SCHOOL OF ENGINEERING AND DESIGN
BRUNEL UNIVERSITY

SEISMOLOGICAL DATA ACQUISITION AND SIGNAL PROCESSING USING WAVELETS

A thesis submitted for the degree of Doctor of Philosophy

BY

GEORGIOS P. HLOUPIS

2008
LONDON, UK

Table of Contents

List of Figures

List of Tables

Abbreviations

ADSL	Asymmetric Digital Subscriber Line
ATM	Asynchronous Transfer Mode
BAMS	Bayesian Adaptive Multiresolution
BBD	Burstable Bandwidth on Demand
BFTR	Bayesian Hypothesis testing
BGP	Border Gateway Protocol
BLOCKJS	Block Thresholding
BPF	Band Pass Filter
C(*l*)	Coiflet filter of length *l*
CF	Compact Flash
CIR	Committed Information Rate
CPU	Central Processing Unit
CRC	Central recording Centre
CV	Cross Validation
CWT	Continuous Wavelet Transform
D(*l*)	Daubechies filter of length *l*
DAC	Digital to Analog Converter
DAS	Data Acquisition System
DECOMPSH	Deterministic/Stohastic
DSL	Digital Subscriber Line
DWT	Discrete Wavelet Transform
EDT	Equal Differential Time
EEWS	Earthquake Early Warning System
EMSC	European Mediterranean Seismological Centre
FDR	False Discovery Rate
FDSN	International Federation of Digital Seismograph Networks
FEC	Forward Equivalency Class
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FTP	File Transfer Protocol
GI-NOA	Geodynamics Institute of National Observatory of Athens
GPRS	General Packet Radio Service
GPS	Global Positioning System
GST	Global System Table
HPF	High Pass Filter
HSNC	Hellenic Seismological Network of Crete
HT	Hard Thresholding
HVSR	Horizontal to Vertical Signal Ratio
HYBSURE	Hybrid SURE
IANA	Internet Assigned Number Authorities
IFT	Inverse Fourier Transform
IP	Internet Protocol
ISC	International Seismological Committee
ISDN	Intergraded Services Digital Network
IS-IS	Intermediate System to Intermediate System
ISP	Internet Service Provider
IWT	Inverse Wavelet Transform
JCL	Job Control Language
L2TP	Level 2 Tunnelling Protocol
LA(*l*)	Least Asymmetric filter of length *l*

LCD	Liquid Crystal Display
LGS	Laboratory of Geophysics and Seismology
LPF	Low pass Filter
LSP	Label Switched Path
LTA	Long Time Average
MAD	Medium Absolute Deviation
MINH	Minimax estimator - hard thresholding
MINS	Minimax estimator - soft thresholding
MMS	Multimedia Messaging System
MODWT	Maximum Overlap Discrete Wavelet Transform
MPLS	Multi Protocol Label Switching
MRA	Multi Resolution Analysis
MSE	Mean square Error
NEIGHBL	Neighborhood Block
NHLM	New global High Noise Model
NLNM	New Low Noise Model
OS	Operating System
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
PPP	Point to Point Protocol
QTF	Quasi Transfer Function
RSVP	Resource ResrVation Protocol
RTP	Reftek Transmission Protocol
RTS	Real Time Subsystem
SAS	Seismic Analysis Subsystem
SCAD	Smoothly Clipped Absolute Deviation
SMS	Short Messaging System
SNR	Signal to Noise Ratio
SPT	System Process Table
ST	Soft Thresholding
STA	Short Time Average
STD	Standard Deviation
STFT	Short Time Fourier Transform
SURE	Stein's Unbiased Risk Criterion
TCP	Transmission Control Protocol
TE	Traffic Engineering
TEI	Technological Educational Institute
THE	Aristotle University of Thessalonica Seismological Network
THRSDA	Hypothesis testing thresholding
TI	Translation Invariant
TSS	Two Station Subarray
UDP	User Datagram Protocol
VISU	VisuShrink
VoIP	Voice over IP
VPN	Virtual Private Network
VSAT	Very Small Aperture Terminal
WAE	Wavelet Azimuth estimation
WDM	Wavelet Denoising Method
WEpE	Wavelet Epicentral Estimation
WME	Wavelet Magnitude estimation
WT	Wavelet Transform
XML	Extensible Markup Language

Acknowledgments

I am grateful for all the support I have received from a large number of people who contributed for the completion of this thesis. Special thanks and my deepest and most sincere appreciation to my Supervisor, Prof. John Stonham from School of Engineering and Design at Brunel University, for his guidance, encouragement and support that provided to me especially in periods where personal problems decelerate the evolution of this thesis.

I owe a special debt of gratitude to Prof. Filippos Vallianatos from the Department of Natural Resources and Environment at Technological Educational Institute at Crete, for his guidance and for sharing his expert knowledge in the fields of geophysics and seismology.

Deep heart thanks to Prof. Dimos Triantis from Electronics Department of the Technological Educational Institution of Athens for the inspiration that provide to me in order to begin this research. Without his insistency and his exhortation this work would never start.

Special thanks to Prof. John Makris, Mr Ilias Papadopoulos and Ms Asimina Kanta from the Laboratory of Seismology and Geophysiscs at Technological Educational Institute at Crete, for all the fruitful and productive discussions and suggestions that provided to me in order to improve several parts of this thesis.

I would like to thank the National Foundation of Scholarships (IKY) for the financial support during this research.

Abstract

This work deals with two main fields:

a) The design, built, installation, test, evaluation, deployment and maintenance of Seismological Network of Crete (SNC) of the Laboratory of Geophysics and Seismology (LGS) at Technological Educational Institute (TEI) at Chania.
b) The use of Wavelet Transform (WT) in several applications during the operation of the aforementioned network.

SNC began its operation in 2003. It is designed and built in order to provide denser network coverage, real time data transmission to CRC, real time telemetry, use of wired ADSL lines and dedicated private satellite links, real time data processing and estimation of source parameters as well as rapid dissemination of results. All the above are implemented using commercial hardware and software which is modified and where is necessary, author designs and deploy additional software modules. Up to now (July 2008) SNC has recorded 5500 identified events (around 970 more than those reported by national bulletin the same period) and its seismic catalogue is complete for magnitudes over 3.2, instead national catalogue which was complete for magnitudes over 3.7 before the operation of SNC.

During its operation, several applications at SNC used WT as a signal processing tool. These applications benefited from the adaptation of WT to non-stationary signals such as the seismic signals. These applications are:

- HVSR method. WT used to reveal undetectable non-stationarities in order to eliminate errors in site's fundamental frequency estimation.
- Denoising. Several wavelet denoising schemes compared with the widely used in seismology band-pass filtering in order to prove the superiority of wavelet denoising and to choose the most appropriate scheme for different signal to noise ratios of seismograms.
- EEWS. WT used for producing magnitude prediction equations and epicentral estimations from the first 5 secs of P wave arrival.
- As an alternative analysis tool for detection of significant indicators in temporal patterns of seismicity. Multiresolution wavelet analysis of seismicity used to estimate (in a several years time period) the time where the maximum emitted earthquake energy was observed.

1. Seismological Signal Processing – An Introduction

1.1 Historical review

From the ancient years earthquakes have attracted the attention of human nature. This caused because there is no other physical phenomenon that produces so much fear and insecurity as an earthquake. It happens suddenly, it is unexpected and until today its prediction remains an open question. Only the last twenty five years there were many destructive earthquakes (Pakistan 2005, North Algeria 2003, El Salvador 2001, Athens 1999, Turkey 1999, Kozani 1995, Nortdridge 1994, Loma Prieta 1989, Kobe 1995, Mexico City 1985, California 1983) that caused injuries and deaths for a significant number of people. The vast majority of earthquakes were take place in areas at the borders of the plates of the outer crust of the Earth, whilst a small number of earthquakes occur because of the eruption of volcanoes. One of the most seismogenic zones around the world is the Eastern part of the Mediterranean Sea where a ferocious event is taking place, "the sinking of the African plate underneath the Eurasian plate" (Kiratzi & Louvari, 2003). The earthquakes there often exceed magnitude 7 on the Richter scale becoming one of the human's most fearful natural disasters.

Since short term earthquake prediction is not a widely accepted procedure the most practical way for earthquake protection is via knowledge of the local geology as well as its dynamic behaviour during earthquakes. A range of approaches and strategies did by engineers include:

a) Measuring earthquakes' attributes by installation and continuous operation of seismological networks.
b) Site characterization by estimating the influence of local site effects during earthquakes.
c) Earthquake early warning system (EEWS) in order to produce useful alerts before the approach of destructive waves.
d) Study of local regional seismicity in order to identify possible characteristic patterns or relations.

The above procedures try to measure attributes of the earthquake and its occurrence times which both are non-stationary quantities. The widely used Fourier transform, although useful for stationary signals from simple dynamic systems

consist of a linear superposition of few independent, strong, nonevolving periodicities. It has severe drawbacks for analyzing signals from the following two important categories: Signals that include transients or variable periodicities and signals that significantly depart from stationarity, as is the case of intermittent energy releases.

1.2 Aim of this Research

This research focus on the engineering aspects of recently installed seismological network as well as to the application of Wavelet Transform (WT), to seismological data acquired from the aforementioned network. The engineering aspects associated with hardware and software for the real time Hellenic Seismological Network of Crete (HSNC) which is mainly designed and fully built, installed, tested, deployed, maintained and evaluated by the author, are presented. In addition, the applicability of wavelets is investigated in the following areas:

a) As an improvement tool for a method widely used to estimate a site's fundamental frequency (Horizontal to Vertical Spectral Ratio – HVSR) and thus estimate local site effects. WT used to reveal undetectable non-stationarities in order to eliminate errors in site's fundamental frequency estimation.

b) As a denoising scheme for the improvement of earthquake recordings acquired by a recently installed seismological network. Several wavelet denoising schemes compared with the widely used in seismology bandpass filtering in order to prove the superiority of wavelet denoising and to choose the most appropriate scheme for different signal to noise ratios of seismograms .

c) As a real time earthquake magnitude and epicentral estimator in order to provide to an EEWS the necessary rapid information for possible alert generation. Wavelet coefficients used to predict earthquake's magnitude and to estimate epicentral location both in the first 4-6 secs after first P wave arrival.

d) As an alternative analysis tool for detection of significant indicators in temporal patterns of seismicity. Multiresolution wavelet analysis of seismicity used to estimate (in a several years time period) the time where the maximum emitted earthquake energy was released.

1.3 Literature review

In signal processing, the representation of signals is crucial to the potential outcome of the project. For example, the Arabic numeral representation permits one to easily notice if a number is a power of 10, but more difficult to realize a power of 16. With the hex representation, the situation is reversed. Meyer (1993) wrote "*any particular representation makes certain information explicit at the expense of information that is pushed into the background and may be quite hard to recover*".

Non-stationary signals observed in a variety of engineering fields (wind, ocean, and earthquake engineering). The inability of conventional Fourier analysis to preserve the time dependence and describe the evolutionary spectral characteristics of non-stationary processes requires tools which allow time and frequency localization beyond standard Fourier analysis. The spectral analysis of nonstationary signals cannot describe the local transient features due to averaging over the duration of the signal. This makes the Fourier representation inadequate when it comes to analyzing transient signals. In seismological signal processing concentrating on (e.g an earthquake) is a process for selecting the essential information from an overwhelming amount of data. In order to localize both the frequency and the time information in a signal several transforms and bases have been proposed (Mallat 1989a, Vetterli 1995). An FFT based method, the Short Time Fourier transform (STFT, also called the window Fourier transform or the Gabor transform) provides time and frequency localization to establish a local spectrum for any time instant. The key feature of the STFT is the application of the FT to a time varying signal when the signal is viewed through a narrow window centered at a specific time. The signal is multiplied by window function (typically Gaussian) and the Fourier integral is applied to the windowed signal. The local frequency content is then obtained at time. The window is moved to a new time and the process is repeated. High resolution cannot be obtained in both time and frequency domains simultaneously due to Heisenberg's inequality (Pinsky, 2001). The window must be chosen for locating sharp peaks or low frequency features, because of the inverse relation between window length and the corresponding frequency bandwidth (Kareem, 1993). A multi resolution representation of the process can be achieved if there is an attribute that allows

variable resolution in time and frequency. Accordingly, the time-frequency window would narrow automatically to observe high frequency contents of a signal and widen to capture low frequency phenomena.

The projection of signals with continuous or discrete wavelet analyzing functions, that are by definition self-similar and well localized in both time (and space) and frequency (or wavenumber) domains, provides efficient descriptions, in terms of sparseness, of non-stationary, transient, or 'bursty' records. Fourier methods of signal decomposition use infinite sines and cosines as basis functions, whereas the wavelet transform uses a set of orthogonal basis functions which are local. A short duration, high frequency phenomenon is buried in a Fourier representation with the background averaged spectral content, whereas wavelet transformation allows the retention of local transient signal characteristics beyond the capabilities of the infinite harmonic basis functions by allowing a multi-resolution representation of a process. The wavelet transform originated in geophysics in the early 1980s for the analysis of seismic signals (Morlet et al., 1982a, b) and was later formalized by Grossmann and Morlet (1984) and Goupillaud et al. (1984). Important advances were made by Meyer (1992, and references therein), Mallat (1989a, b), Daubechies (1988, 1992), Chui (1992a), Wornell (1995), and Holschneider (1995), among others. These advances then impacted other areas of study and particularly applications in geophysics for process understanding. An increasingly large variety of applications using wavelets: detection of periodic signals in noisy time series (e.g. Otazu et al., 2002), monitoring of period variations (e.g. Frick et al., 1997, Fligge et al., 1999), unevenly sampled period analysis (e.g. Foster, 1996), derivation of multifractal properties (e.g. Arneodo et al., 1988, Argoul et al., 1989, Muzy et al., 1991, 1993, Arneodo et al., 1995, 1998), random multifractal synthesis (e.g. Benzi et al, 1993), localized structure identification (e.g. Lucek & Balogh, 1997, Roux et al., 1999, Mouri et al., 1999), polarization analysis (e.g. Baumjohann et al., 1999), denoising (e.g. Fligge & Solanki, 1997, Komm et al., 1999), signal to noise ratio enhancement (e.g. Zhang & Paulson, 1997), information compression (e.g. Muhlmann & Hanslmeier, 1996), multi-resolution image decomposition (e.g. Mallat, 1989a, Pantin & Starck, 1996), studies of random walks (Arneodo et al., 1996), flow structure analysis (Haynes & Norton, 1993), population classification (Bendjoya et al., 1991), clustering detection (Bijaoui

et al., 1993, Girardi et al., 1997, Lima Neto et al., 1997), fluid turbulence simulations (e.g Schneider et al., 1997).

A seismic signal is often seen as containing noise which obscures the information of interest. This noise may simply be measurement error, or it could be fluctuation details, which are problematic when the underlying trends or discontinuities are being investigated. Many methods have been developed for denoising a signal, in the hope that the noise can be suppressed and the significant patterns retained and revealed. These methods have ranged from simple moving averages or moving medians, to methods of considerable mathematical complexity. Wavelets provide denoising methods which are relatively simple to use, while adapting well and automatically to the form of the signal being denoised. The most common procedure where denoising using wavelets may be achieved is by shrinking individual coefficients and reconstructing a signal from these shrunken coefficients. A full description of the method may be found in Donoho and Johnstone (1994) and Donoho et al. (1995). Pazos et al (2003) used a non linear filter based on WT in order to enhance the SNR of seismograms. Galliana – Merino et al (2004) used WT to deconvovle seismograms from noise while same authors (Galliana – Merino et al, 2003) used WT to denoise short period seismograms. Although these efforts were successful they were limited to wavelet approaches very common in literature.

Site effects are the most important factor in the earthquake damage distribution and refers to the effect of the surface and subsurface geological characteristics, surface and subsurface topography and of the strong lateral variations and discontinuities on the enhancement of seismic motion. The influence of local site effects in the earthquake damage distribution has been studied extensively (Hough et al, 1990; Faccioli, 1991; Borcherdt et.al, 1992; Mucciarelli & Monachesi, 1997; Bard, 1999; Nakamura, 1989). The influence of surface geological characteristics on seismic motion has been well documented and is related to the modulation of wave propagation from the seismic source to the recorded site. Site effects studies are used to estimate dynamic response of various types of geological structures and the tectonic environment of the area under study, to estimate the ground amplification effects and to evaluate local sites effects due to surface geology. The Horizontal to Vertical Spectrum Ratio (HVSR), (Nogoshi-Igarashi, 1971, Nakamura 1989, 2000) is

a widely used technique (since it is rapid, cheap and non destructive) for estimating a site's fundamental frequency and amplification ratio and thus its expected behaviour during earthquake. It is based on ambient noise recordings. One point that receives criticism is that it assumes stationarity of the recorded signals since it uses FT to calculate the required spectrums. The use of WT to overcome this restriction seems natural but there are only a couple of studies that involves WT and HVSR. Specifically, Parolai (2006) used wavelet denoising in order to isolate high amplitude transients from ambient noise and Carniel et al (2007) used WT to improve the spectra peak correlation between horizontal and vertical components.

Over the last few years, there is an increasing number of studies in several active seismic areas of the world regarding experimentation of seismic early-warning. Systems have been developed and implemented in Taiwan, Japan, USA, Romania, Turkey and Mexico where alert signals from dense seismological networks in the earthquake source area are sent to nearby urban settlements. The Earthquake Early warning systems (EEWS) are based on real-time automated analysis of ground motion. The early information, provided by EEWS when the seismic waves are still propagating, can be used, to enable several standby procedures such as shutting down of critical systems. Depending on the network geometry and configuration around the potential seismic source and/or target area, the early warning systems can be distinguished in (Kanamori, 2005)

- regional (dense seismic network deployed on the potential earthquake source area
- on-site (single instrument or array of instruments deployed at the target site, which is distant from the earthquake source area

For regional EEWS systems the earthquake warning begins at the time of first P-wave detection at the network operated in the earthquake source area and lasts a few to several tens of seconds depending on the distance between the source and the target area. For on-site EEWS the time begins also at the at the time of first P-wave detection but last a few seconds depending on the distance between the site and the alerts' processing centre.

In order to provide real-time magnitude estimations for EEW purposes, a method based on the predominant period (which can be calculated from the first 4

secs just after the first P-wave arrival) has been first proposed by Allen and Kanamori (2003). The method has been mainly validated and calibrated on regional recordings of ground motion velocity. The evidence for a scaling relationship between the predominant period and magnitude observed on earthquake records from different seismic regions evaluated in several studies for (Allen, 2004, Olivieri & Schweitzer, 2007, Olivieri et. al., 2008). The main question that arised is if the predominant period estimator allow the prediction of the earthquake size from the measurement of frequency content of early radiated seismic signals. This hypothesis is called in question by Olson and Allen (2005) who argue about the deterministic nature of the rupture process and by Rydelek and Horiuchi (2006) who found no evidence of dominant frequency scaling with magnitude from the analysis of waveform data recorded by the Japanese Hi-net seismic network.

Alternative approaches were presented by Wu and Zhao (2006a) for California earthquake data. They determined an attenuation relationship for low-pass, filtered peak displacement amplitude records at the first three seconds after the arrival of the P-wave. Zollo et. al., (2006) used low-pass filtered peak amplitudes from P and S waves initial arrivals and show that is a robust measurement for estimating the magnitudes of earthquakes and has practical application in earthquake early warning systems.

The current research focus to the executions of the tasks involved the design, the deployment, the implementation and the installation of real time seismological network as well as on the elimination of the shortcomings of the above studies. More specific it is focus on the elimination of the following deficiencies:

a) The majority of the ordinary seismological signal processing tools uses bandpass filtering in order to improve the signal to noise ratio of recorded seismograms. It is well known that these conventional filters introduce some delay to the signal resulting to false time arrivals and to lower SNR seismograms introduce distortion in the shape of the P wave arrival (Scherbaum, 1996). This leads to wrong epicentral estimations (especially in real time and automatic estimation systems) since this procedure is highly dependent from the accuracy of time marking. Non linear denoising tools such as wavelets proposed by some studies and proved superior against

conventional filters regarding the improvement of SNR and elimination of false time arrivals. However, until today there is no study focused on the comparison of different wavelet denoising approaches for different SNR seismograms.

b) The widely used HVSR method for estimating site effects based on the spectral ratio between horizontal and vertical components of ambient noise. This ratio is calculated using conventional Fourier transform which by definition assumes stationarity. When the analyzed ambient noise contains non-stationarities (transients, near field artificial noise) this assumption is not verified. There are a couple of studies (Parolai, 2006, Parolai, et. al., 2009) that support that high amplitude short period transients does not affect the result of HVSR method. Until today there is no study that investigate the effect of low amplitude (which are non detectable by amplitude thresholding techniques) long period transients and the detection of them using non linear tools such as wavelets

c) Real time earthquake magnitude estimation from the first few seconds of P wave arrival seems efficient by using the predominant frequency estimator. This estimator as Wolfe (2006) and Simmons et al. (2006) pointed out, has significant scattering in some cases mainly, caused by its recursive calculation based on a spectral domain relation. The alternative solution is to transfer from time-frequency domain to time-scale domain such as wavelet domain. The use of wavelets in magnitude estimation has introduced by Simmons et al., (2006) but until today is applied only in California's seismic zone. In this research the earthquake magnitude is calculated using different WT and in addition WT is used at the same time to estimate the earthquake's epicentral. We emphasize that the use of a single transform has never been used for EEWS to the best of our knowledge

d) Scaling analysis of the temporal properties in earthquakes sequences has been used to understand the dynamic variability of seismicity by means of scale independent measures. These measures often suffer from non-stationarities that characterize the time dynamics of a seismic sequence. The use of wavelet based measure is a solution that can overcome this shortcoming.

1.4 Organization

This thesis has been divided into eight chapters of which Chapter 1 is the introduction.

Chapter 2 begins with necessary theoretical definitions of wavelet transform (WT) in its different implementations (redundant and non-redundant) and continues with some practical considerations to each one. Transformation of a typical seismological signal by means of WT is also presented along with the results of using different bases

Chapter 3 presents the data acquisition and signal processing system, the seismological network. More specific, a detailed description of installed seismological network along with the necessary software needed for data collection and processing initially presented. The procedures that followed in order to satisfy the different requirements that such a system had along with prototype solutions that produced during this research are also presented. Finally, implementation details of HVSR method are presented in order to clarify the several stages and the chapter finishes with the practical considerations that followed during the data collection procedures.

Chapter 4 deals with the use of WT in a method which estimates site's fundamental frequency (Horizontal to Vertical Spectral Ration method – HVSR) which is important for preliminary ambient noise surveys during the investigation of a station's location. This method involves spectrum calculation and since it uses FFT to implement this, it assumes stationarity. After a short presentation of HVSR method the influence of low amplitude non-stationarities is examined by using a set of well controlled artificial signals. Furthermore, the ability of WT to reveal non-stationary portions of examined signals is presented.

Chapter 5 covers the application of WT for denoising seismological data. It provides a detailed presentation of several wavelet denoising schemes which all initially applied to synthetic seismograms. After the qualification of denoising schemes, those which qualified as successful, applied to real seismograms (of different quality) in order to investigate if and how wavelet denoising schemes could increase seismogram's quality. Finally a comparison between wavelet denoised seismograms with band-passed ones is performed in order to estimate the qualitative

improvement by using wavelets in ordinary seismological operations such as epicentre estimation.

Chapter 6 looks at the application of redundant WT for an Earthquake Early Warning System (EEWS). This is achieved by rapid earthquake magnitude as well as epicentral estimation. The main questions that this chapter tries to approach are on the correlation between wavelet coefficients of seismograms recorder at South Aegean from the first 5secs of an earthquake and magnitude and on the use of wavelet coefficients in a parallel algorithm in order to estimate the epicentral location. Recordings from several seismological stations around Crete Island were examined and show that is possible to use the redundant WT for EEWS purposes.

Chapter 7 investigates the use of multiresolution wavelet analysis to seismic catalogue. Scaling analysis of the temporal properties in earthquakes sequences using the standard deviation of wavelet coefficients applied to two seismic catalogues from Western Mediterranean and in both cases, before strongest events, characteristic patterns were revealed.

Chapter 8 is the concluding chapter of the thesis, which highlights to the engineering aspects of the recently installed seismological network as well as to the four distinct areas of signal processing work:

- Wavelet denoising schemes are qualified as superior to ordinary bandpass filters as they preserve the shape and time localization of P wave arrivals and they performed better in seismic signal determination from noisy recordings .
- The WT can be used to identify non stationary disturbances in ambient noise recordings especially when their amplitude is comparable.
- The use of WT coefficients from the first few seconds after P wave arrival as an efficient magnitude and epicentral estimator.
- Multiresolution wavelet analysis of the standard deviation of WT coefficients reveals a two fold feature: at lower scales it decreases in correspondence with the occurrence of the strongest earthquake while at high scales it is characterized by oscillating behaviour, which is a typical for background seismicity.

In addition, recommendations for future work and suggestions for the expansion of proposed techniques are clearly stated.

2. Introduction to wavelets

2.1 Introduction

This section describes some key concepts of wavelet transform (WT) that are going to be used. Basic concepts and theory can be found in many books and papers at many different levels of exposition. Some of the standard books are Chui (1992a, 1992b), Daubechies (1992), Mallat (1998), Meyer (1993), Vetterli (1995). Introductory papers include Graps (1995), Strang (1994) and Cohen (1996).

Wavelet transforms (WT) can be considered as the localized equivalent of the Fourier transform. Their principle is that all signals can be decomposed from sets of other signals with constant shape but varying scale and amplitude. Just as Fourier Transform (FT) decomposes a signal into harmonic components, the WT decomposes a signal into wavelet components. Each wavelet component called *level* and by adding the complete set of levels it is possible to recover the initial signal. In case of FT, where bases are trigonometric sine and cosine functions, the signal is transformed into the frequency domain. But this is done at the cost of time resolution. With wavelets analysis, we may use different sets of bases as long as they follow some properties. Wavelets allow more flexibility for the signal analysis. In fact, with wavelets, the signal is transformed into a time scaled wavelet domain, where the bases are dilations and translations of the original mother wavelet. The latter works as a band pass function which limits the signal, both in frequency and time. Therefore, we can say that, while conserving the time dimension of the original data, the wavelets allow observing a time series at different scales.

2.2 Multiresolution analysis

In order to facilitate how wavelet analysis can be used to extract information at different scales basic aspects of multiresolution analysis (MRA) will be presented here. MRA presented by Mallat (1989a, 1989b, 1989c) based on the idea that a signal can be projected in a nested sequence of subspaces if some requirements (that define the multiresolution analysis) satisfied. Relationships between the wavelet and its complementary transforms can be established through the multiresolution analysis. The latter is guided from following axioms adopted by Daubechies (1988)

A multiresolution analysis consists of a sequence of embedded closed subspaces:

$$...V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2}... \tag{2.1}$$

Such that they have the following properties:

i. Upward completeness

$$\bigcup_{m\in\mathbb{Z}} V_m = L_2(\mathbb{R}) \tag{2.2}$$

ii. Downward completeness

$$\bigcap_{m\in\mathbb{Z}} V_m = \{0\} \tag{2.3}$$

iii. Scale invariance

$$f(t)\in V_m \Leftrightarrow f\left(2^m t\right)\in V_0 \tag{2.4}$$

iv. Shift invariance

$$f(t)\in V_0 \Rightarrow f\left(t-n\right)\in V_0, \forall n\in\mathbb{Z} \tag{2.5}$$

v. Existence of an orthonormal basis

$$\left\{\phi\left(t-n\right)\middle| n\in\mathbb{Z}\right\} \tag{2.6}$$

For V_o, where $\varphi \in V_o$

The function $\varphi(t)$ in Eq. 2.5 is called scaling function

According to Eq. 2.6 the function $\varphi(t)$, as a scaling function, forms an orthogonal basis for V_o. Using the scale and shift invariance we can write orthonormal basis $\left\{\phi_{m,n}\middle| n\in\mathbb{Z}\right\}$ for V_m (Daubechies 1988)

$$\phi_{m,n}(t) = 2^{-m/2}\phi(2^{-m}t-n), \quad \text{for } m,n\in\mathbb{Z} \tag{2.7}$$

The factor $2^{-m/2}$ in Eq.2.7 ensures that the norm $\left\|\phi_{m,n}(t)\right\|$ is always unity independent of m,n which means also that $\left\|\phi(t)=1\right\|$.

It is also proved (Daubechies 1988) that whenever a sequence of spaces satisfy Eq.2.2 – 2.5 there exists an orthonormal basis for $L_2(R)$

$$\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t-n), \quad \text{for } m,n\in\mathbb{Z} \tag{2.8}$$

Such that $\{\psi_{m,n}\}$ n ∈ Z, is an orthonormal basis for a space W_m which is orthogonal complement of the space V_m in V_{m-1}

$$V_m = V_{m+1}\oplus W_{m+1} \tag{2.9}$$

The function $\psi(t)$ in Eq.2.8 is called the wavelet function. A graphic representation of how a scaling and a wavelet function, acting as basis, can span a subspace depicted in Fig. 2.1

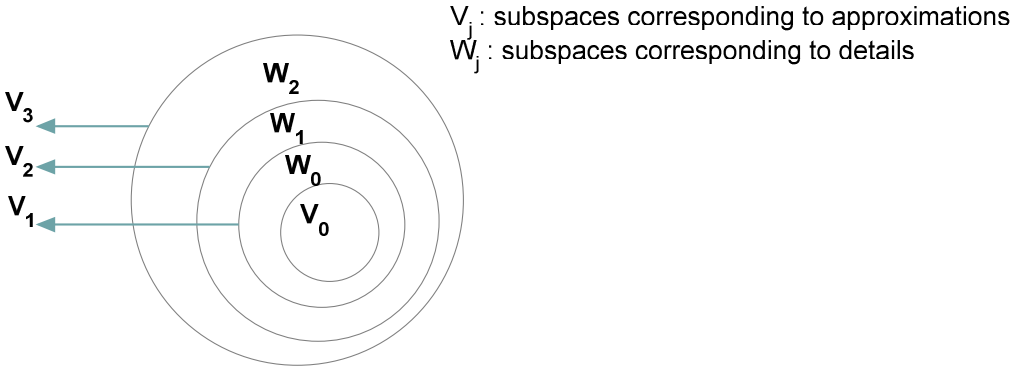


Figure 2.1. Nested vector subspaces spanned by wavelet and scaling basis

Function $\varphi(t)$can be represented as a linear combination of the function $\varphi(2t-n)$

$$\phi(t) = \sqrt{2}\sum_{n\in Z} h(n)\phi(2t-n) \tag{2.10}$$

where coefficients h are defined as :

$$\|h(n)\| = 1, \forall \mathrm{n} \in \mathrm{Z} \tag{2.11}$$

Equation 2.10 is called the *dilation equation*

The scaling function acts as a low-pass filter. It's easy to shown that $\int_{-\infty}^{\infty} \phi(t)dt = 1$

Hence its FT for ω=0 is $\varphi(0)$=1. In a similar way we can represent the wavelet function $\psi(t)$ as a linear combination of function $\varphi(2t-n)$

$$\psi(t) = \sqrt{2}\sum_{n\in Z} g(n)\phi(2t-n) \tag{2.12}$$

where coefficients g for orthonormal basis are defined as:

$$g_1(n) = (-1)^n g_0(-n+1) \tag{2.13}$$

Equation 2.12 is called the *wavelet equation*

In summary MRA provides a framework to approximate a function at different resolutions and determine the difference of information between successive approximations. Scaling function relates the approximations at different scales while wavelet function relates the detail information across scales.

2.3 Fourier and Wavelet transforms

For a signal *f(t)* FT has a great ability to capture signal's frequency as long as *f(t)* is composed of a limited number of stationary components. However, if non-stationary components exists at *f(t)* they are going are to spread over the whole frequency band. This is the initial limitation of FT: it offers excellent frequency resolution without offering time information about it because from the one dimensional domain (time), signal is transformed to another one dimensional domain (frequency). WT could overcome this limitation by providing both time and frequency localization since it transforms the signal from one dimensional domain to a two dimensional domain. Necessary concepts and a brief comparison between them will present.

In general, an event or physical process can be described either in the time domain as *f (t)* or in the frequency domain as *f (ω)* (Press et al., 1995). The FT is a useful tool in determining the frequency contents of a process *f (t)*:

$$\hat{f}(\omega) = (Ff)(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \qquad (2.14)$$

The original process *f(t)* can be reconstructed from the frequency components $\hat{f}(\omega)$ using the inverse Fourier transform (IFT), defined as:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{-i\omega t} d\omega \qquad (2.15)$$

The Fourier transform is generally suitable for analyzing a process whose frequency content does not change over time because it can only provide the average information of the frequency content. A windowed Fourier transform introduced by Gabor (1946) in order to determine time-localized information in the frequency domain:

$$(F^{window} f)(\omega, b) = \int_{-\infty}^{\infty} f(t) w(t-b) e^{-i\omega t} dt \qquad (2.16)$$

The function, *w(t)* (called the windowing function) can be any compactly supported function. The translations *w(t - b)* of the function *w(t)* ensure that the transform includes only the information around the center of the time of interest (*t=b*).

The main disadvantage of the windowed transform is about fixed window size; Once a *w(t)* is chosen, its size remain constant. This is not an ideal solution for a

nonstationary process which may consist of high frequency (rapid changes) and low frequency content (long period trends) requiring different window sizes. A wider window is preferred for low frequency phenomena while a narrow window should be used to analyze high frequency content

On the other hand, the wavelet transform can utilize different window sizes in order to provide both time-localized and frequency-localized information. A comparative visualization depicted at Fig. 2.2

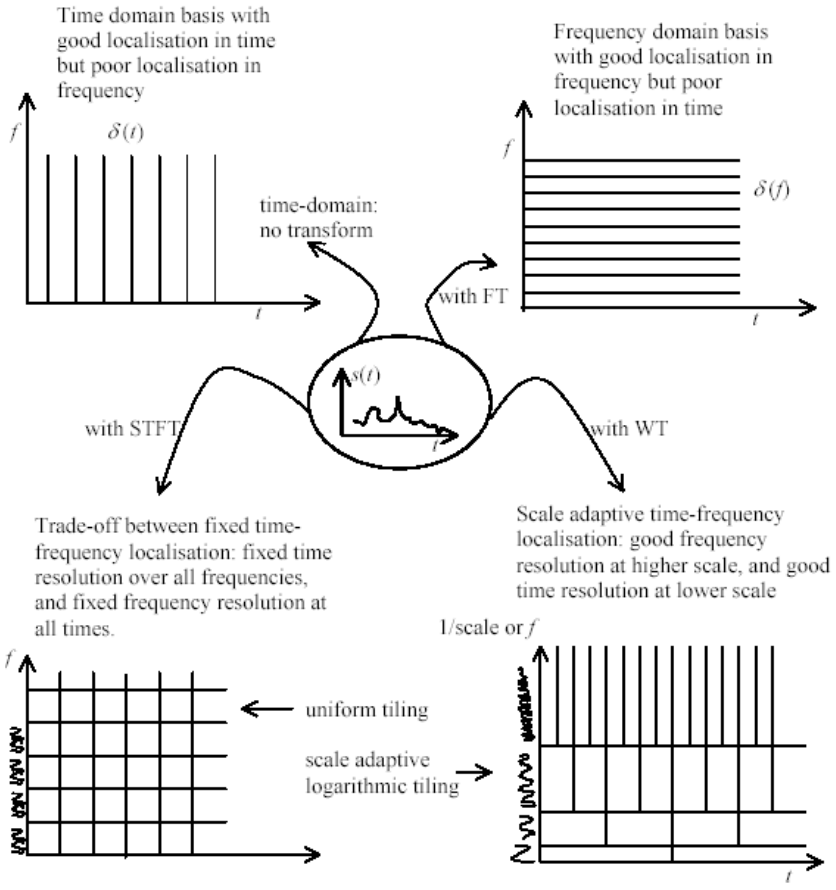


Figure 2.2. Comparative visualization of time-frequency representation of a non-stationary signal in time domain (top left), frequency domain (top right), STFT domain (bottom right) and WT domain (bottom right). (Shukla, 2003)

The wavelet theorem was first formally introduced by Grossmann and Morlet (1984). The wavelet transform at time $t = b$ and scale a is defined as:

$$Wf(a,b) = \frac{1}{\sqrt{2a}} \int_{-\infty}^{\infty} f(t) \psi \left(\frac{t-b}{a} \right) dt \qquad (2.17)$$

The function $\psi(t)$ is called a wavelet. A wavelet is a function that waves through the t-axis such that:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \tag{2.18}$$

The function $\psi(t)$ is called a wavelet if and only if its Fourier transform $\hat{\psi}(\omega)$ satisfies the admissibility condition (Heisenberg,1929):

$$\int_0^{\infty} \frac{\left|\hat{\psi}(\omega)\right|^2}{\omega} d\omega = \int_{-\infty}^{0} \frac{\left|\hat{\psi}(\omega)\right|^2}{\omega} d\omega = C_{\psi} < \infty \tag{2.19}$$

The coefficients produced by $Wf(a,b)$ are called *wavelet coefficients*. The parameter a is called *dilation* or *scale* and is equivalent to the window size in the windowed Fourier transform. By changing a, we can adjust this window size adjusting in this way the analyzing information contained in different frequencies of the data. A small value of a may be used to analyze high frequency information while a larger value of a may be used to analyze a lower frequency process. The parameter b is called *translation* and is equivalent to shifting in a windowed transform. If $a, b \in \Re$ then the WT is called Continuous Wavelet Transform (CWT).

Wavelets are constructed to meet certain mathematical properties in order to be used as basis functions to transform other functions. Depending on application extra restrictions may apply when constructing specific wavelets. The simplest form of wavelet is the Haar wavelet, which is described by the Haar function (Haar, 1910) defined as:

$$\psi(t) = \begin{cases} 1 & 0 \le t < 1/2 \\ -1 & 1/2 \le t < 1 \\ 0 & otherwise \end{cases} \tag{2.20}$$

The Haar function defined in Eq. 2.20 is called a *mother* wavelet. A family of wavelets can be constructed from the mother wavelet by dilations using the dilation parameter a and translations using parameter b. Fig. 2.3 displays a few members of the Haar wavelet families.

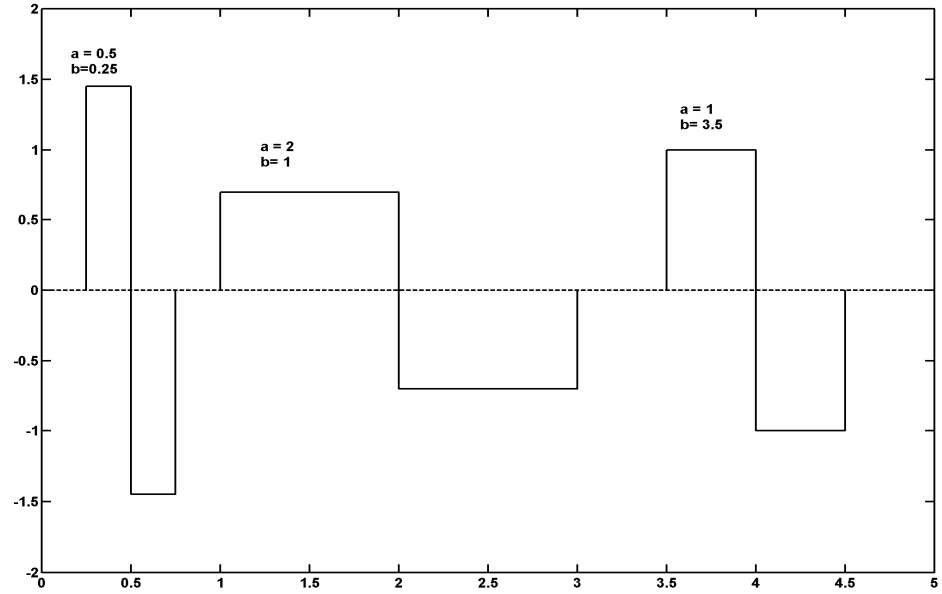


Figure 2.3. Wavelets constructed from Haar basis function.

The graphic representation of CWT is illustrated by the scalogram which is defined as the square magnitude of the CWT coefficients $|W_f(\alpha,b)|^2$. Typical example can be seen at Fig.2.4

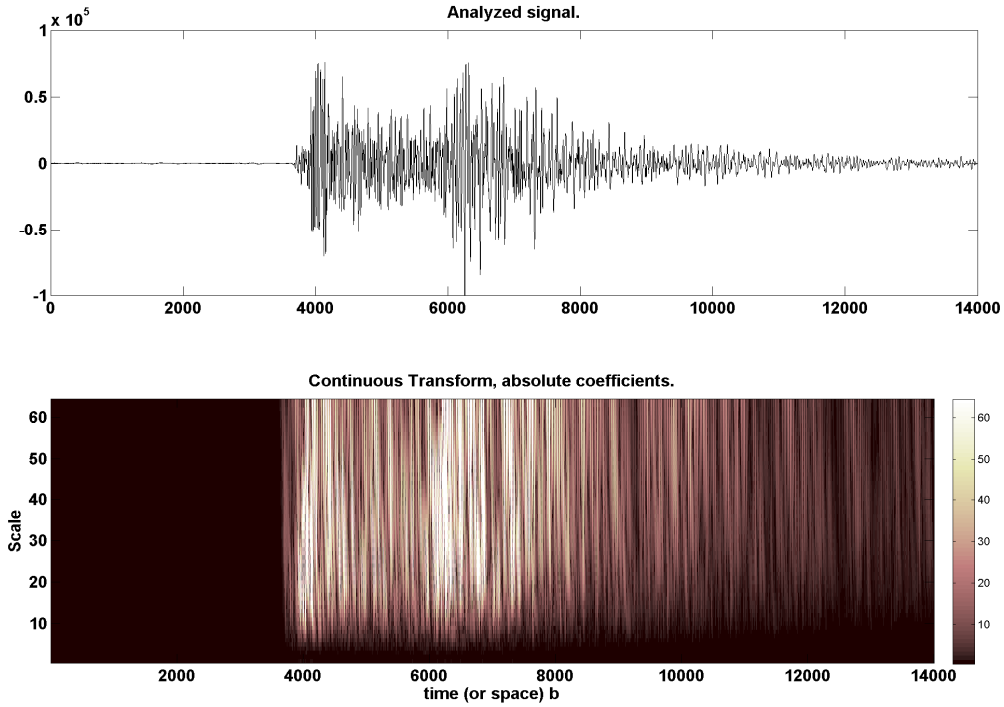


Figure 2.4. Vertical component of a seismic motion recording (top) and is associated CWT scalogram (bottom). Colour scale depicts the squared magnitude of associated wavelet coefficient (brighter scales: higher values, darker scales: lower values)

WT as a linear transform holds the following properties similar to other linear transforms such as FT (these properties can easily be shown by using the Haar wavelet as basis and unit impulse function as transformed signal):

- *Linearity.* As the CWT is defined by an inner product it holds its linearity
- *Time invariance*. The CWT of a time delayed version of a signal is a time-delayed version of its CWT. This important property is lost for discrete sampling schemes.
- *Energy conservation*. The CWT conserves energy in a similar way that FT did using Parceval's theorem

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{a^2} |Wf(\alpha, b)|^2 d\alpha \cdot db \qquad (2.21)$$

- Scaling property. For an *f(t)* with CWT estimated as $W_f(a,b)$ a scaled signal $f'(t) = \frac{1}{\sqrt{\alpha'}} f\left(\frac{t}{\alpha'}\right)$ leads to a CWT defined as $W_{f'}(\alpha, b) = W_f\left(\frac{a}{a'}, \frac{b}{a'}\right)$. Scaling property illustrated in Fig.2.5 showing the CWT of the same signal scaled by a factor of 2.

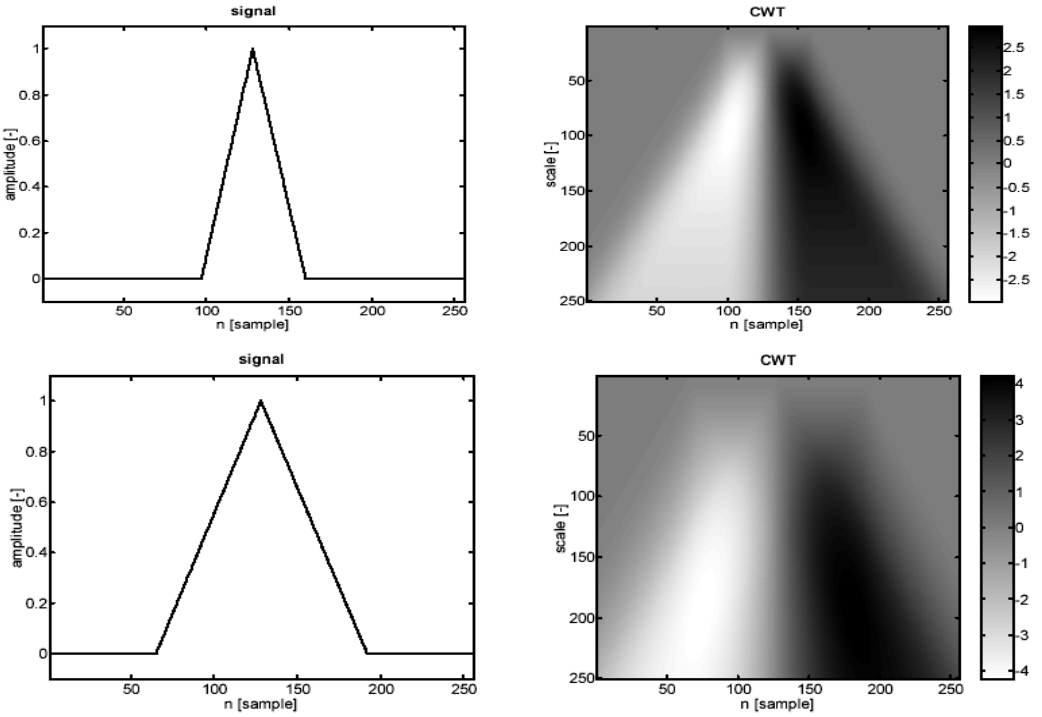


Figure 2.5. Signal *f(n)* (left top), its CWT scalogram using Haar wavelet (right top), signal *f(n)* scaled by a factor of 2 (left bottom) and its CWT using Haar wavelet (right bottom) (Provaznic, 2001).

2.4 Discrete Wavelet Transform

The basic tool needed for analyzing a signal using wavelets is the Discrete Wavelet Transform (DWT) that will be presented here. It plays a role analogous to that of the Discrete Fourier Transform in spectral analysis. DWT derived from the discretization of CWT parameters α and b.

CWT based analysis yields a potential wealth of information. When presented in a plot such as the bottom plot of Fig.2.3 or the right plots of Fig.2.5 is an excellent tool for exploratory data analysis because it helps picking out features of interest. Recall that the CWT leads a one-dimensional signal (from time domain) to a two-dimensional domain (scale, position), it is obvious that there is a lot of redundancy.

In general DWT are generated by sampling (in time-scale domain) a corresponding CWT. The CWT cannot be computed using finite precision discrete machines. However, approximation to the CWT can be made to almost arbitrary precision through dense sampling of the time-scale plane. Such a way the DWT is specified by selecting a time-scale sampling set and an analyzing wavelet. If the DWT must satisfy some properties these selection cannot be made arbitrarily. Discretization could be made according to the following formula:

$$W_{j,k}(t) = \frac{1}{\sqrt{a_j}} W\left(\frac{t - b_k}{a_j}\right) \tag{2.22}$$

where $a_j = a_0^j$ and $b_k = k b_0 a_0^j$ where $j, k \in \mathrm{Z}, a_0 > 1, b_0 \neq 0$. Here j controls the dilation (scale) and k controls the translations (shift). Although this approach could lead to infinity of transforms, the term DWT is commonly used with a particular choice of time-scale sampling: the dyadic sampling (Fig.2.6). In addition the analyzing wavelets are restricted to those that generate orthogonal or biorthogonal bases.

In dyadic lattice the discretization can be achieved by setting α_0=2 and b_0=1. In this case Equation 2.22 becomes:

$$W_{j,k}(t) = \frac{1}{\sqrt{2^j}} W\left(\frac{t - 2^j \cdot k}{2^j}\right) \tag{2.23}$$

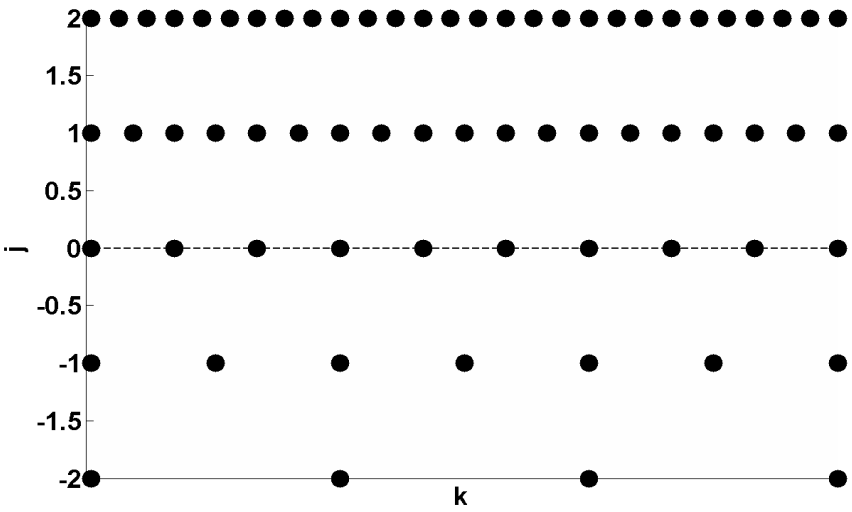


Figure 2.6. Dyadic discretization lattice of α and b for CWT

Wavelets for DWT are usually selected to be orthogonal and have unit energy. This means that decomposed signal has no redundancy unlike the CWT as shown in Fig.2.7.

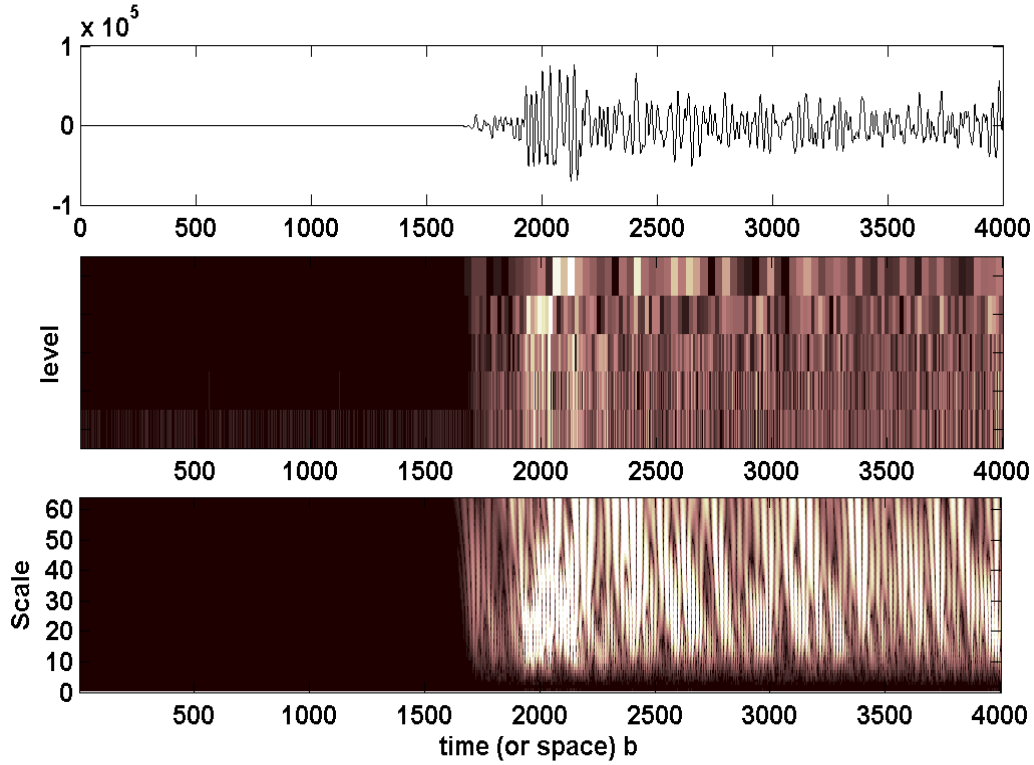


Figure 2.7. DWT scalogram (middle) and CWT scalogram (bottom) of the first secs of an earthquake recording (top).

