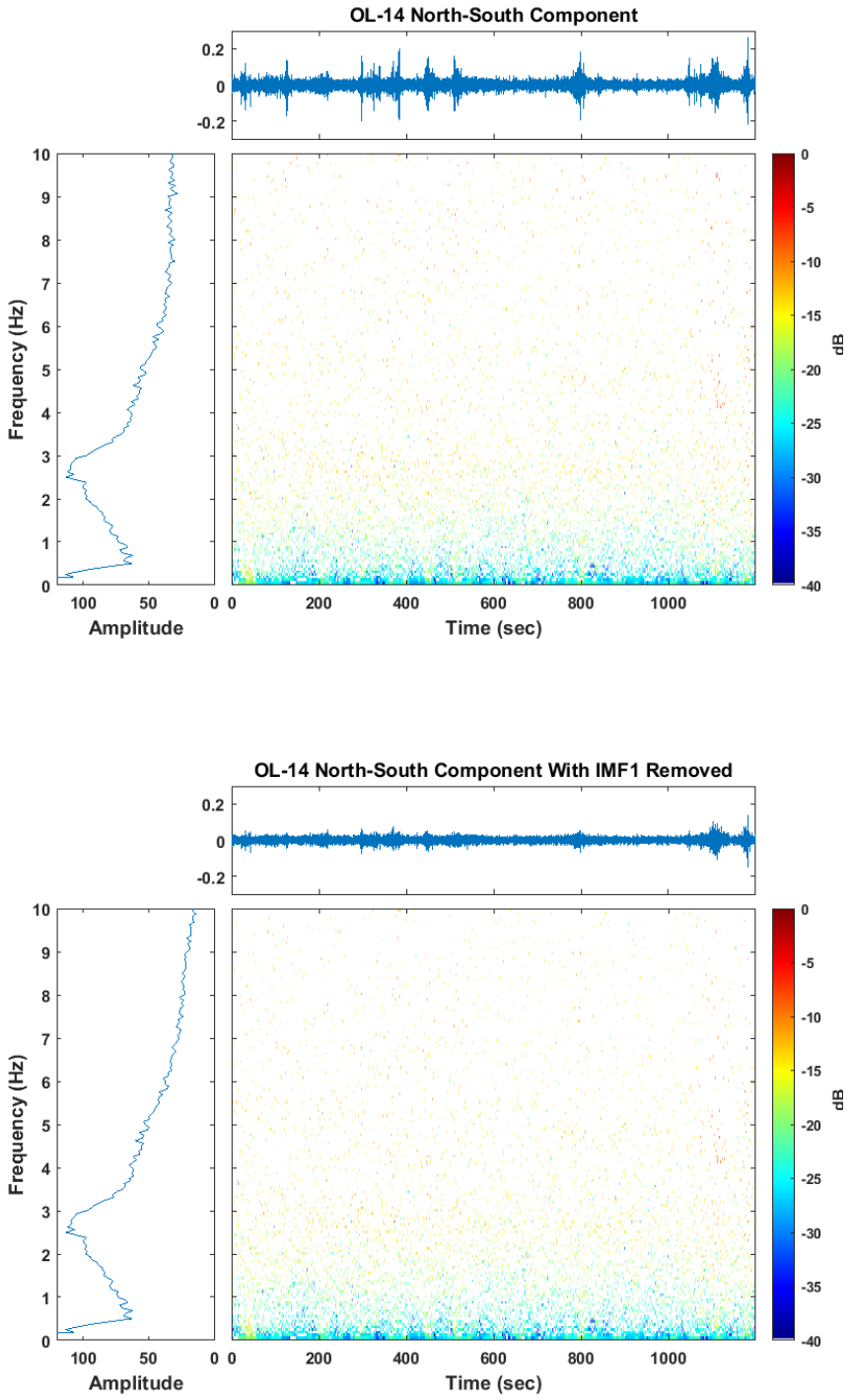

Figure 3.9: HHS of the original North-South component data at OL14 and the HHS of the same component with IMF1 removed.

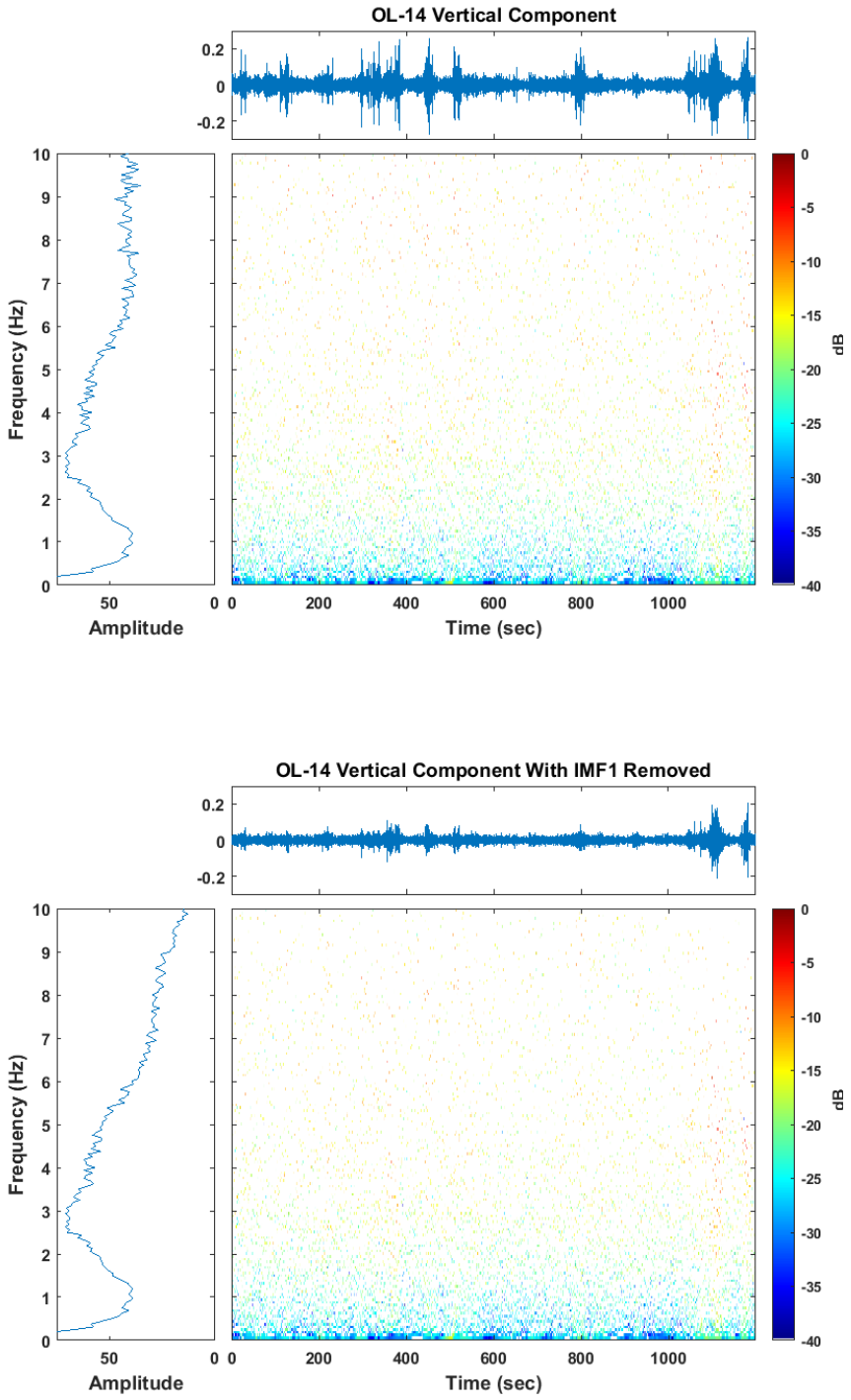

Figure 3.10: HHS of the original vertical component data at OL14 and the HHS of the same component with IMF1 removed.

Using the frequency data generated from the HHS analysis and processing it using the H/V analysis procedure shows a sharper peak resonance then the FFT analysis. Figure 3.11 shows this result. Unlike the FFT analysis, the HHT's instantaneous frequency calculation does not share energy with the higher frequencies. The sharper, more definitive peak improves the sediment thickness calculation.

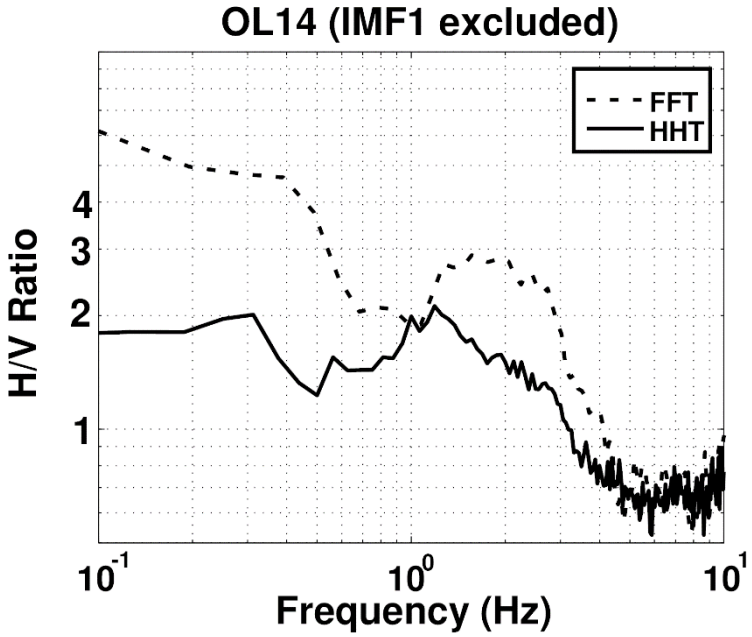

Figure 3.11: The H/V results from the FFT (dashed line) and the HHT (solid line) algorithms with the elimination of IMF1.

From the comparison of the H/V results shown in Figure 3.11, we can see a sharper peak for the HHT-derived H/V ratio than for the FFT-derived ratio. Consequently, it implies a higher resolution in determining the fundamental resonance frequency, which is critical to infer sediment thickness with a better constraint [*Parolai et al*., 2002]. A borehole drilled as a water well in this area (ϕ=40.0167°, λ=116.35°), 8 km away from site OL14, revealed the thickness of the Quaternary overburden to be 92 m [*Liu et al*., 1989]. The sediment conditions are similar to an area near Cologne, Germany, therefore the frequency-sediment thickness relationship used by Parolai, et al. [2002]:

$$h = 108\, f_r - 1.551 \qquad (2.14)$$

is suitable for correlation at the Beijing area site [*Chen et al*., 2008; *Wang et al*., 2008]. Using the HHT peak resonance of 1.2 Hz, the computed sediment thickness is 82 m. Using the FFT peak resonance of 1.6 Hz, the computed sediment thickness is 52 m. In general, the HHT result is more consistent with the depth to bedrock collected in a borehole 8 km away from OL14. In addition, sediment thickness inferred from microtremor H/V analysis at other sites with higher quality data are all in the same range [*Chen et al*. 2008].

Using EMD signal analysis techniques have greatly improved the results from microtremor data collected in the Beijing, China area. The source for microtremor data is ambient noise and by nature noise is nonlinear and non-stationary therefore Fourier based analysis would never full transform data in a truly meaningful way without adding spurious data.

Chapter 4: Application of Empirical Mode Decomposition to Continuous Magnetotelluric Data in Japan

Empirical Mode Decomposition can be used to remove unwanted signals contained in magnetotelluric data. Magnetotelluric (MT) fields are large scale magnetic and induced electric (telluric) fields that occur naturally. MT surveys are widely used in investigations of the Earth's structure at both regional scale for tectonic research and local scale for mineral and energy resource explorations. Incoming solar and cosmic radiation interact with gases in the upper atmosphere; this interaction releases electrons creating a layer of ions called the ionosphere [*Johnson*, 1965]. The variations in the electromagnetic (EM) field in the ionosphere induce eddy currents in the relatively conductive Earth's crust and mantle; these are magnetotelluric (MT) currents [*Telford et al*., 1990]. The ionosphere also acts as a reflector for some of the EM radiations. It traps the EM waves with a frequency lower than 178 MHz (in the very high frequency (VHF) range, i.e., 30 - 300 MHz) by reflecting them back to the Earth in the day time with a frequency of about 17.8 MHz (in the high frequency (HF) range) [*Johnson*, 1965]. This chapter addresses the processing of MT data to look for the natural radiation occurring at an extremely low frequency (ELF, 0 to 100 kHz).

4.1 The Ionosphere and the Schumann Resonance

Interactions between the upper atmospheric gases and incoming solar radiations creates the ionosphere. The ionosphere contains two main layers, the E-layer (90 to 150 km) and the F-layer (150 to 500 km) [*Johnson*, 1965]. The ionosphere changes between day and night with the creation of the D-layer (60 km) and the F-layer splitting into two distinct layers in the daytime [*Johnson*, 1965]. Changes in the flux of solar radiation causes the

diurnal variation in the ionosphere. The layers act as a barrier to ELF and HF EM radiation refracting their energy and acting as a closed waveguide allowing some radiation to resonant in the space between the conductive sphere of the Earth and the conductive spherical shell of the ionosphere [*Bilitza*, 2001].

In 1952, Winfried Otto Schumann (1888 - 1974) mathematically predicted a resonance of ELF radiation within the aforementioned waveguide. In 1960, measurements performed by Balser and Wagner recorded the Schumann Resonance (SR). The Schuman Resonance has a fundamental frequency about 7.8 Hz, with harmonics spaced at 6.5 Hz (14.3, 20.8, 27.3, and 33.8 Hz) [*Galejs*, 1965].

Lightning from thunderstorms provide the majority of the source energy for Schumann Resonance. Geographically, there are three main sources for lightning on Earth and they are Africa, South Asia and the Americas. There is also a seasonal lightning source located near Tahiti that contributes in the southern hemispheric summer, but is not nearly as strong as the other three. Table 4.1 shows the locations and times of the different lightning sources. All the times are in Coordinated Universal Time (UTC) which is essential because it provides a common time base for global events.

Table 4.1: Location, time of activity, distance from Africa and generalized geographic coordinates of major lightning source areas.

