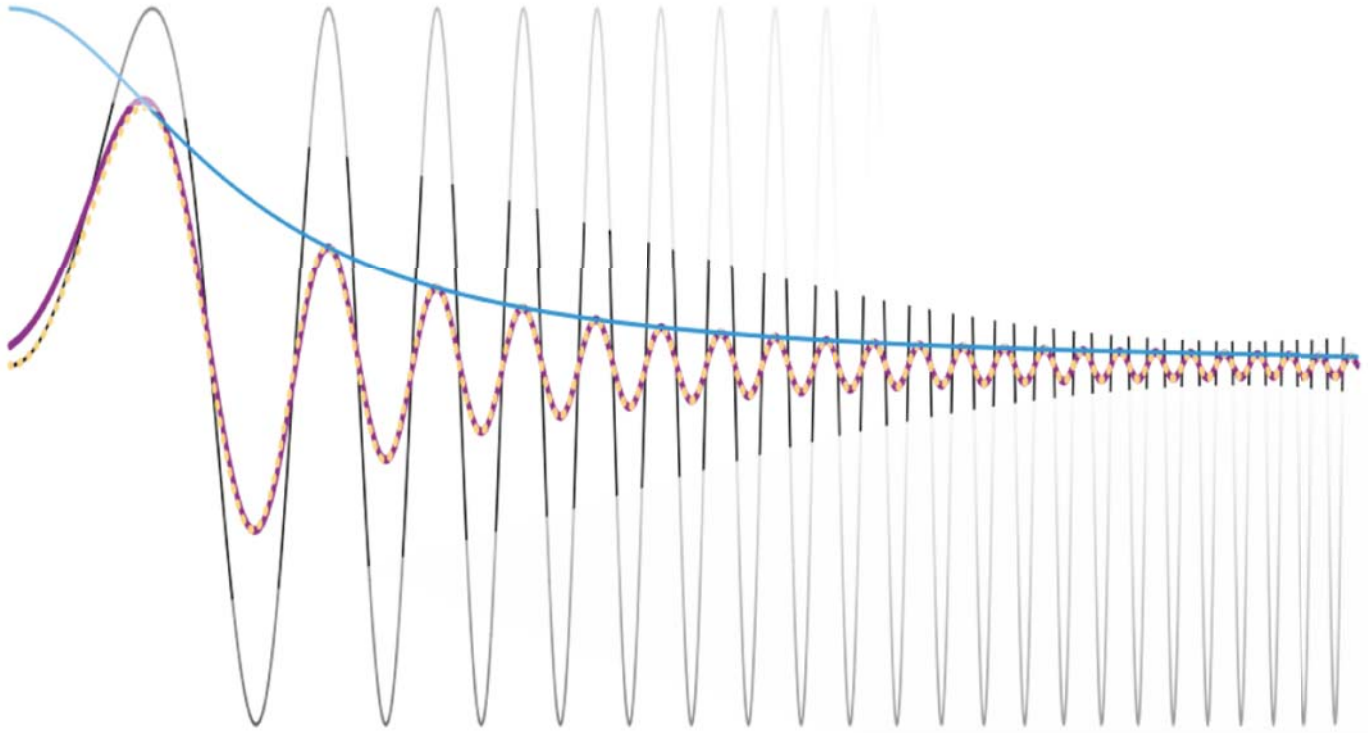
Non Stationary Magnetotelluric Data Processing

Maik Neukirch

Non Stationary Magnetotelluric Data Processing

Maik Neukirch



Non Stationary Magnetotelluric Data Processing

A thesis submitted to

University of Barcelona

in partial fulfilment for the award of the degree of

Doctor of Sciences

presented by

Maik Neukirch

supervised by

Xavier Garcia

Institut del Cienciès de Mar, CSIC, Barcelona, Spain

tutored by

Juanjo Ledo

University of Barcelona, Spain

examined by

Rob Evans

Woods Hole Oceanographic Institution, MA, USA

Marion Jegen

GEOMAR - Helmholtz Centre for Oceanographic Research, Kiel, Germany

Pilar Queralt

University of Barcelona, Spain

23rd April 2014

Abstract

Studies have proven that the desired signal for Magnetotellurics (MT) in the electromagnetic (EM) field can be regarded as 'quasi stationary' (i.e. sufficiently stationary to apply a windowed Fourier transform). However, measured time series often contain environmental noise. Hence, they may not fulfill the stationarity requirement for the application of the Fourier Transform (FT) and therefore may lead to false or unreliable results under methods that rely on the FT. In light of paucity of algorithms of MT data processing in the presence of non stationary noise, it is the goal of this thesis to elaborate a robust, non stationary algorithm, which can compete with sophisticated, state-of-the-art algorithms in terms of accuracy and precision. In addition, I proof mathematically the algorithm's viability and validate its superiority to other codes processing non stationary, synthetic and real MT data.

Non stationary EM data may affect the computation of Fourier spectra in unforeseeable manners and consequently, the traditional estimation of the MT transfer functions (TF). The TF estimation scheme developed in this work is based on an emerging nonlinear, non stationary time series analysis tool, called Empirical Mode Decomposition (EMD). EMD decomposes time series into Intrinsic Mode Functions (IMF) in the time-frequency domain, which can be represented by the instantaneous parameters amplitude, phase and frequency. In the first part of my thesis, I show that time slices of well defined IMFs equal time slices of Fourier Series, where the instantaneous parameters of the IMF define amplitude and phase of the Fourier Series parameters. Based on these findings I formulate the theorem that non stationary convolution of an IMF with a general time domain response function translates into a multiplication of the IMF with the respective spectral domain response function, which is explicitly permitted to vary over time.

Further, I employ real world MT data to illustrate that a de-trended signal's IMFs can be convolved independently and then be used for further time-frequency analysis as done for MT processing. Lastly, a discussion is opened on parallels between the Hilbert-Huang Transform (HHT, the conversion from time series to instantaneous parameters via EMD) and the Fourier Transform with respect to the time-frequency domain.

In the second part of my thesis, I apply the newly formulated theorem to the MT method. The MT method analyses the correlation between the electric and magnetic field due to the conductivity structure of the subsurface. For sufficiently low frequencies (i.e. when the EM field interacts diffusively), the conductive body of the Earth acts as an inductive system response, which convolves with magnetic field variations and results in electric field variations. The frequency representation of this system response is commonly referred to as MT TF and its estimation from measured electric and magnetic time series is summarized as MT processing. The main contribution in this thesis is the design of the MT TF estimation algorithm based on EMD. In contrast to previous works that employ EMD for MT data processing, I (i) point out the advantages of a multivariate decomposition, (ii) highlight the possibility to use instantaneous parameters, and (iii) define the homogenization of frequency discrepancies between data channels. In addition, my algorithm estimates the transfer functions using robust statistical methods such as (i) robust principal component analysis and (ii) iteratively re-weighted least squares regression with a Huber weight function. The code can be used with and without aid of any number of available remote reference stations. Finally, TF uncertainties are estimated by iterating the complete robust regression, including the robust weight computation, by means of a bootstrap routine. The code further contains a routine that calculates the transfer function from noise caused by instrument motion for Marine MT studies and removes this noise accordingly.

The proposed methodology is applied to synthetic and real data (from southern Africa) with and without non stationary character and the results are compared with other processing techniques. I conclude that non stationary noise can heavily affect Fourier based MT data processing but the presented non stationary approach is nonetheless able to extract the impedances correctly even when the other methods fail.

Zusammenfassung

Studien zeigen, dass die für Magnetotelurische (MT) Messungen wichtigen elektromagnetischen (EM) Quellen als quasistationär angesehen werden können, sodass eine Fourier Transformation mit geeigneter Fensterfunktion angewendet werden kann. Die gemessenen Zeitreihen enthalten jedoch nebst dem Signal auch Rauschen. Dieses erfüllt nicht notwendigerweise die zwingende Bedingung der Stationarität für die Fourieranalyse und kann daher zu falschen oder unzuverlässigen Ergebnissen führen. Existierende Lösungen für das Verarbeiten von MT Messreihen mit nichtstationärem Verhalten sind unzureichend entwickelt, da das Signal selbst als stationär angenommen werden kann. In dieser Doktorarbeit wurde ein statistisch robustes, nichtstationäres Verfahren entwickelt, welches mit hoch entwickelten, renommierten Algorithmen, die auf die Fourieranalyse aufbauen, in Zuverlässigkeit und Präzision vergleichbar ist. Zusätzlich wird die mathematische Grundlage für das verwendete Verfahren bewiesen und dessen Überlegenheit im Vergleich zu alternativen Algorithmen anhand echter und synthetischer MT Messreihen, mit und ohne nichtstationären Eigenschaften, bestätigt.

Nichtstationäre EM Messreihen können die Ermittlung ihres Fourier Spektrums auf unvorhersehbare Weise beeinflussen und damit auch die Abschhätzung der MT Übertragungsfunktion (TF für Englisch: Transfer Function) zwischen elektrischen und magnetischen Feldern. Diese Arbeit entwickelt ein Verfahren zur Abschätzung der MT TF basierend auf einem neuen, nichtlinearen und nichtstationären Algorithmus, bekannt als Empirical Mode Dekomposition (EMD, zu Deutsch: Aufspaltung in empirisch ermittelte Modi). Die EMD spaltet die Zeitreihen im Zeit- und Frequenzraum in Funktionen auf (IMF für Englisch: Intrinsic Mode Function), welche über Parameter, wie Amplitude, Phase und Frequenz, dargestellt werden können. Diese sogenannten Momentanparameter sind zeitabhängig. Zu Beginn der

Arbeit wird mathematisch gezeigt, dass IMFs zu jedem Zeitpunkt auch als Fourierreihe betrachtet werden können, wobei die Momentanparameter der IMF die Parameter Amplitude und Phase der Fourierreihe bestimmen. Dies führt zu der Formulierung eines neuen Theorems, welches beschreibt, dass nichtstationäre Faltung einer IMF mit einer allgemeinen Übertragungsfunktion im Zeitraum gleich der Multiplikation der IMF mit der Übertragungsfunktion im Frequenzraum ist. Dabei kann die Übertragungsfunktion im Frequenzraum auch seitabängig sein. Das Theorem wird anschliessend auf reale MT Messreihen angewandt um zu zeigen, dass dies selbst für Zeitreihen gilt, welche in IMFs aufgespalten werden können. Abschliessend wird diskutiert inwieweit man vor dem Hintergrund des Theorems Parallelen zwischen der Hilbert-Huang und der Fourier Transformation ziehen kann.

Das MT Messverfahren analysiert die Korrelation zwischen den elektrischen und magnetischen Feldern, die aufgrund unterirdischer Strukturen der elektrischen Leitfähigkeit auftritt. Sobald sich das EM Feld diffus ausbreitet, wirkt, im Bereich ausreichend kleiner Frequenzen, die leitfähige Erde als induktive Übertragungsfunktion. Diese faltet sich mit den zeitlichen Variationen des magnetischem Feldes und erzeugt dadurch messbare Variationen im elektrischen Feld. Im Frequenzraum ist diese Funktion als MT TF bekannt und ihre Abschätzung aus magnetischen und elektrischen Messreihen wird als Verarbeitung von MT Daten bezeichnet. Der Hauptbeitrag der vorliegenden Arbeit ist der Entwurf eines Verfahrens zur Abschätzung der MT TF basierend auf den EMD Algorithmus. In Gegensatz zu vorhergehenden Arbeiten zu diesem Thema, werden die Vorteile einer multivariaten Variante von EMD aufgezeigt und die Möglichkeit hervorgehoben die Momentanparameter direkt zu verwenden. Weiterhin wird einen Frequenzmittelwert für den Fall ermittelt, dass sich die Momentantfrequenzen zwischen Feldkomponenten unterscheiden. Das vorgeschlagene Verfahren basiert auf robusten statistischen Methoden wie robuste Hauptkomponentenanalyse und die schrittweise, gewichtete Methode der kleinsten Quadrate mit einer Huber Gewichtsfunktion. Durch die Hauptkomponentenanalyse kann das Verfahren Daten anderer Stationen zur Rauschminderung nutzen. Abschliessend wird die Zuverlässigkeit des Verfahrens mittels Bootstrapping abgeschätzt. Zusätzlich kann das Verfahren auch Rauschen eliminieren, welches durch die Bewegung des Messinstrumentes hervorgerufen wurde,

falls die Winkelbewegungen verfügbar sind. Dies ist insbesondere für marine Messungen oftmals hilfreich.

Synthetische und gemessene Daten (aus dem südlichen Afrika), welche sowohl stationäre als auch nichtstationäre Eigenschaften aufweisen, werden mit dem vorgeschlagenem Verfahren verarbeitet und die Ergebnisse mit alternativen Algorithmen verglichen. Anhand der Beispiele wird gezeigt, dass nichtstationäres Rauschen die Verarbeitung von MT Messreihen stark behindern kann, wenn Methoden verwendet werden, welche auf die Fourieranalyse aufbauen. Die Verwendung des vorgeschlagene Verfahrens führt hingegen zu besseren Ergebnissen.

Contents

List of Figures

List of Acronyms

MT Magnetotelluric . . . 1

QC Quality Control . . . 3

AMT Audio-Magnetotelluric . . . 3

EM Electromagnetic . . . 4

FT Fourier Transform . . . 11

HHT Hilbert-Huang Transform . . . 14

EMD Empirical Mode Decomposition . . . 15

IMF Intrinsic Mode Function . . . 15

AM Amplitude modulation . . . 16

PM Phase modulation . . . 16

PCA Principal Component Analysis . . . 18

TF Transfer function . . . 19

EDI Elecrical Data Interchange . . . 22

RHS Right hand side . . . 35

LHS Left hand side . . . 36

EMT Empirical mode decomposition based Magnetotelluric Data Processing . . . 59

MEMD Multivariate Empirical Mode Decomposition . . . 59

IF Instantaneous frequency . . . 62

IA Instantaneous amplitude . . . 67

IP Instantaneous phase . . . 67

IID Identically and independently distributed . . . 69

CHAPTER 1

Introduction

> "We are either progressing or retrograding all the while.
> There is no such thing as remaining stationary in life."

James Freeman Clarke (1810–1888), an american theologian, captured nicely what is true not only for us but nature in general. Whatever we observe in nature, we notice that it is at a continuous change, be it very fast or very slow. Similarly varies the rate of change itself and its rate in turn. Carrying this reasoning into infinity concludes that truly no process remains constant forever and everything always changes. A remarkably poetic representation of this philosophy has been created by Numhauser (1982).

1.1 The Magnetotelluric (MT) Method

The MT method is a geophysical exploration tool in which the variations of electric and magnetic fields are measured in order to investigate the in-depth conductivity structure of the subsurface. The method was first introduced in the 1950's and, since then, enjoyed ever-rising academic and commercial attention. Electromagnetic fields are predominantly affected by conductivity structure and thus, complement well other geophysical exploration methods like seismic or gravity which provide mechanical properties. Applications of the MT method have a wide range, because its investigation depth is solely limited by the frequency range of the observed signal and even though it exhibits reduced spatial resolution with increasing depth (in contrast to e.g. seismic) its complement value is highly rated.

1.1.1 Origin of the MT Method

Schröder and Wiederkehr (2000) review the early historic research on electromagnetic variations in the Earth, which is mostly based on the works of Carl Friedrich Gauss (1838) and James Clerk Maxwell (1873). Gauss was first in interpreting mathematically the effects of terrestrial magnetism, where Maxwell laid out the theoretical relation between electricity and magnetism. These findings fueled the desire to better understand the electromagnetic system Earth and promptly developed into the description of the atmosphere as an 'electromagnetic machine' where, through changes in the terrestrial magnetism, induced electric currents 'would have an effect on conductive layers in the Earth' (Stewart (1880) as cited in Schröder and Wiederkehr, 2000, page 330, section 7). Stewart (1877) recognized the relation between diurnal cycles and geomagnetic phenomena originating from the atmosphere and later, Schuster (1886) supported these ideas qualitatively. The electromagnetic field inside a conductive Earth has been investigated by Lamb (1883) (homogenous sphere and periodic fields), Price (1930) (aperiodic fields) and Lahiri and Price (1939) (conductivity as function of distance to center-point). From these efforts emerged the technique of MT (proposed independently by Rikitake (1948); Tikhonov (1950) and Cagniard (1953)), which defines the electric impedance of a location on the surface of a conductive body as the spectral transfer function between the electric and magnetic field observed at the same location.

1.1.2 History of MT Data Processing

After Cagniard (1953) laid out the theoretical basis for magnetotelluric exploration, Sims and Bostick Jr (1969); Sims et al. (1971) provided the solution for the impedance tensor from MT measurements in a mean-square sense but noted that the presence of noise could induce bias on their estimates. Gamble et al. (1979) showed that inclusion of remote measurements can completely remove this bias if the noise between station and remote is not correlated, a method adopted from econometrics and first proposed by Reiersøl (1941). Weidelt (1972) lists invariant constraints that the solution must obey and proposes to reject data that does not fulfil these fundamentals and thus, opens the field for data quality control in MT data

processing. Reddy and Rankin (1974) extends this idea and proposes that statistical coherences inherently control data quality and only proceed processing with concentration on the seemingly better parts. Finally, Egbert and Booker (1986); Chave et al. (1987) and Jones et al. (1989) demonstrate the supremacy of robust statistical procedures over ordinary least-squares and Egbert (1997) extend the remote reference method to an unlimited number of stations by selecting the two most energetic principal components of all available remote stations as reference. From there on, efforts have been made to reduce the effect of correlated noise in the final estimations, but the proposed solutions are generally data set dependent and cannot be applied without intensive knowledge about the to-be-processed data set.

A concise summary of further data Quality Control (QC) measures is given by Weckmann et al. (2005) who propose a general workflow to streamline manual editing (noise removal) in time and spectral domain. Since manual editing of large amounts of MT time series and spectra is common practice but impractical, Manoj and Nagarajan (2003) propose to teach artificial neural networks.

Oettinger et al. (2001) concentrate on noise removal by estimating a magnetic transfer function between a noisy site and a noise free site in order to use the residual magnetic field (that cannot be explained by the remote reference) as additional regressor to a second quiet remote reference site. This procedure reduces the variance of the estimated MT transfer functions if the remote reference sites do not contain correlated noise. However, the idea is conceptually similar to the multivariate approach proposed by Egbert (1997).

Garcia and Jones (2002) study Audio-Magnetotelluric (AMT) sources and conclude that the signal in the AMT dead band (1 to 5 kHz) is drastically reduced during daytime and therefore, advocate to concentrate measurements for the AMT method during nighttime in high latitudes. Posteriorly, Garcia and Jones (2005), propose the use of a new methodology for high latitudes in which a remote site is measuring continuously MT data, but a telluric survey is carried out at daytime. Through the calculation of the telluric-telluric transfer functions, they obtain the responses at each site.

(a) *Low frequency MT source.* **(b)** *High frequency MT source.*

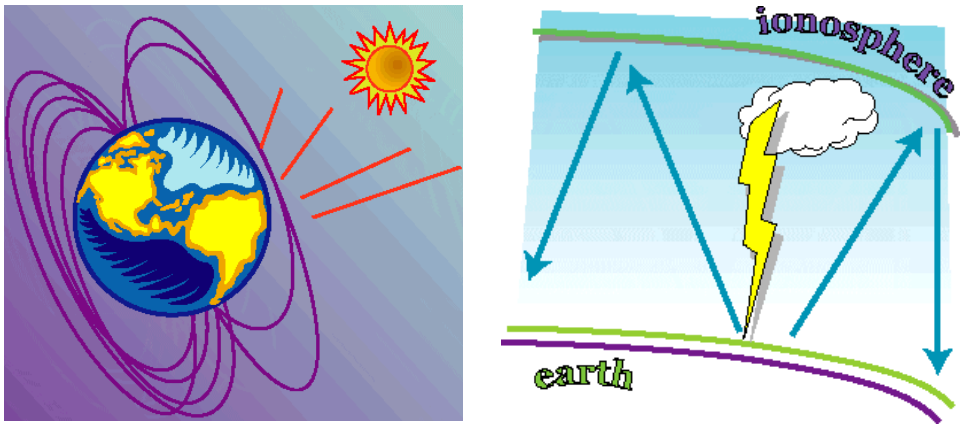


Figure 1.1: *MT signal source types are displayed, (a): magnetospheric fluctuations by solar wind constitute to signals of frequencies* $< 8\,\mathrm{Hz}$ *and (b): energy of frequencies* $> 8\,\mathrm{Hz}$ *stem from lightning activity channeled between ionosphere and Earth. Illustrations are taken from Christopherson (1998).*

1.1.3 Magnetotelluric Sources and Noise

Natural Electromagnetic (EM) field variations are caused by two major working mechanisms: lightning activity at high frequencies ($> 8\,\mathrm{Hz}$) and magnetospheric currents excited by solar wind at low frequencies ($< 8\,\mathrm{Hz}$) (Garcia and Jones, 2002; Christopherson, 1998, cp. Figure 1.1). Rakov and Uman (2007) summarize the electromagnetic lightning discharge to three modes: (a) fast and transient leader-return-stroke sequences, (b) slow and quasi stationary continuing currents and (c) perturbations and surges on the continuing currents. The longest lasting and most abundant in an electromagnetic time series measurements are the perturbed continuing currents, which may be viewed as being stationary on a section with a dynamic length due to recurrent transient strokes. Liu and Fujimoto (2011) conclude that the magnetospheric current is non linearly driven by the dynamic solar wind but behaves in a static manner for high magnetospheric pressure conditions. Both of these EM sources are naturally non stationary, since both, lightning strokes and magnetospheric pressure conditions, are very dynamic and strictly limit the duration of the stationary electromagnetic signal.

In accordance with almost all MT practitioners, Chave and Jones (2012) argue that the magnetotelluric signal is quasi stationary and therefore advo-

cate that the Fourier Transform is the preferred tool to transform data into the frequency domain, even though the Fourier Transform requires a strictly stationary signal. Nevertheless, for the quasi stationary nature of MT source signal, this procedure might be correct for processing the signal that we are after but does not necessary hold for noise sources which certainly can be non stationary (e.g. a moving train) and are usually inseparably mixed with the desired MT source in the collected MT data. So, the MT signal may behave sufficiently stationary but the contained noise clearly does not need to. If we cannot succeed in separating this non stationary noise, it might very well leak throughout the Fourier Spectra and hamper any impedance estimation. In order to diminish this effect, state-of-the-art processing techniques break the MT time series into smaller parts and perform the Fourier Transform on each part, hoping that the non stationary effects of both, signal and noise, are negligible in the major part of the data and can be sorted out by robust statistical routines.

In practice this procedure works very well for data with high signal-to-noise ratios but frequently encounters problems in the presence of strong electromagnetic noise (which would include non stationary signal). Detailed information about electromagnetic noise characteristics is reviewed and described concisely by Szarka (1988) and Junge (1996). Table 1.1 summarizes active and passive sources of noise that can be observed in MT measurements. Passive sources act typically as superficial anomalies and rather lead to wrong interpretations of the results than really introduce electromagnetic noise to the measurements, however, even though they do not add noise directly, in particular the conductive passive noise sources may reduce the signal strength considerably to the point of being undetectable. Active sources, on the other hand, add electromagnetic energy to the measurements and may completely overwhelm the natural signal. Normally, active electromagnetic noise can be assumed in the near field and thus is often recognized (although not removed). Figures 1.2a, 1.2b and 1.3 illustrate noise effects in MT measurements on three examples, figure 1.2a captures a very obvious EM noise from a passing radar sweep, figure 1.2b compares the spectra of a quiet site with the spectra of a site close to a large city, and figure 1.3 exemplifies how difficult it can be to identify hidden noise in the time series. In these examples the noise effects are known and well understood, serving

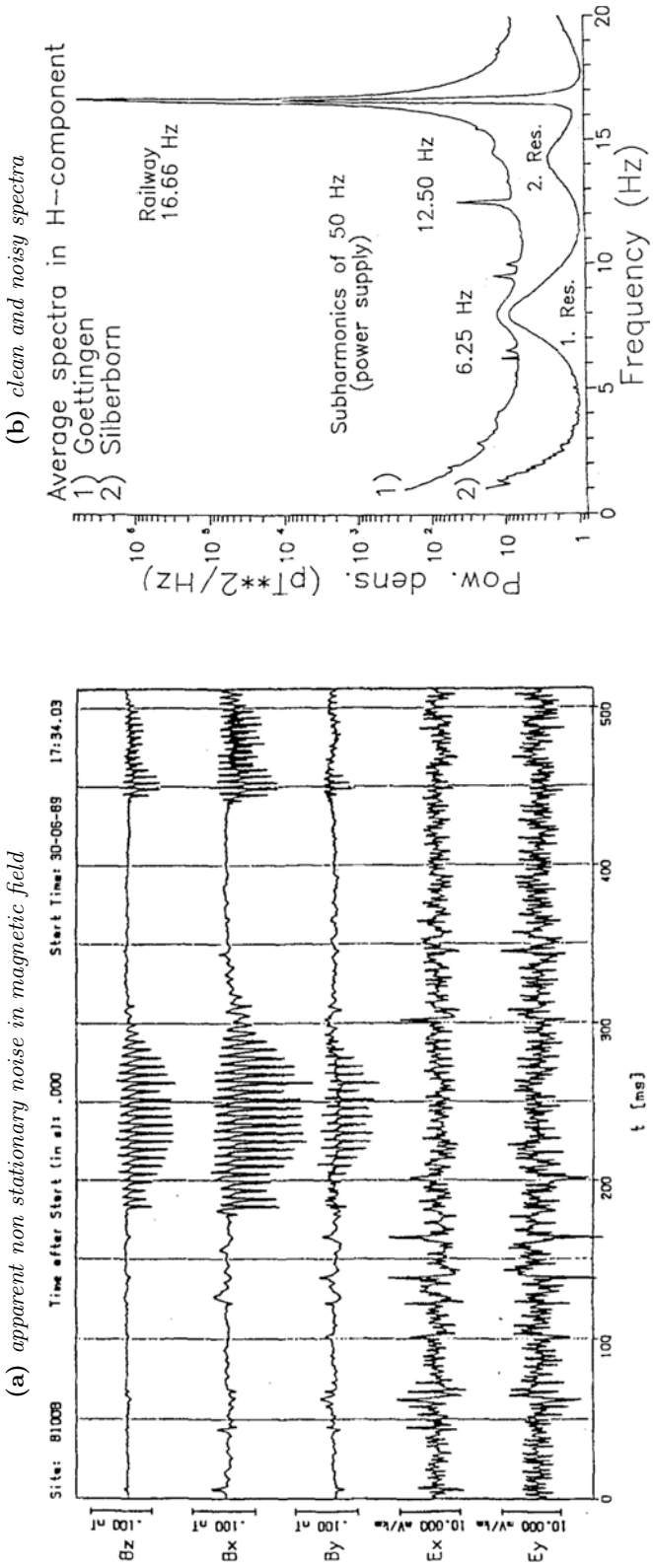


Figure 1.2: *Examples of noise effects in the MT method.* **(a)** $512\,\mathrm{ms}$ *interval of MT time series containing two radar sweeps (Brasse, 1993).* **(b)** *High resolution spectra of the* B_x *component at two sites, one near Göttingen and the other at 30 km distance (Silberborn) for the frequency range between* 1 *and* $20\,\mathrm{Hz}$ *(Füllekrug, 1994). Illustrations are taken from Junge (1996, Figures 2 and 4 respectively).*

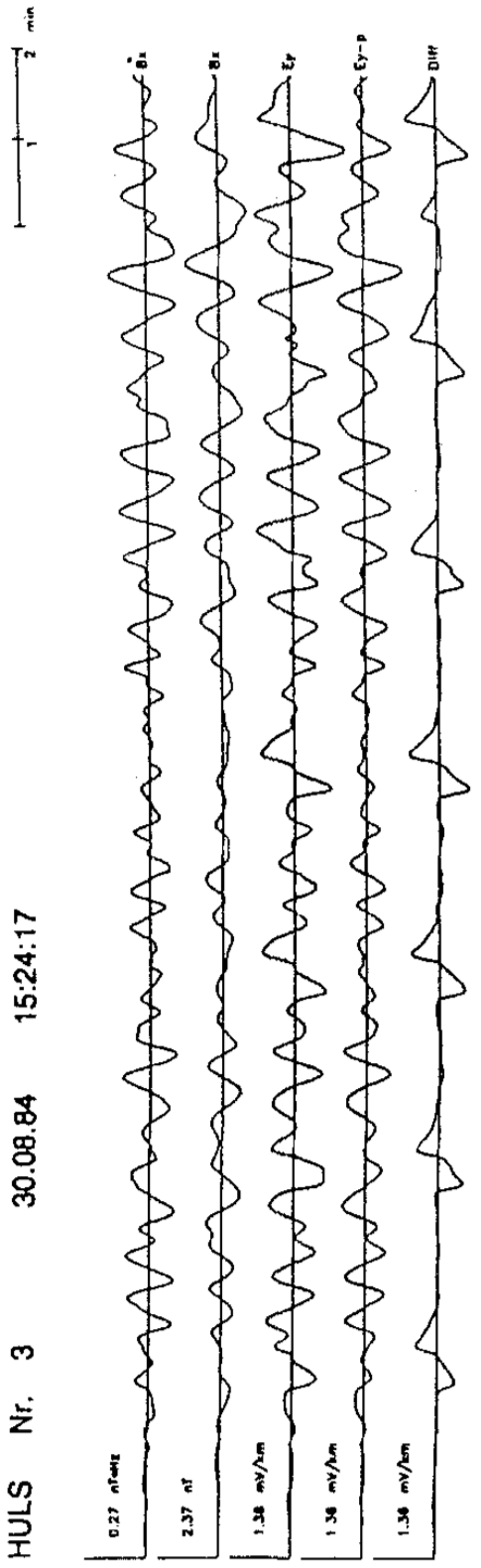


Figure 1.3: *Hidden noise in the telluric field E_y at a field site in Germany (Theissing, 1988). $\dot{B}_x$ and E_y measured field components, B_x integrated magnetic field, E_{y-p} predicted electric field, $\mathrm{Diff} = E_y - E_{y-p}$. Illustration is taken from Junge (1996, Figure 9).*

Table 1.1: *Man-made EM noise in geophysical measurements, taken from Szarka (1988).*

Passive	Active	Other local man-made effects
(affect as compact or elongated superficial resistivity inhomogeneities)	(produce regular or irregular parasitic EM field in the earth)	
• conductive: – power lines – pipelines – telephone lines – cable systems – fences, etc. • resistive: – roads – ditches, etc.	• electric power transmission lines • rectifiers • DC traffic substations • AC railway lines • arc furnaces • complicated mining electric networks • anti-corrosion systems • EM wave transmitters, etc.	• magnetic or EM effects of vehicles • EM induction due to vibrations of man-made origin • piezomagnetic anomaly due to artificial ground loading, etc.
• interaction with natural phenomena (magnetic storms, lightning, wind and ionisation of the air, etc.)		

for illustrative purposes, but it is easy to think of any more complicated form of noise (e.g., non stationarity, aperiodicity, combination of noise) that could be present in measurements.

Despite the continuos effort and progress in MT processing research, many obtained data sets are still heavily affected by noise up to the point that the interesting features are masked and the initial objectives for a survey cannot be met. When MT sites are located close to industry, agriculture or populated areas, cultural noise easily overwhelms the natural signal, therefore, for MT data processing, it is of utmost interest to identify and reduce the noise influence to improve accuracy and precision of impedance estimates and allow MT surveys in urban areas or during times of low signal strength.

1.1.4 Principles of MT Data Processing

Magnetotelluric measurements log the natural variation of magnetic and electric (telluric) fields at the Earth's surface and these measured time series can be statistically analyzed to obtain the relative spectral relation of the electrical to the magnetic field (cp. Figure 1.4 and Vozoff, 1972; Schmucker and Weidelt, 1975). The subsurface conductivity structure enforces a particular distribution of underground currents, which alter the external natural electromagnetic field of the Earth and, therefore, it allows us to derive that structural information of the subsurface conductivity from the analysis of the electromagnetic field on the surface.

The EM field propagation is described by Maxwells equations (Maxwell, 1873). Considering (1) harmonic fields with constant frequency ω, (2) the current density $\mathbf{j}$ is proportional to the electric field $\mathbf{E}$ at a constant conductivity tensor $\boldsymbol{\sigma}$ ($\mathbf{j} = \boldsymbol{\sigma} \cdot \mathbf{E}$) and (3) the field propagates only by diffusion (neglecting displacement currents and surface charges: $i\omega\epsilon\mathbf{E} \ll \boldsymbol{\sigma} \cdot \mathbf{E}$ and $q = 0$), these equations are:

$$\nabla \times \mathbf{E} = -i\omega\mathbf{B}, \tag{1.1a}$$

$$\nabla \times \mathbf{B} = \mu\left(i\omega\epsilon\mathbf{E} + \mathbf{j}\right) = i\omega\mu\epsilon\mathbf{E} + \mu\boldsymbol{\sigma} \cdot \mathbf{E} \approx \mu\boldsymbol{\sigma} \cdot \mathbf{E}, \tag{1.1b}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.1c}$$

$$\nabla \cdot \mathbf{E} = \frac{q}{\epsilon} \approx 0 \tag{1.1d}$$

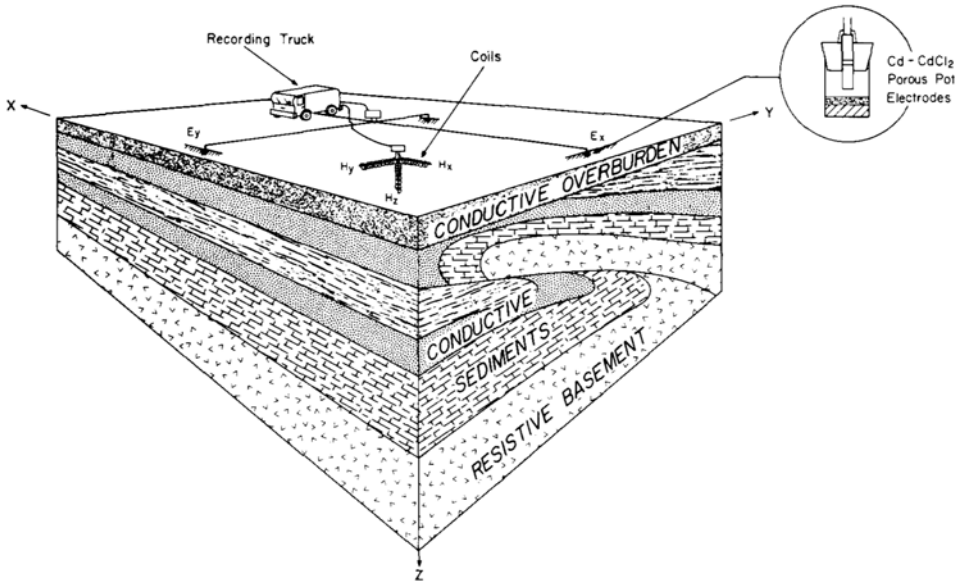


Figure 1.4: *Typical field setup for MT measurements, taken from Vozoff (1972).*

with μ describing the magnetic susceptibility as product between the magnetic susceptibility of vacuum μ_0 and the given medium μ_r. The penetration of the field $\mathbf{F} = \mathbf{E}, \mathbf{B}$ in a homogeneous body is described by applying the curl operator ($\nabla\times$) to Maxwell's equations and solve for each field $\mathbf{F}$:

$$\nabla^2\mathbf{F} = i\omega\mu\sigma\mathbf{F} = k^2\mathbf{F} \tag{1.2}$$

The term $k^2 = i\omega\mu\sigma$ is the diffusion parameter, which describes the complex penetration depth (skin depth) $\delta = \mathfrak{Re}\left(\frac{1}{k}\right)$ of the EM field (Schmucker and Weidelt, 1975) and represents the impedance $Z(\omega) = \frac{i\omega}{k}$ for a homogenous earth. Equations (1.1) for a three dimensional, anisotropic earth are solved in appendix A.1 and special cases for which the impedance tensor elements simplify are derived in appendix A.2. Equations (A.7) readily suggest a relationship between the horizontal electric and magnetic fields and justify the definition of the Tikhonov-Cagniard impedance tensor $\mathbf{Z}$ (in the following only referred to as the impedance $\mathbf{Z}$) as the following relation:

$$\begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix} = \begin{pmatrix} Z_{xx}(\omega) & Z_{xy}(\omega) \\ Z_{yx}(\omega) & Z_{yy}(\omega) \end{pmatrix} \cdot \begin{pmatrix} B_x(\omega) \\ B_y(\omega) \end{pmatrix}. \tag{1.3}$$

Mathematically, $\mathbf{Z}$ can be understood as the system response tensor of the Earth, $\mathbf{B}$ the input vector and $\mathbf{E}$ the output vector of a convolution, thus

$$\mathbf{E}(\omega) = \mathbf{Z}(\omega) \cdot \mathbf{B}(\omega) \tag{1.4}$$

or in time domain

$$\mathbf{e}(t) = \mathbf{z}(t) * \mathbf{b}(t). \tag{1.5}$$

Electric and magnetic fields are recorded as time series and need to be transformed into the frequency domain in order to solve for the impedance, because the impedance tensor is only defined in the frequency domain. Under ideal conditions, the electromagnetic field varies quasi stationary, meaning the spectral composition changes sufficiently slow that a windowed Fourier Transform (FT) can be performed, but for sites closer to inhabited or industrial areas, cultural noise often affects the measurements severely. Cultural noise can be of any kind and is most often non stationary, therefore, measurements of $\mathbf{e}(t)$ and $\mathbf{b}(t)$ are often disturbed by non stationary variations, since the physical measurements contain both, natural signal and cultural noise. However, assuming ideal measurements for the electric and magnetic fields, the system of equations in (1.3) can be solved for the impedance $\mathbf{Z}$ by any suitable optimization algorithm.

1.1.5 Importance of Processing for Interpretation

The magnetic and electrical time series acquired for a MT study cannot be interpreted in their crude form, they always have to be processed to obtain interpretable information. Even after processing the time series, the resulting impedance is still not very descriptive. It only indicates how subsurface resistivity appears for a range of frequencies. In order to obtain a real, structured resistivity model in spatial coordinates, it is necessary to find such a model by data forward modeling and inversion. It is clear, that the success of such procedures heavily depend on the input data quality. Therefore, studies on the improvement of signal-to-noise ratios, noise identification and signal filtering are necessary in order to drive the areas of MT operations closer to urban regions or to extend measurements in times of low signal strength.

1.2 Non Stationary Signal Analysis in the MT Method

In digital signal processing, time series convolution is often related to the FT and therefore implies stationary and linear assumptions on the data. The reason for this prominence lies within the convolution theorem which allows to exchange a weighted integral expression to a simple multiplication, which results in much shorter computation times (Smith, 1997). Margrave (1998) introduced the theory for non stationary convolution filters based on the Fourier Transform arguing that a continuous function is completely described by its Fourier Transform and, therefore, non stationary filtering should be possible in the frequency domain. Huang et al. (1998) show that the frequency information of non stationary signals might describe the signal entirely but gives biased information with respect to the physics of the signal, concluding that misinterpretation of the Fourier Transform cannot be ruled out in a truly non stationary setting.

1.2.1 Fourier Transform and its Limitation

While scrutinizing diffusive heat flow, Fourier (1822, translated to English: Fourier and Freeman, 1878) defined one of the most renown transformation in time series analysis to date, the Fourier Transform. The transformation relates two variables x and ξ with a Fourier integral pair such that

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi} dx = \mathcal{F}(f), \qquad \forall \xi \in \mathbb{R}, \tag{1.6a}$$

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i \xi x} d\xi = \mathcal{F}^{-1}\left(\hat{f}\right), \qquad \forall x \in \mathbb{R} \tag{1.6b}$$

and states that f can be reconstructed from $\hat{f}$ by the Fourier Transform $f = \mathcal{F}\left(\hat{f}\right)$. However, the integration limits indicate that complete (infinite) knowledge of $\hat{f}$ would be required. Since periodic functions contain complete information in a sufficiently long section, adequate for physical measurements (e.g., in time or space), periodic functions are predestined for analyses with the Fourier Transform. Often physical relations can be approximated by periodic functions and thus, qualify more or less for such analyses themselves with small, mostly known errors known as spectral leakage. These assumptions are generally known as stationary (periodic) or quasi stationary (sufficiently periodic) conditions.

1.2.2 Convolution Theorem and its Convenience

Wikipedia (2011) provides the following definition of the convolution operator $*$ for functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$

$$h(t) = (f * g)(t) = \int_{-\infty}^{\infty} f(\tau)\, g(t - \tau)\, d\tau \tag{1.7}$$

and its well known properties with respect to the Fourier Transform

$$\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g) \qquad \text{and} \qquad \mathcal{F}(f) * \mathcal{F}(g) = \mathcal{F}(f \cdot g) \tag{1.8}$$

In conclusion, the theorem states that if a convolution is carried out in its Fourier domain (viz. the frequency domain if the convolution is formulated in the time domain), then the convolution operator transforms to a simple multiplication instead of the infinite weighted integral. Since many physical interactions can be described by convolution this theorem is used widely. However, especially because of its power, practitioners are often tempted to assume too lightly physical signals to be sufficiently stationary, which in turn is necessary for the Fourier transform of the signals in question.

1.2.3 Frequency Band-passed Processing as Workaround

Non stationary time series analysis must abstain from the Fourier Transform, which has become a very basic procedure in MT data processing. A few early attempts to process MT data entirely in the time domain have been driven by Berdichevsky (Berdichevsky et al., 2002, and references therein) but have not been developed further into a state-of-the art processing algorithm. One approach without an explicit need for quasi stationary data is described by Berdichevskiy et al. (1973). In their work, they propose to limit the frequency content of magnetotelluric time series ($\mathbf{e}$(t) and $\mathbf{h}$(t)) by consequent mathematical bandpass filtering subsequently for any desired number of frequencies. From these sets of filtered time series, the analytic signal ($\mathbf{e}_a(t)$ and $\mathbf{h}_a(t)$) of each set can be calculated if the spectral range $\bar{\omega} \pm \Delta\omega$ of the bandpass is sufficiently narrow. The impedance is not sensitive to little frequency variations and thus, each set of analytic signals of the filtered time series can be used to estimate the impedance of the average

frequency in that set:

$$\begin{pmatrix} e_{a,x}(t, \bar{\omega} \pm \Delta\omega) \\ e_{a,y}(t, \bar{\omega} \pm \Delta\omega) \end{pmatrix} = \begin{pmatrix} Z_{xx}(\bar{\omega}) & Z_{xy}(\bar{\omega}) \\ Z_{yx}(\bar{\omega}) & Z_{yy}(\bar{\omega}) \end{pmatrix} \cdot \begin{pmatrix} h_{a,x}(t, \bar{\omega} \pm \Delta\omega) \\ h_{a,y}(t, \bar{\omega} \pm \Delta\omega) \end{pmatrix}. \quad (1.9)$$

Since the impedance $\mathbf{Z}$ is assumed to be constant over time, it can be estimated by regression in the same way as it is done for equation (1.3). Because mathematical filtering does not require stationarity, this approach offers to process MT data when it is contaminated by non stationary noise or when the source signal is non stationary.

However, in theory, the frequency band should tend to zero ($\Delta\omega \to 0$), otherwise the analytic signal may not be correct and meaningful (Huang et al. (1998)). Since, this approach applies a constant bandpass $\Delta\omega$ over a constant centre frequency $\bar{\omega}$, the limit of $\Delta\omega \to 0$ effectively leads to the Fourier transform. In the following section, I will review a novel filter algorithm that is fully data adaptive and thus, it allows to efficiently decompose non stationary time series into functions that meet - per definition - the restrictions for constructing the analytic signal correctly.

1.2.4 The Hilbert-Huang Transform

Huang et al. (1998) introduced the Hilbert-Huang Transform (HHT), which is a new method to transform time series into a time-frequency domain without any assumptions on stationarity and linearity on the signal. The method has been extensively tested since then and successfully applied to different fields (Men-Tzung et al., 2008; Jackson and Mound, 2010; Zeiler et al., 2011; Chen et al., 2012), although a rigorous mathematical foundation is not yet available. The definition of HHT is empirical and data dependent, which on one hand provides a tool that works on non stationary, non linear bases but, on the other hand, denies a profound understanding of the method solely based on its definition.

Despite of the lack of a classical explicit mathematical basis, extensive tests have validated HHT and suggest that it improves time series analysis, in particular in the presence of non stationary or non linear effects (Huang et al., 2009). Furthermore, these tests confirm that a Fourier Transform cannot reliably represent the frequency information in a non stationary signal,

which require non stationary treatment.

Very often time series include non stationary and non linear effects and sometimes it is not desirable or not feasible to remove them. For instance, measurements of natural signals like the Earth's magnetic field are stationary for sufficient long periods of time, but measurements may include environmental noise which can be non stationary (Egbert, 2002; Chave and Thomson, 2004; Garcia and Jones, 2008). In this case, neither the exact noise signal nor the exact desired signal are known. Therefore, the desired stationary part cannot be isolated and the non stationary combination of both must be analysed.

Huang et al. (1998) introduced HHT and described thoroughly its application, restrictions and direct results. Later, Huang et al. (2009) updated the procedure by dividing its results into true amplitude and phase modulations. The most fundamental part of the HHT is called Empirical Mode Decomposition (EMD) which is a recipe for finding a set of orthogonal and complete Intrinsic Mode Function (IMF) from any given de-trended time series $x(t) = x_{\text{total}}(t) - x_{\text{trend}}(t)$.

One can conclude that the technique evolves the Fourier series:

$$x(t) = \sum_n a_n \cdot e^{i\omega_n t} = \sum_n a_n \cdot e^{i\phi_n(t)} \tag{1.10}$$

to a transformation with an amplitude and phase term that are dependent on time:

$$x(t) = \sum_j \hat{a}_j(t) \cdot e^{i\int_{-\infty}^{t} \hat{\omega}_j(t')dt'}. \tag{1.11}$$

Note, that the range of the indices n and j depend on the definition of a_n and $\hat{a}_j$, respectively, as usual for the Fourier Transformation. If a_n is real valued n ranges from $-\infty$ to ∞ without 0 and if a_n is complex n begins from 1. Consequently, we only use the latter definition. The exact same convention applies to $\hat{a}_j$.

The IMF $m_j(t)$ of $x(t)$ are defined as

$$m_j(t) = \hat{a}_j(t) \cdot e^{i\hat{\phi}_j(t)} \tag{1.12}$$

with $\hat{\phi}_j(t) = \int_{-\infty}^{t} \hat{\omega}_j(t')dt'$ and by the following properties:

1. In the whole dataset, the number of extrema and the number of zero-crossings must either equal or differ at most by one, and

2. at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

It is noteworthy, that in the frame of the Fourier expansion, a_n describes the complex amplitude of the mono-frequency part (with ω_n) of the signal $x(t)$ for the entire process. Whereas, $\hat{a}_j(t_0)$ is the complex amplitude of the signal at $t = t_0$ which exhibits the frequency $\hat{\omega}_j(t_0)$.

In other words, the HHT separates time-varying Amplitude modulation (AM) and Phase modulation (PM) in form of oscillatory modes from the data and provides them as IMFs. The values of the PM are in the open interval between -1 and $+1$ and are defined such that they are locally zero-mean functions (number of extrema and number of zero-crossings differ by not more than one, for detailed information, confer Huang et al. (2009)). Huang et al. (1998) argue that phase functions with these properties can be subjected to the Hilbert Transform in order to acquire their analytic signal and that these functions exhibit a physical meaningful instantaneous frequency. The Hilbert Transform of a suitable function $y(x)$ is defined by

$$H(y)(x) = \frac{\text{p.v.}}{\pi} \cdot \int_{-\infty}^{\infty} \frac{y(\tau)}{x - \tau} d\tau \tag{1.13}$$

with $p.v.$ as the Cauchy principal value. We can construct the analytic signal by

$$y_a(x) = y(x) + iH(y)(x) \tag{1.14}$$

where $H(y)(x)$ is the Hilbert Transform as of equation (1.13) and, thus, obtain the signal's phase with

$$\Phi(x) = \arctan\left(\frac{H(y)(x)}{y(x)}\right) \tag{1.15}$$

Ultimately, the instantaneous frequency is defined as time derivative of the phase:

$$\Omega = \frac{d\Phi(x)}{dx}. \tag{1.16}$$

In conclusion, the HHT provides the complete set of instantaneous parameters (amplitude, phase and frequency) as functions of time for any given time series and formulates a mean to obtain complex spectral information from time series without any stationarity assumptions. Therefore, HHT seems to be a natural choice for an in-depth investigation on its performance related to non stationary magnetotelluric processing.