2.5 Pyramidal Algorithm

Following the conclusions of 2.2, it is obvious that wavelet analysis initially lead to a decomposition of examined signal into approximations and details at multiple levels of resolution. In this section, the computation of approximation and detail wavelet coefficients will be discussed. This computation derived from MRA relationships through scaling and wavelet functions. In addition the decomposition (analysis) and reconstruction (synthesis) scheme using filters will also present.

Applying the relations of MRA to discrete case the discrete values of detail signal (coefficients through wavelet function ψ) at level j and time k can be calculated as:

$$d_{jk} = \int_{-\infty}^{\infty} f(t)\psi_{jk}(t)dt \tag{2.24}$$

Equation 2.24 can be seen as the inner product between f and ψ.

Similarly, the discrete values of approximated signal (coefficients through scaling function φ are:

$$\alpha_{jk} = \int_{-\infty}^{\infty} f(t)\phi_{jk}(t)dt \tag{2.25}$$

which also can be seen as the inner product between f and φ.

Equations 2.24 & 2.25 hold under general single scale decomposition. In order to calculate wavelet coefficients multiple scale decomposition relationships must de defined. For approximated signal the following relation holds (detailed proof can be found at Athichanagorn, 1999, pp25-29):

$$a_{j+1,k} = a_j * \bar{h}(2k) \tag{2.26}$$

where $\bar{h}$ sequence can be interpreted as the low-pass filter as described in Equation 2.11. Equation 2.26 shows that the approximated signal at level $j+1$ is computed by taking every second sample from the result of convolution between a_j and $\bar{h}$. This leads to an approximated signal at level $j+1$ which is subsampled by two.

Under similar approach, the detail signal can be expressed as:

$$d_{j+1,k} = d_j * \bar{g}(2k) \tag{2.27}$$

where $\bar{g}$ sequence can be interpreted as the high-pass filter as described in Equation 2.13. Respectively, detail signal at level $j+1$ can be computed by taking every second

from the result of convolution between d_j and $\bar{g}$ which means also that detail signal at level *j+1* is subsampled by two.

As the decomposition level increases the sample sizes of approximated and detail signals decrease (by a factor of two). This computation approach is called *pyramidal* algorithm or the *nonredundant* algorithm (since the data samples are different at each level). Figure 2.8 shows a graphic representation for 3 scales.

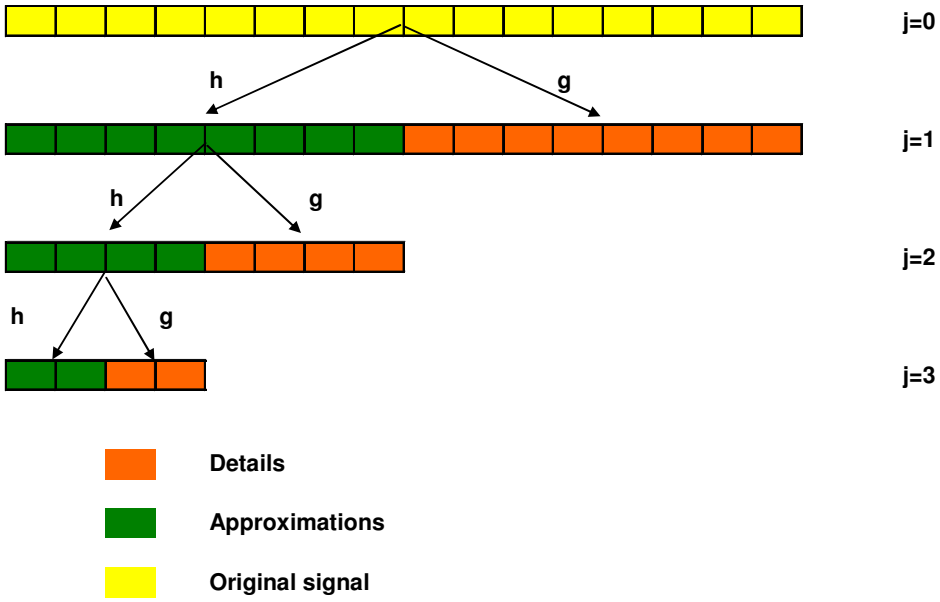


Figure 2.8. Wavelet decomposition using pyramidal algorithm

DWT is commonly implemented using dyadic multirate filter banks (consisting of $\bar{g}$ and $\bar{h}$ filters discussed earlier) which divide the signal frequency band into subbands as shown in Fig.2.9.

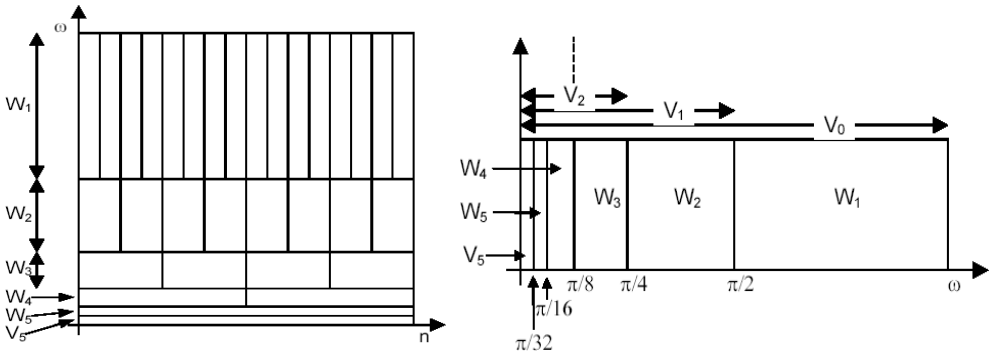


Figure 2.9. DWT for *j*=5 scales. Division of index-spectrum plane (left) and division of frequency spectrum using filter banks (right) (Provaznic, 2001).

At each scale, detail coefficients are generated from the output of high-pass filter since approximation coefficients are outputs from low-pass filters. In other words at each scale approximation captures low frequency trends since detail captures high frequency components. So the connection between wavelets and filter banks is that high-pass filter leads to wavelet function and low-pass to scaling function. A typical structure for signal decomposition is depicted in Fig.2.10 where a three scale DWT is performed using octave-band filter bank.

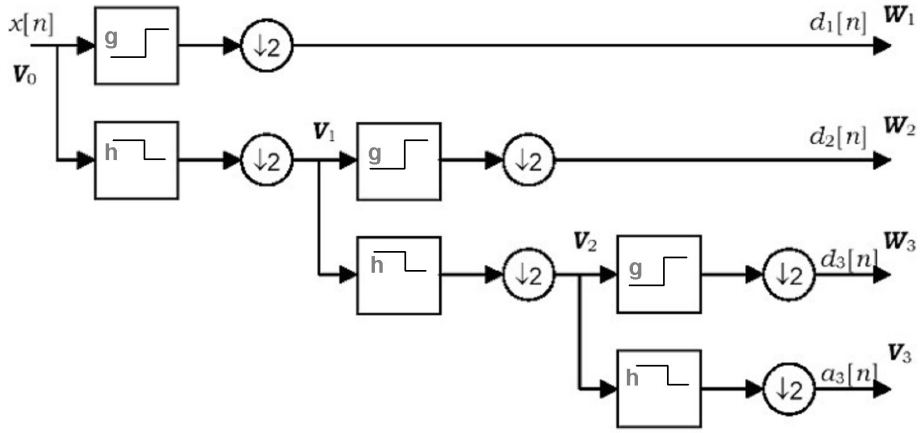


Figure 2.10. A three scale DWT decomposition

Inverse DWT (IDWT) reconstructs a signal from approximation and detail coefficients produced from decomposition. As a reverse procedure the steps that required are upsampling by 2 and filtering. Upsampling by 2 is achieving by inserting zeros between samples. For a multi-level reconstruction a relation for successive approximation must be defined. It is proved that the following relationship holds (Athichanagorn, 1999):

$$a_{j,k} = \breve{a}_{j+1} * h(k) + \breve{d}_{j+1} * g(k) \qquad (2.28)$$

Equation 2.28 shows that approximated signal $\alpha_{j,k}$ can be reconstructed by upsampling coefficients α_{j+1} and d_{j+1} , convolve them with filters h and g respectively and add these convolutions. Figure 2.11 depicts these steps where a filter bank representation of reconstruction is shown at Figure 2.12

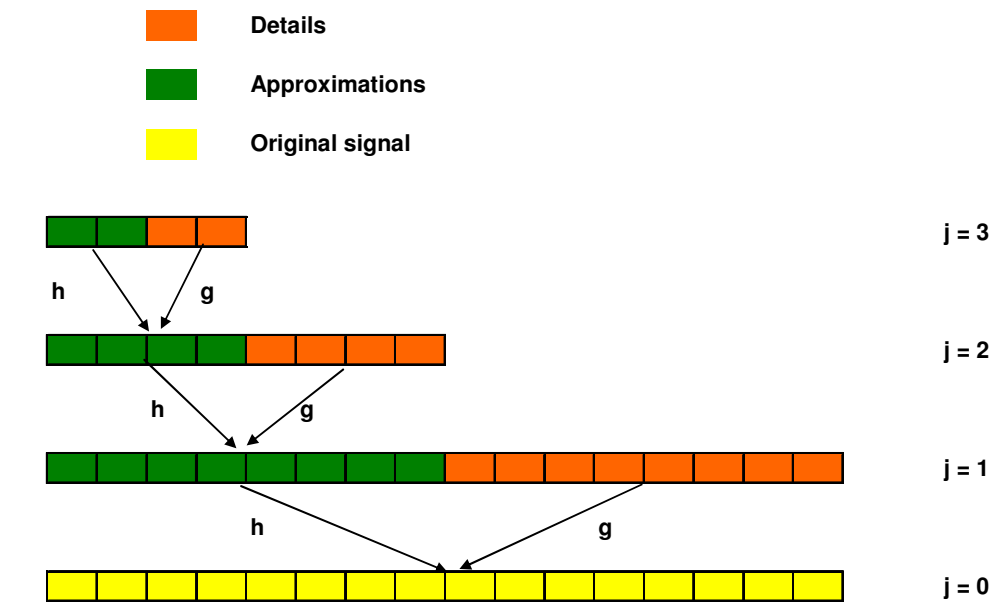


Figure 2.11. Wavelet reconstruction using pyramidal algorithm

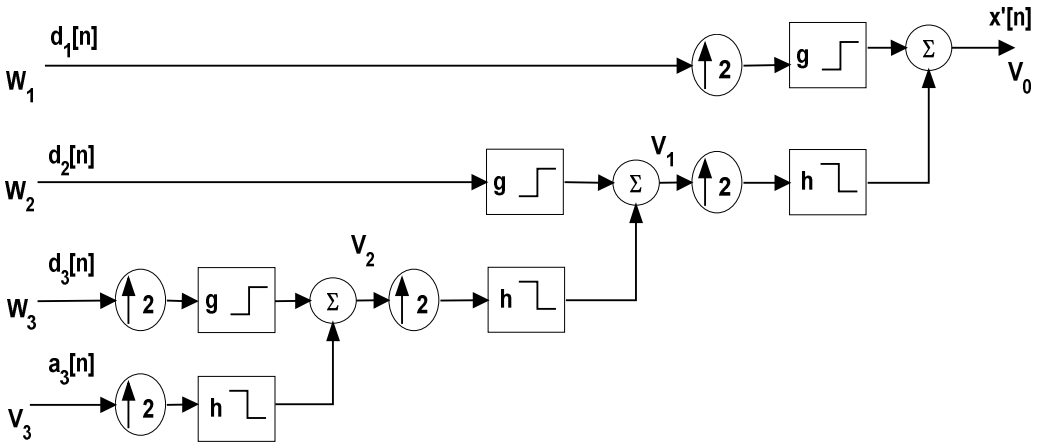


Figure 2.12. A three scale DWT reconstruction

The defining properties of a wavelet filter are not sufficient to yield a DWT whose coefficients can be interpreted as approximations and details at certain levels. Daubechies (1992) keeping in mind certain regularity conditions, defined a set of wavelet filters, all of which are able to localize differences of adjacent weighted averages. By definition a filter of even length *L* has a squared gain function given by:

$$H^{[D]}(f) \equiv D^{\frac{L}{2}}(f) A_L(f) \tag{2.29}$$

where A_L is the squared gain function for the difference filter {1,-1} and $D(f) \equiv 4sin^2(\pi f)$. It can be shown (Percival & Walden, 2000) that $H^{[D]}(f) = H^{[D]}\left(f + \frac{1}{2}\right) = 2$ for all f and any filter $\{h_l\}$ with squared gain function $H^{[D]}$ is a wavelet filter. Equation 2.29 implies that Daubechies wavelet filters can be interpreted as the complex filter consists of *L/2* filter (yielding the differencing operation) in series with a low-pass filter (yielding the averaging procedure).

Similarly, the corresponding scaling filter $\{g_l\}$ has a squared gain function:

$$G^{[D]}(f) \equiv H^{[D]}\left(\frac{1}{2} - f\right) \tag{2.30}$$

For L=2 the Haar wavelet filter $\left\{h_0 = \frac{1}{\sqrt{2}}, h_1 = -\frac{1}{\sqrt{2}}\right\}$ is constructed and for L=4 the Daubechies 4 {D(4)} filter $\left\{h_0 = \frac{1-\sqrt{3}}{4\sqrt{2}}, h_2 = -\frac{3-\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_4 = -\frac{1+\sqrt{3}}{4\sqrt{2}}\right\}$. Squared gain functions for D(4) are shown in Fig.2.13.

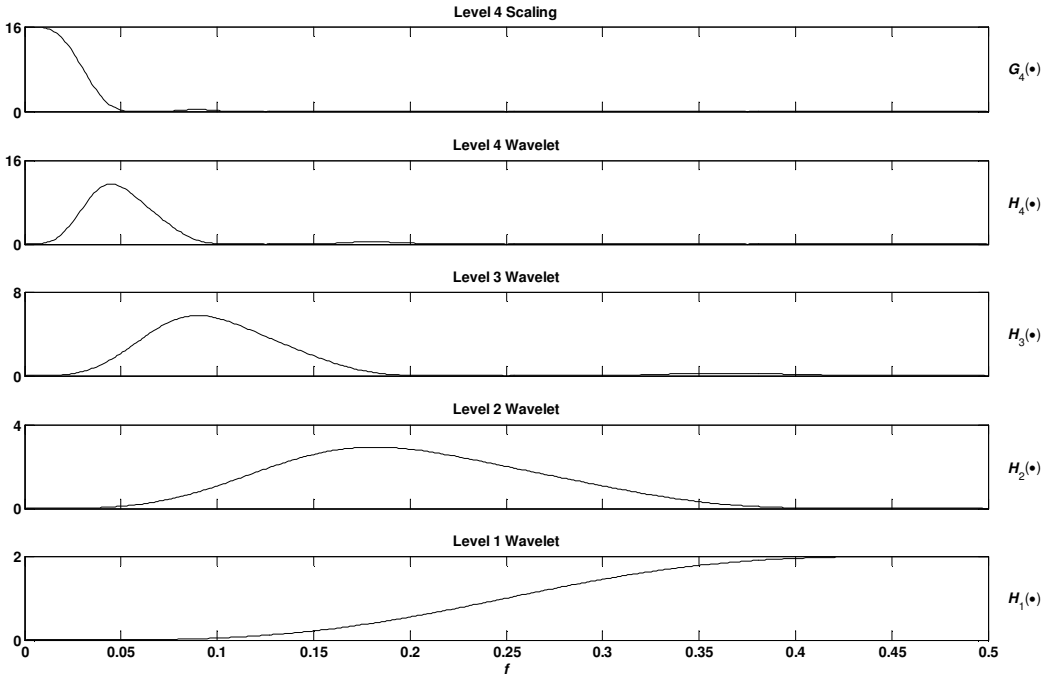


Figure 2.13. Squared gain functions for D(4) filters $\{h_{j,l}\}$for j=1…4 and $\{g_{4,l}\}$

In general, there are many real-valued wavelets based on type $\{h_0,h_1,\ldots.h_{L-1}\}$ with squared gain function $H^{[D]}$. As L increases the difficulty of selecting the appropriate pair of wavelet and scaling filters also increases. Daubechies (1992) defined two additional criteria in order to establish this selection: first, selected scaling filters must have minimum delay than other candidates and second the phase of the scaling filter's transfer function must be as close as zero (similar to linear phase filter's). The class of filters that satisfies the first criterion is the Daubechies (D) family of filters with L=2,4,6,…. since the second criterion is satisfied by the class of filters called *Least Assymetric* (LA) with L=6,8,10,…. A third candidate class of filters are the *Coiflets* (C) with lengths L=6,12,18,24 and 30 which are alternatives to Daubechies filters providing also phase closer to zero than LA. Their main disadvantage is that they provide fewer differencing operations for a given filter width (L/3 versus L/2 for D filters) (Percival & Walden, 2000). Fig.2.14 shows a 5-level D(4) filter since Fig.2.15 shows a LA(8) filter.

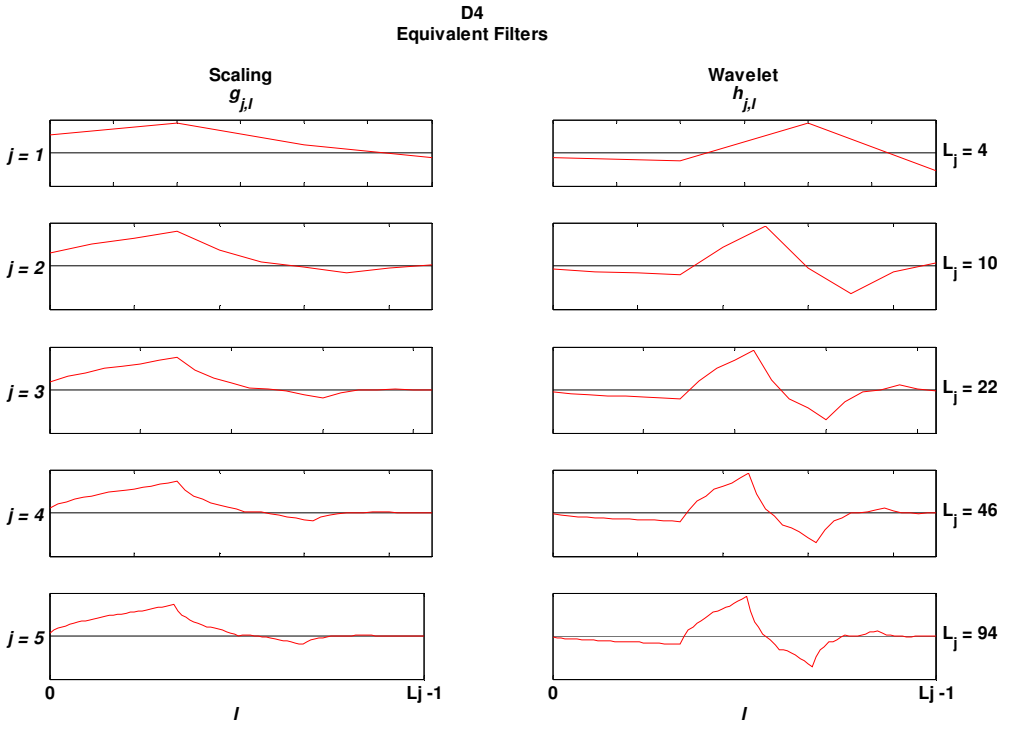


Figure 2.14. D(4) scaling $\{g_{j,l}\}$ and wavelet $\{h_{j,l}\}$ filters for j=1…,5 scales. L_j denotes the length of the filter

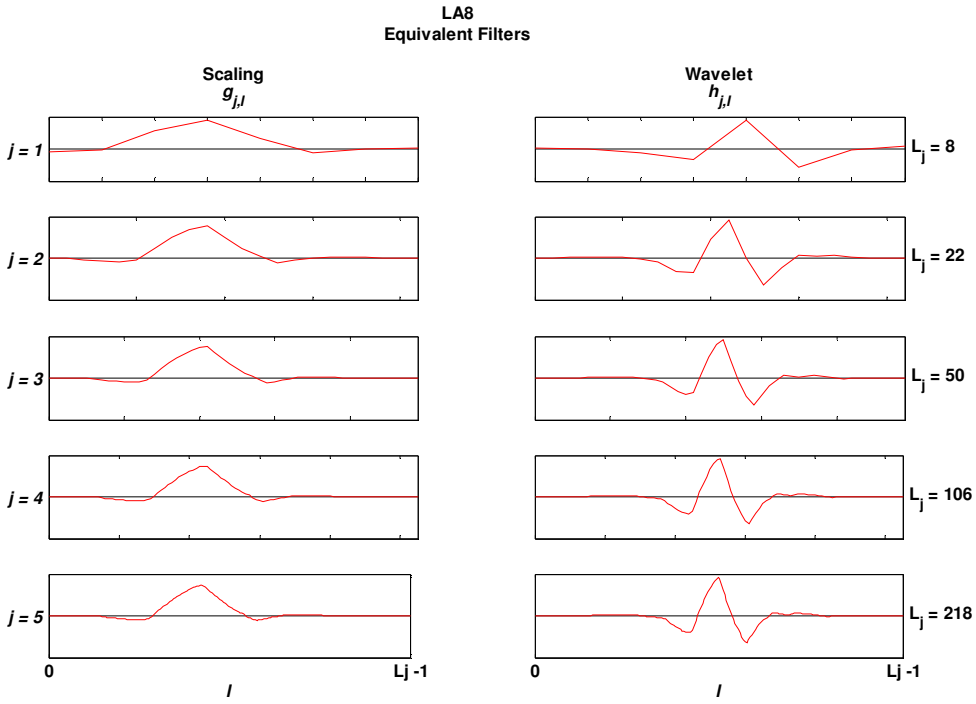


Figure 2.15. LA(8) scaling $\{g_{j,l}\}$ and wavelet $\{h_{j,l}\}$ filters for j=1…,5 scales. L_j denotes the length of the filter

In summary, *pyramidal* algorithm using a subsequent convolution of approximations and details with low and high-pass filters respectively, is able to decompose as well as to reconstruct a signal under the frame of wavelet analysis. The important point is that between successive levels the data samples are downsampled (at decomposition) or upsampled (at reconstruction) by a factor of two. From many possible wavelet filters Daubechies, Least Assymetric and Coiflets are those that are going to be used to next sections.

2.6 Practical considerations for DWT

When one begins a DWT analysis several issues arised since there is no "blind rule" that must be followed in order to get the results. Depending on the application and the nature of examined signal some consideration must be addressed. In this section four of them will be discussed: Boundary conditions, the number of levels, choice of wavelet filter and sample size that is not power of two.

2.6.1 Boundary conditions

DWT (full or partial) implemented by using filters. A filtering operation close the boundaries (the beginning and the end) of the time series $\{X_n: n=0,1,\ldots N-1\}$

assumes that is a periodic sequence with period *N*. This means that is expected for X_{N-1}, X_{N-2}.… to be successful surrogates for unsampled values of process X_{-1}, X_{-2} .. that described by *X* as depicted in Figure 2.16. This assumption holds only when the sample size is chosen appropriately otherwise is problematic especially when there is discontinuity between X_{N-1} and X_0. For example, a time series that presents annual, seasonal and monthly variations would not affected by boundary conditions only if the sample size is equal to an integer multiple of a year.

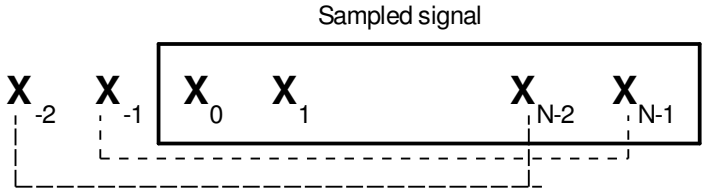


Figure 2.16. Corresponding surrogates for unobserved (X_{-1}, X_{-2}) and sampled (X_0,.…X_{N-1}, X_{N-2}) signal

Percival and Walden (2000) calculated the number of boundary coefficients L_j' (coefficients that are influenced by boundary conditions) for a given filter length for several scales. Table 2.1 list the results which hold for wavelet as well as scaling filters (based on assumption that $L_j' \le N_j$ where $N_j = N/2^j$ is the length of approximation or details at level *j*.

L	L_1'	L_2'	L_3'	L_4'	$L_{j\ge 5}'$
2	0	0	0	0	0
4	1	2	2	2	2
6	2	3	4	4	4
8	3	5	6	6	6
10	4	6	7	8	8
12	5	8	9	10	10
14	6	9	11	12	12
16	7	11	13	14	14
18	8	12	14	15	16
20	9	14	16	17	18

Table 2.1. Number of boundary coefficients based on wavelet filter with length *L*. The calculated number refers to affected coefficients for both sides of time series (Percival and Walden, 2000)

The main result that derived from Table 2.1 is that L_j' is increased as scale j increases and does not depend on N. So, as the level of decomposition becomes bigger the influence of boundary conditions affects more and more the details and approximations. This fact must always keep in mind before the selection of a wavelet filter. An important fact also that Haar wavelet (L=2) does not suffer from any kind of boundary conditions. A relation that denotes the results of Table 2.1 is:

$$\frac{L}{2}-1 \le L_j' \le L-2 \qquad (2.31)$$

An effective technique to reduce boundary effects is to replace the time series X with a time series with length 2N consists of X along with the time reversed version of X as shown:

$$X^{[2N]} = X_0, X_1, X_2, ..., X_{N-2}, X_{N-1}, X_{N-1}, X_{N-2}, ..., X_2, X_1, X_0 \qquad (2.32)$$

Using $X^{[2N]}$ constructed time series instead of X has the effect that the surrogates of unobserved signal's samples X_{-1}, X_{-2} .. are now the X_1, X_2,… of sampled signal, which is obvious that they present a lower mismatch possibility than before. A graphic representation of this fact is depicted in Figure 2.17.

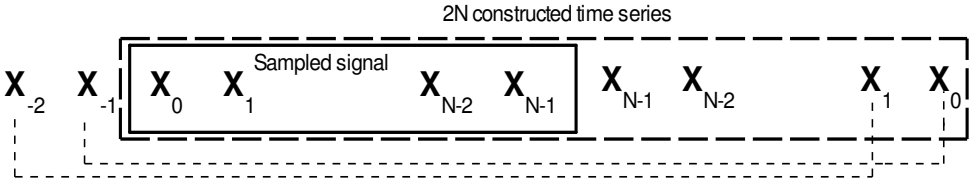


Figure 2.17. Corresponding surrogates for unobserved (X_{-1}, X_{-2}) and 2N constructed time series (X_0,.…X_{N-1}, X_{N-2}, X_{N-1}….. X_0).

When DWT applied to X time series then it yields to decomposition using "*circular boundary conditions*" while applied to $X^{[2N]}$ is a decomposition using "*reflection boundary conditions"*.

2.6.2 Number of levels

For a time series $\{X_n: n=0,1,\ldots N-1\}$ with sample size $N=2^j$ the pyramid algorithm for DWT will completed after j levels. Depending on the application, there is also a great possibility that the algorithm stop at level $J_0 < j$. In this case the DWT is called *partial DWT at level j.* Partial DWT are commonly used in practice because they provide the flexibility to specify a scale J_0 beyond which wavelet analysis of larger scales has no real interest.

Recall that W_j {j=1,2,….J_0} is the vector space of detail coefficients at scale j and V_{J_0} is the vector space of approximation, X_n will be decomposed as:

$$X = \sum_{j=1}^{J_0} W_j + V_{J_0} \tag{2.32}$$

V_{J_0} will contain $N/2^{J_0}$ coefficients of X which represents (as approximations) averages over the scale τ_{J_0}.

A reasonable selection of level J_0 is highly dependant from the application. Main idea under successful picking of J_0 is to set it in a way that scaling coefficients could have a meaningful association with physical scales. Setting J_0 too high will lead to large scales that will not provide useful information. On the other hand, setting J_0 too low will produce approximations that could not promote interesting features.

Another factor that will impact the choice of J_0 is the width L of the wavelet filter. This is because large L produces large filter widths for the higher level of $\bar{h}$ and $\bar{g}$ filters. For this reason an upper bound is $L_{J_0} \le N \le L_{J_0+1}$. This upper bound also ensures that details and approximation will decrease their influence by boundary conditions.

2.6.3 Choice of wavelet filter

As already mentioned the choice of wavelet filter is not a direct procedure and is always dependable from the application in use. A main factor for the selection of the wavelet filter is to define the aim of the analysis and then choose the appropriate wavelet filter. Although there are many studies using wavelets there are only few studies that provide step-by-step procedures (Torrence & Compo, 1999, Breiman & Peters, 2000).

In practice, two main considerations must be taken into account. First, there are many cases where wavelet filters with short widths (L=2,4 or 6) introduce artifacts in results producing unrealistic coefficients (a typical example is presented at section 2.11). On the other hand, wavelet filters with $L > 6$ has two main disadvantages:

- As L increases the number of coefficients that influenced from boundary conditions also increase
- Computational requirement increase also

The objective is the choice of the wavelet filter with smallest L that produces the most reasonable results. In other words there must be a pre-analysis procedure where the selection begins with quite small L and increases until an artifact free coefficient set is produced.

2.6.4 Handling non 2^J sample sizes

As already discussed DWT (full or partial) designed to work with sample sizes of specific length. More specific, full DWT needs sample size N to be a power of two $\{N=2^J\}$. Similarly, partial DWT of level J_0, needs sample size to be an integer multiple to 2^{J_0}. Of course these requirements usually not hold in practice. In order to fulfill the above requirements two solutions were proposed and presented below.

First solution based on padding which is a usual technique at FFT algorithms. The idea is to create a new time series, say X'_n, from original one X_n with new length $N' > N$ and then take the DWT of the X'_n. Padded values are equal to the sample mean $\overline{X}$ of X_n.

Second solution based on truncation of X_n to shorter series whose lengths are a multiple of 2^{J_0}. Let $N'' < N$ *an* integer multiple of 2^{J_0} then two subseries of X_n can be defined: $X^{[1]} \equiv [X_0,....X_{N''-1}]$ and $X^{[2]} \equiv [X_{N-N''},....X_{N''-1}]$ both of length N'', It is proved (Coifman & Donoho,1995) that this scheme leads to scale based decomposition of weighted sum of squares with half weights being attached to the $N - N''$ samples at the beginning and end of X_n.

2.7 Maximal Overlap DWT

In this chapter a modified version of DWT will be discussed. This transform is called *maximum overlap DWT* (MODWT) and focus on the elimination of DWT shortcomings that discuss earlier. MODWT belongs to the group of redundant and non-orthogonal wavelet transforms that met in the literature under names "undecimated DWT" (Shensea, 1992), "shift invariant DWT" (Beylkin,1992), "translation invariant DWT" (Coifman & Donoho, 1995), "Stationary DWT" (Nason & Silverman,1995) and "non-decimated DWT" (Bruce & Gao, 1996). The main differences of MODWT from full or partial DWT will be discussed also.

The MODWT differs from the DWT in that it is a highly redundant, nonorthogonal transform (Percival and Walden, 2000). The MODWT retains

downsampled values at each level of the decomposition that would be otherwise discarded by the DWT. The MODWT is well-defined for all sample sizes N, whereas for a complete decomposition of J levels the DWT requires N to be a multiple of 2^J . The MODWT offers several advantages over the DWT. The redundancy of the MODWT facilitates alignment of the decomposed wavelet and scaling coefficients at each level with the original time series, thus enabling a ready comparison between the series and its decomposition. Coefficients derived using the MODWT are not influenced by circular shifting of the input time series, whereas values derived using the DWT depend upon the starting point of the series. Finally, the redundancy of the MODWT wavelet coefficients modestly increases the effective degrees of freedom (EDOF) on each scale and thus decreases the variance of certain wavelet-based statistical estimates. Since the MODWT is energy conserving, it is well suited for analyzing the scale dependence of variability.

Decomposing X_n using MODWT to J_0 levels theoretically involves the application of J_0 pairs of filters. The filtering operation at the j_{th} level consists of applying a wavelet (high-pass) filter $\{\tilde{g}_{j,k}\}$to yield a set of wavelet coefficients

$$\overline{W}_{j,n} = \sum_{k=0}^{L_j-1} \tilde{g}_{j,k} X_{n-k} \tag{2.33}$$

and a scaling (low-pass)filter $\{\tilde{h}_{j,k}\}$ to yield a set of scaling coefficients

$$\overline{V}_{j,n} = \sum_{k=0}^{L_j-1} \tilde{h}_{j,k} X_{n-k} \tag{2.34}$$

The equivalent wavelet $\{\tilde{h}_{j,k}\}$ and scaling $\{\tilde{g}_{j,k}\}$ filters for the j_{th} level are a set of scale-dependent localized differencing and averaging operators, respectively, and can be regarded as stretched versions of the base (j = 1) filters. The j_{th} level equivalent filter coefficients have a width $L_j=(2^j-1)(L-1)+1$, where L is the width of the j = 1 base filter. In practice, the filters for j >1 are not explicitly created because the wavelet and scaling coefficients can be generated sequentially using an elegant algorithm (which will discussed later) that involves just the j = 1 filters operating on the j_{th} level scaling coefficients to generate the j+1 level wavelet and scaling coefficients (Percival and Walden, 2000). The j_{th} level wavelet coefficients characterize those components of the signal with fluctuations matching the unitless scale 2^{j-1}. In addition, MODWT

coefficients for different scales are approximately uncorrelated and are hence useful statistical measures for partitioning variability by scale.

2.8 *A trous* algorithm

In this section, we discuss a redundant, shift-invariant algorithm that computes the approximated and detail signals at the same spacing across scales using nonorthonormal wavelet bases. This algorithm is called the *a trous algorithm* and it was first proposed by Holschneider et.al (1989). Shensa (1992) later studied the relationship between the *a trous* algorithm and the pyramidal algorithm and concluded that the *a trous* algorithm is in fact a nonorthonormal multiresolution algorithm for which the discrete wavelet transform is exact.

In order to keep the original sampling frequency (in which $2^j = 2^0 = 1$), the wavelet function is now discretized as:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2}}\psi\left(\frac{t-k}{2^j}\right) \tag{2.35}$$

and the scaling function as:

$$\phi_{j,k}(t) = \frac{1}{\sqrt{2}}\phi\left(\frac{t-k}{2^j}\right) \tag{2.36}$$

Mallat (1998) further showed that the approximated signal $a_{j+1,k}$ and detail signal $d_{j+1,k}$ can be computed using these multiscale relationships through low-pass filter $\bar{h}$ and high-pass filter $\bar{g}$. Obviously, the filters $\bar{h}$ and $\bar{g}$ are different from $\bar{h}$ and $\bar{g}$ defined in the section 2.5 in which the wavelets are orthogonal. For any filter sequence $x(n)$, $x_j(n)$ is denoted as the filter sequence obtained by inserting 2^{j-1} zeros between each sample of $x(n)$. The insertion of zeros into the sequence creates holes, thus the algorithm is called the *a trous* algorithm (*trous* means holes in French). The approximated and detail signals are respectively determined by

$$\alpha_{j+1,k} = a_j * \overline{h_j}(k) \tag{2.37}$$

and

$$d_{j+1,k} = a_j * \overline{g_j}(k) \tag{2.38}$$

Therefore, the approximated signal at level j+1 can be calculated by inserting 2^{j-1} zeros between each sample of the low-pass filter $\bar{h}$ and performing the discrete convolution between a_j and the resulting filter. The detail signal at level j +1 can be

computed in a similar fashion using the high-pass filter $\bar{g}$. A diagram of the *a`trous* algorithm for wavelet decomposition is depicted in Fig.2.18. The size and spacing of the original data are preserved across scales while signals at higher levels of decomposition provide information with coarser contents.

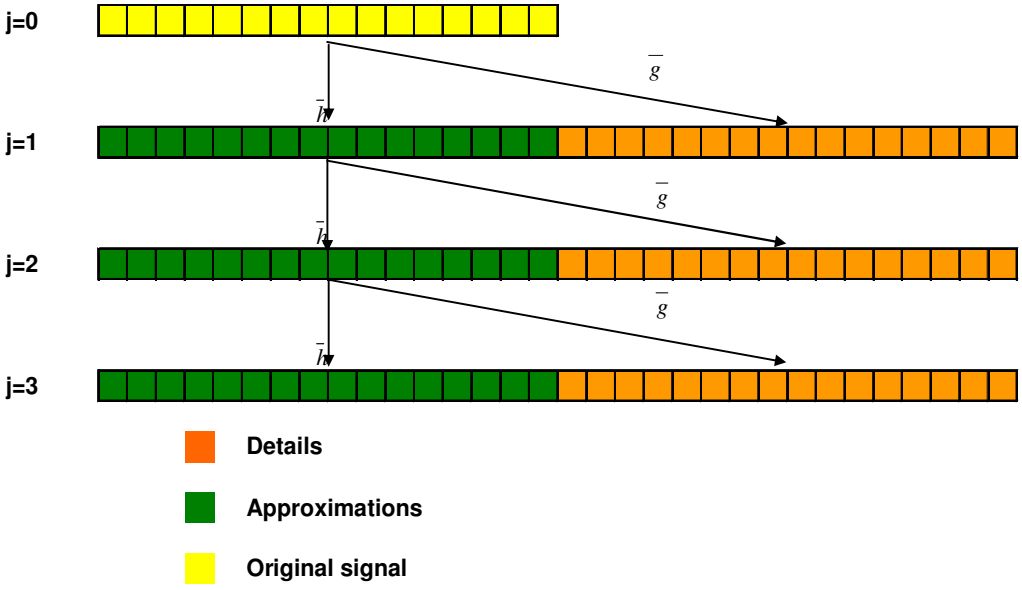


Figure 2.18. Wavelet decomposition using *a'trous* algorithm

In order to reconstruct signal at lower levels of decomposition, Mallat (1998) defined the *dual scaling function* $\tilde{\phi}(t)$ and *dual wavelet* $\tilde{\psi}(t)$ using frame theory introduced by Duffin & Schaeffer (1952). For an orthonormal wavelet basis, the basis and its dual are the same. The multiscale relationships for the dual scaling function $\tilde{\phi}(t)$ and the dual wavelet $\tilde{\psi}(t)$ exist in the Fourier domain as:

$$\hat{\tilde{\phi}} = \frac{1}{\sqrt{2}} \hat{\tilde{h}}\left(\frac{\omega}{2}\right) \hat{\tilde{\phi}}\left(\frac{\omega}{2}\right) \tag{2.39}$$

and

$$\hat{\tilde{\psi}} = \frac{1}{\sqrt{2}} \hat{\tilde{g}}\left(\frac{\omega}{2}\right) \hat{\tilde{\psi}}\left(\frac{\omega}{2}\right) \tag{2.40}$$

where $\tilde{h}$ is called the *low−pass dual filter*, and $\tilde{g}$ is called the *high−pass dual filter*. The reconstructed signal $a_{j,k}$ is then computed as:

$$a_{j,k} = \frac{1}{2}\left(a_{j+1} * \tilde{h}_j(k) + d_{j+1} * \tilde{g}_j(k)\right) \qquad (2.41)$$

Using Equation 2.41, the approximated signal $a_{j,k}$ can be reconstructed by first inserting 2^{j-1} zeros between each sample of the dual filters $\tilde{h}$ and $\tilde{g}$ and then adding together the discrete convolution between a_{j+1} and the resulting low-pass dual filter and the convolution between d_{j+1} and the resulting high-pass dual filter. Figure 2.19 depicts a schematic diagram of signal reconstruction using the `*a trous* algorithm. Starting from any level of decomposition, approximated signals at subsequently lower levels of decomposition and the original signal can be reconstructed over the same grid.

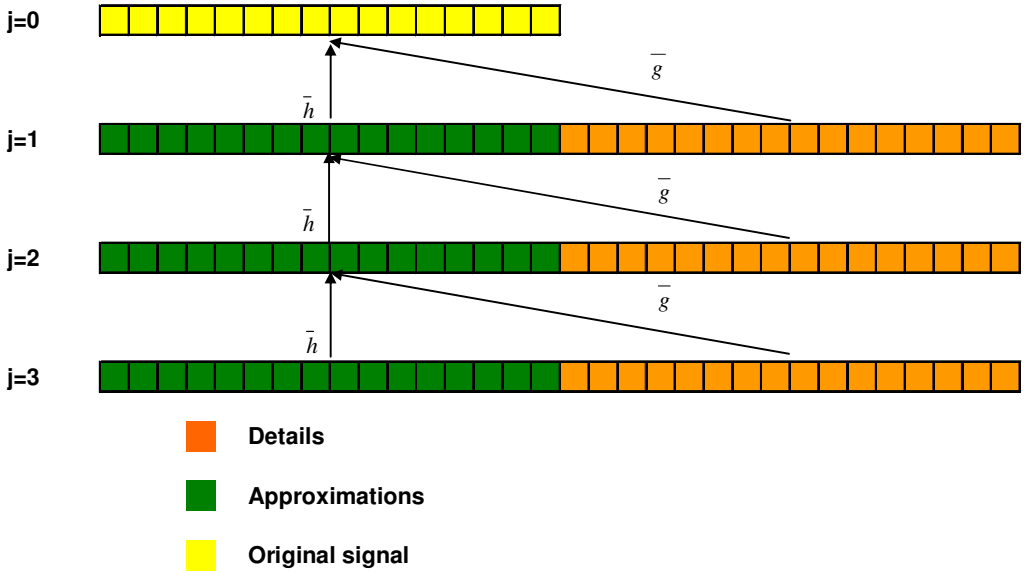


Figure 2.19. Wavelet reconstruction using *a'trous* algorithm

2.9 Practical considerations for MODWT

Following the same procedure as in 2.6 practical considerations about MODWT will be discussed. It must be noted that many of the subjects that presented for DWT also holds for MODWT. However, since there are significant differences between them and the discussed subjects will focus on these differences.

The first issue that does not need special consideration is the handling of samples sizes that are not power of two. As already discussed in Section 2.7 MODWT is defined for all sample sizes so (unlike the DWT) there is no need for special

adaptation to handle certain sample sizes. This is one of the main differences between DWT and MODWT

The second issue is about boundary conditions. Since MODWT uses circular filters for its implementation, results from Section 2.7.1 will also hold here. Percival & Walden (2000) showed that boundary regions of MODWT are larger than DWT. This is not a serious shortcoming of MODWT since, on one hand, the difference is quite small and on the other hand the advantages from using MODWT against DWT are more important.

Therefore, the issues that left for discussion is the choice of wavelet filter and the choice of level J_0

2.9.1 Choice of wavelet filter

The Daubechies class of wavelets possesses appealing regularity characteristics and produces transforms that are effectively localized differences of adjacent weighted averages (Daubechies, 1992). The Least Symmetric subclass has approximate linear phase and exhibits near symmetry about the filter midpoint. This linear phase property means that events and sinusoidal components in the wavelet and scaling coefficients at all levels can be aligned with the original time series. For the MODWT, this alignment is achieved by circularly shifting the coefficients by an amount dictated by the phase delay properties of the basic filter. LA filters are available in even widths *L*. The optimal filter width is dependent on the characteristics of the signal and problem domain of interest. A wider filter is smoother in appearance and reduces the possible appearance of artifacts in the calculated coefficients due to the filter shape. It also results in better uncorrelatedness between wavelet coefficients across scales for certain time series, which is useful for deriving confidence bounds from certain wavelet-based estimates. However, using a wider filter results in many more boundary coefficients, especially at higher levels. LA filters yield coefficients that are approximately uncorrelated between scales while having a filter width short enough such that the impact of boundary conditions is tolerable.

Compared to DWT, MODWT is less dependant upon choice of wavelet filter but not so much so that a particular filter proposed for every wavelet analysis. A preliminary study of the investigated features of examined signal is still needed.

2.9.2 Choice of level J_0

A time series can be completely or partially decomposed into a number of levels. For complete decomposition of a series of length $N = 2^J$ using the DWT, the maximum number of levels in the decomposition is J. In practice, a partial decomposition of level $J_0 < J$ suffices for many applications. A J_0 level DWT decomposition requires that N be an integral multiple of $2^J{}_0$. The MODWT can accommodate any sample size N and, in theory, any J_0. In practice, the largest level is commonly selected such that J_0 preclude decomposition at levels heavily influenced by boundary conditions.

In particular, for alignment of wavelet coefficients with the original series, the condition L_{J_0}, i.e. the width of the equivalent filter at the J_0-th level is less than the sample size, should be satisfied to prevent multiple wrappings of the time series at level J_0. Selection of J_0 determines the number of octave bands and thus the number of scales of resolution in the decomposition.

If, as in DWT, there is a need that at least one MODWT coefficient is not a boundary coefficient, then the following condition must apply

$$J_0 < \log_2\left(\frac{N}{L-1}+1\right) \tag{2.42}$$

if Equation 2.42 seems too tight, alternative upper bounds are

$$J_0 < \log_2(N) \tag{2.43}$$

and in rare cases

$$J_0 < \log_2(1.5N) \tag{2.44}$$

Of course, the above three upper bounds are suggested only as a guide; the final choice of J_0 is application-dependant.

2.10 Field example

In this section wavelet decompositions will be presented using both DWT and MODWT in order to derive their already discussed differences. An example from field measurement will used. More specific a wavelet transform (partial DWT and MODWT) will take place. The examined signal is the vertical component of a seismogram as depicted in Fig. 2.20.

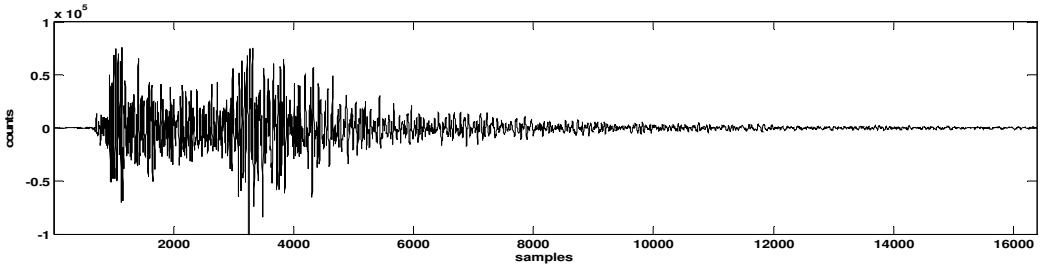


Figure 2.20. Vertical component seismogram. Y-axis refers to output counts from a Δ-Σ DAC; X-axis refers to samples

There are N=16384 samples measured in units of counts (which can easily be converted in Volts by multiply them with corresponding nominal bit weight) and collected at a rate of 125 samples per second, so the sampling rate is Δt=0.008 seconds. Fig.2.21 through 2.24 show the partial DWT coefficients for J_0=5 using Haar, D(4), C(6) and LA(8) filters.

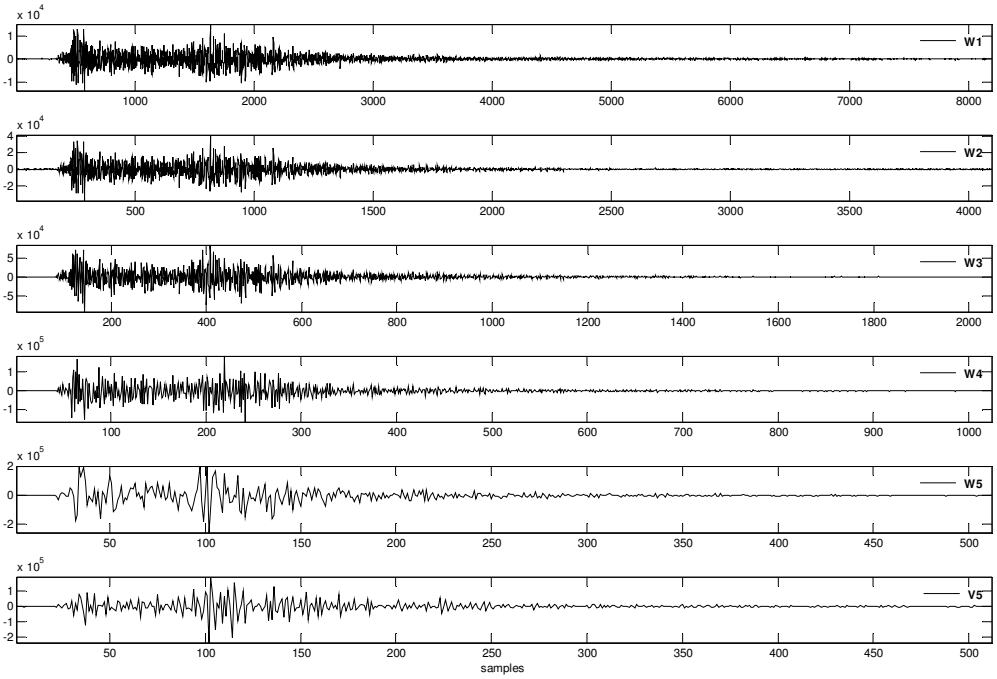


Figure 2.21. Partial DWT coefficients of level J_0=5 using Haar filter.

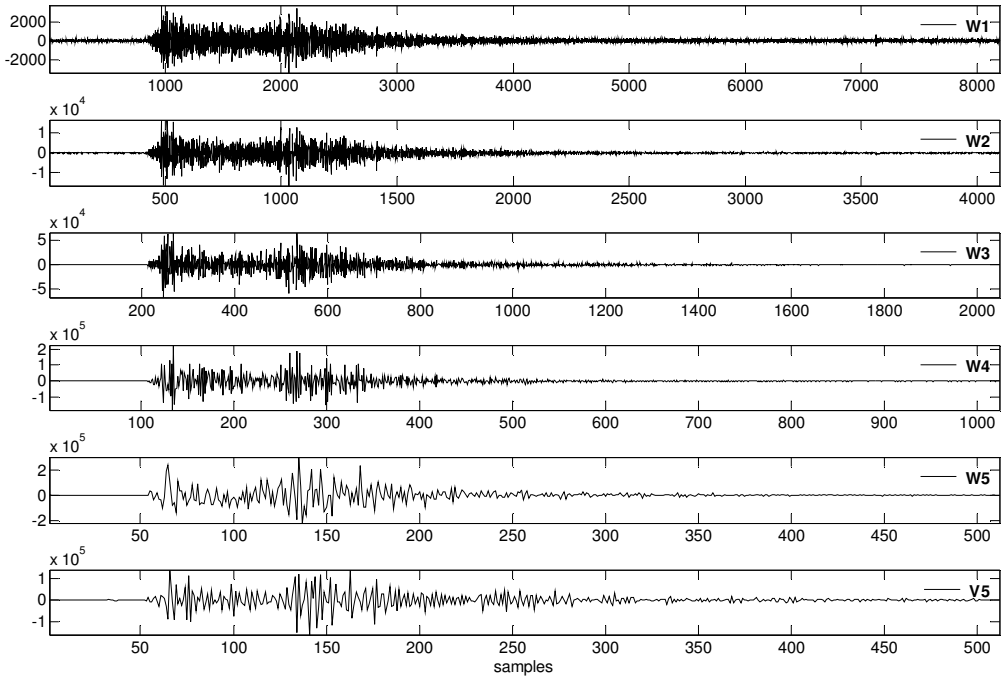


Figure 2.22. Partial DWT coefficients of level J_0=5 using D(4) filter.

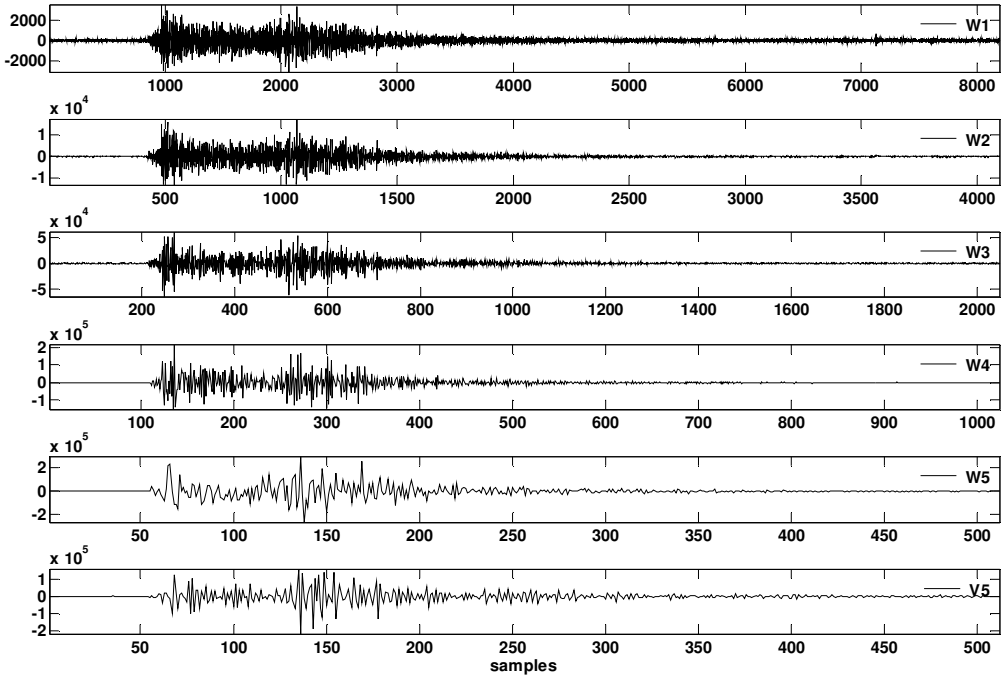


Figure 2.23. Partial DWT coefficients of level J_0=5 using C(6) filter.