Location	Time (UTC)	Coordinates
Africa	1430 – 1730	0°N, 25°E
South Asia	0800 – 1100	0°N, 100°E
Americas	1730 – 2230	0°N, 60°W
Central Pacific	2000 - 2300	17°S, 150°W

The table shows the centers on the equator because seasonally the locations shift north or south depending on which hemisphere is experiencing summer. The mid-latitudes loosely

confine the extent for this variation. Figure 4.1 displays the five-day average lightning density from September 09, 2010 to September 13, 2010. The image clearly shows the three major lightning locations and that these locations are north of the equator because it is the end of summer in the northern hemisphere.

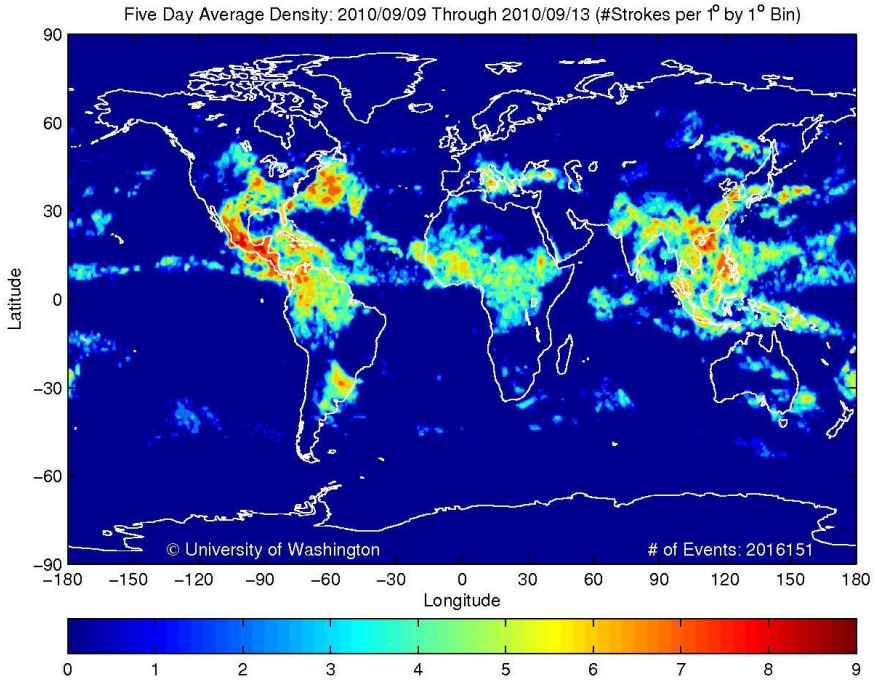

Figure 4.1: Five-Day average lightning density from 09/09/2010 to 09/13/2010. The colors show the number of strikes per 1° grid. The three lightning locations over the continents are clear. (Source: http://wwlln.net/five_day_density_average.jpg)

Lightning usually occurs in the lowest 10 km of the atmosphere (Troposphere). As the major excitation source for the SR, there are about 100 lightning events occurring globally every second. In the three major lightning source regions, tropical thunderstorms develop predominantly in the local afternoon. The released lightning energy is large. A typical lightning bolt bridges a voltage of several hundred million volts, and develops a peak current of up to 10 kA. Each strike can generate several hundred megawatts of electrical power. With this level of the transient current, the maximum magnetic

induction can instantaneously reach a level of 10^{-4} Tesla, comparable to the strength of the major constituent of the static geomagnetic field.

4.2 Esashi Magnetotelluric Data

The data used for this chapter was collected at the Esashi magnetotelluric observatory. The Esashi observatory (ESA) is located in Iwate Prefecture in the northern part of Japan's main island (39° 14' 13" N, 141° 21'17" E, 395 m). The Esashi observatory uses the Phoenix Geophysics System 2000 survey equipment measuring two electric (E_x and E_y, north-south and east-west, respectively) and three magnetic channels (H_x, H_y and H_z, north-south, east-west, and vertical, respectively). The data collected at ESA have three sampling rates: 15 Hz, 150 Hz and 2400 Hz. With duty cycles of continuous collection for the 15 Hz, 16 seconds on and 4 minutes off for the 150 Hz, and 1-second on and 4 minutes off for the 2400 Hz data. For this thesis, the 150 Hz sampled data was analyzed. Figure 4.2 shows the locations of the station in relation to Japan's main island and the proximity to the March 11, 2011 Tohoku earthquake.

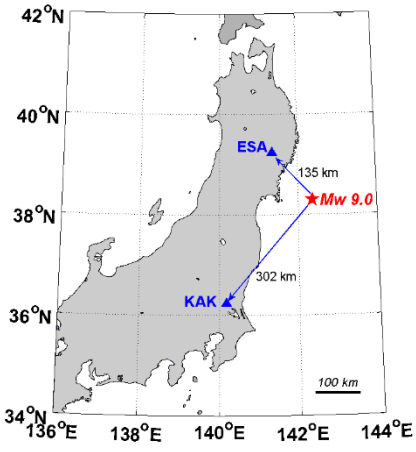

Figure 4.2: Location of the ESA site relative to the island nation of Japan and the proximity to the 2011 Tohoko earthquake.

The Schumann resonances form a standing wave in the cavity between the ionosphere and the Earth's surface. The approximation of the Schumann resonances and their modes can be done by analyzing the Earth's cavity as a spherical shell [*Toledo-Redondo et al*., 2010]. The amplitude response is a function of angular distance between the source and receiver. Table 4.2 shows the angular distance between the Esashi station and the individual major areas of lightning sources and amplitude responses for modes 1 to 3. The amplitude response values are approximations based on the work of Toledo-Redondo et al. [2010]. The azimuth angle of the incoming plane wave determines which sensor component measures the incoming wave. As the magnetic perturbation vector is aligned perpendicularly to the propagation wave vector, sensors orthogonal to the propagation vector record the wave energy more effectively [*Manu et al*., 2015].

Table 4.2: Table showing the location of the Esashi station relative to the lightning centers, the theoretical amplitude response and sensor component that measure the largest signal.

Location	Angular Distance	Azimuth Angle	Amplitude Response			Sensor Component
			Mode 1	Mode 2	Mode3	
Africa	1.97 Rad	260°	0.35	0.23	0.04	NS
Americas	2.80 Rad	100°	0.17	0.15	0.1	NS
South Asia	0.86 Rad	185°	0.27	0.27	0.2	EW

The seasonal variation of thunderstorms affects the measured amplitude of the SR. As the summer and winter move between the hemispheres, the locations of the thunderstorms move in latitude. This change in latitude changes the azimuth angle of the plane waves and the different sensor components measure this change. As storms move from South America to North America, the azimuth angle relative to the Esashi station changes from 120° to 90°, respectively. The expectation is to see a decrease in the amplitude of the SR

for the NS component in southern summer, since the thunderstorms are now located around 30°S as opposed to 30°N.

The Esashi data are freely available for download from the GSI geomagnetic surveying web page (http://vldb.gis.go.jp/sokuchi/geomag).

4.3 Data Processing, Analysis and Interpretation

Data processing followed the basic data processing algorithm with the resultant data reduced to hourly, monthly and seasonal blocks. The basic data processing algorithm is as follows:

1. Read hourly data from file
2. Processed with EMD to generate IMFs
3. Sum IMFs 1 – 4
4. Generate FFT of new signal
5. Reduce data to mean values and standard deviations

Prior to my processing, the only apparent processing done by the station was the removal of the 50 Hz power-line noise. For the magnetic components, the filtering response appears to have leaked or otherwise introduced periodic noise above 40 Hz. For the purpose of this thesis, no data above 35 Hz is used. The data did exhibit a large amount of power in the 0 to 5 Hz range and was contaminating the analysis. While other researchers used Hanning windows [*Sinha et al.*, 2008; *Toledo-Redondo et al.*, 2010] or Lorentzian curves [*Manu et al.*, 2015], the filtering technique used for the thesis was the EMD technique similar to the noise reduction techniques used in the last chapter. The EMD algorithm returned between 10 – 15 IMFs from the original signal. Fourier transforms computed the frequency spectrum for each IMF and when the majority of the energy was

below 5 Hz, the processing algorithm discarded the remaining IMFs. Typically, IMFs 1 – 4 contained all the energy required for the analysis. Figure 4.3 shows the original data (top left), the frequency spectrum for that data (top right) then the IMFs from 1 to 5 in descending order (left side) and their associated frequency spectra (right side).

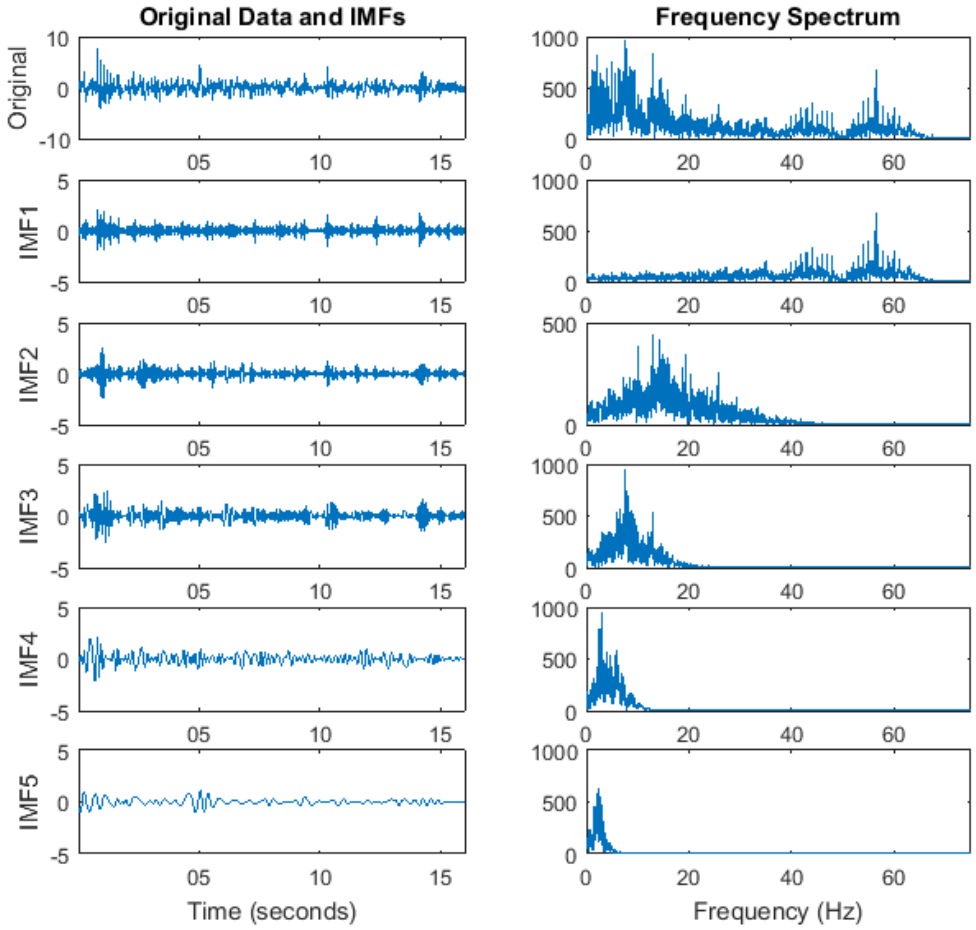

Figure 4.3: The left column contains the IMFs of the original signal. The right column contains the frequency spectrum.

Another processing step was a basic five point smoothing filter to average out data spikes. The window size varied based on the sampling rate and the amount of noise present in the signal. To perform the amplitude and frequency extraction, the processing algorithm used a windowing function focused around the SR frequencies. Window sizes

for modes 1, 2, 3, 4, and 5 were 6.5 – 8.5 Hz, 13 – 15 Hz, 19 – 22 Hz, 25.5 – 28.5 Hz, and 31 – 35 Hz, respectively. For these frequencies, the algorithm extracted the max amplitude within the window. Only the horizontal magnetic components were processed. Schumann resonances visible in the data were the fundamental mode and four harmonic modes. The fourth mode, at times, had a broad peak faint appearance.

Schumann Resonance amplitudes change as the seasons change because the locations of the thunderstorms move north and south relative to the equator. Processing a year's worth of data and looking at each month separately, it is clear by the disappearance of resonance patterns from the components that the source of the plane waves has changed. Figure 4.4 shows the standard SR banding for the north-south component with an increase in amplitude between 1300 and 1600 hours, this coincides with the African storms. The 2^{nd}, 3^{rd}, 4^{th}, and 5^{th} modes show an increase in amplitude at this time block. In addition, the figure shows a small rise in the amplitude between 2100 and 2300 hours, this coincides with the South American storms. Figure 4.5 show the standard SR banding for the east-west component with an increase in amplitude between 0600 and 1000 hours, which coincides with storm over Malaysia. As waves from Malaysia propagate northward, the magnetic field varies in the east-west direction. With the wave traveling from Africa and South America (traveling more in an east-west direction), the magnetic field varies in the north-south direction.

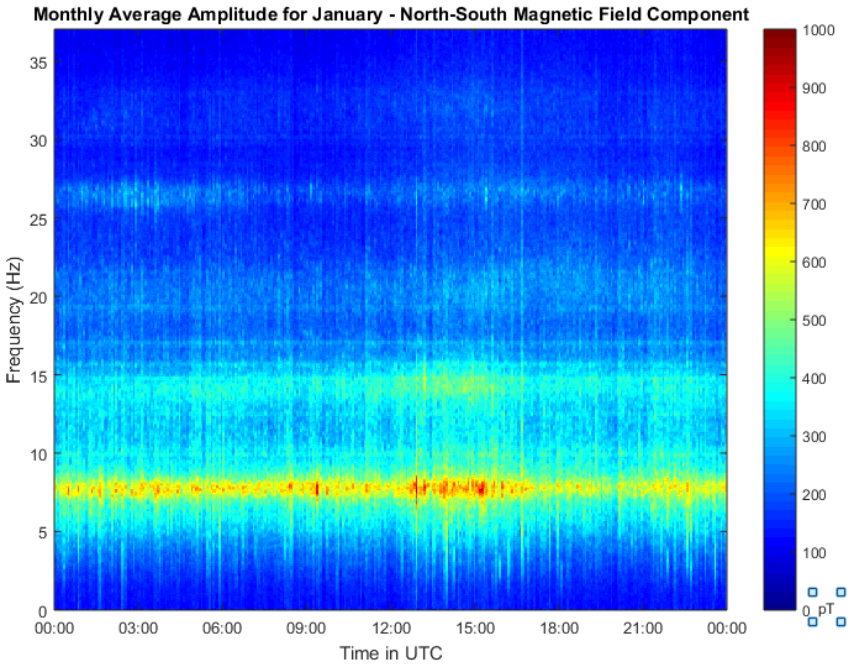

Figure 4.4: North-South magnetic field component for the month of January.