As a final remark, note that the electromagnetic fields in the approach of Berdichevskiy et al. (1973, see equation (1.9) in previous section) need to be narrowly band-passed and thus, represent an approximation of an IMF with the restriction, that the center-frequency (instantaneous frequency) is constant. HHT decomposes signals into IMFs which can be described by their instantaneous frequency and thus, IMFs could be considered as dynamically, band-passed versions of the signal itself.

1.3 Statistical Transfer Function Estimation

The magnetotelluric transfer function $\mathbf{Z}(\omega)$ is defined through coefficients that relate the horizontal electric with the horizontal magnetic field (see equation (1.3) for reference). Expanding this relation results in two equations, one for E_x and one for E_y, each of which are represented through B_x and B_y:

$$E_x(\omega) = Z_{xx}(\omega)B_x(\omega) + Z_{xy}(\omega)B_y(\omega), \tag{1.17a}$$

$$E_y(\omega) = Z_{yx}(\omega)B_x(\omega) + Z_{yy}(\omega)B_y(\omega). \tag{1.17b}$$

Each of these two equations can be regarded as a multiple linear regression problem, which has a unique solution when there are two electric and magnetic field measurements for the frequency ω. For more than two measurements of the electric and magnetic field at a certain frequency (ω), a least squares solution can be found that minimizes the residual difference between the measured and estimated electric field. However, it has been shown that the least squares solution is not suitable for real MT data (Jones et al., 1989), since the measurements often contain noise, which leads to a strong bias in the least squares solution. Therefore robust statistical estimators are the widely preferred solution. A simple robust least squares estimator, such as the one applied here, is an algorithm that iteratively weights the

residuals and finds a solution for which most measurements have the lowest possible residual. All data with higher residuals are dismissed (Huber, 2005). Unfortunately, the robustness of this estimation is only towards errors in the predictor (here the electric field) and does not take into account possible errors in the observable (here the magnetic field), thus the estimation will be biased for noise variance in the magnetic fields. The following sectionadresses this problem.

1.3.1 Principal Component Analysis and Two Step Regression

Egbert (1997) investigates the electromagnetic field in a multivariate sense and reports that all natural electric and magnetic fields share two common polarization vectors. He concludes that the Principal Component Analysis (PCA) of all EM field components of nearby stations would result in two major principal components, if there was no correlated noise present in the signal. This finding is the corner stone of his multivariate Errors-in-Variables processing algorithm, which incorporates intrinsically and thus generalizes the remote-referencing technique introduced to the MT community by Gamble et al. (1979).

Assume that the Principal Components of any given number of horizontal EM field components $\mathbf{E}^i = [E^i_x, E^i_y]$ and $\mathbf{B}^i = [B^i_x, B^i_y]$ can be computed:

$$\mathbf{PC} = \mathcal{PCA}(\mathbf{E}^1, \mathbf{B}^1, \ldots, \mathbf{E}^i, \mathbf{B}^i), \tag{1.18}$$

where $\mathbf{PC}$ are the principal components of the EM field components at site i returned by $\mathcal{PCA}$, which symbolizes a (preferably robust) principal component analysis such as documented by Hubert and Verboven (2003). If there is no correlated noise present between all the measurements, the two major principal components should contain the EM signal polarization vectors, which represent the entire EM signal to first order (Egbert, 1997, 2002; Smirnov and Egbert, 2012). Since the principal components are statistical entities and not measurements, they can be assumed error-free (free of random and uncorrelated noise) and thus can be regarded better observables than the measurements of the magnetic field itself. For each station, we can compute the regression parameters of both, the electric field $\mathbf{E}$ and the

magnetic field $\mathbf{B}$ on the two major principal components $\mathbf{PC}_{1,2}$:

$$\mathbf{E}(\omega) = \mathbf{R}^e(\omega) \cdot \mathbf{PC}_{1,2}(\omega), \tag{1.19a}$$

$$\mathbf{B}(\omega) = \mathbf{R}^h(\omega) \cdot \mathbf{PC}_{1,2}(\omega). \tag{1.19b}$$

These reformulated regression problems allow a robust estimation of the new transfer functions $\mathbf{R}^e$ and $\mathbf{R}^h$. Errors in the electric and the magnetic field can be accounted for by the robust regression scheme described previously.

Multiplying (1.19b) by the inverse of $\mathbf{R}^h$ and inserting the product in (1.19a), shows that the impedance $\mathbf{Z}(\omega)$ is the product of $\mathbf{R}^e$ and the inverse of $\mathbf{R}^h$, viz.

$$\mathbf{Z}(\omega) = \mathbf{R}^e(\omega) \cdot \left(\mathbf{R}^h(\omega)\right)^{-1}. \tag{1.20}$$

The estimation of $\mathbf{Z}(\omega)$ with this two step regression yields robust results insensitive to random gaussian noise present in the electric and magnetic field.

1.3.2 Estimation Uncertainty

A comprehensive interpretation of results from statistical analyses requires that they are accompanied by some measure of uncertainty. There are essentially two main branches of uncertainty measures, They can either be derived from parametric modeling or from non parametric estimation of a statistical model. Parametric modeling tries to fit an ideal stationary statistical model and indicates the goodness of fit through descriptive parameters (e.g. residual variance analysis, curve fitting). Non parametric methods analyze the behavior of the data under statistical conditions and offer descriptive parameters on the characteristics of data (e.g. bootstrapping, jackknife, qq-plots, Theil-Sen estimator).

Since this thesis concentrates on non stationary data analysis, the uncertainty of the data may vary over time and thus, I will use the non parametric method 'bootstrapping' (Efron, 1979; Liu et al., 1988) to estimate the MT Transfer function (TF) and its uncertainty.

1.4 Objectives of the Thesis

Empirical Mode Decomposition (EMD) is a novel technique, which is designed to transform non stationary data into a time-frequency domain, and might very well substitute and qualitatively outperform Fourier Transform based estimation schemes. Some advances with this technique in the field of MT processing have been reported (Cai et al., 2009; Chen et al., 2012) but their method uses EMD as a filter rather than a Fourier Transform substitute and do not take full advantage of the inherent properties of EMD, which is the full decomposition of the time series into a complete and orthogonal set of its instantaneous amplitude, phase and frequency. Therefore, this thesis focusses on the use of EMD as a true Fourier Transform substitute by developing the theory for non stationary convolution and how it can be embedded in MT data processing. Following this general introduction, the chapters 2 to 4 lead stepwise towards a newly proposed MT processing scheme that is entirely freed of the stationarity assumption. Subsequently, the thesis concludes with a concise summary and suggestions on how the work can be continued in chapter 5.

- **Chapter 2: Non Stationary Time Series Convolution**
  In this chapter, I develop the mathematical foundation for using the HHT for traditional transfer function analysis, like MT data processing. The HHT decomposes time series into IMF, which are in the time and in the frequency domain, and this chapter demonstrates that time slices of IMFs equal time slices of Fourier Series, where the instantaneous parameters of the IMF define the parameters amplitude and phase of the Fourier Series. This finding leads to the formulation of the theorem that non stationary convolution of an IMF with a general time domain response function translates into a multiplication of the IMF with the respective spectral domain response function, which is - in addition - explicitly permitted to vary over time. The theorem is verified by comparison with convolution by mathematical filtering on both, stationary and non stationary, synthetic data.

  The chapter concludes with an example of MT data processing verifying that a de-trended signal's IMFs can be convolved independently and then be used for further time-frequency analysis. Finally, a discus-

sion is opened on parallels in between HHT and the Fourier Transform with respect to the time-frequency domain.

- **Chapter 3: Frequency Shift Caused by Convolution**
  By using the Hilbert-Huang Transform, a non stationary time series can be represented by a number of modes, which are complex time series with instantaneous amplitudes, phases and frequencies. Following the non stationary convolution theorem from the previous chapter, I analyze analytically the characteristics of a convolved time series and caution that through non stationary convolution the instantaneous frequency of a signal may change. Finally, I quantify the frequency shift and argue that this difference may hamper any attempt to deconvolve non stationary signals. The argument is illustrated by synthetic data.

- **Chapter 4: Non Stationary MT Data Processing**
  Because non stationary electromagnetic noise affects the computation of Fourier Spectra and therefore, the estimation of the MT TF as outlined in the introduction, this chapter introduces a TF estimation scheme based on the EMD and the non stationary convolution theorem from chapter 2. In contrast to previous works that employ EMD for MT data processing, I argue the necessity of a multivariate analysis, highlighting the possibility to use instantaneous parameters and define the homogenization of frequency discrepancies between data channels. The presented scheme uses the robust statistical estimation of transfer functions based on principal component regression and can be used with and without aid by any number of available remote reference stations. Data errors are estimated by enveloping the entire scheme with a bootstrap routine. The algorithm is verified on synthetic and real data (Southern Africa) with and without non stationary character and the results are compared with different processing techniques.

  Furthermore, the algorithm is used to assess the effects of non stationary noise in MT data processing. I compare the presented algorithm with traditional and efficient processing codes based on the Fourier Transform. The benchmark is carried out on synthetic, non stationary data and fair real data with and without added synthetic, non stationary noise in order to show how easily a quasi stationary method is

compromised by a non stationary source and that the EMT algorithm is able to treat the situation correctly.

Lastly, I describe a complementary, free MatLab program that computes non stationary data for the magnetotelluric method. The synthetic source field can be defined by arbitrary frequency and amplitude functions or by the default parameters from the given example in this chapter. For the impedance model, a one dimensional conductive earth model can be computed or any impedance can be imported from files with the Elecrical Data Interchange (EDI) standard.

Bibliography

Berdichevskiy, M., Bezruk, I., Chinareva, O., 1973. Magnetotelluric sounding using mathematical filters. Vol. 9. American Geophysical Union.

Berdichevsky, M., Dmitriev, V., Keller, G., 2002. Magnetotellurics in the Context of the Theory of Ill-posed Problems. Investigations in Geophysics, 11. Society of Exploration Geophysicists.
URL `http://books.google.com/books?id=X50pVCKGq1AC`

Brasse, H., 1993. Audiomagnetotellurische Tiefensondierungen in Nordost-Africa. Ph.D. thesis, TU Berlin.

Cagniard, L., 1953. Basic theory of the magneto-telluric method of geophysical prospecting. Geophysics 18 (3), 605.

Cai, J.-H., Tang, J.-T., Hua, X.-R., Gong, Y.-R., 2009. An analysis method for magnetotelluric data based on the Hilbert–Huang Transform. Exploration Geophysics 40 (2), 197.

Chave, A. D., Jones, A. G., 2012. The Magnetotelluric Method: Theory and Practice. Cambridge University Press.

Chave, A. D., Thomson, D. J., 2004. Bounded influence magnetotelluric response function estimation. Geophysical Journal International 157 (3), 988.

Chave, A. D., Thomson, D. J., Ander, M. E., 1987. On the robust estimation of power spectra, coherences, and transfer functions. Journal of Geophysical Research: Solid Earth (1978–2012) 92 (B1), 633.

Chen, J., Heincke, B., Jegen, M., Moorkamp, M., 2012. Using empirical mode decomposition to process marine magnetotelluric data. Geophysical Journal International 190 (1), 293.
URL http://dx.doi.org/10.1111/j.1365-246X.2012.05470.x

Christopherson, K. R., 1998. MT Gauges Earths Electric Fields. AAPG Explorer.

Efron, B., 1979. Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics 7 (1), 1.

Egbert, G. D., 1997. Robust multiple-station magnetotelluric data processing. Geophysical Journal International 130 (2), 475.

Egbert, G. D., 2002. Processing and interpretation of electromagnetic induction array data. Surveys in geophysics 23 (2-3), 207.

Egbert, G. D., Booker, J. R., 1986. Robust estimation of geomagnetic transfer functions. Geophysical Journal International 87 (1), 173.

Fourier, J., 1822. Théorie analytique de la chaleur. Chez Firmin Didot, père et fils.
URL http://books.google.es/books?id=TDQJAAAAIAAJ

Fourier, J., Freeman, A., 1878. The Analytical Theory of Heat. The University Press.
URL http://books.google.es/books?id=-N8EAAAAYAAJ

Füllekrug, M., 1994. Schumann-Resonanzen in den Magnetfeld-Komponenten. Cuvillier Verlag, Göttingen, Germany.
URL http://opus.bath.ac.uk/12968/

Gamble, T., Goubau, W. M., Clarke, J., 1979. Magnetotellurics with a remote magnetic reference. Geophysics 44 (1), 53.

Garcia, X., Jones, A. G., 2002. Atmospheric sources for audio-magnetotelluric (AMT) sounding. Geophysics 67 (2), 448.

Garcia, X., Jones, A. G., 2005. A new methodology for the acquisition and processing of audio-magnetotelluric (AMT) data in the AMT dead band. Geophysics 70 (5), G119.

Garcia, X., Jones, A. G., 2008. Robust processing of magnetotelluric data in the AMT dead band using the continuous wavelet transform. Geophysics 73 (6), F223.

Gauss, C. F., 1838. Allgemeine Theorie des Erdmagnetismus. Resultate aus den Beobachtungen des Magnetischen Vereins im Jahre 1838.

Huang, N., Wu, Z., Long, S., Arnold, K., Chen, X., 2009. On instantaneous frequency. Advances in Adaptive Data Analysis 1 (2), 177.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., Liu, H. H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 454 (1971), 903.

Huber, P. J., 2005. Robust Statistics. John Wiley & Sons, Inc.

Hubert, M., Verboven, S., 2003. A robust PCR method for high-dimensional regressors. Journal of Chemometrics 17 (8-9), 438.

Jackson, L. P., Mound, J. E., 2010. Geomagnetic variation on decadal time scales: What can we learn from Empirical Mode Decomposition? Geophysical Research Letters 37 (14).

Jones, A. G., Chave, A. D., Egbert, G., Auld, D., Bahr, K., 1989. A comparison of techniques for magnetotelluric response function estimation. Journal of Geophysical Research: Solid Earth (1978–2012) 94 (B10), 14201.

Junge, A., 1996. Characterization of and correction for cultural noise. Surveys in geophysics 17 (4), 361.

Lahiri, B., Price, A., 1939. Electromagnetic induction in non-uniform conductors, and the determination of the conductivity of the earth from terrestrial magnetic variations. Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 509.

Lamb, H., 1883. On Electrical Motions in a Spherical Conductor.

Liu, R. Y., et al., 1988. Bootstrap procedures under some non-iid models. The Annals of Statistics 16 (4), 1696.

Liu, W., Fujimoto, M., 2011. The Dynamic Magnetosphere. Springer.

Manoj, C., Nagarajan, N., 2003. The application of artificial neural networks to magnetotelluric time-series analysis. Geophysical Journal international 153 (2), 409.

Margrave, G., 1998. Theory of nonstationary linear filtering in the Fourier domain with application to time-variant filtering. Geophysics 63 (1), 244.

Maxwell, J., 1873. A Treatise on Electricity and Magnetism (Clarendon, Oxford, 1873).

Men-Tzung, L., Kun, H., Yanhui, L., C-K, P., Vera, N., 2008. Multimodal pressure-flow analysis: Application of Hilbert Huang transform in cerebral blood flow regulation. EURASIP journal on advances in signal processing 2008.

Numhauser, J., 1982. Todo Cambia. on Album "Todo Cambia", distributed by Philips.

Oettinger, G., Haak, V., Larsen, J., 2001. Noise reduction in magnetotelluric time-series with a new signal–noise separation method and its application to a field experiment in the Saxonian Granulite Massif. Geophysical Journal International 146 (3), 659.

Price, A. T., 1930. Electromagnetic Induction in a Conducting Sphere. Proceedings of the London Mathematical Society s2-31 (1), 217.

Rakov, V. A., Uman, M. A., 2007. Lightning. Physics and Effects. Cambridge University Press.

Reddy, I., Rankin, D., 1974. Coherence functions for magnetotelluric analysis. Geophysics 39 (3), 312.

Reiersøl, O., 1941. Confluence Analysis by Means of Lag Moments and Other Methods of Confluence Analysis. Econometrica 9 (1), 1.

Rikitake, T., 1948. Notes on the Electromagnetic Induction within the Earth., distributed through mtnet.dias.ie.

Schmucker, U., Weidelt, P., 1975. Electromagnetic Induction in the Earth.

Schröder, W., Wiederkehr, K. H., 2000. A history of the early recording of geomagnetic variations. Journal of Atmospheric and Solar-Terrestrial Physics 62 (5), 323.

Schuster, A., 1886. On the diurnal period of terrestrial magnetism. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 21 (131), 349.

Sims, W., Bostick Jr, F., Smith, H., 1971. The estimation of magnetotelluric impedance tensor elements from measured data. Geophysics 36 (5), 938.

Sims, W. E., Bostick Jr, F., 1969. Methods of magnetotelluric analysis. Tech. rep., DTIC Document.

Smirnov, M. Y., Egbert, G., 2012. Robust principal component analysis of electromagnetic arrays with missing data. Geophysical Journal International 190 (3), 1423.

Smith, S. W., 1997. The scientist and engineer's guide to digital signal processing. California Technical Publishing, San Diego, CA, USA.

Stewart, B., 1877. On the Variations of the Daily Range of the Magnetic Declination as Recorded at the Kew Observatory. Proceedings of the Royal Society of London 26 (179-184), 102.

Szarka, L., 1988. Geophysical aspects of man-made electromagnetic noise in the earthA review. Surveys in Geophysics 9 (3-4), 287.

Theissing, J. T., 1988. Eine Methode zur Trennung yon Signal- und Störanteilen in magnetotellurischen Zeitreihen, diploma thesis at Geophysik Univ. Münster.

Tikhonov, A., 1950. On determining electrical characteristics of the deep layers of the earths crust. In: Sov. Math. Dokl. Vol. 2. p. 295.

Vozoff, K., 1972. The magnetotelluric method in the exploration of sedimentary basins. Geophysics 37 (1), 98.

Weckmann, U., Magunia, A., Ritter, O., 2005. Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme. Geophysical Journal International 161 (3), 635.

Weidelt, P., 1972. The inverse problem of geomagnetic induction. Journal of Geophysical Research 38, 257.

Wikipedia, 2011. Convolution — Wikipedia, The Free Encyclopedia. Accessed July 18th 2011.
URL `http://en.wikipedia.org/wiki/Convolution`

Zeiler, A., Faltermeier, R., Brawanski, A., Tomé, A. M., Puntonet, C. G., Górriz, J. M., Lang, E. W., 2011. Brain status data analysis by sliding EMD. In: Proceedings of the 4th international conference on Interplay between natural and artificial computation: new challenges on bio-inspired applications - Volume Part II. IWINAC'11. Springer-Verlag, Berlin, Heidelberg, p. 77.
URL `http://dl.acm.org/citation.cfm?id=2009542.2009551`

CHAPTER 2

Non Stationary Time Series Convolution: On the Relation between Hilbert-Huang and Fourier Transform

published in Journal of Advances in Adaptive Data Analysis
23 April 2013

Maik Neukirch and Xavier Garcia

Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar, CSIC
Pg. Maritím de la Barceloneta 37-49, 08003 Barcelona, Spain

Abstract

The Hilbert-Huang Transform (HHT) decomposes time series into Intrinsic Mode Function (IMF) in time-frequency domain. We show that time slices of IMFs equal time slices of Fourier Series, where the instantaneous parameters of the IMF define the parameters amplitude and phase of the Fourier Series. This leads to the formulation of the theorem that non stationary convolution of an IMF with a general time domain response function translates into a multiplication of the IMF with the respective spectral domain response function which is explicitly permitted to vary over time. We conclude and show on a real world application that a de-trended signal's IMFs can be convolved independently and then be used for further time-frequency analysis. Finally, a discussion is opened on parallels in HHT and the Fourier Transform with respect to the time-frequency domain.

2.1 Introduction

In digital signal processing, time series convolution is often related to the Fourier Transform (FT) and therefore implies stationary and linear assumptions on the data. The reason for this prominence lies within the convolution theorem which allows to exchange a weighted integral expression to a simple multiplication, which results in much shorter computation time (Smith, 1997). Margrave (1998) introduced the theory for non stationary convolution filters based on the Fourier Transform arguing that a continuous function is completely described by its Fourier Transform and, therefore, non stationary filtering should be possible in the frequency domain. Huang et al. (1998) show that the frequency information of non stationary signals might describe the signal entirely but gives biased information with respect to the physics of the signal, concluding that misinterpretation of the Fourier Transform cannot be ruled out in a truly non stationary setting.

Huang et al. (1998) introduced the HHT, which is a new method to transform time series into a time-frequency domain without any assumptions on stationarity and linearity on the signal. The method has been extensively tested since then and successfully applied to different fields (Men-Tzung et al., 2008; Jackson and Mound, 2010; Zeiler et al., 2011; Chen et al., 2012), although a rigorous mathematical foundation is not yet available. The definition of HHT is empirical and data dependent, which on one hand provides a tool that works on non stationary, non linear bases but, on the other hand, denies a profound understanding of the method solely based on its definition.

Despite of the lack of a classical explicit mathematical basis, extensive tests have validated HHT and suggest that it improves time series analysis, in particular in the presence of non stationary or non linear effects (Huang et al., 2009). Furthermore, these tests confirm that a Fourier Transform cannot reliably represent the frequency information in a non stationary signal which, hence, require non stationary treatment.

Very often time series include non stationary and non linear effects and sometimes it is not desirable or not feasible to remove them. For instance, measurements of natural signals like the Earth's magnetic field are stationary for sufficient long periods of time, but measurements may include en-

vironmental noise which can be non stationary (Egbert, 2002; Chave and Thomson, 2004; Garcia and Jones, 2008). In this case, neither the exact noise signal nor the exact desired signal are known. Therefore, the desired stationary part cannot be isolated and the non stationary combination of both must be analyzed. In order to solve that exact problem we discuss how the convolution filter affects non stationary signals and extend the convolution theorem to non stationary signals.

In this work we present the non stationary convolution based on HHT which does not imply assumptions on the stationarity of the signal. Since results of the HHT are neither exclusively in the time nor frequency domain, we cannot readily generalize the established convolution theorem for an analysis based on HHT but we can show, that there are fundamental similarities between the FT and the HHT with respect to convolution and use those similarities to find a new formulation for the non stationary convolution. Due to the nature of non stationary signals and how the frequency information can be recovered by HHT, we will argue that a non stationary convolution based on HHT does not necessarily have an uniquely defined inverse, or a deconvolution operator resulting in the original signal, but we will briefly discuss resulting implications for the deconvolution of such signals.

The paper starts with a brief review of the HHT, highlighting the instantaneous parameters which are the backbone of our theorem. Then, it continues by presenting the formulation of the non stationary convolution theorem and the lemma required for the subsequent proofs. The theorems are interpreted physically and their implications on the relation between FT and HHT are laid out. The paper concludes with two numerical examples on a stationary and a non stationary test signal, and an example of a genuine geophysical application with real world data. It follows a discussion on our findings and suggestions for further work especially with respect to non stationary deconvolution.

2.2 Hilbert-Huang Transform (HHT)

Huang et al. (1998) introduced the HHT and described thoroughly its application, restrictions and direct results. Given a de-trended time series

$f(t) = f_{\text{total}}(t) - f_{\text{trend}}(t)$, the HHT generalizes the Fourier series

$$f(t) = \sum_n F_n \cdot e^{2\pi i \nu_n t} = \sum_n F_n \cdot e^{i\phi_n(t)}, \tag{2.1}$$

where the phase is defined as $\phi_n(t) = 2\pi\nu_n t$, to a series with an amplitude $\hat{F}(t)$ and frequency $\hat{\nu}(t)$

$$f(t) = \sum_j \hat{F}_j(t) \cdot e^{2\pi i \int_{-\infty}^{t} \hat{\nu}_j(t')dt'} = \sum_j \hat{F}_j(t) \cdot e^{i\hat{\phi}_j(\hat{\nu}_j(t),t)}, \tag{2.2}$$

with the phase $\hat{\phi}_j(\hat{\nu}_j(t), t) = 2\pi \int_{-\infty}^{t} \hat{\nu}_j(t')\, dt'$. Note, that the range of the index $n \in \mathbb{Z}$ depends on the definition of the amplitudes F_n as usual for the Fourier Transform. For $F_n \in \mathbb{R}$: $n \in \mathbb{Z}$ and for $F_n \in \mathbb{C}$: $n \in \mathbb{N}^0$ with $\mathbb{N}^0$ being the natural numbers including zero. Both definitions are equivalent, so let us concentrate on the complex definition for F_n in this work. On the other hand, $\hat{F}_j \in \mathbb{R}$ always is a real amplitude of the signal and is defined for $j \in \mathbb{N}^+$ for an infinite long function $f(t)$ with $\mathbb{N}^+$ being the natural numbers exclusive zero.

The Intrinsic Mode Function (IMF) $m_j(t)$ of $f(t)$ are defined as

$$m_j(t) = \hat{F}_j(t) \cdot e^{i\hat{\phi}_j(\hat{\nu}_j(t),t)} \tag{2.3}$$

by the following properties:

1. In the whole dataset, the number of extrema and the number of zero-crossings must either equal or differ at most by one, and

2. at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

Note, that within the frame of the Fourier expansion, F_n describes the constant complex amplitude of the mono-frequency part (with ν_n) of the signal $f(t)$ for the entire process. Whereas, $\hat{F}_j(t)$ is the real amplitude of the IMF j which exhibits a frequency $\hat{\nu}_j(t)$, which both can vary over time.

In other words, the Hilbert-Huang Transform (HHT) separates narrow-bandwidth Amplitude modulations (AMs) $a = \hat{F}_j(t)$ and Phase modulations (PMs) $p = e^{i\hat{\phi}_j(t)}$ from the data and provides them in form of real-valued Intrinsic Mode Functions (IMFs) Huang et al. (2009). This process is called

Empirical Mode Decomposition (EMD) of $f(t)$ and yields the corresponding, real-valued IMFs, which represent the real part of equation (2.3). The AM values a describe time varying signal power, whereas the PM p only contain pure oscillations. The real values of p are in the open interval between -1 and $+1$ and are defined such that they are locally zero-mean functions, which means that the number of extrema and number of zero-crossings do not differ by more than one (for detailed information, see Huang et al. (2009)). Huang et al. (1998) argue that phase functions with these properties can be Hilbert-transformed to acquire their analytic signal and that they exhibit a physical meaningful instantaneous frequency. The Hilbert Transform of a suitable function $p(t)$ is defined by

$$H(p)(t) = \frac{\text{p.v.}}{\pi} \cdot \int_{-\infty}^{\infty} \frac{p(\tau)}{t-\tau} d\tau, \tag{2.4}$$

where $p.v.$ indicates Cauchy's principal value. We can construct the analytic signal by

$$m(t) = a\left(p(t) + iH(p)(t)\right), \tag{2.5}$$

where $H(p)(t)$ is the Hilbert Transform equation (2.4) and, thus, obtain the signal's phase using

$$\hat{\phi}(t) = \arctan\left(\frac{H(p)(t)}{p(t)}\right). \tag{2.6}$$

Ultimately, the instantaneous frequency is defined as the time derivative of the phase:

$$\hat{\nu}(t) = \frac{d\hat{\phi}(t)}{2\pi dt}. \tag{2.7}$$

2.3 Non Stationary Convolution Under HHT

Let us consider an integrable function $f : t \in \mathbb{R} \to \mathbb{R}$ in the integral formulation of the FT pairs (taken from Wikipedia (2011) with references therein):

$$F(\nu) = \int_{-\infty}^{\infty} f(t) \cdot e^{-2\pi i \nu t} dt = \mathcal{F}(f(t)) \tag{2.8}$$

$$f(t) = \int_{-\infty}^{\infty} F(\nu) \cdot e^{2\pi i \nu t} d\nu = \mathcal{F}^{-1}(F(\nu)) \tag{2.9}$$

with $\phi(\nu, t) = 2\pi\nu t$, and a general IMF $m : t \in \mathbb{R} \to \mathbb{R}$:

$$m\left(t\right) = \hat{M}(t) \cdot e^{i\hat{\phi}(t)} \tag{2.10}$$

with $\hat{M}(t)$ representing the real-valued, instantaneous amplitude of m. Let us define a complex amplitude function $\tilde{M}(\hat{\nu}, t) = \hat{M}(t) \cdot e^{i\tilde{\phi}(\hat{\nu},t)}$ and a phase function $\tilde{\phi}(\hat{\nu}(t), t) = \hat{\phi}(t) - 2\pi\hat{\nu}t$ to rewrite the IMF m as

$$m\left(t\right) = \tilde{M}(\hat{\nu}, t) \cdot e^{2\pi i\hat{\nu}t}. \tag{2.11}$$

2.3.1 Theorem 1

Let $m(t)$ be an IMF with instantaneous frequency $\hat{\nu}(t)$, $F^{\tau}(\nu)$ a member of the group of Fourier Transforms of $f^{\tau}(t)$ and $\tau \in \mathbb{R}$ the parameter which describes each member, then:

$$m(t) \cdot F^{\tau}(\hat{\nu}(t)) = m(t) * f^{\tau}(t). \tag{2.12}$$

The frequency wise multiplication of m with F equals the convolution of m with f.

2.3.2 Lemma 1

Let $m(t)$ be an IMF with instantaneous frequency $\hat{\nu}(t)$, then the convolution of $m(t)$ with the delta distribution $\delta(t)$ is

$$\left(m * \delta\right)\left(t_0\right) = \left(\tilde{M}_{t_0} \cdot e^{2\pi i\hat{\nu}(t_0)t} * \delta(t)\right)\left(t_0\right), \tag{2.13}$$

with $\tilde{M}_{t_0} = \hat{M}(t_0) \cdot e^{i\hat{\phi}(t_0) - 2\pi i\hat{\nu}(t_0)t_0} = \tilde{M}(\hat{\nu}(t_0), t_0)$ being the complex amplitude of a monochromatic oscillation with frequency $\hat{\nu}(t_0)$.

The proof of this lemma is trivial, but we include it to stress that this property of the IMF can be used to find more properties of the HHT with the help of well-known properties of the FT. Figure 2.1 provides a graphical illustration of this Lemma.

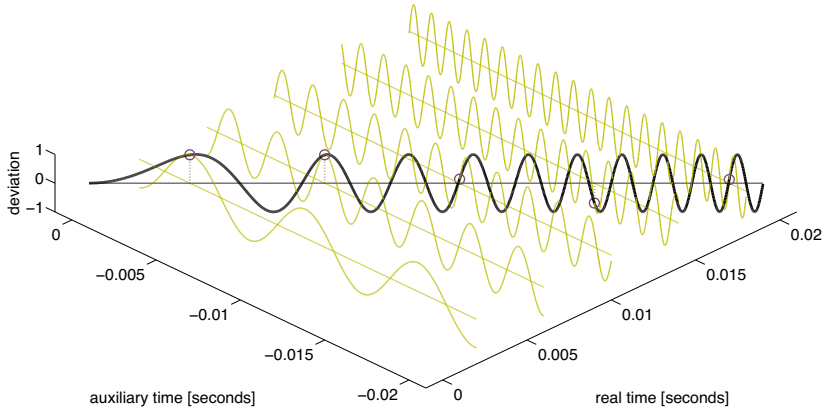


Figure 2.1: *This figure illustrates Lemma 1. The black curve is an IMF (a chirp function with linearly increasing frequency and constant amplitude, cp. section 2.8) and the light curves are sine curves with the frequency, phase and amplitude chosen to be identical with the IMF at the intersection points marked by reddish circles. The IMF evolves along the actual time axis, whereas the sine curves are displayed in an auxiliary time domain whose sole purpose is to visualise them. The vertical axis describes the deviation for both, the IMF and the sine curves.*

2.4 Proof for Theorem 1

Starting on the Right hand side (RHS) from the following identity:

$$\left(m(t) * f^{\tau}(t)\right)(t) = \left(\delta * m * f^{\tau}\right)(t), \tag{2.14}$$

focusing on an isolated time instant t_0

$$\left((\delta * m) * \left(f^{\tau}\right)\right)(t_0) \tag{2.15}$$

and using the sifting property of the delta function with Lemma 1, the RHS yields

$$\begin{aligned}\left((\delta * m) * f^{\tau}\right)(t_0) &= \left[\left(\delta * \tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t}\right) * f^{\tau}(t)\right](t_0) \\ &= \left[\delta * \tilde{M}_{t_0} \cdot \mathcal{F}^{-1}\left(\delta(\hat{\nu}(t_0) - \nu)\right) * \mathcal{F}^{-1}\left(F^{\tau}(\nu)\right)\right](t_0)\,.\end{aligned} \tag{2.16}$$

When we resolve the convolution with the help of the convolution theorem for Fourier Transforms, we obtain

$$\begin{aligned}((\delta * m) * f^{\tau}(t))(t_0) &= \left[\delta * \tilde{M}_{t_0} \cdot \mathcal{F}^{-1}\left(\delta(\hat{\nu}(t_0) - \nu) \cdot F^{\tau}(\nu)\right)\right](t_0) \\ &= \left[\delta * \tilde{M}_{t_0} \cdot \int_{-\infty}^{\infty} \delta(\hat{\nu}(t_0) - \nu) \cdot F^{\tau}(\nu) \cdot e^{2\pi i \nu t} d\nu\right](t_0) \\ &= \left[\delta * \tilde{M}_{t_0} \cdot F^{\tau}(\hat{\nu}(t_0)) \cdot e^{2\pi i \hat{\nu}(t_0) t}\right](t_0) \\ &= \tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t_0} \cdot F^{\tau}(\hat{\nu}(t_0)).\end{aligned} \tag{2.17}$$

Now, the Left hand side (LHS) of Theorem 1 can be written as:

$$\begin{aligned}(\delta * m(t) \cdot F^{\tau}(\hat{\nu}(t)))(t_0) &= \left(\delta * \hat{M}(t) \cdot e^{i\hat{\phi}(t)} \cdot F^{\tau}(\hat{\nu}(t))\right)(t_0) \\ &= \left(\delta * \hat{M}(t) \cdot e^{i\hat{\phi}(t) - 2\pi i \hat{\nu}(t) t + 2\pi i \hat{\nu}(t) t} \cdot F^{\tau}(\hat{\nu}(t))\right)(t_0) \\ &= \hat{M}(t_0) \cdot e^{i\hat{\phi}(t_0) - 2\pi i \hat{\nu}(t_0) t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t_0} \cdot F^{\tau}(\hat{\nu}(t_0)),\end{aligned} \tag{2.18}$$

for $\tilde{M}_{t_0} = \hat{M}(t_0) \cdot e^{i\left(\hat{\phi}(t_0) - 2\pi \hat{\nu}(t_0) t_0\right)}$ as required for Lemma 1, LHS and RHS of Theorem 1 are equivalent for all times t_0.

2.5 Proof for Lemma 1

The proof is a straight forward application of the sifting property of the delta distribution in a convolution and the insertion of a zero term. The LHS can be reformulated as follows by using equation (2.11):

$$\begin{aligned}(m * \delta)(t_0) &= \left(\tilde{M}(\hat{\nu}(t), t) \cdot e^{2\pi i \hat{\nu}(t) t} * \delta(t)\right)(t_0) \\ &= \tilde{M}(\hat{\nu}(t_0), t_0) \cdot e^{2\pi i \hat{\nu}(t_0) t_0}.\end{aligned} \tag{2.19}$$

$\tilde{M}(\hat{\nu}, t)$ can be understood as a complex-valued, instantaneous amplitude which incorporates an instantaneous starting phase in order to linearize the phase term of the IMF m. By reformulating the RHS of Lemma 1, we find

$$\left(\tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t} * \delta\right)(t_0) = \tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t_0}. \tag{2.20}$$

and readily see that both sides are equivalent for $\tilde{M}_{t_0} = \tilde{M}\left(\hat{\nu}(t_0), t_0\right)$.

2.6 Physical Interpretation

2.6.1 Theorem 1

Assuming we know exactly the spectral response $F(\nu)$ of a physical measurement device, this theorem states that we can simply multiply the known spectral response $F(\nu)$ by a known time signal $m(t)$ in order to obtain the signal $n(t)$ measured by the device, if the signal $m(t)$ is an IMF. Per definition, $n(t)$ is exactly the convolution of the time domain response function $f(t)$ of the device with the incoming signal $m(t)$. Our reformulation of a non stationary convolution to a simple multiplication leads to a better understanding of the behavior of physical systems in a non stationary set up and further increases the application range of the HHT.

An important note is that a de-trended signal $x(t)$ should convolve in the same manner as if each of its IMFs $m_j(t)$ are convolved independently to $n_j(t)$ and then are summed over to the total convolved signal $y(t)$. Unfortunately, the non stationary character of IMFs cannot guarantee that the convolution of an IMF results in another IMF; thus, it may not be allowed to sum over n_j in order to form the total convolved signal y. If the present non stationarity is too severe in m_j or in the transfer function $F(\nu, t)$ the convolution of m_j cannot yield another IMF, because the convolution may introduce new extrema without additional zero crossings to the function which is not permitted in the definition of an IMF. In such a case it remains an open question whether n_j still are base functions of y. Certainly, the convolved signal y can not decompose into n_j if not all n_j qualify as IMFs. This restriction on the inverse to our theorem depends very much on the non stationary phase-time relation of signal and transfer function and may be discussed in detail in another work. Here, we only want to stress that the convolution results of IMFs do not need to be IMFs and may not always be summed up in order to compose a total convolved signal of a general time series. However, the theorem will always apply to a signal that is an IMF all by itself, even though the convolved result may or may not be an IMF after the convolution.

Later we will discuss an example for which we can add up the convolved

IMFs of a signal in order to get the total convolved signal. Moreover, in that example we will use the spectral information given by the IMFs of a signal and its convolution in order to estimate the system's transfer function. Therefore, we claim that the convolution of a signal's IMFs may well describe physical properties of signal convolution.

2.6.2 Lemma 1

This Lemma states that the instantaneous parameters of the IMFs at any time $t = t_0$ can be used as parameters of a sine curve to fully describe the IMF at that time. It provides a link between the HHT and the FT and can likely be used to find more properties of the HHT with the help of well-known properties of the FT, since a sine curve is the fundamental base of the FT. Note, that the FT is defined on an infinite time axis and that the time axis of this "virtual" sine curve is not equivalent to the one of the original signal but rather to an imaginary, infinite one, therefore even a piece of a continuous time signal can be described by these virtual sine curves and anything that applies to the entire virtual sine curve also applies to the IMF at time t_0. The Hilbert spectrum is the common mean to visualize the time-frequency behavior of an IMF and therefore, we suggest to refer to the Hilbert spectrum as two-dimensional, time-evolving Fourier spectrum.

2.7 On the Relation between HHT and FT

First of all, we would like to propose the term "time-varying Fourier Transform", which we define as the RHS of equation (2.13). Virtually every function $f(t) : t \in \mathbb{R} \rightarrow \mathbb{C}$ can be represented by a Fourier amplitude in this way but it is worth noting that the implications coming along with the definitions of an IMF, like that it must have a physically meaningful instantaneous frequency, give meaning to the time-varying Fourier amplitudes as physical representations of the IMF in the time and frequency domain.

When we apply Lemma 1 (equation 2.13) to equation (2.2), which is the original formulation of the HHT as taken from Huang et al. (1998) ,we find that the function $f(t)$ is represented by a series of time-varying Fourier

Transforms:

$$(f * \delta)(t_0) = \sum_j (m_j * \delta)(t_0) = \sum_j \left(\tilde{M}^j_{t_0} \cdot e^{2\pi i \hat{\nu}_j(t_0) t} * \delta \right)(t_0), \qquad (2.21)$$

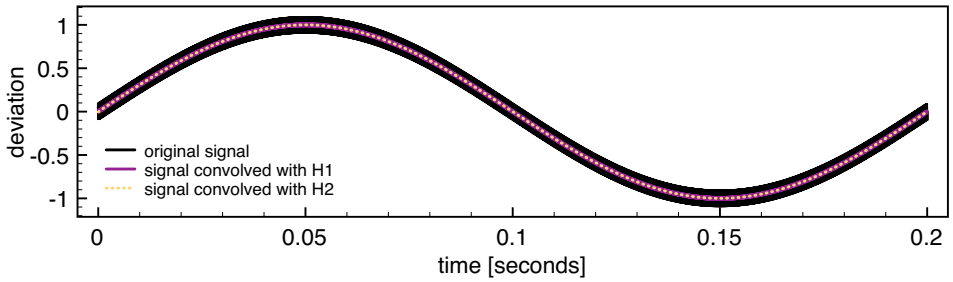
where $j \in \mathbb{N}^+$ is the order of the IMF and $\hat{\nu}_j$ assumes the spectral coordinates of the signal f. This formulation represents (in time slices) how the time-frequency information, obtained from the HHT, is commonly displayed: the Hilbert spectrum.

2.8 Two sandbox examples - Sine and Chirp

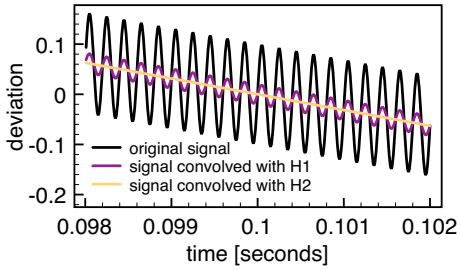
In order to demonstrate Theorem 1, we designed two synthetic time series, the first superposes two sine curves with distinct frequency values and the second is a chirped sine with a linearly increasing frequency. Both of these signals are then subjected to the convolution with a 1^{st}-order Butterworth low pass filter.

Stationary signal - two sine curves The first example is stationary and validates that Theorem 1 holds for the conventional convolution theorem, which states that the convolution of two time series is the inverse FT of the multiplication of the Fourier Transform of each time series. The superposition of two sine functions, here one with a frequency of 5 Hz and another with 5 kHz, is decomposed equally by the HHT and FT. Both Transforms find the exact same two oscillations with the constant amplitude from the signal. The signal is displayed as a black line in Figures 2.2a and 2.2b highlighting its slow and fast oscillation, respectively. In the same graphs are the convolution results with two distinct Butterworth filter. The filter are set up as 1^{st} order low pass filter with normalized cut-off frequencies at 0.05 and 0.005. The convolution with both filters has been applied in forward and reverse direction in order to achieve zero phase filter with amplitudes as shown in Figure 2.2c. Figures 2.2a and 2.2b show that the amplitude of the low frequency oscillation is not affected as both filter are in the pass band, whereas the high frequency part is damped according to the amplitude value of the corresponding frequency and filter. We tested three methods, FT based filtering, time series convolution filtering and the non stationary

(a) *The first synthetic test signal is a superposition of two sine functions with a frequency of* 5 Hz *and* 5 kHz*, respectively. The colored lines show the same signal filtered by* 1^{st} *order low pass Butterworth filter with normalized cut-off frequencies at* 0.05 *and* 0.005*.*



(b) *This zoom-in around* 0.1 s *of figure 2.2a highlights the details of the test signal.*



(c) *The amplitude spectra of the filter that are used here have distinct values at* 5 kHz*.*

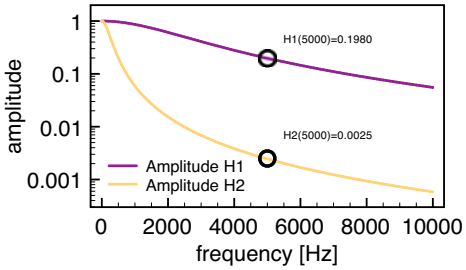


Figure 2.2: *This figure illustrates the first test signal represented by a superposition of two sine functions and its convolution with a* 1^{st} *order low pass filter. The signal is completely stationary and the convolution can be carried out in the time domain, with the FT or with our theorem and yields the exact same results.*

convolution theorem as presented in this work. All three methods yield exactly the same result.

Non stationary signal - chirped sine with linear frequency The second example on synthetic data is on a pure, non stationary signal in form of a chirped sine function with a linearly increasing frequency, which is per definition an IMF. The signal is plotted in Figure 2.3 as a black line with its frequency axis at the top and the time axis at the bottom. Note, that the very same signal illustrates Lemma 1 in Figure 2.1. To perform a convolution, we use the Butterworth filter with the cut-off frequency at 0.05 as described for the previous example. The filter is again set up as zero phase filter with the amplitude displayed as blue line in Figure 2.3. The convolution is carried out via the time series convolution and via the

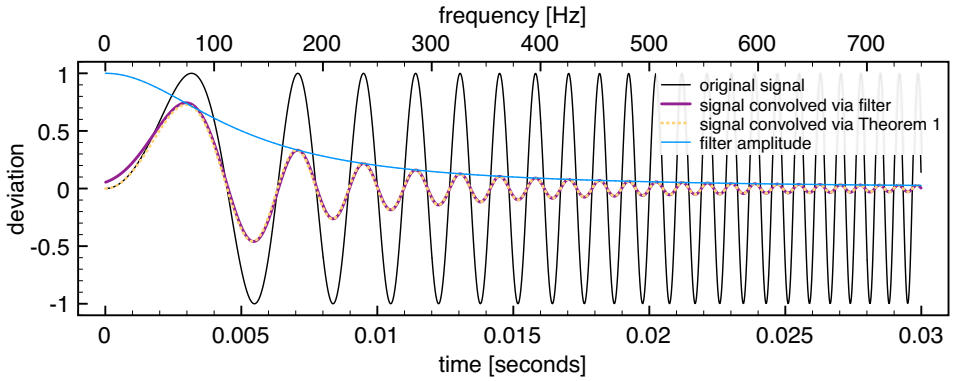


Figure 2.3: *The second test signal is a chirp function with a linearly increasing frequency over time and constant amplitude. This figure compares the convolution of the chirp with a* 1^{st} *order low pass Butterworth filter applied with two methods: time domain convolution and theorem 1 from this paper. Besides the numeric errors at the edge from the numeric convolution algorithm, both methods yield the same result.*

non stationary convolution theorem. Both results are displayed in Figure 2.3 and both are almost identical. The only difference is that the time series convolution algorithm can not deal with the beginning of the time series, since it is defined as a weighted sum that requires values around the location where it calculates the convolution but there exist no values lower than $t = 0$ so the algorithm assumes zero-padding and experiences "edge problems". The calculation based on the non stationary convolution theorem uses only local information and therefore is exact as well at the edges, where the frequency is very low and there is virtually no effect of the low pass filter to be expected. In fact, we tested different filter types (Chebyshev I and II, Elliptic and Bessel analog filter) with the order up to 10 and for several cut-off frequencies, and the convolution theorem presented here gave accurate results for any sampling rate as it is defined on the local, instantaneous parameters, whereas numeric filter procedures depend much on sampling as it relies on weighted sums.

2.9 A Real World Example - Magnetotellurics

In this section we want to present an application of Theorem 1. The authors' field of expertise lies in Magnetotelluric (MT) exploration and, hence, the theorem has been developed in light of processing magnetotelluric data. Magnetotelluric measurements log the natural variation of magnetic and

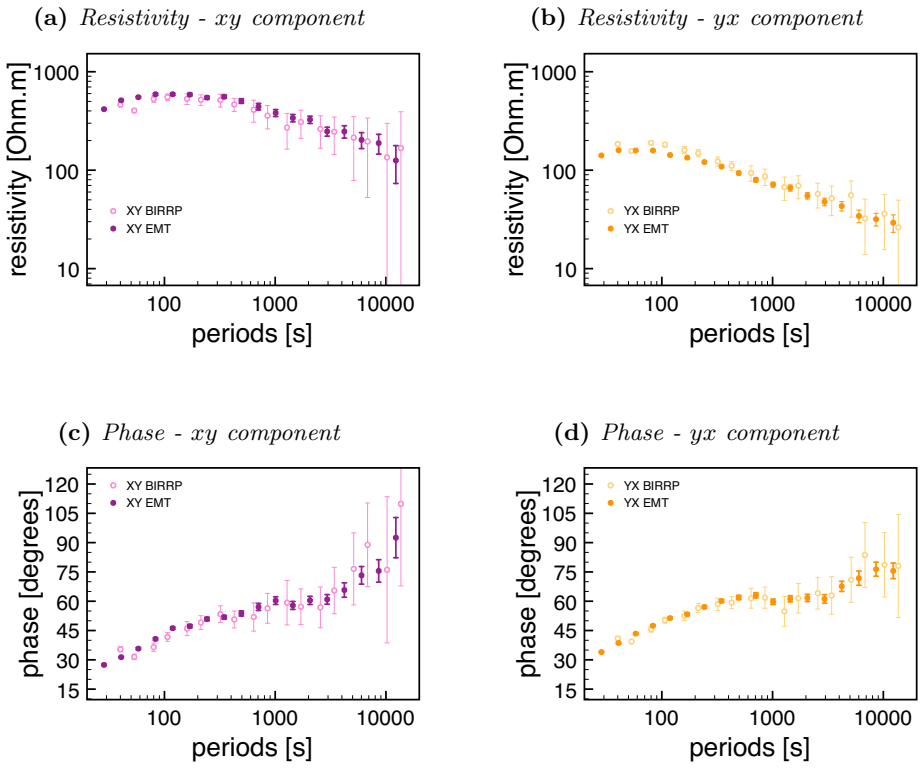


Figure 2.4: *Magnetotelluric data is given as an example application of the non stationary convolution theorem with the electric impedance as the Earth's system response function relating magnetic to electric fields. The main entries of the impedance tensor* $\mathbf{Z}$ *are displayed as amplitude and phase over a range of periods. It describes the subsurface electric conductivity structure and can be used to detect electric anomalies like phase boundaries, ore minerals or water deep inside the Earth.*

electrical (telluric) fields at the Earth's surface and these measured time series can be statistically analyzed to obtain the relative spectral relation of the electrical to the magnetic field Vozoff (1972). The subsurface conductivity structure enforces a particular distribution of underground currents, which alter the external natural electromagnetic field of the Earth and, therefore, it allows us to derive that structural information of the subsurface conductivity from the analysis of the electromagnetic field on the surface.