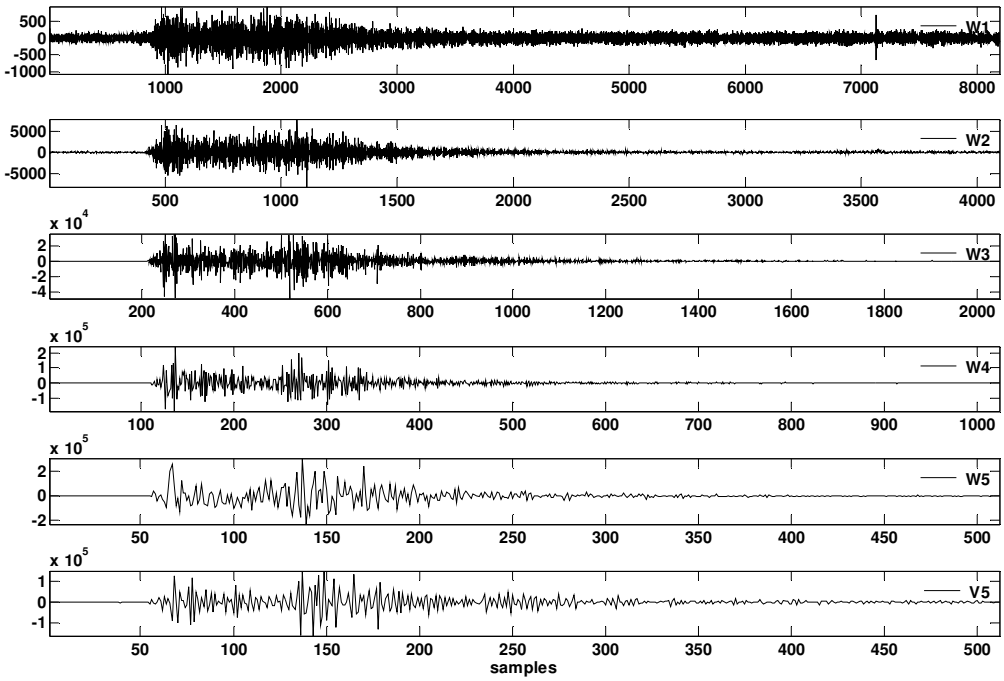


Figure 2.24. Partial DWT coefficients of level J_0=5 using LA(8) filter.

The number of coefficients in each subvector W_j is $N/2^j$ therefore W_1 has 8192, W_2 has 4096, W_3 has 2048, W_4 has 1024, W_5 has 512 and V_5 has 512 also. Number of scales set to 5 because over this number there is no physical meaning for produced coefficients. The coefficients are related with averages over scales of 0.016 secs for W_1 until 0.2560 secs for W_5. In frequency terms, coefficients are related over scales of 62.5Hz for W_1 until 3.9Hz for W_2.

In order to compare the previous results WT reapplied with the same parameters (J_0, filters) but using MODWT. Results are presented in Fig.2.25 through 2.28. For each level, vertical lines add in order to delineate the region outside of which boundary conditions have great influence. Coefficients also are plotted after proper circular shifting in order to be aligned with original time series. This shift is denoted by vector $T^{-s_j}\tilde{W}_j$ where s_j declares the number of samples that corresponding coefficient shifted and defined as (Cornish, 2004):

$$s_j^{[H]} = \frac{-L_j}{2} \tag{2.45}$$

for wavelet coefficients

$$s_j^{[G]} = \frac{(L_j - 1)(L-2)}{2(L-1)} \tag{2.46}$$

for scaling coefficients.

Above Equations hold when L/2 is even. Corresponding equations for odd L/2 can be found in Percival & Walden (2000).

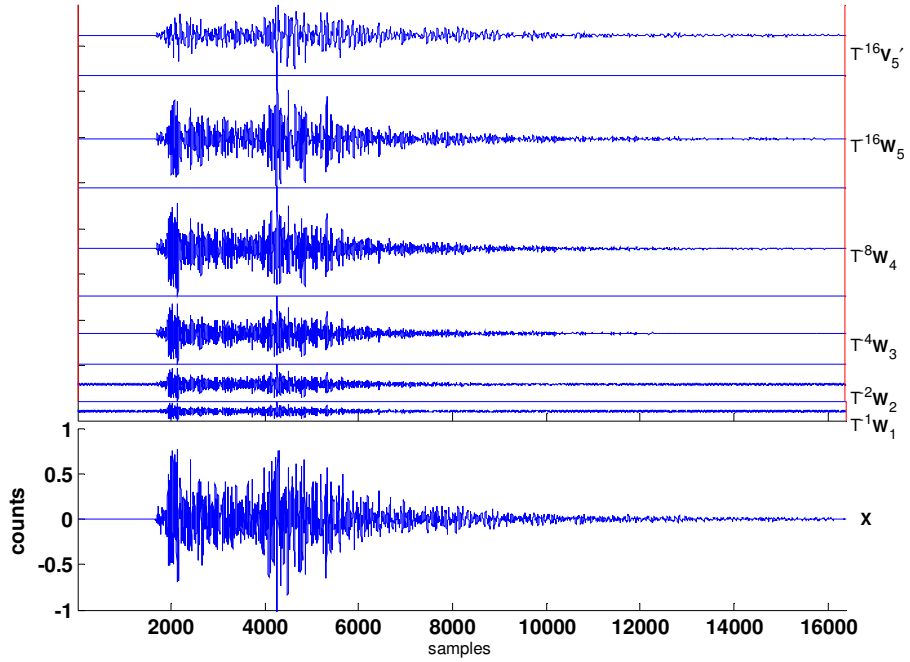


Figure 2.25. MODWT coefficients of level J_0=5 using Haar filter.

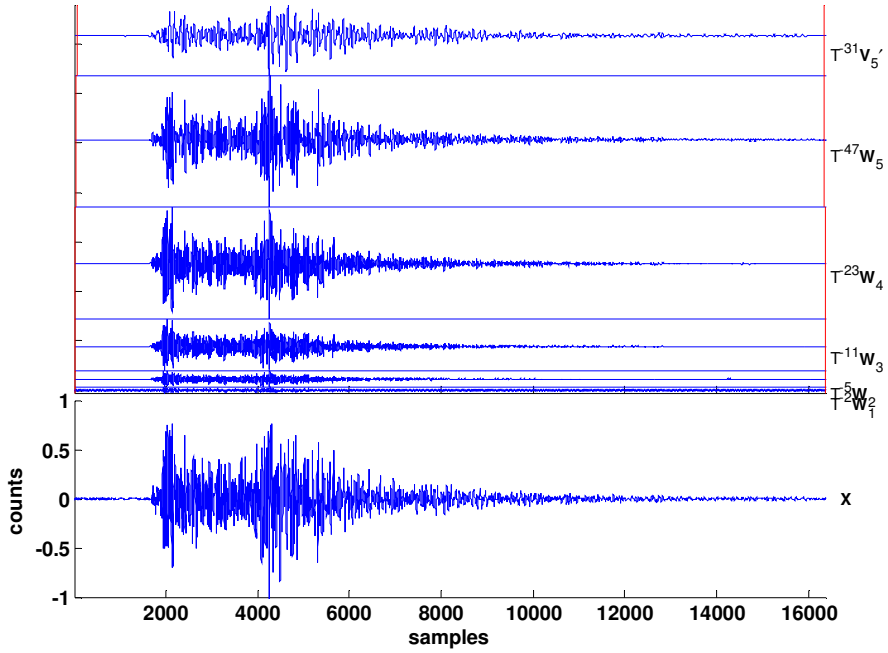


Figure 2.26. MODWT coefficients of level J_0=5 using D(4) filter.

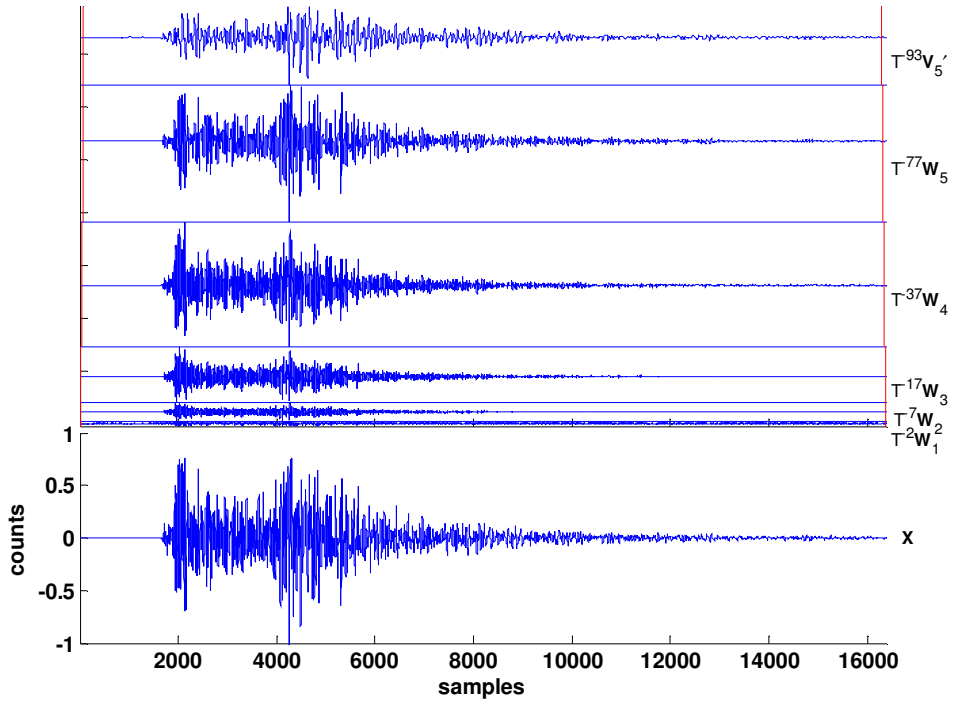


Figure 2.27. MODWT coefficients of level J_0=5 using C(6) filter.

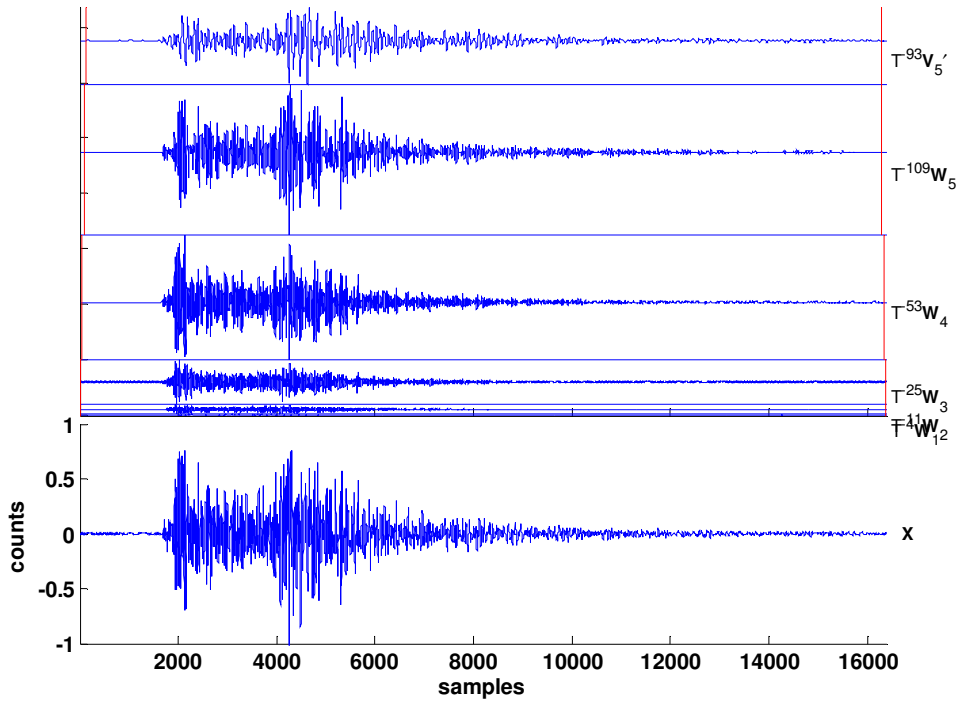


Figure 2.28. MODWT coefficients of level J_0=5 using LA(8) filter.

At the forthcoming sections, the majority of MODWT will based on Daubechies and Least Asymmetric filters. As already mentioned the choice of a wavelet filter is application-dependent. The above two classes of filters seems to be the best candidates for studied applications and this fact will be revealed concisely from a detailed examination at Fig.2.25 until Fig.2.28. An 8 sec (*N=700)* portion (focus on the beginning of the seismic event) is presented in Fig.2.29 until Fig.2.32 for the four already mentioned filters {Haar, D(4), C(6), LA(8)}

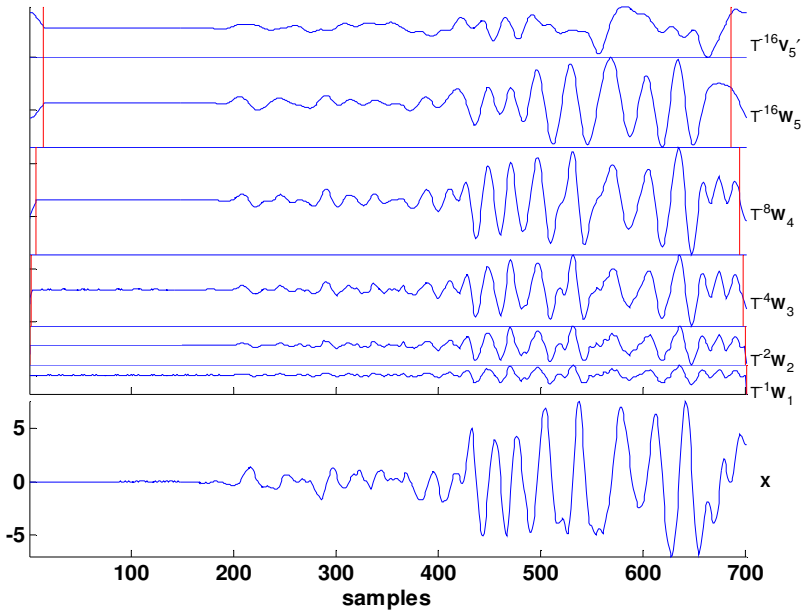


Figure 2.29.Expanded view of Figure 2.25

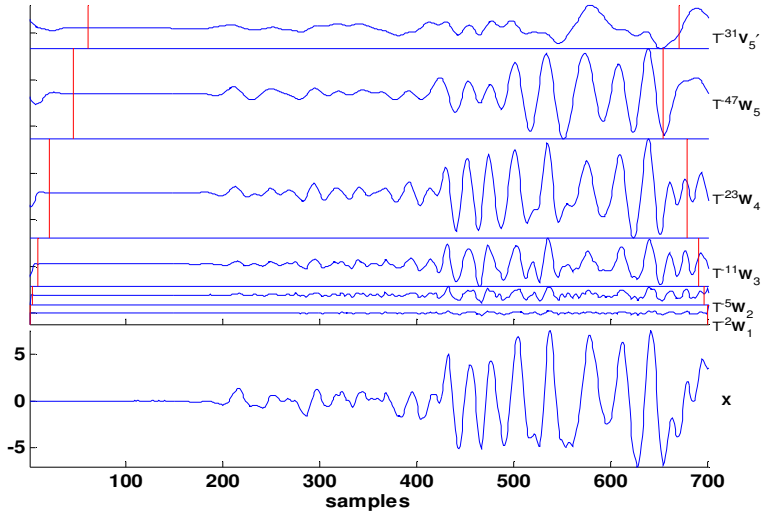


Figure 2.30.Expanded view of Figure 2.26

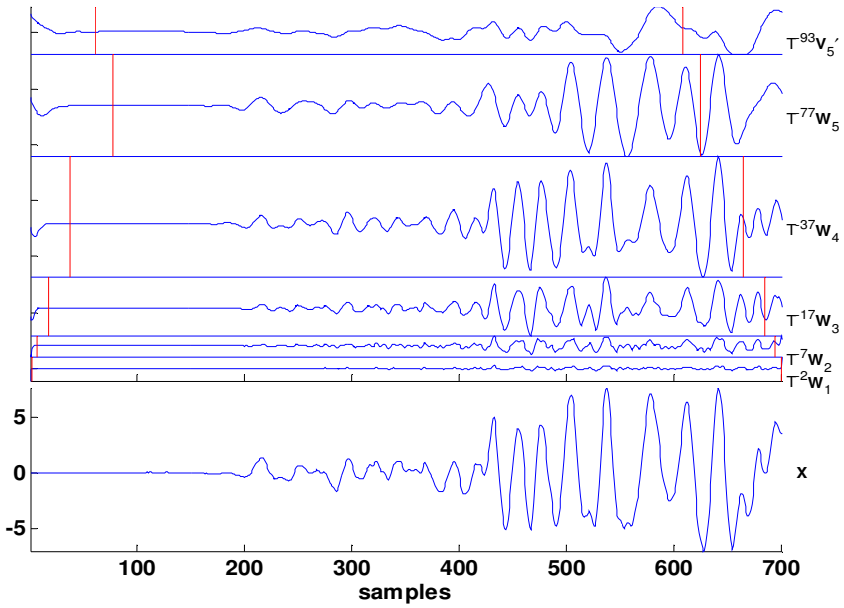


Figure 2.31.Expanded view of Figure 2.27

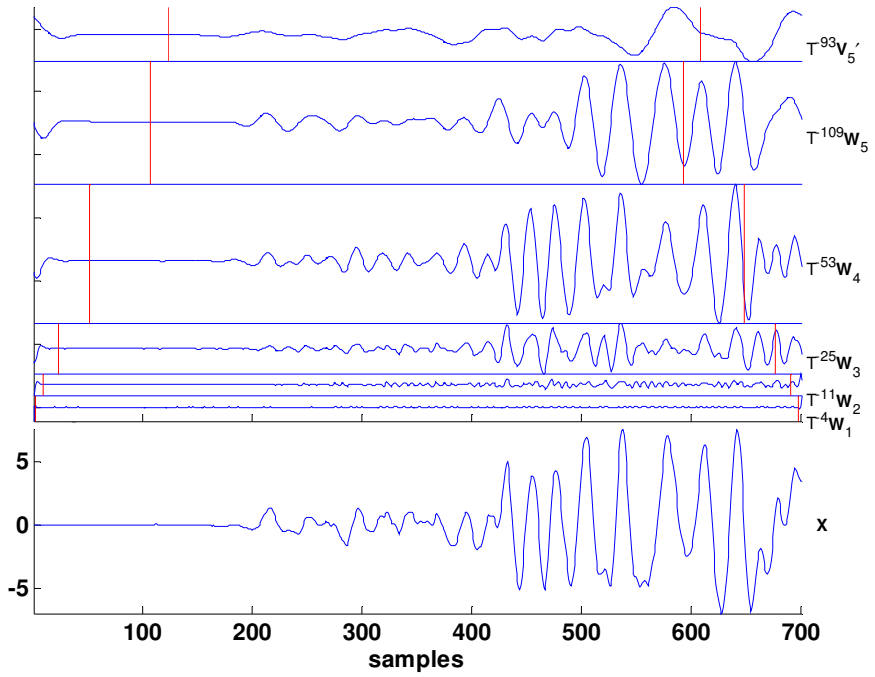


Figure 2.32.Expanded view of Figure 2.28

2.11 Summary

This chapter provides the essential framework needed from the wavelet literature to support the application of wavelet methods in forthcoming sections. More specific, initially the relation between MRA and wavelets is described in order to derive the essential link of using wavelet analysis to perform MRA. Then the advantages of WT over FT according to time-frequency representations are presented and as an example the CWT is described. Dyadic discretization of CWT leads to DWT which can easy be implemented using high & low pass filters which presented along with typical implementation algorithm. Practical considerations of DWT including boundary conditions, choice of last level for a partial DWT and handling sample sizes that are not power of two are also discussed. An extension of DWT, the MODWT is presented along with its implementation algorithm and corresponding practical considerations. The chapter finishes with a field example (a seismic event) which is analyzed using the aforementioned wavelet approaches.

3. Design of Seismological Data Collection Network

3.1 Introduction

This section describes the hardware and software aspects for the Hellenic Seismological Network of Crete (HSNC) of the Laboratory of Geophysics and Seismology (LGS) at Technological Educational Institute (TEI) at Chania, which is mainly designed, built, installed, tested, deployed and maintained by the author in order to enable the research reported in this thesis to be carried out.

The Aegean region (34–42°N, 19–29°E) which comprises the Hellenic arc and the adjacent areas of the Greek mainland, the Aegean Sea and western Turkey, is one of the most seismically active zones of the world and the most active in western Eurasia due to the convergence between the African and Eurasian lithospheric plates. The seismic activity especially in the southern Aegean area is very intense and extends up to a depth of about 180 km. The seismicity of South Aegean (34–38°N, 21–28°E) is extremely high and is characterised by the frequent occurrence of large shallow and intermediate depth earthquakes (Fig.3.1 and Fig.3.2).

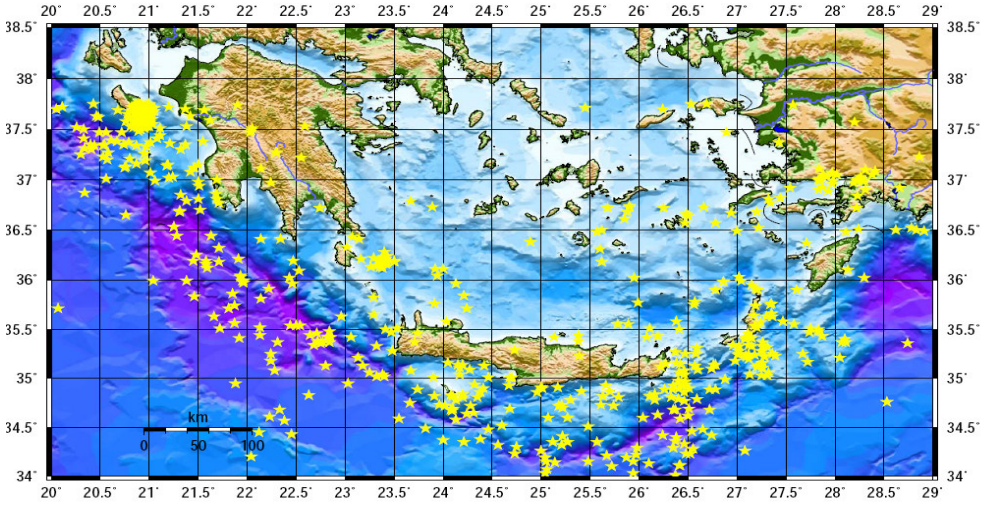


Figure 3.1. Geographic distribution of epicenters for earthquakes with magnitude M_L>=4 during for 1/5/2005 until 30/9/2006. Data extracted from preliminary earthquake catalogues of HSNC and GI-NOA (http://www.gein.noa.gr/services/cat.html).

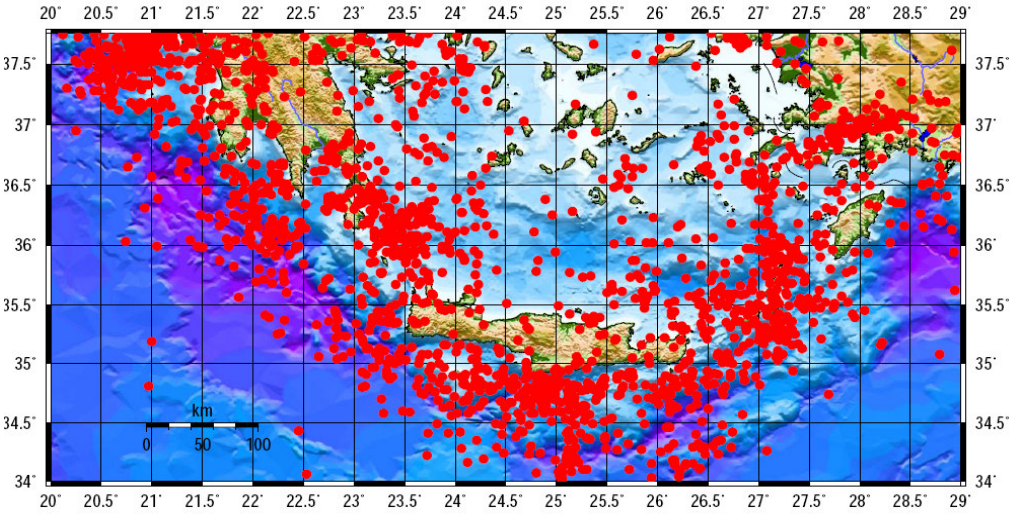


Figure 3.2. Geographic distribution of epicenters for earthquakes with magnitude $M_L<4$ during for 1/5/2005 until 30/9/2006. Data extracted from preliminary earthquake catalogues of HSNC and GI-NOA (http://www.gein.noa.gr/services/cat.html).

Main geotectonic feature of the region consists of the Hellenic Trench, where the Eastern Mediterranean oceanic lithosphere (front part of the African plate) is subducted under the Aegean microplate. In historical times, this high seismicity has produced earthquakes of magnitude as large as 8 (Papazachos, 1990; Papazachos & Papazachou, 1997) pointing out the great seismogenetic potential of the area.

Such an area demands a dense seismological network in order to achieve three main objectives:

- The most possible accurate estimation of earthquakes' attributes in order to increase the knowledge of area's geotectonic structure.
- Investigation of microseismic activity in order to reveal possible unidentified faults or to study the aftershock sequence after strong earthquakes
- Deployment of a regional EEWS focused on the protection of specific areas inside it.

The deployment of public available seismological stations before the installation of HSNC (i.e. 2003) is depicted in Fig. 3.3. The installed stations with public data availability belonged to GEOFON network (http://geofon.gfz-potsdam.de/geofon/). This network wasn't sufficient for the requirements described

above because of two main reasons: the average distance between stations was over 200km and the collecting data center was held in Germany, from where the data were available. The result was that if a research based on these stations only, the results would be rather inaccurate. In addition if a local authority wanted to operate an EEWS based on these stations except their limiting number there would be introduced unnecessary delays from the data transmission to Germany and back (through analog leased lines). The elimination of these shortcomings is the primary objective for the design and installation of SNC that will be described in the next sections.

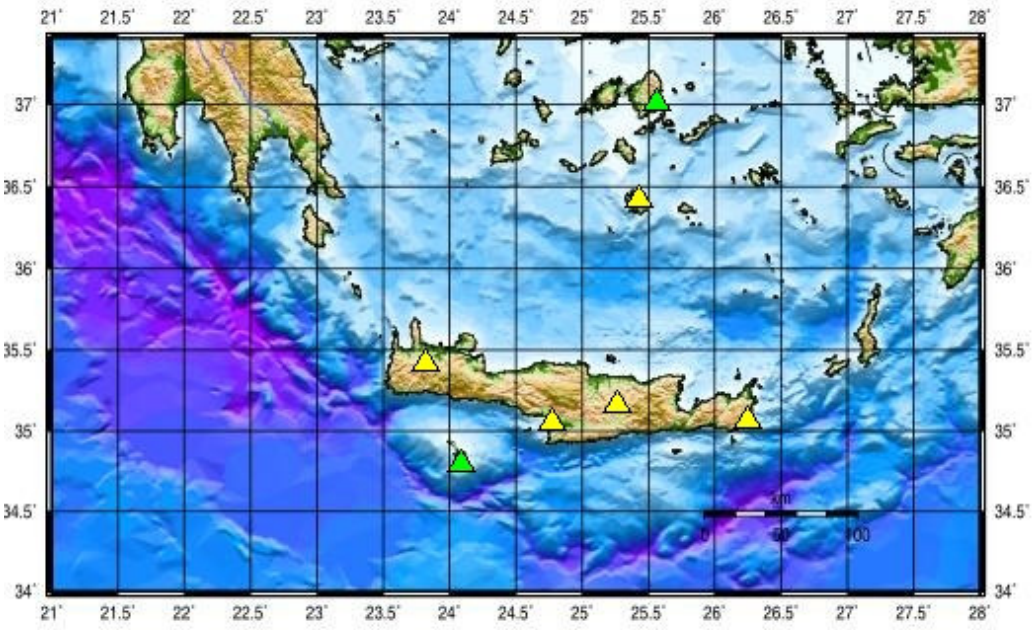


Figure 3.3. Geographic topology of the installed available stations before the installation on Seismological Network of Crete. Yellow triangles denote real time data transmission and green triangles denotes non-real time data transmission.

3.2 Seismological stations network

3.2.1 Requirements analysis

After extensive analysis of the requirements the following issues are critical for the successful long-term operation of a seismological network:

a) Geographical spreading to the whole Crete and to broader South Aegean area.

b) Autonomous, fully automated, high resolution and dynamic range measuring field stations that will incorporate advanced techniques and technologies of analog and digital measurements
c) Wired and/or Wireless data transfer by implementing contemporary communication technology (optical fibres, ISDN, DSL, satellite connection, internet etc)
d) Real Time Telemetry and continuous, error-free, data transmission
e) Central Station surveillance and massive data storage
f) Secure and continuous access (through internet, e.g. on-line or ftp) to the multi-parametric database of the network from certified users.
g) Hierarchical data processing seismological software, expandable and upgradeable.
h) Possibility for automated dissemination of selected data and/or information to the authorities.

3.2.2 Geographical Location selection

When setting up new stations, they can be either part of a local or regional network, or each one can be a single station. In our case the design is a regional network. Studying a map will then give an ideal location, often a remote one, which however could be severely constrained by practical considerations:

- Power: The power required by the equipment can vary between 1 W, to more than 100 W depending on type of station. High power stations need access to the power grid while low power stations might get by using solar cells, in which case a careful consideration on the size of batteries and solar cell to use must be made (see section later on power). It is worth compromising other aspects of station installation in order to be able to use a public power line.
- Communication: If the station is to transmit by radio, it often has to be located near mountain tops. Satellite communications do not have this limitation, but use more power. Land line (leased lines private networking) communications require the station to be near to a network service provider hub.
- Noise: The station must be away from noise sources, particular human made.
- Security: Unfortunately, vandalism and theft is not uncommon and some areas are worse than others. There is not much point in putting up a nice low noise

station, if the solar cells and batteries disappear within a week. So apart from making a solid construction, the geographical location can be important to avoid these problems.

- Access: Ideal locations might require a helicopter to get there. Since few seismologists can afford that in the long run, easy access to a site is essential to ensure long-term maintenance.
- Weather: Areas with humid condition, extreme cold or warm areas or with a high probability of lightning were avoided. Within a given geographical region, there might not be much variation in these parameters.
- Topography: It is generally advised to avoid rough topography, which might modify seismic waveforms, however this conflicts with putting stations on hilltops for better communication.
- Geology: It is well known that a sensor on soft soil is noisier than a sensor on solid rock, so soft soil is avoided.

The above requirements are all conflicting since an ideal station is likely to be in a remote area with no power, communication or security. Our primary goal was to ensure low cost and long-term operational stability because it is usually easier to get initial installation funds than operational funds. This often means that in some cases, the noise level will be higher than ideal and fewer earthquakes will be recorded. However since the operational stability will be higher, the total amount of data might not be reduced compared to a station with unstable operation. This fact is the motivation beyond the chapter 3 where a comparison between wavelet denoising methods took place in order to find the appropriate ones.

A seismic station located on solid basement rock can always be expected to have a lower noise level than a station of soft sediments, even if there is no clear noise source nearby. However, there can be unknown noise sources or the geology can be different from what it appears to be on the surface. What appears to be bedrock, might just be a big boulder. Before making the final selection of a site for a seismic station, preliminary ambient noise studies performed. As we are going to see (Chapter 6), noise above 1 Hz is mostly originating in the near field, while lower frequency noise might originate far away. The ambient noise survey had a dual purpose: first by running a short duration (~1h) ambient noise survey the site's fundamental frequency and amplification ratio were estimated by means of HVSR method as described in

Chapter 4. Results from FRMA station depicted in Fig.3.4 where for other stations are presented in Appendix B

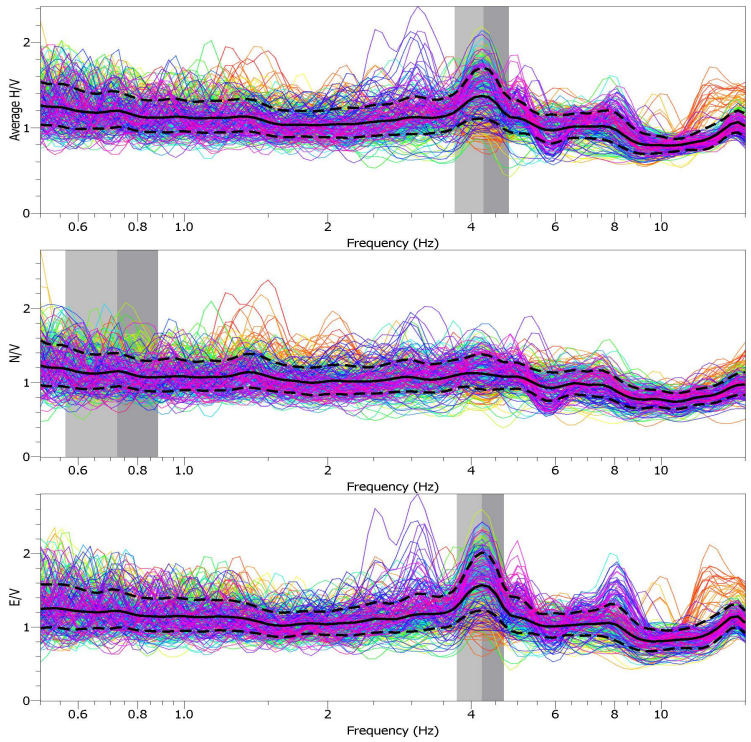


Figure 3.4. Ambient noise survey for seismological station FRMA.

The noise survey consisted of recording continuously for e.g. 24h. Then time intervals of e.g. 10 min, took out and average noise spectrums were calculated which are then compared to the Peterson (1993) curves as depicted in Fig.3.5 for the station FRMA. Peterson curves show the New global High Noise Model (NHLM) and the New Low Noise Model (NLNM). These curves represent upper and lower bounds of a cumulative compilation of representative power spectral densities determined for noisy and quiet periods at 75 worldwide distributed digital stations. These so-called Peterson curves have become the standard, by which the noise level at seismic stations is evaluated. Typically the average spectra of a seismic station must be fall between these curves.

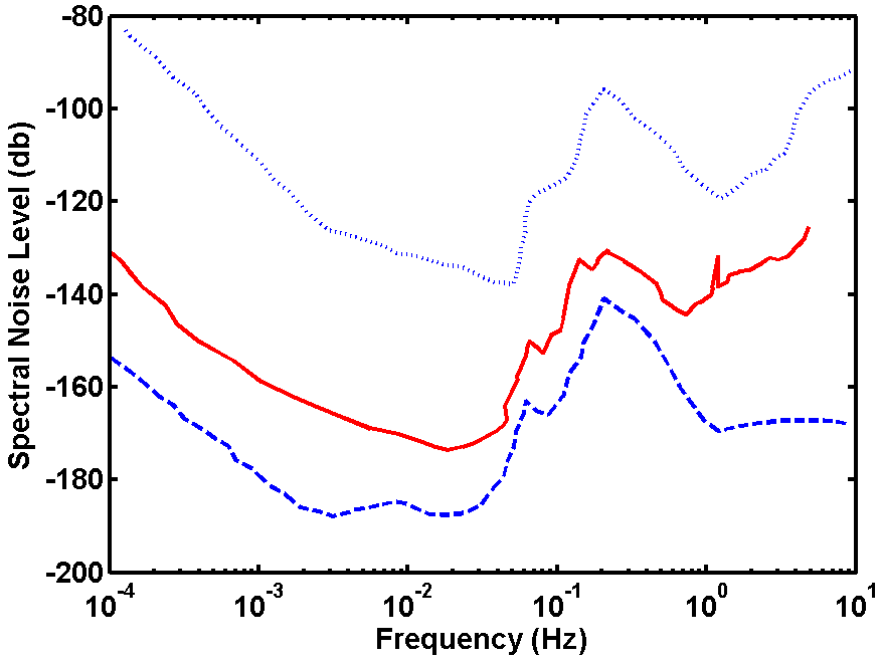


Figure 3.5: Peterson noise curves (NHNM – blue dotted line, NLNM – blue dashed line) and the noise spectral level of Z component (red solid line) at station FRMA

3.2.3 Sensors

Two types of seismometers used in SNC: GURALP CMG-T40 (short period and broadband) and SERCEL L4-3D velocimeters which have 1Hz eigenfrequency as depicted at Fig. 3.6 (Sercel) and 3.7 (Guralp). Both are passive seismometers and contained three independent sensors (Vertical-Z and two horizontals – NS, EW) featuring low noise, large dynamic range and easy installation and use.

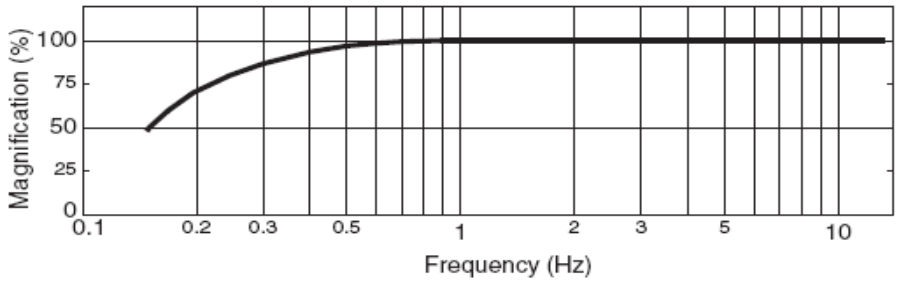


Figure 3.6. Frequency response for Sercel seismometer (www.sercel.com)

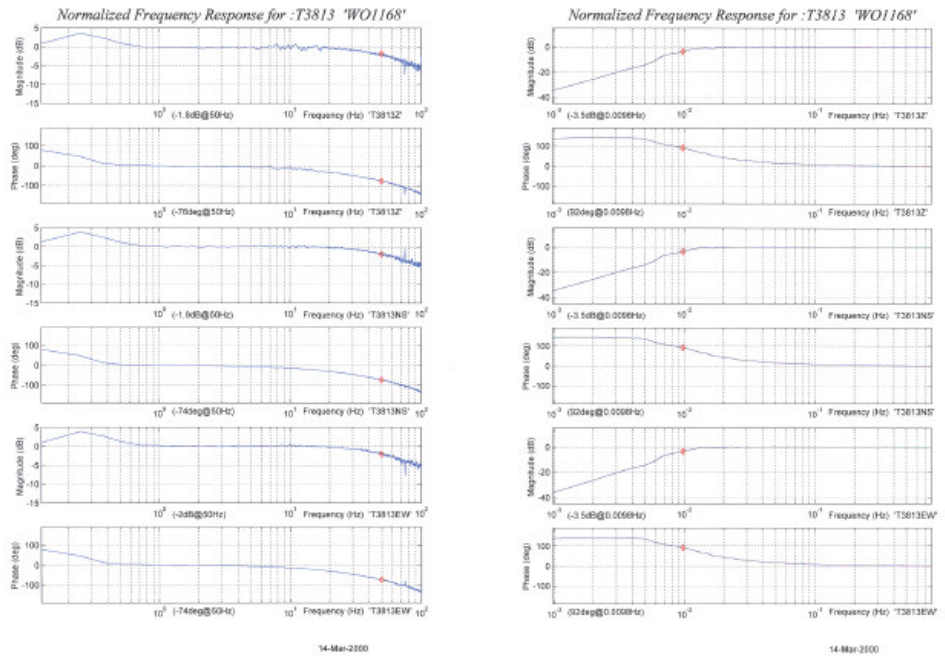


Figure 3.7. Normalized frequency response (high frequencies – left, low frequencies – right) for Guralp CMG-40T seismometer (www.guralp.com).

3.2.4 Data Acquisition System (DAS)

The DAS system that used is the DAS-130-01 by Reftek inc. (Fig. 3.8). It 6 input/output connectors, an LCD display, and a removable lid on top of the case. There are two Channel input connectors (each one can record from a 3-component sensor), a Terminal connector for setup and control, a Net connector combing Ethernet and Serial PPP for network access, a 12 VDC Power connector, and a GPS receiver connector.

Figure 3.8. The DAS installed at SNC (Reftek inc)

Its LCD display allows the user to monitor the status of various subsystems within the 130 without having a terminal device attached. The Crystal Semiconductor 24-bit/digital filter chip set provides more than 135 dB of broadband dynamic range. The digitizing rate is software-selectable from 1000, 500, 250, 200, 125, 100, 50, 40, 25 and 20, 10, 5, and 1 sps (per channel). At any given sample rate, wideband performance is achieved through the use of digital Finite Impulse Response (FIR) digital filters, which provide passband to 80% of the Nyquist. The gain of each channel is software-selectable to either one (x1), thirty-two (x32) or one hundred (x100). The removable CF media allows back up of the data when the link is down for a long period of time, so no data will be lost. If the telemetry link is down for a very short time, then all the data remains in RAM and will be transmitted to the central acquisition station as soon as the link comes up. Data recorded on the CF card then can be *ftp'd* to the central station at the user's request. Two CF II cards of 2 GB capacity totaling 4 GB allow back up of the data for at least 200 days' recording, three channels at 100 sps in Steim2 compressed mode. Two CF II cards also allow the 130-01 Recorder to operate in an "auto-wrap" mode. In this mode, the first disk is filled with data; once it is full the second CF II card is formatted and then filled with data. Once the second disk is filled with data, the first disk is then formatted and data is then recorded to the first disk. When the first disk is again full, the second disk is formatted and data is recorded to it. This process repeats forever. In this mode the user always has at least one CF II card worth of the latest data, plus whatever has been recorded to the current disk. When data is also being recorded to a telemetry link while also recording to the disks in auto-wrap mode, if the telemetry link goes down, the 130 will finish filling the disk it is currently writing to, and then it will format the other disk and write to it until it is full.

For a seismological field station the power consumption is one of the crucial factors. The 130-01 uses 3.3 VDC logic instead of 5V DC logic. In addition, it is utilized the duty cycling of microcomputer subsystems to eliminate unnecessary power drain. For example, the ADC outputs data to a shared memory device so that the CPU can be asleep during acquisition. For a typical application, the central processor unit (CPU) will be active less than 1% of the time in order to move ADC samples into non-volatile memory. Field station recording equipment is very low consumption DC powered equipment. Extra power consumption is produced when

using telemetry and active broadband sensors. The maximum measured values for all the DASes used are: 210 mA (continuous telemetry and data transmission, GPS in cycle mode, sampling at 125Hz, gain x1, passive 3 component sensor attached) plus 108 mA if an active broadband sensor is attached providing a total of 318mA. That means around 7.7Ah per day needed. The calcium lead-acid battery that installed with each DAS is 65Ah which means that can hold up to 6 days (assuming 80% battery availability due to charging cycles) without charging. If there isn't active sensor connected then this period can extend to 9 days.

Accurate timing is provided by a high precision crystal oscillator and an external reference provided by the Global Positioning System (GPS). There are two two possible commercial GPS modules tested : Garmin's and Trible's . It is found that Garmin GPS receiver maintains time an order of magnitude better than the Trimble GPS (100 microseconds compared to 1 millisecond). When power is applied, the 130 comes up with its internal time set to the base year. Current time can be manually set by the operator or automatically by the GPS receiver. The 130-01 will monitor the output of the GPS until a valid 3-D fix has been obtained, then update the frequency of the local oscillator and set time. If the GPS receiver has been recently active (warm start), this process will only take about two minutes. If the GPS has been moved a great distance since its last activity, or if it has not been operated for several months (cold start), this process will take approximately five minutes. The frequency setting is maintained so that the oscillator frequency is always within spec, compensating for aging and temperature drift.

For back up data retrieval, the internal hard disks can be hot swapped in the field or data can be copied via ftp to the CRC. In the field the user can swap an empty formatted disk with the one that has recorded data after verifying in the LCD message the RAM space available before the data is dumped to the disk. The disks can be swapped between dump cycles. Thus, the disk may be removed while the 130 continues to acquire data. Empty formatted disks can be used for data exchange in the field. This approach is very practical because it consumes only a few minutes at a field site. A disk full of data can be then carried to the Lab or Central Station for connection to your data processing PC via standard commercially available CF reader. Another way to get the data from the 130 is to ftp from the central processing PC.

A schematic showing field installation of 130-1 is depicted in Fig. 3.9

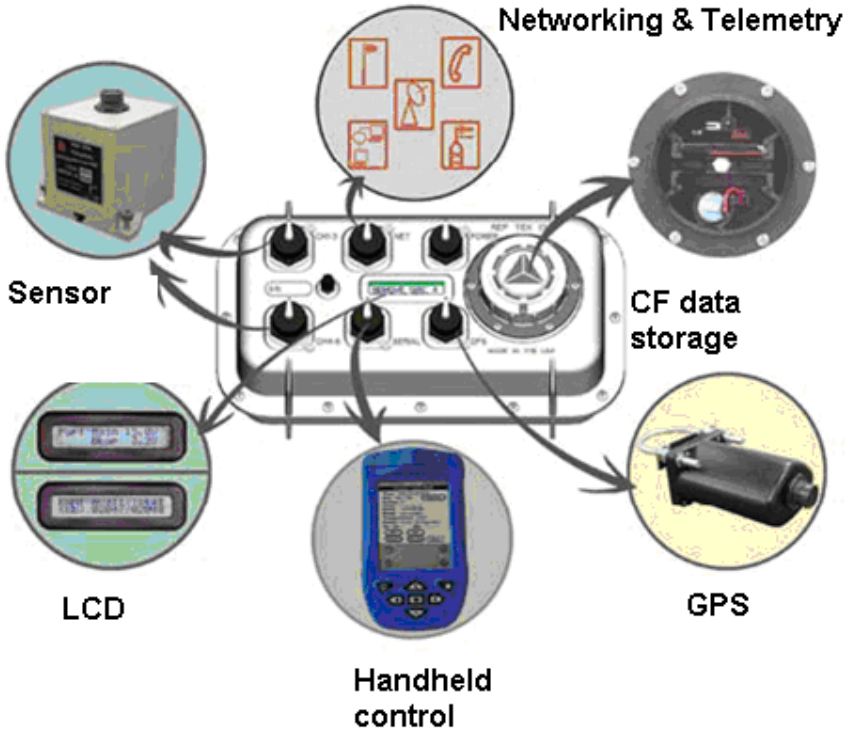


Figure 3.9: General configuration of a field seismological station

Each DAS has built-in hardware to support IP communications. It employs TCP and UDP other either Ethernet or PPP Serial interfaces. Each interface has its own IP address, address mask, gateway, and target host IP address

3.2.5 Installation of the seismic station

The installation consists of setting up sensor and possibly recorder and/or transmission equipment. The installation will depend on type of sensor, and whether local or remote recording is done.

Sensors are usually posed on aluminium enclosures to avoid stray currents and corrosion of the sensor feet when in contact with cement. The horizontal sensors were always oriented NS and EW such that a motion N or E gave a positive signal. The most common way to orient the sensors is to use a magnetic compass, taking magnetic declination into consideration. Inside buildings or vaults this might be unreliable and the direction must be taken outside and transferred inside. The orientation of the sensors should be within ±2 degrees.

Short period sensors usually do not require thermal insulation since they are little sensitive to temperature changes. Broadband sensors on the other hand require a very good thermal insulation.

All the sensors were installed on basement rock. The depth of the installation is then only important to the extent of eliminating low frequency noise caused mainly by temperature changes.

For a telemetered station in a remote area, more security had to be provided rather than the simple construction and there also be a need to install more equipment. In a construction, as this shown in Fig. 3.10 sensor, DAS, communication equipment and battery are placed in the house, while the GPS is outside.

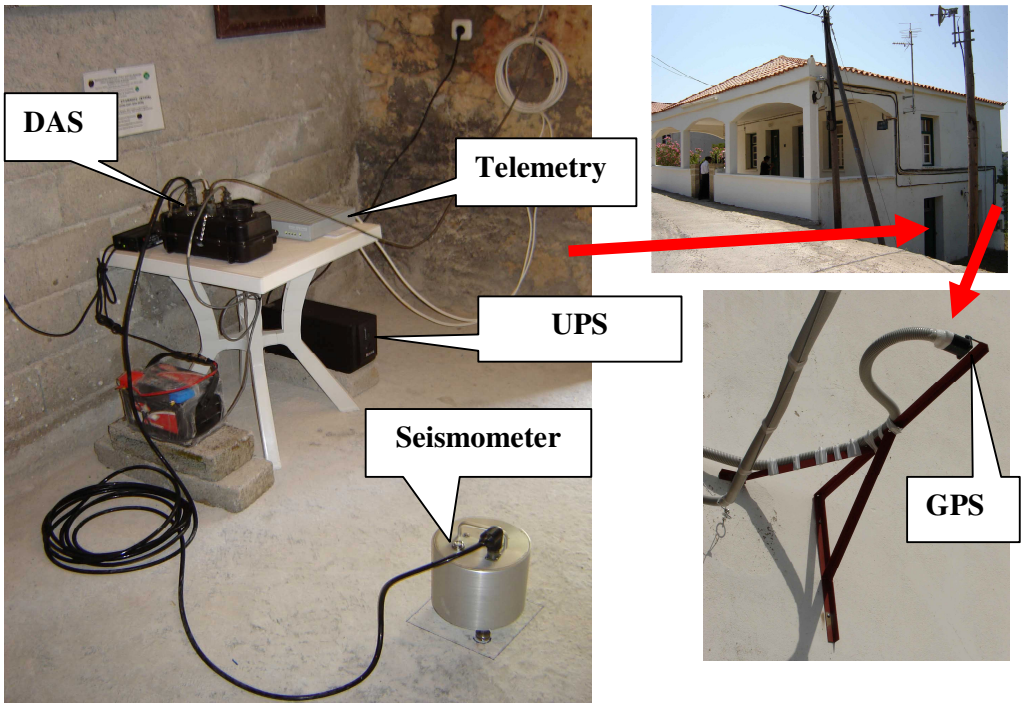


Figure 3.10: Typical installation of a short period seismological station: inside view (left), outside view (right)

Broad band sensors require special care in installation in order to get a good low frequency performance. The lower the frequency, the more demanding and costly is the installation and the sensor, so it should be carefully considered what the purpose of the station is before starting an elaborate construction. The main problem with low frequency noise is tilt, which mainly affects the horizontal sensors. A deformation of 1 μm over a distance of 3 km oscillating with a period of 10 min gives a vertical

acceleration of 10^{-10} m/s^2 in the vertical direction and 10^{-8} m/s^2 in the horizontal direction (Bormann, 2002). These values are well above the NLNM. Tilt can be caused by temperature deformations of the ground or atmospheric pressure changes. A good installation should therefore be as deep as possible and in hard rock. The BB sensor itself will also be affected by temperature fluctuations, so short term temperature changes (with smaller duration than periods to be measured) should be avoided. Different sensors have different sensitivity to temperature changes. Typically, the sensor should be protected from day-night temperature changes. The 24-hour sensor variation in temperature should be less then 1 degree. This stability is mainly achieved by insulation. A simple insulation that used is the one depicted in Fig.3.11

Figure 3.11: Broadband seismometer thermal insulation (inside installation)

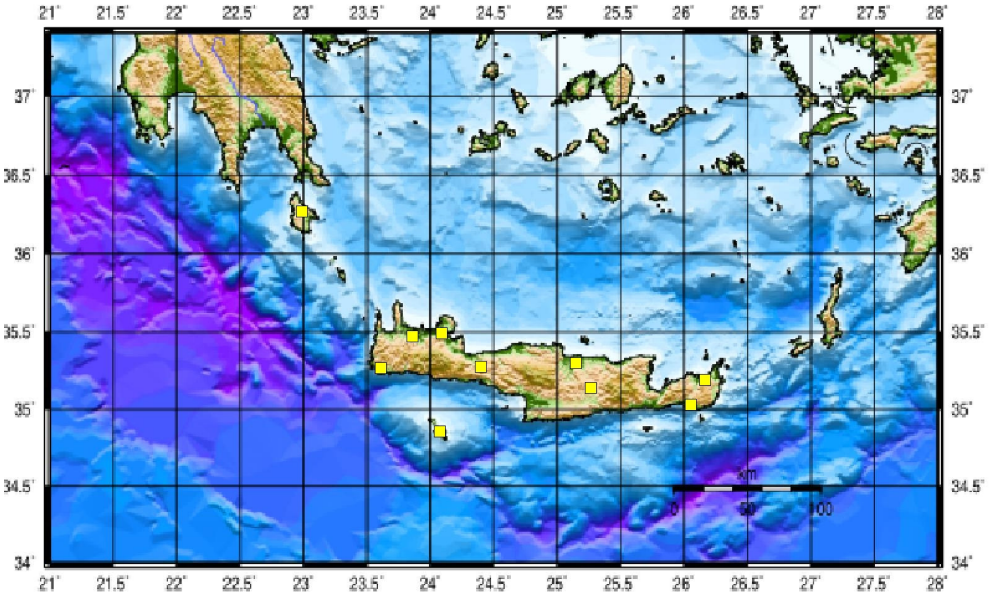


Figure 3.12: Locations of the currently installed SNC stations

The current status of the network is depicted in Fig.3.12. and presented in Table 3.1. All the stations are registered to International Seismological Committee (ISC) and to International Federation of Digital Seismograph Networks (FDSN).