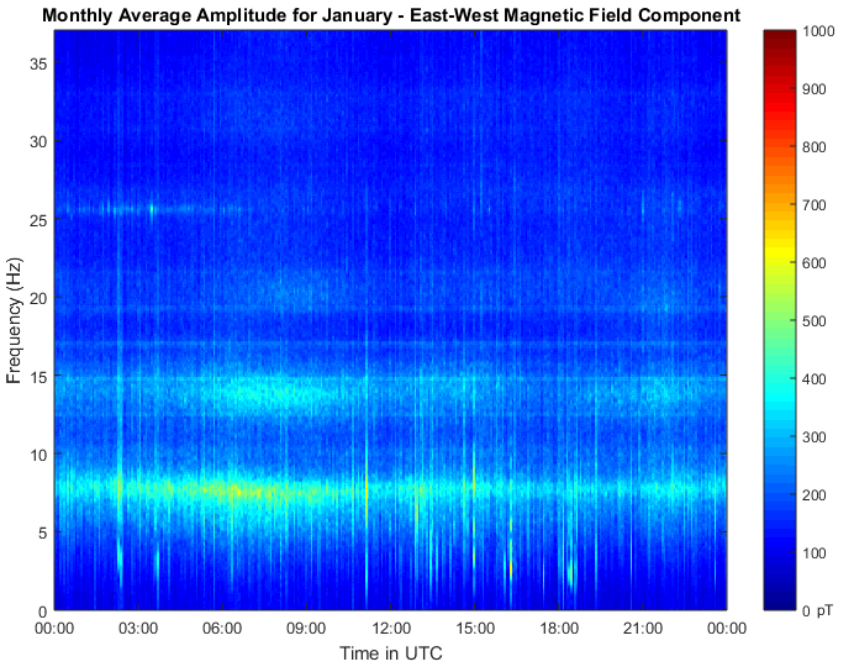

Figure 4.5: East-West magnetic field component for the month of January.

When the seasons switch, the incidence angles of the incoming magnetic wave change because the storms moved northward. Figure 4.6 shows the north-south component and the data contains more energy around the 0600 hour. The interpretation of this extra energy is attributed to the movement of the wintertime Malaysian storms northward, which are now located in Southeast Asia. However, there still exists energy at the same time in the east-west component, meaning the storm's locations are at some odd interval of $\pi/4$. There is also energy present in the 2200 time block in the north-south component; this is the energy from the North American storms, see Figure 4.6. Figure 4.7 shows the East-West component for the month of July. Most of the energy is at the 0600 and 2200 hours with a small increase at 1400. Since the Asian storms have moved north, there is an East-West component to those storms. However, the component response is not as significant as the North-South component (the storms are farther north). There is a decrease in the energy around the 1400 hour for the African storms. Work by others [*Sinha et al*., 2008, *Toledo-Redondo et al*., 2010, *Ortega et al*., 2014, *Manu et al*., 2015] suggests that global weather patterns affect the SR. As thunderstorm activity decreases, the amplitude of the SR decreases. The lack of any visible amplitude change (from the base amplitude) of the SR in July around the 1400 hour suggests either Africa had less lightning in July of 2011 or the location of the storms were close to a node. Given the distance between the Esashi observatory and the three storm centers, the fundamental mode of the SR should always be present and it is not. The SR amplitude is a function of the strength of the storms producing the lightning. Therefore, if the storms are not very storm then there would not be any noticeable increase in the amplitude.

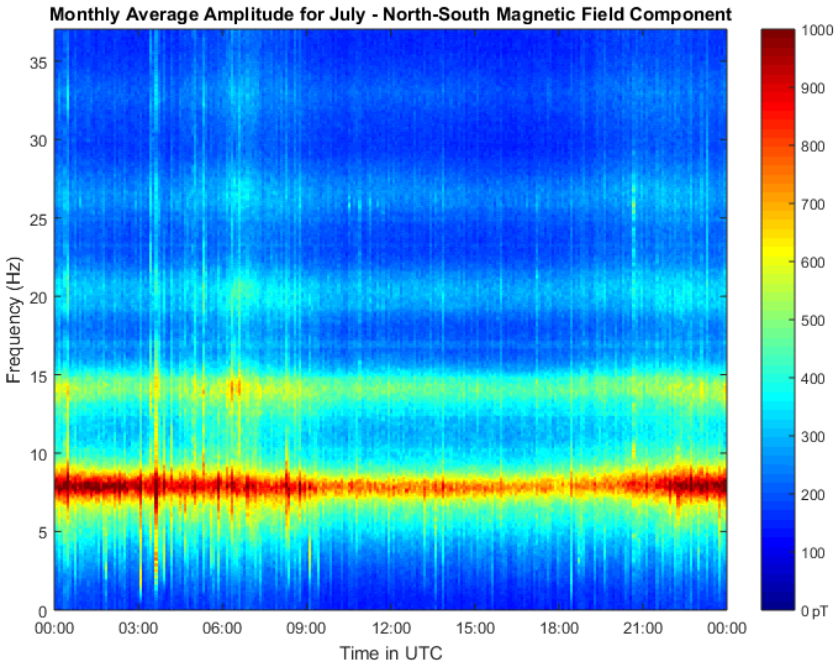

Figure 4.6: North-South magnetic field component for the month of July.

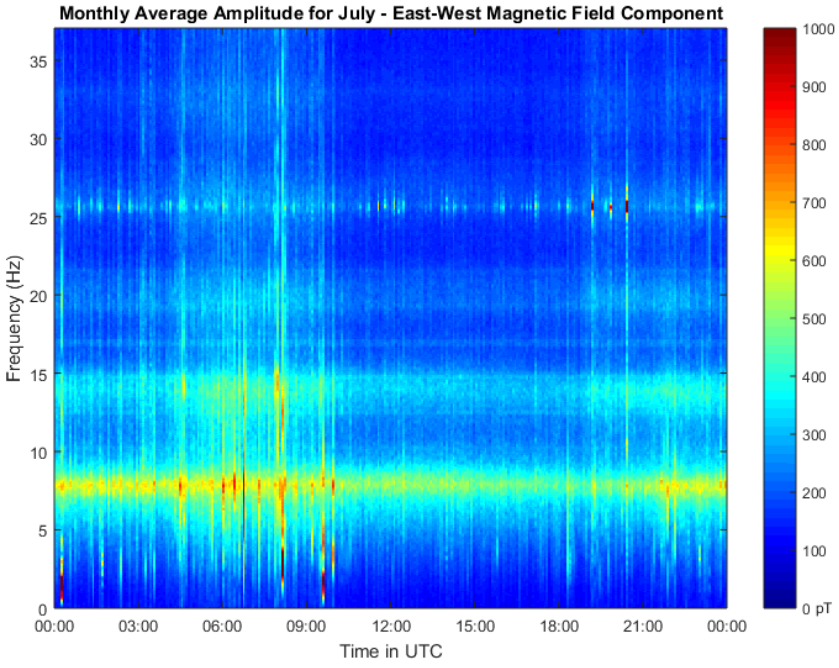

Figure 4.7: East-West magnetic field component for the month of July.

Figure 4.8 shows the north-south component of the magnetic field for September with small amplitude increases for all storm locations but more around the 2100 hour. September is the start of the equinox and the storms start to head south. Figure 4.9 shows the east-west component for the same month. Hours 0800, 1400 and 2100 display increases in amplitude in the fundamental frequency with some energy in the second mode. As the months' progress, the energy slowly moves back to where the analysis started in January. Figure 4.10 shows the north-south component for the month of December. There is an increase in the amplitude around 1400 hour (African storms) with not much energy in the other time blocks. Figure 4.11 shows the east-west component and while there are increases in all time blocks, the Malaysian time block shows the brightest amplitude.

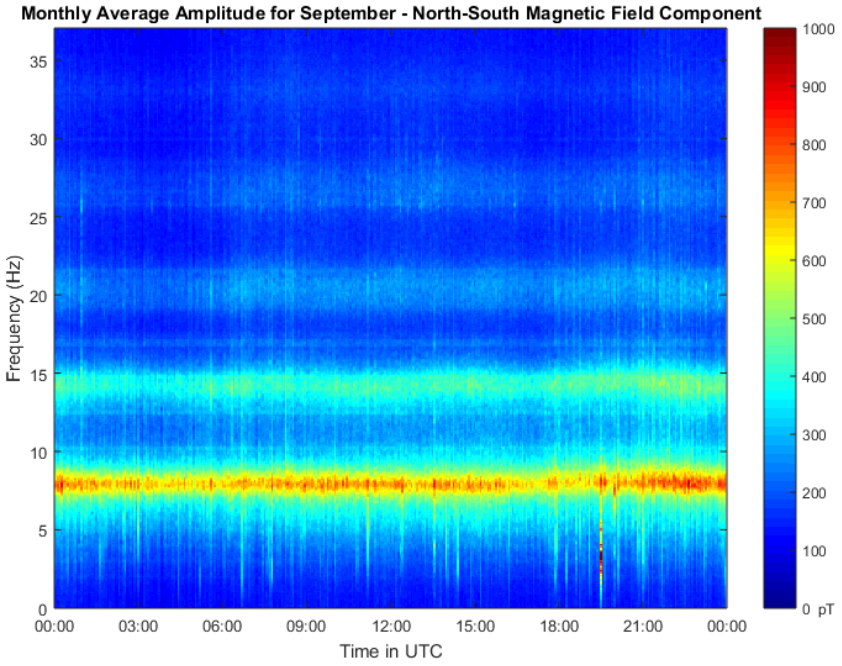

Figure 4-8: North-South magnetic field component for the month of September.

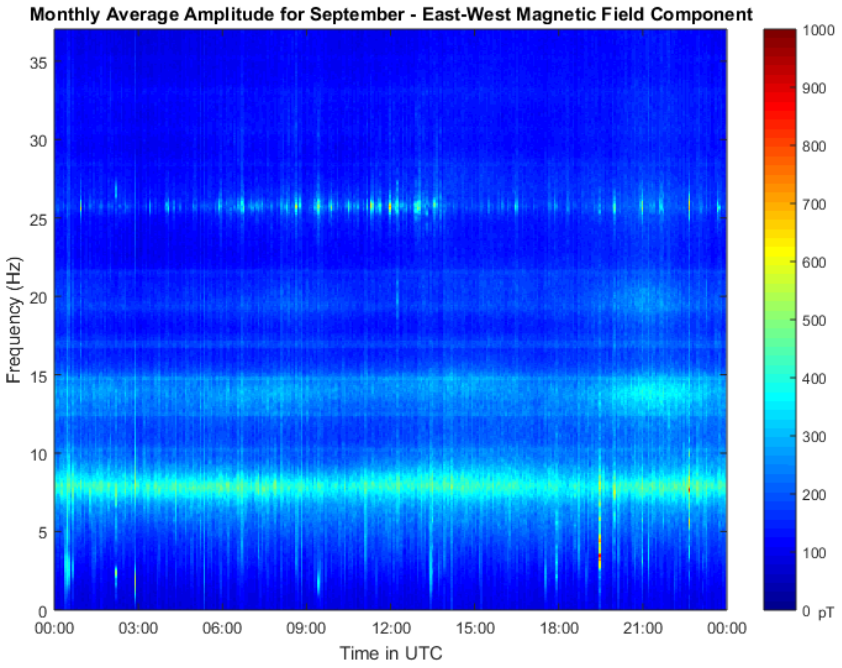

Figure 4.9: East-West magnetic field component for the month of September.

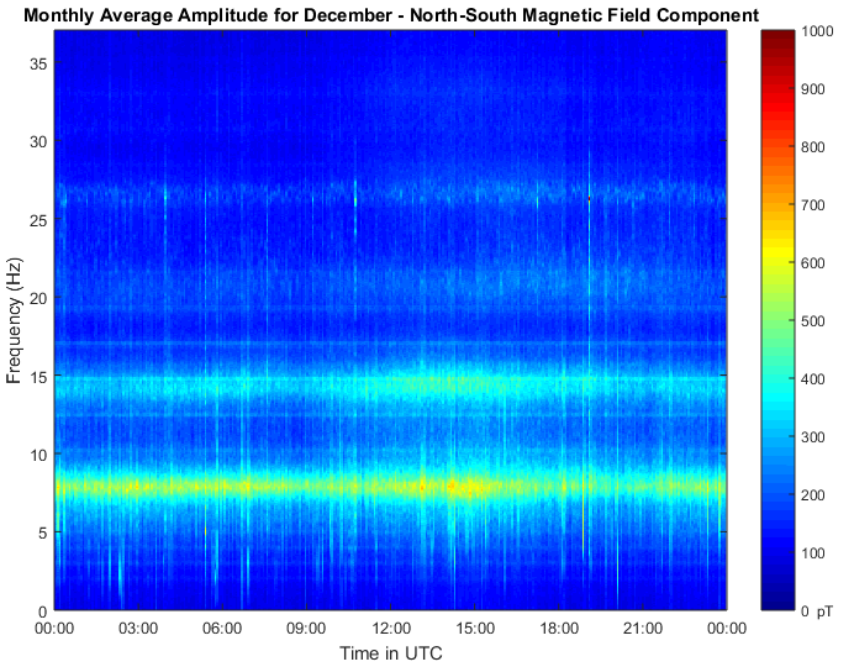

Figure 4.10: North-South magnetic field component for the month of December.