Knowing the spectra of the surface electrical field $\mathbf{E}(\omega)$ and the spectra of the surface magnetic field $\mathbf{B}(\omega)$, we can write the relation between the

horizontal spatial field components as

$$\begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix} = \begin{pmatrix} Z_{xx}(\omega) & Z_{xy}(\omega) \\ Z_{yx}(\omega) & Z_{yy}(\omega) \end{pmatrix} \cdot \begin{pmatrix} B_x(\omega) \\ B_y(\omega) \end{pmatrix} \tag{2.22}$$

where $\mathbf{Z}$ is the impedance which describes the subsurface conductivity volume for a given frequency ω. With respect to the work described in this paper, Z can be simply understood as the system response function of the Earth, $\mathbf{B}$ the input and $\mathbf{E}$ the output of the convolution. Electrical and magnetic fields are recorded as time series and need to be transformed into the frequency domain in order to solve for the impedance, because the impedance tensor is only defined in the frequency domain. Under ideal conditions, the electromagnetic field varies quasi stationary, meaning the spectral composition changes sufficiently slow that a windowed FT can be performed, but for sites closer to inhabited or industrial areas, cultural noise often affects the measurements severely. Cultural noise can be of any kind and is most often non stationary, therefore, measurements of $\mathbf{e}(t)$ and $\mathbf{b}(t)$ are often disturbed by non stationary variations, since the physical measurements contain both, natural signal and cultural noise.

Clearly, the problem described here is not exactly in the format of the theorem where we know input and system response and seek the output but it is similar; we do know input and output and need to find the system response by an optimization procedure. Thus, in this example we also have to assume Theorem 1 to hold in order to search for the optimal solution.

The algorithm that solves for this magnetotelluric data is too complex to be discussed here in detail and will be treated in its entirety in a different work, but we do need to stress that equation (2.22) states a multivariate optimization problem, which requires the use of a special EMD introduced by Rehman and Mandic (2010) and designed for multivariate data but for our purpose it performs an EMD no different than the univariate EMD, only that it ensures data channel correlation within the index of the IMFs (e.g. IMF number two of channel e_x will be at a similar time scale as any other channel's IMF number two).

Looking at the MT problem from the point of this work, $\mathbf{E} = \mathcal{F}(\mathbf{e})$ is the output or result of the convolution (electrical field spectra), $\mathbf{B} = \mathcal{F}(\mathbf{b})$ is the convolution input (magnetic field spectra) and the system response is

$\mathbf{Z}(\nu) = \mathcal{F}(\mathbf{z}(t))$ (Impedance) with the unknown temporal system response function $\mathbf{z}(t)$

$$\mathbf{E} = \mathbf{Z} \cdot \mathbf{B} \tag{2.23}$$

or in time domain

$$\mathbf{e} = \mathbf{z} * \mathbf{b}. \tag{2.24}$$

Both, $\mathbf{e}$ and $\mathbf{h}$, are then decomposed into their respective IMFs with the algorithm by Rehman and Mandic (2010) ,which ensures that for both signals the time scales remain correlated throughout the decomposition process. Theorem 1 suggests that

$$m_{\mathbf{e}}^j = \mathbf{z} * m_{\mathbf{b}}^j = \mathbf{Z} \cdot m_{\mathbf{b}}^j \tag{2.25}$$

with m^j being the respective IMFs for input and output.

Using equation (2.25) and a statistical optimization, we find an optimal solution for $\mathbf{Z}(\nu)$ for the instantaneous parameters given by $m_{\mathbf{b}}^j$ and $m_{\mathbf{e}}^j$. The results of the Impedance tensor for the test data set is presented in Figure 2.4 in dark color. The curves in bright color correspond to a Fourier-Transform-based algorithm processing the same data set.

Note, that this last example shown here also suggests that the convolved IMFs from a signal can be added to produce the complete convolved signal if the results still are IMFs. Here, we used measured time series of the signal and its convolution, $\mathbf{b}$ and $\mathbf{e}$, respectively, and carried out the convolution on their respective IMFs to find the system response function, thus it proofs that the sum of the convolved IMFs indeed reproduce the convolved measurement, even though the convolution has been carried out on each IMF independently. Our example only uses stationary transfer functions with a relatively constant phase, therefore it shall be said that other more complex cases may yield a different experience. For example, let there be a high degree of non stationarity in the phase-time relation between the IMF and the transfer function, then the IMF-system characteristics may preclude the convolution to result in another IMF, because some situations may alter the rotational sense of the phase and, hence, introduce new extrema without additional zero-crossings. However, the observation is interesting enough that we feel it would deserve a more thorough investigation in the next chapter.

2.10 Conclusion

- The convolution between an Intrinsic Mode Function and a temporal system response can be translated into a multiplication.
- The Hilbert-Huang spectra can be represented as a Fourier spectra with time varying complex amplitude.

Bibliography

Chave, A. D., Thomson, D. J., 2004. Bounded influence magnetotelluric response function estimation. Geophysical Journal International 157 (3), 988.

Chen, J., Heincke, B., Jegen, M., Moorkamp, M., 2012. Using empirical mode decomposition to process marine magnetotelluric data. Geophysical Journal International 190 (1), 293.

Egbert, G. D., 2002. Processing and interpretation of electromagnetic induction array data. Surveys in geophysics 23 (2-3), 207.

Garcia, X., Jones, A. G., 2008. Robust processing of magnetotelluric data in the AMT dead band using the continuous wavelet transform. Geophysics 73 (6), F223.

Huang, N., Wu, Z., Long, S., Arnold, K., Chen, X., 2009. On instantaneous frequency. Advances in Adaptive Data Analysis 1 (2), 177.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., Liu, H. H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 454 (1971), 903.

Jackson, L. P., Mound, J. E., 2010. Geomagnetic variation on decadal time scales: What can we learn from Empirical Mode Decomposition? Geophysical Research Letters 37 (14).

Margrave, G., 1998. Theory of nonstationary linear filtering in the Fourier domain with application to time-variant filtering. Geophysics 63 (1), 244.

Men-Tzung, L., Kun, H., Yanhui, L., C-K, P., Vera, N., 2008. Multimodal pressure-flow analysis: Application of Hilbert Huang transform in cerebral blood flow regulation. EURASIP journal on advances in signal processing 2008.

Rehman, N., Mandic, D. P., 2010. Multivariate empirical mode decomposition. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 466 (2117), 1291.

Smith, S. W., 1997. The scientist and engineer's guide to digital signal processing. California Technical Publishing, San Diego, CA, USA.

Vozoff, K., 1972. The magnetotelluric method in the exploration of sedimentary basins. Geophysics 37 (1), 98.

Wikipedia, 2011. Convolution — Wikipedia, The Free Encyclopedia. Accessed July 18th 2011.
URL http://en.wikipedia.org/wiki/Convolution

Zeiler, A., Faltermeier, R., Brawanski, A., Tomé, A. M., Puntonet, C. G., Górriz, J. M., Lang, E. W., 2011. Brain status data analysis by sliding EMD. In: Proceedings of the 4th international conference on Interplay between natural and artificial computation: new challenges on bio-inspired applications - Volume Part II. IWINAC'11. Springer-Verlag, Berlin, Heidelberg, p. 77.
URL http://dl.acm.org/citation.cfm?id=2009542.2009551

CHAPTER 3

Non Stationary Time Series Convolution: Frequency Shift Caused by Convolution

(submitted to Journal of Advances in Adaptive Data Analysis)

Maik Neukirch and Xavier Garcia

Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar, CSIC
Pg. Maritím de la Barceloneta 37-49, 08003 Barcelona, Spain

Abstract

By using the Hilbert-Huang Transform, a non stationary time series can be represented by a number of modes, which are complex time series with instantaneous amplitudes, phases and frequencies. Following the non stationary convolution theorem which allows to translate a convolution into a multiplication, we analyze the characteristics of a convolved time series and show that through convolution the instantaneous frequency may change. We quantify the frequency shift and argue that this difference greatly hampers any attempt to deconvolve non stationary signals.

3.1 Introduction

The Hilbert-Huang Transform (HHT, Huang et al., 1998) is a novel tool to analyze non stationary time series and describes them with their instantaneous, spectral information. HHT decomposes a time series into a number of zero-mean, oscillatory modes, called Intrinsic Mode Function (IMF), in

order to ensure existence of an interpretable analytic signal of each IMF and to express the analytic signal in terms of time series of the instantaneous parameters: amplitude, phase and frequency. The IMF reside in the time and frequency domain and are described by amplitude and phase as functions of time, where the time derivative of the phase constitutes the frequency.

Neukirch and Garcia (2013) present a non stationary convolution theorem that is similar to the convolution theorem for Fourier transform but that does not imply assumptions on the stationarity of the signal since it is based upon the definitions of the IMFs of the Hilbert-Huang Transform. However, they argue that such a non stationary convolution does not necessarily have an uniquely defined inverse, or a deconvolution operator resulting in the original signal, and we wish to continue this discussion focussing on some resulting implications for the deconvolution of non stationary signals.

Since the convolution of a non stationary time series with a response function in the time domain can be transformed into a basic algebraic formulation, in this work, we focus on the repercussions of a non stationary convolution by analyzing the instantaneous phase and its time derivative. Most notably, we find that there can be a frequency shift in the resulting signal with respect to the original signal depending on the degree of non stationarity. This finding may be important for non stationary time series, which are filtered by a system response for technical reasons, as it is often the case for physical measurements.

Table 3.1: *These conventions are used in the course of this chapter.*

(a) *Table of Functions*

t	time
$\phi(t)$	phase function
$\omega(t)=\dot{\phi}$	instantaneous angular frequency
$m(t)$	Intrinsic Mode Function (IMF)
$s(t)$	temporal system response
$S(\omega_m, t)$	spectral system response
$x(t)$	convolution result of s and m
$\Delta\omega(t)$	frequency shift

(b) *Table of Subscripts*

m	relates to original signal
s	relates to system response
x	relates to convolved signal
0	identifies amplitudes
j	order of IMF

3.2 HHT and Non Stationary Convolution

In the Hilbert-Huang Transform (Huang et al., 1998, 2009), the Intrinsic Mode Function (IMF) $m_j(t)$ of $x(t)$ are defined as

$$m_j(t) = m_{0,j}(t) \cdot e^{i\phi_j(t)}. \tag{3.1}$$

with $\phi_j(t) = \int_{-\infty}^{t} \omega_j(t')\, dt'$.

In essence, the HHT separates narrow-bandwidth Amplitude modulation (AM) $m_j(t)$ and Phase modulation (PM) $\phi_j(t)$ from the data and provides them in form of IMFs. This process is called Empirical Mode Decomposition (EMD). Since the phase of the signal is well defined, the instantaneous frequency can be derived from the phase by:

$$\omega_m(t) = \frac{d\phi_m(t)}{dt}. \tag{3.2}$$

Neukirch and Garcia (2013) show that the convolution of an IMF $m_j(t)$ with any time domain system response function $s(t)$ translates into a complex multiplication of the IMF with the frequency domain representation $S(\omega_m, t)$ of that response function.

$$m_j(t) * s(t) = m_j(t) \cdot S(\omega_m(t), t) \tag{3.3}$$

Equation (3.3) simplifies non stationary convolution drastically and one may infer that the same is true for the equally interesting deconvolution of signals. Unfortunately, the simplicity of the inverse operation for multiplication is misleading here and distracts from the fact that there can be involved quite different and unknown phase functions. Let us shed light on the problematic with an analysis of the phase function and allow us to ignore the amplitude functions in this work.

3.3 Phase Analysis of Convolved Time Series

We separate the complex values $m(t)$ and $S(\omega_m(t), t)$ of equation (3.3) for a single IMF in the amplitudes $m_0(t) \in \mathbb{R}^+$ of the IMF m and $S_0(\omega_m(t), t) \in \mathbb{R}^+$ of the response function S, and in the phases $\phi_m(t) \in \mathbb{R}$ of m and $\phi_s(\omega_m(t), t) \in \mathbb{R}$ of S. Note, that both, the amplitude and phase of the

response function, are functions of the instantaneous frequency $\omega_m(t) = \frac{d\phi_m(t)}{dt} = \dot{\phi}_m(t)$ of m and the time $t \in \mathbb{R}$. Then the convolution $x(t) = m(t) * s(t)$ writes:

$$x_0 \exp(i\phi_x) = m_0 \exp(i\phi_m) \cdot S_0\left(\dot{\phi}_m, t\right) \exp\left(i\phi_s(\dot{\phi}_m, t)\right) \tag{3.4}$$

with the following amplitude and phase functions:

$$x_0 = m_0 \cdot S_0\left(\dot{\phi}_m, t\right), \tag{3.5}$$

$$\phi_x = \phi_m + \phi_s\left(\dot{\phi}_m, t\right). \tag{3.6}$$

The observed amplitude x_0 is a function of amplitude and phase of the IMF m whereas the observed phase ϕ_x is independent of the amplitude, therefore, in this work, we restrict the analysis to the phase and leave the amplitudes for another time. The time derivative of equation (3.6) yields:

$$\dot{\phi}_x = \dot{\phi}_m + \frac{\partial\phi_s\left(\dot{\phi}_m, t\right)}{\partial t} + \frac{\partial\phi_s\left(\dot{\phi}_m, t\right)}{\partial\dot{\phi}_m}\ddot{\phi}_m. \tag{3.7}$$

Hence, we find that $\dot{\phi}_x = \dot{\phi}_m$ only for the case that either

1. the phase of the spectral system response ϕ_s is constant over time for a certain frequency $\dot{\phi}_m$ and one of both, ϕ_s is constant for a varying $\dot{\phi}_m$ or $\dot{\phi}_m$ is constant over time, or

2. the second two summands cancel each other.

In all other cases, $\dot{\phi}_x$ will differ from $\dot{\phi}_m$ by the frequency shift

$$\Delta\omega = \frac{\partial\phi_s\left(\dot{\phi}_m, t\right)}{\partial t} + \frac{\partial\phi_s\left(\dot{\phi}_m, t\right)}{\partial\dot{\phi}_m}\ddot{\phi}_m. \tag{3.8}$$

This observation tells us, that in a non stationary convolution a different instantaneous frequency may be observed than the one that the underlying process m had before it convolved with the response function s. The difference will depend on the nature of the response function (phase behavior over frequency and time) and on the signal itself (swiftness of changes in the instantaneous frequency). Furthermore, $x(t)$ will only retain the status of

qualifying as an IMF like m if and only if the frequency shift is larger than the negative instantaneous frequency of m:

$$\Delta\omega > -\dot{\phi}_m. \tag{3.9}$$

If not, the phase ϕ_x will run backwards introducing new extrema without zero crossings and prohibiting x to fall into the definition of an IMF. Even if the frequency shift allows the convolved signal to fall into the definition of an IMF its mere presence may easily cause mode mixing in a time signal that contains more than one IMF, since the instantaneous frequency of one IMF can become larger/smaller than its predecessor/successor.

3.4 A Representative Example

Figure 3.1 illustrates numerically the theoretical findings of the last section. There we define an IMF as a chirped function with a linearly increasing frequency and constant amplitude (see Figure 3.1 (1) to (3), blue dotted line) and a system response function with a quadratic frequency-phase relation and decreasing amplitude (see Figure 3.1 (1) to (3), red line). Naturally, the IMF is defined as a time series but since the time-frequency relation is linear in this example, we can equally use the abscissa for both, time and frequency. Then, the system response function is defined as a spectra, but again, since the time-frequency relation is linear and unique, the same reasoning applies for the abscissa for the system response function. The convolution of both is computed via equation (3.3) and displayed in Figure 3.1 (1) to (3) as purple dashed line. Note here, that in contrast to IMF and system response function the abscissa of the convolution represents the true time but the original frequency of the IMF and not the shifted frequency due to the convolution. The first plot illustrates how the varying amplitude of the system response function envelopes the convolution because the unitary chirp constitutes nothing to the multiplication of amplitudes in equation (3.3). Plot number two and three represent the addition of phases and frequency, respectively. And lastly, the frequency shift is plotted in the fourth diagram in form of the ratio between the frequency shift and the frequency of the IMF. The frequency shift in this example increases up to the value of the original frequency, effectively doubling the observed frequency from before to after the convolution.

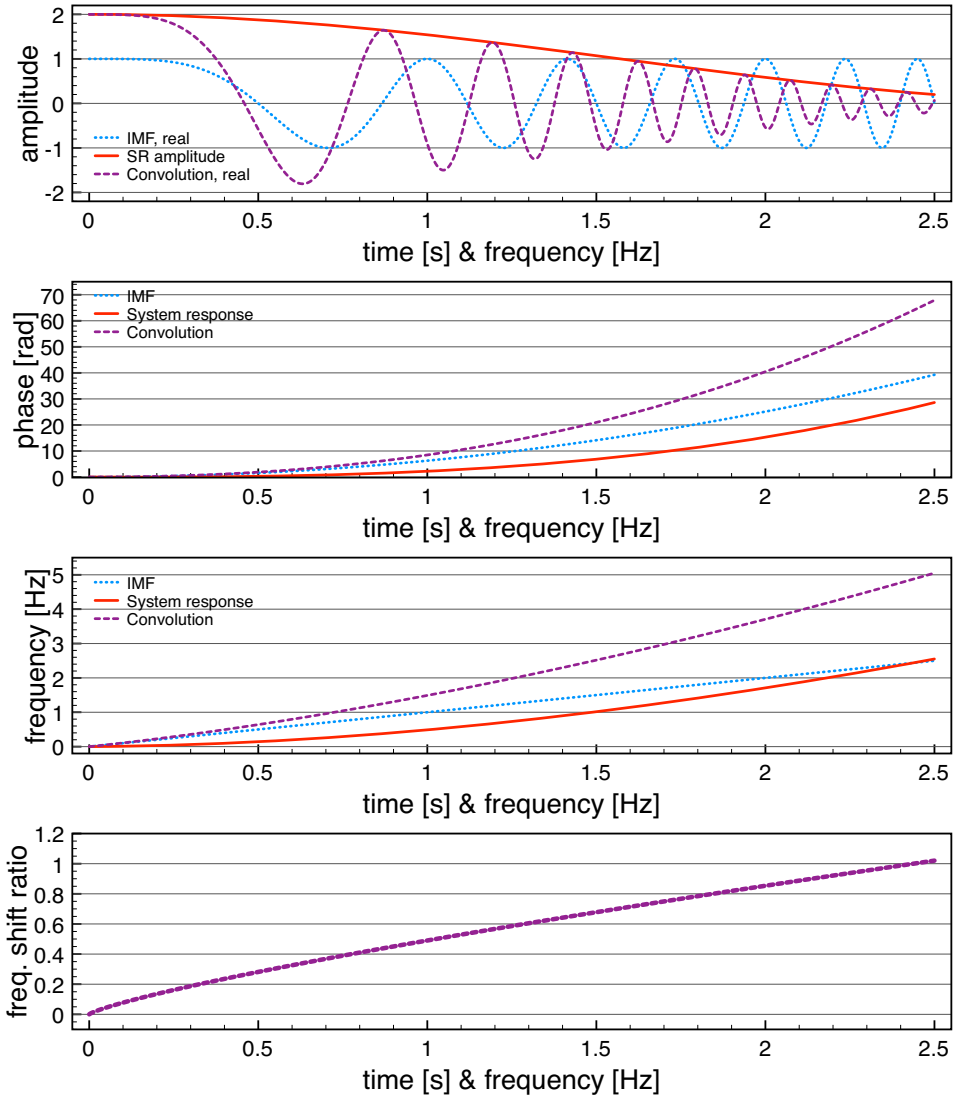


Figure 3.1: *From top to bottom: (1) a Intrinsic Mode Function (unitary chirp), the amplitude of a spectral system response and the convolution of IMF and SR, (2) phase functions, (3) phase gradients and (4) the ratio between the frequency shift and original frequency.*

3.5 Remarks on Deconvolution

For solving a non stationary deconvolution knowing only x and S, we would need to solve equation (3.7) for ϕ_m to recover the phase of m, ignoring the amplitudes for now. Clearly, if equation (3.8) is not equal to zero, solving equation (3.7) will be challenging and might only yield a solution via an iterated optimization algorithm. Furthermore, since x does not need to qualify for an IMF in theory, it may be impossible to find the correct ϕ_x directly from a convolved time series x by means of HHT.

3.6 Conclusion

Convolution of non stationary time series with a general system response function may alter the characteristic time scale of the time series and introduce a shift in the instantaneous frequency depending on the characteristics of the convolved system response function and the instantaneous frequency of the original signal. This frequency shift renders any deconvolution attempt difficult as such that no analytic solution exists but optimization may be successful.

Bibliography

Huang, N., Wu, Z., Long, S., Arnold, K., Chen, X., 2009. On instantaneous frequency. Advances in Adaptive Data Analysis 1 (2), 177.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., Liu, H. H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 454 (1971), 903.

Neukirch, M., Garcia, X., 2013. Non Stationary Time Series Convolution: On the Relation Between Hilbert-Huang and Fourier Transform. Advances in Adaptive Data Analysis. 5 (1).

CHAPTER 4

Non Stationary Magnetotelluric Data Processing With Instantaneous Parameters

published in Journal of Geophysical Research (21 March 2014)

Maik Neukirch and Xavier Garcia

Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar, CSIC
Pg. Maritím de la Barceloneta 37-49, 08003 Barcelona, Spain

Abstract

Non stationarity in electromagnetic data affects the computation of Fourier Spectra and therefore, the traditional estimation of the magnetotelluric (MT) transfer functions (TF). We provide a TF estimation scheme based on an emerging non linear, non stationary time series analysis tool, called Empirical Mode Decomposition (EMD) and show that this technique can handle non stationary effects with which traditional methods encounter difficulties.

In contrast to previous works that employ EMD for MT data processing, we argue the advantages of a multivariate decomposition, highlight the possibility to use instantaneous parameters and define the homogenization of frequency discrepancies between data channels. Our scheme uses the robust statistical estimation of transfer functions based on robust principal component analysis and a robust iteratively re-weighted least squares regression with a Huber weight function. The scheme can be applied with and without aid of any number of available remote reference stations. Uncertainties

are estimated by iterating the complete robust regression, including the robust weight computation, with a bootstrap routine.

We apply our scheme to synthetic and real data (southern Africa) with and without non stationary character and compare different processing techniques to the one presented here. As a conclusion, non stationary noise can heavily affect Fourier based MT data processing but the presented non stationary approach is nonetheless able to extract the impedances.

4.1 Introduction

Natural electromagnetic (EM) field variations are caused by two major working mechanisms: lightning activity at high frequencies ($> 8\,\mathrm{Hz}$) and magnetospheric currents excited by solar wind at low frequencies ($< 8\,\mathrm{Hz}$) (e.g. Garcia and Jones, 2002; Viljanen, 2012). Rakov and Uman (2007) summarize the electromagnetic lightning discharge to three modes: (a) fast and transient leader-return-stroke sequences, (b) slow and quasi stationary continuing currents and (c) perturbations and surges on the continuing currents. The longest lasting and most abundant in an electromagnetic time series measurements are the perturbed continuing currents, which may be viewed as being stationary on a section with some dynamic length confined by the recurrent transient strokes. Liu and Fujimoto (2011) conclude that the magnetospheric current is non linearly driven by the dynamic solar wind but behaves in a static manner for high magnetospheric pressure conditions. Both of these EM sources are naturally non stationary, since both, lightning strokes and magnetospheric pressure conditions, are very dynamic and thus, strictly limit the duration of any stationary electromagnetic signal.

Practitioners argue that the magnetotelluric (MT) signal is quasi stationary (stationary on reasonably long time windows) and, thus, justify the application of the windowed Fourier Transform. In practice this procedure works very well for data with high signal-to-noise ratios but frequently encounters problems in the presence of electromagnetic noise (clearly what is called noise here would include non stationary signal, cp. Junge (1996)). A concise treatise of sophisticated MT signal processing based on the Fourier Transform is given by Chave (2012) in which non stationarity is listed as

one of the problems that affect transfer function estimation.

For instance, if there would be a non stationary electric discharge, the window (data segment) of this event would not qualify as containing stationary data and such a window would have to be considered noise in a windowed Fourier transform algorithm. Moreover, noise sources (which do not include non stationary signal) can be of any kind and do not need to be quasi stationary (e.g. imagine a road with irregularly passing cars near the instruments, cp. Adam et al. (1986)). All non stationary noise sources (may also include non stationary signal) will affect the (windowed) Fourier Transform in unpredictable ways just because the data breaks the necessary assumption for the Fourier Transform at least in the relevant windows. This is not an issue when there are few affected windows, but it would become a problem when non stationary effects are frequent. A more concise treatment of electromagnetic noise and its characteristics is given by Szarka (1988) and Junge (1996), where both acknowledge non stationary noise sources and the aforementioned difficulties. Therefore, we argue, that even though the MT signal may behave sufficiently stationary, the contained noise in the data clearly cannot always be assumed quasi stationary as it would be required for the application of the Fourier Transform.

The isolation or separation of noise has been studied intensively since the introduction of the MT method and the two major noise counteract breakthroughs date back to the 1980s. Gamble et al. (1979) propose the use of a remote station to apply the technique of instrumental variables (Reiersøl, 1941) in order to drastically reduce bias by uncorrelated noise. Later Jones and Jödicke (1984), Egbert and Booker (1986) and Chave et al. (1987) advocate robust regression procedures for transfer function estimation to reduce the influence of unlikely but highly influential data points. Besides these two milestones, there has been much effort in reducing noise influence further by either trying to estimate and remove the noise directly (e.g. Egbert, 1997; Oettinger et al., 2001) in the frequency domain or by filtering, or extracting quiet data sections in the time domain by visual inspection (Garcia et al., 1997) and in the time-frequency domain (e.g. Weckmann et al., 2005, and references therein). The latter procedures are reported to be effective for particular data sets but require intense user attention and good, detailed knowledge about the data. Moreover, noise identification,

separation and/or removal is not always successful, sometimes practitioners encounter data from which it is seemingly impossible to extract reasonable transfer functions. This could be, partly due to the fact that EM data (the combination of signal and noise) is not as quasi stationary as required for the (windowed) Fourier Transform. A very simplistic example would be the presence of a spike in the data, which would compromise the particular data segment (or window) in which it is present. Clearly, the presence of a moderate number of spikes is easy to counteract (through interpolation, e.g. Jones et al., 1989; Junge, 1996), but we argue that the same principle applies to other non stationary effects which might not be as easily identified and mitigated.

Huang et al. (1998) introduce Empirical Mode Decomposition (EMD) in the framework of the Hilbert-Huang Transform (HHT), a novel time series analysis tool, which is data adaptive and suitable for non linear and non stationary data. The decomposition provides data modes (called Intrinsic Mode Functions, IMF) which are defined such that they can be represented as a single oscillation. Thus, Huang et al. (1998) argues that the definition of the IMF allow for a meaningful computation of its instantaneous parameters, like amplitude, phase and frequency, with the Hilbert Transform. In practice however, Huang et al. (2009) demonstrate that the Hilbert Transform often is numerically unstable and advocate a more practical routine to obtain the instantaneous parameters, which first separates amplitude and oscillation and then acquires the instantaneous phase by direct quadrature.

EMD has been tried and applied in several fields, including geophysics and the magnetotelluric method (Battista et al., 2007; Zhang et al., 2003; Cai et al., 2009; Chen and Jegen, 2008). In particular for MT, Cai et al. (2009) present how EMD could be used to separate obvious noise from the signal. Later, Cai (2013) attempts to substitute the Fourier Transform in favor of HHT in MT processing but the segmentation and averaging of data in order to construct marginal spectra (comparable to Fourier Spectra) is unnecessary and limits the potential strength of EMD. In the same year, Chen et al. (2012) present an estimation scheme for the transfer functions in MT data by using the instantaneous parameters (in contrast to marginal spectra). However, they conclude that the implementation of remote reference processing and robust statistics can further improve their approach, because

both techniques are very often required to estimate transfer functions from regular field data.

This work follows Chen et al. (2012) by using directly the instantaneous parameters obtained from EMD but in contrast to their work, here the multivariate variant of EMD by Rehman and Mandic (2010) is discussed and applied. Robust procedures are introduced to estimate instantaneous parameters and a data selection scheme is proposed to ensure independent data. For transfer function estimation a robust regression is advocated, which uses regressors defined by the two major robust principal components (Robust Principal Component Analysis described by Hubert et al. (2009)) of all remote data sets or for single site processing, all the available channels. Effectively this procedure excludes the site channels from the regressors if remote data is available in order to further reduce the risk of propagating correlated noise from between site channels into the principal components. Synthetic examples demonstrate the effect of non stationarity of the source on traditional processing schemes. Semi-synthetic signals, which consist of real signal and synthetic (non stationary) noise, present the effects due to non stationary of noise in real data and lastly, real world data sets verify the power of the algorithm for regular data and most notably, data in which non stationary noise is suspected. Additionally, a MatLab routine is presented, which creates (non) stationary synthetic MT data (or noise).

4.2 Outline of the EMT Algorithm

Figure 4.1 outlines the workflow chart of the algorithm that we have developed to process MT data using the EMD technique. We call the following scheme Empirical mode decomposition based Magnetotelluric Data Processing (EMT). Here we present the outline of the code, the following sections will describe each of the steps thoroughly:

1. Decompose time series with MEMD. The Multivariate Empirical Mode Decomposition (MEMD) method is used to decompose the multivariate data of all available channels (station and remotes) into oscillatory modes.

2. Compute instantaneous Parameters. Separate amplitude and oscillatory phase functions of the modes with Amplitude Phase Demodula-

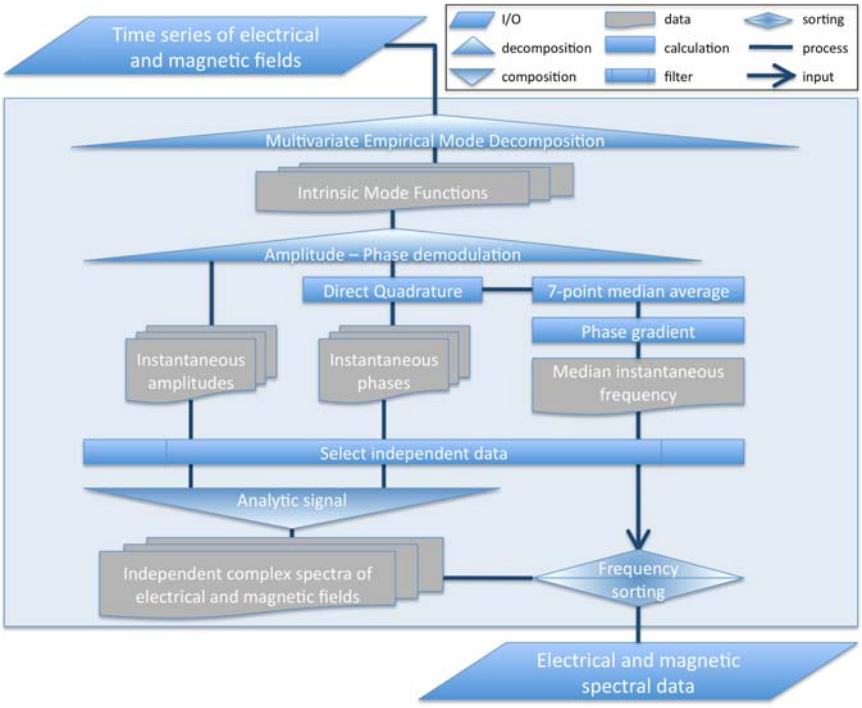


Figure 4.1: *EMT workflow chart to compute spectra.*

tion according to Huang et al. (2009). Generate the complex IMFs from amplitude and oscillatory phase for each channel to permit the computation of the instantaneous phase and the instantaneous frequency defined as time derivative of the phase.

3. Gather independent data points. We ensure linear independence of the data points by defining a time scale of data dependency.

4. Organize data in frequency domain. The data points are collected in wide bins, typically 5 to 10 bins per decade, ensuring enough estimates per decade and statistical stability of the impedance estimation for each bin by exploiting the fact that the MT transfer functions vary slowly with frequency.

5. Estimate transfer functions

   (a) Compute the two major robust principal components from data.

   (b) Robust regression of each channel on principal components.

   (c) Estimate confidence intervals by bootstrapping the regression.

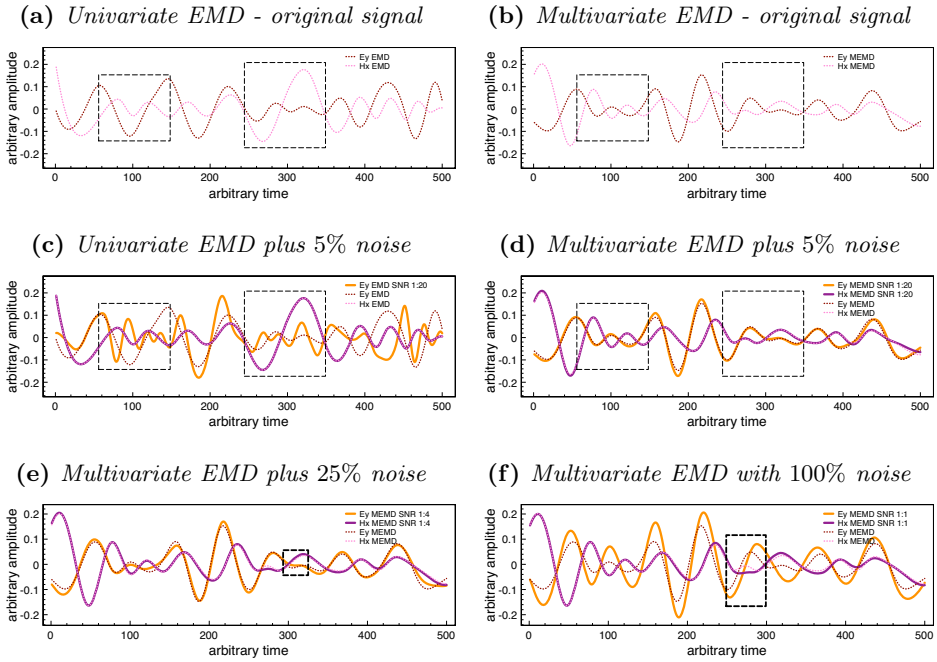


Figure 4.2: *The related modes for two channels of a short example signal from Southern Africa data are compared with respect to time scale consistency for different EMD algorithms and added gaussian noise variance. The presented mode is* #8 *out of* 11 *IMFs obtained from* $1,000$ *data points, however, note, that the actual mode number is irrelevant for the example being representative, since gaussian noise is equally present in all modes. The dashed boxes emphasize where there are major differences in the time scale between the two channels for the univariate case and then, for the multivariate algorithm, the boxes highlight the differences due to the added noise in* E_y*. Panel (e) shows that the time scale in the mode of* E_y *is affected by loosing one oscillation when adding* 25% *noise variance and in Panel (f) the noise variance added to* E_y *even begins to affect the time scale in the modes of* H_x*. Percentage of noise refers to the noise variance for gaussian noise relative to the average signal amplitude.*

4.3 Step 1: Multivariate Empirical Mode Decomposition

Huang et al. (1998) only present the application of their technique to univariate data, but MT data consists of at least four data channels, which depend on each other. Using a univariate EMD each signal is sifted and Hilbert Transformed independently, decreasing the likelihood that the signals remain dependent throughout all modes due to the possible channel-independent noise characteristics. For example, if there would be a high frequency noise in one channel which is absent in the other three, the first IMF of the first channel would contain that noise and start with the rest

of the signal from mode two on, whereas the other channels would contain signal from mode one on, resulting in different time scales for all modes (Figure 4.2). In this example, without any previous knowledge of this noise, the corresponding modes of different channels could never be used jointly for a linear least squares approach, since they do not contain the signal of the same frequency range. For that reason Chen and Jegen (2008) and Cai (2013) suggest to calculate the marginal spectra for each channel and use those in a similar manner as it would be done with the Fourier spectra. This approach has been shown to work very similarly to the usual Fourier approach and to provide novel noise control mechanism but does not take full advantage of the possibilities that EMD offers, namely the instantaneous parameters. Chen et al. (2012) circumvent this problem by only taking into account the data points of a time instant when they find a match for the Instantaneous frequency (IF) for each channel but in any mode. This procedure certainly solves the problem given in the simple example above, but a procedure that only uses data points where the IF matches (arguably within a certain limit) might run into problems as soon as the channels are more seriously distorted by noise and hence, the frequency computation for one or more channels is rather poor, ultimately decreasing the number of valid data points. In this section we discuss a multivariate decomposition algorithm that alleviates that problem by forcing all channels to decompose into correlated IMFs, or in other words, into IMFs of similar time scales, so that we can attribute one common IF value to all channels.

Rehman and Mandic (2010) developed a scheme to analyze multi-variate signals and compute IMFs of each of the signal's components such that they remain correlated in their time scale as much as possible. The algorithm is summarized in the following:

1. Project the multi-variate signal on an orthogonal n-dimensional hypersphere (basis functions defined by Hammersley sequences). The dimensions of the hypersphere represents different time scales much like the orthogonal sine functions in the Fourier Transform.

2. Locate the extrema of each projection (n projections in total).

3. Interpolate the multivariate signal by using the projection extrema locations for each dimension, to obtain a distinct upper and lower en-

velope of the multivariate signal for each dimension of the hypersphere.

4. Average the means of upper and lower envelope for each channel over all dimension.

5. Subtract the average envelope mean from the data and repeat to convergence to obtain the multivariate IMF.

MEMD provides a set of IMFs for each channel and retains the dependency in between those with respect to a most similar time scale (frequency) in all channels. It is also worth noting that for a source in EM field theory all components of the electric and magnetic field have the same frequencies present at all times, meaning that if there is an electric source of 10 Hz it will be accompanied by a magnetic field of 10 Hz. Therefore, MEMD does not at all introduce additional assumptions on the field components but rather ensures a fundamental property inherent in EM field theory for each IMF and thus, it decomposes the MT data into IMFs which can be conceived as independent data sets.

MEMD decomposes the data set into a number of IMFs, which have the information of instantaneous amplitude, phase and frequency at each time step and each IMF is a time series with a dynamic and locally narrow-banded, IF (Flandrin et al., 2004). Each IMF is inter-channel dependent and each time step fulfills the MT equation for its IF in the same way as narrow frequency-banded time series do (Berdichevskiy et al., 1973; Chen et al., 2012; Neukirch and Garcia, 2013). However, real data will always contain noise in all channels and the effect of the noise on the IMFs will largely depend on the (timely-) local signal-to-noise ratio and can easily span from subtle effects (e.g. some noise is present in one of many clean channels) to affecting the amplitude in (originally) clean channels (e.g. half the channels are corrupted by coherent noise and affect the clean ones) to even introduce false information in all channels (e.g. severe noise introduces new extrema). As an example for noise effects, panels (c) to (f) in Figure 4.2 illustrate data with added gaussian noise to a single channel.

This effect is conceptually comparable to how noise leaks in an ordinary Fourier Transform where the signal-to-noise ratio distorts the true (noise-free) spectra, but in the EMD case the effect is local and only affects the signal at some distance around the noise occurrence whereas the Fourier

Spectra is always affected in the whole segment, since it is formulated as an integral.

The Fourier transform is a univariate algorithm and noise in different channels cannot affect each other. Further, obviously non stationary effects can be reduced if the time series are broken in windows (windowed FFT). However, any non stationary noise in a data window will affect the entire Fourier spectrum of that window and often, robust procedures will drop exactly those spectra entirely regardless whether or not there shorter good data sections in that window. For an MEMD-based algorithm, the decision of excluding spectral information can be made for each individual time step instead of entire windows, if desired. However, care has to be taken, because even though spectral estimates are delivered at each time step, the real time frequency resolution is much lower and depends largely on the extrema in the corresponding IMF, but let us defer discussion on this matter to section 5: Independent Data Points.

The most important point, which can be observed in Figure 4.2 is that channels influence each other already during the MEMD. Apparently, noise spreads throughout channels and clean channels may be affected by noise, becoming biased. This noise spreading across channels occurs because the algorithm does not assume that one of the channels can be affected by noise while the others are not, it simply finds the best correlated signal for all modes and accounts the noise as a distortion of the total electromagnetic wave field. It becomes clear that this multivariate decomposition excels with the number of provided clean channels, which aid stabilizing the mode sifting and reduce noise in noisy channels by spreading it over all channels. For this reason the mode of the E_y component in panel (e) and (f) in Figure 4.2 appears to contain less noise than one would expect from adding 25% and 100% variance of gaussian noise, respectively. Naturally, it seems undesirable to spread noise from one channel to the others (which could be entirely avoided with a univariate EMD algorithm as Chen et al. (2012) propose), because we should preferably extract the best undistorted signal possible from our data. But, since MT is an intrinsically multivariate problem, we always need the information of all channels (of the site of interest) for the final TF estimation and the more data points we loose due to large deviations (in, for instance, the IF, which is a data selection criteria by Chen

et al. (2012)) in only one heavily distorted channel, the more difficult it will be to find an accurate transfer function. Using MEMD instead of EMD and enforcing a similar time scale on all channels, robustifies the decomposition procedure and yields more spectral data points which can be evaluated in the regression step at the cost of spreading the multivariate noise and thus increasing noise in some channels.

Usually a good portion of the noise is not correlated between the channels and therefore, affects the channels unequally, resulting in instantaneous parameters that depart from their correct values depending on the noise. Although this is certainly not appreciated for parameters like amplitude and phase, it does come in handy for the frequency computation, which we assume to be constant between the channels. Any deviation of the IF between channels must be due to any of the following:

1. the modes do not fulfill sufficiently the definition of IMFs (having a locally zero-mean),

2. the signal (channel) has been contaminated by noise (heavier contamination will result in larger deviations),

3. the frequency has been altered by non stationary convolution with the system response of the receiver.

The first problem is a very common issue for the first modes in EMD, since the data is always sampled on some rate and the location of the extrema in the data depend much on the sampling rate (in a real non stationary situation, the extrema can be anywhere in between the measurement directly before and directly after the recorded extrema). Routinely applied low-pass filters may alleviate much of this problem but the exact location of the true extrema is the most crucial information for calculating the instantaneous parameter from IMFs and this is usually not well defined for frequencies close to the sampling rate. However, in our experience the uncertainty on the location of the extrema only disperses the instantaneous parameters and does not usually introduce bias, the larger scatter in the regression is not problematic due to the larger number of data points for the higher frequencies in a data set. The second point is almost always an issue in MT and it is broadband, meaning it is found in all frequency ranges and thus,

all IMFs. But, since we know that the frequencies between the channels should be equal, we could use deviations between them as a selective quality marker or down weights in the later regression (similar to Chen et al. (2012)), however, we have not tested this idea in the present work. The last point is a rather new conclusion derived from the non stationary convolution theorem in Neukirch and Garcia (2013) and will be discussed thoroughly in another work. The problem only occurs for non stationary data convolved with a system response that varies over frequency, just like the instrument system responses for MT equipment usually do. It is not present during stationary sections and therefore, a minor issue for most MT data but fairly complicated to analyze, therefore it is out of the scope of this article. In any case, these disturbances are listed for sake of completeness as they will also affect amplitude and phase and thus, can introduce undesired bias to the transfer function estimation if not removed from the data or being accounted for.

Before we continue with the subject of IF, we need to focus on the recovery of the amplitude and phase from the IMFs in the following section.

4.4 Step 2: **Computing Instantaneous Parameters**

(Huang et al., 2009) thoroughly discuss the computation of instantaneous parameters from an IMF and Chen et al. (2012) continues the discussion with respect to an application in MT. We mostly follow their suggested instructions, since the IMFs of MEMD are methodically no different from the ones obtained from univariate EMD. Essentially, Huang et al. (2009) advise to separate amplitude and oscillatory phase with a procedure called Amplitude Phase Demodulation from the IMF. Then the instantaneous phase can be computed by direct quadrature from the separated, oscillatory phase function. In contrast to the original idea (Huang et al., 1998) of using the Hilbert Transform to obtain the phase, the direct quadrature method does not guarantee a strict analytic signal, but the routine performs well in practice and estimates the correct phase of the underlying signal more robust than the Hilbert Transform.

Focussing on the differences between this work and previous studies (Huang et al., 2009; Chen et al., 2012), examples of Instantaneous Parameters are given in Figure 4.3, which feature two modes of a short section of a

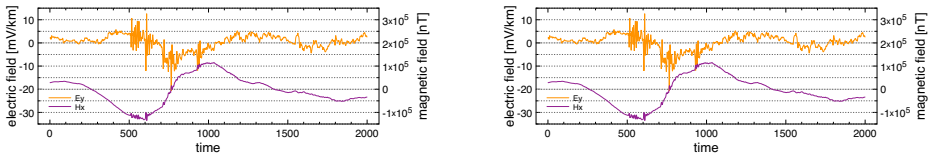
real data set from Southern Africa. Panels (d) and (e) display the Instantaneous amplitude (IA), panels (f) and (g) the Instantaneous phase (IP), and panels (h) and (i) the instantaneous frequency (IF).

By definition, the direct quadrature method divides by very small numbers at the extrema of the phase modulation function leading to numerical instability at those points, which additionally amplifies uncertainties and noise. Since we apply the direct quadrature method (Huang et al., 2009; Chen et al., 2012), the IP usually contain small numerical errors. Especially, these numerical instabilities have a great impact on the time derivation of the phase function and are depicted by simple poles in the IF (see panels (h) and (i) in Figure 4.3).

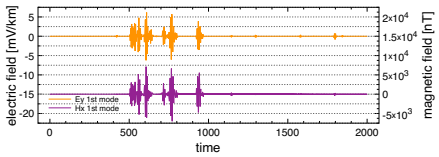
The poles are of first order and almost cancel each other out when summed over, which is why the phase function itself still looks smooth and the mean average over a sufficiently long time range is hardly affected. We found that a 7-point-median-average filter applied on the phase function before differentiating is a sufficient counteract and does not restrict the signal much more than the cubic spline interpolation already did during the sifting procedure, but produces a much more stable IF (cp. panels (j) and (k) in Figure 4.3).