Station Code	Station Name	Latitude (deg)	Longtitude (deg)	Elevation (m)	Sensor
STIA	Sitia, Lasithi	35,2020	26,0909	89	Guralp CMG-40T
KSTL	Kastelli, Heraklio	35,3037	25,0707	70	Sercel L-4-3D
KNDR	Koundoura, Chania	35,2340	23,6857	25	Sercel L-4-3D
PRNS	Prines, Rethymno	35,3649	24,5015	51	Sercel L-4-3D
KLDN	Kaloudiana, Chania	35,4870	23,6905	44	Guralp CMG-40T
FRMA	Ierapetra, Chania	35,0144	25,7334	33	Guralp CMG-40T
CHAN	Chania	35,5192	24,0425	34	Guralp CMG-40T
HERA	Heraklio	35,0722	24,7656	17	Guralp CMG-40T
GVDS	Gavdos	34,8435	24,0902	164	Guralp CMG-40T
KTHR	Kythira	36,2606	23,0042	315	Guralp CMG-40T

Table 3.1 – SNC permanent station list

During the development stage there were some temporary installations that carried out. The locations of these installations are depicted in Fig.3.13.

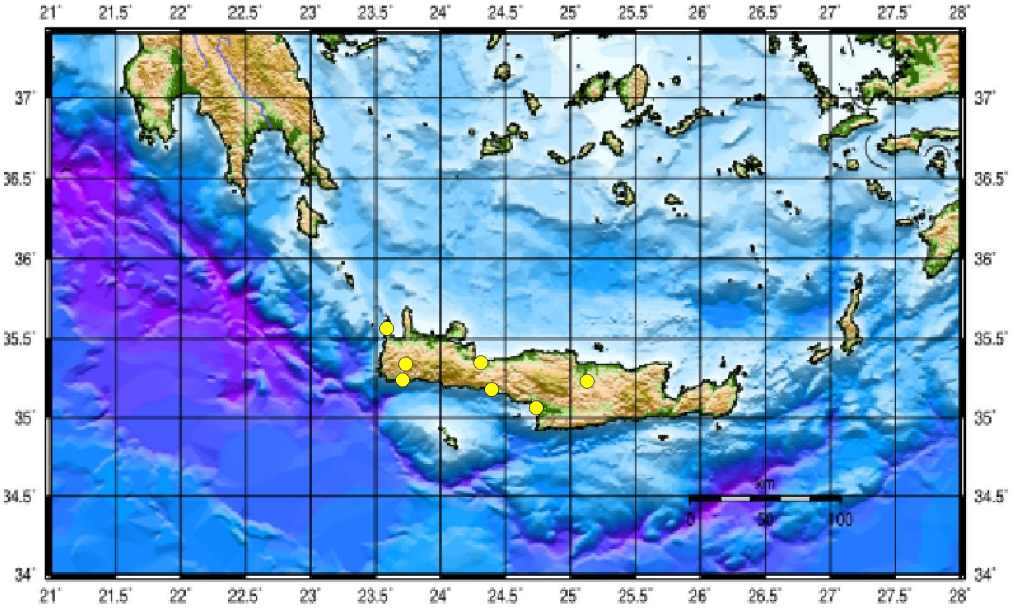


Figure 3.13: Locations of previously installed SNC stations (yellow circles).

3.2.6 Data Transmission and Telemetry

The most important factor for the design of SNC was the reliable and continuous data transmission to the Central Recording Centre (CRC) along with the

highest level of security, redundancy and failover performance in the most possible automated way. The network design based on the client-server model where each element of the network can pass requests and data to hosts inside the network. Each network element can be act as client and server at the same time but for different purposes. For example a DAS is the client that sends data to Data server at CRC. The same DAS can act as server if a certified user wants to download specific data from its CF disk using an ftp client.

3.2.6.a Communication protocol

Since each DAS has the ability for IP communications the obvious choice is a TCP/UDP implementation. We had need for a relatively simple packet-oriented reliable transport protocol that can be easily implemented in embedded systems. We had also the need for transport these packets across Internet Protocol (IP) networks reliably, that is, in an ordered and error-free way. There were several approaches that one might take to accomplish this. The TCP/IP protocol suite provides two basic protocols at the transport layer that are accessible to the application through the Sockets API. These are the Transport Control Protocol (TCP), and User Datagram Protocol (UDP). A third protocol, the Reliable Data Protocol (RDP), has been defined that would generally meet our requirements but this protocol is not typically implemented in the IP stacks in most Operating Systems (OS) available today.

TCP provides reliable, connection-oriented, byte stream service to the application and readily meets our requirements. However, TCP is relatively very large and complex and building a stable implementation is a daunting process to undertake. TCP provides much more functionality than is required to simply move 1 kilobyte packets across the network reliably and the size and complexity of implementation is simply too great. UDP, on the other hand, provides a very simple connection less, packet (datagram) oriented service that is unreliable. Because moving data are already in packets and these packets are 1 kilobyte or less in size, we seek for a protocol that could provide the additional functionality that we require at the application layer and use UDP as the underlying transport.

A second, yet important, requirement is that a client application must be able to connect to the server application without requiring any configuration information. It must connect to the server without any knowledge of its own or the servers addresses

and so forth. UDP provides this capability while TCP, because it is connection-oriented, does not.

Unfortunately the widely used TCP/IP stack cannot satisfy other requirements also:

- Platform independency. All data values must stored in network byte order. Thus, it must be implemented on any hardware platform or OS that provides an IP protocol stack.
- Encapsulation of the application data completely without having any dependency on the contents of any particular application data package
- Self-contained and self-configuring at the client. That is, the protocol stack must discover or be assigned all necessary parameters to operate from the network
- Both the server and client must initiate the connection on-demand. This means that when the application has data to send the connection will be established if needed simply by submitting the data packet to be sent.
- Recover from loss of connection without data loss. This means that if a client is sending data to the server and the connection is lost momentarily, it will reestablish the link and resume sensing the data.
- Dealing with long, thin, pipes effectively. That is to say that it must be capable of high utilization (>90%) of slow (9.6 kbps), high latency (>1 second), connections such as VSAT links.
- Small and relatively simple implementation.
- Function at the application layer using the standard UDP Sockets API on the server system.
- All network traffic must be UDP diagram to/from specific port number.

Our choice was the Ref Tek Protocol (RTP). RTP is designed to provide the application with a full-duplex, packet-oriented, reliable, transport over UDP network connections. RTP is typically used in server-client fashion although this is by no means required. Typically there will by a server application running on a IP host somewhere on the network. Clients will attach themselves to the server to send and receive data. The client is typically an embedded system that attaches to the network through an asynchronous serial interface using Point-to-Point protocol however other interfaces will be implemented in the future. The server is typically an application

program running on a host and accesses the network via UDP sockets provided by the IP stack on the local operating system. The real-time data transmission is performed using RTP client. The RTP client uses UDP/IP on port 2543, which is registered with the Internet Assigned Number Authorities (IANA). The payloads of RTP datagrams are 130 commands, responses and recording package. A typical network configuration using RTP is shown in Fig.3.14 and its correspondence to OSI model , in Fig. 3.15

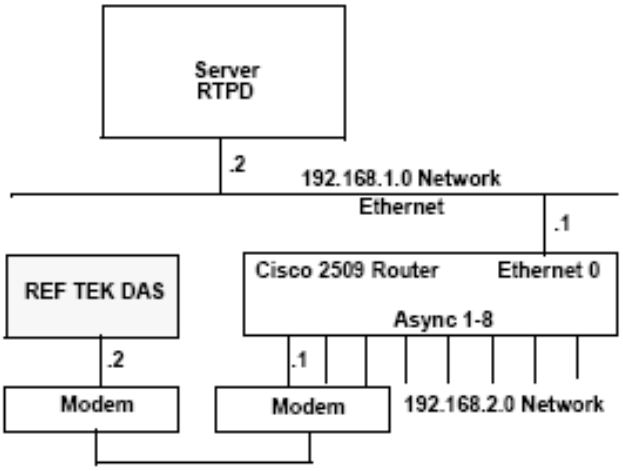


Figure 3.14: Typical network configuration using RTP

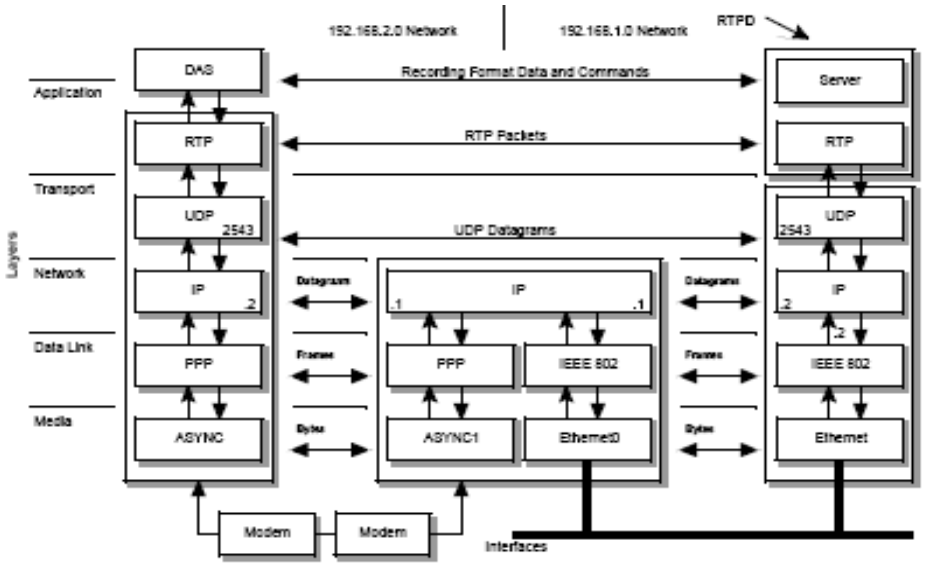


Figure 3.15: Layers and interfaces at OSI view for RTP

Once the DAS is powered up, the RTP client generates the discovery packet, which includes the DAS ID number, data collection host IP address, port number and

send it to the interface programmed by the operator for real-time data transmission. If the DAS is connected to the IP network, the discovery packets are sent to the IP network. The data collection host resided at the CRC is "listening " the port 2543, and as soon as it receives the discovery packet from the DAS with the unique ID number, performs the hand shaking with the RTP client and establishes the connection with the DAS. If the DAS is programmed for a continuous data recording to a designated communication interface (Ethernet or Serial port), data collection host begins receiving the data. The error-correction is performed by the positive acknowledgment from the host. In the event of the communication link failure, the data is acquired on the RAM while RTP client periodically sends the discovery packets. After the communication link comes back up, the data collection host will again receive the discovery packet from the remote DAS, re-establish the communication and resume the data acquisition automatically.

3.2.6.b Wired Links

For the needs of wired links from seismological stations to CRC a Virtual Private Network (VPN) was implemented. The VPN based on Multi Protocol Label Switching (MPLS) technology over conventional Asynchronous Digital Subscriber Lines (ADSL).

MPLS is defined by IETF (Internet Engineering Task Force). It combines the scalability of IP protocol, high speed switching from OSI's 2nd level and the ability to manage the available bandwidth using Traffic Engineering (TE) methods. It is compatible with the main routing protocols that used by Internet Service Providers (ISP) today such as OSPF, RSVP, MP-BGP, IS-IS and advanced services such as IP-VPNs, L2TPv3 providing security levels similar to ATM and Frame-Relay technologies with lower cost. Briefly speaking, the advantages of MPLS are:

- Security. The installed VPN was totally independent from other VPN existing between seismological stations. This is very important in the case where data exchange between collaborating seismological networks is planned. In a case that SNC want to send the data from the stations that had only at the West part of Crete, a new VPN between SNC and other institute can be defined including only these stations. The two VPNs will be isolated even if they include some common stations. To achieve this author used independent

routing tables for each VPN. The innovating part is that in each routing table the VPN ports are defined exclusively without leaving the ISP to handle this. Fig. 3.16 depict this situation

- Flexibility. The use of different topologies (star, partial mesh, full mesh) by using several accessing methods (leased lines, aDSL, Dialup, GPRS) is not forbidden. This is important for two cases: first if a station faces access problems the use of an alternative link to the nearest station can solve the problem without changing the configuration of the VPN. Second, for the use of permanent networks (e.g a mobile seismological network for aftershock studies) only one point of access will be needed in order all the stations of the permanent network to reach the VPN. The above cases are depicted in Fig. 3.17.
- Quality of Service: The use of priorities policy is required in order to ensure that data are always be available. In the case of SNC the design gives higher priority to the seismological stations and lower priority to clients such as certified users that received data or events. This is dictated in order to ensure that in the case of a strong event where a malfunction in telecommunication facilities happens, the available bandwidth will be primarily used for the data transmission to CRC.

In networks with ordinary routing, packets routed from the source to the destination hop-by-hop. Every intermediate router extracts the header of corresponding OSI layer, the information needed to forward it to the next available router. This process is repeated for every hop. The routing decision depends from several factors such as:

- Generation and maintenance of dynamic routing tables.
- Sorting of IP packets in Forward Equivalency Classes (FEC) based on the destination address.
- FEC assignment in the next hop.

The packets that belong to the same FEC forwarded through the same route. As the packets transferred into the network, every hop checks the packet and assigned to the corresponded FEC. By using MPLS the packet assignment did only once; at the packet's entry to the network. FEC is coded with a constant length label which is embedded into packet. The next hop, instead of checking the contents of IP header

checks only the label in order to decide where to forward the packet. At every hop the old label is replaced by a new one. Data packets are forwarded through specific Label Switched Paths (LSP). Using this approach every router knows in advance what it has to do for packets with specific labels.

The VPN that built for the purposes of SNC is shown in Fig. 3.18. For backup purposes the routers that used has the option to connect to the VPN through ISDN lines. This is selected in order to ensure that in case of ADSL failure the data connection keeps active.

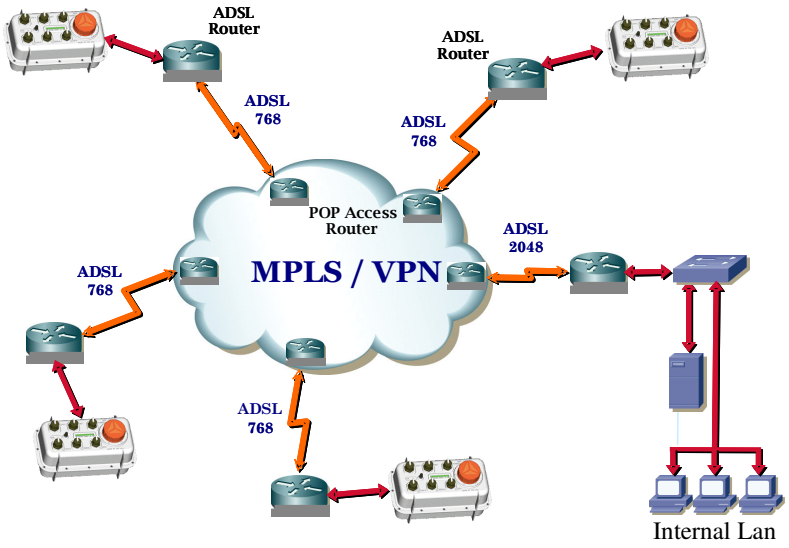


Figure 3.18: VPN topology for the wired links used by SNC (shown 4 from 8 stations)

After the installation of the first six stations an availability survey (for one year operation) were carried out in order to estimate the total availability of the wired link. Results are shown in Fig.3.19. As it can be seen one station presents medium availability while the remaining ones have availability over 92%.

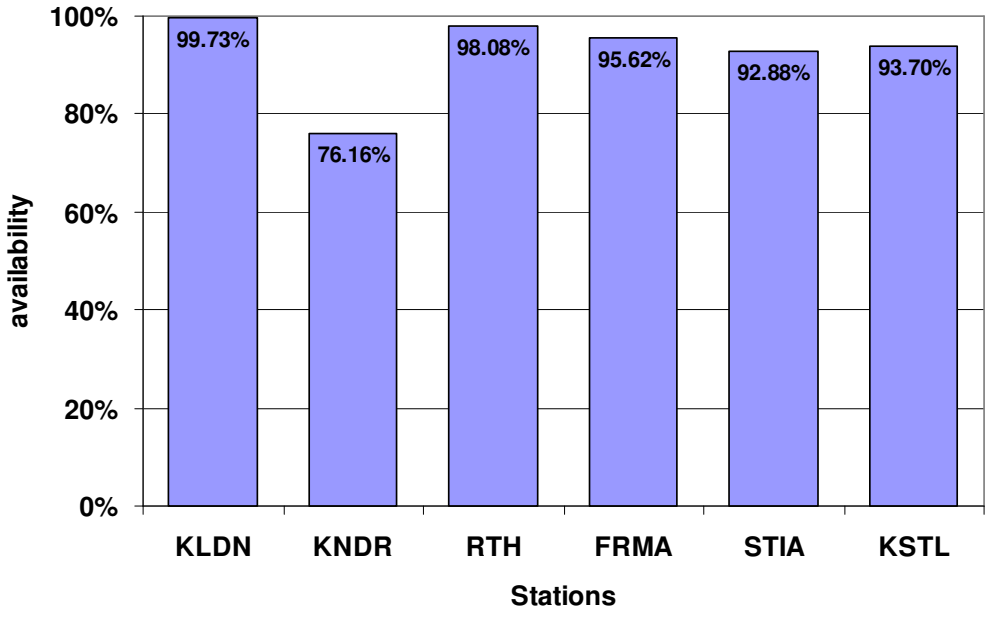


Figure 3.19: Wired link connection availability survey for six SNC stations over 2007

3.2.6.c Satellite Link

When the distance from the nearest point of network access is a forbidden factor the solution of wireless link becomes irreplaceable. Since the distance from the remote stations to CRC is over 200km in some cases the Satellite link (VSAT) is chosen as the only reliable solution as explained below.

In order to fit the satellite link in our needs the following requirements are described during the design stage:

- Quick and easy development of VSAT terminals. The majority of the cases where a VSAT system is going to be installed will be in locations difficult to approach.
- Support of low data rate in VSAT terminals. For sampling 3 channels at 125Hz, each DAS needed around 8 Kbps Transmission rate and 2 Kbps Receive rate (including event and state-of-health packets). These values are rare in the VSAT links because today satellite links allocate obligatory a minimum bandwidth for each terminal (minimum Committed Information Rate (CIR) / remote). If some of the stations exceed this minimum bandwidth demanding more, extra bandwidth is allocated by the remaining available

(Burstable Bandwidth on Demand - BBD). The minimum CIR/remote is usually 20 Kbps. Selection of such a VSAT system for 10 stations would lead to requested bandwidth of 200Kbps when only 80Kbps needed which means over doubled costs.

The only satellite platform (and the one that selected) that had the ability of configurable minimum CIR / remote in a reasonable cost was the iNFINITI platform by iDirect (Dallas, USA). The minimum CIR/remote could be decreased until 16Kbps for general data packets which was the double bandwidth from what we expected. For this reason a different approach was used: the data packets from each DAS defined as voice packets and each link from DAS to CRC denoted as VoIP link. Under this approach the ability of using voice services at the satellite platform was revealed. This is important because the specific platform designed keeping in mind that voice packets must be handled different from data packets. That means that voice packets expected to be smaller than data packets and must have higher priority in order possible losses cannot be identified by human ear. In technical terms, by using voice packets, the minimum CIR / remote was able to decreased until 300bps which was an acceptable value. In addition the data from DAS to CRC as they treated as voice packets, they had the highest priority among any other traffic. Finally BBD mode selected for each station in order to cover the possibility that some station needs higher bandwidth due to retransmission of lost data packets.

The satellite carrier that selected was the Hellas-Sat. Based on EuroStar E2000+ launched at 2003 at 39^{O}E and expected to have 18 years of operation. Its coverage is depicted in Fig.3.20

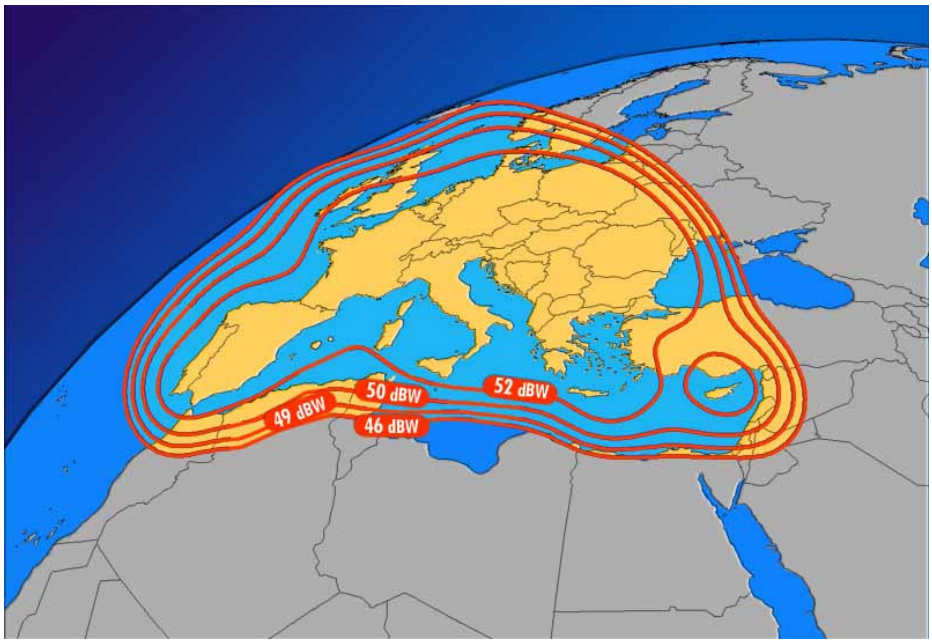


Figure 3.20: Hellas-Sat coverage for F2 beam

The western location that a station is expected to install is the Rhodes island. For this reason a link budget for this location is estimated as shown in App.4b. Translation of the results leads to the following values for the equipment in CRC as well as for the remote stations (Table 3.3). Schematic representation of the installation presented in Fig. 3.21

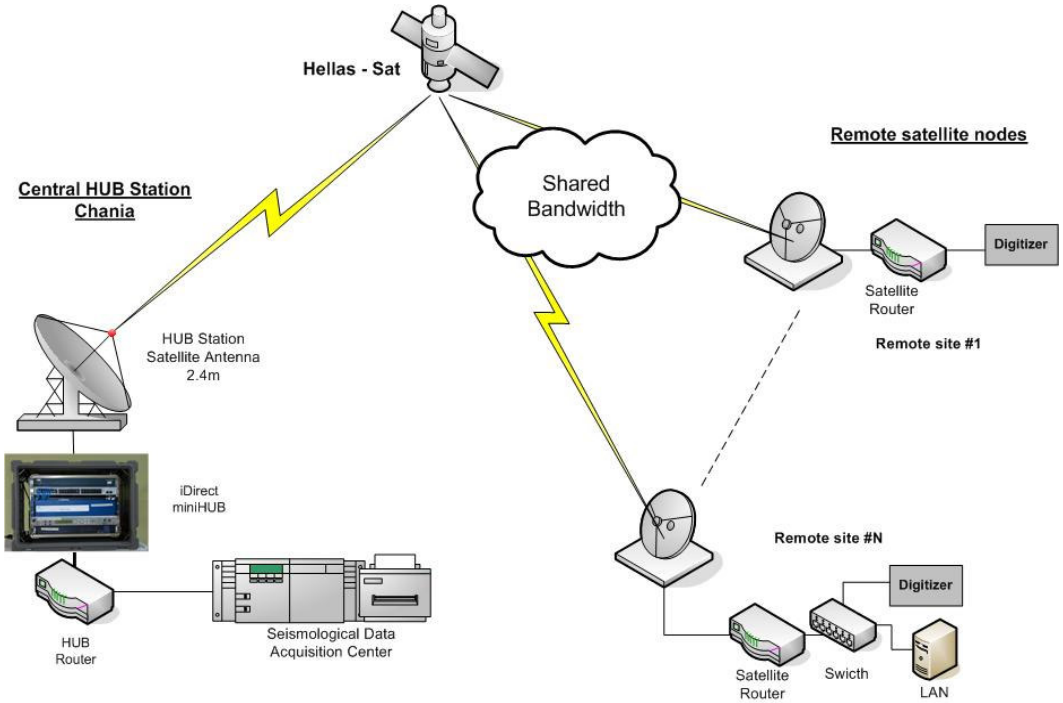


Figure 3.21 – Satellite network topology

CRC main station	*Remote station*
2.4m diameter VSAT antenna	1.2m diameter VSAT antenna
Amplifier 4W (requirements dictates 1W but increased for future use)	Amplifier 2W

Table 3.3 : VSAT equipment details

3.3 Data Collection and Monitoring

The network topology that installed at LGS guided from the following main necessities:

- Continuous data recording
- Stability and easy of use for the testing new algorithms
- Adaptation to the different conditions of exploitation and hardware architecture.

These lead to a scheme with three data service point where the first one receives and stores all the data from SNC and collaborating networks (Primary server), the second one acts as a backup to the first one (Secondary server) and the third (Diskloop) is the point where the seismological signal processing is done as depicted in Fig. 3.22.

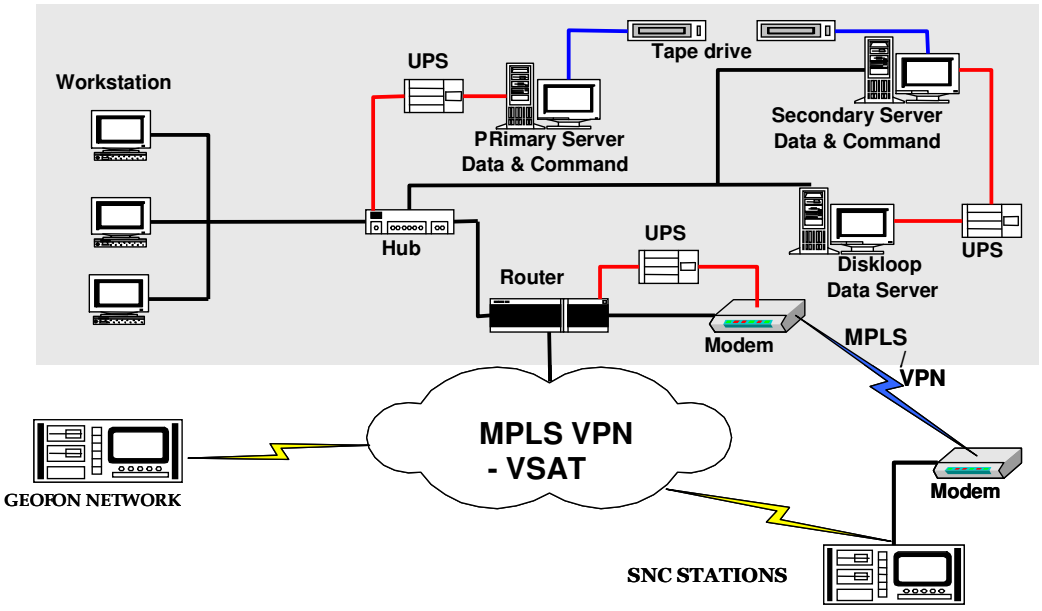
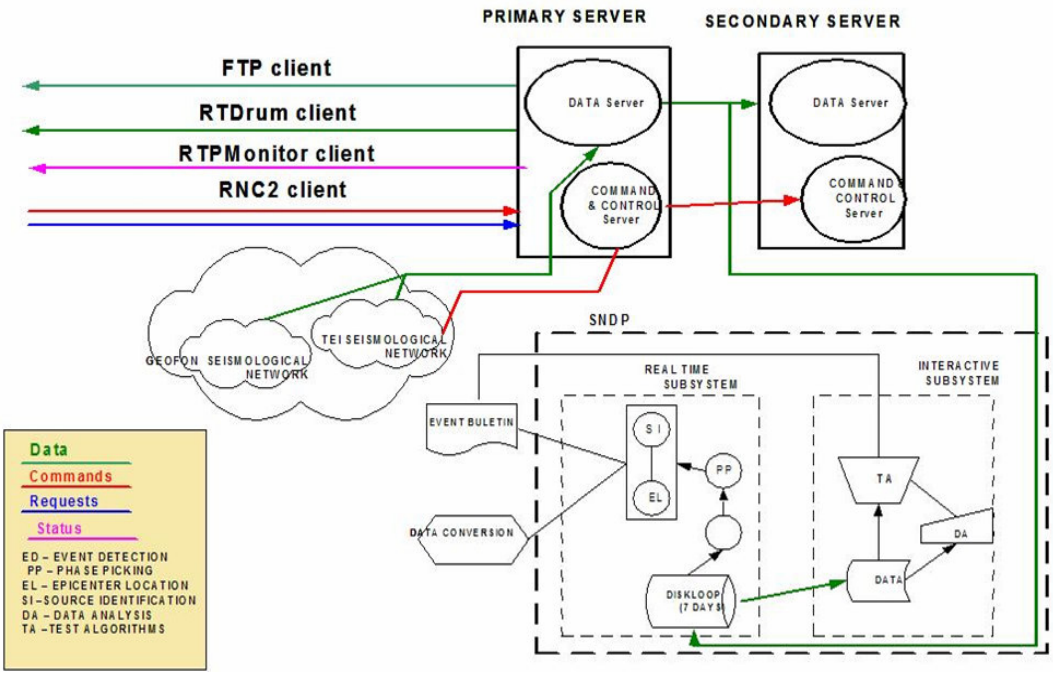


Figure 3.22. Block diagram of the LGS configuration located at Chania

The data collection software is a commercial product (*rtpd* from Reftek inc) which is the implementation of RTP protocol. It runs on Linux OS and is responsible for collecting and archiving the data from the stations, performing quality control of them as well as sending acknowledgement packets back to stations. As a server it can respond to requests from clients, receive data from the stations, resend the data to other collaborating data centers and providing the monitoring packets for further processing. Communication with the DASes is performed by using the port 2543 for data transfer and port 5000 for commands and monitoring purposes. That means that DASes are able to send the data while receiving commands or sending at the same time their status. This scheme except the obvious speed benefits (since it has a dedicated to the data, channel) provides an alternative way to reach the remote DAS (in case of a data send or RTP failure) by using the port 5000 with conventional TCP/IP connection. All these are depicted in Fig. 3.23.



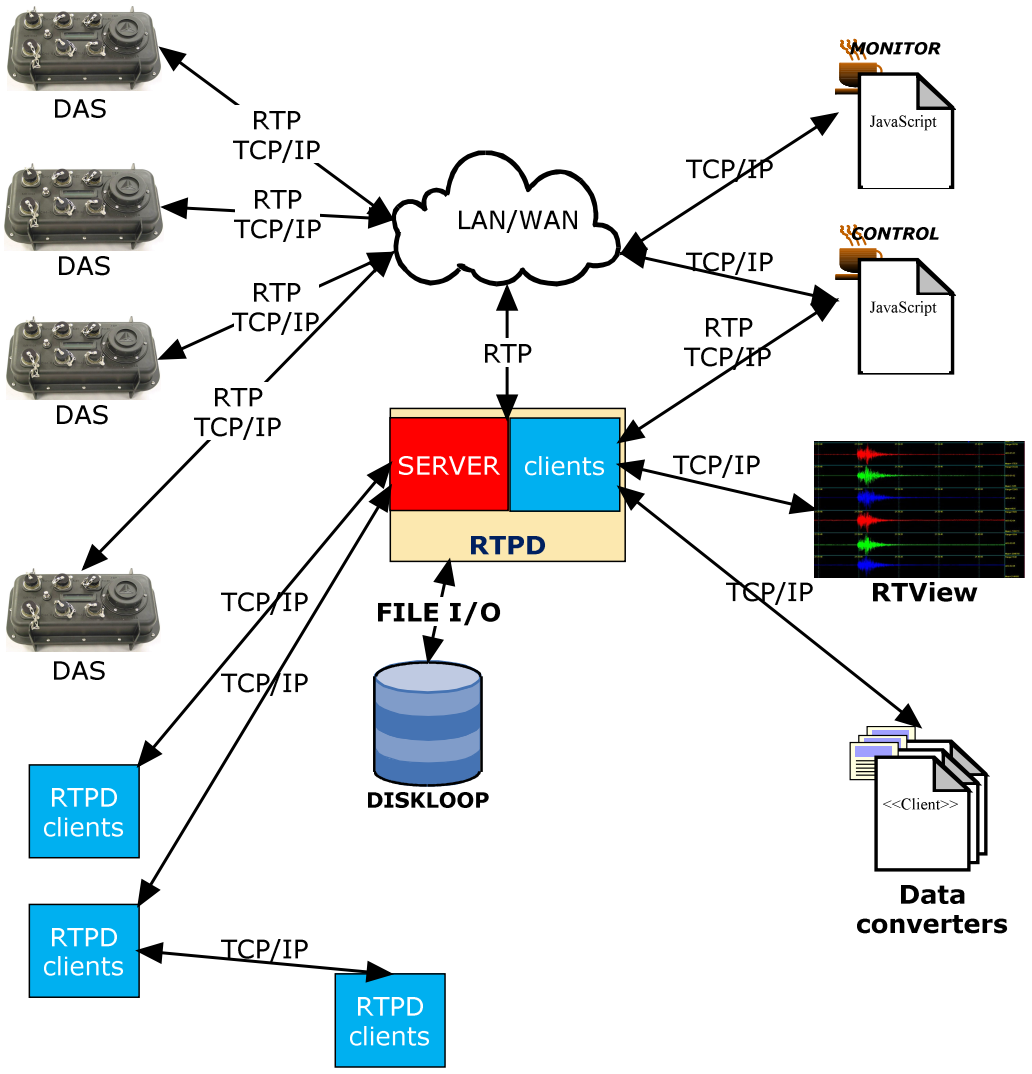


Figure 3.23 Schematic representation of the connectivity features (this page) and processes involved at CRC of LGS (previous page) for *rtpd* server.

The software that provided by the manufacturer for monitoring and controlling the DASes was a specialized Java application (RNC2) as well as a command line translator (rtcc). Both are not very user friendly so there was a need for developing simple tools for control and monitor. The new applications that written by the author are based on dynamic web pages which can pass the necessary command line arguments to the *rtcc* translator by using XML statements. A typical monitoring web page is shown is Fig.3.25 while flowchart diagram shown in Fig. 3.24.

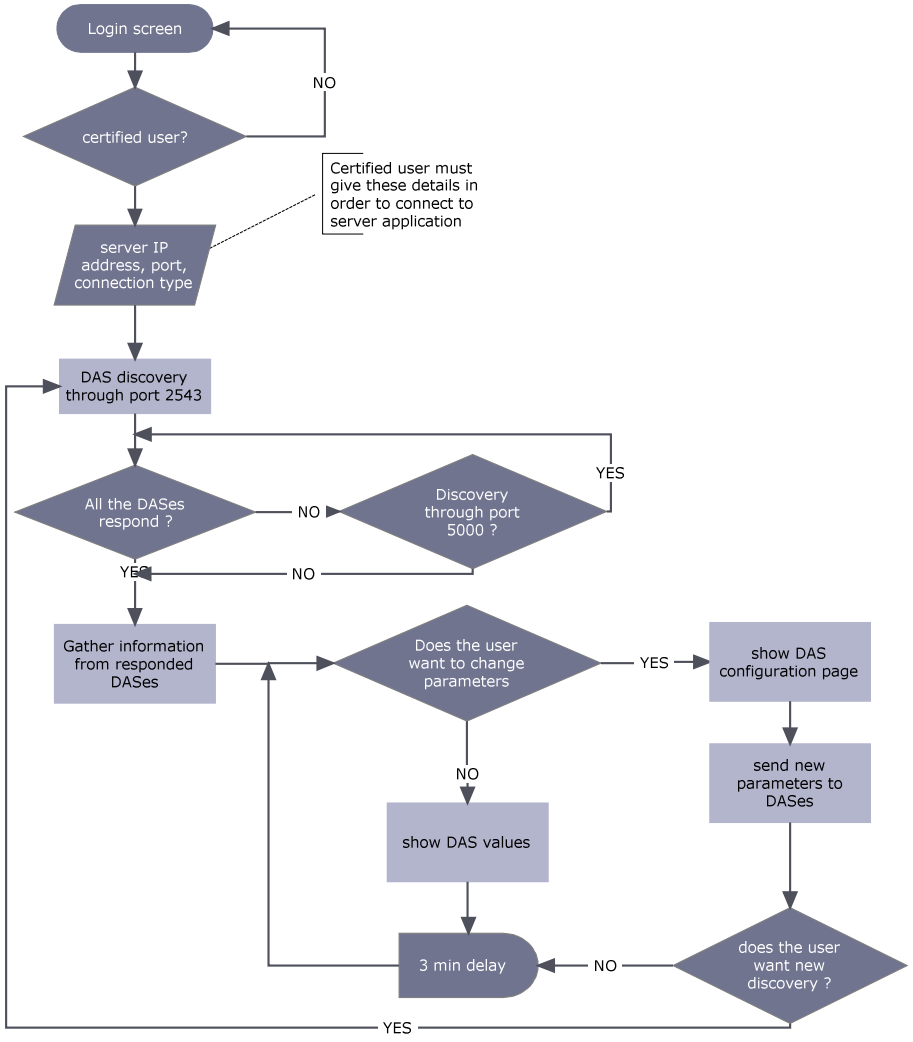
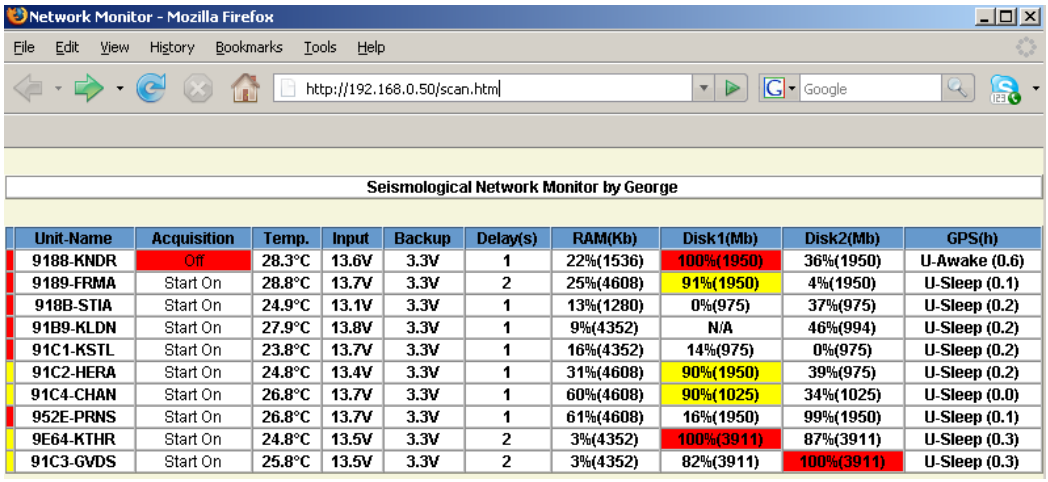


Figure 3.25: Flowchart of monitoring and configuration software module



Unit-Name	Acquisition	Temp.	Input	Backup	Delay(s)	RAM(Kb)	Disk1(Mb)	Disk2(Mb)	GPS(h)
9188-KNDR	Off	28.3°C	13.6V	3.3V	1	22%(1536)	100%(1950)	36%(1950)	U-Awake (0.6)
9189-FRMA	Start On	28.8°C	13.7V	3.3V	2	25%(4608)	91%(1950)	4%(1950)	U-Sleep (0.1)
918B-STIA	Start On	24.9°C	13.1V	3.3V	1	13%(1280)	0%(975)	37%(975)	U-Sleep (0.2)
91B9-KLDN	Start On	27.9°C	13.8V	3.3V	1	9%(4352)	N/A	46%(994)	U-Sleep (0.2)
91C1-KSTL	Start On	23.8°C	13.7V	3.3V	1	16%(4352)	14%(975)	0%(975)	U-Sleep (0.2)
91C2-HERA	Start On	24.8°C	13.4V	3.3V	1	31%(4608)	90%(1950)	39%(975)	U-Sleep (0.2)
91C4-CHAN	Start On	26.8°C	13.7V	3.3V	1	60%(4608)	90%(1025)	34%(1025)	U-Sleep (0.0)
952E-PRNS	Start On	26.8°C	13.7V	3.3V	1	61%(4608)	16%(1950)	99%(1950)	U-Sleep (0.1)
9E64-KTHR	Start On	24.8°C	13.5V	3.3V	2	3%(4352)	100%(3911)	87%(3911)	U-Sleep (0.3)
91C3-GVDS	Start On	25.8°C	13.5V	3.3V	2	3%(4352)	82%(3911)	100%(3911)	U-Sleep (0.3)

Figure 3.24: DAS monitoring web page (top) and configuration page (bottom)

By hyperselecting each DAS the user has the ability to derive the individual settings for each one and then (if needed) to pass the new parameters to DAS flash EPROM. Since the whole procedure is based in HTML methods the obvious advantage is that it can be completed from any device is capable to run a web browser. This is very important in the cases where the administrator cannot have physical access to the CRC because he can control the network from outside using only a web based device (laptop, handheld, palmtop, mobile phone).

3.4 Data processing

3.4.1 Introduction

For the purposes of seismic signal processing, a multitasking, scalable, problem–oriented and open–architecture software designed for UNIX computers/networks called SNDP was installed. The SNDP is intended for interactive and automatic processing of multichannel seismological data recorded by the seismic arrays or regional networks. It's comprised mostly of programming and graphic facilities similar to the well-known program packages for seismic monitoring (e.g. the EP system [Fyen, 1989] at NORSAR or the IMS [Intelligent Monitoring System]; Bach et al., 1990 at the Center for Monitoring Research).It is based on problem-oriented configuration of the system with a user-friendly interface for easy and quick installation of new user's data analysis procedures, thus implementing the open architecture concept. It includes a Job Control Language (JCL), a problem-oriented language, which enables the user to perform complex processing sessions in a manner of an Expert System.

3.4.2 Description

The SNDP consists of two subsystems:

- The Real-Time Subsystem (RTS) for multi-channel data collection, storage in a diskloop, and signal processing of the data flow. The major part of RTS is DP (real-time detectors) module.
- An off-line Seismic Analysis Subsystem (SAS) for automated interactive processing of registered seismograms with more sophisticated and time-consuming procedures (for example: statistically optimal multi-channel data analysis).

The block diagram of SNDP is shown in Fig.3.26

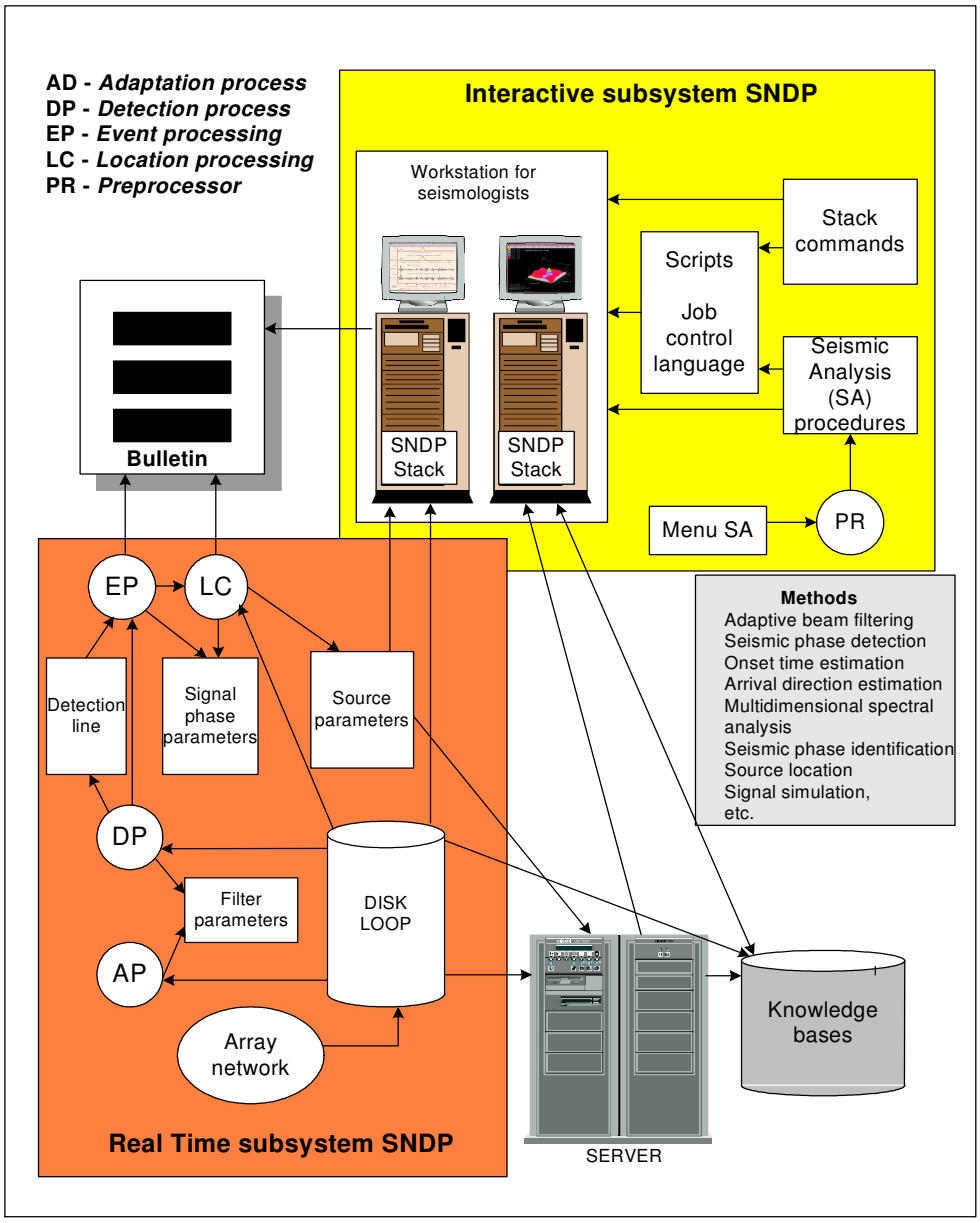


Figure 3.26: Block diagram of seismic processing system (www.synapse.ru)

The main problem, arising in designing such a system as described in the introduction, is to provide the reliable performance of application procedures comprised in the system in conditions of variable intensity of data streams and activity of system users. In the RTS of the SNDP this problem is solved by utilizing modern programming methodology. This involves the following general approaches and facilities:

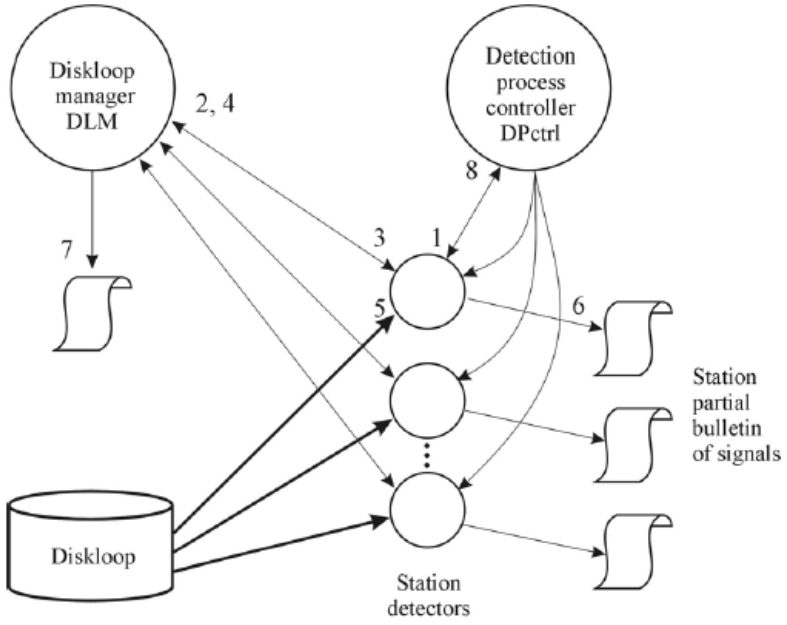
- Up-to-down system design, i.e. dividing the system into a set of abstract levels. As a result, a clear logic, free of local details, has been provided for each level.
- Free asynchronous communication between different system processes is based on the port method. Each process has to know nothing about other system processes. It should possess a single input port with a queue facility and one or several output ports. The process scans its input queue, executes requested actions, and sends results to the output ports.
- Semaphore facilities to control all of the shared computer resources.
- A unique System Process Table (SPT) in which all process ports and their priorities are defined. To install new processes in the system, one needs only to expand the SPT.
- Provision of a single way for modification of the system. This avoids dualism in the system. Structures of the system, sharing memory fields, values of system's parameters, etc., must be defined in macros. Modification of the macros provides the correct changes of all system programs and processes. For example, the SPT may also be contained in its own macro.
- A specific Global System Table (GST) provides flexibility of system handling and easy modification of the system while including into it new tasks, methods, algorithms, and facilities. The GST is contained in the system's shared memory and comprises an arbitrary amount of named structures. A catalogue is placed at the beginning of the GST. It provides a direct access to the named structures. Each structure is defined by a special macro, so as a result the GST is a set of such macros. Each individual system process uses only those macros that it needs for execution. When started, the user's program requests addresses of the named GST structures from the system. To install a new user's program it is sufficient to expand the GST by including the new program macro in it.

The processing stages that involved the above characteristics are:

- Detection of seismic events
- Identification and parameter estimation of seismic phases
- Location of event sources
- Identification of event types

3.4.3 Detection of seismic events

Real time detectors, based on traditional STA/LTA algorithm, are intended for processing of continuous seismic data, incoming to LGS, from field stations as depicted in Fig.3.27



1 Start of the detection processes
2 Report about start
3 Messages from DLM about accumulation of the next data portion
4 Report about signals
5 Loading of the data fragment into the buffer
6 Printing out the station partial results onto console
7 Printing out the station common detected signal bulletin
8 Report to the calling process about procid

Figure 3.27: Schematic representation of detection process (www.synapse.ru)

- Detection of a signal is performed successively and independently within specified frequency bands.
- If the signal is found, just the first onset is selected within one frequency band, the subsequent phases are not considered.
- If one signal is detected by several independent methods, provision is made for selecting that one, which has the greater SNR.