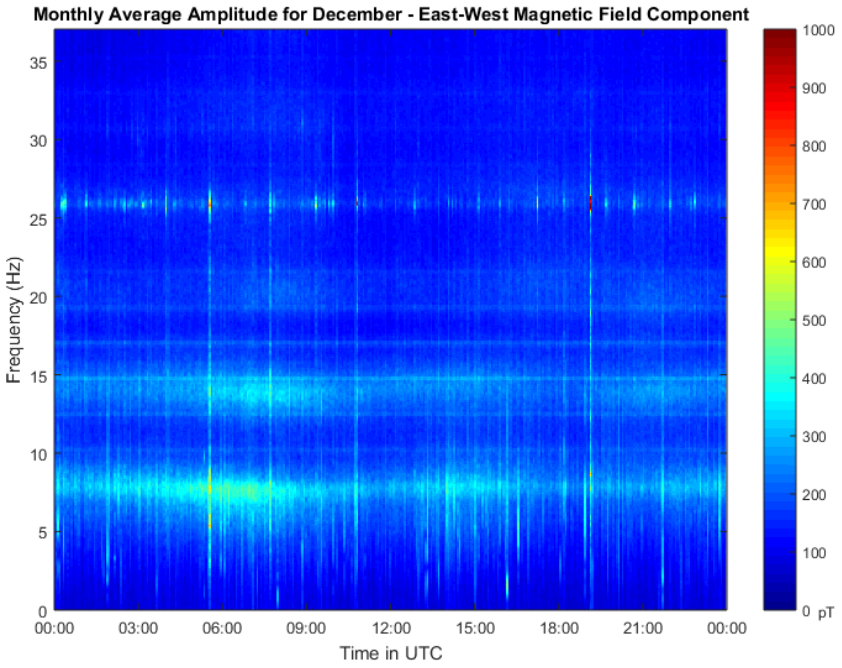

Figure 4.11: East-West magnetic field component for the month of December.

To better contrast the differences between the summer and winter season, Figure 4.12 shows the changes between seasons. In the summer season, the source of the amplitude spike at the 0800 hour is the Southeast Asian storms. The amplitude change relative to the baseline is about the same between the components. This means the plane wave from this source is at a bearing of 225° from the Esashi observatory. The incident angle of the wave affects both the East-West and North-South components equally. In the winter season, the North-South component nearly disappears for the Asians storms. As the storms moved south, the incident plane wave's bearing is closer to 180°. The African storms share equal relative amplitudes in both seasons. The bearing from Esashi to the African locations does not change as drastically as it does for the Asian storm center so the seasonal amplitude change is not as large.

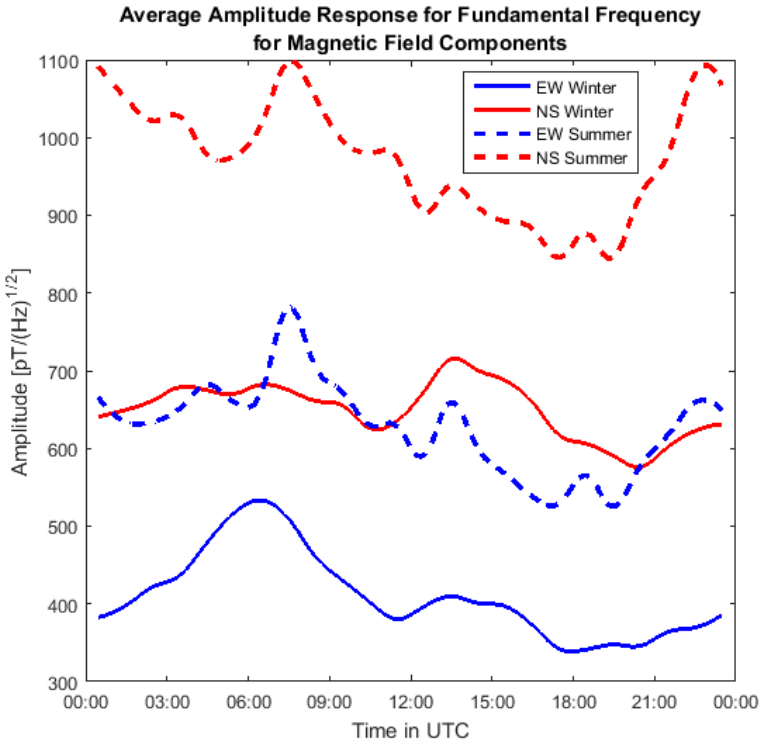

Figure 4.12: Seasonal variation of the fundamental mode of the Schumann Resonance for two components of the magnetic field. Summer and winter seasons are for the northern hemisphere.

The American storms have a small contribution around 1900 hours in the summer components. In the winter season, there is no contribution to either component. The reason is the relative location of the South American storms to the Esashi station. The angular distance between the two is 2.8 radians, which means the expected response is very close to a node of the fundamental frequency. As the storms move north in the summer season, the amplitude response improves as the angular distance reduces.

The peak around 2300 hours in the summer season is nearly nonexistent in the winter season. The cause of this peak is the North American storm center is in the Great Plains. Since there are no strong storms that are lightning producing in North America in the winter season, a peak is not expected. South America is located farther east than North America so the storms in South America happen 3 to 4 hours earlier.

The geometry of the ionosphere changes throughout the day as the Earth rotates about its axis. As was stated in the beginning of the chapter, the Sun creates the D layer [*Johnson*, 1965]. The interaction between the increased solar wind and the gases in the upper atmosphere adds more energy and allows for more free ions. The D layer is 30 km below the E layer [*Johnson*, 1965]. This decrease in height should increase the resonance frequency as the thickness of the waveguide decreases. The Schumann Resonance measured during the day should be higher than measured at night with the peak frequency occurring around noon (local time) and the lowest frequency around midnight (local time).

Table 4.3: The time of the sun rises, sun sets and sun zenith in local time and in UTC (in parentheses).

Month	Sun Rise	Sun Set	Sun Zenith
March	5:50AM (2050)	5:50PM (0850)	11:50AM (0250)
June	4:26AM (1926)	7:00PM (1000)	11:43AM (0243)
September	5:20AM (2030)	5:40PM (0840)	11:35AM (0235)
December	6:50AM (2150)	4:30PM (0730)	11:39AM (0239)

An analysis of the fundamental mode's frequency shows a diurnal change. The North-South component of the magnetic field shows the change more than the East-West component. Figure 4.13 and 4.14 shows the fundamental mode of the North-South and East-West components of the magnetic field for January and June, respectively. The processing algorithm is as follows:

1. Process the data into IMFs.
2. Sum IMFs 1 through 4.
3. Compute the FFT.
4. Filter the FFT amplitudes with a mean filter with a window of 7.

5. Find the mean amplitude for 1 hour of data around the frequency range of the fundamental mode.
6. Find the frequency corresponding to the maximum of the last step.
7. Filtered the frequencies response with a mean filter with a window of 5.
8. Remove any noisy responses.
9. Averaged the five-day blocks together.

The output of the algorithm is a matrix of the peak amplitude and its frequency for each hour of the day for each block of days, which is five days. An attempt to perform this analysis on a month of data produced very confusing results. Processing the data in five-day blocks then averaging those responses proved to be a better algorithm. This process allowed for the review of the data prior to averaging to remove any outlying blocks. Generally speaking, the data from Esashi were good, though some days had very high amplitudes and these required removal prior to any frequency analysis.

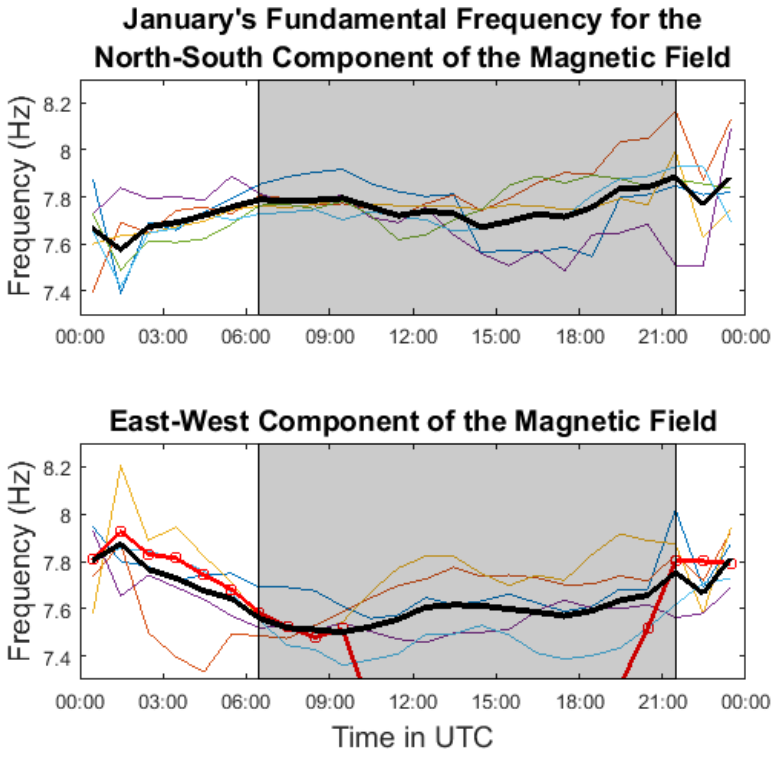

Figure 4.13: The frequency of the fundamental mode for North-South (top) and East-West component of the magnetic field verses the time of day (UTC) for the month of January. The gray box represents nighttime. The red line shows the eliminated data. The black line is the average of the used data.

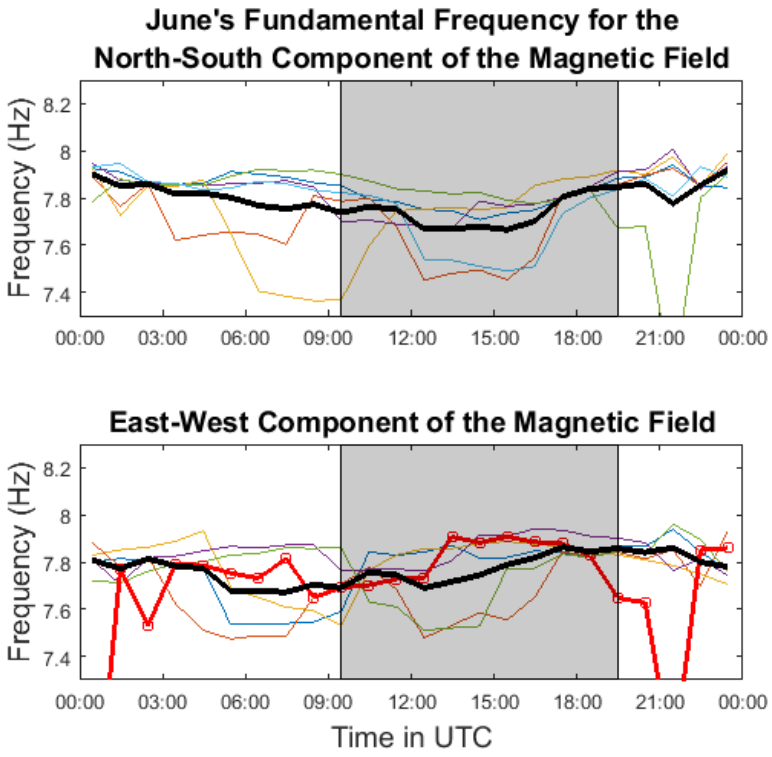

Figure 4.14: The frequency of the fundamental mode for North-South (top) and East-West component of the magnetic field verses the time of day (UTC) for the month of June. The gray box represents nighttime. The red line shows the eliminated data. The black line is the average of the used data.

The colored lines in the figures show the fundamental frequency for each of the five day blocks, with the bold red line being data that were removed as an outlier and the black line the average of the good data. The grey box shows when it was nighttime at the Esashi observatory. The time axis is in UTC, not in local Japan time. The general trend is higher frequencies during the day with lower frequencies at night, though this is not always the case because the size of the ionosphere is not constant. The ionosphere changes and shifts as a function of incoming radiation. The trend for the January North-South component and the June East-West component do not strictly follow the trend of their orthogonal component, their trend is more flat than U-shaped. The central frequency for January and June is around 7.7 Hz and 7.8 Hz, respectively. The location and power

of the storms influence the central frequency. With Asian storms closer to the observation point, this may account for the small increase in the central frequency.

These results are in agreement with results from Toledo-Redondo et al. [2010] and Manu et al. [2015], in that there is an oscillation around the central frequency but the frequency of the resonance is affected not only by the ionosphere geometry but also by the spatial size of the main storms producing the resonances.

4.4 Earthquake Precursors

There have been many studies looking at anomalies in Schumann Resonances that occur days prior to a major seismic event. One paper looked that the Tohoku earthquake and observations from mainland China [*Zhou et al*., 2013]. Their conclusion was that they could see an anomaly in the higher modes of the SR on March 8th or three days prior to the March 1lth earthquake. Changes in the SR is a result of changes to the Earth conductivity due to changes in the subsurface [*Xu et al*., 2013].

Having looked at numerous days of frequency spectrum plots, the spectrum on March 6th 2011, stood out. Between the hours of 0900 and 1500, there is new power in the 3rd and 4th modes (22 and 27 Hz). This change in amplitude and frequency is not seen in any of the data for any other day. March 6th is five days prior to the March 11th earthquake. The average response from March 1st to March 5th was computed and subtracted from the March 6th spectrum. Figure 4.15 shows the response. This anomalous region is a departure from the days prior and is different than what was computed from Zhou et al. [2013] but could be a result of the proximity of the ESA observatory to the epicenter of the earthquake. Since the observatory is closer, the effects are stronger. Further analysis

is required to determine if this anomaly is an earthquake precursor or from a man-made source.