In addition to the numerical instability associated with the direct quadrature method, the particular noise in each channel may cause differences in the IF between channels, where we would expect an electromagnetic field to have the very same frequency in all of its components (electric field and magnetic field) at a given time. However, we can use this fact to find a likely estimate for the common IF for all channel by using its mean or median average. The IF average is a physical meaningful representation of the true frequency of the electromagnetic signal (which is represented by all channels together) for a given time and mode. Heavy outliers from that mean average can be counteracted by using the median average and may be used to identify problematic data sections and can contribute to data quality control as mentioned in the section above. We found the median average to provide us with better estimates of the IF because of the frequent instabilities produced by the direct quadrature. These large irregularities in the IF usually do not occur in all channels at the same time, because of the impedance related phase shift between channels (refer to Huang et al. (2009) for a discussion

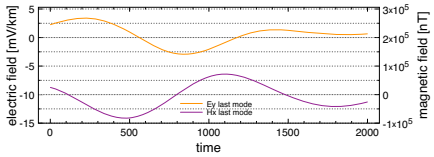
(a) *Data section example, plotted twice for comparison's sake*



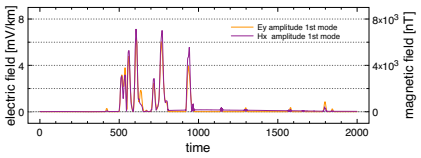
(b) *High: intrinsic mode function*



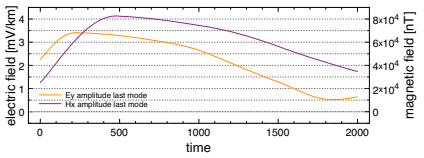
(c) *Low: intrinsic mode function*



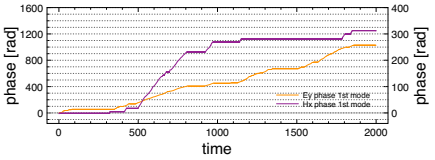
(d) *High: instantaneous amplitude*



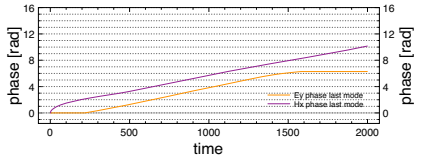
(e) *Low: instantaneous amplitude*



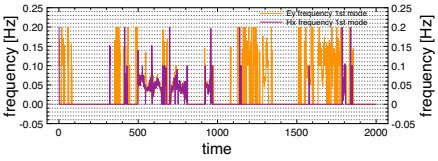
(f) *High: instantaneous phase*



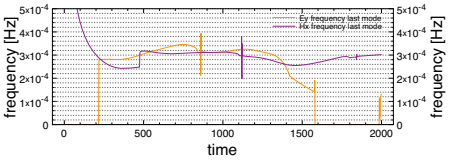
(g) *Low: instantaneous phase*



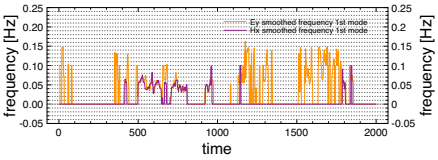
(h) *High: instantaneous frequency*



(i) *Low: instantaneous frequency*



(j) *High: smoothed instantaneous frequency*



(k) *Low: smoothed instantaneous frequency*

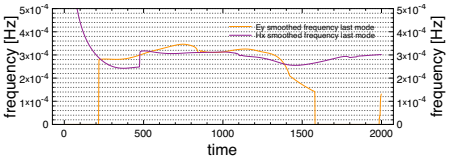


Figure 4.3: *Examples of instantaneous parameters are displayed. Left: high frequency mode (IMF #2). Right: low frequency mode (IMF #5). Both modes are extracted from the same data [site 072], however, panel (a) only displays the sum of the two modes in the spirit of comparability.*

on the nature of these numerical instabilities), but occur very frequently and thus, the median average compensates this problem, whereas the mean average would be drawn towards the outlier regularly.

All three instantaneous parameters: IA, IP, IF, and time form data quadruples and fully describe the original data. The IA and IP can be combined to form the representation of the complex spectra for a given time and frequency. Neukirch and Garcia (2013) lay out the fundamentals for signal-system convolution in a HHT context and provide proof that the convolution of complex, non stationary IMFs with a system response in time domain can be reformulated as the multiplication of the complex, non stationary IMFs with the system response function in the frequency domain. Therefore, when processing MT time series we can use the complex IMFs in the very same way as a Fourier spectra and carry out a statistical analysis in order to find the spectral physical relation between the channels, known as transfer functions.

For the sake of meaningful statistics with linear regression, we should try to ensure that (1) the data errors are independent (estimation accuracy) and that (2) the errors are identically distributed (accuracy and precision of estimation). Starting with the second, since we explicitly allow for non stationarity in our scheme, it is clear that our spectral data cannot be assumed to be drawn from a single distribution. The parameters of any distribution from which the data might start with will likely change during time, this is exactly what non stationarity means. However, the data decomposed by EMD is represented by oscillating modes which are bound to their definition and therefore always are locally zero mean functions. Thus, the definition of the IMFs ensures that the centre (location) of the data distribution is zero for whatever time-varying distribution it follows. Liu et al. (1988) discussed the importance of data being Identically and independently distributed (IID) in statistical system analysis with non parametric methods and came to the conclusion that when the bootstrap algorithm is used, the requirement of the data being IID can be somewhat relaxed, such that it is sufficient to ensure data point independency and that the underlying distributions of the data have a common location. They argue that the non parametric nature of the bootstrap algorithm includes a robustness towards dissimilar distributions in the data as long as the locations of the distributions are very similar (in

our case even equal).

The requirement that the data points are independent is more involved and has not yet been discussed in literature for an EMD setting, therefore, we dedicate the following section to that issue, then we will return to the discussion of the statistical analysis.

4.5 Step 3: **Independent Data Points**

Data independency is an important criteria for the statistical analysis described in this work, which, if left unconsidered, may bias accuracy and/or precision of the used methods. Besides, the understanding of the dependency of data points allows to draw inference about the time frequency resolution.

In our case we need to understand how data points interact and depend on each other in the total framework of HHT. Both, IA and IP derive from an analysis of the inner structure of the corresponding IMF. Each IMF is constructed by a loose sifting procedure based on the signal's extrema and guided by the required properties based on the IMF definition, a highly data adaptive procedure. The subsequent amplitude phase demodulation and the computation of the complex IMF do not rely any more on any data characteristics. For the demodulation the amplitude function and oscillatory phase function are already defined through the IMF definition and it only strips the two apparently different signals apart. Then, the direct quadrature uses the oscillatory phase function to recover locally the argument of the assumedly complex oscillation. The demodulation procedure is comparable to calculating the argument and absolute value of a complex number, which does not change or add any data dependencies, but only changes the way data are described (via the complex IMF which does not introduce information to the data). Therefore, we focus on the mode decomposition itself, when looking for dependencies in the data.

First of all, keep in mind, that per definition all IMFs of a signal are theoretically locally orthogonal, which implies that one mode to the next is linearly independent and uncorrelated. However, independency is by no means guaranteed along a mode. Since the IMFs are solely defined by a subset of points of the entire signal, namely the extrema, the IMF itself cannot have more degrees of freedom than number of extrema. All data points of

an IMF between two extrema usually share a third order interpolation polynomial, a cubic spline, which defines these data points based on the same set of extrema. Therefore, all these points between the same two extrema are dependent, whereas points that base on different sets of extrema are independent (even if just one extrema is different). Hence, it is important to only take into account one single data value for each span between two extrema to impose independency between the final data points. Naturally, the lack of independency in the definition of an IMF compromises greatly the time domain resolution suggested by IMFs but indicates that HHT does not provide a higher spectral resolution than what would be expected by the observed frequency (thus we still need a complete oscillation to meaningful describe spectral data). Furthermore, since the cubic spline requires the closest four extrema at each data point, the distance of influence of every extrema is about two full oscillations and represents some measure of time frequency resolution.

Since only one inter-extrema data point is independent, we have to pick the one which represents the entire range. Each data point should be equally valid since they are dependent. However, noise characteristics can make some points be a poorer choice than others (be reminded on the numeric instabilities due to the direct quadrature discussed in the section above). For the moment we have not designed a selection criteria based on data quality, so we simply take one point per half oscillation defined by the location of the extrema of the function

$$P = \sin\phi \cos\phi. \tag{4.1}$$

Since MT processing is multivariate, we suggest to use equation (4.1) with the integral of the common IF ω_c between the channels, thus

$$\phi(t) = \int_{-\infty}^{t} \omega_c(t')dt'. \tag{4.2}$$

This integral is basically the inverse of the time derivative of the phase used to obtain the IF in the first place, only that now the integrand is the common IF, which results in some sort of common phase for the EM data in equation (4.2), and provides an oscillatory function in equation (4.1) according to the

intrinsic oscillation of the EM data. The choice of this particular function is mainly because of its fairly random selection, if we would choose data points with certain properties (e.g. low/high amplitudes), we could easily introduce bias to the transfer functions, which is not the case for this general function. However, a more careful or sophisticated selection criteria (like a weighted average) for this point could help to reduce numeric or perhaps, even electromagnetic noise and could be discussed elsewhere.

4.6 Step 4: **Frequency Sorting**

As noted above, EMD results in a distinct frequency value for each channel. The average of those values for a given time and mode over all channels is a physical meaningful but biased representation of the true frequency of the electromagnetic signal (which is represented by all channels together). The bias should be lower for data points which have a similar frequency value and may even be considered for data quality control as we stated before. Keeping in mind that we use a common frequency function for all channels defined by the median average between them, in the following we will assume the median frequency as the common frequency between the data channels.

Remember that the instantaneous frequency (IF) is the time derivative of the phase of the complex IMF and does not yield equidistant (as for example the Fourier Transform) but rather continuous frequency values which vary with time and thus, along a mode. For this reason, it is unlikely that we can find two data points (each with two electrical (**e**) and two magnetic (**b**) components) with the very exact frequency value (ω_0), but this would be necessary in order to find a unique estimate for the transfer function tensor (**Z**), which is only defined at a constant frequency:

$$\begin{pmatrix} e_x(\omega_0,t) \\ e_y(\omega_0,t) \end{pmatrix} = \begin{pmatrix} Z_{xx}(\omega_0) & Z_{xy}(\omega_0) \\ Z_{yx}(\omega_0) & Z_{yy}(\omega_0) \end{pmatrix} \cdot \begin{pmatrix} h_x(\omega_0,t) \\ h_y(\omega_0,t) \end{pmatrix}. \tag{4.3}$$

Note, that this equation deviates from the traditional form as it includes time variance for the electromagnetic fields, since the complex IMFs of the data channels are still time series. A similar form of this time variant formula has been introduced by Berdichevskiy et al. (1973) and recycled by Chen et al. (2012), until this form has been proofed for the EMD context by Neukirch

and Garcia (2013). However, even though equation (4.3) suggests that the MT impedance equation is valid at each time instant for the IMFs of the electromagnetic field, the impedance itself cannot be solved for unless there are at least two independent measurements for the same frequency value. But, since the electrical impedance only changes smoothly with frequency (Cagniard, 1953), we can group similar frequency values to increase the amount of measurements available around a certain centre frequency. For this procedure, we select the independent data points based on equation (4.1) and arrange them according to the common IF, omitting time dependency of the data by considering the time axis rather as index for measurements than physical time. The data reorganization in these frequency bins follows the proposed method by Chen et al. (2012), only that we do not allow IMF mixture for the reasons discussed in section 4.3.

Following this reorganization, we form an over-determined system of equations that we can solve for the transfer function tensor at distinct frequency values. The estimation procedure is a bootstrapped, robust principal component regression and will be discussed in detail in the following section.

4.7 Step 5: **Robust Principal Component Regression**

Egbert (1997) shows that MT sources are well described by two electromagnetic field polarizations. Practically, this means that the entire data vector space of all channels in a data set can be represented by the combination of two polarization vectors. Theoretically, the high dimensional data (electric, magnetic and all remote channels) can be described by a fundamental two-dimensional polarization space that contains all the variance of the data. Such a reduction of dimensionality of data vectors can be achieved by a (robust) principal component analysis (PCA), which provides the inherent components of the data vector, ordered by its eigenvalues. The two most dominant Principal components (PC) are the magnetotelluric source vectors since they should be present in all channels and contribute most to the variance of the data (cf. Egbert, 1997). However, in practice MT data are often contaminated by noise and source field effects, which limit this procedure (Egbert, 1997, 2002; Smirnov and Egbert, 2012) such that there are more than two dominant eigenvalues which contain a mixture of

source polarization vectors and correlated noise. In order to separate the dominant principal components in such cases a much more sophisticated multi site analysis is required and described by Smirnov and Egbert (2012), which should be followed for data sets with coherent noise contamination, however the discussion or incorporation of such an analysis is beyond the scope of this work, although, it could be implemented in our algorithm if desired. For this work we assume that the first two principal components are a sufficiently good estimate of the MT source polarization vectors but we restrict the data used for the PCA to remote channels only, if at least two are available. If not, the site channels can be used as usual.

A robust principal component analysis tool is provided by Hubert and Verboven (2003) within the frame of the free Library for Robust Analysis (LIBRA) package (Verboven and Hubert, 2005) for MatLab and referred to as `robpca.m`. Smirnov and Egbert (2012) compare this code for consideration of its usage in the aforementioned multisite analysis of MT data and acknowledge its power, but prefer a self-made solution for its flexibility. Since we do not attempt a multisite data analysis and assume two principal components to be sufficient, the algorithm from LIBRA appears the most reasonable solution at this stage of our algorithm.

After the computation of the two dominant PCs (say $\mathbf{r} = (r_1, r_2)$), we formulate four (or five if the vertical magnetic field is provided) two-variate regression problems in order to separately deal with the noise distributions in each data channel. Assume the North-South electric field e_1, the East-West electric field e_2, the North-South magnetic field b_1, the East-West magnetic field b_2 and, if available, the vertical magnetic field b_3 as data channels. For each data channel $\mathbf{x}$, the regression writes in a matrix notation

$$\mathbf{x} = \mathbf{r} \cdot \mathbf{R}_\mathbf{x} + \boldsymbol{\sigma}_\mathbf{x} \tag{4.4a}$$

$$\mathbf{Z} = (\mathbf{R}_{e_1}; \mathbf{R}_{e_2}) \cdot inv(\mathbf{R}_{b_1}; \mathbf{R}_{b_2}) \tag{4.4b}$$

$$\mathbf{T} = (\mathbf{R}_{b_3}) \cdot inv(\mathbf{R}_{b_1}; \mathbf{R}_{b_2}). \tag{4.4c}$$

$\mathbf{R}_\mathbf{x}$ is a row vector and denotes the regression parameter for vector $\mathbf{x}$ on the PCs $\mathbf{r}$, $\boldsymbol{\sigma}_\mathbf{x}$ represents the noise in $\mathbf{x}$ and $(\cdot\,;\cdot)$ refers to the column wise combination of two row vectors. The tensor $\mathbf{Z}$ is the electric impedance according to equation (4.3) and $\mathbf{T}$ the tipper, which is the magnetic transfer

function between the horizontal and vertical magnetic fields. The $inv(\ )$ operator produces the inverse matrix and the dot operator denotes the matrix multiplication. The formulation of the regression is slightly different from the one that is usually applied in MT, but not as much as it seems at first. Actually, for an ordinary least squares solution for, say $\mathbf{Z}$, this formulation yields exactly equation (4.3), which is the original formulation if time only indicates measurements. The idea behind this alternative formulation is, that the regressors $\mathbf{r}$ result from a robust statistical procedure, which describe a part of the variance in the data, and thus, do not contain outliers that deviate from the dominant inherent information provided by the data. Originally, the regression is carried out on data channels directly, which firstly, may bias the result by correlated noise and secondly may contain highly influential outliers as discussed by Chave and Thomson (2004) and Chave (2012). In our solution, influential outliers in the regressor are unlikely unless they represent a repeated feature in most channels, which would only be the case for correlated noise, but if correlated noise would be present, only a careful and sophisticated data analysis (e.g. a multi site analysis (Smirnov and Egbert, 2012) or noise identification (Weckmann et al., 2005)) can mitigate the influence of this kind of noise. In any way, such noise would be removed, if possible, before any regression attempt and thus, validates the assumption that such noise is not present in the regressors.

We divide the total regression problem in sub steps to separate the expected noise from all channels (compare equations (4.4)) in order to avoid a direct effect of coherent noise between channels. The regressions themselves are carried out robustly with an iteratively re-weighted Huber weight function by calling the MatLab intrinsic function `robustfit.m`, only specifying the desired weight function. Other weight functions are possible (refer to the MatLab documentation for a discussion on the options) and we experimented with each one, concluding that the results obtained with the Huber weight function were most accurate and precise. The robust regression only accounts for outliers in the data channels and not for any possible outlier in the PCs, which have been computed robustly in the PCA and have disregarded bad influence points already.

EMT bootstraps the entire robust regression step in order to compute a data dependent distribution of impedance values and estimate the data in-

trinsic errors of the procedure. Furthermore, as discussed before, the bootstrap operation also relaxes the requirement for statistical regressions for which data should be identically distributed and therefore reflects more reliably the estimates in case of non stationary data. Empirically, we found $1,000$ iterations a sufficient trade off between accuracy and computation time to estimate the uncertainty of our results.

4.8 SynDat - Computing (Non) Stationary Synthetic Data

Availability of synthetic data is fundamental for hypothesis testing in many areas of applied science, since it offers a simple and noise-free mean of acquiring test data, which could be expensive, difficult or time consuming in the laboratory or in the field, and it allows to design easily custom-made properties of test data, which often help to spotlight both, important problems and findings in a hypothesis.

We use the MatLab program SynDat to generate (non) stationary synthetic data in the course of this work. It allows to define freely the frequency and amplitude time series of numerical remote channels (as given in equation (4.6)), which are used to compute synthetic MT data (as given in equation (4.5)) by means of the non stationary convolution theorem according to Neukirch and Garcia (2013). Additionally, the synthetic data can be modeled for any impedance by importing the respective transfer function (TF) from files of the EDI format, or be computed for the impedance of a one-dimensional conductivity model (Keller and Frischknecht, 1966).

This program is freely available upon request to the author.

4.9 Example Datasets

In this section we compare the processing scheme outlined above with the state-of-the-art processing algorithms *Bounded Influence Robust Reference Processing (BIRRP)* by Chave and Thomson (2004), *Electromagnetic Transfer Function (EMTF)* by Egbert (1997) and the LIMS data acquisition processing algorithm (LIMS) by Jones and Jödicke (1984). The four algorithms are applied to a number of synthetic, half synthetic/half real and real data sets. We start with two synthetic data sets, one based on white noise as source signal and the other on a purely non stationary wave form. These

two examples will shed light on the differences between a quasi stationary and non stationary processing scheme. Then we present two examples of real-world data (Southern Africa (Evans et al., 2011)) to compare performance of the processing algorithms on a natural problem. In order to illustrate the effect of non stationary disturbances in the data, we add the electric fields from the second synthetic test to the electric fields of fairly good real data, which effectively introduces non stationary noise in the electric fields but leaves the magnetic fields completely unaffected. Lastly, we process one real data set in which non stationary noise sources are known to interfere and demonstrate the supremacy of *EMT* in such a situation. All plots contain the data and estimated confidence intervals for 95% of the data (doubled standard deviation).

4.9.1 Synthetic Data based on White Noise

Using the SynDat program to create MT synthetic data, in this first example, we prepared two complex remote spectra $\mathbf{s} = (s_x, s_y)$ from white noise

$$s_x = n_x^{w,\text{real}} + i \cdot n_x^{w,\text{imag}} \qquad \text{and} \qquad s_y = n_y^{w,\text{real}} + i \cdot n_y^{w,\text{imag}}.$$

The number of frequencies is $N_f = 12,500$ with a step size of $df = .25$ Hz to obtain a time series of $25,000$ samples with a sampling rate of $dt = 4\ s$. The data $\mathbf{E} = (E_x, E_y)$ and $\mathbf{B} = (B_x, B_y)$ is computed in the frequency domain from $\mathbf{s} = (s_x, s_y)$ by:

$$\mathbf{E} = \mathbf{s} \cdot \mathbf{Z}^{\frac{1}{2}} \qquad \text{and} \qquad \mathbf{B} = \mathbf{s} \cdot inv(\mathbf{Z}^{\frac{1}{2}}) \tag{4.5}$$

with $\mathbf{Z}^{\frac{1}{2}}$ defined as the principal square root of any matrix $\mathbf{Z}$ such that $\mathbf{Z}^{\frac{1}{2}} \cdot \mathbf{Z}^{\frac{1}{2}} = \mathbf{Z}$ in order to fulfill $\mathbf{E} = \mathbf{B} \cdot \mathbf{Z}$. Let us assume the following impedance model

$$\mathbf{Z} = \begin{pmatrix} 0 & 3000 \\ 1000 & 0 \end{pmatrix} * \exp\left(i \begin{bmatrix} 0 & -\frac{\pi}{4} \\ \frac{\pi}{4} & -\pi \end{bmatrix} \right).$$

Note, that here the asterisk operator denotes the element wise multiplication of the matrices, and exp() refers to the exponential of the matrix, element-by-element. The results of processing this synthetic data are displayed in

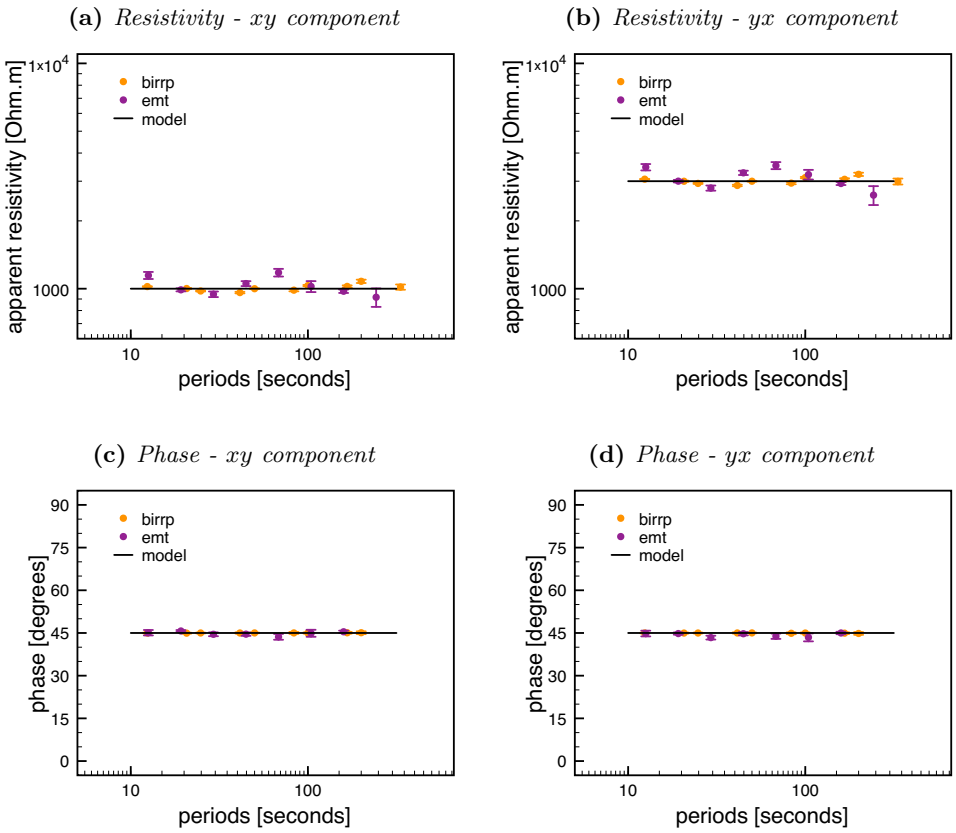


Figure 4.4: *The* EMT *algorithm is compared with* BIRRP *for synthetic stationary data based on spectral white noise as source signal. The assumed homogenous impedance model is defined as* $Z_{xy} = 10^3 \cdot \exp(i\frac{\pi}{4})$ *and* $Z_{yx} = 3 \cdot 10^3 \cdot \exp(-i\frac{\pi}{4})$ *and is plotted as a black line.*

Figure 4.4 for the processing algorithms *BIRRP* and *EMT*. Both algorithms resemble the model fairly well but *BIRRP* has the edge. We explain this by the fact that this synthetic source does not have any wave form and therefore the (M)EMD algorithm struggles to find correlated modes which it can relate to each other. On the other hand, *BIRRP* uses the spectral characteristics of the time series which are, per source definition, well defined.

4.9.2 Synthetic Data based on a Chirp

In order to clearly demonstrate the difference of the processing schemes, the synthetic data discussed here is completely non stationary. Again using SynDat, we define each of two orthogonal magnetic source fields $\mathbf{s} = (s_x, s_y)$ by

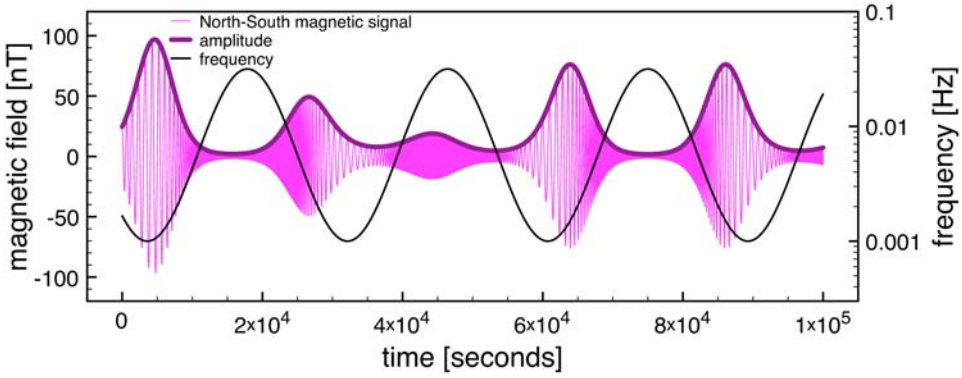


Figure 4.5: *The chirp signal shown here is used as non stationary synthetic signal. The North - South magnetic field component is illustrated with its non stationary amplitude and frequency. Note, that the frequency of the computed signal ranges from* 1 *to* 30 mHz *and therefore, if data is processed outside of that range it can only contain numerical noise.*

a logarithmic frequency oscillation f and a logarithmic amplitude oscillation $\mathbf{a} = (a_x, a_y)$:

$$\begin{aligned} \mathbf{s} &= \mathfrak{Re}\left(\mathbf{a} \cdot \exp(i \int f dt)\right), \\ \log(f) &= A + B\cos(F_f t), \\ \log(\mathbf{a}) &= \mathbf{C} + \mathbf{D}\sin(F_{\mathbf{a}} t)). \end{aligned} \tag{4.6}$$

The parameters A, B, $\mathbf{C} = (C_x, C_y)$ and $\mathbf{D} = (D_x, D_y)$ define frequency and amplitude range and the parameters F_f and $F_{\mathbf{a}}$ control the degree of non stationarity by the oscillation rate of f and $\mathbf{a}$, respectively. The time axis t is sampled at a rate of 4 s for a total length of $100,000$ s or $25,000$ samples. Figure 4.5 displays the magnetic North - South component with its respective amplitude and frequency function. By design the signal is a locally zero mean function to ensure that it complies with the definition of the IMFs, even without the need to apply (M)EMD. As in the last example, the impedance $\mathbf{Z}$ is assumed to be homogenous with

$$\mathbf{Z} = \begin{pmatrix} 10 & 3000 \\ 1000 & 30 \end{pmatrix} * \exp\left(i \begin{bmatrix} \frac{\pi}{4} & -\frac{\pi}{4} \\ \frac{\pi}{4} & -\frac{\pi}{4} \end{bmatrix}\right).$$

The electric and magnetic fields are computed according to equation (4.5). Figure 4.6 compares the results of processing the electric and magnetic data with *BIRRP* and *EMT*. Both algorithms are called with their respective default parameters to compare the results assuming no a priori knowledge

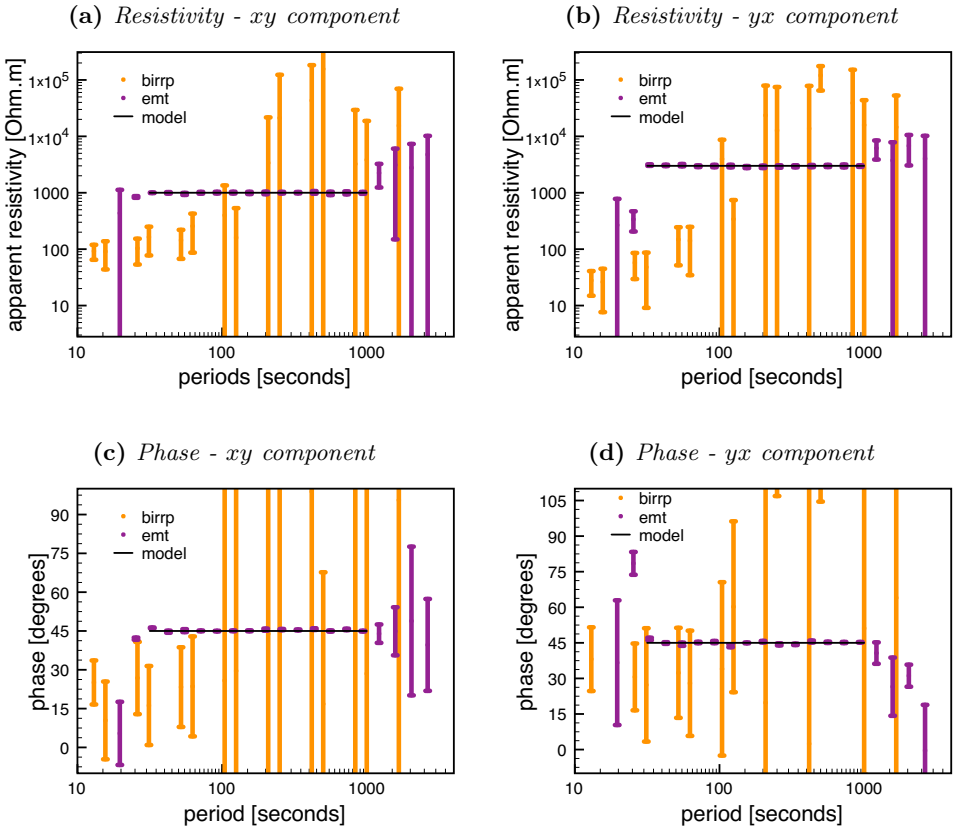


Figure 4.6: *The* EMT *algorithm is compared with* BIRRP *for synthetic non stationary data based on a chirp signal as shown in Figure 4.5. The assumed homogeneous impedance model is defined as* $Z_{xy} = 10^3 \cdot \exp(i\frac{\pi}{4})$ *and* $Z_{yx} = 3 \cdot 10^3 \cdot \exp(-i\frac{\pi}{4})$ *and is plotted as a black line. Note, that the frequency of the computed signal ranges from* 1 *to* 30 mHz *and therefore, the processed data outside that range can only contain noise, however, inside the range only* EMT *is successful in recovering the model.*

about the data. *EMT* successfully recovers the model in the frequency range of the computed data but *BIRRP* fails processing the data, which can only be addressed to the strict non stationarity of the signal and exemplifies that Fourier Transform based methods are not suitable for strictly non stationary signals, even those that apply the windowed Fourier Transform. However, this example is not a fair comparison as this kind of signal is not natural and treatises of the physics of typical MT sources (see Rakov and Uman, 2007; Liu and Fujimoto, 2011) suggest that the sources are not as non stationary as this example for an extended period of time and instead, can be treated as quasi stationary. This example serves only illustrative purposes and is

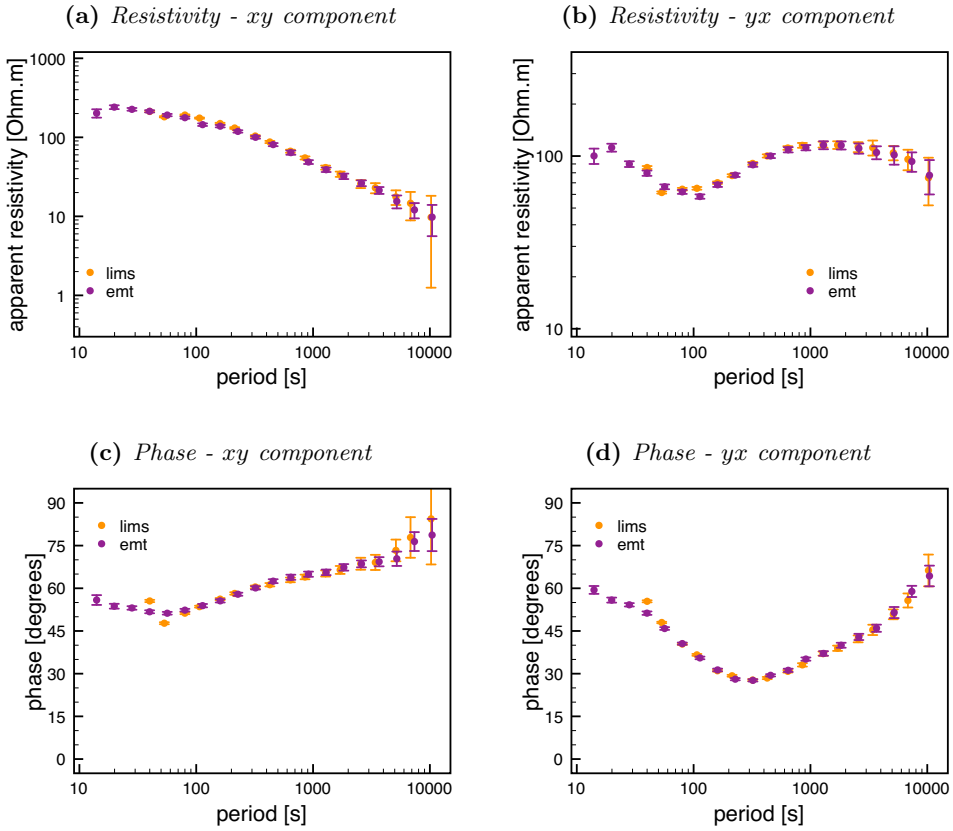


Figure 4.7: *Good example data from Southern Africa (site 072). The LIMS results are the original results from the SAMTEX experiment (Evans et al., 2011).*

designed to highlight the strength of EMD, to expose the weakness of the Fourier methods and most of all, to demonstrate clearly how non stationarity appears in the results of Fourier methods. In the following section, we present more practical examples using real data.

4.9.3 Fairly Good Real Data from Southern Africa

Now, let us compare the algorithms using three real data sets from Southern Africa which correspond to the sites 027 and 072 with site 045 as remote reference, and 042 with 027 as remote reference for long period data and 043 as remote reference for the short period data (Evans et al., 2011). The first two time series have approximately $500,000$ samples on a sampling rate of 5 s and we only consider the horizontal magnetic fields as remote reference,

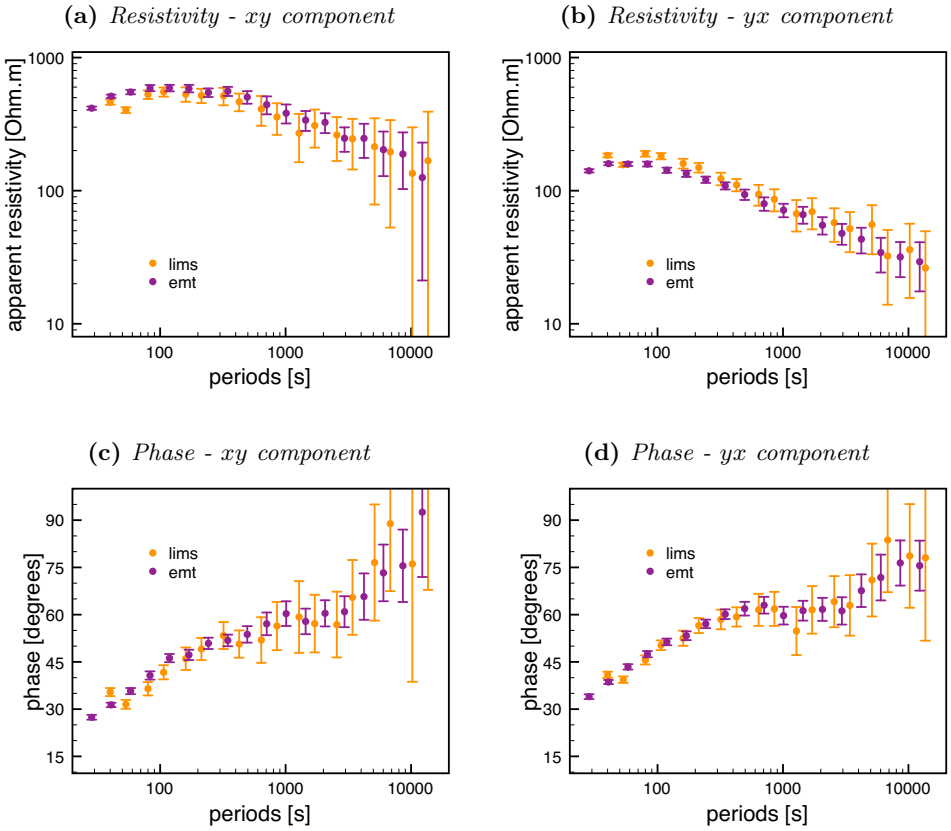


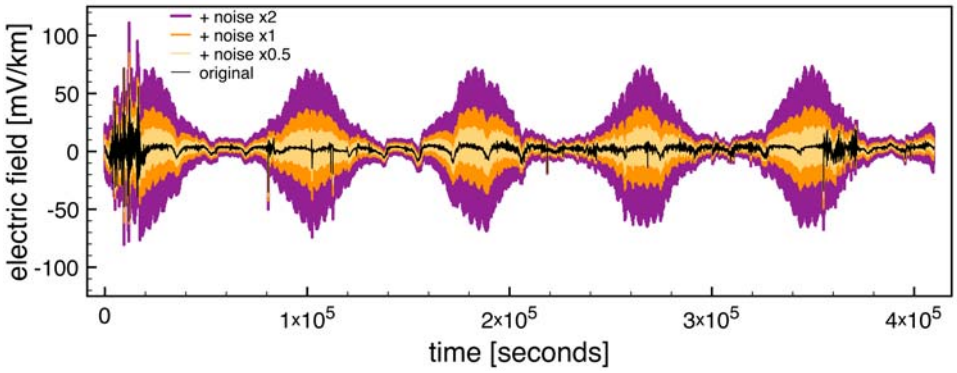
Figure 4.8: *Fair example data from Southern Africa (site 027). The LIMS results are the original results from the SAMTEX experiment (Evans et al., 2011).*

since they have proven to be sufficiently efficient in removing coherent noise from the local fields. The last example has up to 2 million samples for the high frequencies at 2560 Hz and around 500,000 samples on a sampling rate of 5 s. The high frequency data only has one remote reference site and for the long periods we selected the best suitable one.

The first example (site 072) is considered good for MT processing purposes when processed with the available remote magnetic channels (of site 045). Figure 4.7 displays the processing results for the *LIMS processing algorithm* (original results) and *EMT* and both algorithms agree very well.

The second example (site 027) contains somewhat more noise even when processed with the available remote magnetic channels (site 045). Figure 4.8 compares the *LIMS processing algorithm* (original results) with *EMT* and

(a) *Electric Field with Synthetic Noise - North/South component*



(b) *Synthetic Noise Section*

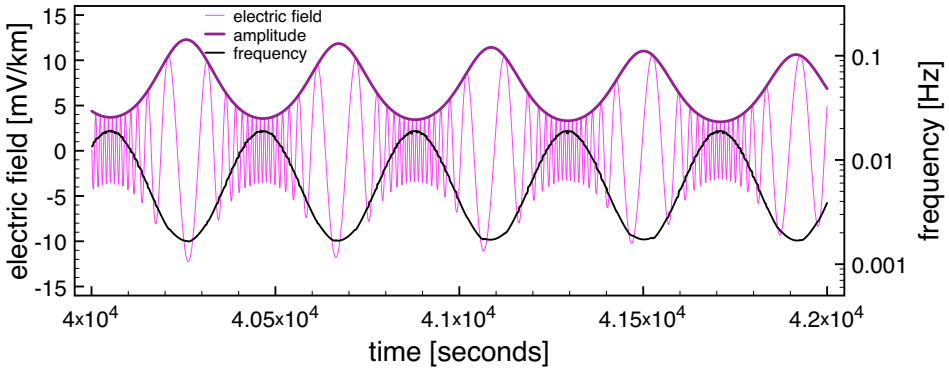


Figure 4.9: *Southern Africa data (site 072) jammed with synthetic non stationary electric noise. Panel (a) displays the electric field channel with and without the added noise to illustrate the impact of the noise compared to the data. Panel (b) zooms into a section of the added noise to emphasize that both amplitude and frequency content of this noise is clearly non stationary.*

shows that there are only marginal differences. Both algorithms agree well with the phase but there is a slight difference in the amplitudes. Concluding this example, *EMT* appears to obtain similar results but the smaller confidence intervals are less conservative or suggest higher precision.

4.9.4 Real Data Jammed with Synthetic Non Stationary Noise

As a semi-synthetic test, we combine the good real data set (site 072) from the previous section with synthetic, non stationary noise. The goal of this test is to learn how easily a quasi stationary source can be compromised by non stationary noise and to test if our algorithm is able to treat the

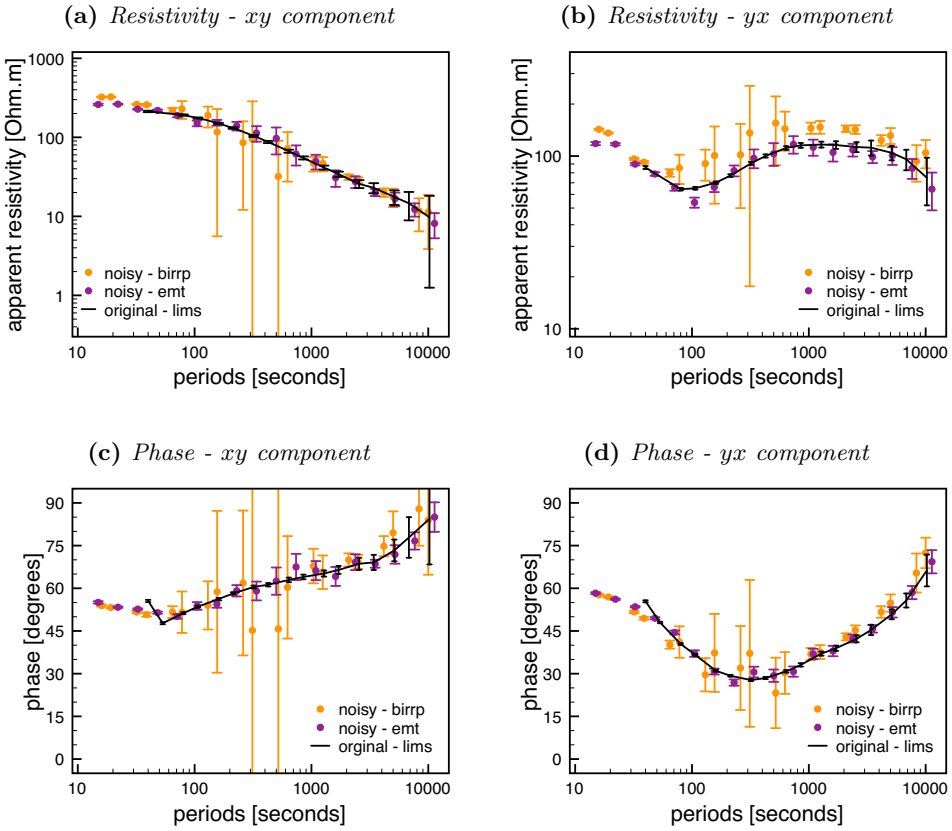


Figure 4.10: *Southern Africa data (site* 072*) jammed with synthetic, non stationary electric noise of low amplitude.*

situation correctly. We consider the non stationary noise to be present in the electric fields only and leave the magnetic and remote channels completely unaffected. This way we can see if the computation of the spectra via Fourier Transform fails or succeeds, since stationary noise in the electric channels could be cleaned by the unaffected magnetic and/or remote reference channels by the remote referencing technique.

As data, we use the data set shown in Figure 4.7 and add independent, purely non stationary noise as defined in equation (4.6) to each electric field channel

$$(e_{x,\text{ with noise}}, e_{y,\text{ with noise}}) = (e_x, e_y) + (s_1, s_2).$$

Then, we try to recover the original impedance (Figure 4.7) by processing the altered data with *BIRRP* and *EMT* to study the effects of the added, non

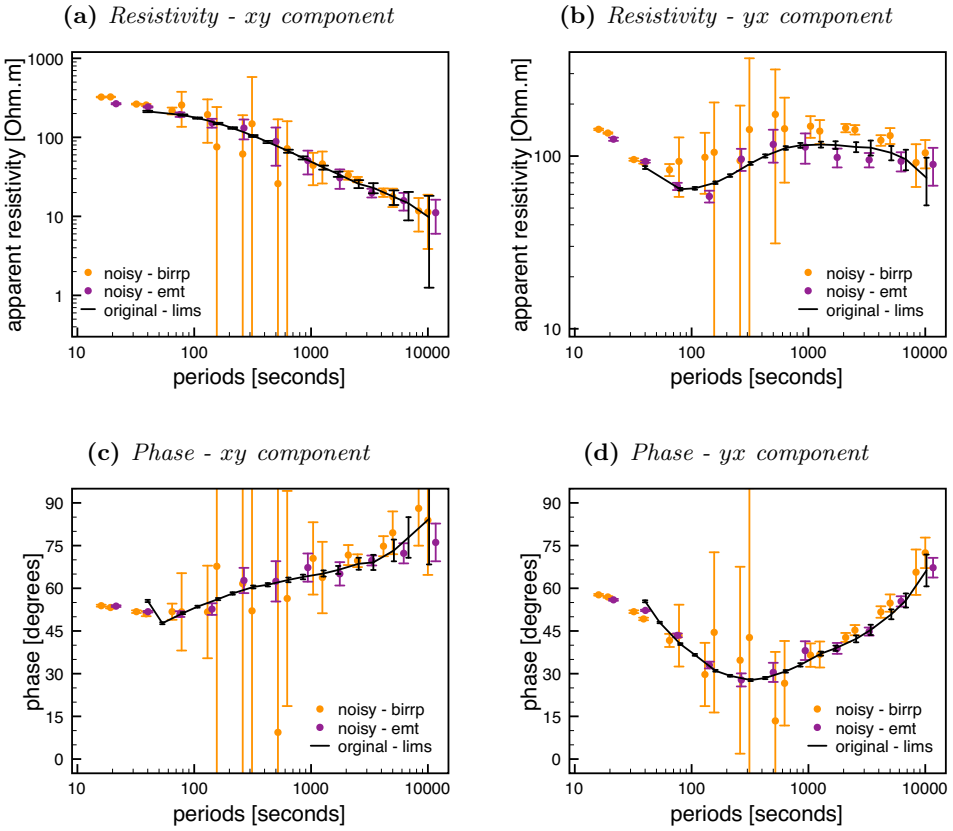


Figure 4.11: *Southern Africa data (site 072) jammed with synthetic, non stationary electric noise of medium amplitude.*

stationary noise. The test is performed trice, first with a certain noise level, then again with the noise doubled and quadrupled. Figure 4.9a displays the electrical field North-South component with and without the added noise for all three tests and as an example, a section of the added noise is illustrated in Figure 4.9b with its parameters amplitude and frequency. The spectral range of the noise is set between 1.7 mHz and 19 mHz, respective 52 s and 610 s, so we expect to see the biggest impact on the data processing results in that range.

Figures 4.10 to 4.12 display the estimated impedances with the increasing impact of the non stationary noise. Where in Figure 4.10 the noise only raises the confidence limits for *BIRRP*, it camouflages the estimates in their entirety for larger noise amplitudes in Figures 4.11 and 4.12 so much that the

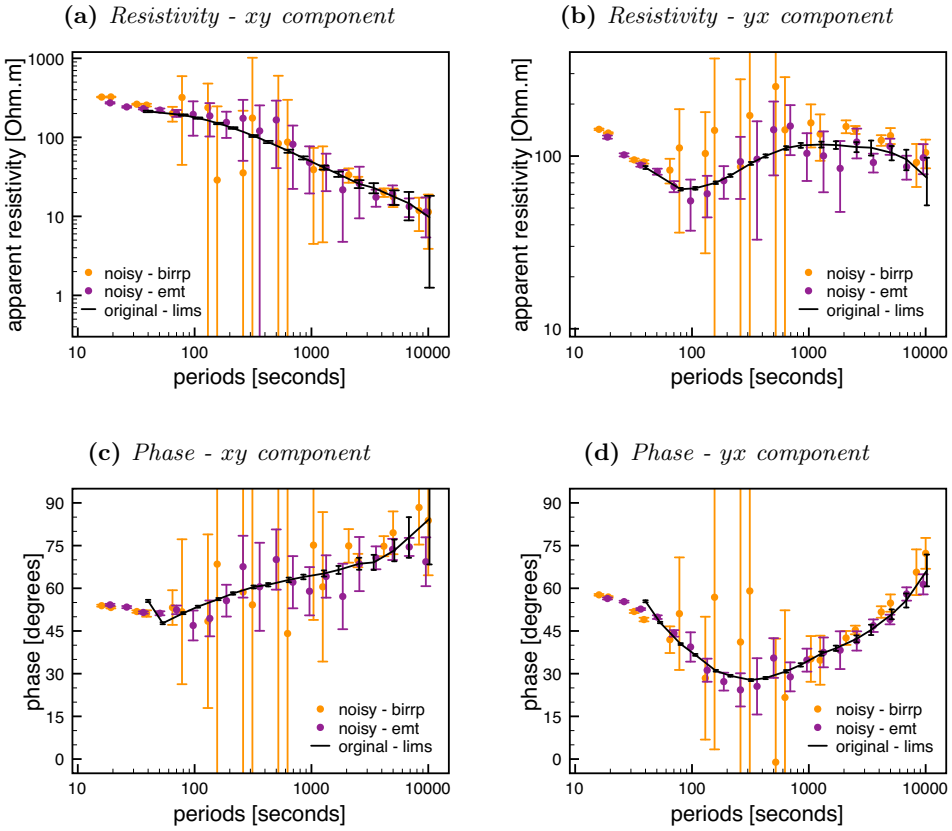


Figure 4.12: *Southern Africa data (site* 072*) jammed with synthetic, non stationary electric noise of high amplitude.*

impedance cannot be retrieved. On the other hand, *EMT* is barely affected by the lowest and medium amplitude noise, and still provides interpretable results with the highest noise amplitude.