3.4.4 Identification and parameter estimation of seismic phases

This process completed by Estimator module which is integrate the following tasks:

- Extraction of seismic phases of signals found by the Detector.
- Exact estimation of seismic phase parameters.
- Preliminary identification of seismic phase types to assist in procedures of:
  - Phase association to source
  - Determine epicenter coordinates
  - Depth
  - Magnitude of the source

3.4.5 Location of event sources and identification of event types

This complex process performs extraction and location of seismic sources according to results of detection and parameter estimation of seismic phases by stations of a regional network. After successful location each event is associated as regional or far source. The discrete steps that followed by this process are:

- Regional association of phases from stations by the node sorting method.
- Refinement of source coordinates by successively narrowing the area of analysis and by rejection of the wrong phase onsets.
- Estimation of source depth by scanning over the depth using depth travel time tables of regional phases.
- All the sources detected at the stage of regional association are represented in the form of a preliminary bulletin of events.
- Analysis of sources from the preliminary bulletin to reveal intersection of data obtained from several stations.
- Extracting and exclusion of partial solutions.
- Forming the final bulletin of detected events.

3.5 Dissemination of results

The software suites that discussed already had one major lack: they didn't provide tools or modules for the publication of results. For this reason a complete set of dissemination tools is designed and implemented. The main requirements were:

- Publication of results using conventional messaging techniques (email, sms)

- Produce of real time maps either as static images or at public platforms (Google Maps)
- Storage of the events as well as the relevant seismograms
- Notification of other collaborating institutes and public bulletins

As long as SNDP software provides the flexibility to implement new algorithms running in parallel with real time system, the scripts that created to fulfil the above requirements are implemented in SNDP as below.

Initially, there was a modification in regional associator module of SNDP (3.4.4) in order to produce output files according to our needs. By doing this, after each event two files were produced: a complete report (named *bull_out*) showing the results that need to be spread and a unique event id. The complete code including the modification is presented in App.4e.

A new relational database, named *seismos* (i.e. "earthquake" in Greek) is designed and implemented in order to hold the results of each detected event. The database built in MySQL and had two main tables as shown in Fig.3.28

Table Events	
Field name	*Data type*
Event_id	Int32
Date	Date/time
Time	Date/time
Lat	Int32
Lon	Int32
Depth	Int32
Mag	Int16
Auto_report	boolean

Table seismograms	
Field name	*Data type*
Event_id	Int32
Seismo_file	file
complete	boolean

Figure 3.28: Structure of relational database Seismos

The Events table holds the information needed in order an event be published in a bulletin. The table seismograms hold 4min of data for each event in order to provide the ability for manual processing by human analysts in a later time. The two tables were separated in order the fastest search speed when a external request about event's info is passed into. Usually other clients want to retrieve only the information

about an event and not the seismogram. The database is updated automatic and manually. When an event occurs, the program *steile* (which will be explained below) retrieves the values from *bull_out* file that needed to fill the tables. In this case the field auto_report receives the value 'Y' which means that is a record created by automatic estimation. Later, when a human analyst is going to provide a more accurate analysis, his results are stored also to the table, having the auto_report valued as 'N'. This is very useful because there is one database scheme that stores all the results which can be discriminated by checking the auto_report field. In secondary server there is also a replica of the main database.

The program *steile* (i.e. "send" in Greek) is responsible for the results handling. More specific, right after the estimation of an event, the *steile* receives the results and produces a series of actions with the following order:

i. An email, reporting all the details of the seismic event (Latitude, Longitude, Magnitude, time), the parameters used for the estimation as well as the detected phase arrivals, is generated. This email is sent to specific users. Typical form of this email is shown in Fig.3.29

Event N 10476

Date	Orig.t	var_t	Lat	Lon	h	N_st	N_as	N_ph	Rls	Dist	XMag
2007/1/7	05:17:41.690	1.72	32.61	19.66	35.0	5	7	7	88	537	3.52

St	Chan	Dist	Phase	Time	TRes	Azim	Slow	SNR	Ampl	Per	Xmag
KNDR	HHZ	468	Pn	2007/1/7 05:18:44.279	0.3	59.6	4.4	2.2	481.6	0.10	3.52
KNDR	HHZ	468	Sn	2007/1/7 05:19:26.938	-2.4	45.9	8.6	2.1	501.7	0.07	-
KLDN	HHZ	490	Pn	2007/1/7 05:18:45.328	-1.4	168.9	13.7	7.5	18775.6	0.09	5.12
KTHR	HHZ	506	Pn	2007/1/7 05:18:49.607	0.8	264.4	14.0	2.5	270.4	0.04	3.48
PRNS	HHE	534	Pn	2007/1/7 05:18:54.677	2.3	120.1	16.2	1.8	365.3	0.08	3.46
PRNS	HHE	534	Sn	2007/1/7 05:19:45.985	2.1	187.3	18.1	2.8	733.5	0.07	-
FRMA	HHN	632	Pn	2007/1/7 05:19:01.917	-1.6	161.6	13.1	6.7	26720.0	0.07	5.48

==========

Laboratory of Geophysics & Seismology (LGS)
Seismological Network of Crete (SNC)
Email alert with automatic Preliminary solution

Figure 3.29: Email reporting an event

ii. Then, another email is formed and send it to European-Mediterranean Seismological Centre (www.emsc-csem.org) sending only the phases in order to produce the most accurate hypocentral location estimation since it collects data from more than 35 seismological networks around Europe. The form of this email has the structure presented in Fig.3.30

```
SNC-LGS reports:

KNDR  HHZ  Pn    2007/1/7 05:18:44.279
KNDR  HHZ  Sn    2007/1/7 05:19:26.938
KLDN  HHZ  Pn    2007/1/7 05:18:45.328
KTHR  HHZ  Pn    2007/1/7 05:18:49.607
PRNS  HHE  Pn    2007/1/7 05:18:54.677
PRNS  HHE  Sn    2007/1/7 05:19:45.985
FRMA  HHN  Pn    2007/1/7 05:19:01.917
```

Figure 3.30: Email reporting an event to EMSC

iii. Next, the database seismos is updated automatically. At this stage, first the table events is filled and then the table seismograms. There is a possibility that not all the stations are operative during the event or some station presents a delay. In this case the seismograms table will not complete and the complete field is going to receive the value 'N'. Later, when the analyst is going to perform manual analysis, the analysis software is configured to first check the value of report filed. If is 'Y' (which means that all the data are present) it proceed to the usual procedures, otherwise it requests data retransmission from the diskloop. The result that the analyst produces is written to the database with different event_id (it has the form YYYYMMDDDHHMMSS) and with auto_report field set to 'N'. This is useful when on wants to discriminate the results produced from automatic system and human analyst.

iv. Next the mapping application receives the results. The mapping application consists of two main parts: static and dynamic. The static mapping application produces a jpeg graphic file which can be embedded to MMS messages or to static web pages. A typical file of this type depicted in Fig.3.31.

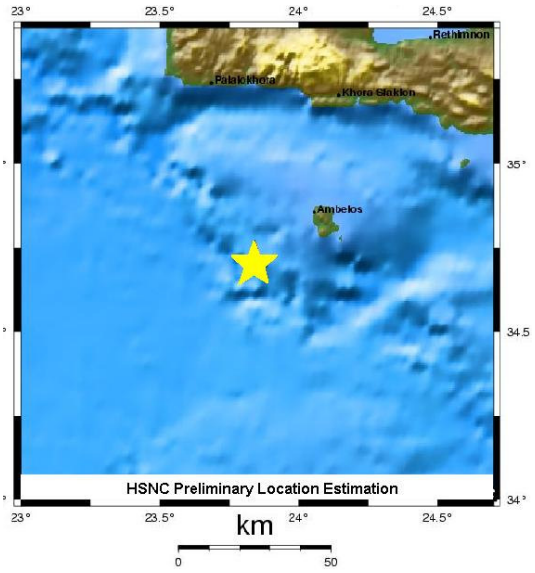
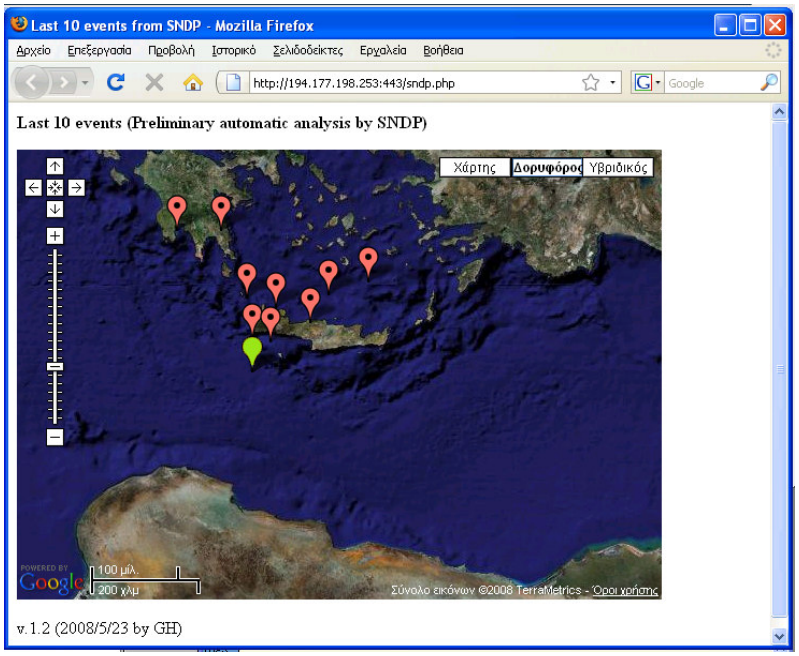


Figure 3.31 : Static graphic file produced by dissemination software

The dynamic mapping consists of automatically producing the code needed for the generation of a Google Map (maps.google.com). The code is written in Java and PHP. The application connects to database seismos, requests the last 10 events and produce them in a Google Map. The last event is indicated and every one of the 10 events can present its details (Lat, Lon, e.t.c) when the corresponding icon is pressed. Two snapshots of the application is shown in Fig.3.32.



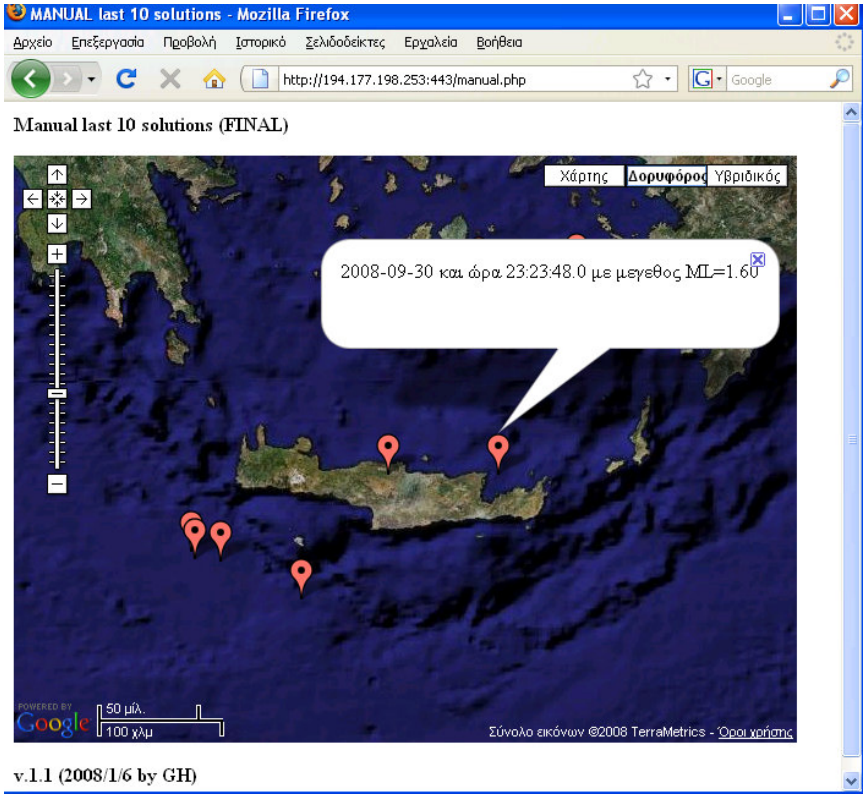


Figure 3.32: Dynamic mapping application (GoogleMaps) in general view (top) and in interactive (bottom) view.

v. At last step, a SMS message is generated and send to predefined users. The text of the message gives very briefly the main details of the event such as : `HSNC-LGS Earthquake alert (PRELIMINARY) <--> Date=2006/1/8, Time=22:01:43, Lat=34.23, Lon=22.12, Mag(L)=4.7.`

3.6 Experimental details on ambient noise measurements

This section describes the experimental details of field measurements for ambient noise, the data acquisition scheme as well as basic guidelines that followed through measurements. Most of the guidelines are derived from the results of SESAME project (see http://sesame-fp5.obs.ujf-grenoble.fr/SES_TechnicalDoc.htm) and Mucciarelli (1998).

3.6.1 Data acquisition guidelines

The experimental parameters play a significant role at HVSR technique because up to now there is not a general framework for the data collection and processing of ambient noise measurements. In this section the guidelines from some complete studies (Mucciarelli 1998, Parolai et.al. 2001, SESAME project 2004) will be presented.

The experimental parameters that affect the HVSR data collection procedure are classified in three main categories:

- Parameters related to data acquisition
- Parameters related to field measurement
- External parameters

Studies that examined the above parameters concluded to the following guidelines that followed during our measurements:

- Use of accelerometers is not preferable because they are very sensitive to frequencies below 1Hz
- Sampling frequency of 50Hz is enough since the maximum frequency that is useful for the engineers is up to 25Hz.
- Cable length between seismometer and data acquisition system must be less than 100m
- Use of seismometers with eigenperiod higher than 20sec is not preferable because they demand high relaxation time without providing better recordings
- Investigation of a site with expected frequency f_0 must be done with seismometers with eigenfrequency lower than f_0
- In order to have a reliable estimation of HVSR the following relation must hold when splitting the ambient noise recording into several time windows

$$f_0 > \frac{10}{t_w} \qquad (3.1)$$

where t_w is the length of each window in sec. It is recommended that the number of iterations $n_c = t_w \cdot n_w \cdot f_0$ should be more than 200 (i.e for an f_0 at 0.5Hz n_w=10 and t_w=40 or n_w=20 and t_w=20 but not n_w=40 and t_w=10). This fact is clearly depicted in Table 3.4

f_0 (Hz)	Minimum length of t_w (sec)	Minimum number of iterations n_c	Minimum number of windows n_w	Minimum useful signal (sec)	Minimum recording time (min)
0.5	20	200	10	400	20
1	10	200	10	200	10
2	5	200	10	100	50
5	5	200	10	40	3
10	5	200	10	20	2

Table 3.4 – Proposed parameter values for HVSR survey (SESAME project, 2004)

- A site's resonance frequency is characterized as clear peak if the amplitude of HVSR (A_0) at this point is more than 2

3.6.2 Field measurement guidelines

- Seismometers must be fully leveled and with direct coupling with the ground.
- The placement of seismometer at concrete or asphalt has no effect since no artificial peaks produced from 0.2Hz to 20Hz (Fig.3.33)
-

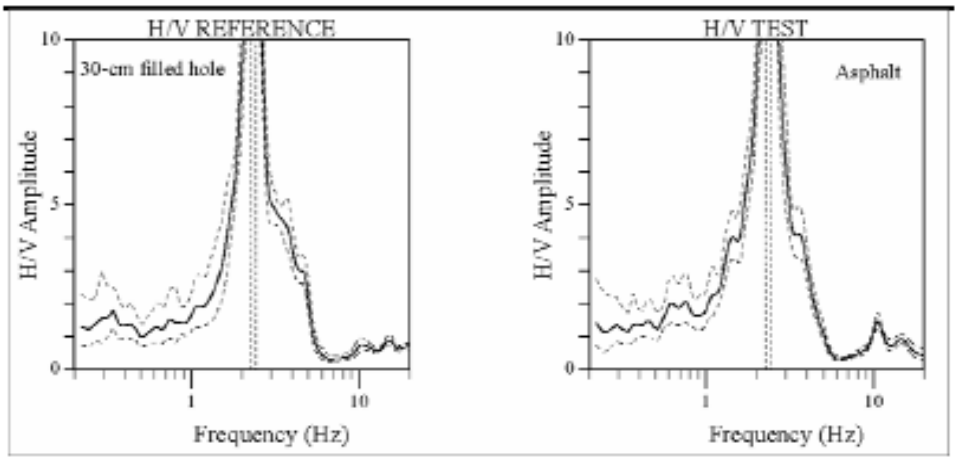


Figure 3.33. Comparison of HVSR curves with ambient noise that recorded with asphalt (left) and without asphalt (right) at the same site (http://sesame-fp5.obs.ujf-grenoble.fr)

- Measurements at soft grounds (mud, turf, wet ground) must be avoided.
- At grounds covered by snow or ice the seismometer must be installed above a metallic plate in order to avoid possible inclination due to liquefaction
- Measurements near buildings, trees or other similar structures must be avoided when wind with velocity larger than 5m/s is blow. Under these conditions the

movement of these structures could introduce low frequency domination into HVSR results (Fig. 3.34)

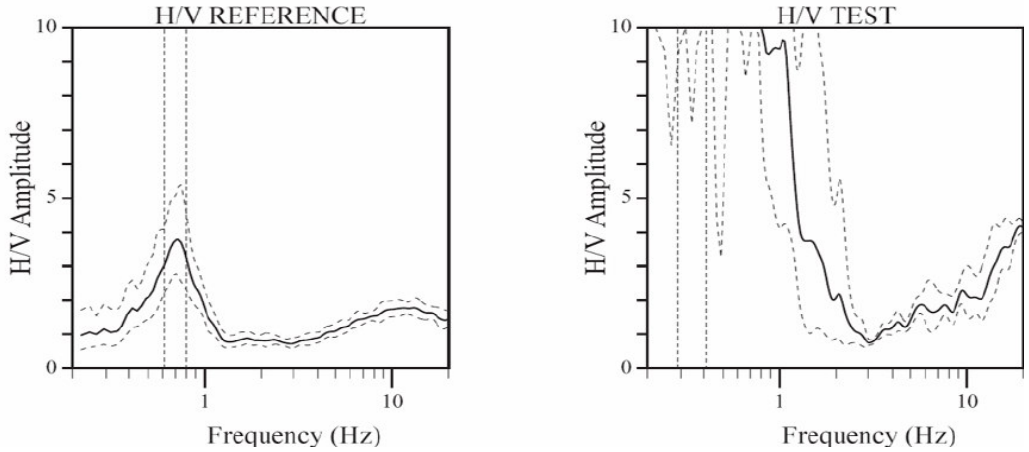


Figure 3.34. Comparison of HVSR curves with ambient noise that recorded with wind (left) and without wind (right) at the same site (http://sesame-fp5.obs.ujf-grenoble.fr)

- Measurements above garages, tunnels, large tubes and other similar basement structures must be avoided because we have undesirable vitiation of fundamental's frequency amplification value (Fig. 3.34).

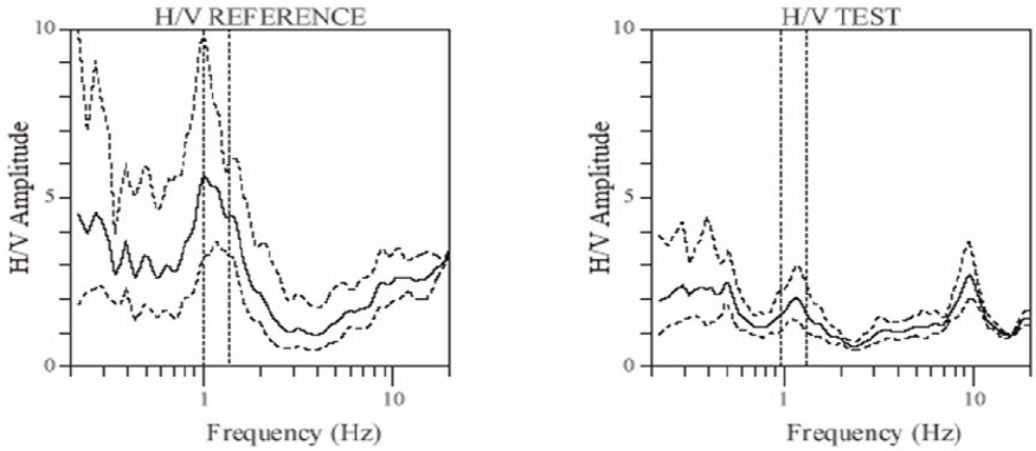


Figure 3.35. Comparison of HVSR curves with ambient noise that recorded 30m away from a basement structure (left) and above it (right) at the same site (http://sesame-fp5.obs.ujf-grenoble.fr)

3.6.3 External parameters

- Seismometer must be protected from wind even for low velocities (>5m/s).
- Measurements under strong raining must be avoided. Under light raining no problem is reported as long as the seismometer is covered.(Fig. 3.36)

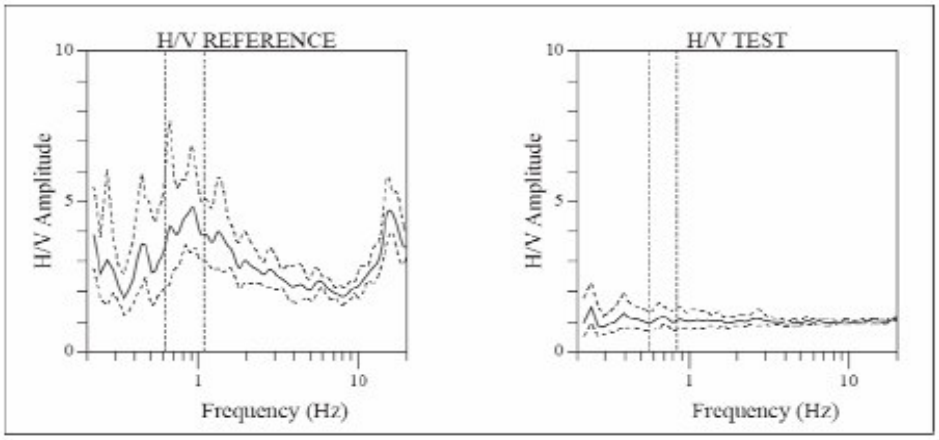


Figure 3.36. Comparison of HVSR curves with ambient noise that recorded under strong rain (left) and without rain (right) at the same site (http://sesame-fp5.obs.ujf-grenoble.fr)

- Measurements near single-colored noise sources (industrial units, generators, pumps, motors) must be avoided.
- All the temporary small period perturbations (human walk, cars and motorcycles) could influence the HVSR results. In case that these perturbations could not rejected the duration of the measurement must be increased (Fig.3.37)

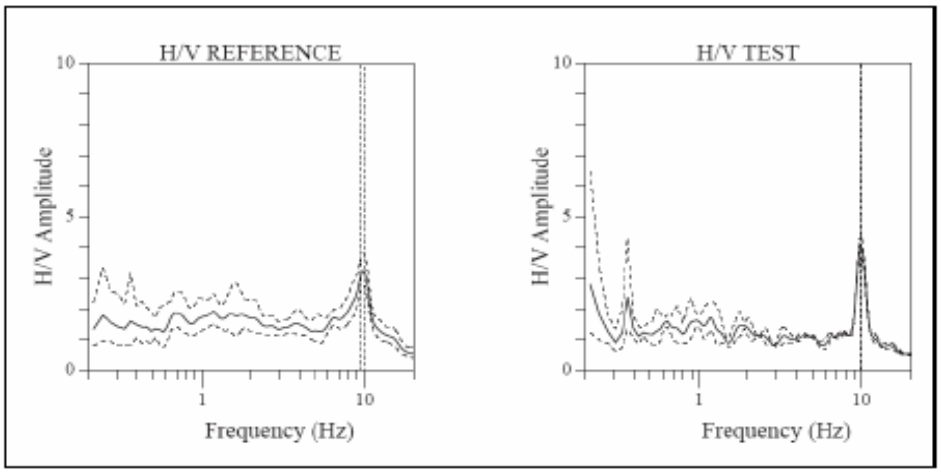


Figure 3.37. Comparison of HVSR curves with ambient noise that recorded with people walking at different distances near seismometer (left) and without people (right) at the same site (http://sesame-fp5.obs.ujf-grenoble.fr)

3.6.4 Data processing

In the current section the necessary steps for the estimation of HVSR will be presented. These steps will be performed by public available (http://www.geo.uib.no/seismo/REPORTS/SESAME/J-SESAME/) J-SESAME software (Atakan et al., 2004a,b, Koller et. al 2004) which is also described.

A key part to HVSR estimation is the selection of the most stationary (aka "quiet") windows N_i in order to ensure that the estimation includes only far away sources. Short period transients, high amplitude non-stationary noise and any other perturbations as described before must be excluded form the ambient noise signal. There are two approaches to achieve this: manually and automatic by using an antitrigger algorithm.

Besides the manual selection directly from the screen, which is often the most reliable, but also the most time consuming, an automatic window selection module has been used in view of processing large amounts of data. The objective is to keep the most stationery parts of ambient noise and to avoid the transients often associated with specific sources (walks, close traffic). This objective is exactly the opposite of the usual goal of seismologists who want to detect signals, and have developed specific "trigger" algorithm to track the unusual transients. As a consequence, we have used here an "antitrigger" algorithm which is exactly the opposite: it detects transients but it tries to avoid them.

The procedure to detect transients is, very classically, based on a comparison between the short term average (STA), i.e., the average level of signal amplitude over a short period of time, (typically around 0.5 to 2.0 s), and the long term average "LTA", i.e., the average level of signal amplitude over a much longer period of time, denoted (typically several tens of seconds). When the ratio STA/LTA exceeds an a priori determined threshold (typical values are between 3 and 5), then an "event" is detected.

In our case, we want to select windows without any energetic transients: it means that we want the ratio STA/LTA to remain below a small threshold value S_{max} (typically around 1.5 – 2) over a long enough duration. Simultaneously, we also want to avoid noise windows with anomalously low amplitudes: we therefore also introduce a minimum threshold S_{min}, which should not be reached throughout the selected noise window.

There are also two other criteria that may be optionally used for the window selection:

- One may wish to avoid signal saturation – as saturation does affect the Fourier spectrum. J-SESAME looks for the maximum amplitude over the whole noise recording, and automatically excludes windows during which the peak amplitude exceeds 99.5 % of this maximum amplitude.
- In some cases, there might exist long transients (for instance related to heavy traffic, trains and machines) during which the STA/LTA will actually remain within the set limits, but that may be not representative of real seismic noise. Another option has therefore introduced to avoid "noisy windows", during which the LTA value exceeds 80% of the peak LTA value over the whole recording.

The software automatically looks for time windows satisfying the above criteria; when one window is selected, the program looks for the next time window, and allows two subsequent windows to overlap by a specified amount (default value is 10%). A typical example using data collected in the frame of SE-RISK project, is shown in Fig. 3.38

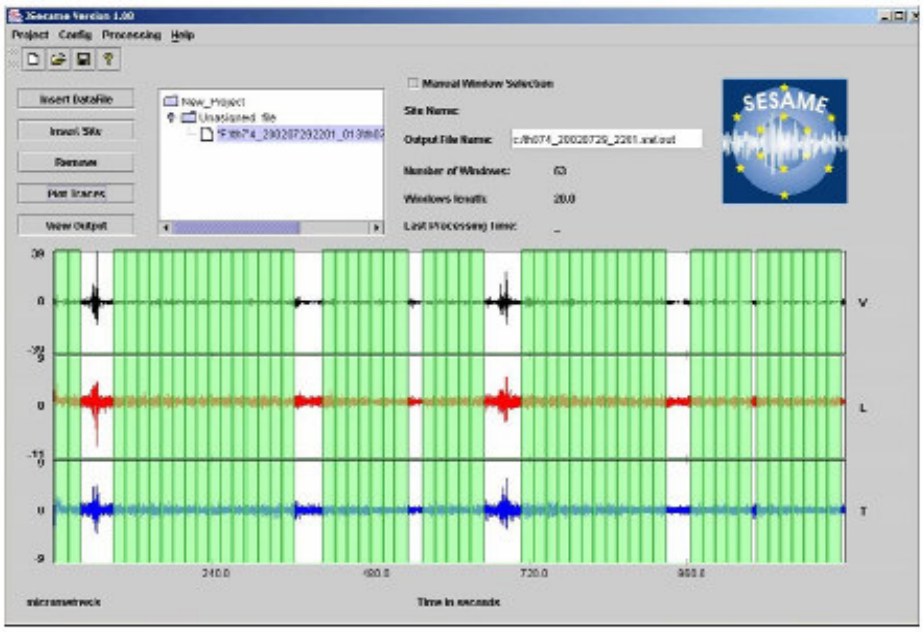


Figure 3.38. Selected “quiet” windows of ambient noise (at green)

After the selection of appropriate windows the following estimation procedure is followed where the typical values are those that suggested as defaults:

- DC offset removal
- Cosine tapering of each window (usually 5%)
- Spectrum calculation of each component NS, EW (horizontals) and V (vertical) for each window N_i. At Fig.3.39 we plot the results from the spectrum calculation for window N_{11} (which is highlighted with red color).

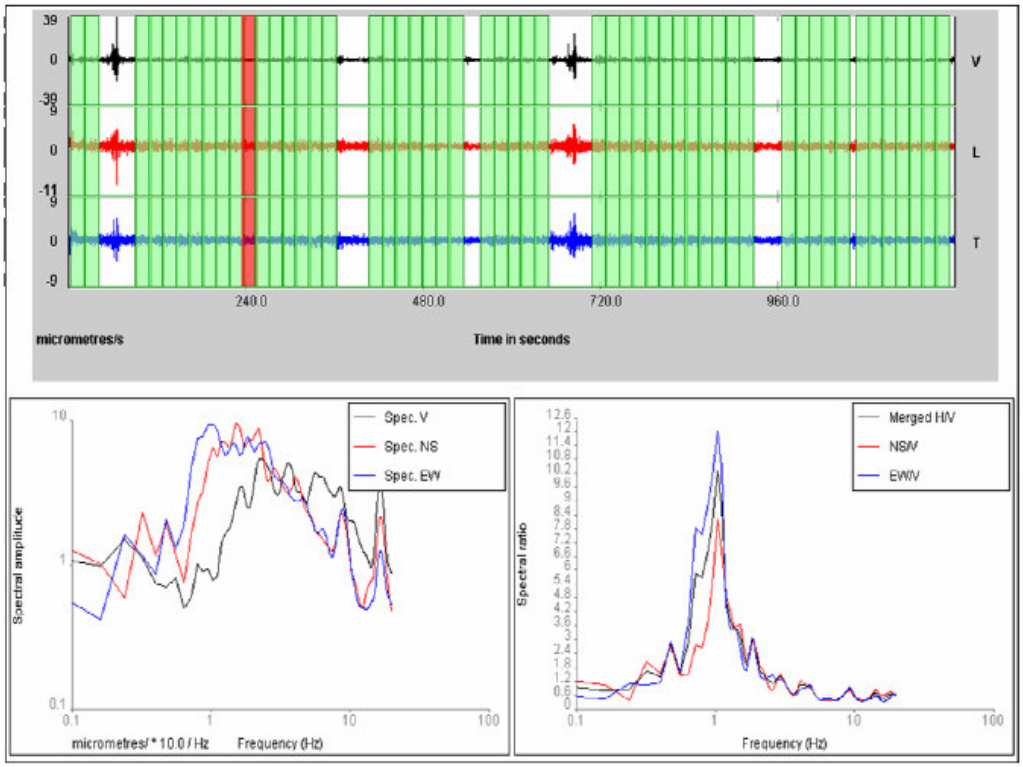


Figure 3.39. HVSR results for window N_{12} (upper part - highlighted red). Spectrum of each component NS, EW, V (lower part-left) and their estimated ratios H/V, NS/V, EW/V (lower part-right)

- Spectrum Smoothing. The raw values of Fourier spectra are smoothed (for each window N_i) before the computation of H/V ratio using the Konno-Ohmachi (1998) method. This specific method is preferred because the smoothing window function can be adapted to the frequency range that we are interested providing a smoothing window function constant in a variable logarithmic frequency range. When the smoothing window requires values of the raw spectrum outside the range 0 - $F_{Nyquist}$, then the smoothed spectrum will be restricted to the significant frequency interval.

- Calculation of geometric mean of horizontal components $H(N_i)$spectrum. For each window N_i the mean is calculated as :

$$H(N_i) = \sqrt{EW(N_i) \bullet NS(N_i)} \qquad (3.2)$$

- Calculation of HVSR for each window N_i :

$$HVSR(N_i) = \frac{H(N_i)}{V(N_i)} \qquad (3.3)$$

- Calculation of HVSR for all the selected windows :

$$HVSR = 10^{\frac{\sum \log_{10} HVSR(N_i)}{n'_w}} \qquad (3.4)$$

Where n'_w is the number of selected windows.

- Estimation of site's fundamental frequency f_0. At Figure 3.41 the f_0 depicted by grey vertical line among with its standard deviation (STD)

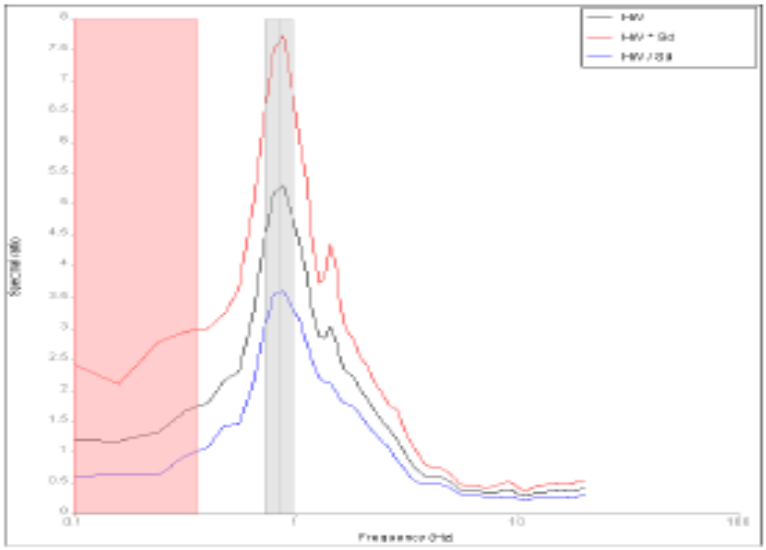


Figure 3.41. HVSR (black line), HVSR*STD (red line), HVSR/STD(blue line). f_0 (grey line) and f_0±mean STD (grey zone). Pink zone defines the frequency range where possible results are unreliable due to window length.

- Visual estimation of HVSR amplitude (A_0) that corresponds to site's fundamental frequency.

3.7 Summary

This chapter deals with the experimental details of seismological signal processing as well as some practical aspects and rules followed during this research. The author of this thesis was responsible for the design, testing, installation, deployment, maintenance and expansion of SNC at LGS. Design requirements, deployment processes and methods, implementation details, testing procedures and evaluation actions are presented. Today, SNC operates in real time and is able to provide automatic and manual estimations of earthquake events at South Aegean area. Its installation, improved the detection capability of the whole area since the interstation distance was decreased and the lower magnitude of completeness from 3.7 before installation, today is 3.2.

4. Wavelet-based Applications on Ambient Noise

4.1 Introduction

Disasters that can be caused by strong earthquakes have direct relation with local geological conditions due to ground motion. For better understanding of ground conditions direct results of ground motion during an earthquake are the best data. However this kind of approach has the disadvantage that can be applied mainly in areas with high seismicity. For this reason methodologies that rely on analysis of ambient noise recordings are very useful in order to estimate the influence of surface geological conditions.

Among these methodologies the Horizontal to Vertical Spectrum Ratio (hereafter HVSR) or Quasi Transfer Function, (QTF) that proposed by Nakamura (Nakamura, 1989) received great attention from the scientific community (although some different ideas support the theoretical background of the method on the wave type of ambient noise). This caused from the simple, rapid and economical data collection as well as from the direct results that can be produced according to the dynamic characteristics of the near surface geological structure. For this reason, HVSR method can be used for the characterization of a site before the installation of a seismological station.

This chapter describes the main concepts of HVSR technique and how it can be improved by means wavelet transform techniques.More specific, after the presentation of key concepts for HVSR, effects of long duration, low amplitude disturbances will be discussed since this is usually undetectable from thresholding algorithms. The examination of known ambient noise signals contaminated with synthetic disturbances is used in order to investigate the effects of these disturbances. Since are usually undetectable a wavelet projection is proposed as a preliminary test for the identification of these disturbances. The chapter finishes with some real examples and the improvements that obtained for the used of WT.

4.2 Ambient noise basics

4.2.1 Definition of Ambient noise

By the term ambient noise we define the continuous vibration of earth which is not caused by events of small duration like earthquakes or explosions. The spectrum of ground motion that is recorded at one point is the superposition of the characteristics of a source (source effects) and local geological conditions (site effects).

Dividing the sources that cause the ambient noise with respect to their spectrum we can summarize:

1. Long period noise (f < 0.3 until 0.5 Hz), generated by ocean waves at long distances. It is characterized from constant spectrum for several hours and related with climatic condition that existed above oceans.
2. Intermediate period motion (between 0.3-0.5Hz and 1 Hz) produced by sea waves near coast as well as from the wind
3. Small period motion(f > 1 Hz), produced by human activities.

Another separation, between long (T > 1s) και small (T < 1s) periods corresponds to the traditional separation between microseisms of natural generated sources and ambient noise of artificial – human activities. Although the limit of 1s between them can be shifted to higher periods (Seo, 1997) at urban areas which characterized by low frequencies and high impedance differences among subsurface layers.

4.2.2 Relation of ambient noise with surface waves

Ambient noise consists mainly from surface waves. Many researchers (among them Nogoshi & Igarashi, 1971, Field & Jacob, 1993, Lachet & Bard, 1994, Ansary et al., 1995, Horike, 1996, Tokimatsu et al., 1996, Konno & Ohmachi, 1998) agree to the following conclusions (Bard, 1999):

- The shape of the H/V spectral ratio related with ellipticity of Rayleigh waves because they dominate at vertical component
- Ellipticity depends from the frequency and presents a sharp peak near the fundamental frequency for sites that characterized from great impedance difference between upper and deeper layers.

Some researchers endorse the opinion that these results hold good only in simple structures (*Horike*, 1996) but from the above studies arised the ellipticity of curves for Rayleigh waves at sites with different and complex velocity profiles. These results are enough in order to extend above assumptions to every single dimension velocity profile

The transfer function S_T of surface waves defined as

$$S_T = S_{HS} / S_{HB} \qquad (4.1)$$

where S_{HS} is the spectra of horizontal noise component at surface and S_{HB} is the spectra of horizontal noise component that incidence from bedrock layers to surface layers. The S_{HS} mainly affected by surface waves. As long as the ambient noise propagates mainly as Rayleigh waves there is a possibility that S_{HS} influenced by that. The influence of Rayleigh waves must be included at vertical component of noise E_{VS} at surface but not included at vertical component E_{VB} at bedrock. Assuming that the vertical component of ambient noise is not amplificated from upper sedimentary layers, the E_S ratio

$$E_S = E_{VS} / E_{VB} \qquad (4.2)$$

depicts the influence of Rayleigh wave: If there is not such an influence then E_S=1. The E_S ratio has value greater than 1 if Rayleigh wave influence increase.

The S_T/E_S ratio can be considered as a reliable estimation of a transfer function, called S_{TT} defined as follows:

$$S_{TT} = S_T / E_S = R_S / R_B \qquad (4.3)$$

Where $R_S = S_{HS} / E_{VS}$ and $R_B = S_{HB} / E_{VB}$.

At bedrock layers the propagation of seismic movement is the same for all directions. That means R_B=1 and from Eq.4.3 we have $S_{TT}=R_S$. That means that the transfer function of surface layers can be estimated by using only ambient noise.

The same methodology can be followed for the estimation of maximum amplification: the horizontal amplification from the surface layers can be estimation from the ration between horizontals and vertical amplification ratios.

Noise at deeper layers assumed that is the same for all directions because multiple wave motions repeat a local diffraction after noise that generated from an earthquake. For this reason it is possible the estimation of local effects by using the recorded ambient noise at surface.

4.2.3 Relation of ambient noise with body waves

Separating the wavefield of ambient motion into surface waves (with index s) and body waves (index b) we could derive the following relations:

$$S^{NH}(f) = S_b^{H}(f) + S_s^{H}(f) = H_T(f) \,.\, R_b^{H}(f) + S_s^{H}(f) \qquad (4.4)$$

$$S^{NV}(f) = S_b^{V}(f) + S_s^{V}(f) = H_V(f) \,.\, R_b^{V}(f) + S_s^{V}(f), \qquad (4.5)$$

where R_b^{H} (and R_b^{V} respectively) is the horizontal (vertical respectively) spectrum of ambient noise body wave at reference station

After calculations, the H/V ratio between S^{NH} and S^{NV} could be expressed as:

$$A^{NHV} = [\, H_T \,.\, A_r^{NHV} + \beta \,.\, A_s \,] \,/\, [\, V_T + \beta \,] \qquad (4.6)$$

where

- A_r^{NHV} is the H/V spectrum ratio of ground motion in bedrock
- β is the relational analogy of surface noise waves from vertical component namely $\beta = S_s^{V}(f) \,/\, R_b^{V}(f)$
- A_s is the ratio between horizontal to vertical component of surface waves, namely $A_s(f) = S_s^{H}(f) \,/\, S_s^{V}(f)$

Studies of many ambient noise recordings from different stations show that it consist from body waves as well as from surface waves. However, up to now there is no complete theory that explains the nature of ambient noise wavefield.

4.3 The Nakamura method

Determination of dynamic characteristics of surface geological topologies can be achieved by using data from boreholes. Although it offers great accuracy, borehole method has a great disadvantage: in order to cover a wide area we need a number of boreholes, close to each other. This leads to increasing cost, a considerable amount of involved people and finally a quite big time period until the extraction of safe results. Aiming to the development of a rapid and effective method for the estimation of ground response during an earthquake, Nakamura proposed the HVSR (Horizontal to Vertical Spectral Ratio or QTS-Quasi Transfer Spectra) method based on ambient noise recordings at 0.5 – 20 Hz frequency range. This range includes all the noises that generated from natural and human activities under the assumption that the impact of each independent source is limited.

At Fig.4.1 we depict the differences at recordings between different earthquakes at different stations. As we can notice recordings at different stations even for the same earthquake produces different results. Besides that we can observe that recordings from different earthquakes are quit similar for each station because of the geological conditions of each area. In other words, concerning the dynamic characteristics, the influence of surface geology seems to be important than other involved factors.

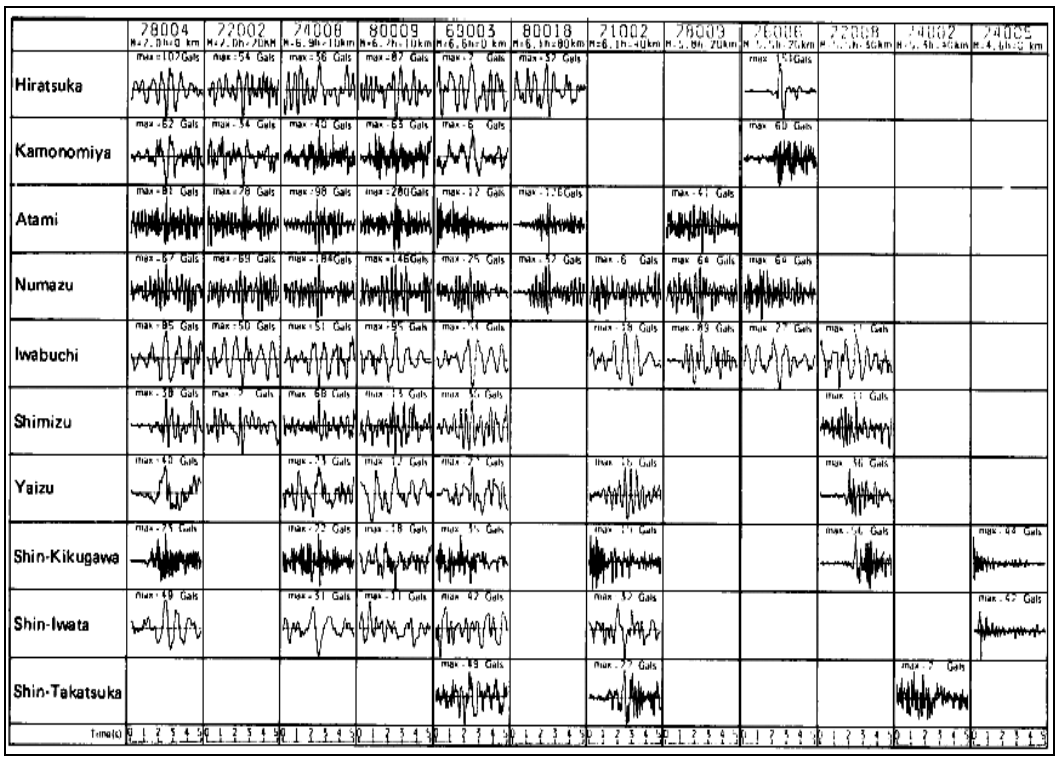


Figure 4.1 Accelerograms from different earthquakes at different stations (Nakamura, 1989).

For the forthcoming analysis (*Nakamura*, 2000) we consider a typical geolocigal structure of a sedimentary basin as depicted in Fig.4.5

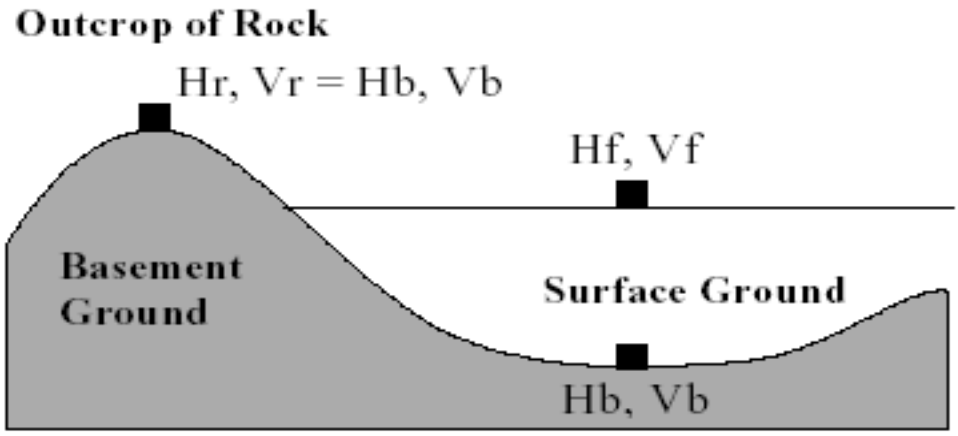


Figure 4.5 Typical geolocical structure of a sedimentary basin (Nakamura, 2000)

We assume that ambient noise consist of Rayleigh and other waves. The horizontal and vertical spectra (H_f ,V_f) of surface ground can be expressed as follows :

$$H_f = A_h * H_b + H_s \qquad (4.7)$$

$$V_f = A_v * V_b + V_s \qquad (4.8)$$

$$T_h = \frac{H_f}{H_b} \qquad (5.9)$$

$$T_v = \frac{V_f}{V_b} \qquad (4.10)$$

where

- A_h and A_v are the amplification factors of horizontal and vertically incident body waves
- H_b and V_b are the spectra of horizontal and vertical motion in the basement under the basin
- H_s and V_s are the spectra of horizontal and vertical propagation direction of Rayleigh waves.
- T_h and T_v are the amplification factors of horizontal and vertical motion of surface geological structures

In a frequency range where the horizontal component amplificate importantly the vertical component cannot (A_v=1). If there is no effect of Rayleigh waves, then $V_f \approx V_b$. Otherwise, if $V_f > V_b$, then we consider that this affected by surface waves. In

such case we calculate the effect of Rayleigh waves from the ratio V_f /V_b (=T_v) and the horizontal amplification can be written as:

$$T_h^* = \frac{T_h}{T_v} = \frac{\frac{H_f}{V_f}}{\frac{H_b}{V_b}} = \frac{QTS}{\frac{H_b}{V_b}} \tag{4.11}$$

where QTS is : $$QTS = \frac{H_f}{V_f} = \frac{A_f * H_b + H_s}{A_v * V_b + V_s} = \frac{H_b}{V_b} \cdot \frac{\left[A_h + \frac{H_s}{H_b}\right]}{\left[A_v + \frac{V_s}{V_b}\right]} \tag{4.12}$$

In Eq.4.12 we have $(H_b / V_b) \approx 1$

From the above equations we can derive the following:

- ✓ H_s/H_b and V_s/V_b relates directly with Rayleigh waves energy. In case of no influence of Rayleigh waves then $QTS=A_h/A_v$.
- ✓ If we have any significant influence then $QTS=H_s/V_s$,
- ✓ If the ground motion at basement V_b is relatively large comparing to Rayleigh waves then $QTS=A_h$.
- ✓ QTS represents first order proper frequency due to multiple reflections of S_H waves and leads to the estimation of amplification factor, independent from the amount of influence caused by Rayleigh waves.

The HVSR technique also applied at earthquake records. At Fig.4.6 the H/V spectral ratio depicted for two different earthquakes that recorded at the same station. Even though the seismograms seem different the H/V spectral ratios are quite similar.

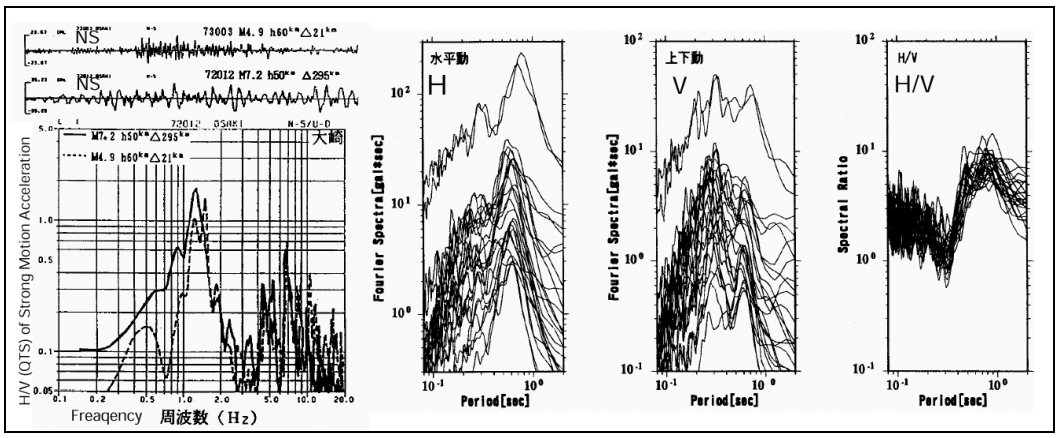


Figure 4.6 H/V spectral ratio of strong motion recordings from different earthquakes that recorded at the same station (*Nakamura*, 1989).

4.4 Estimation of HVSR

The horizontal to vertical spectral ratio (HVSR) is estimated from the three components of the ambient noise recordings (North-South, East-West and Vertical) through the following steps (Magri et al.1994, Mucciarelli 1998):

1) Offset removal: the mean of the entire recording is deducted from each sample.
2) Single window processing: The second step consists in selecting portions of the signal that do not contain transients. This is done manually (by visual examination of the recording) or with the aid of automatic window selection algorithms based on well-known STA/LTA approaches.
3) Process of individual windows:
   i. Bandpass filtering (usually from 0.1 to 20 Hz)
   ii. Cosine tapering with a length of 5% is applied on both side of the window signal of the vertical, North-South and East-West component.
   iii. FFT is applied to each component in order to obtain the spectral amplitudes.
   iv. A smoothing function is applied (usually Konno & Ohmachi window function) to each spectral amplitude (Kohno et al. 1995).
   v. Merging of two horizontal spectral amplitudes with a quadratic or arithmetic mean

vi. Thus in each of the windows the distribution of $\log_{10}$(H/V) is obtained as a function of frequency.

4) H/V is averaged over all window results

The majority of studies (*Bard,* 1999) show spectral ratios with one fundamental frequency at loose soils and nearly flat at rock soils (Fig.4.8). Of course the amplitude and the fundamental frequency differ from site to site. The natural content of fundamental frequency relates directly with the fundamental resonance frequency. However, there are a number of sites that they present a more complex structure with two or more amplification frequencies. This fact may be initially explained by the ellipticity of Rayleigh waves for sites with high impedance differences between two subsequent layers. Finally there are many cases where the spectral ratios may not present detectable peaks or they have unexplained many bumps.