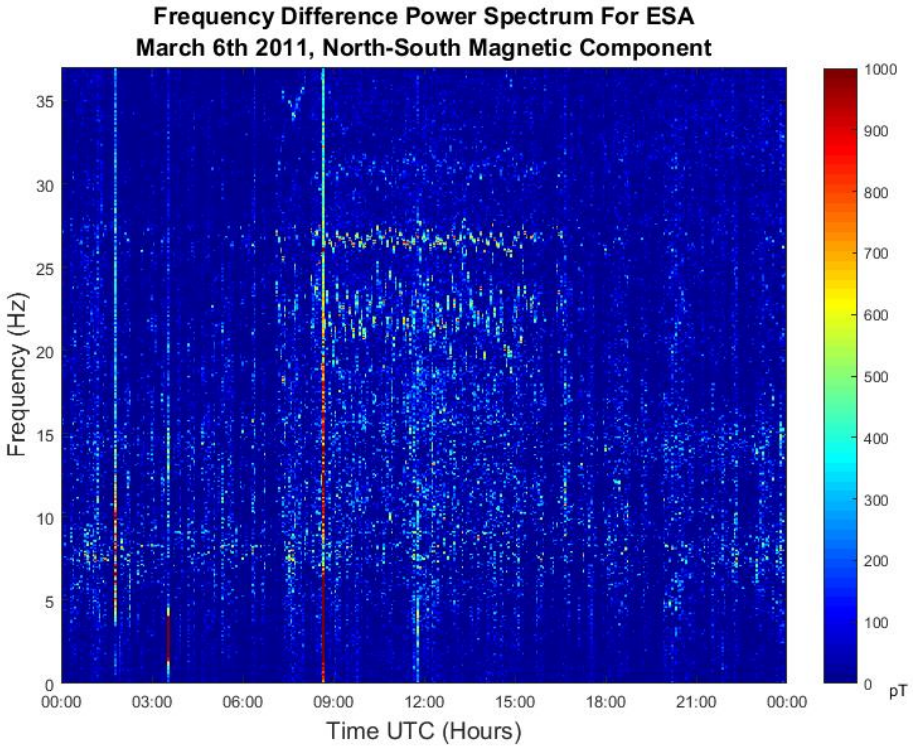

Figure 4.15: Frequency spectrum difference for March 6th, 2011 North-South magnetic component.

EMD has been used to remove unwanted frequencies in the magnetotelluric data record without any adverse effects to the analysis of the Schumann Resonances. More work can be done on the analysis of earthquake precursors and the Hilbert Huang Transforms. Using the Hilbert Huang Spectrum could provide new insights in how the Schumann Resonances change days before a major earthquake.

Chapter 5: Bi-dimensional Empirical Mode Decomposition Application

5.1 Texture Analysis

Texture analysis is a difficult problem in image processing when using Fourier-based analysis because textures are non-stationary. The problem with texture analysis is the vast number of primitive textures that compose larger textures not only in the grey scale but also in the color or multi-band. Texture analysis is used in all types of applications from remote sensing, medical, fingerprint and document analysis, photographic effects, manufacturing quality assurance, pattern recognition, and target detection. The intent of this section is not to discuss texture analysis in great depth but to provide an example of how Bi-dimensional Empirical Mode Decomposition (BEMD) decomposes images that are composed of different textures.

Prior to BEMD, signal processing techniques used filter response to characterize textures. These filters were derived from Gabor functions (a special case of the short time Fourier transform) or wavelet transforms, both of which require a certain amount of *a priori* information and are subject to the response of the function and type of wavelet used. Both of these techniques intend to represent the images with both spatial and frequency information present.

To demonstrate BEMD, a number of simple textures were synthesized using a checkerboard pattern of two different sizes with a linear feature running parallel to the X-axis and at a 45° angle. These simple textures have different wavelengths and demonstrate the BEMD response to different wavelengths. The first texture is the small checkerboard (a), horizontal lines (b) and the summation of the two surfaces (c).

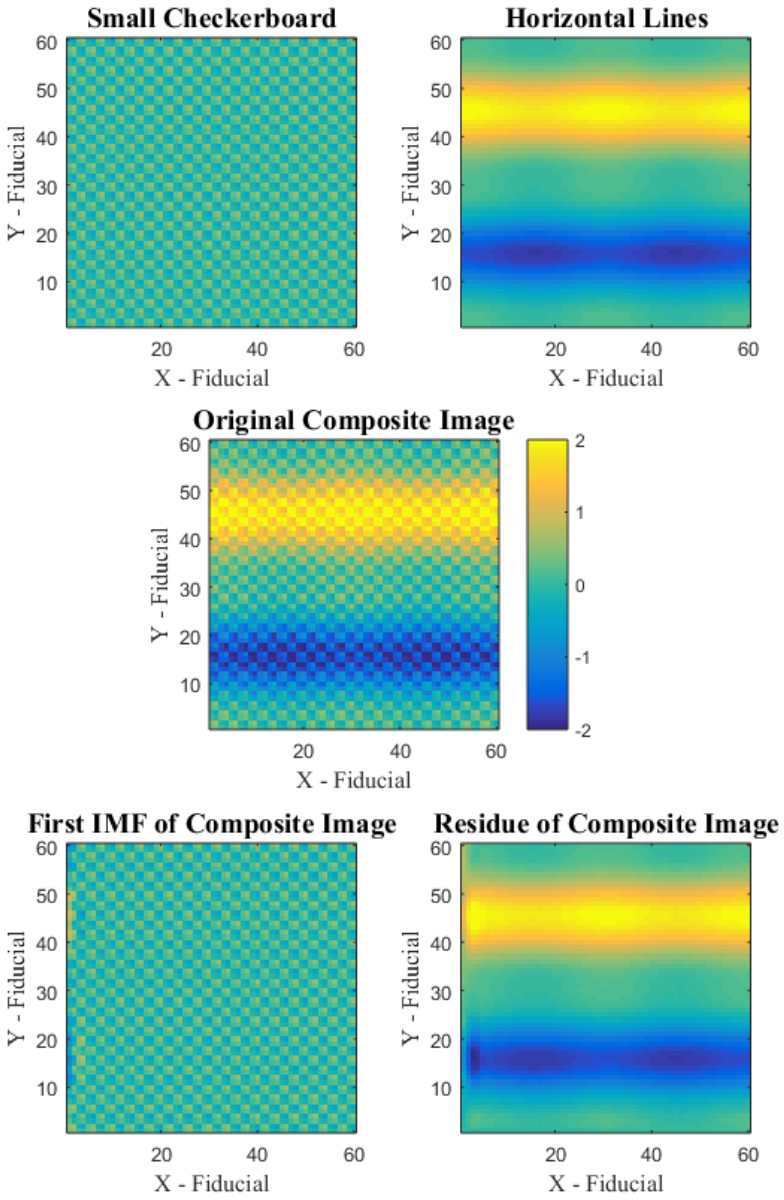

Figure 5.1: Simple texture analysis: a) small checkerboard pattern; b) two horizontal lines with positive and negative values; c) the summation of the checkerboard and the horizontal lines; d) first IMF surface containing all the checkerboard energy; e) the residue surface containing the horizontal line energy.

The first texture is the small checkerboard with horizontal lines of alternating amplitudes.

The summation of the textures is shown in Figure 5.1 (c). BEMD is run on this surface

and instructed to return one IMF and the residue. The reason for this is that the composite image contains only two textures so the decomposition should only return two surfaces. If the result set contained three surfaces (two-IMF and a residue), the residue surface would contain noise and have no physical meaning. The first IMF surface of the composite images, shown in Figure 5.1 (d), contains the wavelengths of the small checkerboard texture and some artifacts caused by the interpolation, however this is only limited to the first four pixels in the X-direction of the image. The residue of the composite image, shown in Figure 5.1 (e), contains all of the energy of the horizontal line texture. There is a certain amount of amplitude fluctuation in the horizontal line; this energy is created from the basis function used to generate the feature. What is remarkable is that the residue contains this energy.

A quick visual inspection of the components and the decomposed images shows that there appears to be subtle differences in the residue, aside from the obvious artifacts on the left side of the decomposed images. Since the first IMF contains the energy from the shorter wavelength, I subtracted the first IMF from the original small checkerboard image (Figure 5.2 (a)) and the noise is within 5% of the original maximum amplitude (not including the left edge effect). Similarly, the residue and horizontal surfaces have the same amount of relative error, which is shown in Figure 5.2 (b). There is a conversation of energy happening in the error surfaces. Wherever there is a negative value on one surface there is a positive value in the exact pixel of the other surface, so overall there is a mean zero error between the original and decomposed images. This was confirmed in subtracting the original surface with the summation of the IMF and residue, as shown in Figure 5.2 (c), where the maximum error is 2.2×10^{-16} or 0. The horizontal line and

checkerboard case was a pretty simple one and shows how well the decomposition works for simple surfaces.

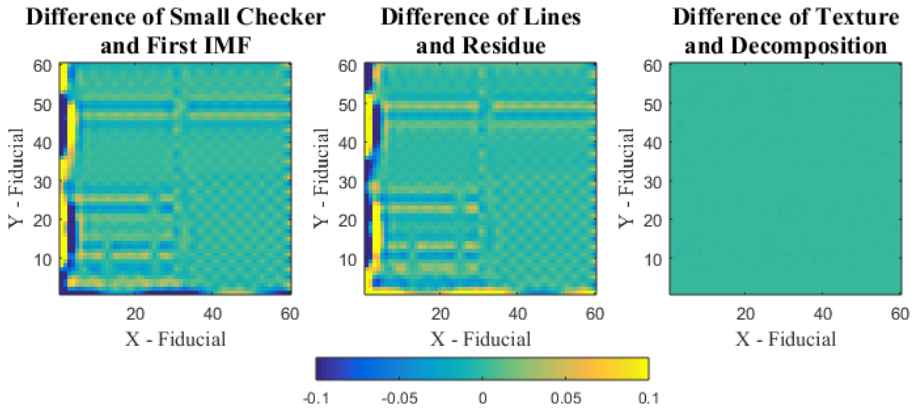

Figure 5.2: Error plot of the two original surfaces with the IMF and residual surface: a) error of small checkerboard and first IMF; b) error of horizontal lines and residual surface; c) error of the original composite surface and summation of IMF and residual surfaces.

In the previous example, the decomposition worked well for an image composed of two different wavelengths. For the next example, I added a third and similar wavelength to the image. Similarly, to the previous example, the final image is composed of a small checkerboard Figure 5.3 (a), a big checkerboard Figure 5.3 (b), and two horizontal lines with a positive and negative response Figure 5.3 (c). The larger checkerboard has a wavelength similar to the wavelength of the horizontal lines.

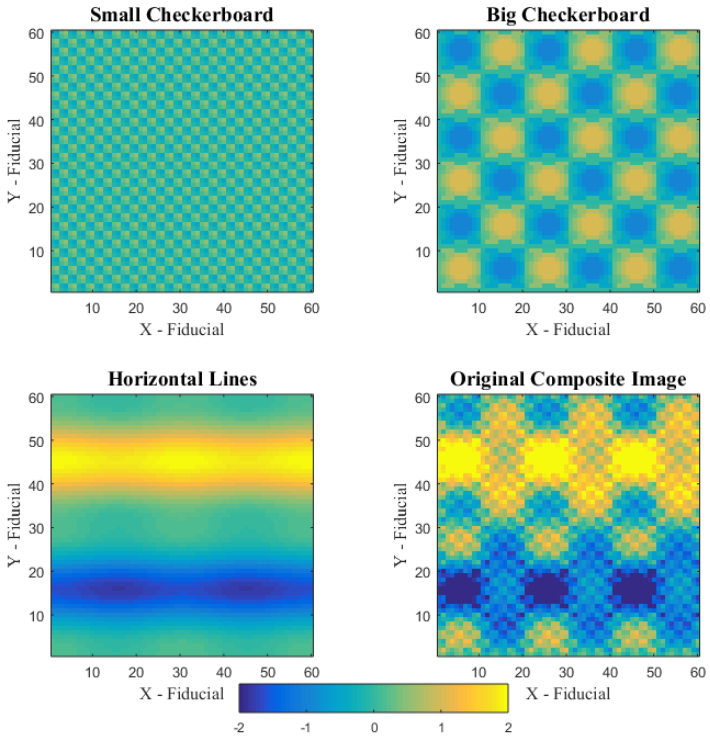

Figure 5.3: Simple three texture summation: a) small checkerboard; b) large checkerboard; c) horizontal lines; d) summation of small and large checkerboard and horizontal line surfaces.

The BEMD results contain the same discontinuities between 10 to 20 and 40 to 50 on the Y-axis while the rest of the images are different than the previous example. The first IMF of the composite image, Figure 5.4 (b), distinctly shows the checkerboard pattern however the image bandwidth is a larger with a cyclical amplitude modulation across the texture. This is caused by a resonance between the smaller checkerboard and the edge of the larger checkerboard patterns. The second IMF surface, Figure 5.4 (c), contains the energy contributed by the larger checkerboard texture. The larger checkerboard pattern is clearly seen, but the amplitude difference looking down alternating checkboard columns of ten pixels is caused by the similarities in wavelength between the checkerboard and the horizontal lines. The residue surface, Figure 5.4 (d), contains the leftover energy associated with the horizontal lines. Since the wavelengths are so similar between the

large checkerboard and horizontal lines, the energy in the residue surface is smeared compared to the previous example.

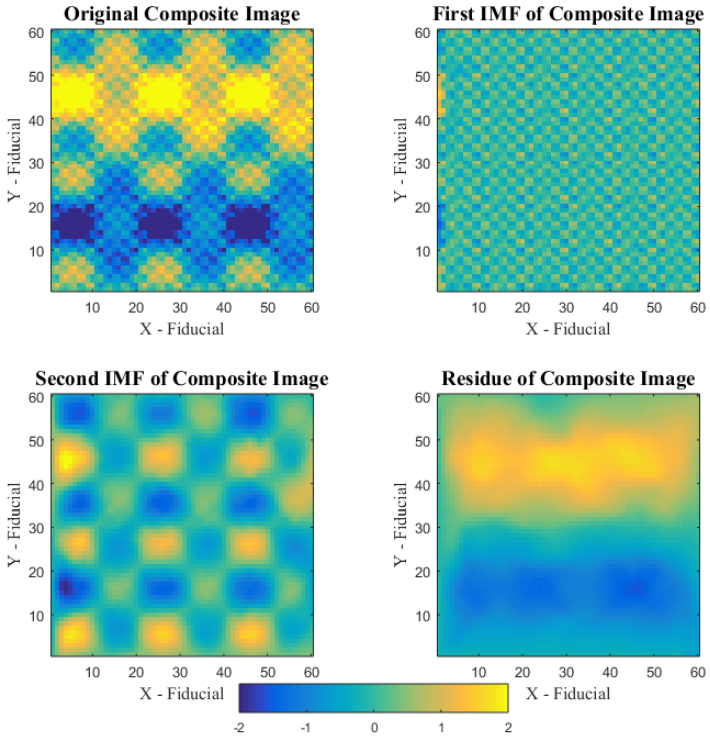

Figure 5.4: Decomposition of three texture surface: a) original texture; b) IMF1 of original texture; c) IMF2 of the original texture; d) residual signal of original texture after IMF extraction.