4.9.5 Problematic Real Data from Southern Africa

This last example is a real, broad band data set and has been acquired in a region where DC trains operate and active mining takes place. Evans et al. (2011) report problems in processing the data in particular due to these noise sources.

We focus on site 042 with 027 as reference for long period measurements ($> 20\,\mathrm{s}$) and with 043 as reference for short period data ($< 20\,\mathrm{s}$). The long period data have been collected with LIMS instruments and the short period data were measured by band 5 of Phoenix Systems' instruments.

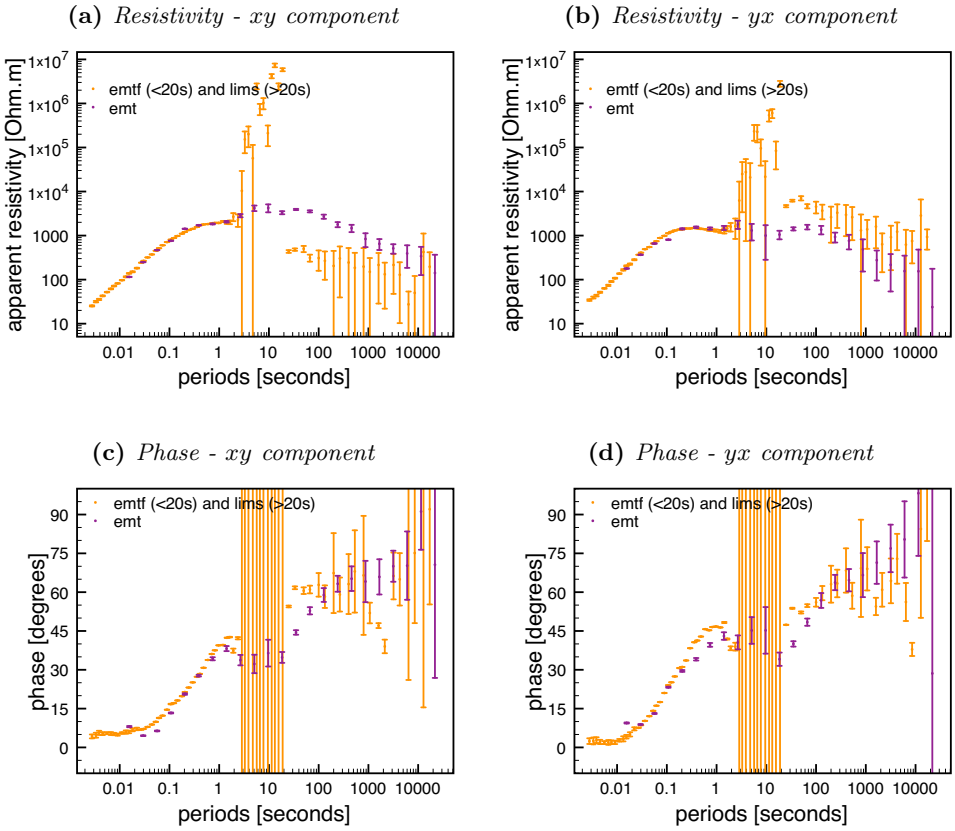


Figure 4.13: *Site* 042 *contains a tremendous amount of noise which complicates interpretation of periods of* 3 s *and more. The* EMTF *(Egbert, 1997) (BBMT) and the* LIMS *(Jones and Jödicke, 1984) (LMT) results are the original results from the SAMTEX experiment (Evans et al., 2011). The merge of BBMT and LMT responses was performed manually and as it still is common practice, the LMT apparent resistivity were shifted to match the BBMT apparent resistivity at the overlapping periods.*

The site contains a tremendous amount of noise which makes interpretation difficult from about 3 s on (Figure 4.13). The data of this site were originally processed with *EMTF* (Egbert, 1997) for short periods (Broad Band Magnetotellurics (BBMT)) and with the *LIMS processing algorithm* (Jones and Jödicke, 1984) for the Long period Magnetotellurics (LMT). The amplitude results from LMT have been scaled by the acquisition team to account for static shift according to the interpretation of the BBMT data, whereby the results from *EMT* are unchanged, since it does not suggest that the the long period data have been affected notably by static shift.

We use originally published data for this plot, because we argue, that

(in time of original publication) the interpretation of the data (that it is affected by static shift) was wrong due to some noise effect. The EMT result is not shifted, because it does not lead to the conclusion that the LMT data requires a shift, which exemplifies the long ranging effect of noise beyond data processing and highlights the strength of the algorithm in this situation.

Besides the apparent noise between 3 s and 20 s, the phase estimations between 1 s and 100 s obtained from *EMT* are consistently 5 ° to 10 ° lower than the results estimated by the other algorithms, which we cannot explain at this point. Two possible reasons for this discrepancy could be due to non stationary spectral leakage in the other algorithms (compare processing of a purely non stationary data set in Figure 4.6) or due to strong correlated noise distorting significantly the first two dominant principal components.

4.10 Conclusion

In the course of this work, we outlined a robust magnetotelluric data processing scheme purely based on non stationary methods and showed that its results compare to state-of-the-art algorithms. In contrast to other groups, our algorithm directly uses the instantaneous parameters of the measured multivariate time series and therefore, naturally handle non stationary sources. In theory, our scheme is less apt to introduce bias from spectral leakage due to this kind of noise and our synthetic and real data examples support this.

The algorithm carefully incorporates the most general and important data quality control measures like remote referencing and robust statistics as countermeasures for uncorrelated noise between occupied sites and control of highly influential but statistically unlikely data points, respectively.

This new methodology operates in a time-frequency domain and, therefore, potentially enables new data quality control measures like controlling instantaneous changes in the parameters amplitude, phase and frequency, which could be investigated in a future work.

The function to select the independent data samples assures that the correct amount of data is selected, but the function of choice is somewhat arbitrary. On one hand it can be seen as an advantage that the samples are drawn arbitrary or random, but on the other hand alternative ways should

be investigated for assessing their performance.

We demonstrated on synthetic and real data that a non stationary approach in MT processing can be fruitful. The synthetic, non stationary source in this work is specifically designed to disturb the Fourier Transform and to break its assumptions, however, the results provide an insight in how bad real, non stationary noise can affect MT measurements and encourage to verify the findings on more real world data sets that are suspected to contain, in particular, non stationary noise, e.g. data that is acquired close to train lines, mining activity or electric fences.

Lastly, we present one such example of real data and find that, at the time of original data processing, even the interpretation of the data has been affected by non stationary noise, because the long period data has been corrected unnecessarily for static shift by the original processing team.

We encourage to reassess more data sets that have been difficult to process in the past in order to investigate for non stationary effects. However, we wish to stress that, at this moment, our proposed algorithm is realized in MatLab and runs rather slow (about 1 day for 10 million data points) on desktop computers. Most of the time, it delivers similar results compared to much faster and more efficient processing algorithms like *BIRRP* (Chave and Thomson, 2004), *EMTF* (Egbert, 1997) or the *LIMS processing algorithm* (Jones and Jödicke, 1984). Therefore, we consider our algorithm a special purpose code for data that is suspected to be contaminated by non stationary effects.

Bibliography

Adam, A., Szarka, L., Verö, J., Wallner, A., Gutdeutsch, R., 1986. Magnetotellurics (MT) in mountains—noise, topographic and crustal inhomogeneity effects. Physics of the Earth and Planetary Interiors 42 (3), 165.

Battista, B. M., Knapp, C., McGee, T., Goebel, V., 2007. Application of the empirical mode decomposition and Hilbert-Huang transform to seismic reflection data. Geophysics 72 (2), H29.

Berdichevskiy, M., Bezruk, I., Chinareva, O., 1973. Magnetotelluric sounding using mathematical filters. Vol. 9. American Geophysical Union.

Cagniard, L., 1953. Basic theory of the magneto-telluric method of geophysical prospecting. Geophysics 18 (3), 605.

Cai, J.-h., 2013. Magnetotelluric Response Function Estimation Based on Hilbert–Huang Transform. Pure and Applied Geophysics 170 (11), 1899.

Cai, J.-H., Tang, J.-T., Hua, X.-R., Gong, Y.-R., 2009. An analysis method for magnetotelluric data based on the Hilbert–Huang Transform. Exploration Geophysics 40 (2), 197.

Chave, A. D., 2012. Estimation of the magnetotelluric response function. Cambridge University Press, Ch. 5, p. 165.

Chave, A. D., Thomson, D. J., 2004. Bounded influence magnetotelluric response function estimation. Geophysical Journal International 157 (3), 988.

Chave, A. D., Thomson, D. J., Ander, M. E., 1987. On the robust estimation of power spectra, coherences, and transfer functions. Journal of Geophysical Research: Solid Earth (1978–2012) 92 (B1), 633.

Chen, J., Heincke, B., Jegen, M., Moorkamp, M., 2012. Using empirical mode decomposition to process marine magnetotelluric data. Geophysical Journal International 190 (1), 293.
URL http://dx.doi.org/10.1111/j.1365-246X.2012.05470.x

Chen, J., Jegen, M., 2008. Empirical mode decomposition and Hilbert-Huang transform (HHT) in MT data processing, extended Abstract at 19. International Workshop on EM Induction in the Earth.

Egbert, G. D., 1997. Robust multiple-station magnetotelluric data processing. Geophysical Journal International 130 (2), 475.

Egbert, G. D., 2002. Processing and interpretation of electromagnetic induction array data. Surveys in geophysics 23 (2-3), 207.

Egbert, G. D., Booker, J. R., 1986. Robust estimation of geomagnetic transfer functions. Geophysical Journal International 87 (1), 173.

Evans, R. L., Jones, A. G., Garcia, X., Muller, M., Hamilton, M., Evans, S., Fourie, C. J. S., Spratt, J., Webb, S., Jelsma, H., Hutchins, D., 2011. Electrical lithosphere beneath the Kaapvaal craton, southern Africa. Journal of Geophysical Research 116 (B4), B04105.

Flandrin, P., Rilling, G., Goncalves, P., 2004. Empirical mode decomposition as a filter bank. Signal Processing Letters, IEEE 11 (2), 112.

Gamble, T., Goubau, W. M., Clarke, J., 1979. Magnetotellurics with a remote magnetic reference. Geophysics 44 (1), 53.

Garcia, X., Chave, A. D., Jones, A. G., 1997. Robust processing of magnetotelluric data from the auroral zone. Journal of geomagnetism and geoelectricity 49 (11-12), 1451.

Garcia, X., Jones, A. G., 2002. Atmospheric sources for audio-magnetotelluric (AMT) sounding. Geophysics 67 (2), 448.

Huang, N., Wu, Z., Long, S., Arnold, K., Chen, X., 2009. On instantaneous frequency. Advances in Adaptive Data Analysis 1 (2), 177.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., Liu, H. H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 454 (1971), 903.

Hubert, M., Rousseeuw, P., Verdonck, T., 2009. Robust PCA for skewed data and its outlier map. Computational statistics & data analysis 53 (6), 2264.

Hubert, M., Verboven, S., 2003. A robust PCR method for high-dimensional regressors. Journal of Chemometrics 17 (8-9), 438.

Jones, A., Jödicke, H., 1984. Magnetotelluric transfer function estimation improvement by a coherence-based rejection technique. 1984 SEG Annual Meeting.

Jones, A. G., Chave, A. D., Egbert, G., Auld, D., Bahr, K., 1989. A comparison of techniques for magnetotelluric response function estimation. Journal of Geophysical Research: Solid Earth (1978–2012) 94 (B10), 14201.

Junge, A., 1996. Characterization of and correction for cultural noise. Surveys in geophysics 17 (4), 361.

Keller, G. V., Frischknecht, F. C., 1966. Electrical methods in geophysical prospecting. Pergamon.

Liu, R. Y., et al., 1988. Bootstrap procedures under some non-iid models. The Annals of Statistics 16 (4), 1696.

Liu, W., Fujimoto, M., 2011. The Dynamic Magnetosphere. Springer.

Neukirch, M., Garcia, X., 2013. Non Stationary Time Series Convolution: On the Relation Between Hilbert-Huang and Fourier Transform. Advances in Adaptive Data Analysis. 5 (1).

Oettinger, G., Haak, V., Larsen, J., 2001. Noise reduction in magnetotelluric time-series with a new signal–noise separation method and its application to a field experiment in the Saxonian Granulite Massif. Geophysical Journal International 146 (3), 659.

Rakov, V. A., Uman, M. A., 2007. Lightning. Physics and Effects. Cambridge University Press.

Rehman, N., Mandic, D. P., 2010. Multivariate empirical mode decomposition. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 466 (2117), 1291.

Reiersøl, O., 1941. Confluence Analysis by Means of Lag Moments and Other Methods of Confluence Analysis. Econometrica 9 (1), 1.

Smirnov, M. Y., Egbert, G., 2012. Robust principal component analysis of electromagnetic arrays with missing data. Geophysical Journal International 190 (3), 1423.

Szarka, L., 1988. Geophysical aspects of man-made electromagnetic noise in the earthA review. Surveys in Geophysics 9 (3-4), 287.

Verboven, S., Hubert, M., 2005. LIBRA: a MATLAB library for robust analysis. Chemometrics and intelligent laboratory systems 75 (2), 127.

Viljanen, A., 2012. Description of the magnetospheric/ionospheric sources. Cambridge University Press, Ch. 3B, p. 96.

Weckmann, U., Magunia, A., Ritter, O., 2005. Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme. Geophysical Journal International 161 (3), 635.

Zhang, R. R., Ma, S., Safak, E., Hartzell, S., 2003. Hilbert-Huang transform analysis of dynamic and earthquake motion recordings. Journal of Engineering Mechanics 129 (8), 861.

CHAPTER 5

Conclusions and Outlook

5.1 Conclusions

The purpose of this thesis is to investigate the necessity of a non stationary time series analysis in the magnetotelluric (MT) method. The introduction outlines that, even though, the MT signal is at most times sufficiently stationary, the inevitable noise can be of any form and therefore, can be non stationary as well. The effort in developing a non stationary method in MT has been little for the popularity of the Fourier Transform (FT) but we highlight one approach by Berdichevskiy et al. (1973) that can be regarded as a precursor to this work and an early attempt on truly non stationary analysis. By emphasizing the limitations of that approach, I set the stage for my approach by introducing the Hilbert-Huang Transform (HHT, Huang et al., 1998) and establishing the necessary, basic foundations.

In the following, I conclude the main parts of this thesis, which lead to the development of a new, non stationary processing scheme for MT data that is based on the HHT and non stationary convolution.

5.1.1 Convolution and Deconvolution of Non Stationary Signals

I demonstrate mathematically that the convolution between an Intrinsic Mode Function (IMF) that is described by its fully decomposed instantaneous parameters and a temporal system response can be translated into a multiplication of the same IMF and the spectral system response. As a result I found, that the Hilbert-Huang spectra can be represented as a Fourier

spectra with time varying complex amplitude.

This finding proofs for the first time that the HHT is a suitable tool for typical transfer function analyses methods like MT data processing and that the instantaneous frequency of non stationary processes obtained through HHT is a physically meaningful entity comparable with the frequency obtained by the Fourier analysis for stationary processes.

By analytically analyzing the results implied by the non stationary convolution theorem, I find that convolution of non stationary time series with a general system response function may alter the characteristic time scale of the time series and introduce a shift in the instantaneous frequency depending on the characteristics of the convolved system response function and the instantaneous frequency of the original signal. This frequency shift renders any deconvolution attempt difficult as such that no analytic solution exists but numeric optimization procedures may be successful. Fortunately, the MT impedance tensor is relatively smooth over frequency and therefore, the change in frequency for a non stationary signal would likely be much smaller than the usual smoothing bandwidth for the impedance tensor estimation.

5.1.2 Applying Non Stationary Convolution to MT

In the course of this work, I outlined a robust magnetotelluric data processing scheme purely based on non stationary methods and showed that its results compare to state-of-the-art algorithms. In contrast to other proposed solutions, my algorithm directly uses the instantaneous parameters of the measured multivariate time series and therefore, naturally handles non stationary sources. In theory, the presented scheme is less apt to introduce bias from spectral leakage due to this kind of noise and the synthetic and real data examples support this.

The algorithm carefully incorporates the most general and important data quality control measures like remote referencing and robust statistics as countermeasures for uncorrelated noise between occupied sites and control of highly influential but statistically unlikely data points, respectively.

I demonstrate on synthetic and real data that a non stationary approach in MT processing can be fruitful. The synthetic, non stationary source in this work is specifically designed to disturb the Fourier Transform and to break its assumptions, however, the results provide an insight in how bad real,

non stationary noise can affect MT measurements and encourage to verify the findings on more real world data sets that are suspected to contain, in particular, non stationary noise, e.g. data that is acquired close to train lines, mining activity or electric fences.

Lastly, one example of non stationary, real data is presented, where, at the time of original data processing, even the interpretation of the data has been affected by the non stationary noise, because the long period data has been corrected unnecessarily for static shift by the original processing team.

5.1.3 Justification for the Non Stationary Approach

I compare the Empirical-mode-decomposition-based MagnetoTelluric processing algorithm (EMT) with *BIRRP* by Chave and Thomson (2004), which is a traditional, efficient and renown processing code based on the Fourier Transform, by performing several synthetic tests, first, synthetic, non stationary data is constructed from two non stationary sources to demonstrate the inability of a Fourier based method to deal with non stationary sources. Then, secondly, these sources are used as a synthetic noise source. The non stationary noise is added only to the electric fields and leaves the magnetic and remote channels completely unaffected. Therefore, I can show that the computation of the spectra via Fourier Transform fails, because uncorrelated stationary noise in the spectra should be cleaned by the remote referencing technique, which is applied in the test. Since any uncorrelated (random) non stationary noise acts as any random stationary noise and does not affect the measurements other than decreasing the confidence in the results (larger error bars), this test shows that the mere fact, that the added noise is non stationary, affects the estimated results by a Fourier Transform based method or even makes it impossible to extract reasonable transfer functions, whereas the EMT algorithm is able to deal with the non stationarity and allows a more precise estimation to a lower signal-to-noise ratio.

Conclusively, these tests demonstrate that non stationary sources can heavily impact on traditional MT processing routines which rely on the Fourier Transform but that this effect can be diminished by relying on a purely non stationary analysis. Even though, the non stationary source is specifically designed to disturb the Fourier Transform and to break its assumptions, the results provide an insight in how bad real non stationary

noise can affect MT measurements. This becomes even more evident, when the real data example exhibits very similar problems in conjunction with other Fourier based method but not with the EMT algorithm.

However, I wish to stress that, at this moment, the proposed algorithm is realized in MatLab and runs rather slow (about 1 day for 10 million data points) on desktop computers. Most of the time, it delivers similar results compared to much faster and more efficient processing algorithms like *BIRRP* (Chave and Thomson, 2004), *EMTF* (Egbert, 1997) or the *LIMS processing algorithm* (Jones and Jödicke, 1984). Therefore, I consider this algorithm a special purpose code for data that is suspected to be contaminated by non stationary effects.

5.2 Outlook

Besides the development of the EMT algorithm, this thesis provides a set of new ideas for further research on the field of non stationary electromagnetic methodology.

5.2.1 Development of Quality Control Measures

The proposed algorithm operates in a time-frequency domain and, therefore, potentially enables new data quality control measures like controlling instantaneous changes in the parameters amplitude, phase and frequency, which could be investigated in a future work. This could include coherence based rejection techniques as well as exclusion of overly non stationary parts of the data (note that the natural sources are rarely non stationary).

5.2.2 Selection of Independent Data Points

The function to select the independent data samples in the presented algorithm assures that the correct amount of data is selected, but the function of choice is somewhat arbitrary. On one hand, it can be seen as an advantage that the samples are drawn arbitrary or random, but on the other hand, alternative ways should be investigated for assessing their performance.

5.2.3 Re-examine Previously Problematic Real World Data

It would be of the utmost importance to the success of the algorithm to verify it on more real world data sets that are suspected to contain non stationary noise, e.g., data that is acquired close to train lines, mining activity or electric fences. More data sets that have been difficult to process in the past could be reprocessed in order to investigate for the possibility of non stationary contamination.

5.2.4 Fortran Conversion, Parallel Computing and Open Access

As mentioned before, the code execution is slow compared to Fourier based methods and under regular conditions, the results are similar. Since the development of the code has mostly finished and in order to make it more competitive, it would be advantageous to convert it to the Fortran programming language and enable parallel computing.

Another critical point would be to allow open access to the code in order to spread its usage. Firstly, this would increase the rate of bug fixing; secondly, yield more (or less) confidence from a larger testing community; thirdly, raise awareness of (or lower uncertainty about) non stationary effects in the data; and fourthly, motivate the algorithm's further development.

5.2.5 Non Stationary Controlled Sources for CSEM

Controlled Source Electromagnetic (CSEM) methods are sensitive to the subsurface conductivity structure and thus had led to its use in resource exploration. Since the frequency for peak sensitivity and the exact location of an exploration target is normally unknown prior exploration, it is desirable to acquire the transfer functions for a broad range of frequencies and in a wide area. Investigations in both directions have been driven by optimizing properties of the Fourier transform in order to enhance the frequency range and the source-receiver distances. My research on Non stationary (NS) time series analysis tools significantly enhanced processing of NS time series, hence the possibility of NS source waveforms could be assessed. For instance, the source could be defined by a chirp function that is highly customizable in amplitude and frequency range in order to accommodate any frequency range in combination with virtually any amplitude for each frequency (e.g.

in order to counter attenuation by decreasing power from low frequencies to increase high frequency power assuming constant energy supply).

Using a chirp as source waveform, we can chose any bandwidth and we could account amplitude attenuation over source-receiver distance, since both the amplitude and frequency can be tailored to specific needs. Both are typical problems for CSEM as discussed by Myer et al. (2010) among others. As an example, by reducing low frequency amplitude in favor for high frequency amplitude, source signal range could be increased for high frequencies assuming a given power generator output. This idea has been proposed by Neukirch and Garcia (2014).

Bibliography

Berdichevskiy, M., Bezruk, I., Chinareva, O., 1973. Magnetotelluric sounding using mathematical filters. Vol. 9. American Geophysical Union.

Chave, A. D., Thomson, D. J., 2004. Bounded influence magnetotelluric response function estimation. Geophysical Journal International 157 (3), 988.

Egbert, G. D., 1997. Robust multiple-station magnetotelluric data processing. Geophysical Journal International 130 (2), 475.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., Liu, H. H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 454 (1971), 903.

Jones, A., Jödicke, H., 1984. Magnetotelluric transfer function estimation improvement by a coherence-based rejection technique. 1984 SEG Annual Meeting.

Myer, D., Constable, S., Key, K., 2010. Broad-band waveforms and robust processing for marine CSEM surveys. Geophysical Journal International 184, 689.

Neukirch, M., Garcia, X., 2014. Non Stationary, broad-band waveforms for CSEM: An analysis with synthetic data, extended Abstract at 76th EAGE Conference.

Acknowledgment

Curiously, even though every doctoral thesis is unique in the work it presents, every author before already came to the same conclusion in one single point: It could not have been realized without the help of some other great people.

First of all, I'm indebted to Xavi, my supervisor. Since the moment we met, he continuously inspired confidence in me by not only devoting his time to my queries at any moment but also by allowing great freedom for me to develop my skills freely. This combination of availability and freedom must be fueled by trust, thus I hope this thesis rewards his trust.

I would like to offer my special thanks to my tutor at the University of Barcelona, Juanjo. Whenever, I came to him at last minute with any administrative query, he welcomed me with a warm and caring smile.

I wish to acknowledge the wonderful time I had with Gerardo Romano, Luciano Telesca (Institute of Methodologies for Environmental Analysis) and the rest of the IMAA staff, as well as Mohamed Rouai (University Moulay Ismaïl) during my stay in June 2011 in Tito, Italy. Awed by Nature's might at Vesuvius and Pompeii and by the admiration for the italian cuisine, the principal idea of the proofs in chapter 2 were inspired. Further, the spontaneous assistance of Agata Siniscalchi and Zaccaria Petrillo (University of Bari, Italy) in finding the elusive "COMDAT" data deserves special thanks.

For the great opportunity to stay at Woods Hole Oceanographic Institution in November 2011, I am particularly grateful. Advice tirelessly given by Alan Chave has been a great help in my understanding of the statistical implications of the EMT algorithm. Also the technical assistance provided by Rob Evans and Jimmy Elsenbeck was greatly appreciated.

Discussions and interchange with Jin Chen, Björn Heincke and Sebastian Hölz (GEOMAR, Kiel) in February 2012 were invaluable to me. Though,

I still can't figure how one person could possibly eat a 28-inch-diameter hamburger, many other problems have been addressed. Special thanks are extended to Marion Jegen for providing the financial support for the visit.

The work for this thesis was funded within the project CODOS financed by REPSOL. I want to thank the whole group at the Barcelona-Center for Subsurface Imaging and in particular the head Cesar R. Ranero for providing the opportunity for this thesis at BCSI. I very much enjoyed my time there, especially when we went together for skiing, kayaking, swimming, snorkeling and playing beach volleyball, not to mention the full tummy after each of the "calçotadas". My special thanks are extended to Manel, who has the innate ability to just cheer you up, Estelle, with best regards to Pierre and little Samuel, Adriá, Ivan and Naiara for sharing sweat and joy during field work in Garrotxa, the many lunches we had together and all the time I could spent with them. Along with my colleagues at BCSI, I would like to express my appreciation to our colleagues at the University of Barcelona: Pilar Queralt, who I admire for her kind devotion to students; Lena, Xenia, Perla and David, with whom I learned the basics of MT in Dublin, and Joan, who introduced me to practical MT measurements and the beauty of Garrotxa.

I would like to express my gratitude to Mireia, who provided me with the energy to initiate this dissertation and journey South, and Daniel for intensively reviewing very early, very messy drafts that almost had derailed motivation. Special thanks are extended to my friends Frank, Denis and Tina, Elisabeth, Anika, René, Paulo and Carlos, who always offered me a safe haven far away from any stress and trouble.

Above all, I'm most grateful to my family for their unconditional support. Der größte Dank gebührt Gabi, Margot, André, Jürgen und dem Rest meiner Familie, die mich in all meinen Vorhaben bedingungslos unterstützt haben. Vielen Dank für Eure Unterstützung, Eure Besuche, Eure Geduld, Euer Verständnis, Eure Liebe, ganz einfach vielen Dank, dass es Euch gibt.

Lastly, I thank my best colleague, my best friend and love, Savitri, for being there when my eyes can't see what truly is important.

APPENDIX A

The Tikhonov-Cagniard Impedance derived from Maxwell's Equations for Plane Wave Sources and a full Apparent Conductivity Tensor

A.1 Derivation of the Impedance

The EM field is described by Maxwells equations (Maxwell, 1873). Considering (1) harmonic fields with constant frequency ω, (2) the current density $\mathbf{j}$ is proportional to the electric field $\mathbf{E}$ at a constant conductivity σ ($\mathbf{j} = \boldsymbol{\sigma}\mathbf{E}$) and (3) the field propagates only by diffusion (neglecting displacement currents and surface charges: $i\omega\epsilon\mathbf{E} \ll \boldsymbol{\sigma} \cdot \mathbf{E}$ and $q = 0$), these equations are:

$$\boldsymbol{\nabla} \times \mathbf{E} = -i\omega\mathbf{B}, \tag{A.1a}$$

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu\left(i\omega\epsilon\mathbf{E} + \mathbf{j}\right) = i\omega\mu\epsilon\mathbf{E} + \mu\boldsymbol{\sigma} \cdot \mathbf{E} \approx \mu\boldsymbol{\sigma} \cdot \mathbf{E}, \tag{A.1b}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0, \tag{A.1c}$$

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{q}{\epsilon} \approx 0 \tag{A.1d}$$

with μ describing the magnetic susceptibility as product between the magnetic susceptibility of vacuum μ_0 and the magnetic susceptibility of the given medium μ_r. The penetration of the field $\mathbf{F} = \mathbf{E}, \mathbf{B}$ in a homogeneous earth is described by applying the curl operator ($\boldsymbol{\nabla}\times$) to Maxwell's equations and solve the equations for each field $\mathbf{F}$:

$$\nabla^2\mathbf{F} = i\omega\mu\sigma\mathbf{F} = k^2\mathbf{F} \tag{A.2}$$

The term $k^2 = i\omega\mu\sigma$ is the diffusion parameter, which describes the complex penetration depth (skin depth) $\delta = \mathfrak{Re}\left(\frac{1}{k}\right)$ of the EM field (Schmucker and Weidelt, 1975) and represents the impedance $Z(\omega) = \frac{i\omega}{k}$ for a homogenous earth. The solution for the fields $\mathbf{F}$ is given by $\mathbf{F} = \mathbf{F}_0 \cdot e^{-\mathbf{k}\cdot\mathbf{x}}$ with $\mathbf{x} = (x, y, z)$ and $\mathbf{k} = (k_x, k_y, k_z)$ as the space dimension and their respective wave numbers. That the wave number is directly related to the impedance becomes clear when one solves equations (A.1) for a three dimensional, anisotropic body with the apparent conductivity tensor $\boldsymbol{\sigma} = \sigma_{ij}, \quad i, j \in [x, y, z]$. Then, the components of equations A.1a and A.1b yield:

$$-k_y E_z + k_z E_y = -i\omega B_x, \tag{A.3a}$$

$$-k_z E_x + k_x E_z = -i\omega B_y, \tag{A.3b}$$

$$-k_x E_y + k_y E_x = -i\omega B_z, \tag{A.3c}$$

$$-k_y B_z + k_z B_y = \mu\sigma_{xx} E_x + \mu\sigma_{xy} E_y + \mu\sigma_{xz} E_z, \tag{A.3d}$$

$$-k_z B_x + k_x B_z = \mu\sigma_{yx} E_x + \mu\sigma_{yy} E_y + \mu\sigma_{yz} E_z, \tag{A.3e}$$

$$-k_x B_y + k_y B_x = \mu\sigma_{zx} E_x + \mu\sigma_{zy} E_y + \mu\sigma_{zz} E_z. \tag{A.3f}$$

Traditionally, the impedance is defined as the relation between the horizontal field components, so we need to eliminate the vertical components. Rearrange equations (A.3c) and (A.3f):

$$B_z = -\frac{i}{\omega} k_x E_y + \frac{i}{\omega} k_y E_x, \tag{A.4a}$$

$$E_z = -\frac{k_x}{\mu\sigma_{zz}} B_y + \frac{k_y}{\mu\sigma_{zz}} B_x - \frac{\sigma_{zx}}{\sigma_{zz}} E_x - \frac{\sigma_{zy}}{\sigma_{zz}} E_y. \tag{A.4b}$$

Since it is only guaranteed for the diagonal elements of $\boldsymbol{\sigma}$ to be larger than zero, we may only divide by the main diagonal elements, hence we need to reformulate equation (A.3d) to yield E_x and equation (A.3e) to yield E_y:

$$E_x = -\frac{k_y}{\mu\sigma_{xx}} B_z + \frac{k_z}{\sigma_{xx}} B_y - \frac{\sigma_{xy}}{\sigma_{xx}} E_y - \frac{\sigma_{xz}}{\sigma_{xx}} E_z, \tag{A.5a}$$

$$E_y = -\frac{k_z}{\mu\sigma_{yy}} B_x + \frac{k_x}{\mu\sigma_{yy}} B_z - \frac{\sigma_{yx}}{\sigma_{yy}} E_x - \frac{\sigma_{yz}}{\sigma_{yy}} E_z. \tag{A.5b}$$

Inserting equations (A.4) in (A.5) and rearranging to isolate the electric field

component yields:

$$\left(1+\frac{ik_y^2}{\omega\mu\sigma_{xx}}-\frac{\sigma_{xz}\sigma_{zx}}{\sigma_{xx}\sigma_{zz}}\right)E_x = -\frac{k_y\sigma_{xz}}{\mu\sigma_{xx}\sigma_{zz}}B_x+\left(\frac{k_z}{\mu\sigma_{xx}}+\frac{k_x\sigma_{xz}}{\mu\sigma_{xx}\sigma_{zz}}\right)B_y + \left(\frac{ik_xk_y}{\omega\mu\sigma_{xx}}-\frac{\sigma_{xy}}{\sigma_{xx}}+\frac{\sigma_{xz}\sigma_{zy}}{\sigma_{xx}\sigma_{zz}}\right)E_y, \quad \text{(A.6a)}$$

$$\left(1+\frac{ik_x^2}{\omega\mu\sigma_{yy}}-\frac{\sigma_{yz}\sigma_{zy}}{\sigma_{yy}\sigma_{zz}}\right)E_y = -\left(\frac{k_z}{\mu\sigma_{yy}}-\frac{k_y\sigma_{yz}}{\mu\sigma_{yy}\sigma_{zz}}\right)B_x+\frac{k_x\sigma_{yz}}{\mu\sigma_{yy}\sigma_{zz}}B_y + \left(\frac{ik_xk_y}{\omega\mu\sigma_{yy}}-\frac{\sigma_{yx}}{\sigma_{yy}}+\frac{\sigma_{yz}\sigma_{zx}}{\sigma_{yy}\sigma_{zz}}\right)E_x. \quad \text{(A.6b)}$$

Multiply equations (A.6a) and (A.6b) by the factors of the left hand side, $\left(1+\frac{ik_x^2}{\omega\mu\sigma_{yy}}-\frac{\sigma_{yz}\sigma_{zy}}{\sigma_{yy}\sigma_{zz}}\right)$ and $\left(1+\frac{ik_y^2}{\omega\mu\sigma_{xx}}-\frac{\sigma_{xz}\sigma_{zx}}{\sigma_{xx}\sigma_{zz}}\right)$, respectively, and separate the electric field components:

$$E_x = \frac{Z_{xx}^n}{\mu Z^d}B_x+\frac{Z_{xy}^n}{\mu Z^d}B_y \quad \text{(A.7a)}$$

$$E_y = \frac{Z_{yx}^n}{\mu Z^d}B_x+\frac{Z_{yy}^n}{\mu Z^d}B_y. \quad \text{(A.7b)}$$

Further, the electric fields in equation (A.7) can be substituted in equation (A.4a) to yield the transfer function that relates the vertical magnetic field to the horizontal magnetic field (commonly known as tipper):

$$B_z = \frac{iZ_{zx}^n}{\mu^2 Z^d}B_x+\frac{iZ_{zy}^n}{\mu^2 Z^d}B_y. \quad \text{(A.8a)}$$

with Z_{ij}^n as the numerator and Z^d as denominator for Z_{ij}:

$$Z_{xx}^n = \mu\omega k_y\sigma_{xy}\sigma_{yz}-ik_xk_y{}^2\sigma_{yz}-\mu\omega k_y\sigma_{xz}\sigma_{yy}+\mu\omega k_z\sigma_{xy}\sigma_{zz} - \mu\omega k_z\sigma_{xz}\sigma_{zy}-ik_xk_yk_z\sigma_{zz}-ik_x{}^2k_y\sigma_{xz} \quad \text{(A.9a)}$$

$$Z_{xy}^n = ik_x{}^2k_y\sigma_{yz}-\mu\omega k_x\sigma_{xy}\sigma_{yz}+\mu\omega k_x\sigma_{xz}\sigma_{yy}+\mu\omega k_z\sigma_{yy}\sigma_{zz} - \mu\omega k_z\sigma_{yz}\sigma_{zy}+ik_x{}^2k_z\sigma_{zz}+ik_x{}^3\sigma_{xz} \quad \text{(A.9b)}$$

$$Z_{yx}^n = \mu\omega k_y\sigma_{xz}\sigma_{yx}-\mu\omega k_y\sigma_{xx}\sigma_{yz}-ik_xk_y{}^2\sigma_{xz}-\mu\omega k_z\sigma_{xx}\sigma_{zz} + \mu\omega k_z\sigma_{xz}\sigma_{zx}-ik_y{}^2k_z\sigma_{zz}-ik_y{}^3\sigma_{yz} \quad \text{(A.9c)}$$

$$Z_{yy}^n = ik_xk_y{}^2\sigma_{yz}+\mu\omega k_x\sigma_{xx}\sigma_{yz}-\mu\omega k_x\sigma_{xz}\sigma_{yx}-k_z\mu\omega\sigma_{yx}\sigma_{zz} + \mu\omega k_z\sigma_{yz}\sigma_{zx}+ik_xk_yk_z\sigma_{zz}+ik_x{}^2k_y\sigma_{xz} \quad \text{(A.9d)}$$

$$
\begin{aligned}
Z^n_{zx} &= {k_y}^2\sigma_{xy}\sigma_{yz} - {k_y}^2\sigma_{xz}\sigma_{yy} + k_xk_y\sigma_{xx}\sigma_{yz} - k_xk_y\sigma_{xz}\sigma_{yx} \\
&\quad + k_xk_z\sigma_{xx}\sigma_{zz} - k_xk_z\sigma_{xz}\sigma_{zx} + k_yk_z\sigma_{xy}\sigma_{zz} - k_yk_z\sigma_{xz}\sigma_{zy} \qquad \text{(A.9e)} \\
Z^n_{zy} &= -{k_x}^2\sigma_{xx}\sigma_{yz} + {k_x}^2\sigma_{xz}\sigma_{yx} - k_xk_y\sigma_{xy}\sigma_{yz} + k_xk_y\sigma_{xz}\sigma_{yy} \\
&\quad + k_xk_z\sigma_{yx}\sigma_{zz} - k_xk_z\sigma_{yz}\sigma_{zx} + k_yk_z\sigma_{yy}\sigma_{zz} - k_yk_z\sigma_{yz}\sigma_{zy} \qquad \text{(A.9f)} \\
Z^d &= \mu\omega\sigma_{xx}\sigma_{yy}\sigma_{zz} - ik_y{}^2\sigma_{yz}\sigma_{zy} - ik_xk_y\sigma_{xz}\sigma_{zy} - ik_xk_y\sigma_{yz}\sigma_{zx} \\
&\quad - ik_x{}^2\sigma_{xz}\sigma_{zx} - \mu\omega\sigma_{xx}\sigma_{yz}\sigma_{zy} - \mu\omega\sigma_{xy}\sigma_{yx}\sigma_{zz} + \mu\omega\sigma_{xy}\sigma_{yz}\sigma_{zx} \\
&\quad + \mu\omega\sigma_{xz}\sigma_{yx}\sigma_{zy} - \mu\omega\sigma_{xz}\sigma_{yy}\sigma_{zx} + ik_y{}^2\sigma_{yy}\sigma_{zz} + ik_xk_y\sigma_{yx}\sigma_{zz} \\
&\quad + ik_xk_y\sigma_{xy}\sigma_{zz} + ik_x{}^2\sigma_{xx}\sigma_{zz} \qquad \text{(A.9g)}
\end{aligned}
$$

A.2 Simplification of Impedance Tensor Elements

The general impedance is a complex relation between the anisotropic conductivity and the spatial wave numbers. However, we can extract some special cases under which the tensor elements reduce complexity.

A.2.1 Isotropic Material

One-dimensional structures One-dimensionality refers to the fact that there are no changes of the conductivity structure along $x-$ and $y-$ direction. For us, this means $k_x = k_y = 0$ because both, the electric and magnetic field of a plane wave, do not vary laterally over a laterally uniform media. Isotropy is defined as directional independence of a measure, which means the conductivity tensor elements are $\sigma_{xx=yy} = \sigma_0$ and $\sigma_{i\neq j} = 0$. Further, the amplitude of the wave number vector k is defined by $k^2 = k_x^2 + k_y^2 + k_z^2$. Hence the tensor elements simplify to:

$$
Z_{xx} = Z_{yy} = Z_{zx} = Z_{zy} = 0, \quad Z_{xy} = -Z_{yx} = \frac{k_z}{\mu\sigma_0} = \frac{i\omega}{k_z}. \qquad \text{(A.10)}
$$

Note that the impedance has no information about the vertical conductivity σ_{zz} since all currents flow horizontal for a plane wave propagating through such a media with variation only along the axis of propagation.

Two-dimensional structures For a two-dimensional subsurface conductivity, the horizontal wave numbers have a principal axis defined by angle ϕ and the lateral diffusion coefficient in strike direction k_0^2, thus the measurements can be obtained in strike direction so that the wave number $k_y = 0$ and the apparent conductivity tensor elements $\sigma_{yx} = \sigma_{xy} = \sigma_{zy} = \sigma_{yz} = 0$. Note, that even though the material in the subsurface itself may be isotropic, when we analyze the impedance for a certain frequency, we cannot directly relate it to that geologic structure with isotropic material, because the impedance only relates to a volume of material and can only describe macroscopic properties of that volume, which, for a two dimensional structure, will appear anisotropic. The apparent conductivity tensor of isotropic material in a two dimensional arrangement will appear anisotropic in the direction of varying isotropic conductivity, that is in our case the $x-$ and $z-$ direction.

$$Z_{xx} = Z_{yy} = Z_{zy} = 0 \tag{A.11a}$$

$$Z_{xy} = \frac{k_x\sigma_{xz} + k_z\sigma_{zz}}{\mu \cdot (\sigma_{xx}\sigma_{zz} - \sigma_{xz}\sigma_{zx})} \tag{A.11b}$$

$$Z_{yx} = -\frac{k_z \cdot \omega}{\mu\omega\sigma_{yy} + ik_x{}^2} \tag{A.11c}$$

$$Z_{zx} = \frac{ik_xk_z}{\mu\omega\sigma_{yy} + ik_x{}^2} \tag{A.11d}$$

A.2.2 Anisotropic Material

Martí (2014) provides a complete summary of the concurrent understanding of anisotropy. The role of anisotropy and how it is interpreted in data is still in debate, in the following are the resulting equations for the impedance tensor elements assuming a homogenous anisotropic media which anisotropy properties are determined by the geologic structure and the material properties as before for the isotropic material special cases.

One-dimensional and Anisotropic For one dimensional geologic models, a plane wave still has no lateral variation and thus, $k_x = k_y = 0$ but the conductivity is directionally dependent, thus $\boldsymbol{\sigma}$ is a full tensor. Since the material does not change laterally, the anisotropic elements of the conduc-

tivity tensor are symmetric, viz. $\sigma_{xy} = \sigma_{yx}$, $\sigma_{xz} = \sigma_{zx}$ and $\sigma_{zy} = \sigma_{yz}$, and hence the impedance simplifies to:

$$Z^n_{xx} = k_z \cdot (\sigma_{xy}\sigma_{zz} - \sigma_{xz}\sigma_{yz}) \tag{A.12a}$$
$$Z^n_{xy} = k_z \cdot \left(\sigma_{yy}\sigma_{zz} - {\sigma_{yz}}^2\right) \tag{A.12b}$$
$$Z^n_{yx} = k_z \cdot \left({\sigma_{xz}}^2 - \sigma_{xx}\sigma_{zz}\right) \tag{A.12c}$$
$$Z^n_{yy} = k_z \cdot (\sigma_{xz}\sigma_{yz} - \sigma_{xy}\sigma_{zz}) \tag{A.12d}$$
$$Z^d = \sigma_{xx}\sigma_{yy}\sigma_{zz} + 2\sigma_{xy}\sigma_{xz}\sigma_{yz} - \sigma_{zz}{\sigma_{xy}}^2 - \sigma_{yy}{\sigma_{xz}}^2 - \sigma_{xx}{\sigma_{yz}}^2 \tag{A.12e}$$
$$Z_{zx} = Z_{zy} = 0 \tag{A.12f}$$
$$\tag{A.12g}$$

Note that $Z_{xx} + Z_{yy} = 0$ as suggested by computations (Vozoff, 1972) and theoretic studies (Kováčiková and Pek, 2002).

Two-dimensional and Anisotropic As for isotropic material, in a two-dimensional, anisotropic subsurface conductivity, the horizontal wave numbers have a principal axis defined by angle ϕ and the lateral diffusion coefficient in strike direction k_0^2, thus the measurements can be obtained in strike direction so that the wave number $k_y = 0$ and the apparent conductivity tensor elements exhibit symmetry along the uniform axis, viz. $\sigma_{xy} = \sigma_{yx}$ and $\sigma_{zy} = \sigma_{yz}$. Then the impedance tensor elements simplify to:

$$Z_{xx} = \mu\omega \cdot k_z \cdot (\sigma_{xy}\sigma_{zz} - \sigma_{xz}\sigma_{yz}) \tag{A.13a}$$
$$Z^n_{xy} = (k_x\sigma_{xz} + k_z\sigma_{zz}) \cdot \left(i{k_x}^2 + \mu\omega\sigma_{yy}\right) - \mu\omega \cdot \sigma_{yz} \cdot (k_x\sigma_{xy} + k_z\sigma_{yz}) \tag{A.13b}$$
$$Z^n_{yx} = -\mu\omega \cdot k_z \cdot (\sigma_{xx}\sigma_{zz} - \sigma_{xz}\sigma_{zx}) \tag{A.13c}$$
$$Z^n_{yy} = \mu\omega \cdot (k_x\sigma_{xx}\sigma_{yz} - k_x\sigma_{xy}\sigma_{xz} - k_z\sigma_{xy}\sigma_{zz} + k_z\sigma_{yz}\sigma_{zx}) \tag{A.13d}$$
$$Z^n_{zx} = ik_xk_z\left(\sigma_{xx}\sigma_{zz} - \sigma_{xz}\sigma_{zx}\right) \tag{A.13e}$$
$$Z^n_{zy} = i{k_x}^2 \cdot (\sigma_{xz}\sigma_{yx} - \sigma_{xx}\sigma_{yz}) + ik_xk_z \cdot (\sigma_{yx}\sigma_{zz} - \sigma_{yz}\sigma_{zx}) \tag{A.13f}$$
$$Z^d = \mu\omega \cdot \sigma_{yz} \cdot (\sigma_{xy}\sigma_{xz} - \sigma_{xx}\sigma_{yz}) + \mu\omega \cdot \sigma_{xy} \cdot (\sigma_{yz}\sigma_{zx} - \sigma_{xy}\sigma_{zz}) + \left(\mu\omega\sigma_{yy} + i{k_x}^2\right) \cdot (\sigma_{xx}\sigma_{zz} - \sigma_{xz}\sigma_{zx}) \tag{A.13j}$$

A.3 Impedance and Apparent Conductivity

From the previous section, we can conclude that the impedance represents an apparent conductivity tensor that the electromagnetic field experiences for a given frequency very similar to the descriptive value of the apparent resistivity. The EM wave experiences the subsurface as a volume and thus the assumed mean anisotropic parameters of a homogenous media in an integrated volume yield an equal impedance as the real conductivity structure of the same volume, be it isotropic or anisotropic and in any dimensionality. From a practical point of view, apparent conductivity tensors for different sites and at a range of frequencies, may be used to estimate real conductivity tensorial structure by inversion, similar as it is achieved in other methods, like Electrical Resistivity Tomography.

Considering the variables $\mathbf{k}$ and σ_{ij} with $i, j \in [x, y, z]$ in their relation given by $\mathbf{Z}(\omega)$, we find that the impedance tensor contains a given number of unknown parameters:

- One-dimensional and isotropic: 3, the complex-valued k_z and the real valued σ_0,
- Two-dimensional and isotropic: 9, the two complex-valued k_z and k_x, and five real-valued apparent conductivity tensor elements σ_{xx}, σ_{xz}, σ_{yy}, σ_{zx} and σ_{zz}.
- One-dimensional and anisotropic: 8, the complex-valued k_z and six real-valued conductivity tensor elements σ_{xx}, σ_{xy}, σ_{xz}, σ_{yy}, σ_{yz} and σ_{zz},
- Two-dimensional and anisotropic: 11, the two complex-valued k_z and k_x, the seven real-valued apparent conductivity tensor elements σ_{xx}, σ_{xy}, σ_{xz}, σ_{yy}, σ_{yz}, σ_{zx} and σ_{zz},
- Three-dimensional: 15, the three complex-valued k_z, k_x and k_y, and all nine real-valued conductivity tensor elements.