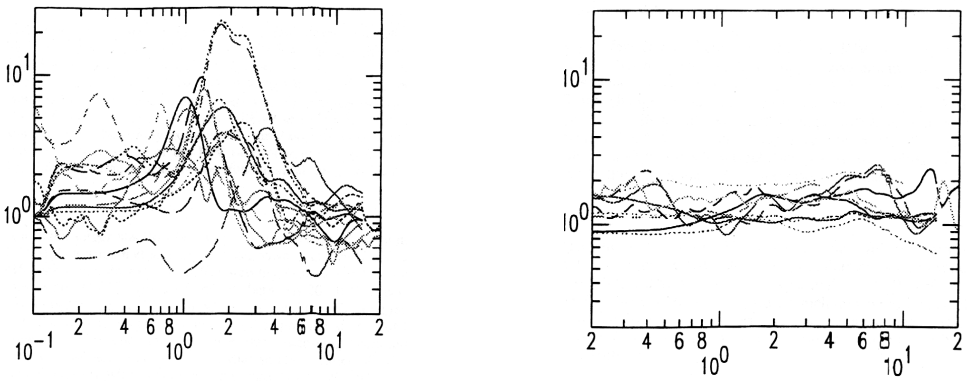


Figure 4.8 – An example of H/V spectral ratios in loose soils (left) and in rock soils (right) (Bard, 1999).

Nevertheless for the majority of cases (over 90%) the higher peak corresponds to the fundamental frequency such as $f_p^{NHV} = f_0^{NHV}$ και $A_p^{NHV} = A_0^{NHV}$. The assumption that the funadamental peak frequencies are the site's fundamental resonance frequency based on the hypothesis that local effects are stronger than any other influence. This hypothesis verified by many experimental studies at long period sites (T > 1s): Mexico (Lermo et al., 1994), Los Angeles (Yamanaka et al., 1993), San Francisco Bay area (Hough et al., 1991). This caused from the existence of very loose gsoils (Mexico, SF Bay) which imply big impedance differences in relation with

depth or where there is not a loose soil site (Los Angeles Basin) implied from the natural source of microseisms at long periods (the spectrum in near white noise levels)

Concerning short periods the results from the studies are not so homogeneous. Hough et al. (1992) estimate a resonant frequency around 2.5Hz at Tiber Valley (Italy), Gitterman et al. (1996) reported a resonant frequency around 3Hz at Hula Valley (North Israel). Yamanaka et al. (1993) reported big variations from man made noise around 3Hz at Los Angeles area since Milana et al. (1996) gather the same conclusion from their studies at Central Italy.

Probably these contradictable observations relate with impedance contrast as a function of depth:

- Strong impedance difference: the body and surface waves trapped and there is a detectable peak at spectrum which corresponds to the fundamental resonance frequency whatever the source of the waves is.
- Weak impedance difference: the waves are not trapped enough leading to ambient noise spectra representative of source characteristics.

4.5 Effects of non-stationary noise in HVSR

The calculation of spectral ratios in HVSR involves FFT methods as mentioned earlier. This lead to the necessary assumption that the used ambient noise must consists from the most stationary parts of measured time-series. This is the expected object from step 2 of HVSR estimation at section 4.4. In this section we will discuss the effects of non-stationary noise in HVSR calculation.

An unanswered question by the seismological community is whether only the stationary part of the recorded signal should be used, or the nonstationary (e.g., man made noise, industrial disturbances, random strong clipped signals) signals could also be included in the analysis. Most authors exclude nonstationary noise (e.g., Horike *et al.,* 2001) while others (Mucciarelli *et al.,* 2003) showed that the H/V ratio of nonstationary noise might not worse the results. In addition, they reported cases where the results are similar in amplitude, to the H/V spectral ratio of earthquakes. Since HVSR is, in general, a rapid approach the use of only the stationary part of the signal would make the method less attractive. This is a usual case in urban areas where (for the selection of “quiet” noise) long measurement periods are required

Transient noise is expected to be generated mainly by close sources and generally affects the noise spectra at frequencies higher than 1–2 Hz (McNamara *et al.,* 2004, Parolai & Galliana-Merino, 2006). The ratio between body and surface waves could compose a different wavefield due to the different energy that required to penetrate different sediments. This event could caused from the effect of different resonant frequencies (above or below 1Hz).

Parolai & Galliana-Merino (2006) studied the influence of short period transients with amplitude one to two times greater that mean amplitude of noise (Fig.4.9). They conclude that transients are dominated by energy at frequencies higher than 1–2 Hz and including these transients in the HVSR calculation does not worsen the results.

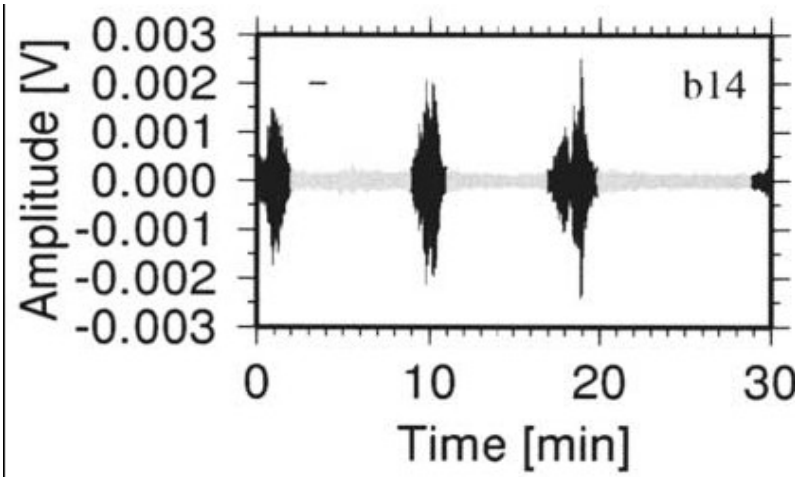


Figure 4.8 – Typical ambient noise recordings used at Parolai & Galliana-Merino (2006).

The question that remains unanswerable is what influence has the non-stationarities (including long period transients) especially when they have amplitude comparable to that of ambient noise. If these non-stationarites exists only in few specific times then the averaging of results (step 4 in HVSR calculation) will eliminate their influence. The problem that we faced is when this type of non-stationarities is present to the majority of a recorded microtremor signal. This is a usual case in urban areas field measurements where artificial noise sources are unpredictable and unavoidable.

4.6 Evaluation of HVSR calculations

There is a number of methods that used by researchers in order to evaluate the reliability of HVSR results but three of them are usually referenced as they can be applied in the majority of studies: geotechnical borehole data, HVSR from S-waves of earthquake data and Spectral Analysis of Surface Waves.

Geotechnical data derived by boreholes that take place in the area of interest. From borehole data we have a clear picture of subsurface soils which lead to an accurate estimation of velocity profile at the subsurface layers. A typical subsurface profile side-by-side with corresponding S-wave velocity profiles from 3 sites are depicted in Fig.4.12

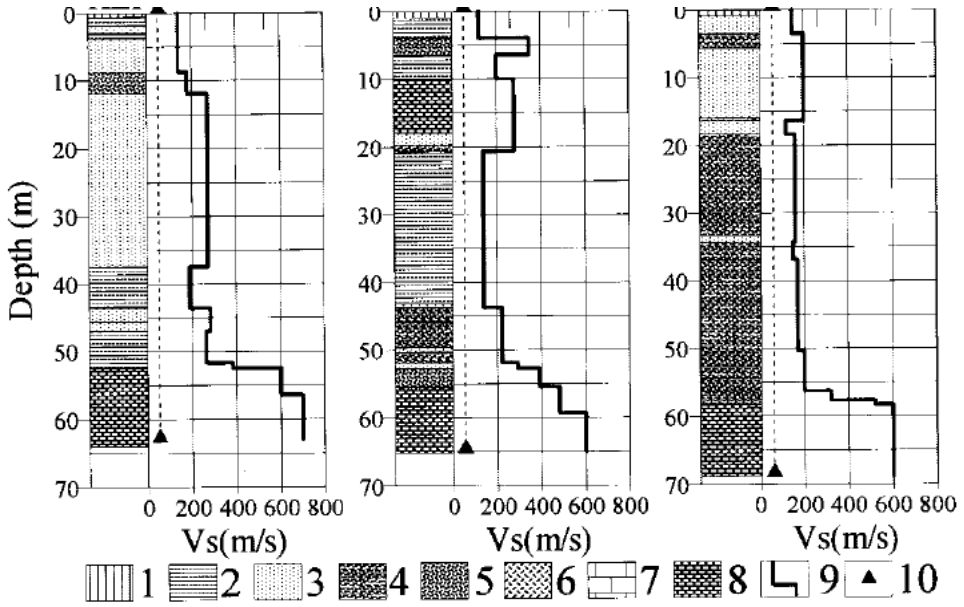


Figure 4.9 – Soil- and S-wave velocity-profiles from three sites (1) Fill, (2) Silt, (3) Sand, (4) Clay, (5) Coarse sand, (6) Loam, (7) Soft rock, (8) Mudstone, (9) S-wave velocity (combined results from Athanasopoulos, 2000, Mastrolorentzo, 2004 and Papadopoulos et. al, 2006)

The S-wave velocity profile is used to estimate site's transfer function which leads to the determination of site's fundamental frequency and amplification ratio. It is obvious that borehole data provide to the researcher the most accurate tool to characterize the subsurface but this method has two main disadvantages: it's a destructive technique and it is costly and time consuming.

Of course there are non-destructive techniques that are able to estimate the S-wave velocity profile. Among them, great attention is gathered by Spectral Analysis of Surface Wave (SASW) using array measurements. According to this technique a circular or near circular array of measurement stations is using to record ambient noise signals. The objective is to estimate the phase velocities among measurement point . For the estimation of phase velocities from array records of ambient noise, two methods have mainly been applied. One method is the frequency-wavenumber (*f-k*) spectral method (e.g., Horike, 1985; Matsushima and Okada, 1990; Kawase *et al.,* 1998), and the other is the spatial autocorrelation method (e.g., Aki, 1957; Okada *et al.,* 1987; Malagnini *et al.,* 1993). For the spatial autocorrelation method, four to six stations per array are sufficient (Okada *et al.,* 1987), and it is necessary to deploy stations in a circular configuration with one common station at the center (Aki, 1957; Okada *et al.,* 1987). On the other hand, more than seven stations per array are necessary for the stable results in *f-k* spectral method (Okada *et al.,* 1990), but there are fewer restrictions on array configurations compared with spatial autocorrelation method. From phase velocities it is possible to calculate an S-wave velocity profile which will lead us to estimation of fundamental frequency. In every case the main objective os to estimate the P and S velocity profiles because from them it is easy to estimate a theoretical model as depicted in Fig.4.10

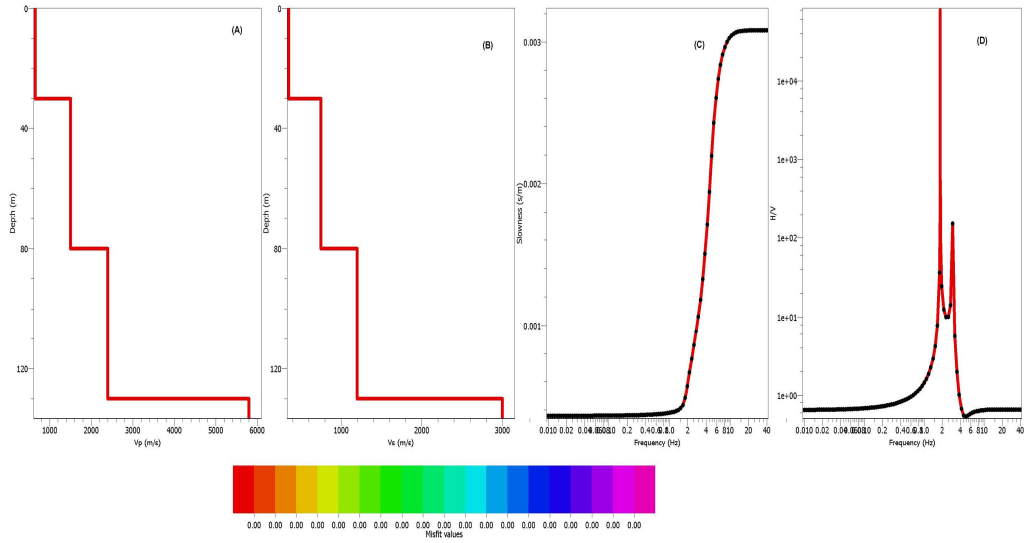


Figure 4.10 – P (a) and S (b) velocity profiles and the calculated theoretical H/V response (d)

The main objective of estimating the fundamental frequency of a site is to predict its response at forthcoming earthquake. It is more than obvious that HVSR from S-waves of earthquakes recorded at the area of interest should be a reference value. The phenomenon of seismic wave amplification due to sedimentary deposits and topographical conditions constitutes a major factor in damage and loss during strong earthquakes. The site response is best determined from recorded ground motion during an actual strong event by means of comparison with recordings at a nearby reference site located on rock (Daragh, 1991, Bocherdt, 1992, Rogers et.al, 1994, Gutiere et. al, 1994). In most cases, mainly in regions where the seismic activity is relatively low this type of analysis is usually impractical. Many investigators evaluated site response functions from moderate to weak motions of earthquakes (King & Tucker, 1984, McGarr et. al, 1991, Liu et, al 1994, Zavlnsky et. al 2000 and references herein). The implementation of this approach, however, still requires the rather frequent occurrence of earthquakes. Some observations after strong earthquakes have shown that the effects of local topography on ground motion might be of great importance (Hartzell et. al., 1994, Spudich, 1996, Bouchon & Barer, 1996).

HVSR can be applied to different regions in a recorded seismogram as marked in Fig.4.11. There are several studies that proposing usually P-waves, S-waves and Coda-waves. Sato & Kawase (2001) produce a comparative study and conclude that best results for earthquake HVSR are obtained by using S-waves and Coda-waves. In our study we are going to use S-waves from strong motion recordings since Coda-waves give us similar or in some case worst results

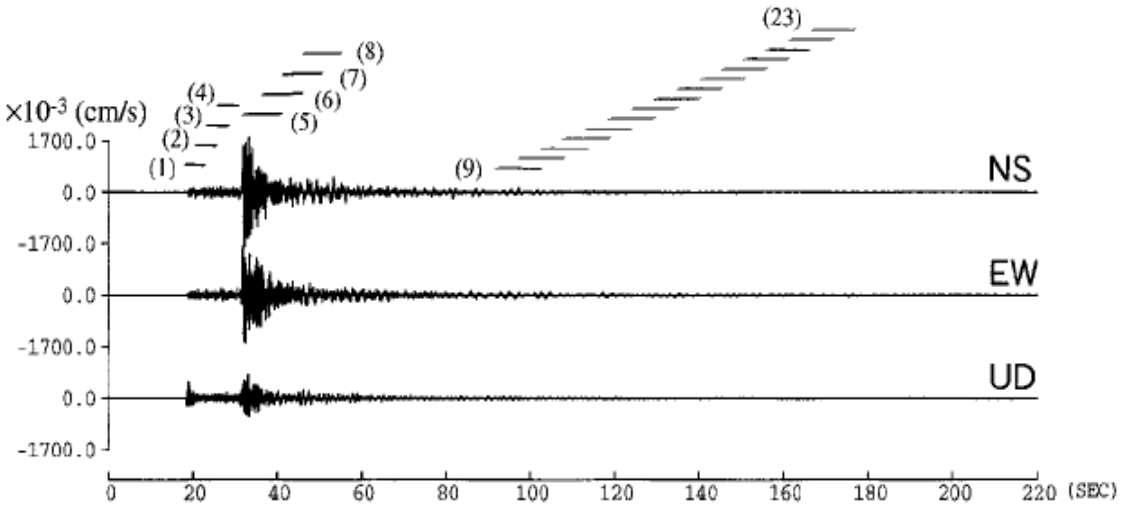


Figure 4.11 - Time windows that can be used in strong motion HVSR (1) *P* wave, (2) 2.5 sec after *P* wave, (3) 5 sec after *P* wave, (4) just before *S* wave, (5) *S* wave, (6) 5 sec after S wave, (7) 10 sec after *S* wave, (8) 15 sec after S-wave, and (9) to (23) late S coda. We will call (2) and (3) early *P* coda and (6) to (8), early *S* coda. The late *S* coda starts after twice the *S*-wave travel time and ends when the signal-to-noise spectral ratio is less than 2 between 1 and 3 Hz. (Sato & Kawase ,2001)

For illustrative purpose two typical HVSR calculations using the above approaches are presented in Fig.4.12

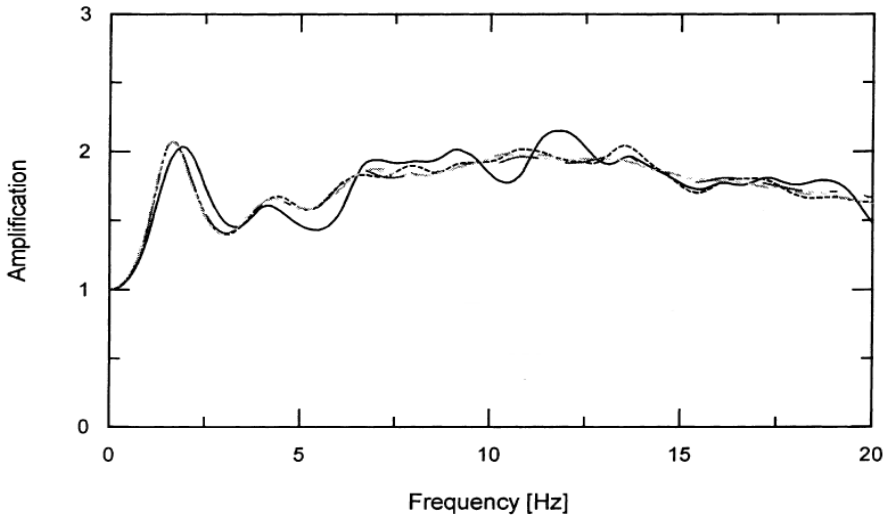


Figure 4.12 – HVSR curves from S-waves of earthquake (solid line), from borehole data (dashed line) and from SASW array method (dotted line) (Sato & Kawase, 2001).

4.7 Evaluation of STA/LTA algorithm using synthetic disturbances

The purpose of this section is to investigate the limits of the widely used Short Term Average – Long Term Average (STA/LTA, Stewart 1977) algorithm. STA/LTA used in ambient noise studies in order to automatically detect the non-stationary parts of the signal and thus to exclude them. This is done by using the algorithm in an antitrigger way: when it finds windows that fulfill the STA/LTA criteria, these are excluded from the processing stream of data (SESAME, 2004). By selecting appropriate parameters one can avoid the non-stationary parts of the signal that can possibly alter the frequency content.

The STA/LTA is, in principle, an amplitude comparison algorithm. In other words it produces results only when differences in amplitudes can identified; it usually has no ability over small differences. To illustrate this fact we run some test cases using artificial (hereafter SX) signals (Fig.4.13 – details at Table 4.1) added to real mictrotremors recordings (hereafter RX).

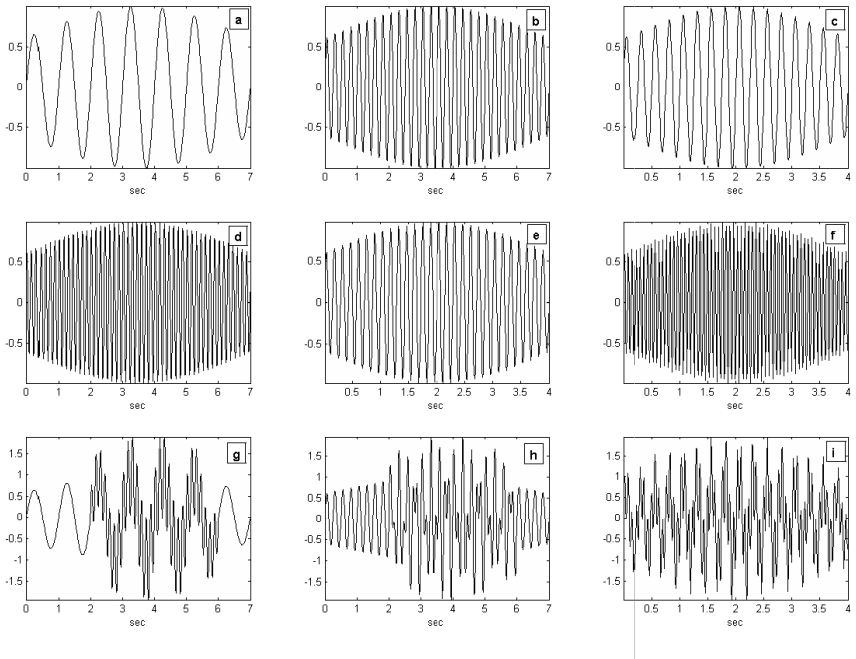


Figure 4.13: Test signals used as artificial disturbances to ambient noise recordings. Details of their attributes can be found in Table 1.

All recordings acquired using Reftek DAS-130 as data logger equipped with Guralp GMT-40T seismometer and with LGIT CityShark equipped with Lenartz 3D

5sec sensor. Sampling rate was 125Hz and recording duration was 20min. In order to avoid as much as possible the urban area disturbances that will be examined later, RX collection sites were located in countryside locations, far away from any human-involved activities. In addition, these recordings are analyzed thoroughly using he method that will be described later in order to verify the absence of important long duration transients.

Signal name	*Duration (secs)*	*Frequency (Hz)*
S1	7	1
S2	7	4
S3	4	4
S4	7	7
S5	4	7
S6	4	15
S7	7 (1^{st}) , 4 (2^{nd})	1 (1^{st}) , 7 (2^{nd})
S8	7 (1^{st}) , 4 (2^{nd})	4 (1^{st}), 7 (2^{nd})
S9	7 (1^{st} and 2^{nd})	4 (1^{st}), 15 (2^{nd})

Table 4.1: Specifications of used test signals

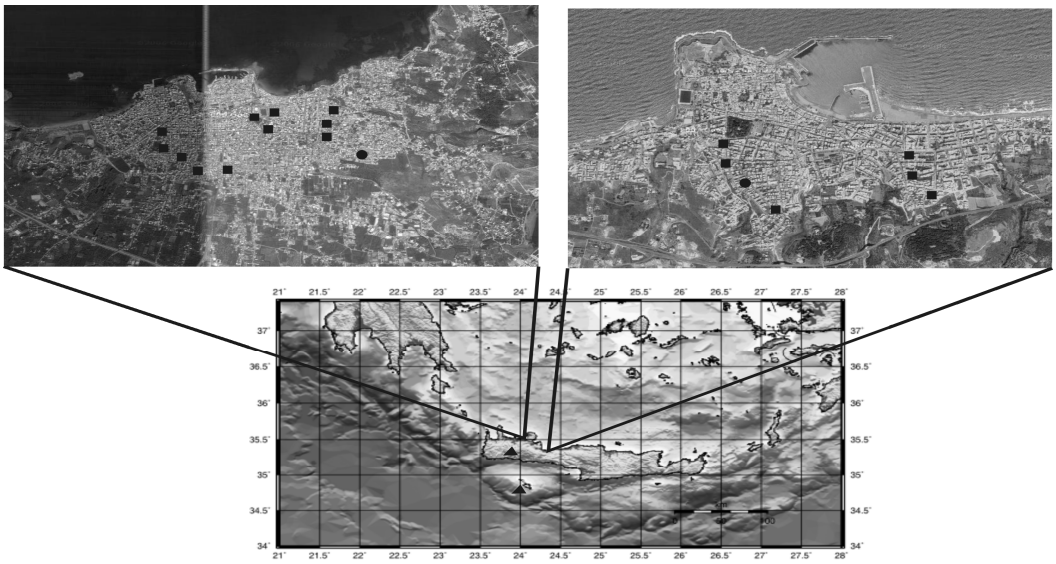


Figure 4.14 Study areas: Chania (top left) and Rethymno (top right) at Crete Island (bottom). Triangles indicate RX sites, circles X sites and squares KRX sites.

We used as detection threshold the value $\gamma = \overline{|X_n|}$ which is defined as the average of absolute values of ambient noise time series X_n (for n samples). Typical examples of thresholds calculated using γ value are depicted in Fig.4.15, for different recordings.

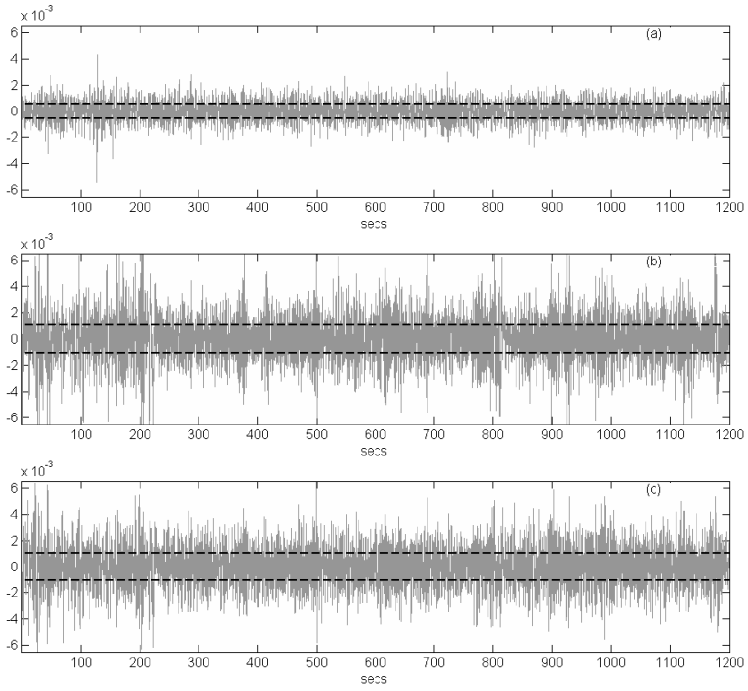


Figure 4.15: Thresholds calculated from γ value (a): Vertical component, (b): N-S component, (c): E-W component.

Data are collected at locations as shown in Fig.4.14. As one can easily identify the proposed value of γ is rather conservative because we want to illustrate the transients with low amplitude (comparable to γ which is by definition a low value attribute). The window selection procedure is tested using the corresponding modules from well known, public available, software packages GEOPSY (www.geopsy.org) and J-SESAME (http://www.geo.uib.no/seismo/software/jsesame/jsesame.html). Typical window selections from several recordings with and without added artificial disturbances are shown in Fig.4.16

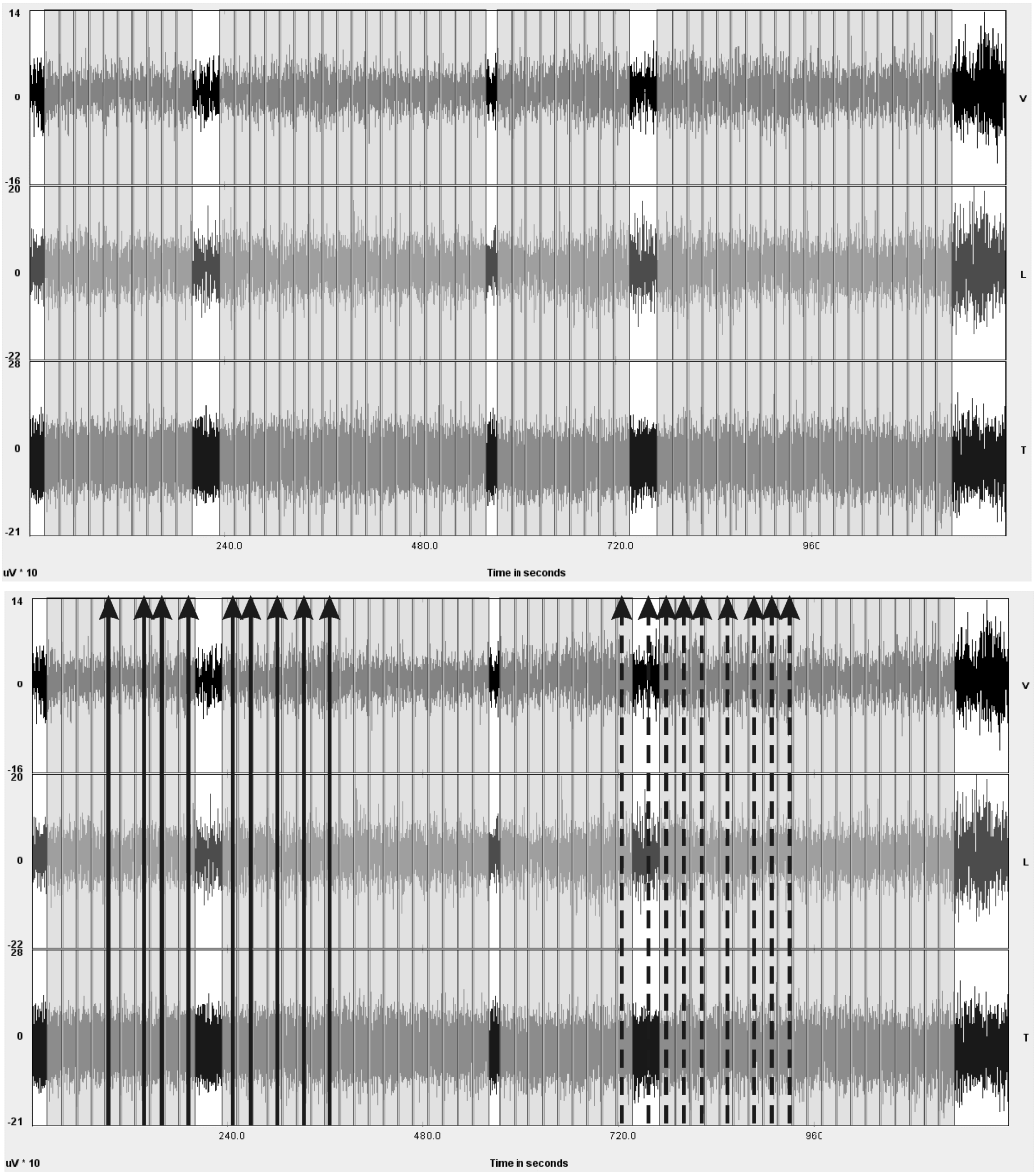


Figure 4.16: Window selections (grey zones) using STA/LTA algorithm for raw ambient noise recordings without artificial disturbances (top plot) and with disturbances (bottom plot) starting at points indicated by arrows. All signal (S1-S9) added consequently two times as indicated by arrow groups (first set – solid arrows, second set dashed arrows). Each individual arrow indicates the starting sample of artificial signal.

Initially we used the window selection procedure to raw recordings in order to determine the windows that will be accepted for the HVSR calculation. Next we contaminate the recordings by adding SX (X:1→ 9) signal to every recording and reprocess them with the window module. Each artificial disturbance is added

independently to each recording and repeated 20 times (starting from 50th sec and repeated every 50 secs). The value of γ remains stable for each recording. We tested γ values from 0.5 to 2.5 with 0.1 resolution. The performance (P_r) of "antitriggering" is calculated using the relation:

$$P_r = \frac{n_{dd}}{n_{ad}} \tag{4.16}$$

Where n_{dd} : number of detected (and excluded) disturbances

n_{ad} : total number of disturbances, here 20

At Fig.4.17 we summarize the performance of "antitriggering" for STA/LTA by plotting the results of P_r in relation of γ.

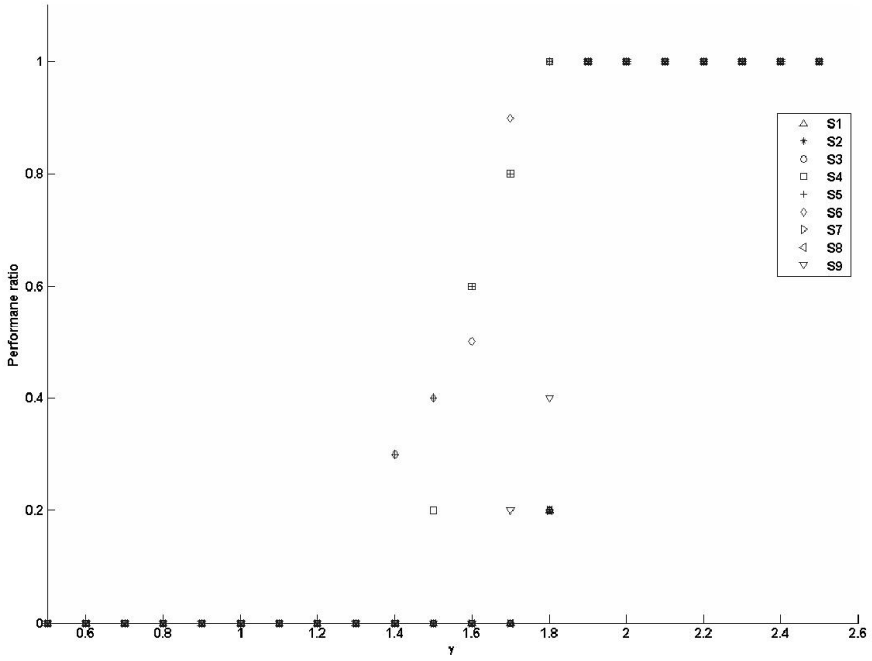


Figure 4.17: Average results of STA/LTA performance according to γ value for all the artificial signals.

We can conclude that if the amplitude of the transients is at least 1.5 times the value of γ it's rather impossible for the STAL/LTA algorithm to detect these transients. Between 1.5 and 1.85 there are some cases where the detection is achieved but these detections achieved to transients with higher frequencies (signals S4, S5,

S6). In addition, the artificial signals that are monochromatic can be detected easier. Of course, this event cannot act as a rule because it can easily be inverted if one selects appropriate STA/LTA window lengths. For γ>1.85 the STA/LTA performs as expected.

4.8 Impact of low amplitude transients to HVSR

The purpose of this section is to investigate if and how, undetectable transients from STA/LTA algorithm can affect the results of HVSR method. In order to facilitate this we use ambient noise recordings (hereafter KRX) from 17 sites with known attributes. These are derived from results of previous studies (boreholes, array techniques) or we calculated 1D models using EERA code (Bardet et.al 2000) based on V_s profiles derived from previous studies also (EMERIC 2005, SE-RISK 2008). We used the recordings with γ<1.85 as described in the previous section and we prepared our dataset by adding the artificial disturbances in a progressive way: we begun by adding each one of nine SX signals in 50th sec of each KRX thus creating 17x9=153 cases. We performed the HVSR calculation and then we add one more instance of each SX at 100th sec (thus creating 153 more cases with recordings that included two parts of the same SX – at 50th and 100th sec) and repeat the calculation. This procedure repeated 25 times. HVSR calculation performed as ordinary with recordings that bandpass filtered (0.2Hz to 20Hz), detrended and baseline corrected, 5% cosine tapered, smoothed using Kohno & Ohmachi window with b=25.

A typical example of how a single undetectable transient can affect the calculated spectrum is illustrated in Fig.4.18. Since the FT provides excellent localization in frequency domain it is expected that the (undetectable by STA/LTA) added disturbance will produce a peak in frequency domain. Of course, it is not expected that a single peak could significant affect site's f_0 since HVSR is an averaging result method. The question that arises is if there is a limit, over which, HVSR could be affected.

At Fig.4.19 results from a site with good agreement between HVSR results and results obtained when parts of artificial disturbance (here S1) is added, are presented. It is obvious that as the number of undetectable disturbances increases the estimated f_0 is not depicted by a clear peak.

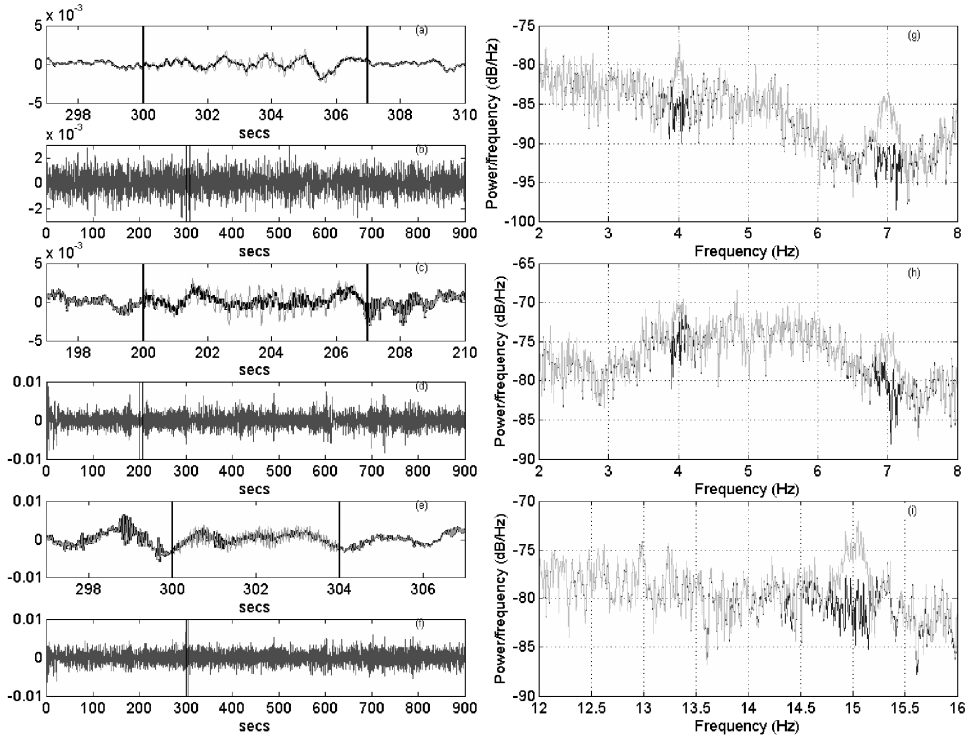


Figure 4.18: Typical example of spectrum contamination when one part of artificial signal is added. (a,b) S8 signal (γ=1) added from 300 until 307sec at raw recording. (c,d) S8 signal (γ=1.2) added from 200 until 207sec at raw recording. (e,f) S6 signal (γ=1) added from 300 until 304sec at raw recording. (g) Spectrum of (b) recording without artificial signal (black line) and with it (grey line). (h) Spectrum of (d) recording without artificial signal (black line) and with it (grey line). (i) Spectrum of (f) recording without artificial signal (black line) and with it (grey line).

At Fig.4.20 the HVSR results using recordings (from different known sites) where 20 parts of artificial disturbances, added. We can identify three general cases: First, when f_0 and artificial disturbances appeared individual then it's difficult to identify which is real and which not. These are the case with 1Hz and 4Hz signals and some of 7Hz (only when f_0 is quite high). Second when f_0 and artificial disturbances are quite close, the f_0 seems shifted or more dispersed. This is the case with 1Hz signals. Finally, with high frequency artificial disturbances (15Hz and some 7Hz) an individual peak appeared at those values.

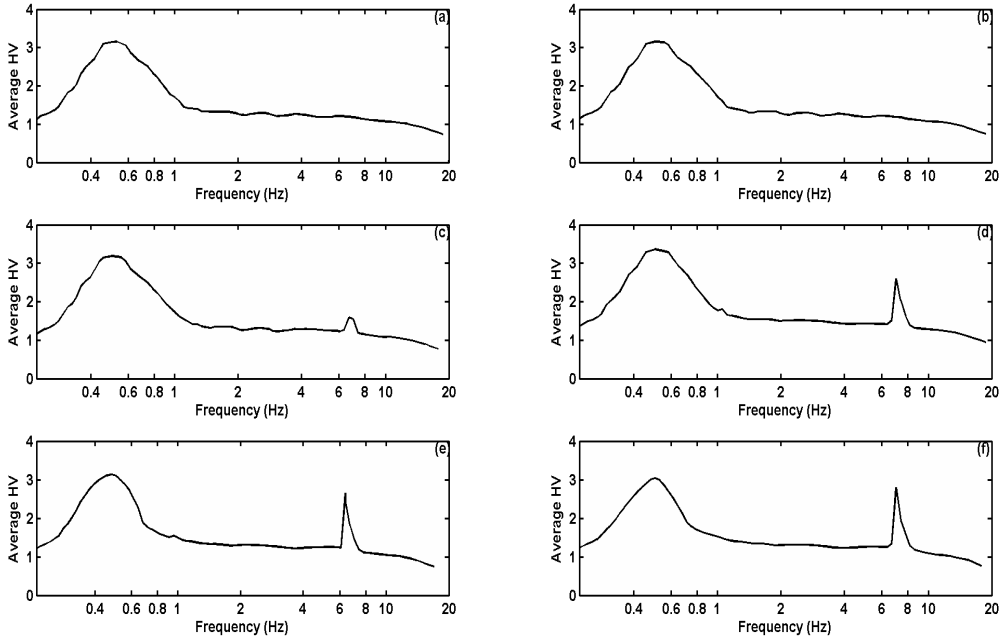


Figure 4.19: HVSR results from a single site with increasing parts of S1 signal added to raw recordings. (a) raw recording, (b) 1 part, (c) 7 parts, (d) 14 parts, (e) 20 parts, (f) 25 parts.

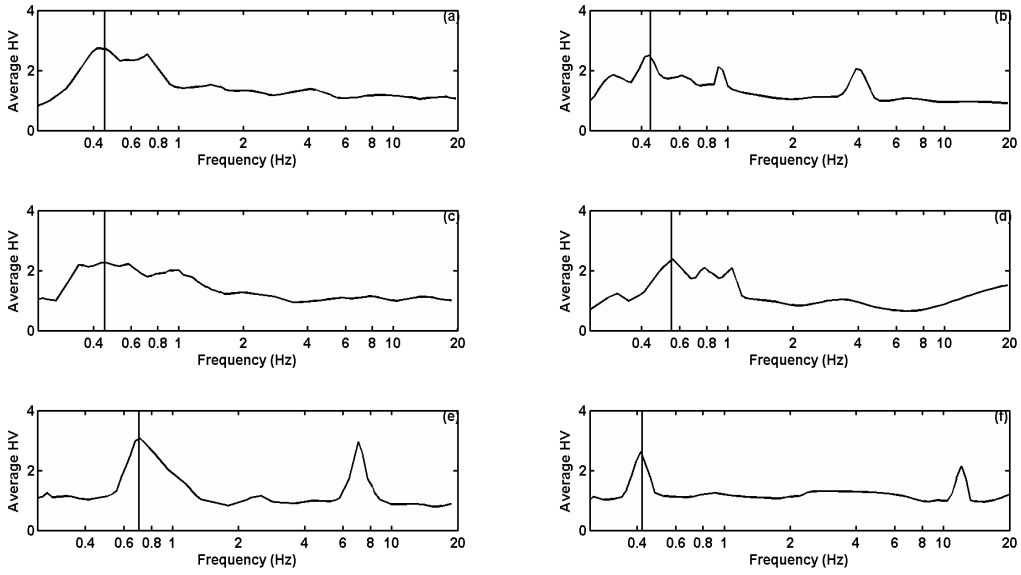


Figure 4.20: HVSR results from different sites using recordings with artificial (non-detected) disturbances added. Vertical line in each plot indicates the calculated f_0 from same raw recordings. Signals used to contaminate the corresponding recordings are: S1 at (a), S7 at (b), S1 at (c), S1 at (d), S5 at (e), S6 at (f).

An experienced analyst, as far there is some knowledge of the local geology, could reject these values but here we seek a more "computational" way to reject it.

From the above it is obvious that there are enough cases where the undetectable transients could produce problems to estimation of f_0. At this point we must mention that the values of artificial disturbances are carefully selected to represent non-stationary disturbances that recorded during the big number of measurements that took place the last four years around big cities at Crete Island. These values are representable only for the urban areas that we investigate and no further extension should be made for other areas without prior acquisition of possible local singularities.

4.9 Application of MODWT for identification of low amplitude transients

In this section the application of MODWT in raw recordings will be discussed along with its application to experimental ones. From what we present so far it is obvious that, in case that we want to detect low amplitude non-stationary components to ambient noise recording, we need a more efficient detector than traditional STA/LTA. Our proposal is to transfer the amplitude detection from 1D time domain to 2D time-frequency domain using wavelets. Briefly speaking, by using time-frequency representation we are able to detect the frequency of non-stationary part as well as its time attribute (i.e. position) inside recordings. For this transformation we used MODWT and not DWT because we don't want to make any prior assumptions about the nature of recordings as explained in previous section.

The MODWT performed using Least Asymmetric (LA) filters with length 8 using the WMTSA software (public available at www.atmos.washington.edu/~wmtsa). The number of levels used for decomposition was $J_0 = 8$ which corresponds to physical period up to 4.098sec. Using the MODWT transform we calculate wavelet and scaling coefficients for consecutive windows with length each recording. We propose the use of MODWT not to whole recording but to subsections of it. The smaller window that a decomposition can be achieved dictated by the number of samples as already mentioned by Equations 2.42 and 2.43. Fig.4.21 illustrates such a decomposition from a recording where we select n=20000 (160sec). It contains one S2 disturbance at 8^{th} sec, one S6 at 50^{th} sec and one S8 at 120^{th} sec. Lower panel illustrates the recording.

The top panel shows the wavelet coefficients $\tilde{W}_{j,t}$ for levels *j*=1-8 and the $\tilde{V}_8$ scaling coefficients plotted on the same vertical scale as the lower panel. For each level solid vertical lines near end points demarcate the boundaries outside of which the coefficients are influenced by boundary conditions. The coefficients are plotted after circular shifting of wavelet coefficients $\tilde{W}_{j,t}$ in order to be properly aligned with original recording. The shifted vector is marked at the right of each subplot in the form of $T^{-v_j}\tilde{W}_j$ where v_j indicates the number of samples the coefficients have been shifted.

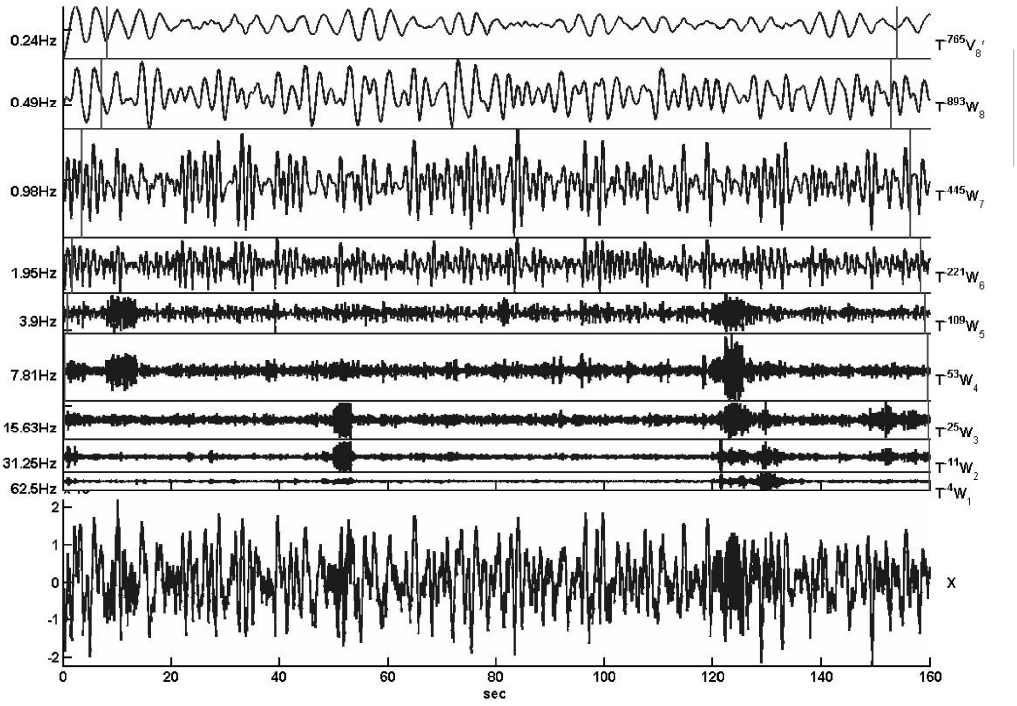


Figure 4.21: MODWT wavelet (levels 1–8) and scaling (J_0=8) coefficients (top panel) and ambient noise recording (bottom panel). Solid vertical lines (top panel) indicate the boundaries outside of which coefficients are influenced by circular shifting. .

Without any prior experience the reader can easily identify the nature and position of S2, S6 and S8 signals at different scales which, as already mentioned, corresponds

to different frequency bands. At this stage the analyst can select manually the portions of the recording that like to exclude or by using an automated procedure.

At final stage we apply the MODWT to recordings from sites with calculated f_0 is not agree with those than calculated from other methods or there were secondary HVSR picks that could confuse the analyst. At Fig.4.22 – 4.28 representative results (from several locations shown at Fig.4.14) with known differences between f_0 from HVSR and 1D linear equivalent model using EERA code, are presented.

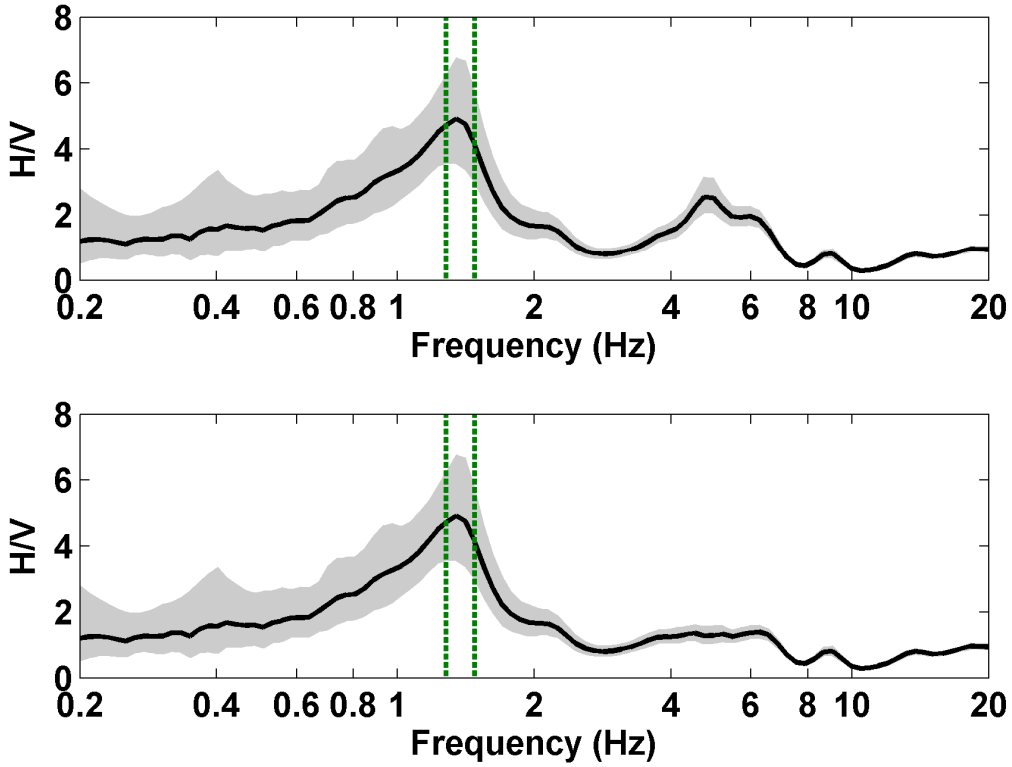


Figure 4.22: Results from site KRX01 without WT non-stationarity detection (top) and with WT detection (bottom). Disturbances detected at around 5Hz. Green dashed lines indicate the f_0 portion from 1D model as calculated from inverted V_s profile (EERA model).