With the introduction of the larger checkerboard pattern, the frequencies of the surface begin to mix. This mixing is called mode mixing. Mode mixing happens when an IMF contains energy from multiple modes, as in this case when the original texture only contains three textures. Figure 5.5 (a) shows some error within the first few pixels but shows little error between the texture and IMF1. The same can be said for images (b) and (c) in that the trend of the error show the leakage of the energy between the IMF and the corresponding texture. Even with this energy leakage between the IMF and textures, the summation of the IMFs and the original image show no obvious errors.

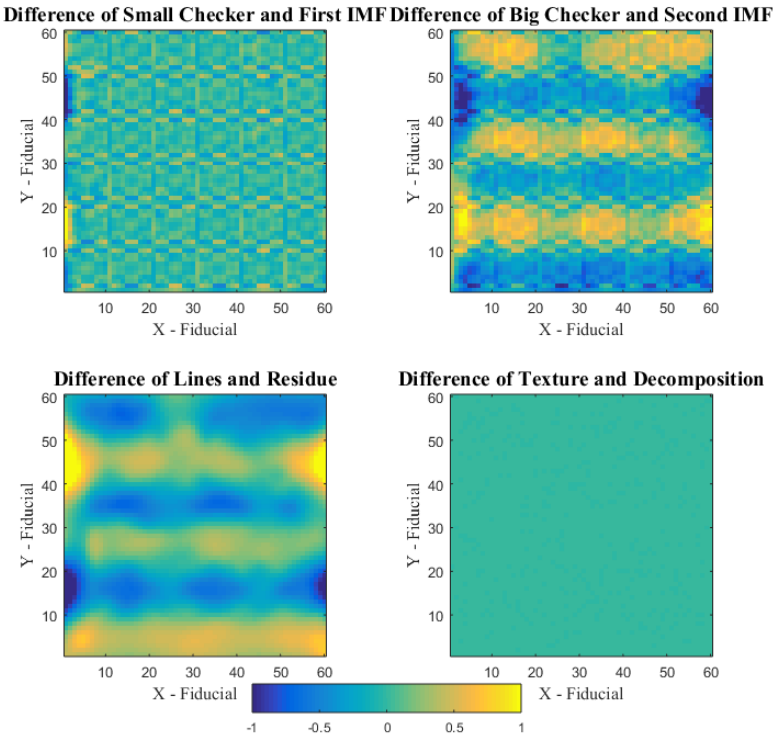

Figure 5.5: Error plot of the three original surfaces with the IMF and residual surfaces: a) error of small checkerboard and first IMF; b) error of the large checkerboard and the second IMF; c) error of horizontal lines and residual surface; d) error of the original composite surface and summation of IMF and residual surfaces.

The third example is similar to the previous two but the horizontal lines were replaced with a diagonal feature, Figure 5.6 (c). The decomposition of this texture is similar as in the previous example. The amplitude of the checkerboard pattern is more consistent in the second IMF with parts being distorted by the decomposition, Figure 5.7 (c). The residue surface contains the energy of the diagonal feature with some effects from the interpolation around the edges, Figure 5.7 (d). Because BEMD is an empirical technique and isolates the localized wavelengths, there will be some leakage between IMF surface and residue when the wavelengths are within 2λ (where λ is the wavelength).

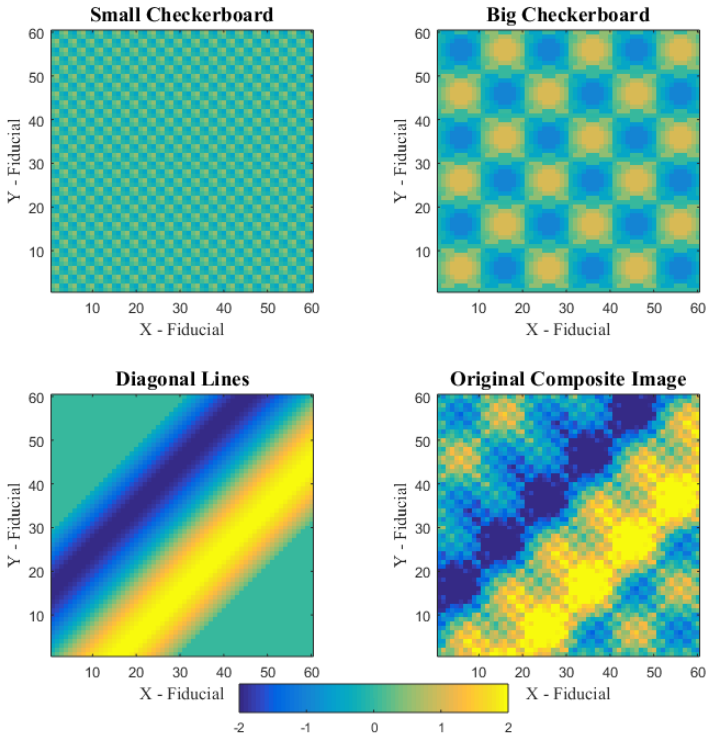

Figure 5.6: Simple texture with checkerboards and diagonal lines: a) small checkerboard surface; b) large checkboard surface; c) diagonal line surface; d) summation of the three textures.

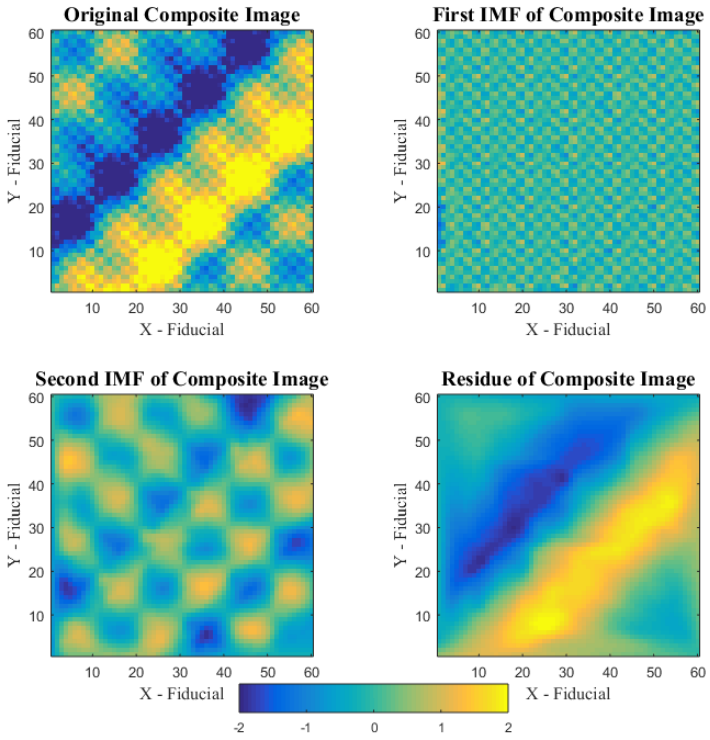

Figure 5.7: The resulting IMF surfaces generated by the BEMD algorithm: (a) the original synthetic image; (b) first IMF representing the short wavelength signal; (c) second IMF representing the medium wavelength signal; (d) the residual representing the long wavelength diagonal lines.

As was the case with the previous example, the error analysis shows little error between the textures and the corresponding IMFs. There is a bit more error associated with the diagonal lines and the residual IMF, but this caused by the edge effects of the interpolation.

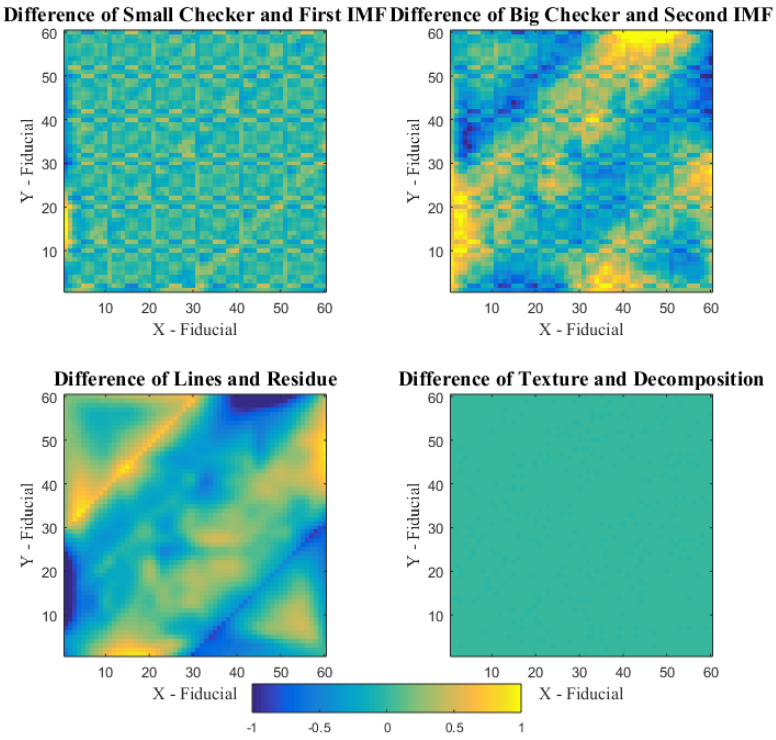

Figure 5.8: Error plot of the three original surfaces with the IMF and residual surfaces: a) error of small checkerboard and first IMF; b) error of the large checkerboard and the second IMF; c) error of diagonal lines and residual surface; d) error of the original composite surface and summation of IMF and residual surfaces.

For the final example, a simple texture has been created using a small checkerboard, a large checkerboard and a gradient fill Figure 5.9 (c). Unlike the previous examples, the gradient fill is not an IMF because it is a monotonic function.

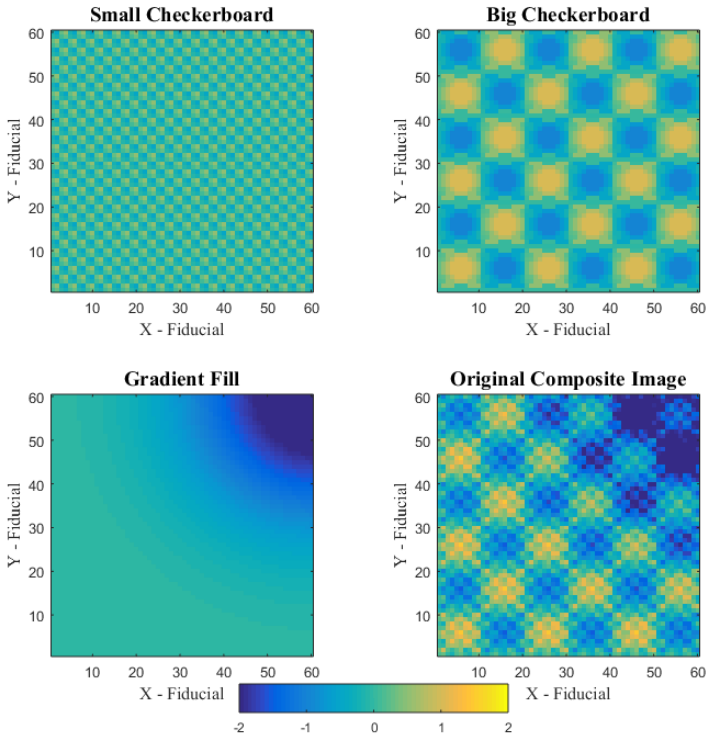

Figure 5.9: A three component synthetic texture: (a) the small wavelength checkerboard; (b) the medium wavelength checkerboard; (c) the long wavelength gradient fill; (d) the summation of the three components.

The first IMF shows the small checkerboard texture with the typical amplitudes seen in the past examples, Figure 5.10 (b). The second IMF surface has consistent amplitude that is not influenced from the gradient fill, Figure 5.10 (c). Finally, the residue surface contains some energy from the second IMF but those amplitudes are around 1% of the original amplitude, also the surface shows a decreasing function from (0, 0) to (60, 60).

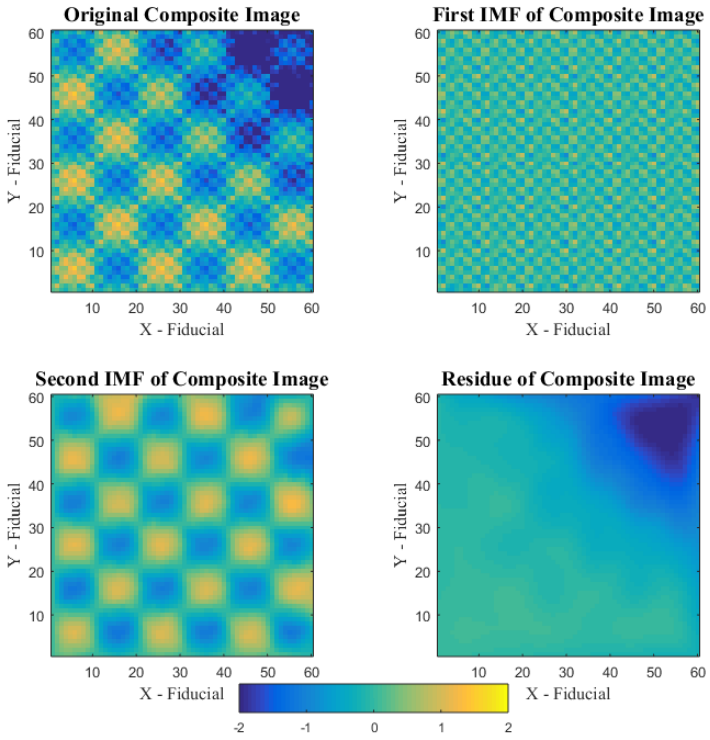

Figure 5.10: The resulting IMF surfaces generated by the BEMD algorithm: (a) the original synthetic image; (b) first IMF representing the short wavelength signal; (c) second IMF representing the medium wavelength signal; (d) the residual representing the long wavelength gradient fill.