Since the impedance tensor only contains up to 8 (12) known variables (the four complex-valued impedance tensor elements, 6 complex-valued elements including the tipper), the information about the apparent conductivity tensor elements are naturally ambiguous and only joint information from more

than one frequency and/or more than one site are necessary to reconstruct the subsurface structure by inversion. This quick analogy also shows that it is fundamentally more difficult to invert for a three-dimensional structure and that, in such a case, it is not possible to simplify the impedance tensor whether or not anisotropy is present.

Bibliography

Kováčiková, S., Pek, J., 2002. Generalized Riccati equations for 1-D magnetotelluric impedances over anisotropic conductors Part : plane wave field model. Earth Planets and Space 54, 473.

Martí, A., 2014. The role of electrical anisotropy in magnetotelluric responses: From modelling and dimensionality analysis to inversion and interpretation. Surveys in Geophysics 35 (1), 179.

Maxwell, J., 1873. A Treatise on Electricity and Magnetism (Clarendon, Oxford, 1873).

Schmucker, U., Weidelt, P., 1975. Electromagnetic Induction in the Earth.

Vozoff, K., 1972. The magnetotelluric method in the exploration of sedimentary basins. Geophysics 37 (1), 98.

APPENDIX B

Original Publications

Non Stationary Time Series Convolution: On the Relation between Hilbert-Huang and Fourier Transform

published in Journal of Advances in Adaptive Data Analysis
23 April 2013

Maik Neukirch and Xavier Garcia

Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar, CSIC
Pg. Maritím de la Barceloneta 37-49, 08003 Barcelona, Spain

Advances in Adaptive Data Analysis
Vol. 5, No. 1 (2013) 1350004 (13 pages)

DOI: 10.1142/S1793536913500040

World Scientific
www.worldscientific.com

NONSTATIONARY TIME SERIES CONVOLUTION: ON THE RELATION BETWEEN THE HILBERT–HUANG AND FOURIER TRANSFORM

MAIK NEUKIRCH* and XAVIER GARCIA†

Barcelona Center for Subsurface Imaging
Institut de Ciències del Mar, CSIC
Pg. Maritím de la Barceloneta 37-49, 08003 Barcelona, Spain
**neukirch@icm.csic.es*
†*xgarcia@icm.csic.es*

Received 12 April 2012
Revised 28 February 2013
Accepted 28 February 2013
Published 23 April 2013

The Hilbert–Huang Transform (HHT) decomposes time series into intrinsic mode functions (IMF) in time-frequency domain. We show that time slices of IMFs equal time slices of Fourier series, where the instantaneous parameters of the IMF define the parameters amplitude and phase of the Fourier series. This leads to the formulation of the theorem that nonstationary convolution of an IMF with a general time domain response function translates into a multiplication of the IMF with the respective spectral domain response function which is explicitly permitted to vary over time. We conclude and show on a real world application that a de-trended signal's IMFs can be convolved independently and then be used for further time-frequency analysis. Finally, a discussion is opened on parallels in HHT and the Fourier transform with respect to the time-frequency domain.

Keywords: Time series; convolution; nonstationary; Hilbert–Huang transform; Fourier transform.

1. Introduction

In digital signal processing, time series convolution is often related to the Fourier transform (FT) and therefore implies stationary and linear assumptions on the data. The reason for this prominence lies within the convolution theorem which allows to exchange a weighted integral expression to a simple multiplication, which results in much shorter computation time [Smith (1997)]. Margrave [1998] introduced the theory for nonstationary convolution filters based on the FT arguing that a continuous function is completely described by its FT and, therefore, nonstationary filtering should be possible in the frequency domain. Huang *et al.* [1998] show that the frequency information of nonstationary signals might describe the signal entirely but

gives biased information with respect to the physics of the signal, concluding that misinterpretation of the FT cannot be ruled out in a truly nonstationary setting.

Huang *et al.* [1998] introduced the Hilbert–Huang Transform (HHT), which is a new method to transform time series into a time-frequency domain without any assumptions on stationarity and linearity on the signal. The method has been extensively tested since then and successfully applied to different fields [Lo *et al.* (2008); Jackson and Mound (2010); Zeiler *et al.* (2011); Chen *et al.* (2012)], although a rigorous mathematical foundation is not yet available. The definition of HHT is empirical and data dependent, which on one hand provides a tool that works on nonstationary, nonlinear bases but, on the other hand, denies a profound understanding of the method solely based on its definition.

Despite of the lack of a classical explicit mathematical basis, extensive tests have validated HHT and suggest that it improves time series analysis, in particular in the presence of nonstationary or nonlinear effects [Huang *et al.* (2009)]. Furthermore, these tests confirm that a FT cannot reliably represent the frequency information in a nonstationary signal which, hence, require nonstationary treatment.

Very often time series include nonstationary and nonlinear effects and sometimes it is not desirable or not feasible to remove them. For instance, measurements of natural signals like the Earth's magnetic field are stationary for sufficient long periods of time, but measurements may include environmental noise which can be nonstationary [Egbert (2002); Chave and Thomson (2004); Garcia and Jones (2008)]. In this case, neither the exact noise signal nor the exact desired signal are known. Therefore, the desired stationary part cannot be isolated and the nonstationary combination of both must be analyzed. In order to solve that exact problem we discuss how the convolution filter affects nonstationary signals and extend the convolution theorem to nonstationary signals.

In this work, we present the nonstationary convolution based on HHT which does not imply assumptions on the stationarity of the signal. Since results of the HHT are neither exclusively in the time nor frequency domain, we cannot readily generalize the established convolution theorem for an analysis based on HHT but we can show, that there are fundamental similarities between the FT and the HHT with respect to convolution and use those similarities to find a new formulation for the nonstationary convolution. Due to the nature of nonstationary signals and how the frequency information can be recovered by HHT, we will argue that a nonstationary convolution based on HHT does not necessarily have an uniquely defined inverse, or a deconvolution operator resulting in the original signal, but we will briefly discuss resulting implications for the deconvolution of such signals.

The paper starts with a brief review of the HHT, highlighting the instantaneous parameters which are the backbone of our theorem. Then, it continues by presenting the formulation of the nonstationary convolution theorem and the lemma required for the subsequent proofs. The theorems are interpreted physically and their implications on the relation between FT and HHT are laid out. The paper concludes with two numerical examples on a stationary and a nonstationary test signal, and

an example of a genuine geophysical application with real world data. It follows a discussion on our findings and suggestions for further work especially with respect to nonstationary deconvolution.

2. The HHT

Huang *et al.* [1998] introduced the HHT and described thoroughly its application, restrictions and direct results. Given a de-trended time series $f(t) = f_{\text{total}}(t) - f_{\text{trend}}(t)$, the HHT generalizes the Fourier series

$$f(t) = \sum_n F_n \cdot e^{2\pi i \nu_n t} = \sum_n F_n \cdot e^{i\phi_n(t)}, \tag{1}$$

where the phase is defined as $\phi_n(t) = 2\pi\nu_n t$, to a series with an amplitude $\hat{F}(t)$ and frequency $\hat{\nu}(t)$

$$f(t) = \sum_j \hat{F}_j(t) \cdot e^{2\pi i \int_{-\infty}^{t} \hat{\nu}_j(t')dt'} = \sum_j \hat{F}_j(t) \cdot e^{i\hat{\phi}_j(\hat{\nu}_j(t),t)}, \tag{2}$$

with the phase $\hat{\phi}_j(\hat{\nu}_j(t), t) = 2\pi \int_{-\infty}^{t} \hat{\nu}_j(t')dt'$. Note that the range of the index $n \in \mathbb{Z}$ depends on the definition of the amplitudes F_n as usual for the FT. For $F_n \in \mathbb{R} : n \in \mathbb{Z}$ and for $F_n \in \mathbb{C} : n \in \mathbb{N}^0$ with $\mathbb{N}^0$ being the natural numbers including zero. Both definitions are equivalent, so let us concentrate on the complex definition for F_n in this work. On the other hand, $\hat{F}_j \in \mathbb{R}$ always is a real amplitude of the signal and is defined for $j \in \mathbb{N}^+$ for an infinite long function $f(t)$ with $\mathbb{N}^+$ being the natural numbers exclusive zero.

The intrinsic mode functions (IMF) $m_j(t)$ of $f(t)$ are defined as

$$m_j(t) = \hat{F}_j(t) \cdot e^{i\hat{\phi}_j(\hat{\nu}_j(t),t)} \tag{3}$$

by the following properties:

(i) In the whole data set, the number of extrema and the number of zero-crossings must either equal or differ at most by one, and
(ii) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

Note that within the frame of the Fourier expansion, F_n describes the constant complex amplitude of the mono-frequency part (with ν_n) of the signal $f(t)$ for the entire process. Whereas, $\hat{F}_j(t)$ is the real amplitude of the IMF j which exhibits a frequency $\hat{\nu}_j(t)$, which both can vary over time.

In other words, the HHT separates narrow-bandwidth amplitude modulations (AM) $a = \hat{F}_j(t)$ and phase modulations (PM) $p = e^{i\hat{\phi}_j(t)}$ from the data and provides them in form of real-valued IMFs [Huang *et al.* (2009)]. This process is called empirical mode decomposition (EMD) of $f(t)$ and yields the corresponding, real-valued IMFs, which represent the real part of Eq. (3). The AM values a describe time varying signal power, whereas the PM p only contain pure oscillations. The real

values of p are in the open interval between -1 and $+1$ and are defined such that they are locally zero-mean functions, which means that the number of extrema and number of zero-crossings do not differ by more than one (for detailed information, see Huang *et al.* [2009]). Huang *et al.* [1998] argue that phase functions with these properties can be Hilbert-transformed to acquire their analytic signal and that they exhibit a physical meaningful instantaneous frequency. The Hilbert Transform of a suitable function $p(t)$ is defined by

$$H(p)(t) = \frac{1}{\pi} p.v. \int_{-\infty}^{\infty} \frac{p(\tau)}{t-\tau} d\tau, \tag{4}$$

where $p.v.$ indicates Cauchy's principal value. We can construct the analytic signal by

$$m(t) = a(p(t) + iH(p)(t)), \tag{5}$$

where $H(p)(t)$ is the Hilbert Transform Eq. (4) and, thus, obtain the signal's phase using

$$\hat{\phi}(t) = \arctan\left(\frac{H(p)(t)}{p(t)}\right). \tag{6}$$

Ultimately, the instantaneous frequency is defined as the time derivative of the phase:

$$\hat{\nu}(t) = \frac{d\hat{\phi}(t)}{2\pi dt}. \tag{7}$$

3. Nonstationary Convolution Under the HHT

Let us consider an integrable function $f : t \in \mathbb{R} \to \mathbb{R}$ in the integral formulation of the FT pairs (taken from Wikipedia [2011] with references therein):

$$F(\nu) = \int_{-\infty}^{\infty} f(t) \cdot e^{-2\pi i \nu t} dt = \mathcal{F}(f(t)), \tag{8}$$

$$f(t) = \int_{-\infty}^{\infty} F(\nu) \cdot e^{2\pi i \nu t} d\nu = \mathcal{F}^{-1}(F(\nu)), \tag{9}$$

with $\phi(\nu, t) = 2\pi\nu t$, and a general IMF $m : t \in \mathbb{R} \to \mathbb{R}$:

$$m(t) = \hat{M}(t) \cdot e^{i\hat{\phi}(t)}, \tag{10}$$

with $\hat{M}(t)$ representing the real-valued, instantaneous amplitude of m. Let us define a complex, time-frequency amplitude function $\tilde{M}(\hat{\nu}, t) = \hat{M}(t) \cdot e^{i\tilde{\phi}(\hat{\nu},t)}$ and a phase function $\tilde{\phi}(\hat{\nu}(t), t) = \hat{\phi}(t) - 2\pi\hat{\nu}t$ to rewrite the IMF m as

$$m(t) = \tilde{M}(\hat{\nu}, t) \cdot e^{2\pi i \hat{\nu} t}. \tag{11}$$

Theorem 1. *Let $m(t)$ be an IMF with instantaneous frequency $\hat{\nu}(t)$ and $F^{\tau}(\nu)$ a member of the group of Fourier Transforms of $f^{\tau}(t)$, where $\tau \in \mathbb{R}$ is the parameter*

which describes each member, then:

$$m(t) \cdot F^{\tau}(\hat{\nu}(t)) = m(t) * f^{\tau}(t). \tag{12}$$

The frequency wise multiplication of m with F equals the convolution of m with f.

Lemma 1. *Let $m(t)$ be an IMF with instantaneous frequency $\hat{\nu}(t)$, then the convolution of $m(t)$ with the delta distribution $\delta(t)$ is*

$$(m * \delta)(t_0) = (\tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t} * \delta(t))(t_0), \tag{13}$$

with $\tilde{M}_{t_0} = \hat{M}(t_0) \cdot e^{i\hat{\phi}(t_0) - 2\pi i \hat{\nu}(t_0) t_0} = \tilde{M}(\hat{\nu}(t_0), t_0)$ being the complex amplitude of a monochromatic oscillation with frequency $\hat{\nu}(t_0)$.

The proof of this lemma is trivial, but we include it to stress that this property of the IMF can be used to find more properties of the HHT with the help of well-known properties of the FT. Figure 1 provides a graphical illustration of this lemma.

Proof. (*for Theorem* 1) Starting on the right-hand side (RHS) from the following identity:

$$(m(t) * f^{\tau}(t))(t) = (\delta * m * f^{\tau})(t), \tag{14}$$

focusing on an isolated time instant t_0

$$((\delta * m) * (f^{\tau}))(t_0) \tag{15}$$

and using the sifting property of the delta function with Lemma 1, the RHS yields

$$\begin{aligned}((\delta * m) * f^{\tau})(t_0) &= [(\delta * \tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t}) * f^{\tau}(t)](t_0) \\ &= [\delta * \tilde{M}_{t_0} \cdot \mathcal{F}^{-1}(\delta(\hat{\nu}(t_0) - \nu)) * \mathcal{F}^{-1}(F^{\tau}(\nu))](t_0).\end{aligned} \tag{16}$$

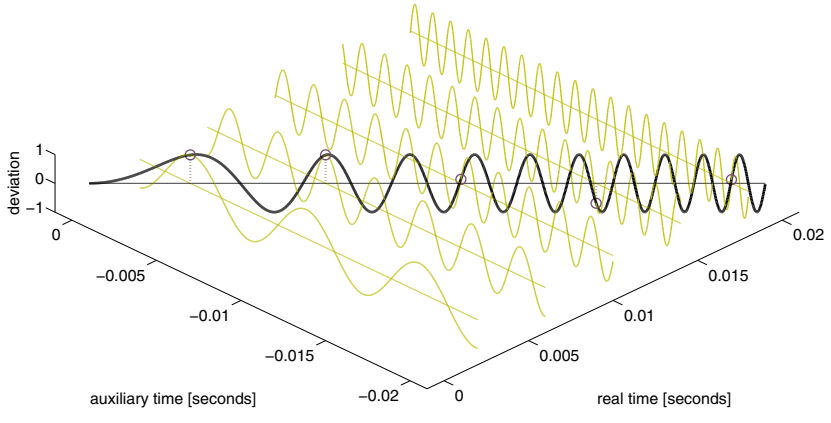


Fig. 1. This figure illustrates Lemma 1. The black curve is an IMF (a chirp function with linearly increasing frequency and constant amplitude, Sec. 6.2) and the light curves are sine curves with the frequency, phase and amplitude chosen to be identical with the IMF at the intersection points marked by reddish circles. The IMF evolves along the actual time axis, whereas the sine curves are displayed in an auxiliary time domain whose sole purpose is to visualize them. The vertical axis describes the deviation for both, the IMF and the sine curves.

When we resolve the convolution with the help of the convolution theorem for FT, we obtain

$$\begin{aligned}((\delta * m) * f^{\tau}(t))(t_0) &= [\delta * \tilde{M}_{t_0} \cdot \mathcal{F}^{-1}(\delta(\hat{\nu}(t_0) - \nu) \cdot F^{\tau}(\nu))](t_0) \\ &= \left[\delta * \tilde{M}_{t_0} \cdot \int_{-\infty}^{\infty} \delta(\hat{\nu}(t_0) - \nu) \cdot F^{\tau}(\nu) \cdot e^{2\pi i \nu t} d\nu\right](t_0) \\ &= [\delta * \tilde{M}_{t_0} \cdot F^{\tau}(\hat{\nu}(t_0)) \cdot e^{2\pi i \hat{\nu}(t_0) t}](t_0) \\ &= \tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t_0} \cdot F^{\tau}(\hat{\nu}(t_0)). \end{aligned} \tag{17}$$

Now, the left-hand side (LHS) of Theorem 1 can be written as:

$$\begin{aligned}(\delta * m(t) \cdot F^{\tau}(\hat{\nu}(t)))(t_0) &= (\delta * \hat{M}(t) \cdot e^{i\hat{\phi}(t)} \cdot F^{\tau}(\hat{\nu}(t)))(t_0) \\ &= (\delta * \hat{M}(t) \cdot e^{i\hat{\phi}(t) - 2\pi i \hat{\nu}(t) t + 2\pi i \hat{\nu}(t) t} \cdot F^{\tau}(\hat{\nu}(t)))(t_0) \\ &= \hat{M}(t_0) \cdot e^{i\hat{\phi}(t_0) - 2\pi i \hat{\nu}(t_0) t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t_0} \cdot F^{\tau}(\hat{\nu}(t_0)), \end{aligned} \tag{18}$$

for $\tilde{M}_{t_0} = \hat{M}(t_0) \cdot e^{i(\hat{\phi}(t_0) - 2\pi \hat{\nu}(t_0) t_0)}$ as required for Lemma 1, LHS and RHS of Theorem 1 are equivalent for all times t_0. □

Proof. (*for Lemma 1*) The proof is a straight forward application of the sifting property of the delta distribution in a convolution and the insertion of a zero term. The LHS can be reformulated as follows by using Eq. (11):

$$\begin{aligned}(m * \delta)(t_0) &= (\tilde{M}(\hat{\nu}(t), t) \cdot e^{2\pi i \hat{\nu}(t) t} * \delta(t))(t_0) \\ &= \tilde{M}(\hat{\nu}(t_0), t_0) \cdot e^{2\pi i \hat{\nu}(t_0) t_0}. \end{aligned} \tag{19}$$

$\tilde{M}(\hat{\nu}, t)$ can be understood as a complex-valued, instantaneous amplitude which incorporates an instantaneous starting phase in order to linearize the phase term of the IMF m. By reformulating the RHS of Lemma 1, we find

$$(\tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t} * \delta)(t_0) = \tilde{M}_{t_0} \cdot e^{2\pi i \hat{\nu}(t_0) t_0} \tag{20}$$

and readily see that both sides are equivalent for $\tilde{M}_{t_0} = \tilde{M}(\hat{\nu}(t_0), t_0)$. □

4. Physical Interpretation

4.1. *Theorem 1*

Assuming we know exactly the spectral response $F(\nu)$ of a physical measurement device, this theorem states that we can simply multiply the known spectral response $F(\nu)$ by a known time signal $m(t)$ in order to obtain the signal $n(t)$ measured by the device, if the signal $m(t)$ is an IMF. Per definition, $n(t)$ is exactly the convolution of the time domain response function $f(t)$ of the device with the incoming signal $m(t)$. Our reformulation of a nonstationary convolution to a simple multiplication leads to a better understanding of the behavior of physical systems in a nonstationary set up and further increases the application range of the HHT.

An important note is that a de-trended signal $x(t)$ should convolve in the same manner as if each of its IMFs $m_j(t)$ are convolved independently to $n_j(t)$ and then are summed over to $y(t)$. Unfortunately, the nonstationary character of IMFs cannot guarantee that the convolution of an IMF results in another IMF; thus, it may not be allowed to sum over n_j in order to form the total convolved signal y. If the present nonstationarity is too severe in m_j or in the transfer function $F(\nu, t)$ the convolution of m_j cannot yield another IMF, because the convolution may introduce new extrema without additional zero-crossings to the function which is not permitted in the definition of an IMF. In such a case it remains an open question whether n_j still are base functions of y. Certainly, the total convolved signal y cannot decompose into n_j if not all n_j qualify as IMFs. This restriction on the inverse to our theorem depends very much on the nonstationary phase-time relation of signal and transfer function and may be discussed in detail in another work. Here, we only want to stress that the convolution results of IMFs do not need to be IMFs and may not always be summed up in order to compose a total convolved signal of a general time series. However, the theorem will always apply to a signal that is an IMF all by itself, even though the convolved result may or may not be an IMF after the convolution.

Later we will discuss an example for which we can add up the convolved IMFs of a signal in order to get the total convolved signal. Moreover, in that example we will use the spectral information given by the IMFs of a signal and its convolution in order to estimate the system's transfer function. Therefore, we claim that the convolution of a signal's IMFs may well describe physical properties of signal convolution.

4.2. *Lemma 1*

This lemma states that the instantaneous parameters of the IMFs at any time $t = t_0$ can be used as parameters of a sine curve to fully describe the IMF at that time. It provides a link between the HHT and the FT and can likely be used to find more properties of the HHT with the help of well-known properties of the FT, since a sine curve is the fundamental base of the FT. Note that the FT is defined on an infinite time axis and that the time axis of this "virtual" sine curve is not equivalent to the one of the original signal but rather to an imaginary, infinite one, therefore even a piece of a continuous time signal can be described by these virtual sine curves and anything that applies to the entire virtual sine curve also applies to the IMF at time t_0. The Hilbert spectrum is the common mean to visualize the time-frequency behavior of an IMF and therefore, we suggest to refer to the Hilbert spectrum as two-dimensional, time-evolving Fourier spectrum.

5. On the Relation between HHT and FT

First of all, we would like to propose the term "time-varying FT", which we define as the RHS of Eq. (13). Virtually every function $f(t) : t \in \mathbb{R} \to \mathbb{C}$ can be represented by

a Fourier amplitude in this way but it is worth noting that the implications coming along with the definitions of an IMF, like that it must have a physically meaningful instantaneous frequency, give meaning to the time-varying Fourier amplitudes as physical representations of the IMF in the time and frequency domain.

When we apply Lemma 1 [Eq. (13)] to Eq. (2), which is the original formulation of the HHT as taken from Huang *et al.* [1998], we find that the function $f(t)$ is represented by a series of time-varying FTs:

$$(f * \delta)(t_0) = \sum_j (m_j * \delta)(t_0) = \sum_j (\tilde{M}^j_{t_0} \cdot e^{2\pi i \hat{\nu}_j(t_0)t} * \delta)(t_0), \tag{21}$$

where $j \in \mathbb{N}^+$ is the order of the IMF and $\hat{\nu}_j$ assumes the spectral coordinates of the signal f. This formulation represents (in time slices) how the time-frequency information, obtained from the HHT, is commonly displayed: the Hilbert spectrum.

6. Two Sandbox Examples — Sine and Chirp: Two Synthetic Signals

In order to demonstrate Theorem 1, we designed two synthetic time series, the first superposes two sine curves with distinct frequency values and the second is a chirped sine with a linearly increasing frequency. Both of these signals are then subjected to the convolution with a 1st-order Butterworth low pass filter.

6.1. *Stationary signal — two sine curves*

The first example is stationary and validates that Theorem 1 holds for the conventional convolution theorem, which states that the convolution of two time series is the inverse FT of the multiplication of the FT of each time series. The superposition of two sine functions, here one with a frequency of 5 Hz and another with 5 kHz, is decomposed equally by the HHT and FT. Both transforms find the exact same two oscillations with the constant amplitude from the signal. The signal is displayed as a black line in Figs. 2(a) and 2(b) highlighting its slow and fast oscillation, respectively. In the same graphs are the convolution results with two distinct Butterworth filter. The filter are set up as 1st-order low pass filter with normalized cut-off frequencies at 0.05 and 0.005. The convolution with both filters has been applied in forward and reverse direction in order to achieve zero phase filter with amplitudes as shown in Fig. 2(c). Figures 2(a) and 2(b) show that the amplitude of the low frequency oscillation is not affected as both filter are in the pass band, whereas the high frequency part is damped according to the amplitude value of the corresponding frequency and filter. We tested three methods, FT based filtering, time series convolution filtering and the nonstationary convolution theorem as presented in this work. All three methods yield exactly the same result.

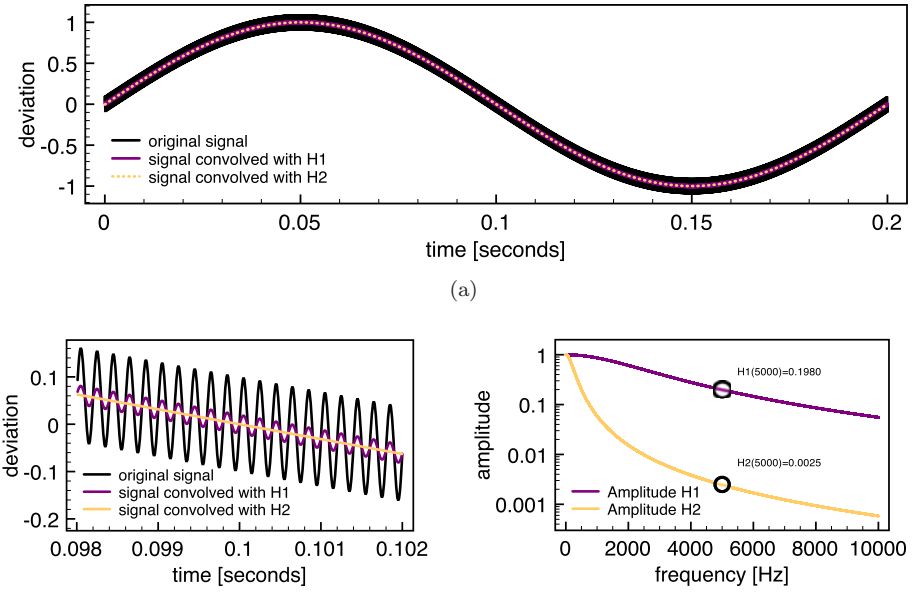


Fig. 2. This figure illustrates the first test signal represented by a superposition of two sine functions and its convolution with a 1st-order low pass filter. The signal is completely stationary and the convolution can be carried out in the time domain, with the FT or with our theorem and yields the exact same results. (a) The first synthetic test signal is a superposition of two sine functions with a frequency of 5 Hz and 5 kHz, respectively. The colored lines show the same signal filtered by 1st-order low pass Butterworth filter with normalized cut-off frequencies at 0.05 and 0.005. (b) This zoom-in around 0.1 s of Fig. 2(a) highlights the details of the test signal. (c) The amplitude spectra of the filter that are used here have distinct values at 5 kHz.

6.2. *Nonstationary signal — Chirped sine with linear frequency*

The second example on synthetic data is on a pure, nonstationary signal in the form of a chirped sine function with a linearly increasing frequency, which is as per definition an IMF. The signal is plotted in Fig. 3 as a black line with its frequency axis at the top and the time axis at the bottom. Note that the very same signal illustrates Lemma 1 in Fig. 1. To perform a convolution, we use the Butterworth filter with the cut-off frequency at 0.05 as described for the previous example. The filter is again set up as zero phase filter with the amplitude displayed as blue line in Fig. 3. The convolution is carried out via the time series convolution and via the nonstationary convolution theorem. Both results are displayed in Fig. 3 and both are almost identical. The only difference is that the time series convolution algorithm cannot deal with the beginning of the time series, since it is defined as a weighted sum that requires values around the location where it calculates the convolution but there exist no values lower than $t = 0$ so the algorithm assumes zero-padding and experiences "edge problems". The calculation based on the nonstationary convolution

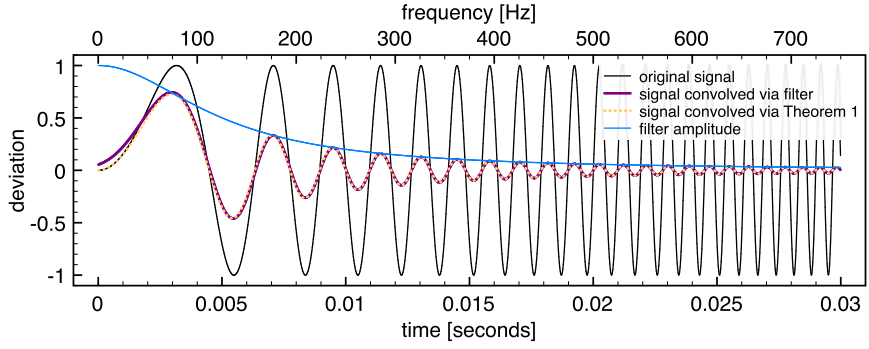


Fig. 3. The second test signal is a chirp function with a linearly increasing frequency over time and constant amplitude. This figure compares the convolution of the chirp with a 1st-order low pass Butterworth filter applied with two methods: time domain convolution and Theorem 1 from this paper. Besides the numeric errors at the edge from the numeric convolution algorithm, both methods yield the same result.

theorem uses only local information and therefore is exact as well at the edges, where the frequency is very low and there is virtually no effect of the low pass filter to be expected. In fact, we tested different filter types (Chebyshev I and II, Elliptic and Bessel analog filter) with the order up to 10 and for several cut-off frequencies, and the convolution theorem presented here gave accurate results for any sampling rate as it is defined on the local, instantaneous parameters, whereas numeric filter procedures depend much on sampling as it relies on weighted sums.

7. A Real World Example — Magnetotelluric Transfer Functions

In this section, we want to present an application of Theorem 1. The authors' field of expertize lies in magnetotelluric (MT) exploration and, hence, the theorem has been developed in light of processing MT data. MT measurements log the natural variation of magnetic and electrical (telluric) fields at the Earth's surface and these measured time series can be statistically analyzed to obtain the relative spectral relation of the electrical to the magnetic field Vozoff [1972]. The subsurface conductivity structure enforces a particular distribution of underground currents, which alter the external natural electromagnetic field of the Earth and, therefore, it allows us to derive that structural information of the subsurface conductivity from the analysis of the electromagnetic field on the surface.

Knowing the spectra of the surface electrical field $E(\omega)$ and the spectra of the surface magnetic field $H(\omega)$, we can write the relation between the horizontal spatial field components as

$$\begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix} = \begin{pmatrix} Z_{xx}(\omega) & Z_{xy}(\omega) \\ Z_{yx}(\omega) & Z_{yy}(\omega) \end{pmatrix} \cdot \begin{pmatrix} H_x(\omega) \\ H_y(\omega) \end{pmatrix}, \tag{22}$$

where Z is the impedance which describes the subsurface conductivity volume for a given frequency ω. With respect to the work described in this paper, Z can be

simply understood as the system response function of the Earth, H is the input and E is the output of the convolution. Electrical and magnetic fields are recorded as time series and need to be transformed into the frequency domain in order to solve for the impedance, because the impedance tensor is only defined in the frequency domain. Under ideal conditions, the electromagnetic field varies quasi-stationary, meaning the spectral composition changes sufficiently slow that a windowed FT can be performed, but for sites closer to inhabited or industrial areas, cultural noise often affects the measurements severely. Cultural noise can be of any kind and is most often nonstationary, therefore, measurements of $e(t)$ and $h(t)$ are often disturbed by nonstationary variations, since the physical measurements contain both, natural signal and cultural noise.

Clearly, the problem described here is not exactly in the format of the theorem where we know input and system response and seek the output but it is similar; we do know input and output and need to find the system response by an optimization procedure. Thus, in this example we also have to assume Theorem 1 to hold in order to search for the optimal solution.

The algorithm that solves for this MT data is too complex to be discussed here in detail and will be treated in its entirety in a different work, but we do need to stress that Eq. (22) states a multivariate optimization problem, which requires the use of a special EMD introduced by Rehman and Mandic [2010] and designed for multivariate data but for our purpose it performs an EMD no different than the univariate EMD, only that it ensures data channel correlation within the index of the IMFs (e.g. IMF number two of channel e_x will be at a similar time scale as any other channel's IMF number two).

Looking at the MT problem from the point of this work, $E = \mathcal{F}(e)$ is the output or result of the convolution (electrical field spectra), $H = \mathcal{F}(h)$ is the convolution input (magnetic field spectra) and the system response is $Z(\nu) = \mathcal{F}(z(t))$ (Impedance) with the unknown temporal system response function $z(t)$

$$E = Z \cdot H \tag{23}$$

or in time domain

$$e = z * h. \tag{24}$$

Both, e and h, are then decomposed into their respective IMFs with the algorithm by Rehman and Mandic [2010], which ensures that for both signals the time scales remain correlated throughout the decomposition process. Theorem 1 suggests that

$$m_e^j = z * m_h^j = Z \cdot m_h^j \tag{25}$$

with m^j being the respective IMFs for input and output.

Using Eq. (25) and a statistical optimization, we find an optimal solution for $Z(\nu)$ for the instantaneous parameters given by m_h^j and m_e^j. The results of the impedance tensor for the test data set is presented in Fig. 4 in dark color. The curves in bright color correspond to a FT-based algorithm processing the same data set.

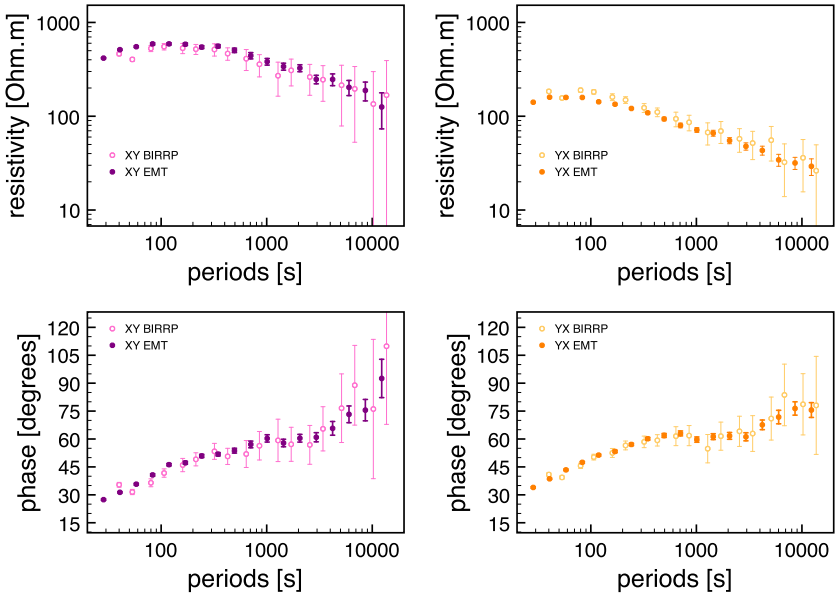


Fig. 4. MT data is given as an example application of the nonstationary convolution theorem with the electric impedance as the Earth's system response function relating magnetic to electric fields. The main entries of the impedance tensor Z are displayed as amplitude and phase over a range of periods. It describes the subsurface electric conductivity structure and can be used to detect electric anomalies like phase boundaries, ore minerals or water deep inside the Earth.

Note that the last example shown here also suggests that the convolved IMFs from a signal can be added to produce the complete convolved signal if the results still are IMFs. Here, we used measured time series of the signal and its convolution, h and e, respectively, and carried out the convolution on their respective IMFs to find the system response function, thus it proofs that the sum of the convolved IMFs indeed reproduce the convolved measurement, even though the convolution has been carried out on each IMF independently. Our example only uses stationary transfer functions with a relatively constant phase, therefore it shall be said that other more complex cases may yield a different experience. For example, let there be a high degree of nonstationarity in the phase-time relation between the IMF and the transfer function, then the IMF-system characteristics may preclude the convolution to result in another IMF, because some situations may alter the rotational sense of the phase and, hence, introduce new extrema without additional zero-crossings. However, the observation is interesting enough that we feel it would deserve a more thorough investigation in another work.

8. Conclusion

- The convolution between an IMF and a temporal system response can be translated into a multiplication.

- The Hilbert–Huang spectra can be represented as a Fourier spectra with time varying complex amplitude.

Acknowledgments

We express our gratitude to Daniel Rudolf from the Mathematical Institute of the Friedrich Schiller University of Jena, Germany, for his extensive reviews and fruitful discussions and thank Savitri Galiana from the Barcelona Center for Subsurface Imaging for her revision and constructive comments. This work has been conducted under the umbrella of the industry collaboration project CO-DOS funded by Repsol.

References

Chave, A. D. and Thomson, D. J. (2004). Bounded influence magnetotelluric response function estimation. *Geophys. J. Int.*, **157**(3): 988–1006.

Chen, J., Heincke, B., Jegen, M. and Moorkamp, M. (2012). Using empirical mode decomposition to process marine MT data. *Geophys. J. Int.*, pp. 293–309.

Egbert, G. D. (2002). Processing and interpretation of electromagnetic induction array data. *Surveys Geophys.*, **23**(2): 207–249.

Garcia, X. and Jones, A. G. (2008). Robust processing of magnetotelluric data in the AMT dead band using the continuous wavelet transform. *Geophysics*, **73**(6): F223.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. R. Soc. A, Math. Phys. Eng. Sci.*, **454**(1971): 903–995.

Huang, N. E., Wu, Z., Long, S. R., Arnold, K. C. and Chen, X. (2009). On instantaneous frequency. *Adv. Adapt. Data Anal.*, **1**(2): 177–229.

Jackson, L. P. and Mound, J. E. (2010). Geomagnetic variation on decadal time scales: What can we learn from empirical mode decomposition? *Geophys. Res. Lett.*, **37**(14).

Lo, M., Hu, K., Liu, Y., Peng, C. K. and Novak, V. (2008). Multimodal pressure-flow analysis: Application of Hilbert Huang transform in cerebral blood flow regulation. *EURASIP J. Adv. Signal Process.*, 2008: 785243, DOI: 10.1155/2008/785243.

Margrave, G. F. (1998). Theory of nonstationary linear filtering in the Fourier domain with application to time-variant filtering. *Geophysics*, **63**(1): 244–259.

Rehman, N. and Mandic, D. P. (2010). Multivariate empirical mode decomposition. *Proc. R. Soc. A, Math., Phys. Eng. Sci.*, **466**(2117): 1291–1302.

Smith, S. W. (1997). The scientist and engineer's guide to digital signal processing. ISBN: 0-9660176-3-3. *California Technical Publishing*. San Diego, CA, USA.

Vozoff, K. (1972). The magnetotelluric method in the exploration of sedimentary basins. *Geophys.*, **37**: 98–141.

Wikipedia (2011). Fourier Transform — Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/wiki/Fourier_transform, accessed July 18th 2011.

Zeiler, A., Faltermeier, R., Brawanski, A., Tomé, A. M., Puntonet, C. G., Górriz, J. M. and Lang, E. W. (2011). Brain status data analysis by sliding EMD. *Proc. 4th Int. Conf. Interplay Between Natural and Artificial Computation*: *New Challenges on Bioinspired Applications*, Vol. Part II. Springer-Verlag, Berlin, Heidelberg, pp. 77–86.

Non Stationary Time Series Convolution: Frequency Shift Caused by Convolution

(submitted to Journal of Advances in Adaptive Data Analysis)

Maik Neukirch and Xavier Garcia

Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar, CSIC
Pg. Maritím de la Barceloneta 37-49, 08003 Barcelona, Spain

Advances in Adaptive Data Analysis

Non Stationary Time Series Convolution: Frequency Shift Caused by Convolution

Maik Neukirch

Barcelona Center for Subsurface Imaging, Institut de Ciènces del Mar, CSIC
Pg. Maritím de la Barceloneta 37-49, 08003 Barcelona, Spain
neukirch@icm.csic.es
http://<www.barcelona-csi.cmima.csic.es>

Xavier Garcia

Barcelona Center for Subsurface Imaging, Institut de Ciènces del Mar, CSIC, Spain
xgarcia@icm.csic.es

Received 04 March 2013

By using the Hilbert-Huang Transform, a non stationary time series can be represented by a number of modes, which are complex time series with instantaneous amplitudes, phases and frequencies. Following the non stationary convolution theorem which allows to translate a convolution into a multiplication, we analyse the characteristics of a convolved time series and show that through convolution the instantaneous frequency may change. We quantify the frequency shift and argue that this difference greatly hampers any attempt to deconvolve non stationary signals.

1. Introduction

The Hilbert-Huang Transform (HHT, Huang et al., 1998) is a novel tool to analyse non stationary time series and describes them with their instantaneous, spectral information. HHT decomposes a time series into a number of zero-mean, oscillatory modes, called Intrinsic Mode Functions (IMF), in order to ensure existence of an interpretable analytic signal of each IMF and to express the analytic signal in terms of time series of the instantaneous parameters: amplitude, phase and frequency. The IMF reside in the time and frequency domain and are described by amplitude and phase as functions of time, where the time derivative of the phase constitutes the frequency.

Neukirch and Garcia (2013) present a non stationary convolution theorem that is similar to the convolution theorem for Fourier transform but that does not imply assumptions on the stationarity of the signal since it is based upon the definitions of the IMFs of the Hilbert-Huang Transform. However, they argue that such a non stationary convolution does not necessarily have an uniquely defined inverse, or a deconvolution operator resulting in the original signal, and we wish to continue this

discussion focussing on some resulting implications for the deconvolution of non stationary signals.

Since the convolution of a non-stationary time series with a response function in the time domain can be transformed into a basic algebraic formulation, in this work, we focus on the repercussions of a non-stationary convolution by analysing the instantaneous phase and its time derivative. Most notably, we find that there can be a frequency shift in the resulting signal with respect to the original signal depending on the degree of non stationarity. This finding may be important for non stationary time series, which are filtered by a system response for technical reasons, as it is often the case for physical measurements.

2. Hilbert-Huang Transform and Non-Stationary Convolution

In the Hilbert-Huang Transform (Huang et al., 1998, 2009), the Intrinsic Mode Functions (IMF) $m_j\left(t\right)$ of $x\left(t\right)$ are defined as

$$m_j\left(t\right) = m_{0,j}\left(t\right) \cdot e^{i\phi_j(t)}. \tag{1}$$

with $\phi_j\left(t\right) = \int_{-\infty}^{t} \omega_j\left(t'\right) dt'$.

In essence, the HHT separates narrow-bandwidth amplitude modulations (AM) $m_j(t)$ and phase modulations (PM) $\phi_j(t)$ from the data and provides them in form of IMFs. This process is called Empirical Mode Decomposition (EMD). Since the phase of the signal is well defined, the instantaneous frequency can be derived from the phase by:

$$\omega_m(t) = \frac{d\phi_m\left(t\right)}{dt}. \tag{2}$$

Neukirch and Garcia (2013) show that the convolution of an IMF $m_j(t)$ with any time domain system response function $s(t)$ translates into a complex multiplication of the IMF with the frequency domain representation $S(\omega_m, t)$ of that response function.

$$m_j(t) * s(t) = m_j(t) \cdot S(\omega_m(t), t) \tag{3}$$

Table 1: These conventions are used in the course of this article.

t	time
$\phi(t)$	phase function
$\omega(t) = \dot{\phi}$	instantaneous angular frequency
$m(t)$	Intrinsic Mode Function (IMF)
$s(t)$	temporal system response function
$S(\omega_m, t)$	spectral system response function
$x(t)$	convolution result of s and m
$\Delta\omega(t)$	frequency shift

(a) Table of Functions

m	relates to original signal
s	relates to system response
x	relates to convolved signal
0	identifies amplitudes
j	order of IMF

(b) Table of Subscripts

Equation 3 simplifies non stationary convolution drastically and one may infer that the same is true for the equally interesting deconvolution of signals. Unfortunately, the simplicity of the inverse operation for multiplication is misleading here and distracts from the fact that there can be quite different phase functions be involved, which may not necessarily be known. Let us shed light on the problematic with an analysis of the phase function and allow us to ignore the amplitude functions in this work.

3. Phase Analysis of Convolved Time Series

We separate the complex values $m(t)$ and $S(\omega_m(t), t)$ of Equation (3) for a single IMF in the amplitudes $m_0(t) \in \mathbb{R}^+$ of the IMF m and $S_0(\omega_m(t), t) \in \mathbb{R}^+$ of the response function S, and in the phases $\phi_m(t) \in \mathbb{R}$ of m and $\phi_s(\omega_m(t), t) \in \mathbb{R}$ of S. Note, that both, the amplitude and phase of the response function, are functions of the instantaneous frequency $\omega_m(t) = \frac{d\phi_m(t)}{dt} = \dot{\phi}_m(t)$ of m and the time $t \in \mathbb{R}$. Then the convolution $x(t) = m(t) * s(t)$ writes:

$$x_0 \exp(i\phi_x) = m_0 \exp(i\phi_m) \cdot S_0\left(\dot{\phi}_m, t\right) \exp\left(i\phi_s(\dot{\phi}_m, t)\right) \tag{4}$$

with the following amplitude and phase functions:

$$x_0 = m_0 \cdot S_0\left(\dot{\phi}_m, t\right), \tag{5}$$

$$\phi_x = \phi_m + \phi_s\left(\dot{\phi}_m, t\right). \tag{6}$$

The observed amplitude x_0 is a function of amplitude and phase of the IMF m whereas the observed phase ϕ_x is independent of the amplitude, therefore, in this work, we restrict the analysis to the phase and leave the amplitudes for another time. The time derivative of Equation (6) yields:

$$\dot{\phi}_x = \dot{\phi}_m + \frac{\delta\phi_s\left(\dot{\phi}_m, t\right)}{\delta t} + \frac{\delta\phi_s\left(\dot{\phi}_m, t\right)}{\delta\dot{\phi}_m}\ddot{\phi}_m. \tag{7}$$

Hence, we find that $\dot{\phi}_x = \dot{\phi}_m$ only for the case that either

(1) the phase of the spectral system response ϕ_s is constant over time for a certain frequency $\dot{\phi}_m$ and one of both, ϕ_s is constant for a varying $\dot{\phi}_m$ or $\dot{\phi}_m$ is constant over time, or
(2) the second two summands cancel each other.

In all other cases, $\dot{\phi}_x$ will differ from $\dot{\phi}_m$ by the frequency shift

$$\Delta\omega = \frac{\delta\phi_s\left(\dot{\phi}_m, t\right)}{\delta t} + \frac{\delta\phi_s\left(\dot{\phi}_m, t\right)}{\delta\dot{\phi}_m}\ddot{\phi}_m. \tag{8}$$

This observation tells us, that in a non stationary convolution a different instantaneous frequency may be observed than the one that the underlying process m had

before it convolved with the response function s. The difference will depend on the nature of the response function (phase behaviour over frequency and time) and on the signal itself (swiftness of changes in the instantaneous frequency). Furthermore, $x(t)$ will only retain the status of qualifying as an IMF like m if and only if the frequency shift is larger than the negative instantaneous frequency of m:

$$\Delta\omega > -\dot{\phi}_m. \tag{9}$$

If not, the phase ϕ_x will run backwards introducing new extrema without zero crossings and prohibiting x to fall into the definition of an IMF. Even if the frequency shift allows the convolved signal to fall into the definition of an IMF its mere presence may easily cause mode mixing in a time signal that contains more than one IMF, since the instantaneous frequency of one IMF can become larger/smaller than its predecessor/successor.

4. A Representative Example

Figure 1 illustrates numerically the theoretical findings of the last section. There we define an IMF as a chirped function with a linearly increasing frequency and constant amplitude (see Figure 1 (1) to (3), blue dotted line) and a system response function with a quadratic frequency-phase relation and decreasing amplitude (see Figure 1 (1) to (3), red line). Naturally, the IMF is defined as a time series but since the time-frequency relation is linear in this example, we can equally use the abscissa for both, time and frequency. Then, the system response function is defined as a spectra, but again, since the time-frequency relation is linear and unique, the same reasoning applies for the abscissa for the system response function. The convolution of both is computed via Equation 3 and displayed in Figure 1 (1) to (3) as purple dashed line. Note here, that in contrast to IMF and system response function the abscissa of the convolution represents the true time but the original frequency of the IMF and not the shifted frequency due to the convolution. The first plot illustrates how the varying amplitude of the system response function envelopes the convolution because the unitary chirp constitutes nothing to the multiplication of amplitudes in Equation 3. Plot number two and three represent the addition of phases and frequency, respectively. And lastly, the frequency shift is plotted in the fourth diagram in form of the ratio between the frequency shift and the frequency of the IMF. The frequency shift in this example increases up to the value of the original frequency, effectively doubling the observed frequency from before to after the convolution.