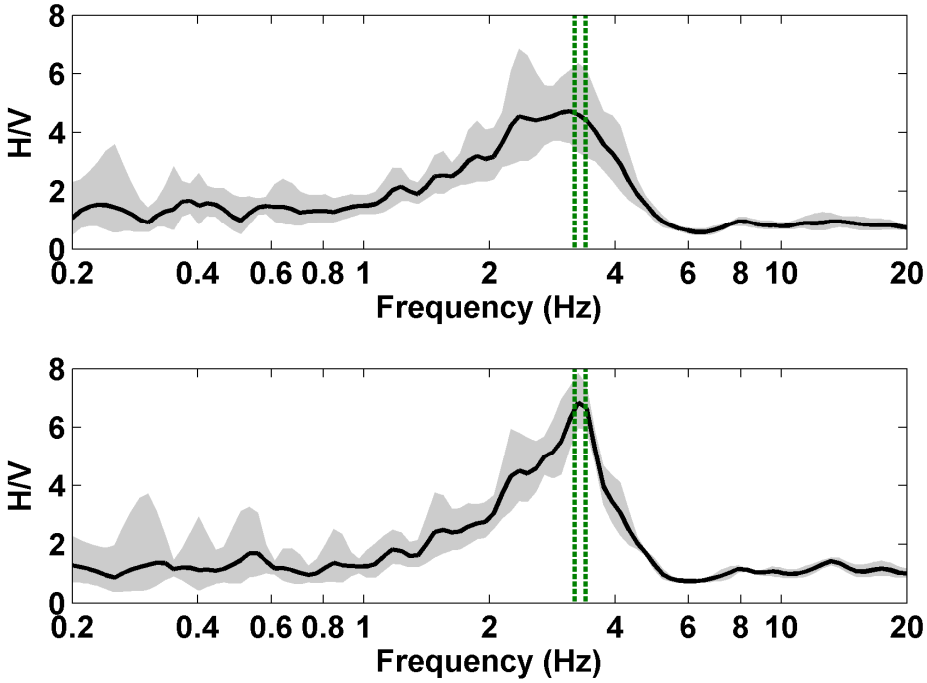


Figure 4.23: Results from site KRX02 without WT non-stationarity detection (top) and with WT detection (bottom). Disturbances detected at around 3.8Hz. Green dashed lines indicate the f_0 portion from 1D model as calculated from inverted V_s profile (EERA model).

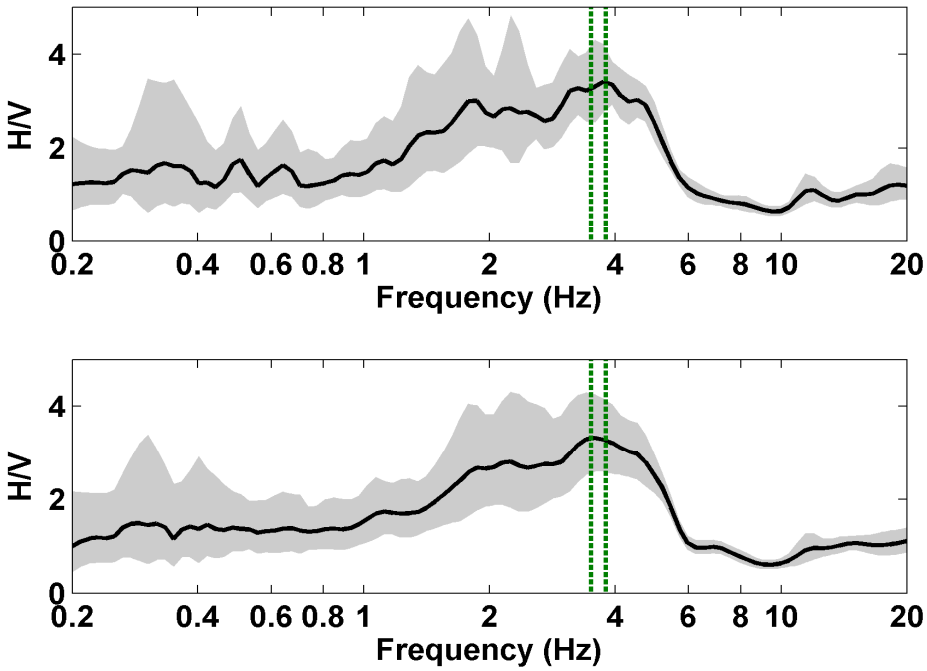


Figure 4.24: Results from site KRX03 without WT non-stationarity detection (top) and with WT detection (bottom). Disturbances detected at around 3.9Hz. Green dashed lines indicate the f_0 portion from 1D model as calculated from inverted V_s profile (EERA model).

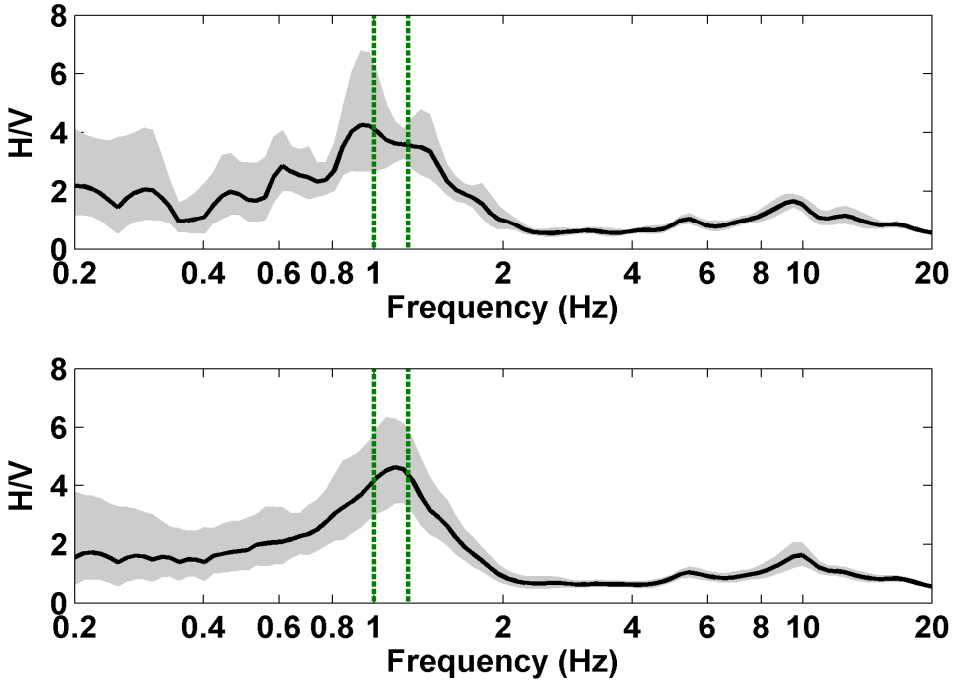


Figure 4.25: Results from site KRX04 without WT non-stationarity detection (top) and with WT detection (bottom). Disturbances detected at around 1.5Hz. Green dashed lines indicate the f_0 portion from 1D model as calculated from inverted V_s profile (EERA model).

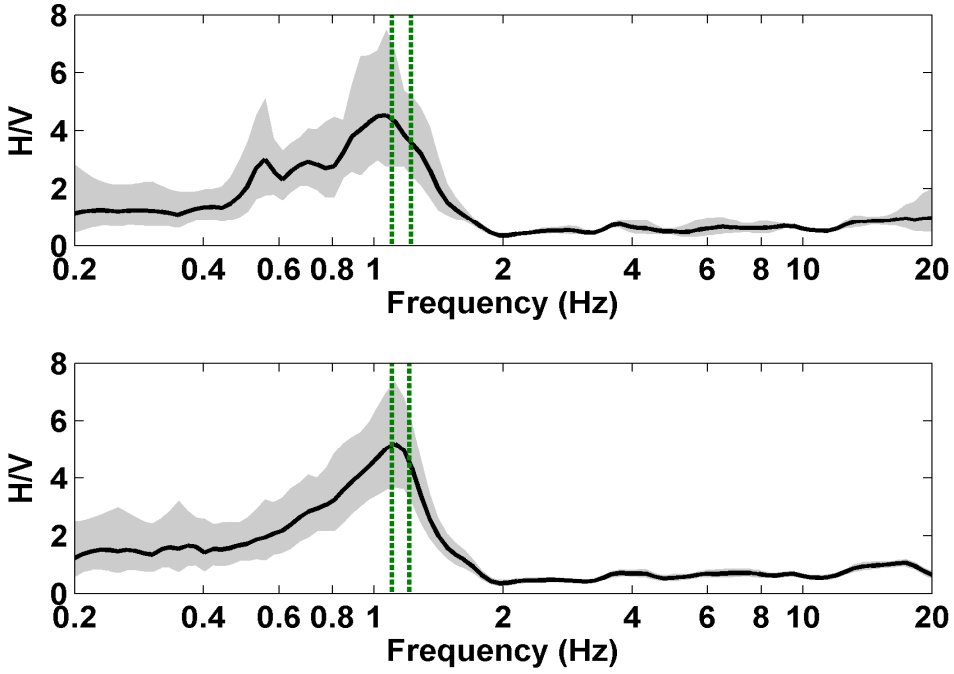


Figure 4.26: Results from site KRX06 without WT non-stationarity detection (top) and with WT detection (bottom). Disturbances detected at around 0.8 and 0.55Hz. Green dashed lines indicate the f_0 portion from 1D model as calculated from inverted V_s profile (EERA model).

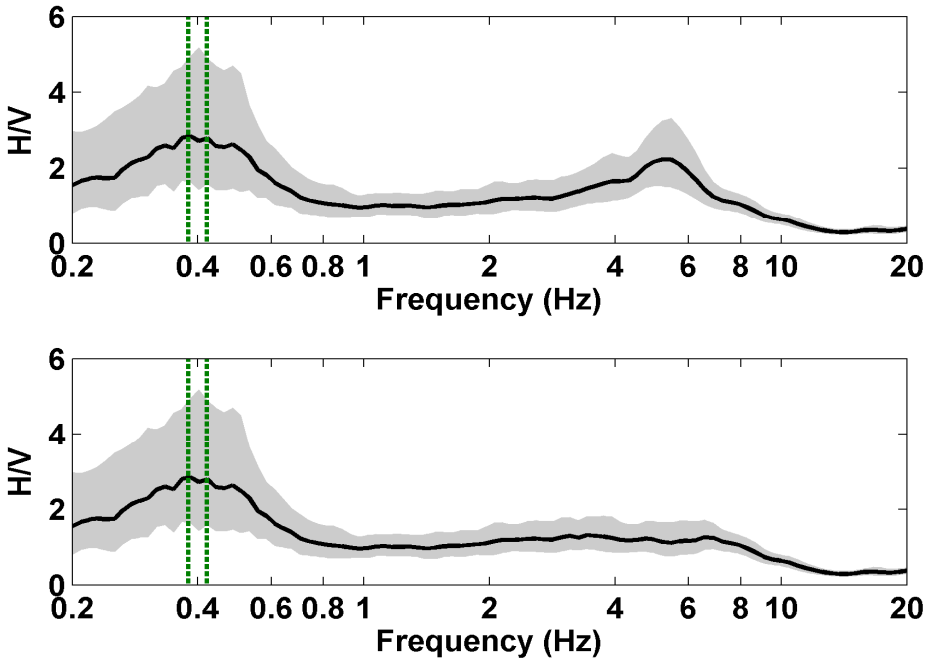


Figure 4.27: Results from site KRX08 without WT non-stationarity detection (top) and with WT detection (bottom). Disturbances detected at around 5.8Hz. Green dashed lines indicate the f_0 portion from 1D model as calculated from inverted V_s profile (EERA model).

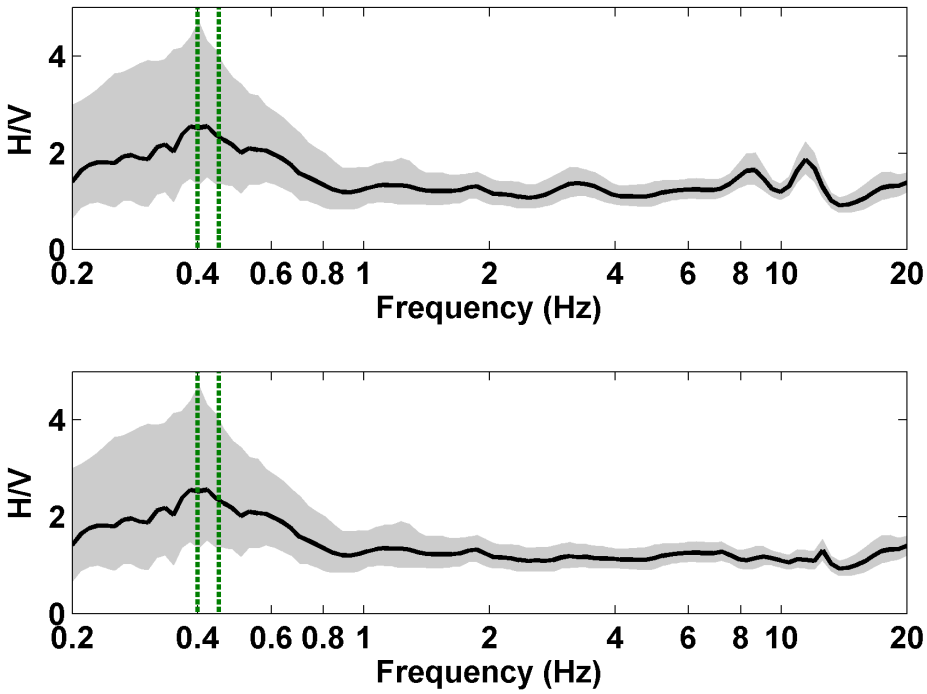


Figure 4.28: Results from site KRX09 without WT non-stationarity detection (top) and with WT detection (bottom). Disturbances detected at around 8.5Hz and 11Hz. Green dashed lines indicate the f_0 portion from 1D model as calculated from inverted V_s profile (EERA model).

The MODWT transformed recordings are shown in Figures 4.29 and 4.30. After detection and rejection of possible disturbed non-stationarities the HVSR results that we produced are illustrated in Fig.4.25.

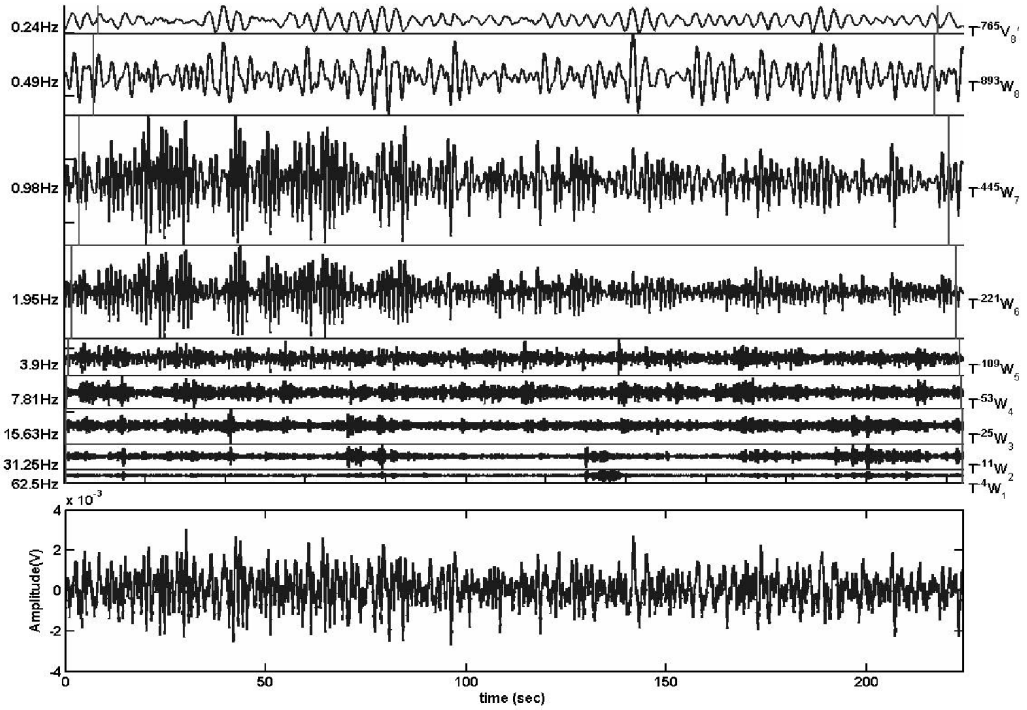


Figure 4.29: MODWT for recording from site KRX04. Significant non-stationary signal detected between 10^{th} and 95^{th} sec at around 1Hz.

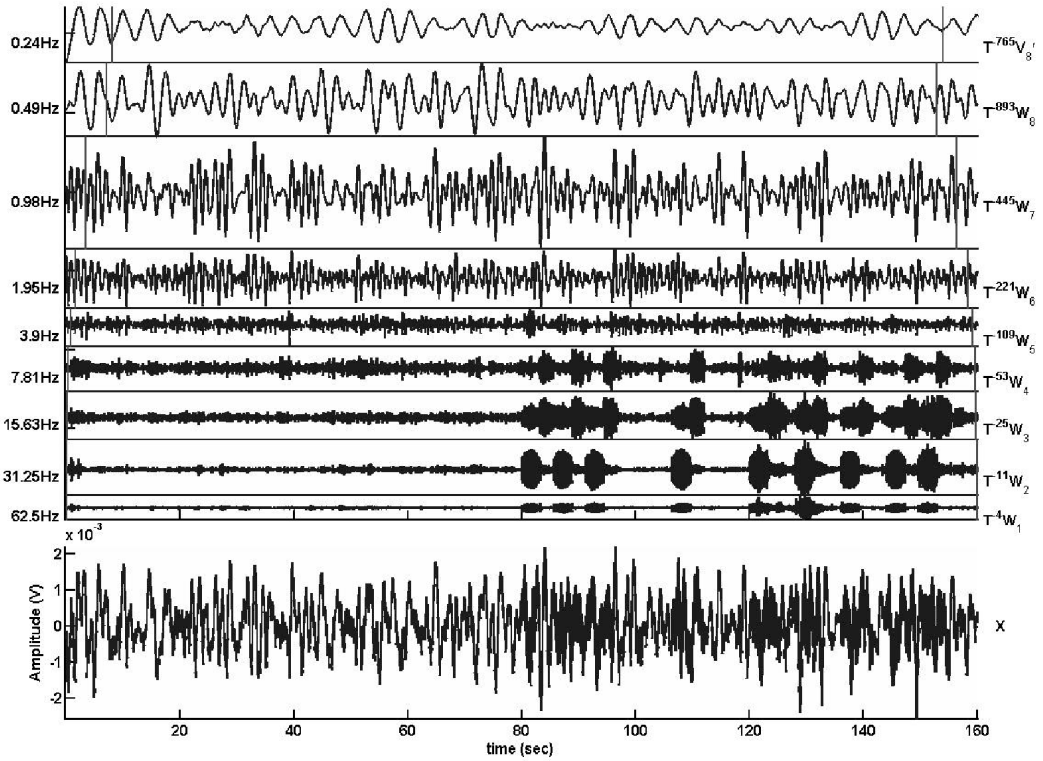


Figure 4.30: MODWT for recordings from site KRX09. Significant non-stationary signal detected at 80th, 85th, 87th, 110th, 120th, 128th, 138th, 145th and 150th secs.

4.10 Summary

Non-stationary, artificial disturbances are a usual situation in urban places where industrial and anthropogenic activities run continuously. At these locations, ambient noise recordings collected even at late hours could suffer from transients with amplitudes comparable of ambient noise. Based on analysis described on the previous sections we may conclude that these transients remain undetectable from STA/LTA algorithm. In case of these transients have long duration there is a strong possibility to contaminate the calculation of f_0. An experienced analyst could probably identify the non-stationarities by examining the FT in each windows but our intension is to focus on an automated signal selection procedure. The use of a WT in order to localise a signal in time-frequency domain and especially the use of a redundant one, such as the

MODWT, can solve the problem of accurate automatic determination of the above transients. If the amount of the transients is significant then they can be excluded since the MODWT is able to provide their time position in the recording. The whole procedure can run as a preliminary test to ambient noise recordings in order to provide a clear view of the existence of transients in ambient noise recordings.

5. Denoising seismic data

5.1 Introduction

During the development of a seismological network the choice of optimum field stations is problematic. It is not unusual that a preselected station's location introduce some undesirable disturbances to recording seismograms after installation. The main disturbance that affects the quality of a seismogram is the ambient noise. In this chapter, wavelet methods for noise elimination will be presented and compared. In addition two main applications that are high related to the quality of a seismogram (P-wave detection and detection of microseismic signals) will examined before and after the application of appropriate wavelet denoising schemes.

In order to derive onset times, amplitudes or other useful characteristic from a seismogram, the usual denoising procedure involves the use of a linear pass-band filter. This family of filters is zero-phase and is useful according to phase properties but their efficiency is reduced when transients are existing near seismic signals (Scherbaum, 1996). An alternative solution is the Wiener filter which proposed by Douglas (1997). This filter is a linear filter that focuses on the elimination of the mean square error between recorded and expected signal. Its main disadvantage is the assumption that signal and noise are stationary. This assumption does not hold for the seismic signals leading to denoising solutions that does not assume stationarity. Solutions based on WT proved effective for denoising problems across several areas.

Here we present recent WT denoising methods (WDM) that will applied later to seismic sequences of Seismological Network of Crete. The contents of this chapter begin with discussion of thresholding rules followed by discussion of classical wavelet estimators and Bayesian approach to wavelet thresholding. The contribution of this chapter is the comparison of WDM in seismological signals and the selection of the most appropriate methods for these type of signals.

5.2 Non-parametric regression using wavelets

Every type of measurement is associated with noise. Denoising is the procedure that targets to the attenuation of the fluctuations in data values in order to derive at least the most reperesentative characteristics of measured signal. Assuming that we have an unknown signal represented from $f(t_i)$ which is measured as:

$$y_i = f(t_i) + \varepsilon_i \quad \text{i} = 1,2,..\text{n} \tag{5.1}$$

where ε_i is the random noise with level σ associated with the measurement. The goal of any denoising method is to minimize the difference between y_i and the true signal *f(t)*. If *f(t)* in known, by running a parametric regression (i.e using least squares) the estimation of regression parameters could be achieved. When the underlying function *f(t)* is unknown (which is the case in seismic signals) this process is called *non-parametric regression.* A usual non-parametric regression solution is to express the unknown function *f(t)* as a Fourier series and then estimate the Fourier coefficients from the measured data. The success of this approach is highly dependant on the solution of the basis function for the series. An appropriate basis could be a function that carries the ability to approximate a large set of possible response functions using only few terms from Fourier expansion. From this point it is clear now that wavelets could be one group of the quested basis functions since they have the ability to express a wide variety of functions by using a limited set of basis.

5.3 Wavelet direct thresholding rules

As we cannot make prior assumptions about the behaviour of the seismic data it is quite clear that denoising seismic measurements is a non-parametric regression problem. An effective approach to this type of problems is the method called *wavelet thresholding.*

Applying WT to the model at Equation 5.1 we have

$$\text{W}_y = \text{W}_f + \text{W}_\varepsilon \tag{5.2}$$

Since WT is orthogonal noise is transformed to noise. If $w_{\text{j,k}}$ are the wavelet coefficients of W*f* we may write

$$y_{\text{j,k}} = w_{\text{j,k}} + \sigma_{\text{j,k}} \tag{5.3}$$

The Equation 5.3 states that the wavelet coefficients of a noisy signal may be expressed as noiseless coefficients plus noise.

The computation scheme of a wavelet thresholding approach consists of three basic stages which also depicted on Fig.5.1:

1. Application of WT to the noisy data in order to derive wavelet approximation and details (coefficients)
2. By using thresholds, coefficients are reduced, nullified those that are below a certain level. At this stage the noise coefficients are eliminated

3. Unknown function *f(t)* is estimated after inverse WT of the remaining coefficients from previous stage.

$$y \xrightarrow{WT} \left(c_{j_0,k}, d_{j,k}\right) \xrightarrow{THRESHOLDING} \left(c_{j_0,k}, \lambda \cdot (d_{j,k})\right) \xrightarrow{InverseWT} f$$

Figure 5.1. Function estimation using thresholding.

As already mentioned, detail coefficients from WT represent the relative changes of measured signal through scales. Since noise is mainly fluctuations in the signal, these will depicted at details coefficients at small scales (high frequencies). To reduce noise, one can simply decrease or eliminate detail coefficients whose magnitude is below a predefined threshold value. This process usually found in the literature under the name *wavelet shrinkage*. This idea initially introduced by Donoho (1994,1995). He proposed two thresholded approaches: hard and soft thresholding. The criterion for this scheme is to replace the detail coefficients with zero whenever its magnitude is smaller than a preset threshold λ. Soft thresholding (ST) is defined as:

$$d_{j,k}^{S} = \begin{cases} 0 & \text{if } \left|\mathrm{d_{j,k}}\right| \le \lambda \\ \mathrm{d_{j,k}} - \lambda & \text{if } \mathrm{d_{j,k}} > \lambda \\ \mathrm{d_{j,k}} + \lambda & \text{if } \mathrm{d_{j,k}} < -\lambda \end{cases} \tag{5.4}$$

Hard thresholding (HT) is defined as:

$$d_{j,k}^{H} = \begin{cases} 0 & \text{if } \left|\mathrm{d_{j,k}}\right| \le \lambda \\ \\ \mathrm{d_{j,k}} & \text{if } \left|\mathrm{d_{j,k}}\right| > \lambda \end{cases} \tag{5.5}$$

The ST is a continuous function and acts under the frame of "preserve" or "kill" rule since the HT is a discontinuous function that acts under "shrink" (by the amount λ) or "kill" rule as shown in Fig.5.2

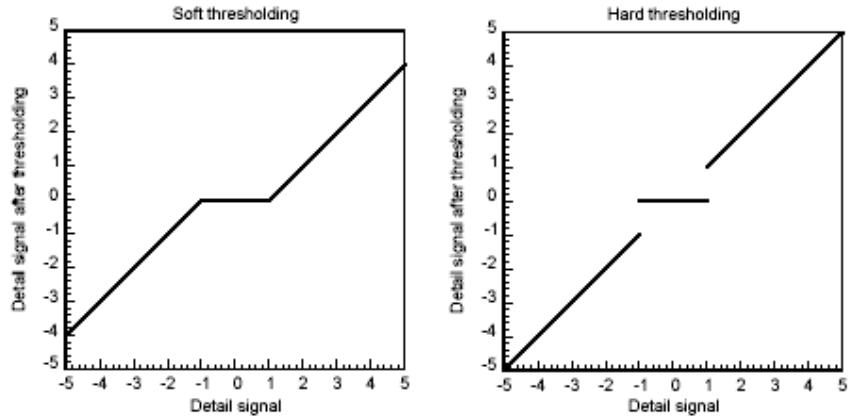


Figure 5.2. Soft and Hard thresholding functions

HT is preferred for applications where the signal contains peaks and discontinuities since ST is preferred for signals where smoothness of the estimated is the leading object. On the other hand, HT (due to its discontinuity) could be very sensitive to small changes of the data (Bruce & Gao, 1996). Also simple values of HT produce large variance at the estimated function. In order to remedy the above drawbacks some solutions such as *firm thresholding* (Bruce & Gao 1997) and *nonnegative garrotte thresholding* (Gao, 1998) proposed. Although the overcome the drawbacks of HT and ST the produce secondary complex requirements that act inhibitory to the estimated function. A quite fair solution is the *smoothly clipped absolute deviation* (SCAD) threshold proposed by Antoniadis & Fan (2001) which defined as:

$$d_{j,k}^{SCAD} = \begin{cases} sign(d_{j,k})\max(0, |d_{j,k}| - \lambda) & \text{if } |d_{j,k}| \le 2\lambda \\ \dfrac{(\alpha\text{-}1)\mathrm{d}_{\mathrm{j,k}} - \alpha\lambda sign(d_{j,k})}{\alpha\text{-}2} & \text{if } 2\lambda < |\mathrm{d}_{\mathrm{j,k}}| \le \alpha\lambda \\ \mathrm{d}_{\mathrm{j,k}} & \text{if } \mathrm{d}_{\mathrm{j,k}} > \alpha\lambda \end{cases} \qquad (5.6)$$

where α suggested to have value near 3.7. In our study (after experimenting) we got the best results in average with α=3.8. SCAD rule acts under the frame of "preserve", "shrink" or "kill". It's a piecewise linear function and does not create excessive bias when the wavelet coefficients are large. This is a useful feature for the first seconds of the seismic recordings where the amplitude monotonically increases.

5.4 Wavelet direct thresholding estimators

Threshold λ can be chosen using a variety of methods according to given situations. One can separate them into two main categories: *global threshold* and *level dependent threshold.* Global threshold means that a single value of λ is applied to wavelet detail coefficients $d_{j,k}$ ($j=j_0,\ldots,j\text{-}1$ for $k=1,2,\ldots,2^{j-1}$). Level dependent threshold means that a possibly different value of λ could be applied at each scale j. A typical representation of both categories depicted in Fig.5.3, where at the top plot a global threshold is selected (3.04) for all the scales instead of the bottom plot where the threshold is different in each scale

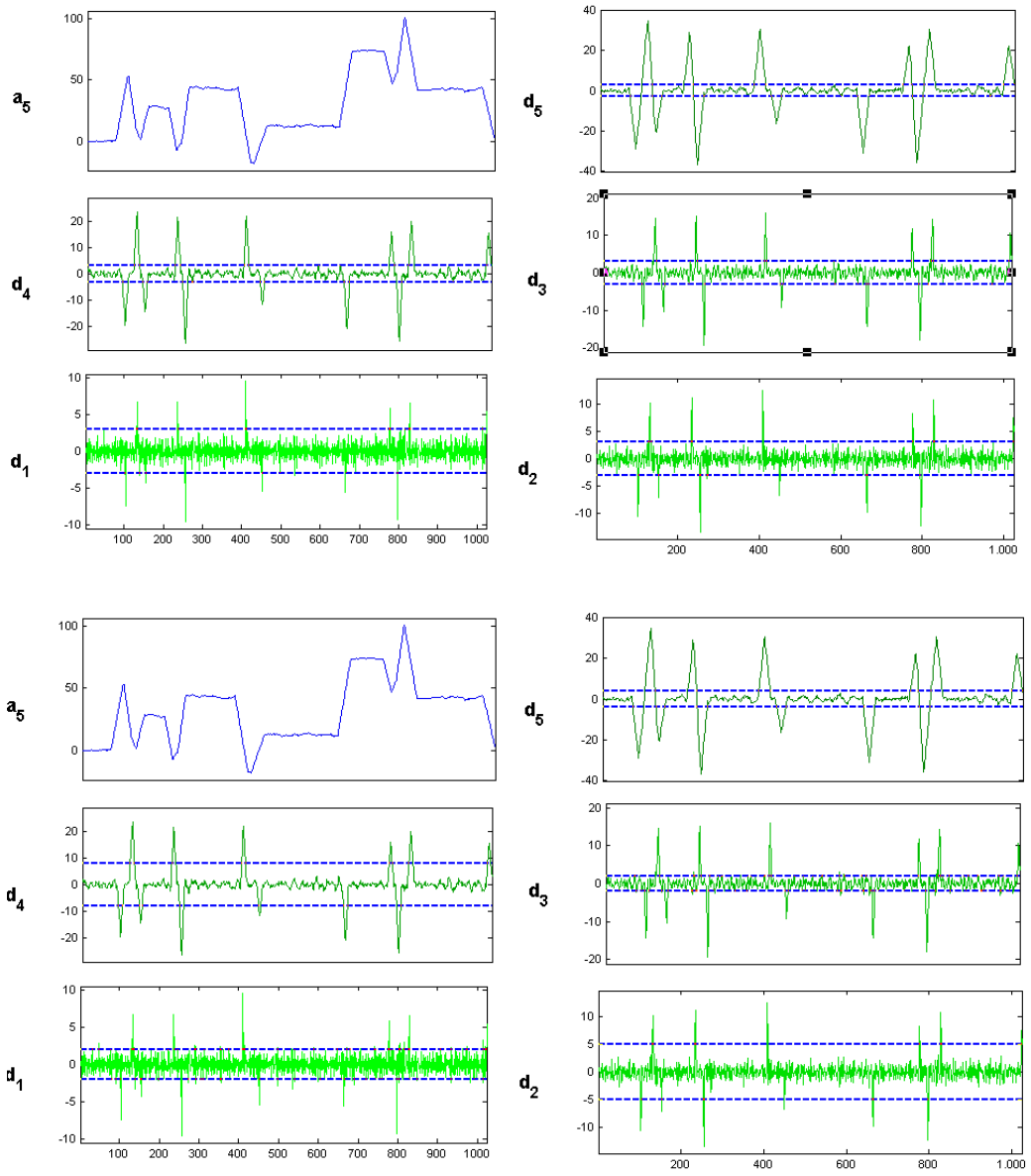


Figure 5.3. Global (top plot) and level dependent (bottom plot) thresholding for a wavelet decomposed signal

Both methods require the estimation of noise level σ. A classical solution to the estimation of σ is the standard deviation of y_i values thus the estimation will take place in time domain. This is acceptable only when the underlying function $f(t_i)$ is without discontinuities. An alternative successful approach is to estimate σ in wavelet domain (Donoho & Johnstone, 1994). Here, σ is estimated using wavelet coefficients of last level (finest) because this is the level that contains more noise than signal. Its value is given by:

$$\sigma = \frac{MAD(d_{j-1,k})}{0.6745} \tag{5.7}$$

where MAD is the Medium Ablosute Deviation of the magnitudes of all the coefficients at finest level and factor 0.6745 rescales σ in order to be a suitable estimator.

5.4.1 Universal threshold

Donoho & Johnstone (1994) proposed a minimax estimator for f in Equation 5.1 based on ST and HT as well as on a parameter λ_n obtained through the minimization of a theoretical upper limit. It is defined as:

$$\lambda_U = \sigma\sqrt{2\log n} \tag{5.8}$$

where σ can be obtained as described earlier and n is the number of samples.

This threshold is called *universal* and its main advantages is that ensures that every sample in WT in which the underlying function is zero will be estimated as zero (Antoniadis et.al, 2001). Also, is not dependent on wavelet scale. Its main drawback is that it leads to estimates that underestimate underlying function, since it has the tendency to eliminate to many coefficients especially at finest scales (Morretin, 1997)

5.4.2 Adaptive (Sureshrink) threshold

This scheme proposed also by Donoho & Johnstone (1995) and uses the wavelet coefficients at each level j in order to choose a value for threshold λ_S. Stein's (1981) unbiased risk criterion (SURE) is employed in order to get an unbiased estimate of the l^2 risk.

If in level j we have n_j coefficients then Sureshrink threshold is defined as:

$$\lambda_S = \arg\min_{0\le t\le\sqrt{2\log(n_j)}} SURE(y_i, t) \tag{5.9}$$

Where

$$SURE(y_j, t) = n_j - 2\sum_{k=1}^{n_j} I\left\{\left|y_j, k\right| \le t\sigma_j\right\} + \sum_{k=1}^{n_j}\left\{\frac{y_j, k}{\sigma_j^2} \wedge t^2\right\} \tag{5.10}$$

The only drawback of Sureshrink is that when the wavelet coefficients are characterized from noticeable sparseness the noise contributed by many coefficients where the signal is zero swamps the information contributed the few coefficients where the signal is non-zero (Donoho, 1995). To overcome this special case a hybrid Sureshrink (HS) is proposed (Donoho & Johnstone, 1995) which is defined as:

$$\lambda_j^{HS} = \begin{cases} \lambda_j^U & \text{if } \sum_{k=0}^{2j-1} \hat{d}_{jk}^2 \le \hat{\sigma}^2 2^{j/2}\left(2^{j/2} + j^{3/2}\right) \\ \lambda_j^S & \text{otherwise} \end{cases} \qquad (5.11)$$

The main idea behind HS is that if the set of wavelet coefficients found to be sparse then the HS replaced by Universal threshold; otherwise SURE criterion is used to select the λ value.

5.4.3 Cross-Validation threshold

Cross-Validation (CV) is a method for estimation the prediction error for a fitted model to the data. The prediction error measures the fit success when it predicts a future sample. In terms of regression the predicted error is defined as:

$$P_{err} = E(y' - y'')^2 \qquad (5.12)$$

Where *y'* is the future sample and *y''* is the prediction using a particular model. An obvious solution is to use the mean square error (MSE) of the residuals as an estimate of P_{err}. In practice it's proved that this is a very optimistic expectation because using the same data to fit and evaluate regression model we lead to an underestimation of P_{err}. CV uses part of data samples to estimate the regression model while the remaining data samples used to evaluate that the estimated model is adequate. A classical algorithm for CV introduced by Efron & Tibshirani (1993):

Assume that we have *n* data samples ($y_1, y_2 \ldots y_n$). For each y_i :

- Fit the model and then revoke this sample
- Compute the predicted value of *i*-th sample, $\hat{y}_i$
- Compute $CV = n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ as an estimation of P_{err}.

Based on this idea Nason (1994) proposed that in order to directly apply the DWT, the original data must initially split into two subsets of equal size: one containing the even-indexed samples, and the other, the odd-indexed. The odd-indexed samples used to predict the even-indexed and vice versa. To be more specific, the goal is to estimate λ by minimizing the mean integrated squared error (MISE) between λ and underlying function *f*, defined as:

$$M(\lambda) = \mathrm{E}\int\left[\hat{f}_\lambda(x) - f(x)\right]^2 dx \qquad (5.13)$$

where $\hat{f}_\lambda$ is the wavelet threshold estimator of f in the model defined in Equation 5.12, if we use the threshold λ. Estimator $M(\lambda)$ is used because f is unknown. The implementation algorithm is:

- Samples y_i with odd indexes are removed form data set. The remaining *m-1* even indexed samples are reindexed from j=1 to 2^{m-1}
- From reindexed y_i an estimator $\hat{f}_t^{\,E}$ calculated
- From the removed odd indexed data, an interpolated version of them is calculated as:

$$\overline{y_i^O} = \begin{cases} \dfrac{y_{2j-1} + y_{2j+1}}{2} & \text{for } j = 1,.....,\dfrac{n}{2} - 1 \\ \dfrac{y_1 + y_{n-1}}{2} & \text{for } j = \dfrac{n}{2} \end{cases} \qquad (5.14)$$

- Replace previous step for $\hat{f}_t^{\,O}$ and $\overline{y_i^E}$ respectively
- An estimate of M(λ) is then calculated as:

$$M(\lambda) = \sum_{\xi} \left[\left(\hat{f}_{t,j}^{\,E} - \overline{y_i^O} \right)^2 + \left(\hat{f}_{t,j}^{\,O} - \overline{y_i^E} \right)^2 \right] \qquad (5.15)$$

$M(\lambda)$ is estimated by using *n/2* samples and therefore a correction for the sample size is needed. Nason (1995) used a heuristic method to estimate *leave-out-half cross-validation* threshold as:

$$\lambda_{CV} = \frac{\lambda_{\min}}{\sqrt{1 - \dfrac{\log 2}{\log n}}} \qquad (5.16)$$

where $\lambda_{\min} = \arg\min_{\lambda \geq 0} \mathrm{M}(\lambda)$.

5.4.4 Multiple Hypothesis Testing approach

The idea that wavelet thresholding can be viewed as a multiple hypothesis testing problem was proposed by Abramovich & Benjamini (1995, 1996). For each wavelet coefficient d_{jk} test is performed in the following hypothesis:

$$H_0: d_{jk} = 0 \text{ against } H_1: d_{jk} \neq 0.$$

If H_0 is rejected, d_{jk} is retained otherwise is discarded.

Let D be the number of d_{jk} coefficients that are retained in the model. Of those D, R are correctly kept and $G=D-R$ erroneously kept. In this case the error is $s=G/D$. The *false discovery rate* (FDR) of d_{jk} reflects the expected amount of erroneously retained coefficients among the ones finally kept; this is defined as the expectation (E) of s. In order to achieve the best possible maximization of d_{jk} coefficients that will retained to the model according rule $Es < \alpha$, Abramovich & Benjamini (1995, 1996) proposed the following algorithm:

- Take $j_0=0$. For each d_{jk} {j=0,1,…J-1; k=0,1,…2^{J-1}} calculate the 2—sided value p_{jk} (for H_0: $d_{jk}=0$)

$$p_{jk} = 2\left(1-\Theta\left(\frac{|d_{jk}|}{\sigma}\right)\right)$$

  where Θ is the cumulative distribution function of a standard normal random variable.

- Order p_{jk} values according to their size (each p_{jk} corresponds to a d_{jk})
- Find $m = \max\left\{i : p_i < \left(\frac{i}{m}\right)a\right\}$. $(0.01 < \alpha < 0.05)$ For this m calculate

$$\lambda_{FDR} = \sigma \cdot \Theta^{-1}\left(1-\frac{p_m}{2}\right) \qquad (5.17)$$

- Use threshold value λ_{FDR} and apply soft or hard thresholding τo d_{jk}

5.4.5 Recursive Hypothesis Testing approach

The λ_{FDR} discussed at 5.4.4 is a global threshold derived from a hypothesis testing that seeks to include as many d_{jk} is possible. Since this approach is not successful with signals that the underlying function has few features covered by noise, Odgen & Parzen (1996) propose an alternative hypothesis testing approach: The d_{jk} will include only when there is strong evidence that they are needed for reconstruction and the calculated threshold will be level-dependent.

Let $X_1,X_2,\ldots,X_s$ (i=1,2,…s) be independent random variables that represent the d_{jk} at level $j=j_0,\ldots,J-1$ with $c=2^J$. Let also Ξ_s represents a non-empty subset of indices {1,2,..,s}. Under these assumptions multiple hypothesis testing problem could be expressed as:

$$H_0 : \mu_i = 0, i \in \Xi_s \text{ against } H_1 : \mu_i \neq 0, \ \forall i \in \Xi_s \ , \mu_i = 0, \ \forall i \notin \Xi_s \qquad (5.18)$$

where $\mu = \{\mu_1, \mu_2, \ldots \mu_s\}$ is the mean vector.

If the number of the elements of Ξ_s in known, say equal to l, then the standard likelihood ratio test would be the sum of squares of l largest X_i elements. In practice, l is usually unknown so a recursive testing approach for Ξ_s using only one element is the proposal of Odgen & Parzen (1996). Using this approach the suitable test statistics is the largest of the squared X_i elements. The critical point α for this distribution is then estimated as:

$$x_s^a = \left\{ \Theta^{-1} \left(\frac{(1-\alpha)^{1/s} + 1}{2} \right) \right\}^2 \tag{5.19}$$

The proposed algorithm consists of the following steps:

i. Check $X_i^2 > x_s^a$

ii. If step i) is false then there is no strong indication that significant signal exists among wavelet coefficients. The threshold λ_{Rj} (at current level j) receives the value of absolute largest remaining X_i

iii. If step i) is true this is an indication that significant signal exists among wavelet coefficients. Remove largest X_i , set $s=s-1$ and return to step i)

The above algorithm could be interpreted as a ST rule using level-dependent thresholds and is called *thrsda*. The critical point is the value of α. A small value would lead to small number of wavelet coefficients included in reconstruction producing a smoother estimate. On the other hand, large value produces less smother estimate since the number of included wavelet coefficients is higher. In this study α=0.06.

5.4.6 Shift Invariant threshold

Wavelet thresholding using DWT sometimes dominated by artifacts of various kinds. These artifacts are more obvious near discontinuities where they exhibit as pseudo-Gibbs phenomena, over and under shooting at certain levels. In order to improve suppression of these artifacts Coifman & Donoho (1995) proposed the translation invariant wavelet thresholding scheme. The basic idea is to correct possible misalignments between important features of underlying function and features of wavelet basis. This idea implemented by applying a range of shifts and then averaging.

The lack of shift invariance is not necessarily a problem for many applications but denoising is not one of them. The problem arises from the fact that the DWT of a signal X_i and the DWT of a shifted version of X_{i+n} are different. They are different not under the scope of coefficients similarity but under the scope of *coherency*. The coherence of the signal is a measure of how compact the signal is represented. Compact representation is the one that needs few coefficients to represent the signal. Difference in coherency means that DWT of X_i needs different number of coefficients than DWT of X_{i+i} in order to represent it compactly. A typical example presented in Figure 5.4a and b

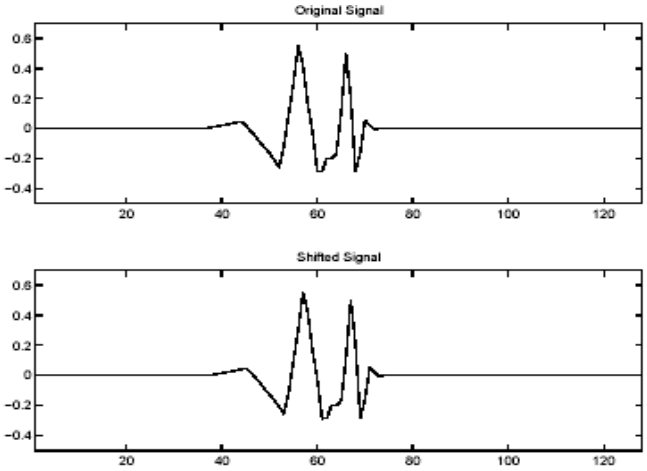


Figure 5.4a. Two similar signals where the bottom is the right shifted version of top

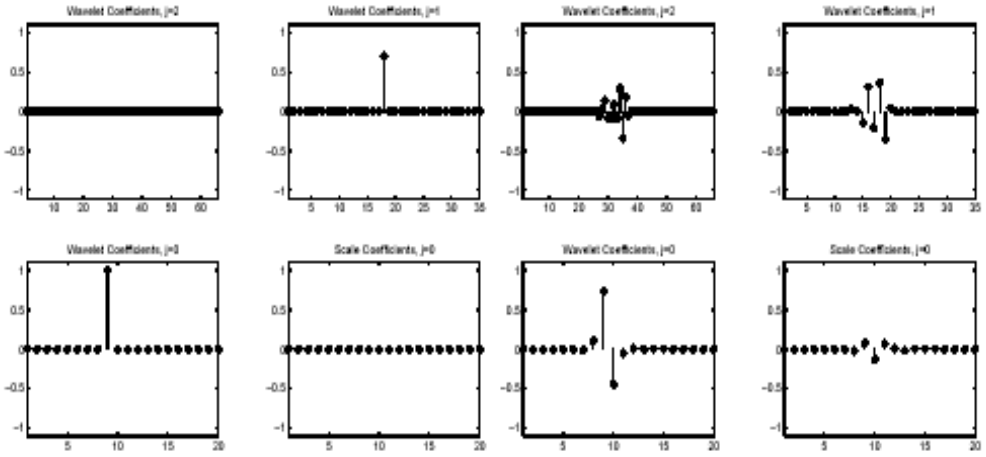


Figure 5.4b. DWT of original signal (four left plots) and DWT of shifted version (four right plots) (Adams et.al., 2001)

White noise is, in general, incoherent. That means that there exists no representation in which white noise can compactly specified. Under this point of view a successful denoising approach is the one that retains the coherent part and reject tha coherent part of a sampled signal. For periodic signals Fourier transforms provide quite clear coherent representations. For non-periodic and more even for non-stationary signals FT fails. This is where DWT comes to fill this gap but without being the ultimate tool. As presented in Figures 5.4a & b the coherence of DWT of a signal is highly dependent on the time shift of this signal. In other words there is a strong possibility that a time shifted version of a signal may be more coherent (in relation with its DWT representation) than the original signal. It is natural to assume that denoising performance could be improved if one found the optimal shift that presents the most coherent representation, denoising the shifted signal and then unshift it. Since there are many algorithms that could find this optimal shift there is no general shift that could be applied everywhere.

Coifman & Donoho (1995) proposal is based on denoising all possible shifts and then averaging. Let S_h represent the shift operator. For a signal $f(k)$ $\{k=1,2,\ldots n\}$ $S_h f(k)=f(k+h \bmod n)$. The denoised signal $\hat{f}$ using *translation invariant wavelet threshold* λ_{TI} is defined as:

$$\hat{f} = \frac{1}{n}\sum_{h=0}^{n-1}(S_h^{-1}\cdot W_f^{-1})\lambda_{\mathrm{TI}}(W_f S_h) f \qquad (5.20)$$

where W_f and W_f^{-1} are the unshift DWT and inverse DWT of f respectively.

5.5 Wavelet Group Thresholding estimators

5.5.1 Introduction

The discussed thresholding estimators designed to apply at wavelet coefficients through term-by-term basis; each d_{jk} is compared with a threshold λ estimated by one of the previous methods. If $|d_{jk}| > \lambda$ then this coefficient is retained otherwise is rejected. Term-by-term thresholding compromise between variance and bias contribution to mean square error. However, it is not the optimal solution because usually it removes too many coefficients resulting to a biased estimator.

An obvious solution is to utilizing information about neighboring wavelet coefficients. Instead of thresholding individual d_{jk} the rules and estimators that

discussed in previous sections could be applied in groups (blocks) of coefficients. This approach provides more "information" about the decision of retain or discard the examined wavelet coefficient. This is done because the amount of information that is available from the signal for estimating the average wavelet coefficient (within a group) is at least two times larger than the term-by-term case. This increase the success rate of threshold decisions since they will be more accurate and thus improve the convergence rates.

5.5.2 A non-overlapping block thresholding estimator

Cai (1999) using the approach of ideal adaptation using oracle proposed the *non-overlapping block thresholding estimator*. At each level $j = j_0,..J-1$, the d_{jk} are grouped together into non-overlapping groups of length L. If the last block is not filled then the first few necessary coefficients are used to fill (which is call *Augmented case*). In contrary, if the last coefficients are quite few then they are not used in the process (which is called *Truncated case*).

Let j_b denote the set of indices of the coefficients in the bth block at level j

$(jb) = \{(j,k) : (b-1)L+1 \le k \le bL\}$

and let S_{jb}^2 denote the L^2 energy of the noisy signal in block (jb). Estimation of d_{jk} , within each block, is done by using the James-Stein rule:

$$\tilde{d}_{jk}^{(jb)} = \max\left(0, \frac{S_{jb}^2 - \lambda L\sigma^2}{S_{jb}^2}\right) d_{jk} \qquad (5.21)$$

In order to be operational Equation 5.21 must have predefined values for L and λ. In this study we follow Cai's (1999) approach who suggests using $L = log(n)$ and λ=4.50524 .

As was the case in term-by-term methods, estimation of the underlying function is obtained by using IDWT at scaling c_{jk} $\{k=0,1,\ldots 2^{jo}-1\}$ and thresholded $\tilde{d}_{jk}^{(jb)}$ wavelet coefficients $\{j=j_0,\ldots J-1;\ k=0,1,\ldots 2^j-1\}$. The resulting estimator denoted as *BlockJS.*

5.5.3 An overlapping block thresholding estimator

Modification of non-overlapping block thresholding estimator proposed by Cai & Silvermann (2001). Their idea was that the treatment of the wavelet

coefficients that take place in the middle of each block is highly dependant from the data of the whole block.

The procedure begins by grouping the wavelet coefficients d_{jk} into non-overlapping blocks (jb) of length L_0 at each level $j=j_0,\ldots.J$-1. Each block is extended by an amount $L_1=\max(1,[L_0/2])$ in both directions in order to construct blocks (jB) with length $L=L_0+2L_1$. Let S_{jB}^2 denote the L^2 energy of the noisy signal in block (jB). Within each block (jb) the appropriate wavelet coefficients are estimated as:

$$\breve{d}_{jk}^{(jb)} = \max\left(0, \frac{S_{jB}^2 - \lambda L\sigma^2}{S_{jB}^2}\right) d_{jk} \qquad (5.22)$$

The typical values of L and λ are:

- $L_0 = \log\frac{n}{2}$ and λ=4.50524 .This scheme called *NeighBlock estimator* and used neighbouring coefficients outside the block of interest in fixing the threshold
- $L_0=L_1=1$ and $\lambda = \frac{2}{3}\log n$. This scheme is called *NeighCoeff estimator* and chooses a threshold for each coefficient by reference to the coefficients itself as well as to its neighbours.

Estimation of underlying function done using IDWT as described in 5.5.2.

5.6 Wavelet Bayesian Thresholding

In recent years, Bayesian non-linear wavelet thresholding has attracted the attention of a number of researchers. A comprehensive review and comparative study of Bayesian wavelet denoising approaches are presented by Antoniadis et al. (2001). Several recent studies (Abramovich et.al, 2002, Abramovich & Sapatinas, 1998, Chipman et. al, 1997, Clyde & George 2000, Vannucci & Corradi, 1999, Vidakovic 1998) have demonstrated that Bayesian wavelet thresholding methods are advantageous over the classical wavelet thresholding approaches in the some situations. In this section a Bayesian approach will be discussed.

Bayesian approaches to choosing the shrinkage method are less ad-hoc than earlier methods, and have been shown to be effective. It is known that, in general, Bayes rules are “shrinkers” and that their shape in many cases has a desirable property

for wavelet shrinkage: it can heavily shrink small arguments and only slightly large arguments. If we use Bayes models for the wavelet coefficients the resulting optimal actions can be very close to thresholding.