The error surfaces show what energy is present in each of the surfaces. The error surface, Figure 5.11, created by the big checkerboard and the second IMF shows that the IMF surface is smooth and the (60, 60) corner had some of the energy removed. The error surface of the residue and the gradient fill shows that some of the big checkerboard energy was present in the residue surface. However, the big checkerboard texture that is present in the residue image is less than 20% of the original checkerboard amplitude. While some energy is present in the residue from the checkerboard pattern, the majority of the energy is also present in the second IMF. From this residual checkerboard energy, the number of IMF surfaces returned from the algorithm should be increased to four. While this increase in the number of surfaces help to ensure the residue is monotonic, it

should also be noted that a certain amount of gradient fill energy would be present in the newly created third IMF.

With energy from the big checkerboard being present in the residue, the remedy was to increase the number of IMF surfaces generated by the BEMD algorithm. However, some of the energy from the gradient fill was extracted into the newly created third IMF, so the question arose as to how many IMFs should be extracted from the data? The answer is subjective and depends on the objectives of the analysis and the type of data being analyzed. For these textures, if the concern is short wavelengths, then two IMF surfaces are enough and if the longer wavelengths are of concern then perhaps three or four IMF surfaces are needed. However, the chosen number should not be too large or the algorithm is then trying to decompose interpolation noise and the results become meaningless.

Since the composition of the image is known to contain three different textures, the BEMD algorithm was set to extract two IMFs and the residual. The resulting IMF surfaces represent the original three components with some distortion. Only two IMFs were generated because by selecting more than two the third IMF is distorted and does not resemble the gradient fill. As a matter of fact, the gradient fill is not even an IMF because it does not oscillate and is a monotonic function. Therefore, the selection of two IMFs in this instance is correct. Prior to executing the BEMD function, the surface should be evaluated to see just how many modes should be extracted or else the algorithm will continue to decompose noisy surfaces which hold no physical meaning.

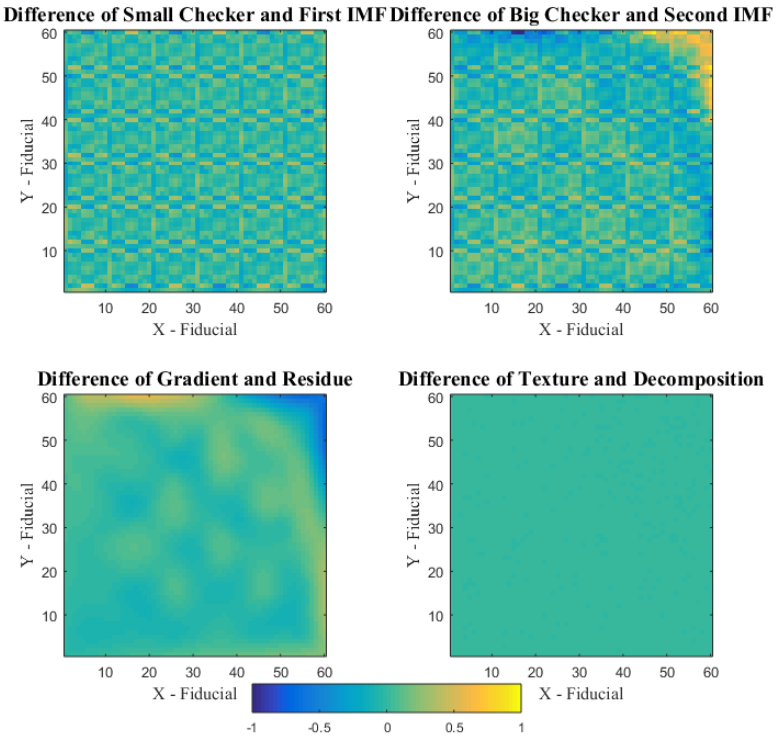

Figure 5.11: Gradient fill error analysis: a) error of small checkerboard and first IMF; b) error of the large checkerboard and the second IMF; c) error of gradient fill and residual surfaces; d) error of the original composite surface and summation of IMF and residual surfaces.

5.2 Helicopter-Borne Total Magnetic Field Data

I applied the BEMD analysis to two different helicopter-based total field magnetic data sets with the goal of removing sensor noise and separating compact ferrous anomalies from the local geology. The first example shows the ability to eliminate the short wavelength ripple noise generated in data acquisition. The second example shows how to highlight the smaller and more interesting anomalies from a complex geological background. The results showed that using BEMD to filter these data was successful after adding noise to the original data in order to remove the sensor noise into the first IMF (the IMF with highest wave number).

Figure 5.12(a) shows the original data where the sensor bias noise causes a ripple across the image in the east direction. This sensor noise is an artifact of representing the 3-D Helicopter borne data in a 2-D image. Subsequent near-overlapping passes collected at disparate altitudes will result in noise of this character. The noise is expressed intermittently with respect to its spatial distribution as the data collection altitude for near-overlapping passes converge and diverge. Figure 5.12(b) shows the data after the square wave noise was added. The first IMF contained the added noise and the sensor ripple along with some energy from the higher wave-number anomalies as seen in Figure 5.12(c). The residue image, Figure 5.12(d) contains the original data with the sensor ripple removed. A quick visual comparison of the original and residue images shows that the general shapes of the anomalies have not changed but the energy associated with the noise has been removed. The residue image clearly shows the details of the eastern side of the data where the sensor noise made the data nearly impossible to interpret. When using the residue data to perform geophysical inversion (which is not a good practice but illustrates a point), the results from the smoothed data are better than the raw data which contained the sensor noise. After performing these inversions, the fit coefficient, the correlation between the original and fit data, for the anomaly at (61, 42) increased from 46% to 96%. Anomalies, such as (9, 60) and (15, 54), had some energy removed as part of the first IMF but the resulting inversion was still favorable with a 4% increase in correlation coefficient and an overall reduction in parameter uncertainty.

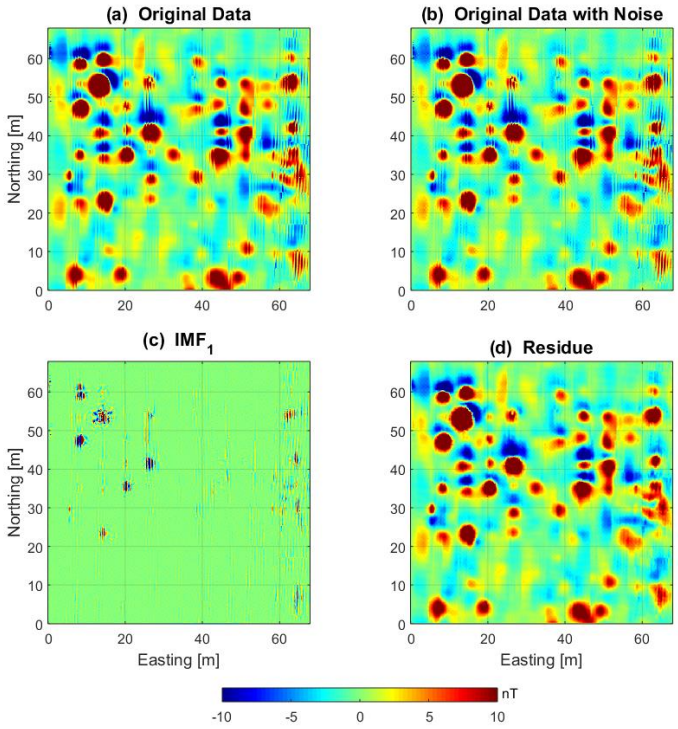

Figure 5.12: Application of BEMD to helicopter-borne total magnetic field data. a) original data; b) original data with square-wave noise added; c) IMF1 of original data and noise; d) residual surface showing noise removed.

As with the second example, BEMD was also applied as a spatial filter to total field magnetic data with a complex geologic background. Figure 5.13(a) shows the original data and the influence of geology can be clearly seen as it starts in the southeast corner and curves gently around to the northeast corner. Sensor noise is present in the data but is not consistent throughout. Square wave noise was added to the data to create a consistent surface for the BEMD algorithm to process the added noise and sensor noise from the image. Figure 5.13(b) shows the first IMF which contains the majority of the sensor noise and all the added noise. The first IMF also contains some energy from compact ferrous anomalies. This energy is associated with the short wavelength aspects of the anomalies. As BEMD continues to process the data, the second and third IMFs, Figure 5.13(c), contain the mid-range wave-number anomalies with a tiny bit of sensor noise and

geology. The sensor noise that is present in the second IMF is mostly associated with the high amplitude anomalies which changed the nature of the wavenumber as the part of the initial surface interpolation. The third IMF contains wavenumbers associated with the geology, which is why the long linear trends are seen in the summation of the second and third IMFs. Finally, the residue image, Figure 5.13(d), shows the long wavenumber anomalies associated with the geology.

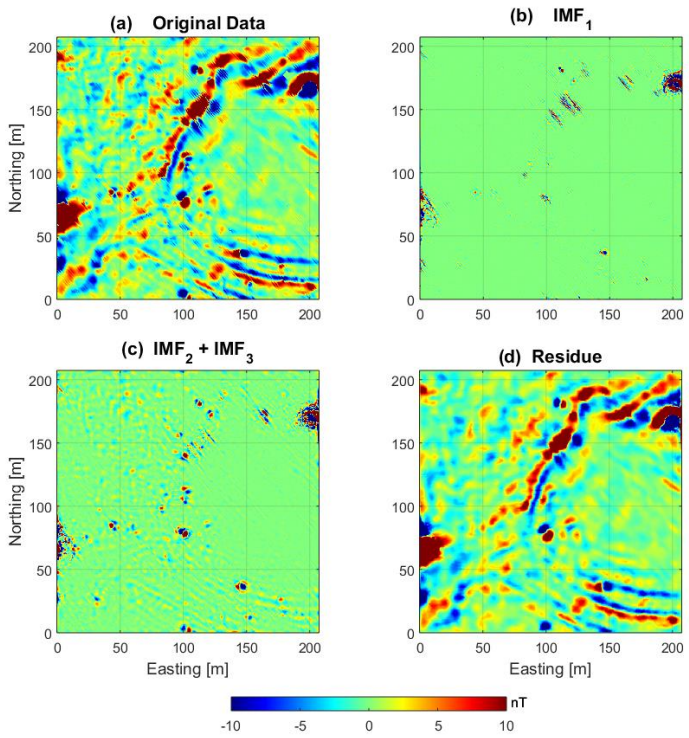

Figure 5.13: Example of BEMD data filtering. a) Original total field magnetometry data with high frequency/low amplitude noise; b) IMF1 of original data; c) summation of IMF2 and IMF3; d) the residual surface after BEMD processing.

Although many of the anomalies in the original image are obvious and would be selected as areas of interest, the selection of these anomalies became easier without the interference of the regional geology. Anomalies like those at (90, 95), (190, 20) and (95, 110) became more clear and the causes are not likely geologic. The large scale anomalies located at (0, 75) and (200, 160) exhibit complexity in the IMFs but are too large to be

considered anomalies of interest. An experienced data processor would flag the anomaly at (125, 175), but in the summation IMF it has significant energy in the mid-range wavenumbers to be obvious even to a less experienced analyst. The sensor noise at (175, 170) distorts the local geologic response and induces an anomaly, but in the first IMF it is clear that the cause of the anomaly is sensor noise. Anomaly selection should not be done using only the summation IMF, but should be used in conjunction with all images to make a proper assessment. By using all four images, a complete set of anomalies were identified.

Sensor noise is intermittently present in the image from both surveys. The goal of the addition was to provide a layer of noise similar in nature to the intermittent sensor noise for the BEMD algorithm to process. Figure 5.14a shows the two-dimensional FFT for the original geology laden image where nearly all the energy is contained below (4, 4) with harmonics extending outward. The decision to use noise with a wave length of 1.5 m is supported by the FFT image of the original data and the sensor spacing, which was 1.5 m. The two-dimensional FFT image for the first IMF shows the energy added by the noise at (4.2, 4.2), shown in Figure 5.14(b). In general, the FFT images of Figure 5.14 show that each IMF continues to remove higher wavenumbers as the number of IMFs increase. The FFT of the summation IMF shows a significant decrease of the energy in the high wavenumber range and a smearing of mid-range wavenumbers, shown in Figure 5.14(c). The residue image contains the longest wavelength anomalies which are clearly seen by Figure 5.14(d) which has high amplitudes in the lower wavenumbers.

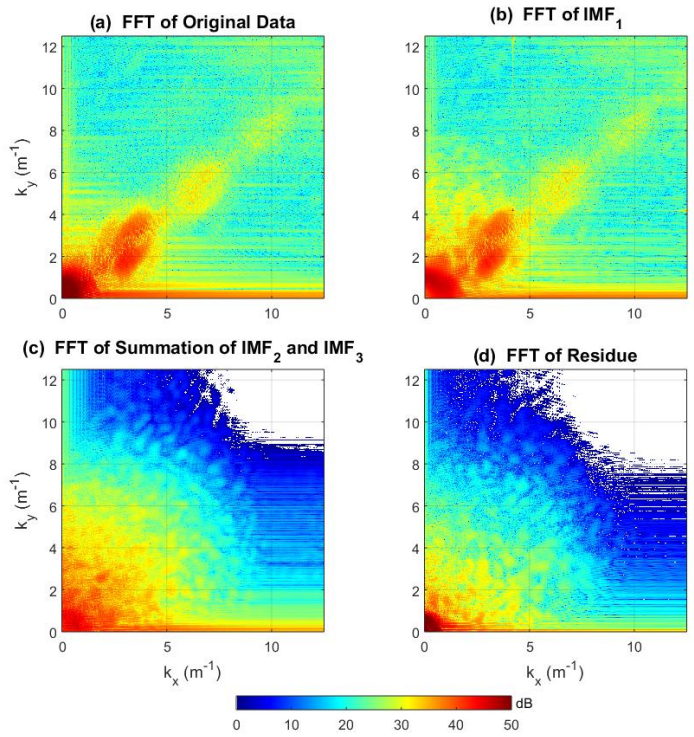

Figure 5.14: FFT analysis of original data and the IMF surfaces, a) FFT of original data; b) FFT of IMF1; c) FFT of the summation of IMF2 and IMF3; d) FFT of the residual surface.