5. Remarks on Deconvolution

For solving a non stationary deconvolution knowing only x and S, we would need to solve Equation (7) for ϕ_m to recover the phase of m, ignoring the amplitudes for now. Clearly, if Equation (8) is not equal to zero, solving Equation (7) will be challenging and might only yield a solution via an iterated optimisation algorithm.

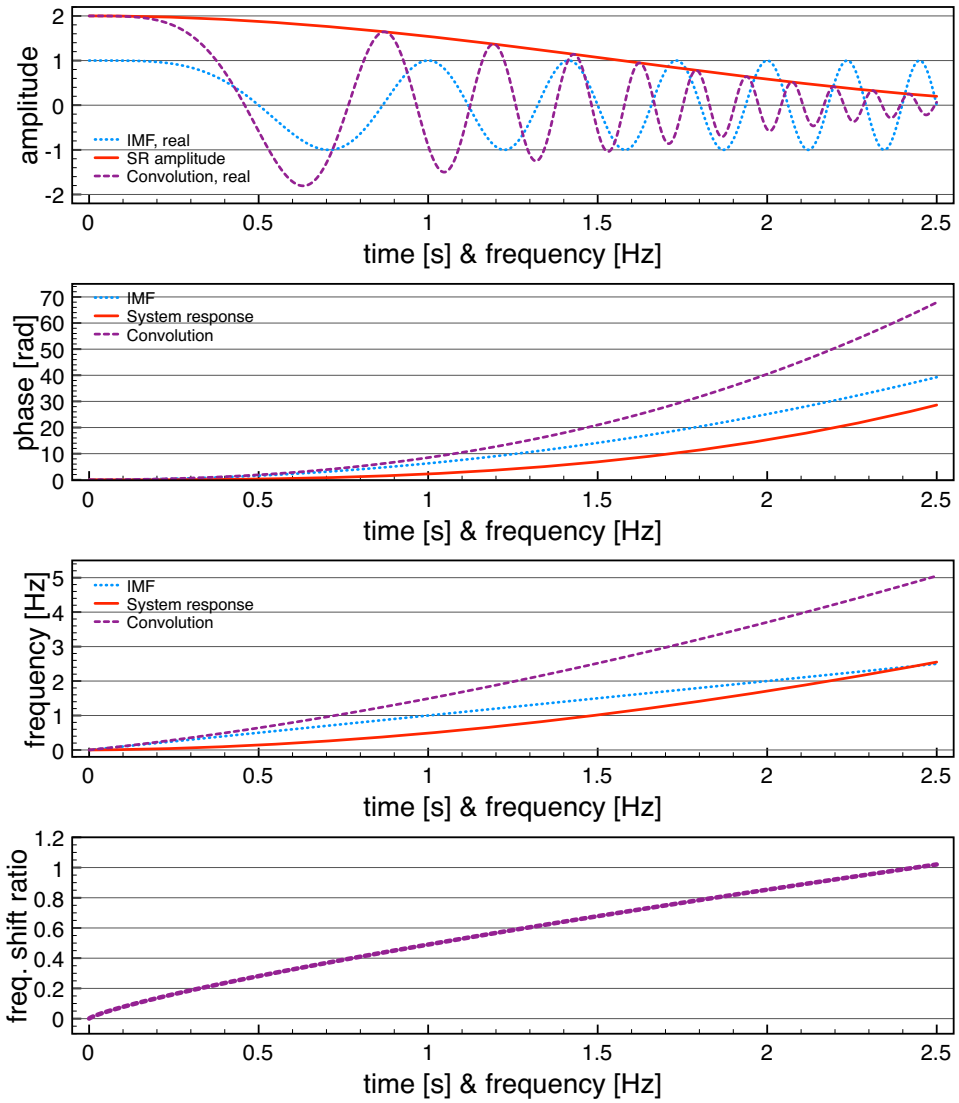


Fig. 1: From top to bottom: (1) a Intrinsic Mode Function (unitary chirp), the amplitude of a spectral system response and the convolution of IMF and SR, (2) phase functions, (3) phase gradients and (4) the ratio between the frequency shift and original frequency.

Furthermore, since x does not need to qualify for an IMF in theory, it may be impossible to find the correct ϕ_x directly from a convolved time series x by means of HHT.

6. Conclusion

Convolution of non stationary time series with a general system response function may alter the characteristic time scale of the time series and introduce a shift in the instantaneous frequency depending on the characteristics of the convolved system response function and the instantaneous frequency of the original signal. This frequency shift renders any deconvolution attempt difficult as such that no analytic solution exists but optimisation may be successful.

Acknowledgments

This work was funded by Repsol under the umbrella of the project CO-DOS.

References

Huang, N., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., Liu, H. H., mar 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 454 (1971), 903–995.

Huang, N., Wu, Z., Long, S., Arnold, K., Chen, X., 2009. On instantaneous frequency. Advances in Adaptive Data Analysis 1 (2), 177–229.

Neukirch, M., Garcia, X., 2013. Non Stationary Time Series Convolution: On the Relation Between Hilbert-Huang and Fourier Transform. Advances in Adaptive Data Analysis. In press.

Non Stationary Magnetotelluric Data Processing With Instantaneous Parameters

published in Journal of Geophysical Research (21 March 2014)

Maik Neukirch and Xavier Garcia

Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar, CSIC
Pg. Maritím de la Barceloneta 37-49, 08003 Barcelona, Spain

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE

10.1002/2013JB010494

Key Points:
- Nonstationary impedance estimation for the magnetotelluric method
- Approach verified with synthetic, real, and nonstationary data

Correspondence to:
M. Neukirch,
neukirch@icm.csic.es

Citation:
Neukirch, M., and X. Garcia (2014), Nonstationary magnetotelluric data processing with instantaneous parameter, *J. Geophys. Res. Solid Earth*, *119*, doi:10.1002/2013JB010494.

Received 5 JUL 2013
Accepted 24 FEB 2014
Accepted article online 27 FEB 2014

Nonstationary magnetotelluric data processing with instantaneous parameter

M. Neukirch[1] and X. Garcia[1]

[1] Barcelona Center for Subsurface Imaging, Institute of Marine Science, CSIC, Barcelona, Spain

Abstract Nonstationarity in electromagnetic data affects the computation of Fourier spectra and therefore the traditional estimation of the magnetotelluric (MT) transfer functions (TF). We provide a TF estimation scheme based on an emerging nonlinear, nonstationary time series analysis tool, called empirical mode decomposition (EMD) and show that this technique can handle nonstationary effects with which traditional methods encounter difficulties. In contrast to previous works that employ EMD for MT data processing, we argue the advantages of a multivariate decomposition, highlight the possibility to use instantaneous parameters, and define the homogenization of frequency discrepancies between data channels. Our scheme uses the robust statistical estimation of transfer functions based on robust principal component analysis and a robust iteratively reweighted least squares regression with a Huber weight function. The scheme can be applied with and without aid of any number of available remote reference stations. Uncertainties are estimated by iterating the complete robust regression, including the robust weight computation, with a bootstrap routine. We apply our scheme to synthetic and real data (Southern Africa) with and without nonstationary character and compare different processing techniques to the one presented here. As a conclusion, nonstationary noise can heavily affect Fourier-based MT data processing but the presented nonstationary approach is nonetheless able to extract the impedances.

1. Introduction

Natural electromagnetic (EM) field variations are caused by two major working mechanisms: lightning activity at high frequencies (>8 Hz) and magnetospheric currents excited by solar wind at low frequencies (<8 Hz) [e.g., *Garcia and Jones*, 2002; *Viljanen*, 2012]. *Rakov and Uman* [2007] summarize the electromagnetic lightning discharge to three modes: (a) fast and transient leader-return stroke sequences, (b) slow and quasi-stationary continuing currents, and (c) perturbations and surges on the continuing currents. The longest lasting and most abundant in an electromagnetic time series measurements are the perturbed continuing currents, which may be viewed as being stationary on a section with some dynamic length confined by the recurrent transient strokes. *Liu and Fujimoto* [2011] conclude that the magnetospheric current is nonlinearly driven by the dynamic solar wind but behaves in a static manner for high magnetospheric pressure conditions. Both of these EM sources are naturally nonstationary, since both, lightning strokes and magnetospheric pressure conditions, are very dynamic and thus strictly limit the duration of any stationary electromagnetic signal.

Practitioners argue that the magnetotelluric (MT) signal is quasi-stationary (stationary on reasonably long time windows) and, thus, justify the application of the windowed Fourier transform. In practice, this procedure works very well for data with high signal-to-noise ratios but frequently encounters problems in the presence of electromagnetic noise (clearly what is called noise here would include nonstationary signal) [cp. *Junge*, 1996]. A concise treatise of sophisticated MT signal processing based on the Fourier transform is given by *Chave* [2012] in which nonstationarity is listed as one of the problems that affect transfer function estimation.

For instance, if there would be a nonstationary electric discharge, the window (data segment) of this event would not qualify as containing stationary data and such a window would have to be considered noise in a windowed Fourier transform algorithm. Moreover, noise sources (which do not include nonstationary signal) can be of any kind and do not need to be quasi-stationary (e.g., imagine a road with irregularly passing cars near the instruments) [cp. *Adam et al.*, 1986]. All nonstationary noise sources (may also include nonstationary signal) will affect the (windowed) Fourier transform in unpredictable ways just because the data breaks the necessary assumption for the Fourier transform at least in the relevant windows. This is not an

issue when there are few affected windows, but it would become a problem when nonstationary effects are frequent. A more concise treatment of electromagnetic noise and its characteristics is given by *Szarka* [1987] and *Junge* [1996], where both acknowledge nonstationary noise sources and the aforementioned difficulties. Therefore, we argue that even though the MT signal may behave sufficiently stationary, the contained noise in the data clearly cannot always be assumed quasi-stationary as it would be required for the application of the Fourier transform.

The isolation or separation of noise has been studied intensively since the introduction of the MT method and the two major noise counteract breakthroughs date back to the 1980s. *Gamble* [1979] propose the use of a remote station to apply the technique of instrumental variables [*Reiersøl*, 1941] in order to drastically reduce bias by uncorrelated noise. Later, *Jones and Jodicke* [1984], *Egbert and Booker* [1986], and *Chave and Thomson* [1987] advocate robust regression procedures for transfer function estimation to reduce the influence of unlikely but highly influential data points. Besides these two milestones, there has been much effort in reducing noise influence further by either trying to estimate and remove the noise directly [e.g., *Egbert*, 1997; *Oettinger et al.*, 2001] in the frequency domain or by filtering or extracting quiet data sections in the time domain by visual inspection [*Garcia et al.*, 1997] and in the time-frequency domain [e.g., *Weckmann et al.*, 2005, and references therein]. The latter procedures are reported to be effective for particular data sets but require intense user attention and good, detailed knowledge about the data. Moreover, noise identification, separation, and/or removal is not always successful, sometimes practitioners encounter data from which it is seemingly impossible to extract reasonable transfer functions. This could be partly due to the fact that EM data (the combination of signal and noise) are not as quasi-stationary as required for the (windowed) Fourier transform. A very simplistic example would be the presence of a spike in the data, which would compromise the particular data segment (or window) in which it is present. Clearly, the presence of a moderate number of spikes is easy to counteract (through interpolation) [e.g., *Jones et al.*, 1989; *Junge*, 1996], but we argue that the same principle applies to other nonstationary effects which might not be as easily identified and mitigated.

Huang et al. [1998] introduce empirical mode decomposition (EMD) in the framework of the Hilbert-Huang Transform (HHT), a novel time series analysis tool, which is data adaptive and suitable for nonlinear and nonstationary data. The decomposition provides data modes (called intrinsic mode functions (IMFs)) which are defined such that they can be represented as a single oscillation. Thus, *Huang et al.* [1998] argues that the definition of the IMF allow for a meaningful computation of its instantaneous parameters, like amplitude, phase, and frequency, with the Hilbert Transform. In practice, however, *Huang et al.* [2009] demonstrate that the Hilbert Transform often is numerically unstable and advocate a more practical routine to obtain the instantaneous parameters, which first separates amplitude and oscillation and then acquires the instantaneous phase by direct quadrature.

EMD has been tried and applied in several fields, including geophysics and the magnetotelluric method [*Battista et al.*, 2007; *Zhang et al.*, 2003; *Cai et al.*, 2009; *Chen and Jegen*, 2008]. In particular, for MT, *Cai et al.* [2009] present how EMD could be used to separate obvious noise from the signal. Later, *Cai* [2012] attempts to substitute the Fourier transform in favor of HHT in MT processing, but the segmentation and averaging of data in order to construct marginal spectra (comparable to Fourier spectra) are unnecessary and limit the potential strength of EMD. In the same year, *Chen et al.* [2012] present an estimation scheme for the transfer functions in MT data by using the instantaneous parameters (in contrast to marginal spectra). However, they conclude that the implementation of remote reference processing and robust statistics can further improve their approach, because both techniques are very often required to estimate transfer functions from regular field data.

This work follows *Chen et al.* [2012] by using directly the instantaneous parameters obtained from EMD but in contrast to their work; here the multivariate variant of EMD by *Rehman and Mandic* [2009] is discussed and applied. Robust procedures are introduced to estimate instantaneous parameters, and a data selection scheme is proposed to ensure independent data. For transfer function estimation, a robust regression is advocated, which uses regressors defined by the two major robust principle components (robust principal component analysis described by *Hubert et al.* [2009]) of all remote data sets or for single site processing, all the available channels. Effectively, this procedure excludes the site channels from the regressors if remote data are available in order to further reduce the risk of propagating correlated noise from between site channels into the principal components. Synthetic examples demonstrate the effect of nonstationarity of the

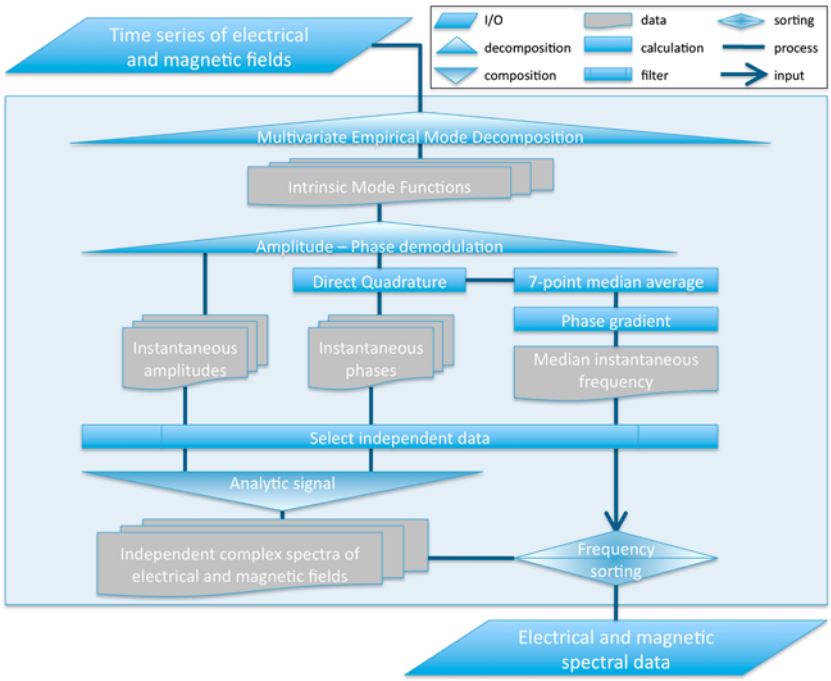


Figure 1. EMT workflow chart to compute spectra.

source on traditional processing schemes. Semisynthetic signals, which consist of real signal and synthetic (nonstationary) noise, present the effects due to nonstationary of noise in real data, and lastly, real-world data sets verify the power of the algorithm for regular data and most notably, data in which nonstationary noise is suspected. Additionally, a MATLAB routine is presented, which creates (non)stationary synthetic MT data (or noise).

2. Outline of the EMT Algorithm

Figure 1 outlines the workflow chart of the algorithm that we have developed to process MT data using the EMD technique. We call the following scheme empirical mode decomposition-based magnetotelluric data processing, in short EMT. Here we present the outline of the code, the following sections will describe each of the steps thoroughly:

1. Decompose time series with multivariate empirical mode decomposition (MEMD). The MEMD method is used to decompose the multivariate data of all available channels (station and remotes) into oscillatory modes.
2. Compute instantaneous Parameters. Separate amplitude and oscillatory phase functions of the modes with amplitude-phase demodulation according to *Huang et al.* [2009]. Generate the complex IMFs from amplitude and oscillatory phase for each channel to permit the computation of the instantaneous phase and the instantaneous frequency defined as time derivative of the phase.
3. Gather independent data points. We ensure linear independence of the data points by defining a time scale of data dependency.
4. Organize data in frequency domain. The data points are collected in wide bins, typically 5 to 10 bins per decade, ensuring enough estimates per decade and statistical stability of the impedance estimation for each bin by exploiting the fact that the MT transfer functions vary slowly with frequency.
5. Estimate transfer functions. (i) Compute the two major robust principal components from data to use as regressor, (ii) robust regression of each channel on principal components, and (iii) estimate confidence intervals by means of bootstrapping the robust regression.

3. Algorithm Step 1: MEMD

Huang et al. [1998] only present the application of their technique to univariate data, but MT data consist of at least four data channels, which depend on each other. Using a univariate EMD, each signal is sifted and

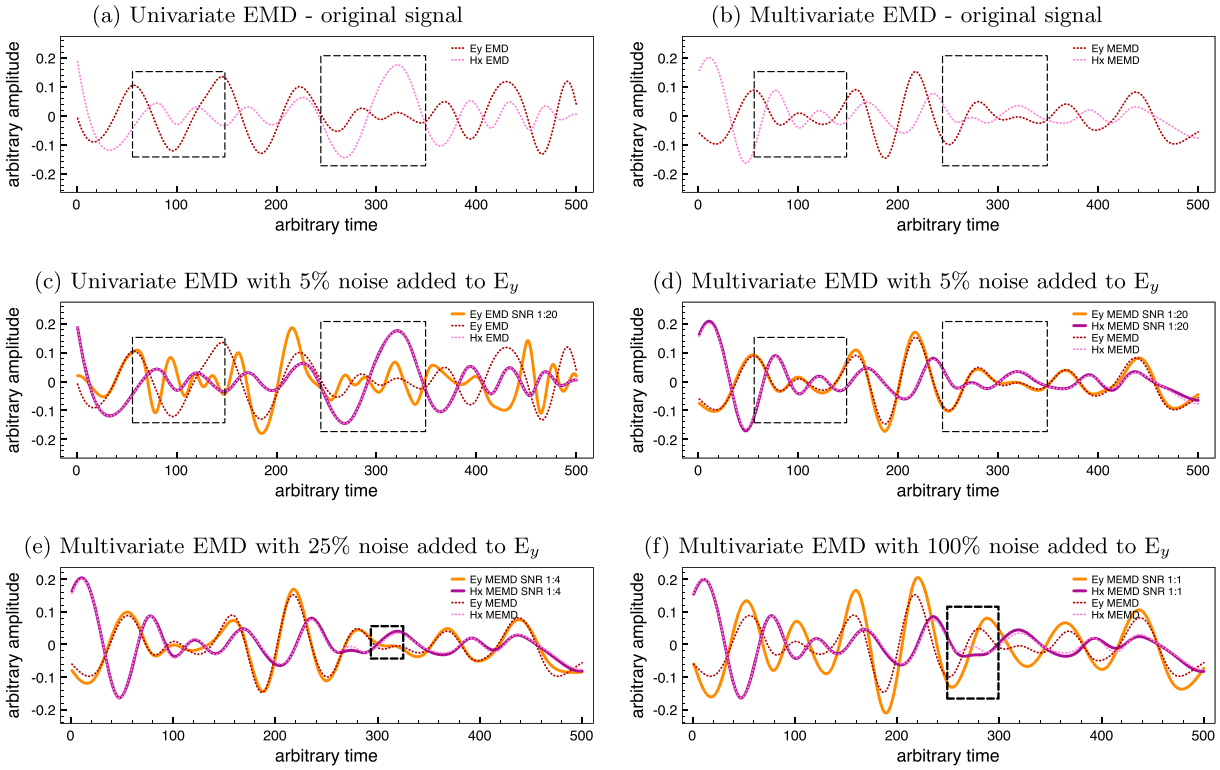


Figure 2. (a–f) The related modes for two channels of a short example signal from Southern Africa data are compared with respect to time scale consistency for different EMD algorithms and added Gaussian noise variance. The presented mode is #8 out of 11 IMFs obtained from 1000 data points; however, note that the actual mode number is irrelevant for the example being representative, since Gaussian noise is equally present in all modes. The dashed boxes emphasize where there are major differences in the time scale between the two channels for the univariate case, and then, for the multivariate algorithm, the boxes highlight the differences due to the added noise. Figure 2e shows that the time scale in the mode of E_y is affected by loosing one oscillation when adding 25% noise variance, and in Figure 2f the noise variance added to E_y even begins to affect the time scale in the modes of H_x. Percentage of noise refers to the noise variance for Gaussian noise relative to the average signal amplitude.

Hilbert transformed independently, decreasing the likelihood that the signals remain dependent throughout all modes due to the possible channel-independent noise characteristics. For example, if there would be a high-frequency noise in one channel which is absent in the other three, the first IMF of the first channel would contain that noise and start with the rest of the signal from mode two on, whereas the other channels would contain signal from mode one on, resulting in different time scales for all modes (Figure 2). In this example, without any previous knowledge of this noise, the corresponding modes of different channels could never be used jointly for a linear least squares approach, since they do not contain the signal of the same frequency range. For that reason *Chen and Jegen* [2008] and *Cai* [2012] suggest to calculate the marginal spectra for each channel and use those in a similar manner as it would be done with the Fourier spectra. This approach has been shown to work very similarly to the usual Fourier approach and to provide novel noise control mechanism but does not take full advantage of the possibilities that EMD offers, namely, the instantaneous parameters. *Chen et al.* [2012] circumvent this problem by only taking into account the data points of a time instant when they find a match for the instantaneous frequency (IF) for each channel but in any mode. This procedure certainly solves the problem given in the simple example above, but a procedure that only uses data points were the IF matches (arguably within a certain limit) might run into problems as soon as the channels are more seriously distorted by noise, and hence, the frequency computation for one or more channels is rather poor, ultimately decreasing the number of valid data points. In this section we discuss a multivariate decomposition algorithm that alleviates that problem by forcing all channels to decompose into correlated IMFs or in other words into IMFs of similar time scales, so that we can attribute one common IF value to all channels.

Rehman and Mandic [2009] developed a scheme to analyze multivariate signals and compute IMFs of each of the signal's components such that they remain correlated in their time scale as much as possible. The algorithm is summarized in the following:

1. Project the multivariate signal on an orthogonal *n* dimensional hypersphere (basis functions defined by Hammersley sequences). The dimensions of the hypersphere represents different time scales much like the orthogonal sine functions in the Fourier transform.
2. Locate the extrema of each projection (*n* projections in total).
3. Interpolate the multivariate signal by using the projection extrema locations for each dimension, to obtain a distinct upper and lower envelopes of the multivariate signal for each dimension of the hypersphere.
4. Average the means of upper and lower envelopes for each channel over all dimensions.
5. Subtract the average envelope mean from the data and repeat to convergence to obtain the multivariate IMF.

MEMD provides a set of IMFs for each channel and retains the dependency in between those with respect to a most similar time scale (frequency) in all channels. It is also worth noting that for a source in EM field theory all components of the electric and magnetic field have the same frequencies present at all times, meaning that if there is an electric source of 10 Hz, it will be accompanied by a magnetic field of 10 Hz. Therefore, MEMD does not at all introduce additional assumptions on the field components but rather ensures a fundamental property inherent in EM field theory for each IMF, and thus, it decomposes the MT data into IMFs which can be conceived as independent data sets.

MEMD decomposes the data set into a number of IMFs, which have the information of instantaneous amplitude, phase, and frequency at each time step, and each IMF is a time series with a dynamic and locally narrow banded IF [*Flandrin and Rilling*, 2004]. Each IMF is interchannel dependent, and each time step fulfills the MT equation for its IF in the same way as narrow frequency-banded time series do [*Berdichevsky and Bezruk*, 1973; *Chen et al.*, 2012; *Neukirch and Garcia*, 2013]. However, real data will always contain noise in all channels, and the effect of the noise on the IMFs will largely depend on the (timely) local signal-to-noise ratio and can easily span from subtle effects (e.g., some noise is present in one of many clean channels) to affecting the amplitude in (originally) clean channels (e.g., half the channels are corrupted by coherent noise and affect the clean ones) to even introduce false information in all channels (e.g., severe noise introduces new extrema). As an example for noise effects, Figures 2c to 2f illustrate data with added Gaussian noise to a single channel.

This effect is conceptually comparable to how noise leaks in an ordinary Fourier transform where the signal-to-noise ratio distorts the true (noise-free) spectra, but in the EMD case the effect is local and only affects the signal at some distance around the noise occurrence, whereas the Fourier spectrum is always affected in the whole segment, since it is formulated as an integral.

The Fourier transform is a univariate algorithm, and noise in different channels cannot affect each other. Further, obviously, nonstationary effects can be reduced if the time series are broken in windows (windowed FFT). However, any nonstationary noise in a data window will affect the entire Fourier spectrum of that window, and often, robust procedures will drop exactly those spectra entirely regardless whether or not there shorter good data sections in that window. For an MEMD-based algorithm, the decision of excluding spectral information can be made for each individual time step instead of entire windows, if desired. However, care has to be taken, because even though spectral estimates are delivered at each time step, the real-time frequency resolution is much lower and depends largely on the extrema in the corresponding IMF, but let us defer discussion on this matter to section 5.

The most important point, which can be observed in Figure 2, is that channels influence each other already during the MEMD. Apparently, noise spreads throughout channels and clean channels may be affected by noise, becoming biased. This noise spreading across channels occurs because the algorithm does not assume that one of the channels can be affected by noise while the others are not; it simply finds the best correlated signal for all modes and accounts the noise as a distortion of the total electromagnetic wave field. It becomes clear that this multivariate decomposition excels with the number of provided clean channels, which aid stabilizing the mode sifting and reduce noise in noisy channels by spreading it over all channels. For this reason the mode of the E_y component in Figures 2e and 2f appears to contain less noise than one would expect from adding 25% and 100% variances of Gaussian noise, respectively. Naturally, it seems undesirable to spread noise from one channel to the others (which could be entirely avoided with a univariate EMD algorithm as *Chen et al.* [2012] propose), because we should preferably extract the best undistorted signal possible from our data. But since MT is an intrinsically multivariate problem, we always need the

information of all channels (of the site of interest) for the final TF estimation, and the more data points we loose due to large deviations (in, for instance, the IF, which is a data selection criterion by *Chen et al.* [2012]) in only one heavily distorted channel, the more difficult it will be to find an accurate transfer function. Using MEMD instead of EMD and enforcing a similar time scale on all channels robustifies the decomposition procedure and yields more spectral data points which can be evaluated in the regression step at the cost of spreading the multivariate noise and thus increasing noise in some channels.

Usually, a good portion of the noise is not correlated between the channels and therefore affects the channels unequally, resulting in instantaneous parameters that depart from their correct values depending on the noise. Although this is certainly not appreciated for parameters like amplitude and phase, it does come in handy for the frequency computation, which we assume to be constant between the channels. Any deviation of the IF between channels must be due to any of the following:

1. The modes do not fulfill sufficiently the definition of IMFs (having a locally zero mean).
2. The signal (channel) has been contaminated by noise (heavier contamination will result in larger deviations).
3. The frequency has been altered by nonstationary convolution with the system response of the receiver.

The first problem is a very common issue for the first modes in EMD, since the data are always sampled on some rate and the location of the extrema in the data depends much on the sampling rate (in a real nonstationary situation, the extrema can be anywhere in between the measurement directly before and directly after the recorded extrema). Routinely applied low-pass filters may alleviate much of this problem, but the exact location of the true extrema is the most crucial information for calculating the instantaneous parameter from IMFs, and this is usually not well defined for frequencies close to the sampling rate. However, in our experience the uncertainty on the location of the extrema only disperses the instantaneous parameters and does not usually introduce bias; the larger scatter in the regression is not problematic due to the larger number of data points for the higher frequencies in a data set. The second point is almost always an issue in MT, and it is broadband, meaning it is found in all frequency ranges and thus all IMFs. But since we know that the frequencies between the channels should be equal, we could use deviations between them as a selective quality marker or down weights in the later regression (similar to *Chen et al.* [2012]); however, we have not tested this idea in the present work. The last point is a rather new conclusion derived from the nonstationary convolution theorem in *Neukirch and Garcia* [2013] and will be discussed thoroughly in another work. The problem only occurs for nonstationary data convolved with a system response that varies over frequency, just like the instrument system responses for MT equipment usually do. It is not present during stationary sections and therefore a minor issue for most MT data but fairly complicated to analyze; therefore, it is out of the scope of this article. In any case, these disturbances are listed for sake of completeness as they will also affect amplitude and phase and thus can introduce undesired bias to the transfer function estimation if not removed from the data or being accounted for.

Before we continue with the subject of IF, we need to focus on the recovery of the amplitude and phase from the IMFs in the following section.

4. Algorithm Step 2: Computing Instantaneous Parameters

Huang et al. [2009] thoroughly discuss the computation of instantaneous parameters from an IMF, and *Chen et al.* [2012] continue the discussion with respect to an application in MT. We mostly follow their suggested instructions, since the IMFs of MEMD are methodically no different from the ones obtained from univariate EMD. Essentially, *Huang et al.* [2009] advise to separate amplitude and oscillatory phase with a procedure called amplitude-phase demodulation from the IMF. Then the instantaneous phase can be computed by direct quadrature from the separated, oscillatory phase function. In contrast to the original idea [*Huang et al.*, 1998] of using the Hilbert transform to obtain the phase, the direct quadrature method does not guarantee a strict analytic signal, but the routine performs well in practice and estimates the correct phase of the underlying signal more robust than the Hilbert transform.

Focusing on the differences between this work and previous studies [*Huang et al.*, 2009; *Chen et al.*, 2012], examples of instantaneous parameters are given in Figure 3, which feature two modes of a short section of a real data set from Southern Africa. Figures 3d and 3e display the instantaneous amplitude (IA), Figures 3f and 3g the instantaneous phase (IP), and Figures 3h and 3i the instantaneous frequency (IF).

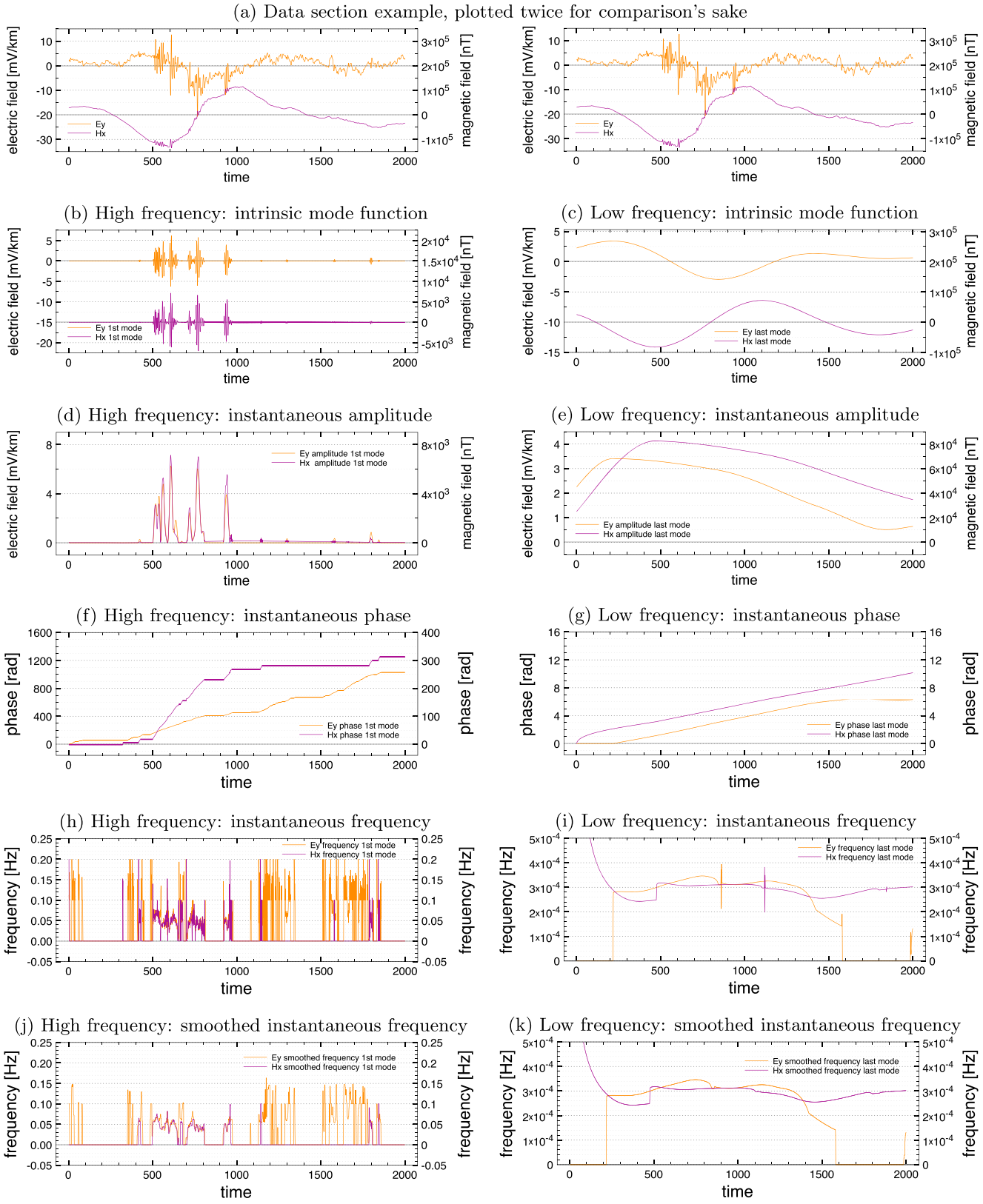


Figure 3. (a) The sum of the two modes in the spirit of comparability. Examples of instantaneous parameters are displayed. (b, d, f, h, and j) High-frequency mode (IMF #2). (c, e, g, i, and k) Low-frequency mode (IMF #5). Both modes are extracted from the same data (site 072).

By definition, the direct quadrature method divides by very small numbers at the extrema of the phase modulation function leading to numerical instability at those points, which additionally amplifies uncertainties and noise. Since we apply the direct quadrature method [*Huang et al.*, 2009; *Chen et al.*, 2012], the IP usually contains small numerical errors. Especially, these numerical instabilities have a great impact on the time derivation of the phase function and are depicted by simple poles in the IF (see Figures 3h and 3i).

The poles are of first order and almost cancel each other out when summed over, which is why the phase function itself still looks smooth and the mean average over a sufficiently long time range is hardly affected. We found that a seven-point median average filter applied on the phase function before differentiating is a sufficient counteract and does not restrict the signal much more than the cubic spline interpolation already did during the sifting procedure but produces a much more stable IF (cp. Figures 3j and 3k).

In addition to the numerical instability associated with the direct quadrature method, the particular noise in each channel may cause differences in the IF between channels, where we would expect an electromagnetic field to have the very same frequency in all of its components (electric field and magnetic field) at a given time. However, we can use this fact to find a likely estimate for the common IF for all channel by using its mean or median average. The IF average is a physical meaningful representation of the true frequency of the electromagnetic signal (which is represented by all channels together) for a given time and mode. Heavy outliers from that mean average can be counteracted by using the median average and may be used to identify problematic data sections and can contribute to data quality control as mentioned in the section above. We found the median average to provide us with better estimates of the IF because of the frequent instabilities produced by the direct quadrature. These large irregularities in the IF usually do not occur in all channels at the same time, because of the impedance-related phase shift between channels (refer to *Huang et al.* [2009] for a discussion on the nature of these numerical instabilities), but occur very frequently, and thus, the median average compensates this problem, whereas the mean average would be drawn toward the outlier regularly.

All three instantaneous parameters: IA, IP, IF, and time form data quadruples and fully describe the original data. The IA and IP can be combined to form the representation of the complex spectra for a given time and frequency. *Neukirch and Garcia* [2013] lay out the fundamentals for signal system convolution in a HHT context and provide proof that the convolution of complex, nonstationary IMFs with a system response in time domain can be reformulated as the multiplication of the complex, nonstationary IMFs with the system response function in the frequency domain. Therefore, when processing MT time series we can use the complex IMFs in the very same way as a Fourier spectrum and carry out a statistical analysis in order to find the spectral physical relation between the channels, known as transfer functions.

For the sake of meaningful statistics with linear regression, we should try to ensure (1) that the data errors are independent (estimation accuracy) and (2) that the errors are identically distributed (accuracy and precision of estimation). Starting with the second, since we explicitly allow for nonstationarity in our scheme, it is clear that our spectral data cannot be assumed to be drawn from a single distribution. The parameters of any distribution from which the data might start with will likely change during time; this is exactly what nonstationarity means. However, the data decomposed by EMD are represented by oscillating modes which are bound to their definition and therefore always are locally zero mean functions. Thus, the definition of the IMFs ensures that the center (location) of the data distribution is zero for whatever time-varying distribution it follows. *Liu* [1988] discussed the importance of data being identically and independently distributed (IID) in statistical system analysis with nonparametric methods and came to the conclusion that when the bootstrap algorithm is used, the requirement of the data being IID can be somewhat relaxed, such that it is sufficient to ensure data point independency and that the underlying distributions of the data have a common location. They argue that the nonparametric nature of the bootstrap algorithm includes a robustness toward dissimilar distributions in the data as long as the locations of the distributions are very similar (in our case even equal).

The requirement that the data points are independent is more involved and has not yet been discussed in literature for an EMD setting; therefore, we dedicate the following section to that issue, then we will return to the discussion of the statistical analysis.

5. Algorithm Step 3: Independent Data Points

Data independency is an important criterion for our statistical analysis, which if left unconsidered may bias accuracy and/or precision of the methods we use in this work. Besides, the understanding of the dependency of data points allows to draw inference about the time-frequency resolution.

In our case we need to understand how data points interact and depend on each other in the total framework of HHT. Both IA and IP derive from an analysis of the inner structure of the corresponding IMF. Each IMF is constructed by a loose sifting procedure based on the signal's extrema and guided by the required properties based on the IMF definition, a highly data adaptive procedure. The subsequent amplitude-phase demodulation and the computation of the complex IMF do not rely any more on any data characteristics. For the demodulation the amplitude function and oscillatory phase function are already defined through the IMF definition and it only strips the two apparently different signals apart. Then, the direct quadrature uses the oscillatory phase function to recover locally the argument of the assumedly complex oscillation. The demodulation procedure is comparable to calculating the argument and absolute value of a complex number, which does not change or add any data dependencies but only changes the way data are described (via the complex IMF which does not introduce information to the data). Therefore, we focus on the mode decomposition itself, when looking for dependencies in the data.

First of all, keep in mind that per definition, all IMFs of a signal are theoretically locally orthogonal, which implies that one mode to the next is linearly independent and uncorrelated. However, independency is by no means guaranteed along a mode. Since the IMFs are solely defined by a subset of points of the entire signal, namely, the extrema, the IMF itself cannot have more degrees of freedom than number of extrema. All data points of an IMF between two extrema usually share a third-order interpolation polynomial, a cubic spline, which defines these data points based on the same set of extrema. Therefore, all these points between the same two extrema are dependent, whereas points that base on different sets of extrema are independent (even if just one extremum is different). Hence, it is important to only take into account one single data value for each span between two extrema to impose independency between the final data points. Naturally, the lack of independency in the definition of an IMF compromises greatly the time domain resolution suggested by IMFs but indicates that HHT does not provide a higher spectral resolution than what would be expected by the observed frequency (thus, we still need a complete oscillation to meaningful describe spectral data). Furthermore, since the cubic spline requires the closest four extrema at each data point, the distance of influence of every extrema is about two full oscillations and represents some measure of time-frequency resolution.

Since only one interextrema data point is independent, we have to pick the one which represents the entire range. Each data point should be equally valid since they are dependent. However, noise characteristics can make some points be a poorer choice than others (be reminded on the numeric instabilities due to the direct quadrature discussed in the section above). For the moment we have not designed a selection criterion based on data quality, so we simply take one point per half oscillation defined by the location of the extrema of the function:

$$P = \sin\phi \cos\phi. \tag{1}$$

Since MT processing is multivariate, we suggest to use (1) with the integral of the common IF ω_c between the channels; thus,

$$\phi(t) = \int_{-\infty}^{t} \omega_c(t')\mathrm{d}t'. \tag{2}$$

This integral is basically the inverse of the time derivative of the phase used to obtain the IF in the first place, only that now the integrand is the common IF, which results in some sort of common phase for the EM data in (2), and provides an oscillatory function in (1) according to the intrinsic oscillation of the EM data. The choice of this particular function is mainly because of its fairly random selection, if we would choose data points with certain properties (e.g., low/high amplitudes), we could easily introduce bias to the transfer functions, which is not the case for this general function. However, a more careful or sophisticated selection criterion (like a weighted average) for this point could help to reduce numeric or perhaps, even electromagnetic noise and could be discussed elsewhere.

6. Algorithm Step 4: Frequency Sorting

As noted above, EMD results in a distinct frequency value for each channel. The average of those values for a given time and mode over all channels is a physical meaningful but biased representation of the true frequency of the electromagnetic signal (which is represented by all channels together). The bias should be lower for data points which have a similar frequency value and may even be considered for data quality control as we stated before. Keeping in mind that we use a common frequency function for all channels defined by the median average between them, in the following we will assume the median frequency as the common frequency between the data channels.

Remember that the instantaneous frequency (IF) is the time derivative of the phase of the complex IMF and does not yield equidistant (as, for example, the Fourier transform) but rather continuous frequency values which vary with time and thus along a mode. For this reason, it is unlikely that we can find two data points (each with two electrical (e) and two magnetic (h) components) with the very exact frequency value (ω_0), but this would be necessary in order to find a unique estimate for the transfer function tensor (Z), which is only defined at a constant frequency:

$$\begin{pmatrix} e_x(\omega_0, t) \\ e_y(\omega_0, t) \end{pmatrix} = \begin{pmatrix} Z_{xx}(\omega_0) & Z_{xy}(\omega_0) \\ Z_{yx}(\omega_0) & Z_{yy}(\omega_0) \end{pmatrix} \cdot \begin{pmatrix} h_x(\omega_0, t) \\ h_y(\omega_0, t) \end{pmatrix}. \tag{3}$$

Note that this equation deviates from the traditional form as it includes time variance for the electromagnetic fields, since the complex IMFs of the data channels are still time series. A similar form of this time variant formula has been introduced by *Berdichevsky and Bezruk* [1973] and recycled by *Chen et al.* [2012], until this form has been proofed for the EMD context by *Neukirch and Garcia* [2013]. However, even though (3) suggests that the MT impedance equation is valid at each time instant for the IMFs of the electromagnetic field, the impedance itself cannot be solved for unless there are at least two independent measurements for the same frequency value. But since the electrical impedance only changes smoothly with frequency [*Cagniard*, 1953], we can group similar frequency values to increase the amount of measurements available around a certain center frequency. For this procedure, we select the independent data points based on (1) and arrange them according to the common IF, omitting time dependency of the data by considering the time axis rather as index for measurements than physical time. The data reorganization in these frequency bins follows the proposed method by *Chen et al.* [2012], only that we do not allow IMF mixture for the reasons discussed in section 3.

Following this reorganization, we form an overdetermined system of equations that we can solve for the transfer function tensor at distinct frequency values. The estimation procedure is a bootstrapped, robust principal component regression and will be discussed in detail in the following section.

7. Algorithm Step 5: Robust Principal Component Regression

Egbert [1997] shows that MT sources are well described by two electromagnetic field polarizations. Practically, this means that the entire data vector space of all channels in a data set can be represented by the combination of two polarization vectors. Theoretically, the high-dimensional data (electric, magnetic, and all remote channels) can be described by a fundamental two-dimensional polarization space that contains all the variance of the data. Such a reduction of dimensionality of data vectors can be achieved by a (robust) principal component analysis (PCA), which provides the inherent components of the data vector, ordered by its eigenvalues. The two most dominant principal components (PCs) are the magnetotelluric source vectors since they should be present in all channels and contribute most to the variance of the data [cf. *Egbert*, 1997]. However, in practice, MT data are often contaminated by noise and source field effects, which limit this procedure [*Egbert*, 1997, 2002; *Smirnov and Egbert*, 2012] such that there are more than two dominant eigenvalues which contain a mixture of source polarization vectors and correlated noise. In order to separate the dominant principal components in such cases, a much more sophisticated multisite analysis is required and described by *Smirnov and Egbert* [2012], which should be followed for data sets with coherent noise contamination; however, the discussion or incorporation of such an analysis is beyond the scope of this work, although it could be implemented in our algorithm if desired. For this work, we assume that the first two principal components are a sufficiently good estimate of the MT source polarization vectors, but we restrict the data used for the PCA to remote channels only, if at least two are available. If not, the site channels can be used as usual.

A robust principal component analysis tool is provided by *Hubert* [2003] within the frame of the free LIBrary for Robust Analysis (LIBRA) package [*Verboven*, 2005] for MATLAB and referred to as `robpca.m`. *Smirnov and Egbert* [2012] compare this code for consideration of its usage in the aforementioned multisite analysis of MT data and acknowledge its power but prefer a self-made solution for its flexibility. Since we do not attempt a multisite data analysis and assume two principal components to be sufficient, the algorithm from LIBRA appears the most reasonable solution at this stage of our algorithm.

After the computation of the two dominant PCs (say $r = (r_1, r_2)$), we formulate four (or five if the vertical magnetic field is provided) two-variate regression problems in order to separately deal with the noise distributions in each data channel. Assume the north-south electric field e_1, the east-west electric field e_2, the north-south magnetic field h_1, the east-west magnetic field h_2, and, if available, the vertical magnetic field h_3 as data channels. For each data channel x, the regression writes in a matrix notation:

$$x = r \cdot R_x + \sigma_x, \quad Z = \mathrm{inv}(R_{h_1}, R_{h_2}) \cdot (R_{e_1}, R_{e_2}) \quad \text{and} \quad T = \mathrm{inv}(R_{h_1}, R_{h_2}) \cdot (R_{h_3}). \tag{4}$$

R_x is a row vector and denotes the regression parameter for channel x on the PCs r; σ_x represents the noise in x; Z is the electric impedance according to (3); and T is the tipper function, which is the magnetic transfer function between the horizontal and vertical magnetic fields. The inv( ) operator produces the inverse matrix, and the dot operator denotes the matrix multiplication. The formulation of the regression is slightly different from the one that is usually applied in MT but not as much as it seems at first. Actually, for an ordinary least squares solution for, say Z, this formulation yields exactly (3), which is the original formulation if time only indicates measurements. The idea behind this alternative formulation is that the regressors r result from a robust statistical procedure, which describe a part of the variance in the data and thus do not contain outliers that deviate from the dominant inherent information provided by the data. Originally, the regression is carried out on data channels directly, which first, contain highly influential outliers as discussed by *Chave and Thomson* [2004] and *Chave* [2012] and second, may contain correlated noise. In our solution, influential outliers in the regressor are unlikely unless they represent a repeated feature in most channels, which would only be the case for correlated noise, but if correlated noise would be present, only a careful and sophisticated data analysis (e.g., a multisite analysis [*Smirnov and Egbert*, 2012] or noise identification [*Weckmann et al.*, 2005]) can mitigate the influence of this kind of noise. In any way, such noise would be removed, if possible, before any regression attempt and thus again validates the assumption that such noise is not present in the regressors.

We divide the total regression problem in substeps to separate the expected noise from all channels (compare (4)) in order to avoid a direct effect of coherent noise between channels. The regressions themselves are carried out robustly with an iteratively reweighted Huber weight function by calling the MATLAB intrinsic function `robustfit.m`, only specifying the desired weight function. Other weight functions are possible (refer to the MATLAB documentation for a discussion on the options), and we experimented with each one, concluding that the results obtained with the Huber weight function were most accurate and precise. The robust regression only accounts for outliers in the data channels and not for any possible outlier in the PCs, which have been computed robustly in the PCA and have disregarded bad influence points already.