Anomaly selection benefited the user by providing an image of the data that was not cluttered with geology. The number of anomalies chosen by an experienced data processor using either the original data or the summation surface did not drastically change; the anomalies became the prominent feature of the summation IMF making the selection process easier. A number of anomalies were obscured by the geology and might have been missed by users with less experience.

Chapter 6: Summary and Conclusions

This thesis explored the application of the Hilbert-Huang Transform (HHT) to different types of data as an advanced data processing technique. A brief description of existing data processing techniques was explored, followed by a description of the empirical mode decomposition for both one- and two-dimensional data. Empirical mode decomposition processes the data into intrinsic mode functions that break the data into signals with well-behaved Hilbert Transforms. The HHT is adaptive and sufficiently processes nonlinear and nonstationary data. Furthermore, the HHT uses the Hilbert Transform to compute the instantaneous frequencies. It is important to note that the HHT is a new technique, but has gained some traction in recent years through use in many different disciplines, as seen in this paper.

The first application of the HHT was to microtremor data collected in Beijing. Seismic sensors placed around Beijing collected ambient noise data over an hour. Microtremor data is used in engineering geophysics to explore the subsurface with a nonintrusive method. Collecting ambient noise data in a busy city with lots of frequencies is a very good example of nonlinear and nonstationary data. For Beijing, the HHT was applied to three problems commonly associated with microtremor data,

1) The elimination of local transients
2) The elimination of monochromatic noise
3) Creating a sharper peak in the H/V analysis using the Hilbert-Huang Spectrum.

Local transient noise dominated the first IMF, yet data analyzed with and without the removal of IMF1 showed that removing IMF1 greatly reduced the influence of the

transient on the STA/LTA ratio. The H/V analysis also improved as a result of removing the noisy transient signals.

EMD can be used as a bandpass filter to remove unwanted frequencies. Monochromatic noise is when a noise pollutes the signal and disrupts the analysis. By transforming the data into IMFs and inspecting the IMFs, it is easy to pick out the IMF that contains the noise. EMD break the signal down into modes that contain similar frequencies with the frequencies profile decreasing as the IMF number increases. For the data presented, IMF3 contained the majority of the monochromatic noise. Removal of this mode improved the H/V analysis, producing a clear peak in the H/V ratio. Use of the bandpass filtering has a similar affect, but unlike bandpass filtering, if the noise mysteriously stops, FFT based bandpass filtering cannot adequately process nonstationary data.

Like the time-frequency-energy plots of the Wavelet Transform, the Hilbert-Huang Spectrum (HHS) is a frequency-time-energy plot that shows the amount of energy present in each of the frequency. Implementation of the HHS to the H/V analysis had a dramatic affect by providing a clear and defined peak H/V ratio. Whereas the FFT analysis showed a broader peak with a higher peak frequency. Using the FFT analysis, the H/V ratio had a frequency of 1.6 Hz resulting in a depth of 52 m. The HHS peak frequency was 1.2 Hz resulting in a depth of 82 m. The HHS result proved more consistent with a local well.

EMD processing was applied to magnetotelluric data collected at the Esashi observatory in Japan. Excluded IMFs from the data improved the analysis by removing the lower frequencies that saturated the frequency spectrum. The Schumann Resonances clearly showed the fundamental mode and modes up to 33 Hz. Processing a year's worth of data into datasets containing the four seasons, the seasonal variations were clearly seen. As the

storms moved from the southern hemisphere to the northern hemisphere the magnetic components recorded different contributions from all the storms centers. Daily variations showed the frequency changes from changing geometry of the ionosphere.

In data from March 6th, five days before the Tohoku earthquake on March11th, there is an anomalous signal. Whether or not this signal is an earthquake precursor has yet to be fully explored and requires more research. Perhaps by comparing data from other MT stations in Japan or by comparing the MT data surrounding other major earthquake events, may show similar anomalous signals around the same time as the one observed at Esashi.

A texture analysis showed how the BEMD method works and indicated shortcomings when the bandwidth of the image is not large enough. Adding noise to the image improved the separation between modes, as long as the added noise was consistent throughout the image and spatial frequency was at least double the bandwidth of the first mode of interest. Adding noise to the data increases the overall bandwidth of the data and prevents mode mixing. Mode mixing occurs when there is not sufficient bandwidth to support the correct separation of modes.

The BEMD method was also applied to helicopter-borne total field magnetometry data as a way to correct sensor overlap and to eliminate the influence of local geology. Inversion results showed an increase in the overall fit of anomalies processed with the BEMD method. Compact ferrous dipoles have a high spatial frequency where dipoles created by geology are broad in nature. The texture analysis separated the geology and the items of interest.

Generally speaking, the use of EMD for advanced signal processing produced very good results. Compared to FFT, the frequency and time resolution of HHT is superior because of the computation of the instantaneous frequency from the intrinsic mode functions. However, there is a significant loss in performance as the EMD computation is intensive. A Fast HHT algorithm should make the HHT a more attractive method for data analysis. The BEMD algorithm is extremely slow when the images are large. If time is not important but results are, the Hilbert-Huang Transform provides excellent data quality with noise reduction.

REFERENCES

Arai, H., and K. Tokimatsu (2004), S-Wave Profiling by Inversion of Microtremor H/V Spectrum, *Bull. Seism. Soc. Am*, 94, 53-63.

Asten, M.W., W.R. Stephenson, and P. Davenport (2005), Shear-wave velocity profile for Holocene sediments measured from microtremor array studies, SCPT, and seismic refraction, *Journal of Engineering and Environmental Geophysics*, 10, 235–242.

Atakan, K., Duval, A.M., Theodulidis, N., Guillier, B., Chatelain, J.-L., Bard, P.Y., and the SESAME-Team (2004), The H/V spectral ratio technique: experimental conditions, data processing and empirical reliability assessment, paper presented at the *Proceedings of the 13th World Conference on Earthquake Engineering*, Vancouver, B.C., Canada, Paper No. 2268.

Bard, P. Y., and Sesame Participants (2004), The SESAME project: an overview and main results, paper presented at the *Proceedings of the 13th World Conference on Earthquake Engineering*, Vancouver, Paper 2207.

Bilitza, D. (2001), International Reference Ionosphere 2000, *Radio Science*, Volume 36 Number 2, Pages 261 – 275.

Bodin, P., and S. Horton (1999), Broadband microtremor observation of basin resonance in the Mississippi embayment, Central US, *Geophysical Research Letters*, 26, 903-906.

Bracewell, R. N. (2000), *The Fourier Transform and Its Applications (3rd ed.)*, Boston, MA, McGraw-Hill, ISBN 0-07-116043-4.

Cara, F., Giulio, G., and Rovelli, A. (2003), A Study on Seismic Noise Variations at Colfiorito, Central Italy: Implications for the Use of H/V Spectral Ratios, *Geophysical Research Letters.*, 30, doi:10.1029/2003GL017807.

Chen, Y., L. Liu, R. Mehl (2007)Towards tomographic imaging with microtremor arrays: the application of Hilbert-Huang Transform, *Eos Trans.* AGU, 88(52), Fall Meet. Suppl., Abstract S43B-1312, 2007.

Chen, Q., L. Liu, W. Wang, and E. Rohrbach (2008), Site effects on earthquake ground motion based on microtremor measurements for metropolitan Beijing, *Chinese Science Bulletin*, in press.

Cooley, J., and J. Tukey (1965), An algorithm for the machines calculation of complex Fourier series, *Mathematics of Computation*, 19, 297-301.

Damerval, C., S. Meignen, and V. Perrier (2005), A fast algorithm for bidimensional EMD, *IEEE Signal Processing Letters*, 12, 701–704.

Galejs, J. (1965), Schumann Resonances, *Radio Science*, 69D(8), 1043 – 1055.

Huang, N. E. (2001), Computer implemented empircial mode decomposition apparatus, method and article of manufacture for two-dimensional signals, Patent 6311130, U.S. Patent and Trademark Office, Washington, D. C.

Huang, N. E., Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, and H. Liu (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceeding of the Royal Society of London A, 454*, 903–995.

Huang, N. E., and Z. Wu (2008), A Review on Hilbert-Huang Transform Method and its Applications to Geophysical Studies, *Reviews Of Geophysics*, 49.

Johnson, F. S 1965, Satellite Environment Handbook, Stanford Univerisity Press, Stanford, CA.

Linderhed, A. (2004), Adaptive image compression with wavelet pakets and empirical mode decomposition, Ph.D. dissertation, Dept. of Electrical Engineering, Linköping Univ., Linköping, Sweden.

Liu, L., E. Roeloffs, and X. Zheng (1989), Seismically induced water level fluctuations in the Wali well, Beijing, China: *Journal of Geophysical Research*, 94, 9453-9462.

Liu, L., Z. Liu, and B. Barrowes (2011), Through-wall bio-radiolocation with UWB impulse radar: observation, simulation and signal extraction, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 791-798.

Liu, Z., L. Liu, and B. Barrowes (2010), The application of the Hilbert-Huang transform in through-wall life detection with UWB impulse radar, *PIERS online*, 6, 695-699.

Liu, L., R. Mehl, and Q. Chen (2008), Applications of the Hilbert-Huang transform to microtremor data analysis enhancement, paper presented at the *Proceedings of the 14th World Conference on Earthquake Engineering*, Beijing, China.

Liu, L., R. Mehl, W. Wang, and Q. Chen (2015), Applications of the Hilbert-Huang transform for microtremor data analysis enhancement, *Journal of Earth Science*, 26(6), 799–806.

Liu, L., K. He, R. Mehl, W. Wang, and Q. Chen (2008), Site effect assessment of earthquake ground motion based on advanced data processing of microtremor array measurements, oral presentation at the American Geophysical Union (AGU) Annual Meeting, San Francisco, California.

Liu, L., R. Mehl, and Y. Chen (2008), Tomographic imaging with microtremor arrays: the application of Hilbert-Huang transform, *Seismological Research Letters*, 79(2), Seismological Society of America annual meeting, SSA08 Abstract: 08-291 (invited), 2008.

Mano, A. R. Rawat, A. K. Sinha, S. Gurubaran, K. Jeeva (2015), Schumann resonances observed at Maitri, Anatarctica: diurnal veriation and its interpretation in terms of global thuderstorm activity, *Current Science*, 109(4), 784 - 790.

Mehl, R., J. Foley, and L. Liu (2004), Improving UXO characterization accuracy by accounting for sensor position and orientation, paper preseneted at the *Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems*, Colorado Springs, Colorado.

Mehl, R., and L. Liu, Applications of bidimensional empirical mode decomposition for spatial filtering of total field magnetic data, *IEEE Geoscience and Remote Sensing Letters*, (in preparation).

Nunes, J. C, Y. Bouaoune, E. Delechelle, O. Niang, and P. Bunel (2003a), Image analysis by bidimensional empirical mode decomposition, *Image and Vision Computing*, 21(12), pp. 1019 – 1026.

Nunes, J. C, O. Niang, Y. Bouaoune, E. Delechelle, and P. Bunel (2003b), Bidimensional empirical mode decomposition modified for texture analysis, *Proceedings Lecture Notes Computer Science*, 2749, 295 – 296.

Ortega, P., A. Guha, E. Williams, G. Satori (2014), Schumann Resonance Obersvations from the Central Pacific Ocean, paper presented at the *XV International Conference on Atmospheric Electricity*, Norman, Oklahoma.

Parolai, S., P. Bormann, and C. Milkereit (2002), New relationships between Vs, thickness of sediments, and resonance frequency calculated by the H/V ratio of seismic noise for the Cologne area (Germany), *Bulletin of the Seismological Society of America*, 92, 2521–2527.

Report of the Defense Science Board Task Force on Unexploded Ordnance (2003), United States Department of Defense.

Sinha, A. K., B. Pathan, S. Gurubaran, C. Selvaraj, K. Jeeva (2008), Characteristics of Schumann resonance as observed in the Indian sector, paper presented at the *DST Workshop on Electrodynmaical Coupling of Atmospheric Regions*.

Telford, W. M., L. P. Geldart, R. E. Sheriff (1990), Applied Geophysics, Cambridge University Press, Cambridge, United Kingdom.

Toledo-Redondo, S., A. Salinas, J. Portí, J. A. Morente, J. Fornieles, A. Méndez, J. Galindo-Zaldíver, A. Pedrera, A. Ruiz-Constán, F. Anahnah (2010), Study of Schumann resonances based on magnetotelluric records from whe western Mediterranean and Antarctica. *Journal of Geophysical Research*, 115, D22114, doi:10.1029/2010JD014316.

Wang, W., L. Liu, Q. Chen, and E. Rohrbach (2008), Site responses and sediment structures along the NW and NE profiles in Beijing area revealed by microtremor H/V spectral ratio studies, *Proceedings of the 14th World Conference on Earthquake Engineering*, Beijing, China, Paper No. 09-01-0059.

Xu, G., P. Han, Q. Huang, K. Hattori, F. Febriani, H. Yamaguchi (2013), Anomalous behaviors of geomagnetic diurnal variations prior to the 2011 off the Pacific coast of Tohoku earthquake (Mw9.0), *Journal of Asian Earth Sciences*, 77.

Xu, Y., B. Liu, J. Liu, and S. Riemenschneider (2006), Two-dimensional empirical mode decomposition by finite elements, *Proceedings of the Royal Society of London A.,* 462, 3081–3096.

Zhou, H., Z. Zhou, X. Qiao, H. Yu (2013), Anomalous phenomena in Schumann resonance band observed in China before the 2011 magnitude 9.0 Tohoku-Oki earthquake in Japan, *Journal Geophysical Research Atmosphere*, 118, 13,338 – 13,345.