EMT bootstraps the entire robust regression step in order to compute a data-dependent distribution of impedance values and estimate the data intrinsic errors of the procedure. Furthermore, as discussed before, the bootstrap operation also relaxes the requirement for statistical regressions for which data should be identically distributed and therefore reflect more reliably the estimates in case of nonstationary data. Empirically, we found 1000 iterations a sufficient trade-off between accuracy and computation time to estimate the uncertainty of our results.

8. Example Data Sets

In this section, we compare the processing scheme outlined above with the state-of-the-art processing algorithms Bounded Influence Remote Reference Processing (BIRRP) by *Chave and Thomson* [2004], *EMTF* by *Egbert* [1997], and the Long period Intelligent Magnetotelluric System (LIMS) data acquisition processing algorithm by *Jones and Jodicke* [1984]. The four algorithms are applied to a number of synthetic, half synthetic/half real and real data sets. We start with two synthetic data sets, one based on white noise as source signal and the other on a purely nonstationary waveform. These two examples will shed light on the differences between a quasi-stationary and nonstationary processing scheme.

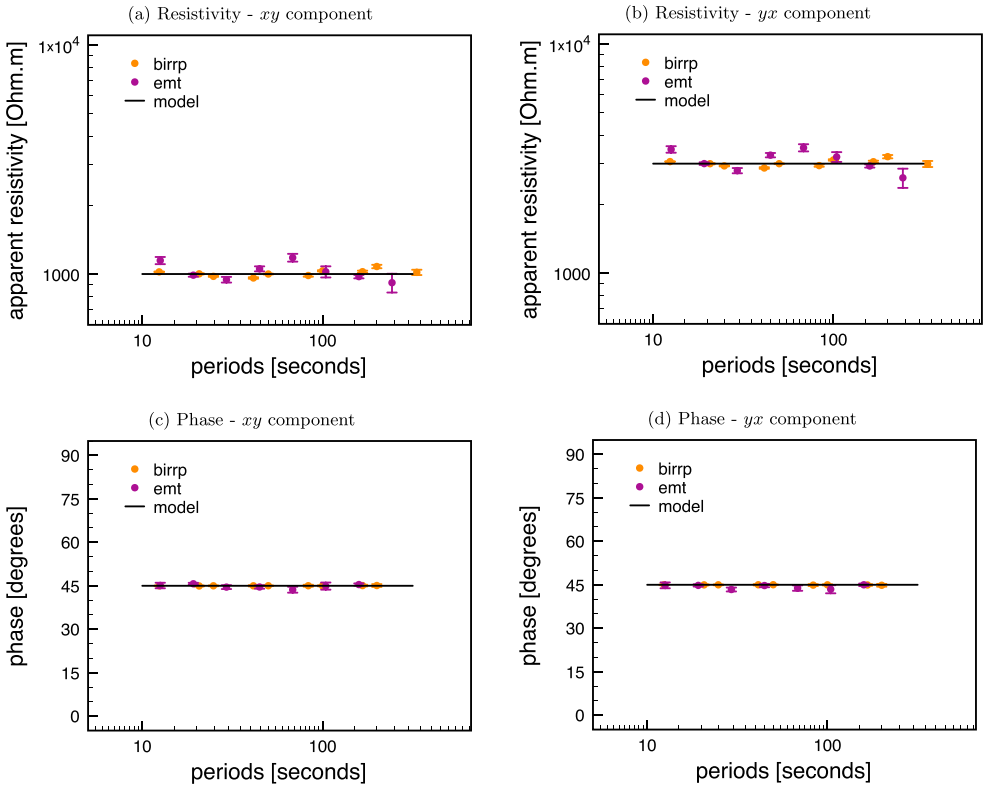


Figure 4. The EMT algorithm is compared with BIRRP for synthetic stationary data based on spectral white noise as source signal. The assumed homogenous impedance model is defined as $Z_{xy} = 10^3 \cdot \exp(i\frac{\pi}{4})$ and $Z_{yx} = 3 \cdot 10^3 \cdot \exp(-i\frac{\pi}{4})$ and is plotted as a black line.

Then we present two examples of real-world data (Southern Africa) [*Evans et al.*, 2011] to compare performance of the processing algorithms on a natural problem. In order to illustrate the effect of non-stationary disturbances in the data, we add the electric fields from the second synthetic test to the electric fields of fairly good real data, which effectively introduces nonstationary noise in the electric fields but leaves the magnetic fields completely unaffected. Lastly, we process one real data set in which nonstationary noise sources are known to interfere and demonstrate the supremacy of EMT in such a situation. All plots contain the data and estimated confidence intervals for 95% of the data (doubled standard deviation).

8.1. Synthetic Data Based on White Noise

Using an auxiliary program to create MT synthetic data (see Appendix A), in this first example, we prepared two complex remote spectra $s = (s_x, s_y)$ from independent white noise:

$$s_x = n_x^{w,\text{real}} + i \cdot n_x^{w,\text{imag}} \qquad \text{and} \qquad s_y = n_y^{w,\text{real}} + i \cdot n_y^{w,\text{imag}}.$$

The number of frequencies is $N_f = 12,500$ with a step size of $\mathrm{d}f = 0.25$ Hz to obtain a time series of 25,000 samples with a sampling rate of $\mathrm{d}t = 4$ s. The data $E = (E_x, E_y)$ and $H = (H_x, H_y)$ are computed in the frequency domain from $s = (s_x, s_y)$ by

$$E = s \cdot Z^{\frac{1}{2}} \qquad \text{and} \qquad H = s \cdot \text{inv}(Z^{\frac{1}{2}}) \tag{5}$$

with $S = M^{\frac{1}{2}}$ as the principal square root S of matrix M such that $S \cdot S = M$ in order to fulfill $E = H \cdot Z$ with the model

$$Z = \begin{pmatrix} 0 & 3000 \\ 1000 & 0 \end{pmatrix} * \exp\left(i \begin{bmatrix} 0 & -\frac{\pi}{4} \\ \frac{\pi}{4} & -\pi \end{bmatrix} \right).$$

Note that here the asterisk operator denotes the element wise multiplication of the matrices, and exp() refers to the exponential of the matrix, element by element. The results of processing this synthetic data

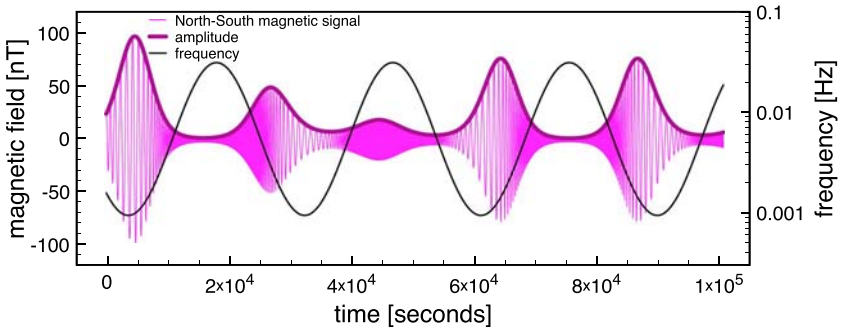


Figure 5. The chirp signal shown here is used as nonstationary synthetic signal. The north-south magnetic field component is illustrated with its nonstationary amplitude and frequency. Note that the frequency of the computed signal ranges from 1 to 30 mHz, and therefore, if data are processed outside of that range, it can only contain numerical noise.

are displayed in Figure 4 for the processing algorithms BIRRP and EMT. Both algorithms resemble the model fairly well, but BIRRP has the edge. We explain this by the fact that this synthetic source does not have any waveform, and therefore, the (M)EMD algorithm struggles to find correlated modes which it can relate to each other. On the other hand, BIRRP uses the spectral characteristics of the time series which are, per source definition, very well defined.

8.2. Synthetic Data Based on a Chirp

In order to clearly demonstrate the difference of the processing schemes, the synthetic data discussed here are completely nonstationary. Again, using SynDat (Appendix A), we define each of two orthogonal

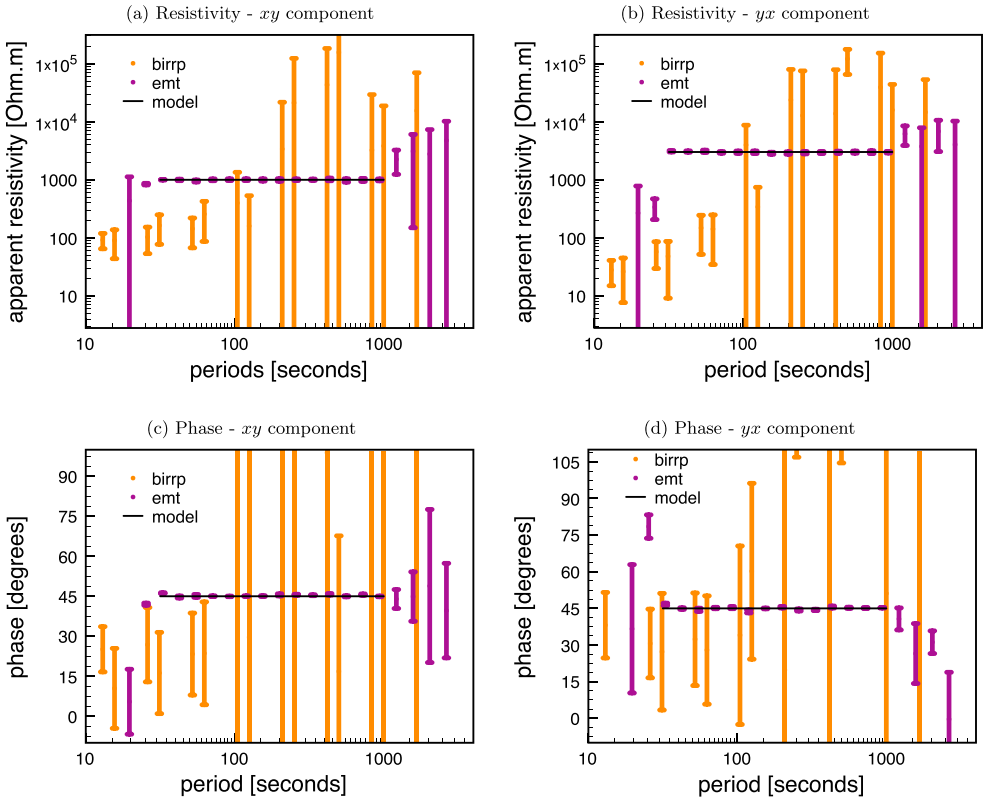


Figure 6. The EMT algorithm is compared with BIRRP for synthetic nonstationary data based on a chirp signal as shown in Figure 5. The assumed homogeneous impedance model is defined as $Z_{xy} = 10^3 \cdot \exp(i\frac{\pi}{4})$ and $Z_{yx} = 3 \cdot 10^3 \cdot \exp(-i\frac{\pi}{4})$ and is plotted as a black line. Note that the frequency of the computed signal ranges from 1 to 30 mHz, and therefore, the processed data outside that range can only contain noise; however, inside the range, only EMT is successful in recovering the model.

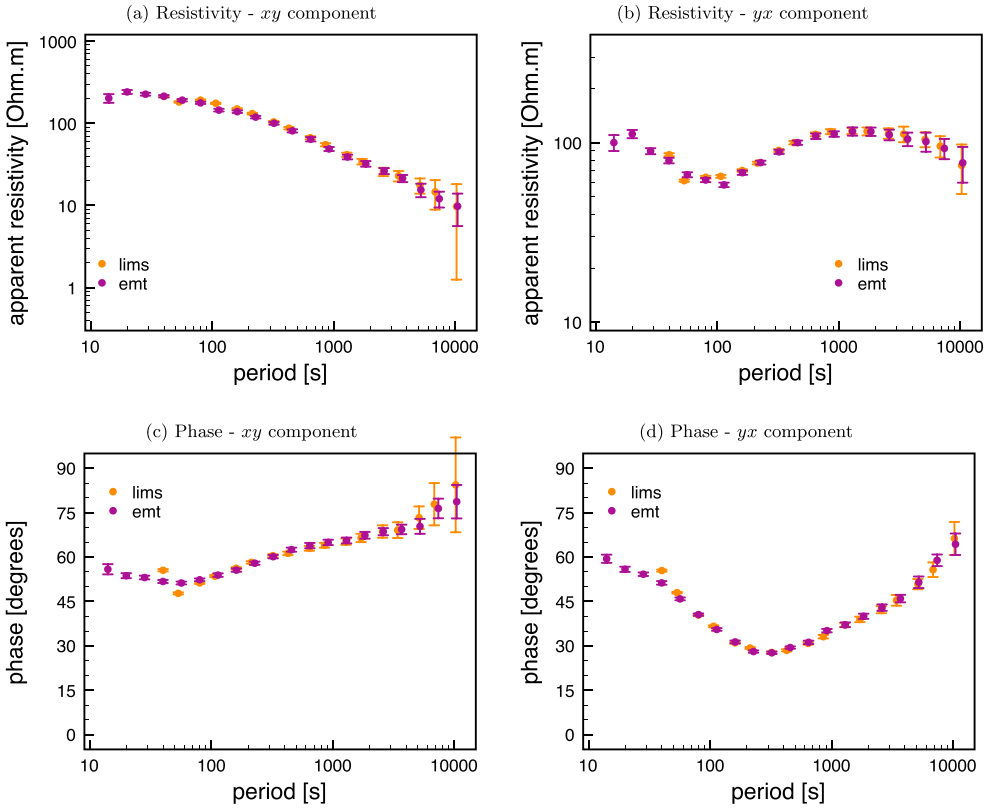


Figure 7. Good example data from Southern Africa (site 072). The LIMS results are the original results from the Southern African Magnetotelluric Experiment (SAMTEX) [*Evans et al.*, 2011].

magnetic source fields $s = (s_x, s_y)$ by a logarithmic frequency oscillation f and a logarithmic amplitude oscillation a:

$$\begin{aligned} s &= \Re\mathrm{e}\left(a \cdot \exp\left(i \int f \mathrm{d}t\right)\right) \\ \log(f) &= A + B\cos(F_f t) \\ \log(a) &= C + D\sin(F_a t)). \end{aligned} \tag{6}$$

The parameters A, B, $C = (C_x, C_y)$ and $D = (D_x, D_y)$ define frequency, and amplitude range and the parameters F_f and F_a control the degree of nonstationarity by the oscillation rate of f and a, respectively. The time axis t is sampled at a rate of 4 s for a total length of 100, 000 s or 25, 000 samples. Figure 5 displays the magnetic north-south component with its respective amplitude and frequency function. By design the signal is a locally zero mean function to ensure that it complies with the definition of the IMFs, even without the need to apply (M)EMD. As in the last example, the impedance Z is assumed to be homogenous with

$$Z = \begin{pmatrix} 10 & 3000 \\ 1000 & 30 \end{pmatrix} * \exp\left(i \begin{bmatrix} \frac{\pi}{4} & -\frac{\pi}{4} \\ \frac{\pi}{4} & -\frac{\pi}{4} \end{bmatrix}\right).$$

The electric and magnetic fields are computed according to (5). Figure 6 compares the results of processing the electric and magnetic data with BIRRP and *EMT*. Both algorithms are called with their respective default parameters to compare the results assuming no a priori knowledge about the data. EMT successfully recovers the model in the frequency range of the computed data, but BIRRP fails processing the data, which can only be addressed to the strict nonstationarity of the signal and exemplifies that Fourier transform based methods are not suitable for strictly nonstationary signals, even those that apply the windowed Fourier transform. However, this example is not a fair comparison as this kind of signal is not natural and treatises of the physics of typical MT sources [see *Rakov and Uman*, 2007; *Liu and Fujimoto*, 2011] suggest that the

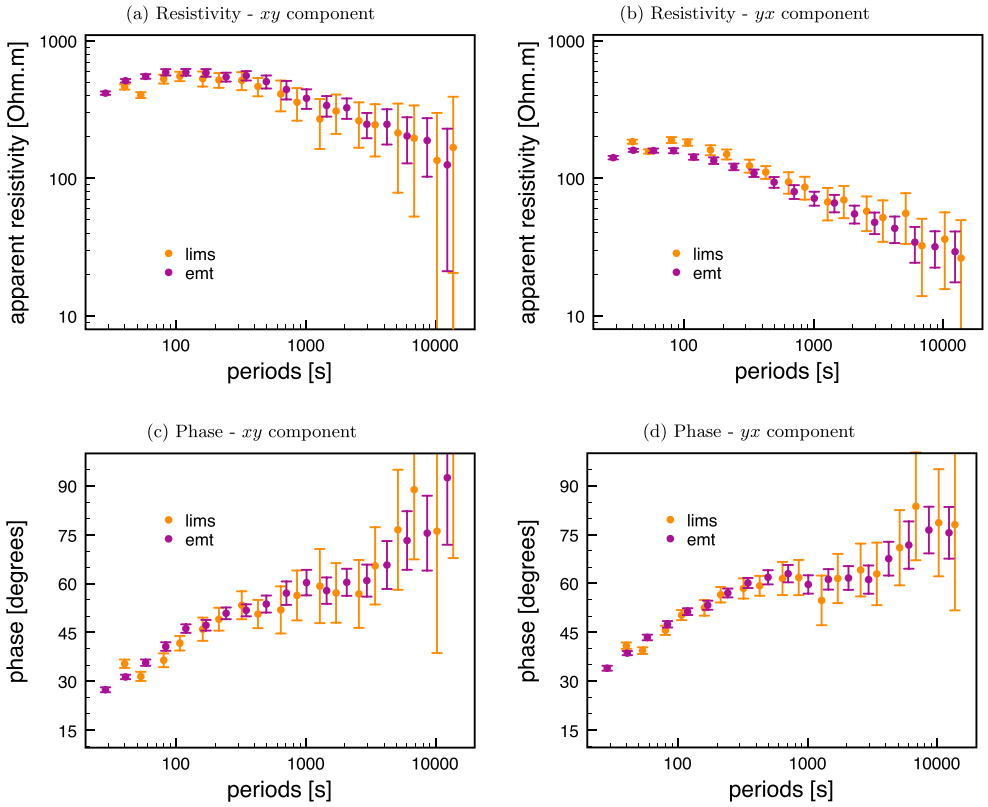


Figure 8. Fair example data from Southern Africa (site 027). The LIMS results are the original results from the SAMTEX [*Evans et al.*, 2011].

sources are not as nonstationary as this example for an extended period of time and instead can be treated as quasi-stationary. This example serves only illustrative purposes and is designed to highlight the strength of EMD, to expose the weakness of the Fourier methods and most of all, to demonstrate clearly how non-stationarity appears in the results of Fourier methods. In the following section, we present more practical examples using real data.

8.3. Fairly Good Real Data From Southern Africa

Now let us compare the algorithms using three real data sets from Southern Africa which correspond to the sites 027 and 072 with site 045 as remote reference, 042 with 027 as remote reference for long-period data, and 043 as remote reference for the short-period data [*Evans et al.*, 2011]. The first two time series have approximately 500, 000 samples on a sampling rate of 5 s, and we only consider the horizontal magnetic fields as remote reference, since they have proven to be sufficiently efficient in removing coherent noise from the local fields. The last example has up to 2 million samples for the high frequencies at 2560 Hz and around 500, 000 samples on a sampling rate of 5 s. The high-frequency data only have one remote reference site, and for the long periods, we selected the best suitable one.

The first example (site 072) is considered good for MT processing purposes when processed with the available remote magnetic channels (of site 045). Figure 7 displays the processing results for the *LIMS processing algorithm* (original results) and EMT, and both algorithms agree very well.

The second example (site 027) contains somewhat more noise even when processed with the available remote magnetic channels (site 045). Figure 8 compares the LIMS processing algorithm (original results) with EMT and shows that there are only marginal differences. Both algorithms agree well with the phase, but there is a slight difference in the amplitudes. Concluding this example, EMT appears to obtain similar results but the smaller confidence intervals are less conservative or suggest higher precision.

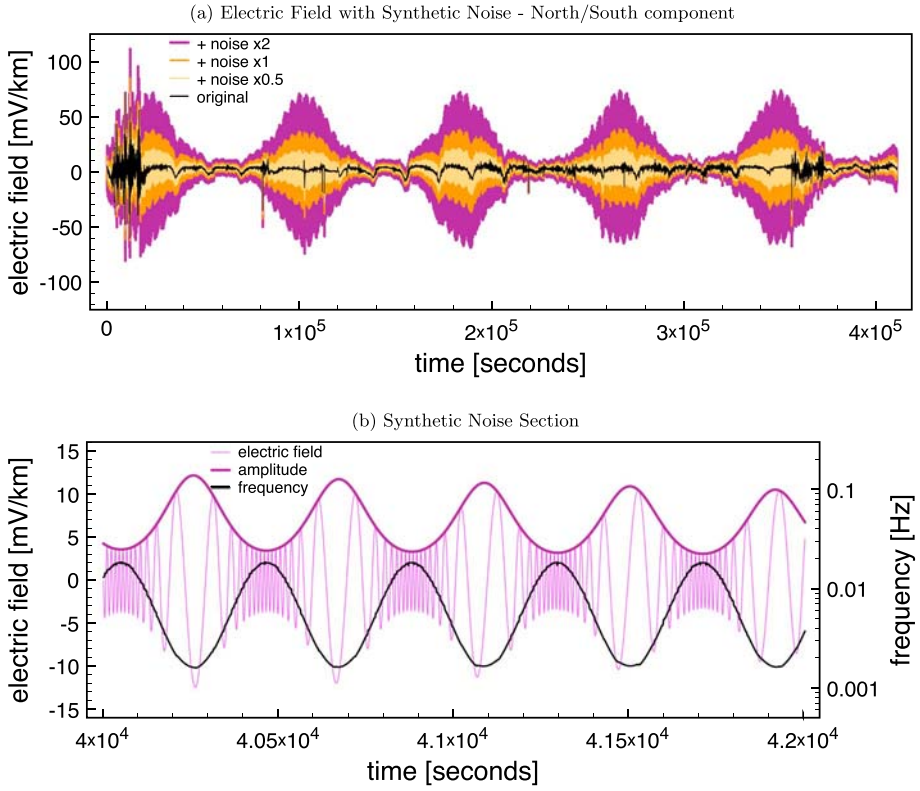


Figure 9. Southern Africa data (site 072) jammed with synthetic nonstationary electric noise. (a) The electric field channel with and without the added noise to illustrate the impact of the noise compared to the data. (b) Section of the added noise to emphasize that both amplitude and frequency contents of this noise is clearly nonstationary.

8.4. Real Data Jammed With Synthetic Nonstationary Noise

As a semisynthetic test, we combine the good real data set (site 072) from the previous section with synthetic, nonstationary noise. The goal of this test is to learn how easily a quasi-stationary source can be compromised by nonstationary noise and to test if our algorithm is able to treat the situation correctly. We consider the nonstationary noise to be present in the electric fields only and leave the magnetic and remote channels completely unaffected. This way we can see if the computation of the spectra via Fourier transform fails or succeeds, since stationary noise in the electric channels could be cleaned by the unaffected magnetic and/or remote reference channels by the remote-referencing technique.

As data, we use the data set shown in Figure 7 and add independent, purely nonstationary noise as defined in (6) to each electric field channel:

$$(e_{x,\text{ with noise}}, e_{y,\text{ with noise}}) = (e_x, e_y) + (s_1, s_2).$$

Then, we try to recover the original impedance (Figure 7) by processing the altered data with BIRRP and EMT to study the effects of the added, nonstationary noise. The test is performed thrice, first with a certain noise level, then again with the noise doubled and quadrupled. Figure 9a displays the electrical field north-south component with and without the added noise for all three tests and as an example; a section of the added noise is illustrated in Figure 9b with its parameters amplitude and frequency. The spectral range of the noise is set between 1.7 mHz and 19 mHz, respectively 52 *s* and 610 *s*, so we expect to see the biggest impact on the data processing results in that range.

Figures 10 to 12 display the estimated impedances with the increasing impact of the nonstationary noise. Where in Figure 10 the noise only raises the confidence limits for BIRRP, it camouflages the estimates in their entirety for larger noise amplitudes in Figures 11 and 12 so much that the impedance cannot be retrieved. On the other hand, EMT is barely affected by the lowest and medium amplitude noises and still provides interpretable results with the highest noise amplitude.

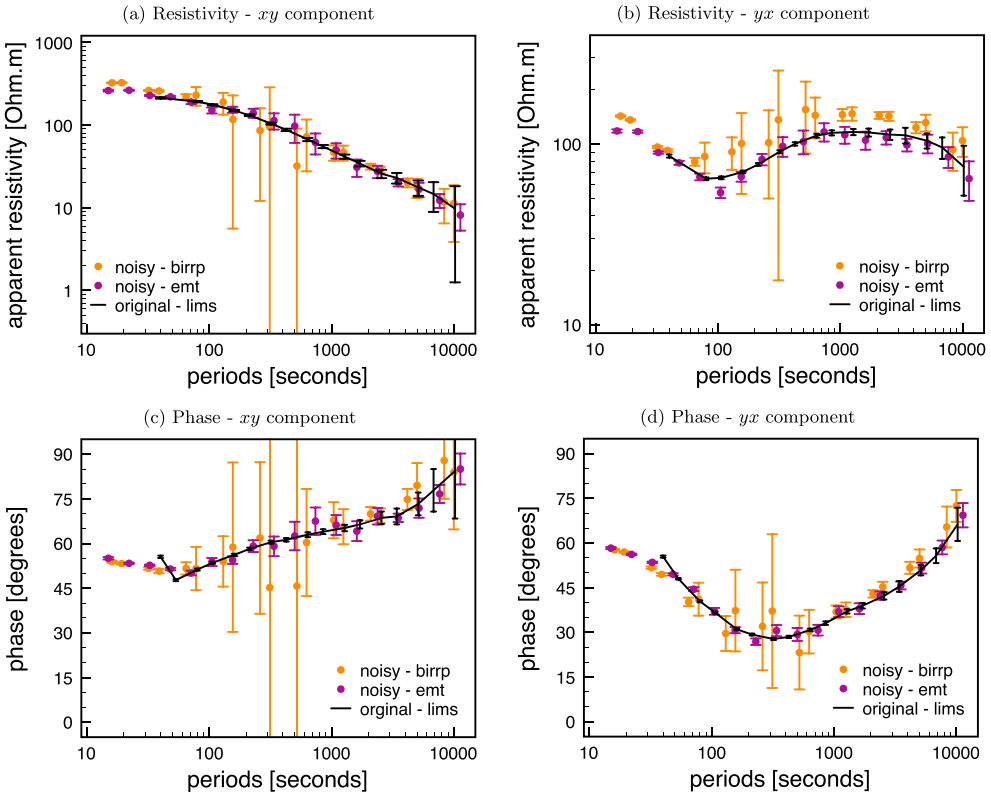


Figure 10. Southern Africa data (site 072) jammed with synthetic, nonstationary electric noise of low amplitude.

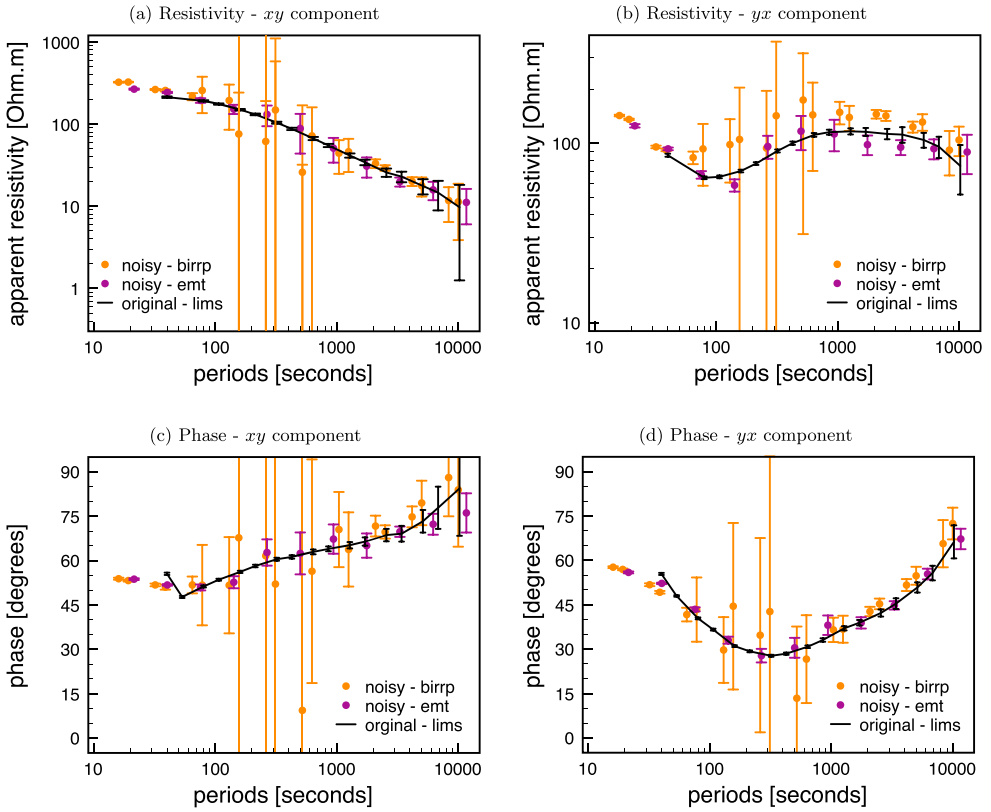


Figure 11. Southern Africa data (site 072) jammed with synthetic, nonstationary electric noise of medium amplitude.

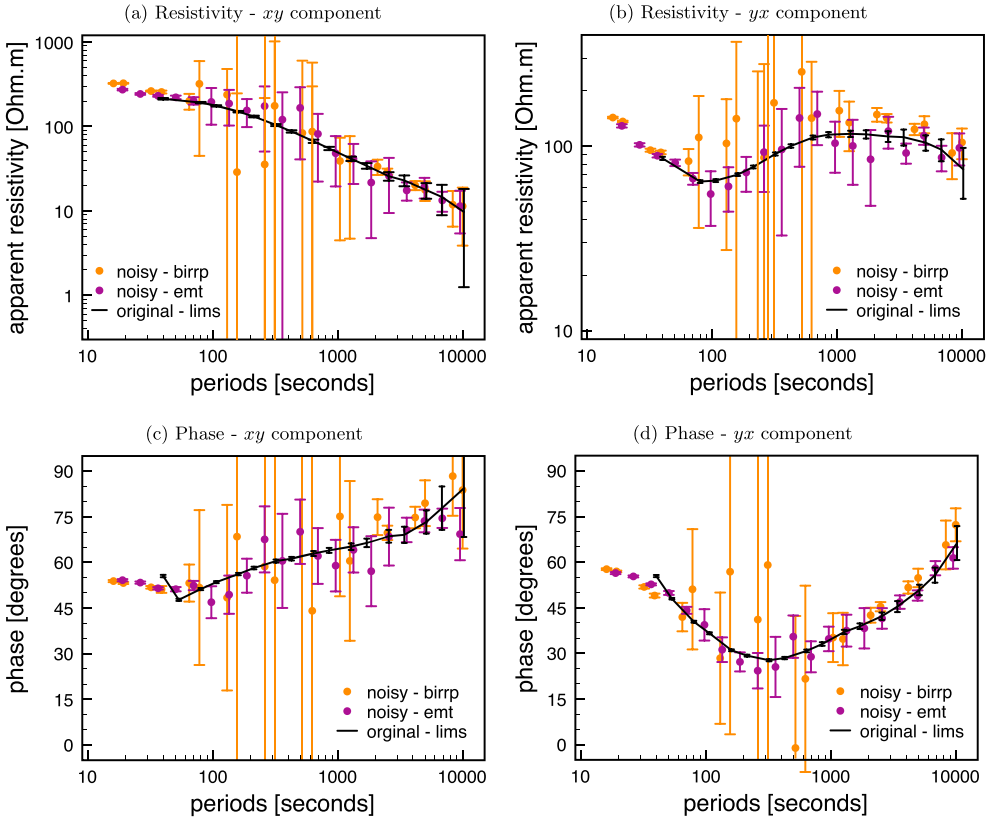


Figure 12. Southern Africa data (site 072) jammed with synthetic, nonstationary electric noise of high amplitude.

8.5. Problematic Real Data From Southern Africa

This last example is a real, broadband data set and has been acquired in a region where DC trains operate and active mining takes place. *Evans et al.* [2011] report problems in processing the data in particular due to these noise sources.

We focus on site 042 with 027 as reference for long-period measurements ($>$ 20 s) and with 043 as reference for short-period data ($<$ 20 s). The long-period data have been collected with LIMS instruments and the short-period data were measured by band 5 of Phoenix Systems' instruments. The site contains a tremendous amount of noise which makes interpretation difficult from about 3 s in Figure 13. The data of this site were originally processed with EMTF [*Egbert*, 1997] for short periods (BBMT) and with the LIMS processing algorithm [*Jones and Jodicke*, 1984] for the long-period data (long-period magnetotelluric (LMT)). The amplitude results from LMT have been scaled by the acquisition team to account for static shift according to the interpretation of the BBMT data, whereby the results from EMT are unchanged, since it does not suggest that the measurements of the long-period data have been affected notably by static shift.

We use originally published data for this plot, because we argue that (in time of original publication) the interpretation of the data (that it is affected by static shift) was wrong due to some noise effect. The EMT result is not shifted, because it does not lead to the conclusion that the LMT data require a shift, which exemplifies the long-ranging effect of noise beyond data processing and highlights the strength of the algorithm in this situation.

Besides the apparent noise between 3 s and 20 s, the phase estimations between 1 s and 100 s obtained from EMT are consistently 5° to 10° lower than the results estimated by the other algorithms, which we cannot explain at this point. Two possible reasons for this discrepancy could be due to nonstationary spectral leakage in the other algorithms (compare processing of a purely nonstationary data set in Figure 6) and due to strong correlated noise distorting significantly the first two dominant principal components.

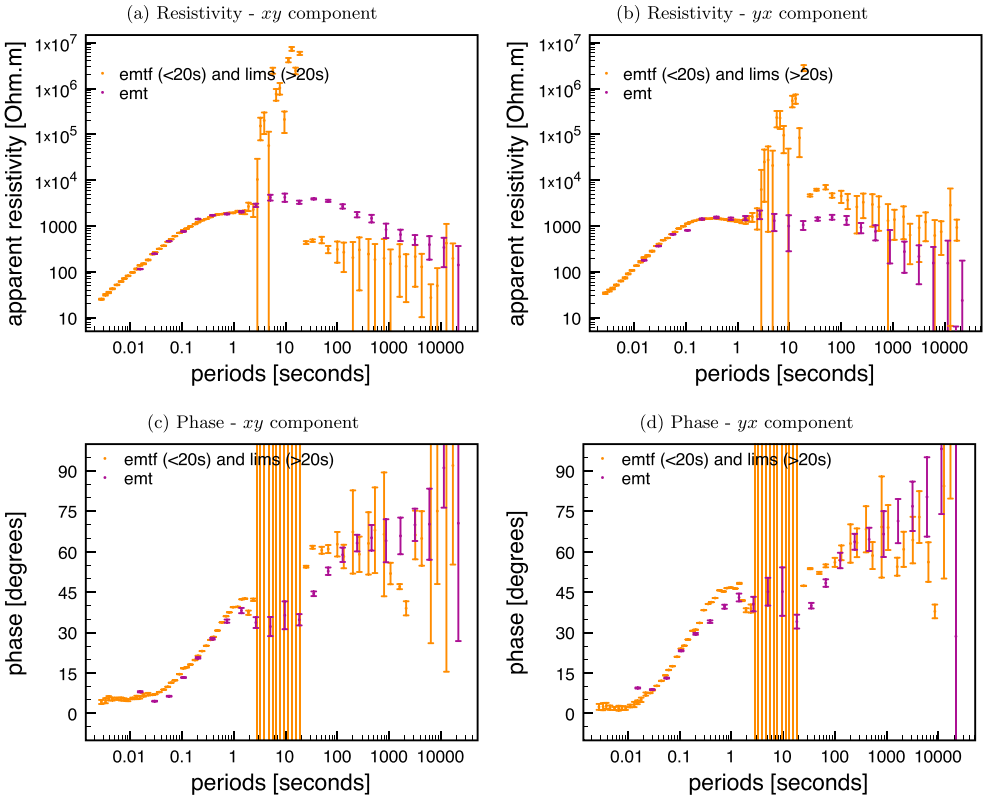


Figure 13. Site 042 contains a tremendous amount of noise which complicates interpretation of periods of 3 s and more. The EMTF [*Egbert*, 1997] (BBMT) and the LIMS [*Jones and Jodicke*, 1984] (LMT) results are the original results from the SAMTEX [*Evans et al.*, 2011]. The merge of BBMT and LMT responses was performed manually, and as it still is common practice, the LMT apparent resistivity were shifted to match the BBMT apparent resistivity at the overlapping periods.

9. Conclusion

In the course of this work, we outlined a robust magnetotelluric data processing scheme purely based on nonstationary methods and showed that its results compare to state-of-the-art algorithms. In contrast to other groups, our algorithm directly uses the instantaneous parameters of the measured multivariate time series and therefore naturally handles nonstationary sources. In theory, our scheme is less apt to introduce bias from spectral leakage due to this kind of noise and our synthetic and real data examples support this.

The algorithm carefully incorporates the most general and important data quality control measures like remote referencing and robust statistics as countermeasures for uncorrelated noise between occupied sites and control of highly influential but statistically unlikely data points, respectively.

This new methodology operates in a time-frequency domain and, therefore, potentially enables new data quality control measures like controlling instantaneous changes in the parameters amplitude, phase, and frequency, which could be investigated in a future work.

The function to select the independent data samples assures that the correct amount of data is selected, but the function of choice is somewhat arbitrary. On one hand it can be seen as an advantage that the samples are drawn arbitrary or random, but on the other hand alternative ways should be investigated for assessing their performance.

We demonstrated on synthetic and real data that a nonstationary approach in MT processing can be fruitful. The synthetic, nonstationary source in this work is specifically designed to disturb the Fourier transform and to break its assumptions; however, the results provide an insight in how bad real, nonstationary noise can affect MT measurements and encourage to verify the findings on more real-world data sets that are suspected to contain, in particular, nonstationary noise, e.g., data that are acquired close to train lines, mining activity, or electric fences.

Lastly, we present one such example of real data and find that at the time of original data processing, even the interpretation of the data has been affected by nonstationary noise, because the long-period data have been corrected unnecessarily for static shift by the original processing team.

We encourage to reassess more data sets that have been difficult to process in the past in order to investigate for nonstationary effects. However, we wish to stress that at this moment, our proposed algorithm is realized in MATLAB and runs rather slow (about 1 day for 10 million data points) on desktop computers. Most of the time, it delivers similar results compared to much faster and more efficient processing algorithms like BIRRP [*Chave and Thomson*, 2004], EMTF [*Egbert*, 1997], or the LIMS processing algorithm [*Jones and Jodicke*, 1984]. Therefore, we consider our algorithm a special purpose code for data that are suspected to be contaminated by nonstationary effects.

Appendix A: SynDat: Computing (Non)Stationary Synthetic Data for MT

Availability of synthetic data is fundamental for hypothesis testing in many areas of applied science, since it offers a simple and noise-free mean of acquiring test data, which could be expensive, difficult, or time consuming in the laboratory or in the field, and it allows to design easily custom-made properties of test data, which often help to spotlight both, important problems and findings in a hypothesis.

We use the MATLAB program SynDat to generate (non)stationary synthetic data in the course of this work. It allows to define freely the frequency and amplitude time series of numerical remote channels (as given in (6)), which are used to compute synthetic MT data (as given in (5)) by means of the nonstationary convolution theorem according to *Neukirch and Garcia* [2013]. Additionally, the synthetic data can be modeled for any impedance by importing the respective transfer function (TF) from files of the Electrical Data Interchange format or be computed for the impedance of a one-dimensional conductivity model [*Keller and Frischknecht*, 1966]. This program is freely available upon request to any of the authors.

Acknowledgments
The authors thank Alan Chave, Woods Hole Oceanographic Institution, for suggesting the use of the autocorrelation for data independency, which ultimately led to our final solution. Further, we appreciate the constructive discussions with Jin Chen and Bjoern Heincke at Geomar, Kiel. Data from Southern Africa are part of the SAMTEX. This work has been funded by REPSOL under the framework of the CO-DOS project. And lastly, we thank Jin Chen, an anonymous reviewer, and the Associate Editor for their very constructive criticism.

References

Adam, A., L. Szarka, J. Verö, A. Wallner, and R. Gutdeutsch (1986), Magnetotellurics (MT) in mountains—Noise, topographic and crustal inhomogeneity effects, *Phys. Earth Planet. Inter.*, *42*(3), 165–177.

Battista, B., C. Knapp, T. McGee, and V. Goebel (2007), Application of the empirical mode decomposition and Hilbert–Huang transform to seismic reflection data, *Geophysics*, *72*(2), H29—H37, doi:10.1190/1.2437700.

Berdichevsky, M. N., and I. A. Bezruk (1973), *Magnetotelluric Sounding With the use of Mathematical Filters*. Izv. Akad. Nauk SSSR.

Cagniard, L. (1953), Basic theory of the magnetotelluric method of geophysical prospecting, *Geophysics*, *18*, 605–635.

Cai, J. H. (2012), Magnetotelluric response function estimation based on Hilbert–Huang transform, *Pure Appl. Geophys.*, *170*, 1899–1911, doi:10.1007/s00024-012-0620-3.

Cai, J. H., J. T. Tang, X. R. Hua, and Y. R. Gong (2009), An analysis method for magnetotelluric data based on the Hilbert–Huang Transform, *Explor. Geophys.*, *40*(2), 197.

Chave, A. (2012), The magnetotelluric method: Theory and practice, in *Estimation of the Magnetotelluric Response Function*, edited by A. D. Chave and A. G. Jones, pp. 165–218, Cambridge Univ. Press, Cambridge, U. K.

Chave, A. D., and D. J. Thomson (1987), On the robust estimation of power spectra, coherences, and transfer functions, *J. Geophys. Res.*, *92*(B1), 633–648.

Chave, A. D., and D. J. Thomson (2004), Bounded influence magnetotelluric response function estimation, *Geophys. J. Int.*, *157*(3), 988–1006, doi:10.1111/j.1365-246X.2004.02203.x.

Chen, J., and M. Jegen (2008), *Empirical Mode Decomposition and Hilbert–Huang transform (HHT) in MT Data Processing*. 19. International Workshop on EM Induction in the Earth.

Chen, J., B. Heincke, M. Jegen, and M. Moorkamp (2012), Using empirical mode decomposition to process marine magnetotelluric data, *Geophys. J. Int.*, *190*(1), 293–309, doi:10.1111/j.1365-246X.2012.05470.x.

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

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

Egbert, G. D., and J. R. Booker (1986), Robust estimation of geomagnetic transfer functions, *Geophys. J. Int.*, *87*(1), 173–194.

Evans, R. L., et al. (2011), Electrical lithosphere beneath the Kaapvaal craton, southern Africa, *J. Geophys. Res.*, *116*(B4), B04105, doi:10.1029/2010JB007883.

Flandrin, P., and G. Rilling (2004), Empirical mode decomposition as a filter bank, *Signal Process Lett.*, *11*, 112–114.

Gamble, T. D. (1979), Magnetotellurics with a remote magnetic reference, *Geophysics*, *44*, 53–68.

Garcia, X., and A. G. Jones (2002), Atmospheric sources for audio-magnetotelluric (AMT) sounding, *Geophysics*, *67*(2), 448–458.

Garcia, X., A. D. Chave, and A. G. Jones (1997), Robust processing of magnetotelluric data from the auroral zone, *J. Geomagn. Geoelec.*, *49*(11–12), 1451–1468.

Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. Phys. Soc. A*, *454*(1971), 903–995.

Huang, N. E., Z. Wu, S. R. Long, K. C. Arnold, and X. Chen (2009), On instantaneous frequency, *Adv. Adapt. Data Anal.*, *01*, 77.

Hubert, M. (2003), A robust PCR method for high-dimensional regressors, *J. Chemom.*, *17*(8-9), 438–452, doi:10.1002/cem.783.

Hubert, M., P. J. Rousseeuw, and T. Verdonck (2009), Robust PCA for skewed data and its outlier map, *Comput. Stat. Data Anal.*, *53*, 2264–2274.

Jones, A. G., and H. G. Jodicke (1984), Magnetotelluric transfer function estimation improvement by a coherence-based rejection technique, paper EMU, 271, presented at 54th Annl. Int. Mtg, Soc. Excl. Geophys., Atlanta, Ga.

Jones, A. G., A. D. Chave, G. D. Egbert, and D. Auld (1989), A comparison of techniques for magnetotelluric response function estimation, *J. Geophys. Res.*, *94*, 14,201–14,213.

Junge, A. (1996), Characterization of and correction for cultural noise, *Surv. Geophys.*, *17*, 361–391.

Keller, G. V., and F. C. Frischknecht (1966), *Electrical Methods in Geophysical Prospecting*, 519 pp., Pergamon Press, Oxford, U. K.

Liu, R. Y. (1988), Bootstrap procedures under some non-IID models, *Ann. Statist.*, *16*, 1696–1708.

Liu, W., and M. Fujimoto (2011), *The Dynamic Magnetosphere*, Springer, Netherlands.

Neukirch, M., and X. Garcia (2013), Non stationary time series convolution: On the relation between the Hilbert-Huang and Fourier transforms, *Adv. Adapt. Data Anal.*, *5*, 1,350,004.

Oettinger, G., V. Haak, and J. C. Larsen (2001), Noise reduction in magnetotelluric time series with a new signal-noise separation method and its application to a field experiment in the Saxonian Granulite Massif, *Geophys. J. Int.*, *146*, 659–669.

Rakov, V. A., and M. A. Uman (2007), *Lightning, Physics and Effects*, Cambridge Univ. Press, Cambridge, U. K.

Rehman, N., and D. P. Mandic (2009), Multivariate empirical mode decomposition, *Proc. Phys. Soc. London, Sect. A*, *466*(2117), 1291–1302.

Reiersøl, O. (1941), Confluence analysis by means of lag moments and other methods of confluence analysis, *Econometrica*, *9*, 1–24.

Smirnov, M. Y., and G. D. Egbert (2012), Robust principal component analysis of electromagnetic arrays with missing data, *Geophys. J. Int.*, *190*, 1423–1438, doi:10.1111/j.1365-246X.2012.05569.x.

Szarka, L. (1987), Geophysical aspects of man-made electromagnetic noise in the earth—A review, *Surv. Geophys.*, *9*, 287–318.

Verboven, S. (2005), LIBRA: A Matlab library for robust analysis, *Chemometr. Intell. Lab.*, *75*, 127–136.

Viljanen, A. (2012), The magnetotelluric method: Theory and practice, in *Description of the Magnetospheric/Ionospheric Sources*, edited by A. D. Chave and A. G. Jones, pp. 96–121, Cambridge Univ. Press, Cambridge, U. K.

Weckmann, U., A. Magunia, and O. Ritter (2005), Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme, *Geophys. J. Int.*, *161*(3), 635–652.

Zhang, R. R., S. Ma, E. Safak, and S. Hartzell (2003), Hilbert-Huang transform analysis of dynamic and earthquake motion recordings, *J. Eng. Mech.*, *129*, 861–875.

Non Stationary Magnetotelluric Data Processing

Studies have proven that the desired signal in the electromagnetic (EM) field for Magnetotelluric (MT) surveys can be regarded as 'quasi stationary' (i.e. sufficiently stationary to apply a windowed Fourier transform). However, the measured time series often contain environmental noise, which may not fulfill this requirement for the application of the Fourier Transform (FT) and therefore may lead to false or unreliable results under methods that rely on the FT. In light of paucity of MT data processing algorithms in the presence of non stationary noise, in this thesis, I elaborate a robust, non stationary algorithm, which can compete with sophisticated, state-of-the-art algorithms in accuracy and precision. In addition, I proof mathematically the algorithm's viability and validate its superiority to other codes processing non stationary, synthetic and real MT data.

Maik Neukirch received a MSc degree in Applied Geophysics from the Joint Master program by TU Delft, ETH Zurich and RWTH Aachen University in 2009. In his Master thesis he studied the theoretical sensitivity of capacitively-coupled Electrical Resistivity Tomography.

He pursued his doctoral studies at the Institute of Marine Sciences (Institut de Cienciès del Mar) in Barcelona, where his work was focused on non stationary time series analysis and its application to the Magnetotelluric (MT) method. The final thesis was presented at the University of Barcelona in 2014.