Suppose that the measured data is contaminated by an additive white Gaussian noise, $\varepsilon(t_i)$ as follows:

$$f(t_i) = g(t_i) + \varepsilon(t_i), \quad i = 1.2.,,,,N \tag{5.23}$$

where $f(t_i)$ and $g(t_i)$ represent the noisy and denoised experimental data, respectively. The noise term $\varepsilon(t_i) \sim N(0,\sigma^2)$ is a vector of independent, identically distributed (IID) errors with the mean of zero and the variance of σ^2 in which the symbol ~ denotes a distribution. The denoising problem is equivalent to a univariate non-parametric regression problem. The goal of the Bayesian denoising approach is to recover the underlying noise-free time series $g(t_i)$ from the experimental measurements $f(t_i)$.

In the wavelet decomposition of time series decomposition of the time series, the noise item $\varepsilon(t_i)$ in Equation 5.23 is also resolved into a series of corresponding noise ε_{jk}: Thus, the decomposed coefficients d_{jk} can be expressed by

$$\hat{d}_{jk} = d_{jk} + \sigma_j \varepsilon_{jk}, \quad j = j_0,.....J-1; k = 0,1,....2^j - 1 \tag{5.24}$$

where $\varepsilon(t_i) \sim N(0,1)$ are independent random variables and d_{jk} are noise free coefficients. The distribution of the decomposed coefficients on d_{jk} and σ_j^2 is expressed as:

$$\hat{d}_{jk} \left| d_{jk}, \sigma_j^2 \right. \sim N(d_{jk}, \sigma_j^2) \tag{5.25}$$

Equation 5.25 is the conditional distribution form of Equation 5.24. Following Vidacovic (1998), a non-informative prior distribution of d_{jk} can be assumed as follows:

$$d_{jk} \left| \gamma_{jk} \right. \sim N(0, \gamma_{jk}, \tau_j^2) \tag{5.26}$$

where γ_{jk} is a binary random variable with independent Bernoulli distribution $\{ P(\gamma_{\xi\kappa} = 1) = 1 - P(\gamma_{\xi\kappa} = 0) = \pi_j \}$. It determines whether the coefficients are zero (γ_{jk}=0) or non zero (γ_{jk}=1). The variance τ_j^2 represents the magnitude of d_{jk} at jth decomposition level. In practical applications the same values of τ_j^2 and π_j are assigned for all coefficients in the jth level. The values of π_j = 0.5 and τ_j = 1 are chosen for the non-informative prior in currect study. The standard deviation σ_j is

estimated from the wavelet coefficients of the *j*th level of wavelet decomposition through dividing the median of the wavelet coefficients by a factor. The value of 0.6745 suggested by Percival and Walden (2000) for the factor will be used.

Based on the Bayes' rule, the posterior distribution is obtained form Equations 5.25 and 5.26 as follows:

$$d_{jk} \mid \gamma_{jk}, \hat{d}_{jk}, \sigma_j^2 \sim N\left(\gamma_{jk} \frac{\tau_j^2}{\sigma_j^2 + \tau_j^2} \hat{d}_{jk}, \gamma_{jk} \frac{\sigma_j^2 \tau_j^2}{\sigma_j^2 + \tau_j^2} \right) \tag{5.27}$$

and the marginal posterior distribution of d_{jk} conditionally on σ_j^2 is given by a mixture model described as follows (Vidacovic, 2000):

$$d_{jk} \left| \hat{d}_{jk}, \sigma_j^2 \right. \sim p(\gamma_{jk} = 1 \left| \hat{d}_{jk}, \sigma_j^2 \right.) \xi(d_{jk}) + \left[1 - p(\gamma_{jk} = 1 \left| \hat{d}_{jk}, \sigma_j^2 \right.) \delta(0) \right] \tag{5.28}$$

where $\xi(d_{jk})$ is the posterior distribution of Equation 5.27 which describes the non-zero d_{jk} with the probability of

$$p(\gamma_{jk} = 1 \mid \hat{d}_{jk}, \sigma_j^2) = \frac{1}{1 + \eta_{jk}} \tag{5.29}$$

Where η_{jk} is the posterior odds ration of $\gamma_{jk} = 0$ versus $\gamma_{jk} = 1$, given by

$$n_{jk} = \frac{1 - \pi_J}{\pi_j} \frac{(\sigma_j^2 + \tau_j^2)^{1/2}}{\sigma_j^2} \exp\left(- \frac{\tau_j^2 \hat{d}_{jk}^2}{2\sigma_j^2 (\sigma_j^2 + \tau_j^2)} \right) \tag{5.30}$$

The decomposed coefficients are thresholded based on the following hypothesis test:

$$H_0 : d_{jk} = 0 \text{ versus } H_1 : d_{jk} \neq 0 \tag{5.31}$$

The following thresholding rule is used to test the hypothesis and threshold the wavelet coefficients (Antoniadis et.al, 2001):

$$\tilde{d}_{jk} = \hat{d}_{jk} \, I(\eta_{jk} < 1) \tag{5.32}$$

where $I(.)$ is an indication function. It is equal to unity when $\eta_{jk} < 1$. Thus, is H_0 rejected and the coefficient d_{jk} is estimated by $\hat{d}_{jk}$. Otherwise is equal to zero and $\hat{d}_{jk}$ is thresholded. Since η_{jk} in Equation 5.30 represents the Bayes factor,

$$P\left(\frac{\left(H_0\middle|\hat{d}_{jk}\right)}{\left(H_1\middle|\hat{d}_{jk}\right)}\right) \tag{5.33}$$

Equation 5.32 is also called *Bayes factor thresholding rule* (BFTR) (Vidacovic, 1998). The thresholded coefficients $\tilde{d}_{jk}$ are then used to reconstruct the denoised function as usual. Figure 5.5 depicts the Bayesian WT denoising approach which is done in five steps:

1) Measured data are decomposed by WT
2) non-informative prior distributions are imposed on the decomposed coefficients
3) posterior distributions of the coefficients are derived based on the Bayes' theorem
4) the noise is removed from the coefficients through a Bayesian hypothesis testing
5) the denoised data is reconstructed through an inverse WT

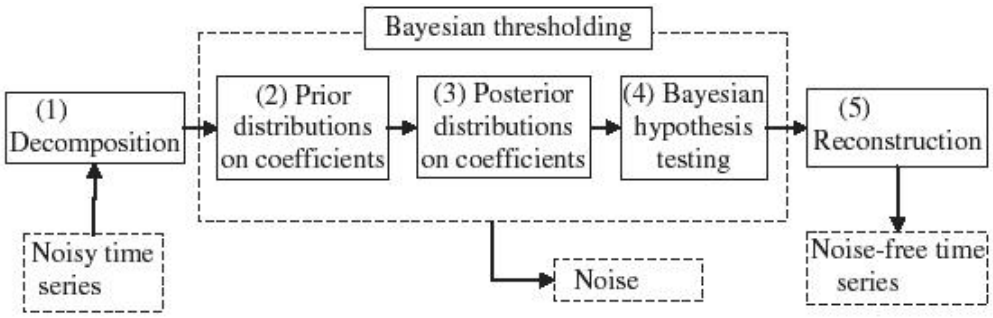


Figure 5.5. Outline of Bayesian WT denoising approach.

Except BFTR, the following Bayesian implementations used in this study:

- *BAMS* (Bayesian Adaptive Multiresolution Smoother) proposed by Vidacovic & Ruggeri (2000).They obtained obtained wavelet shrinkage estimates by putting a distribution on σ^2 and considering a prior distribution for the wavelet coefficients d_{jk} that is similar in spirit to the prior mixture model.
- *Decompsh* (Deterministic – Stohastic Decompositions) proposed by Huang & Cressie (2000). Since all the Bayesian approaches described previously to obtain wavelet shrinkage and wavelet thresholding estimates, assumed a prior

for each wavelet coefficient d_{jk} with zero mean, they relaxed this assumption by assuming that the underlying signal is composed of a piecewise smooth deterministic part plus a zero mean stochastic part.

5.7 Performance measures

In order to measure the performance of the various methods that presented before, the following measures has been used (where $f(t_i)$ and $g(t_i)$ represent the noisy and denoised experimental data, respectively):

- Signal to Noise Ratio (*SNR*).

$$SNR = 10\log_{10}\left(\frac{\sum_{i=1}^{N}\left|f(t_i)\right|^2}{\sum_{i=1}^{N}\left|f(t_i)-g(t_i)\right|^2}\right)[db] \qquad (5.34)$$

- Mean Square Error (MSE)

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left|f(t_i)-g(t_i)\right|^2 \qquad (5.35)$$

The above measures used widely in denoising studies and the information that produce can be characterized as "quantitative" since they provide the numerical proofs for the performance of a denoising method. The "qualitative" approach requires measures that can examine the success of a method regarding its ability to preserve the shape of the seismic signal. More specific, it is of high interest to preserve the shape of the P seismic phase arrival (P onset) since it's the key for the extraction of useful conclusions for the location of an earthquake, i.e. the hypocentral coordinates and the origin time of the seismic event along with seismological network operation, early warning systems and structure of the earth. For this reason, one more performance measure will be used: the P-wave delay Δt_p which is defined as the delay between denoised and original detected initial P wave pulse. The denoised seismogram will be always the one with the better SNR.

It is well known that conventional filters introduce some delay to the signal resulting to false time arrivals (Scherbaum, 1996). In general, wavelet denoising methods improves this drawback and as it can be seen from Table 5.9 there are small differences between them. An illustration of this situation is shown at Figure 5.13, where two denoised signals are overlapped to the original in order to get an illustrative view of the t_p differences.

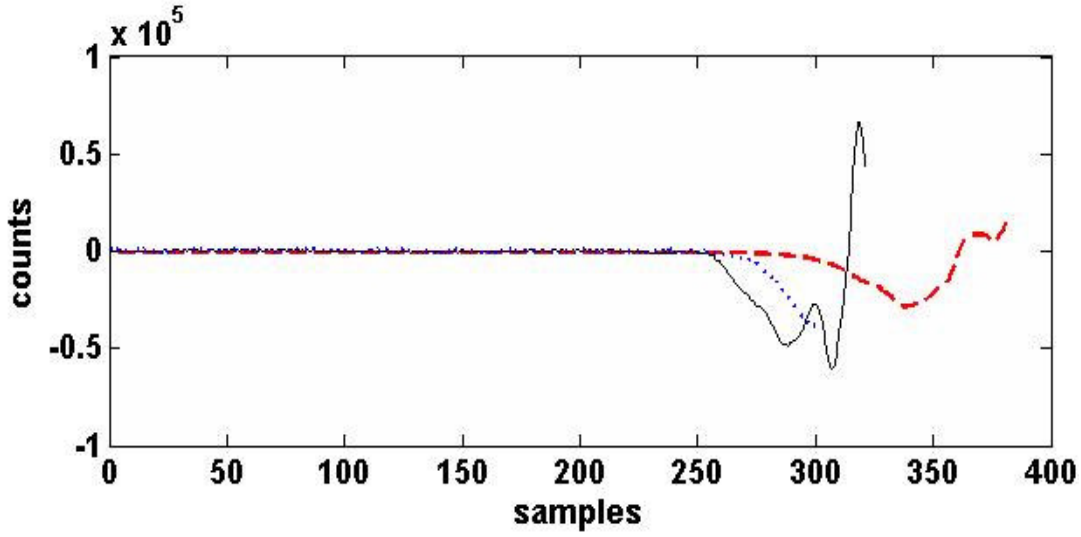


Figure 5.6. Time differences for P-wave arrival among the real signal (solid black line), denoised signal after filtering with a Chebyshev type I filter (red dashed line) and denoised signal after filtering with a wavelet (TI-H) method (blue dotted line).

This time difference could create an important error in earthquake's location estimation. Even it is not used in every day seismological practice it is useful to present the following simplified example that educational demonstrate the influence of P-wave arrivals in the accuracy of earthquake location. The event's distance from each station is empirically calculated by multiplying by 10 the difference between S-wave and P-wave arrivals and then correspond this time difference to distance by using travel-time graphs, such as the one in Fig. 5.8. It is more than obvious that time differences that produced by filtering can alter the estimation of the epicentre by more or less kilometres. A typical example of this situation is the location difference that produced by using the seismograms (from 3 stations) at Fig. 5.6. This situation is depicted in Figure 5.9 where the estimated location from the recordings of three stations which are filtered by using a Chebyshev type I filter varies significantly from the correct location (as this announced by other seismological networks, EMSC and THE).

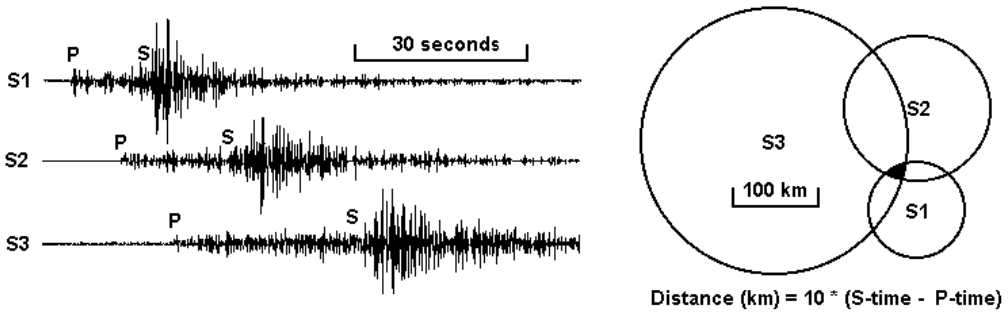


Figure 5.7. Location using the circle method. Stations S1, S2, S3 record an earthquake (left). The estimated epicentre located at inside black common area of three circles (right)

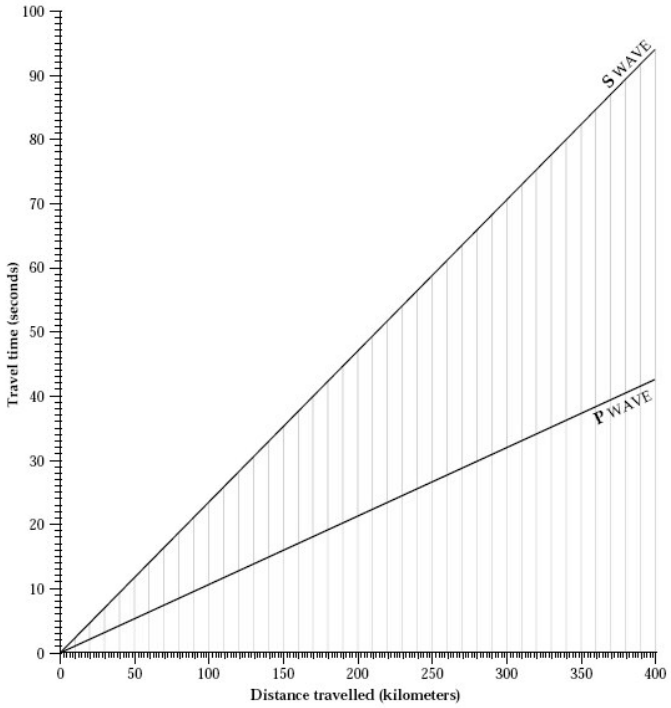


Figure 5.8. Travel time graph

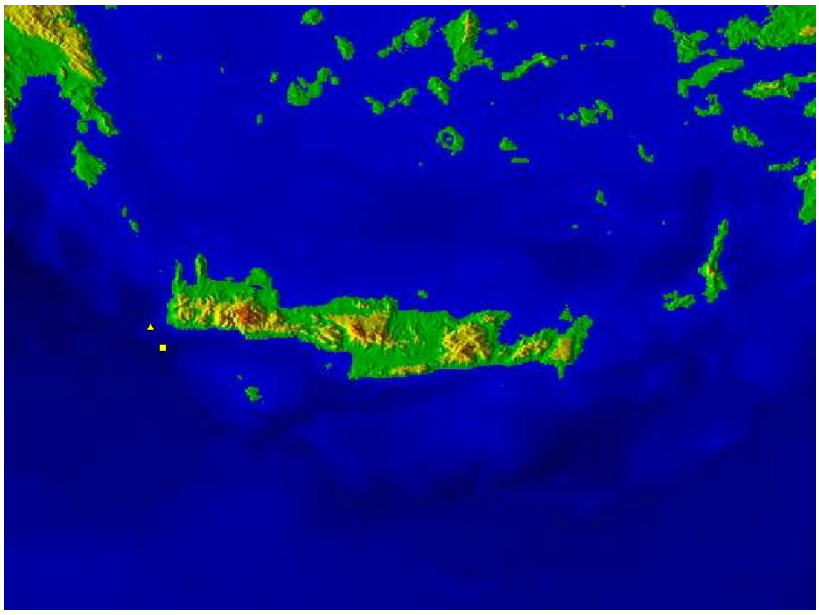


Figure 5.9. Epicentral location difference (triangle: correct location, square: wrong location).

Also important is the amplitude difference that could possibly produced by filtering because the maximum amplitude in a seismogram used in earthquake's magnitude estimation and especially in local Richter magnitude (Lay & Wallace, 1995) as shown in Fig.5.10

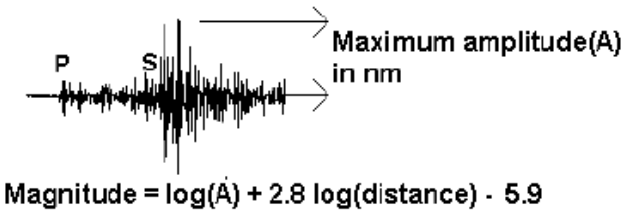


Figure 5.10. Magnitude determination using max amplitude of S-waves

5.8 Synthetic example

In this section all the discussed methods will apply to a set of known seismological signals corrupted by artificial noise. This comparison will reveal the performance of each method and at the same time will lead to the optimization for some critical parameters of each method when needed.

5.8.1 Description of the synthetic signals

The initially signal (Figure 5.11) is a seismogram that follows the typical behaviour of recorded seismograms at Seismological Network of Crete (SNC hereafter).

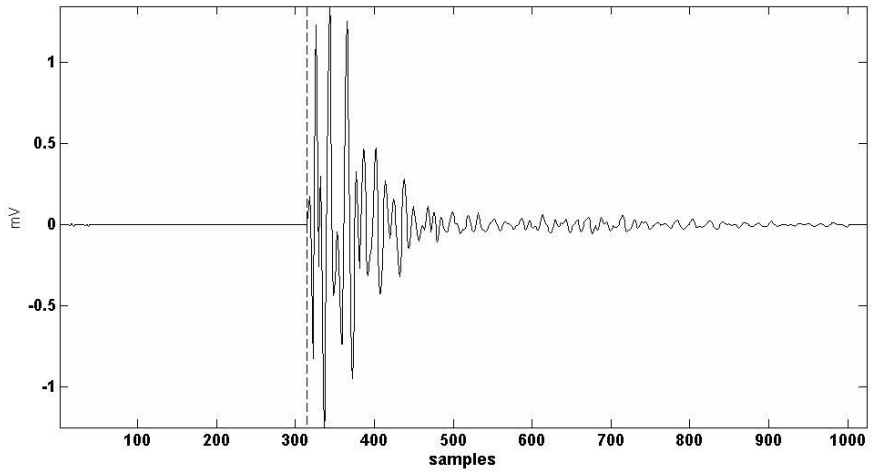


Figure 5.11. Synthetic seismic signal, P-onset time (at sample 315) pointed out by dashed line

In order to explore the validity of the proposed techniques the seismic recording at Figure 5.11 contaminated randomly with different types of noise: a) Gaussian white noise, b) colored noise c) ambient noise recorded from seismic stations of SNC). The contamination applied using different SNR levels (from -10db to +60db). The SNR of the original signal is defined between pre-event portion (0-300 samples) and post-event portion (300-600 samples). Each signal is name-coded using the format *XXn* where *XX* declares the nature of the added noise (GN: white Gaussian, CL: colored noise, AQ: noise without spikes, AS: real noise with spikes) and *n* defines the serial number. Figures 5.12a, 5.12b and 5.12c present three typical noise-contaminated signals and Table 5.1 summarizes the catalog of signals that will be used.

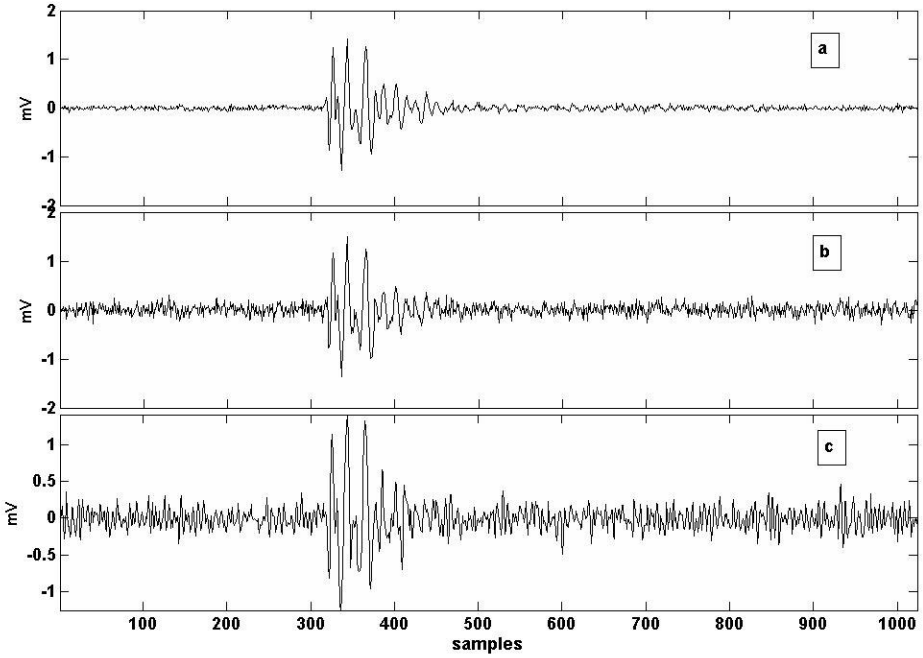


Figure 5.12. Seismic signal contaminated with different types of noise. Subplot *a* is a CL type noise, subplot *b* is a GN type noise and subplot *c* is a AQ type noise (see text for details).

	Type of noise	SNR (db)	Acronym
1	Ambient noise with spikes	-10	AS1
2	Colored	-9	CL1
3	ambient noise without spikes	-8	AQ1
4	white Gaussian	-7	GN1
5	ambient noise with spikes	-6	AS2
6	Colored	-5	CL2
7	ambient noise without spikes	-4	AQ2
8	white Gaussian	-3	GN2
9	ambient noise with spikes	-2	AS3
10	Colored	-1	CL3
11	ambient noise without spikes	0	AQ3
12	white Gaussian	1	GN3
13	ambient noise with spikes	2	AS4
14	Colored	3	CL4
15	ambient noise without spikes	4	AQ4
16	white Gaussian	5	GN4
17	ambient noise with spikes	6	AS5
18	Colored	7	CL5
19	ambient noise without spikes	8	AQ5
20	white Gaussian	9	GN5
21	ambient noise with spikes	10	AS6
22	Colored	12	CL6

23	ambient noise without spikes	12	AQ6
24	white Gaussian	12	GN6
25	ambient noise with spikes	18	AS7
26	Colored	18	CL7
27	ambient noise without spikes	18	AQ7
28	white Gaussian	18	GN7
29	ambient noise with spikes	26	AS8
30	colored	26	CL8
31	ambient noise without spikes	26	AQ8
32	white Gaussian	26	GN8
33	ambient noise with spikes	32	AS9
34	colored	32	CL9
35	ambient noise without spikes	32	AQ9

Table 5.1. Catalog of seismic signals used in simulation study

5.8.2 The influence of the thresholding rule

At section 5.3 three thresholding rules are discussed: soft, hard and SCAD. In this subsection the performance of wavelet denoising in regards of threshold value will be evaluated. More specific soft and hard thresholding values will be examined in accordance of MSE in order to evaluate the denoising performance at different thresholds.

Soft and hard thresholding rules with values from 0 to 8 will be examined. SCAD is not included because in practice is used more as a thresholding estimator rather that a thresholding rule. Denoising is performed using the following basis: Haar, D(4), C(6) and LA(8). The noisy signals are: AQ1, GN3, AQ5 and CL6. Representative results of threshold value vs. MSE are shown at Fig.5.13.

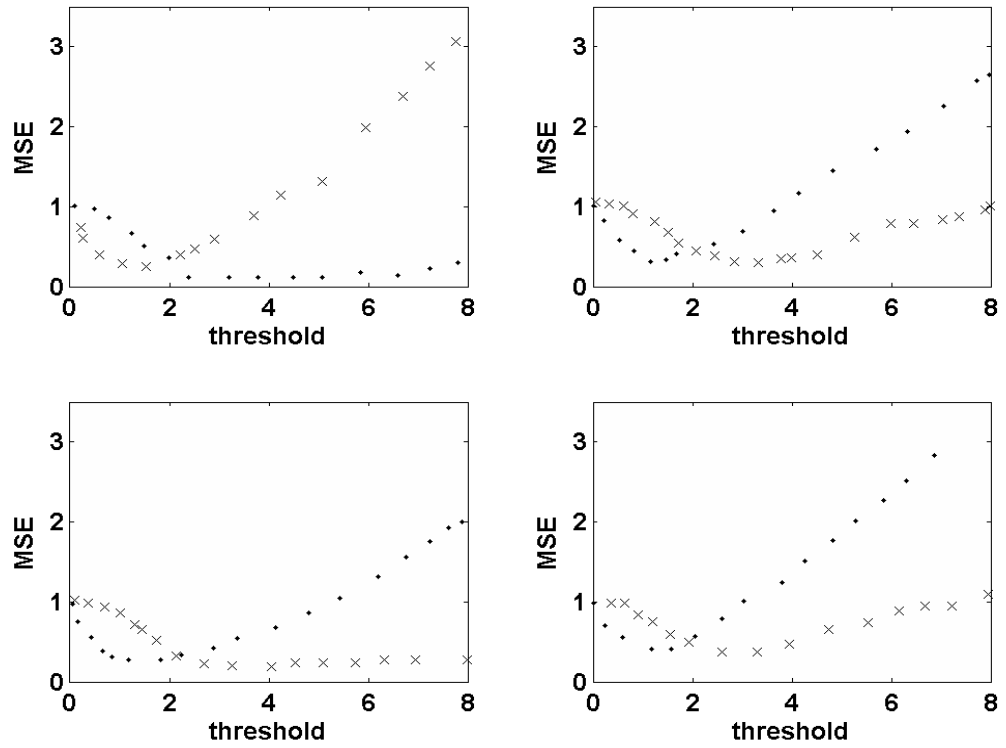


Figure 5.13 Performance evaluation of denoising ('.' for ST, 'x' for HT) using D(4) basis for GN3 (top left), CL6 (bottom left), AQ1(top right) and AQ5 (bottom right)

In each case the MSE between original and denoised signal is calculated and the results are summarized in Table 5.2

	Signal AQ1		**Signal AQ5**		**Signal GN3**		**Signal CL3**	
Basis	*Hard*	*Soft*	*Hard*	*Soft*	*Hard*	*Soft*	*Hard*	*Soft*
Haar	3.8	1.4	2.8	1.2	3.4	1.4	2.8	1.2
D(4)	3.4	1.6	3.2	1.4	3.4	1.2	3.4	1.4
D(6)	3.4	1.6	3.2	1.4	2.8	1.2	3.2	1.4
C(6)	3.2	1.4	3.4	1.4	2.6	1.2	3.2	1.4
LA(8)	3.4	1.6	3.4	1.4	2.8	1.2	3.1	1.4

Table 5.2. Best numerically found thresholds in terms of MSE of denoised signals

The expected theoretical thresholds can be estimated by using the Eq.5.8. The theoretical calculated value is 3.7 which is close to the experimental hard thresholding values but not to the soft thresholding ones. This is a well known unsolved problem in wavelet literature (Lorenz, 1998, Graham, 1999, Adams et al, 2001) originating from

the fact that Eq.5.8 does not distinguish between soft and hard thresholds. Up to now there are no equations that can be used to calculate different soft and hard thresholds theoretically and for this reason a preliminary survey such the one here is useful for estimating the limits of the forthcoming soft and hard thresholds according to the nature of noisy signals.

5.8.3 The influence of the DWT basis

Another choice that affects globally the denoising performance, except threshold, is the choise of the DWT basis that used for signal's decomposition. In this section the effect of DWT basis will be evaluated.

The settings that will be used for this simulation are:

- The original and noisy signals are the same as in section 5.8.2
- In each denoising procedure, the best threshold (as obtained in 5.8.1) will be used.
- Performance will be measured in terms of MSE

The results presented in Table 5.3

	Signal RQ1		**Signal RQ5**		**Signal GN3**		**Signal CL3**	
Basis	*Hard*	*Soft*	*Hard*	*Soft*	*Hard*	*Soft*	*Hard*	*Soft*
Haar	0.163	0.2263	0.4014	0.3732	0.0587	0.2264	0.2761	0.3215
D(4)	0.0472	0.1482	0.1771	0.2438	0.1557	0.2785	0.1878	0.2250
D(6)	0.0518	0.1527	0.1328	0.2001	0.1897	0.2471	0.1875	0.2081
C(6)	0.05541	0.1475	0.1559	0.2118	0.1987	0.3021	0.1987	0.2878
LA(8)	0.05146	0.1430	0.09931	0.1228	0.1554	0.2417	0.1331	0.1839

Table 5.3. MSE of denoised signals, by soft & hard thresholding, with different DWT basis used.

The answer to the obvious question "why one basis could not used for all ?" is that the noise component manifests itself as small amplitude coefficients in DWT and amplitude is the discriminative factor. Therefore, it's clear that the more concentrated the signal energy is in small set of DWT coefficients, the easier is to distinguish signal and noise components. In cases where the DWT coefficients of the signal are small and widely spreaded out, these coefficients get mixed with those of the noise and the

wavelet denoising schemes yield poor performance. Therefore the best basis as interpreted from Table 5.3 gives their corresponding classes of signals the most energy concentrated DWTs.

5.8.4 Application of Wavelet Denoising to Synthetic Signal Examples

Signals presented in Table 5.1 will denoised using the methods that described earlier. For the ease of reference these methods will presented at Table 5.3 along with the thresholding rule that used with. Results from the comparison between different wavelet methods and conventional bandpass filters are presented in Appendix F.

From the results below it is obvious that there is no wavelet denoising approach that "does it all". Although some of them have a high performance in the majority of cases cannot be used *a-priori*. In contrast, we can justify as unsuccessful the Cross-Validation method (soft and hard versions) as well as the Minimax as they does not present better performance against conventional bandpass filtering especially in terms of P-wave SNR and MSE.

Index	*Acronym*	*Estimator Type*	*Thresholding rule*
1	SCAD	SCAD	Not specific
2	VISU-S	VisuShrink	Soft
3	VISU-H	VisuShrink	Hard
4	SURE	Sure (Classical)	Not specific
5	HYBSURE	Sure (Hybrid)	Not specific
6	CV-H	Cross Validation	Hard
7	CV-S	Cross Validation	Soft
8	MINS	Minimax	Soft
9	MINH	Minimax	Hard
10	TI-H	Translation Invariant	Hard
11	TI-S	Translation Invariant	Soft
12	NEIGHBL	NeighBlock	Not specific
13	BLOCKJS_A	Block Thresholding	Augment
14	BLOCKJS_T	Block Thresholding	Truncate
15	THRSDA	Hypothesis testing	Soft
16	FDR_H	False Discovery Rate	Hard
17	FDR_S	False Discovery Rate	Soft
18	DECOMPSH	Deterministic/Stohastic	Not specific
19	BFTR	Bayesian Hypothesis testing	
20	BAMS	Bayesian Adaptive Multiresolution	

Table 5.4. Acronyms for the set of denoised methods used.

5.9 Evaluation with real seismic data

In order to evaluate the results obtained from previous section, the aforementioned filtering approaches were applied to a set of real seismic records acquired from the SNC. These records correspond to events with various magnitudes and were selected as representatives for the specific seismological network and classified in five categories as will be shown later. The performance will be measured by using the Δt_p initially and then by estimating the epicentral error that produced due to erroneous P onset times. The MSE is rejected because, in the case of recorded data, the real denoised data is unknown.

A database of 1250 selected events (from 01/01/2004 until 30/06/2008) used for this study. Events are selected in order to fulfill two main criteria:

a) The corresponding events must be detected and bulleted by another global Seismological Institute.
b) The seismological stations' recordings must present noticeable differences in their quality.

The necessity of condition a) based on the event that we need valid time estimation for P wave arrivals in order to calculate the difference value in Δt_p calculation. This calculation performed by using the HYPO2000 (http://earthquake.usgs.gov/research/software/index.php#Hypo) software (Klein, 2002) and an inverse approach as will be presented later. The b) condition ensures that the analysis will cover the majority of typical recordings that SNC collects and analyzes.

The recordings database consists of five groups of real seismic data, graded from A to E in relation with their quality which is expressed by SNR as shown in Table 5.5.

Category	*SNR*	*Comments*
A	>30	Very good quality. Analyst can clearly identify the P wave arrivals
B	(10,30)	Good quality. Analyst needs a filter in order to identify the P wave arrivals.
C	(3,10)	Medium quality. Amplitude of the seismic event is comparable to noise amplitude. Analyst needs a filter to identify the P wave arrivals
D	(0,3)	Poor quality. There is no indication of seismic event. Analyst can hardly detect the P wave arrival after extensive filtering.
E	*	Existence of high amplitude transients and other non homogeneous disturbances

Table 5.5. Attributes of used seismic data recordings.

Each recording tested with each one of the 16 remaining denoising methods (as shown in Table 5.4). Results from representative recordings that belong to separate categories (as presented in Table 5.5) for each one of the denoising methods are presented below. For each method a combined plot between category's typical recording and denoised recording is presented. In addition for the best two methods (for each category) detailed combined plots around P arrival time are presented also. In the latter plots the trigger times from an STA/LTA detector are also depicted in order to have the ability to compare the accuracy improvement in P-wave arrival before and after denoising application.

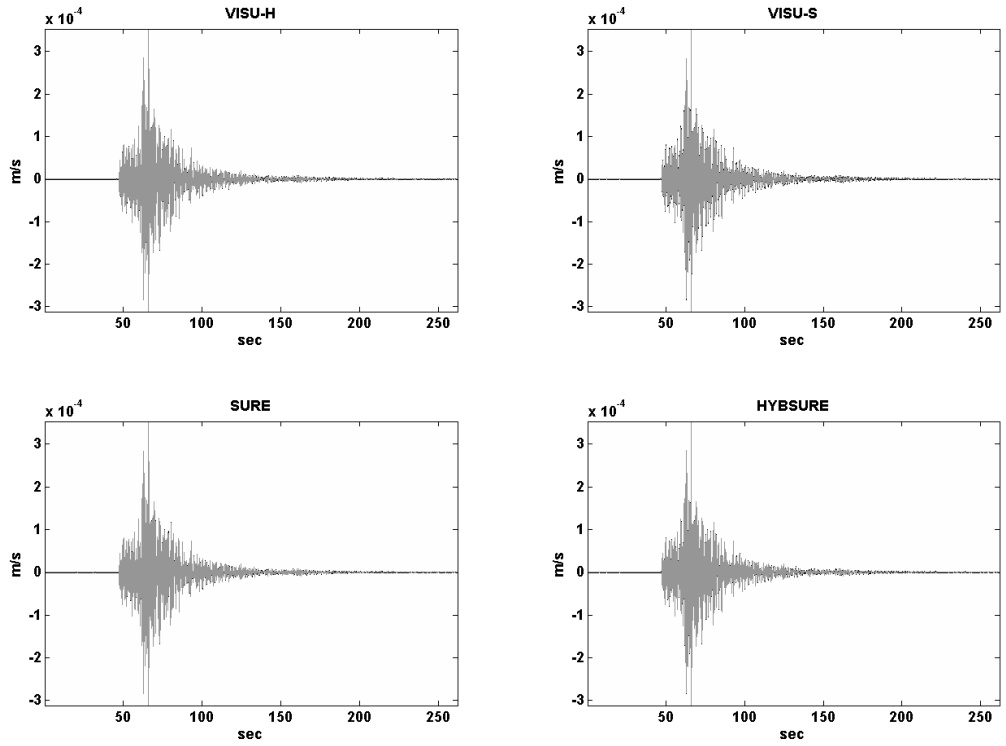


Figure 5.14 Original (black line) and denoised (grey line) recording from A category using VISU-H (top left), VISU-S (top right), SURE (bottom left) and HYBSURE (bottom right).

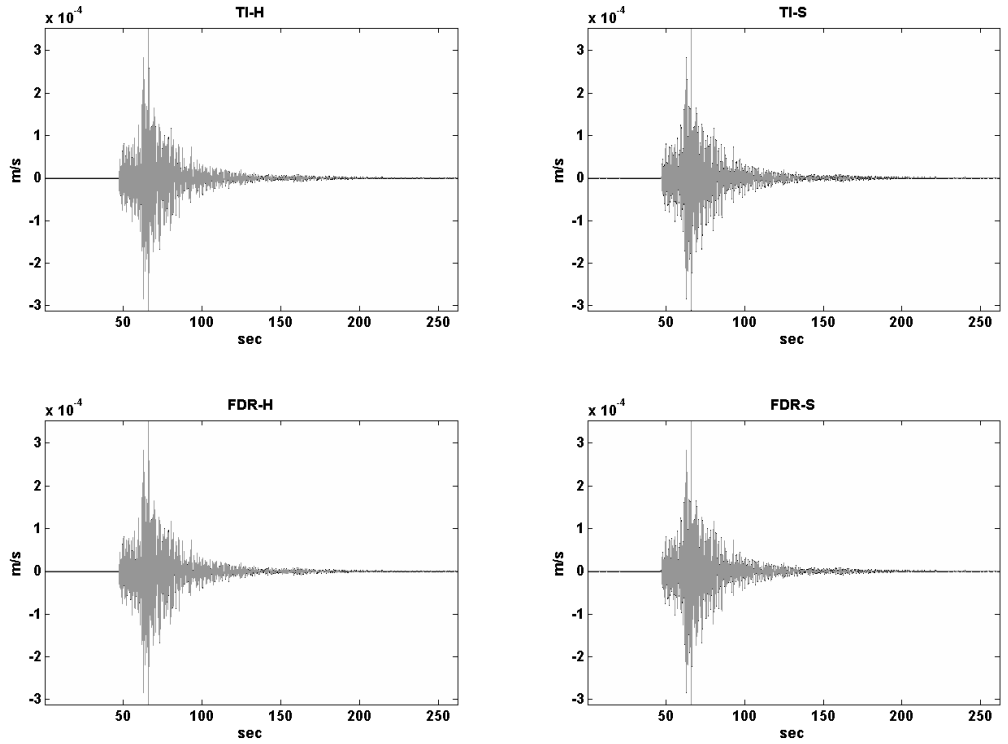


Figure 5.15 Original (black line) and denoised (grey line) recording from A category using TI-H (top left), TI-S (top right), FDR-H (bottom left) and FDR-S (bottom right).

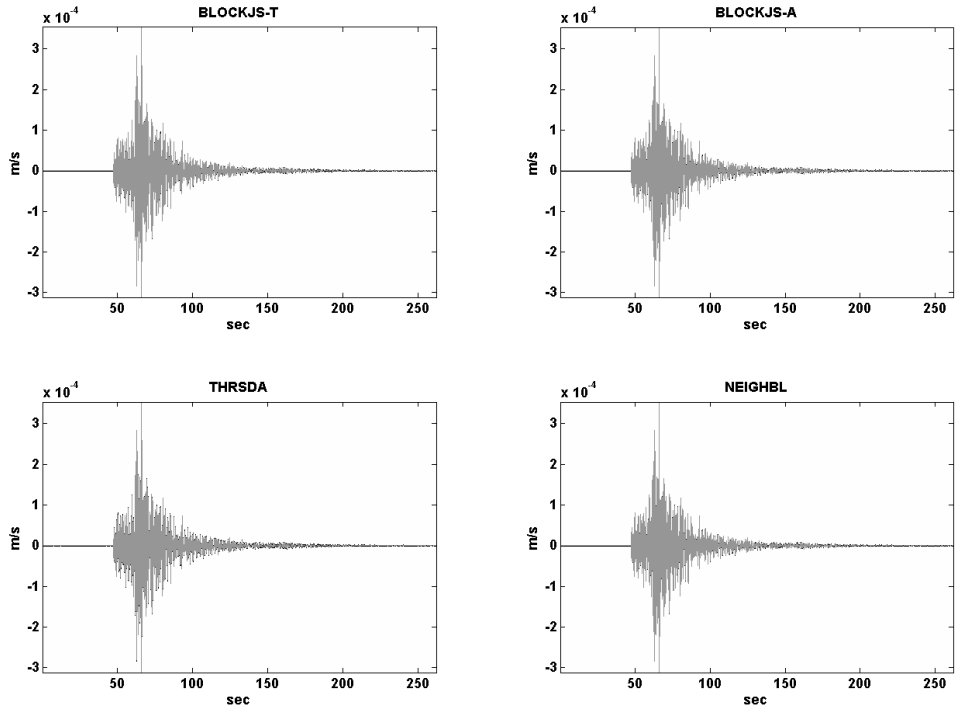


Figure 5.16 Original (black line) and denoised (grey line) recording from A category using TBLOCKJS-T (top left), BLOCKJS-A (top right), THRSDA (bottom left) and NEIGHBL (bottom right).

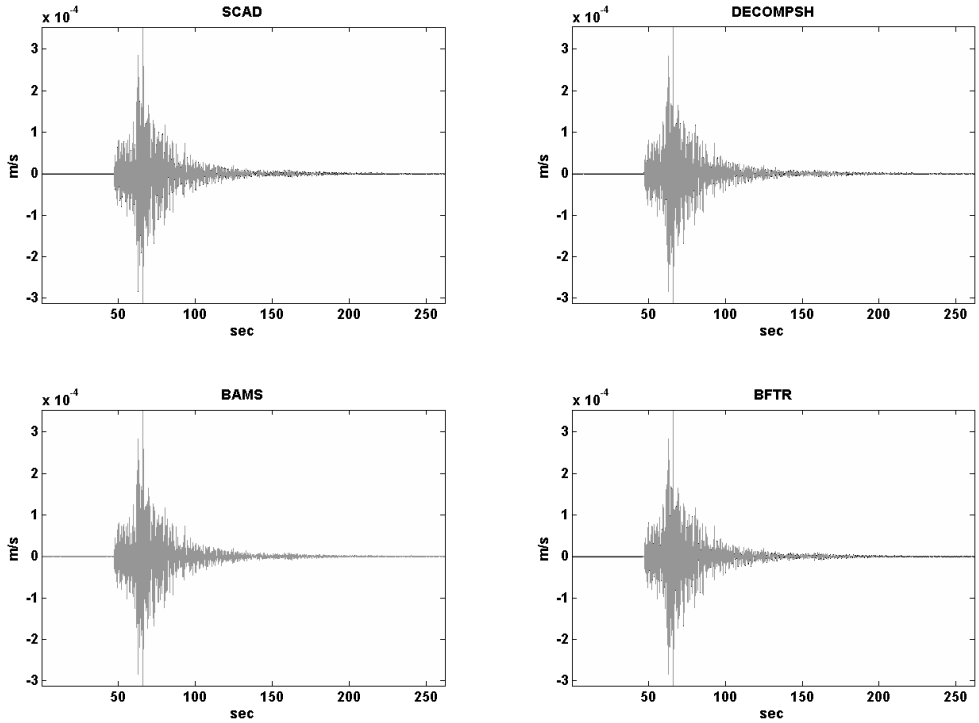


Figure 5.17 Original (black line) and denoised (grey line) recording from A category using SCAD (top left), DECOMPSH (top right), BAMS (bottom left) and BFTR (bottom right).

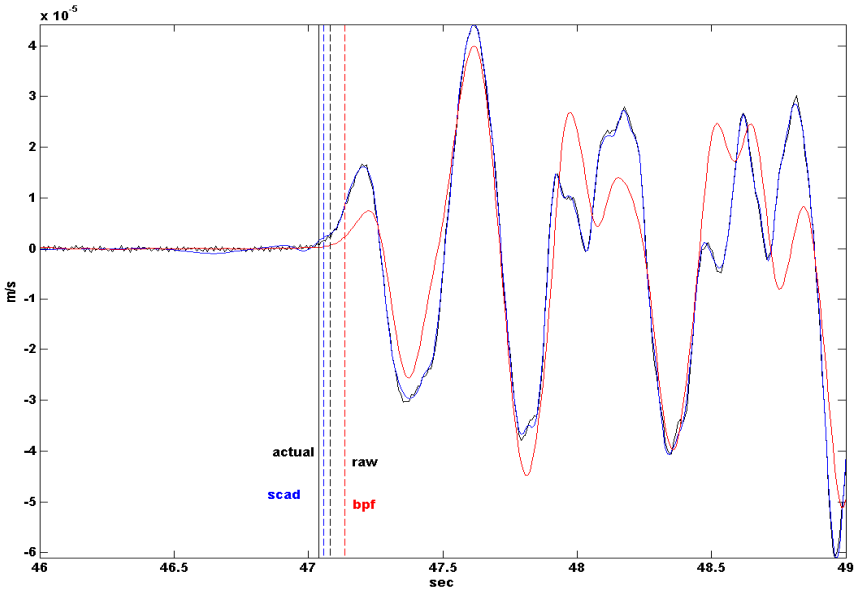


Figure 5.18 P-wave arrival of the original (black line), wavelet (SCAD) denoised (blue line) and bandpass filtered (red line) category A recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, unfiltered recording – black dashed line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

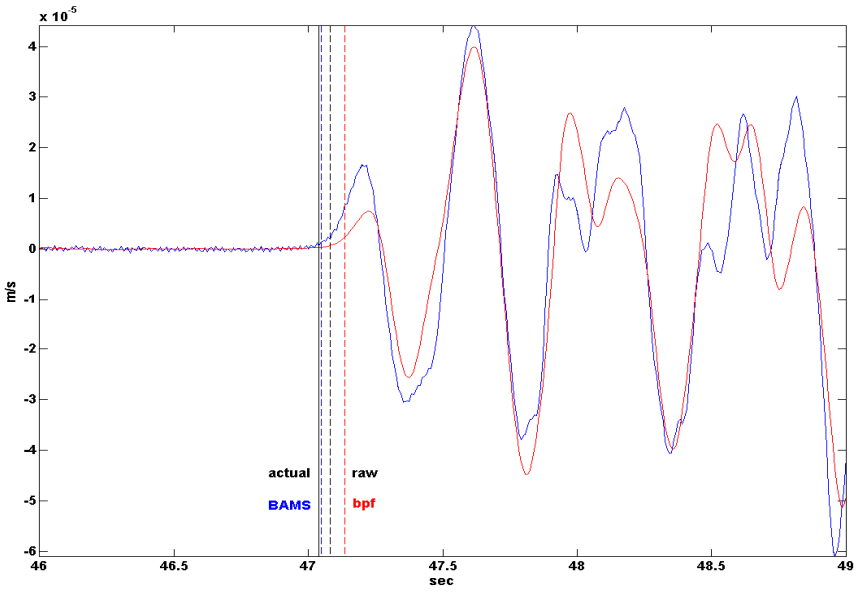


Figure 5.19 P-wave arrival of the original (black line), wavelet (BAMS) denoised (blue line) and bandpass filtered (red line) category A recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, unfiltered recording – black dashed line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

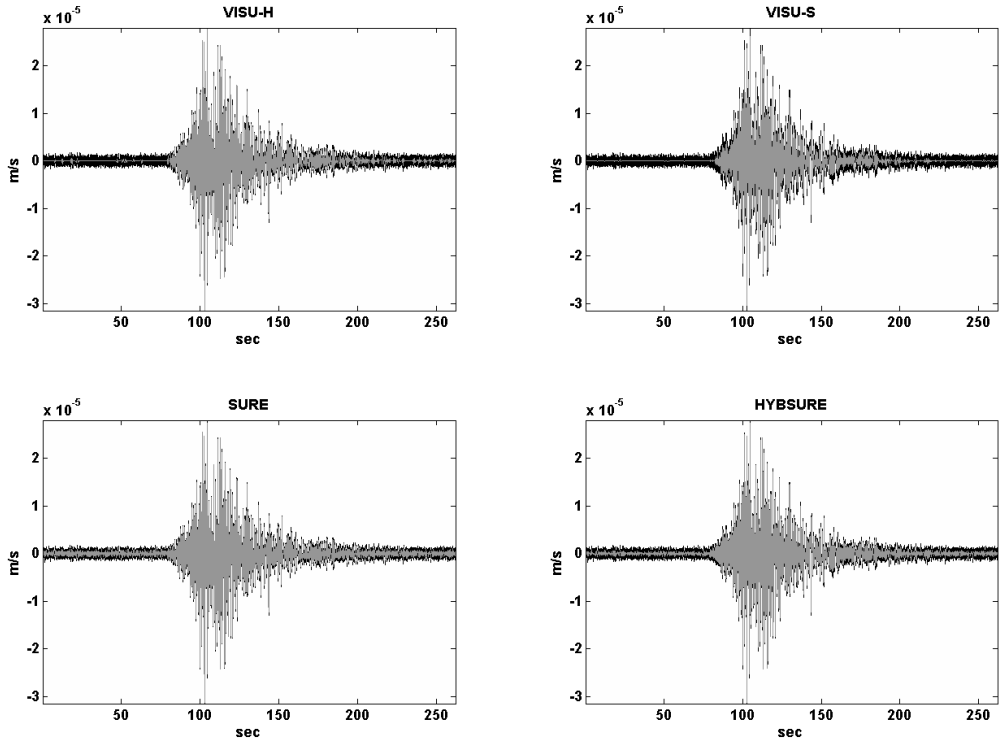


Figure 5.20 Original (black line) and denoised (grey line) recording from B category using VISU-H (top left), VISU-S (top right), SURE (bottom left) and HYBSURE (bottom right).

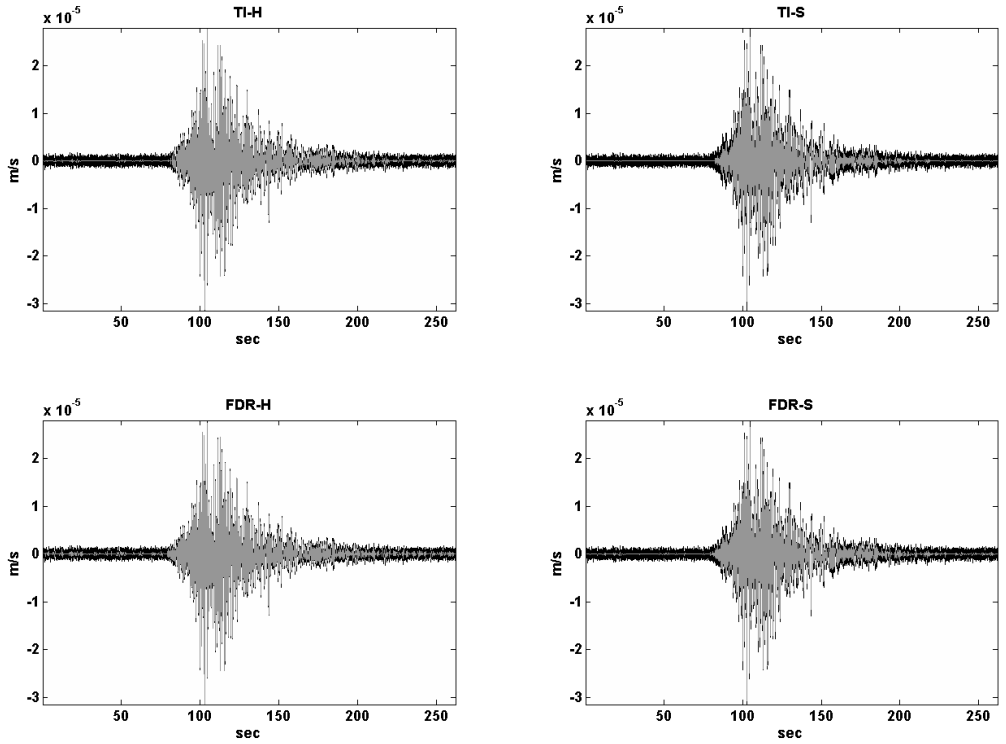


Figure 5.21 Original (black line) and denoised (grey line) recording from B category using TI-H (top left), TI-S (top right), FDR-H (bottom left) and FDR-S (bottom right).

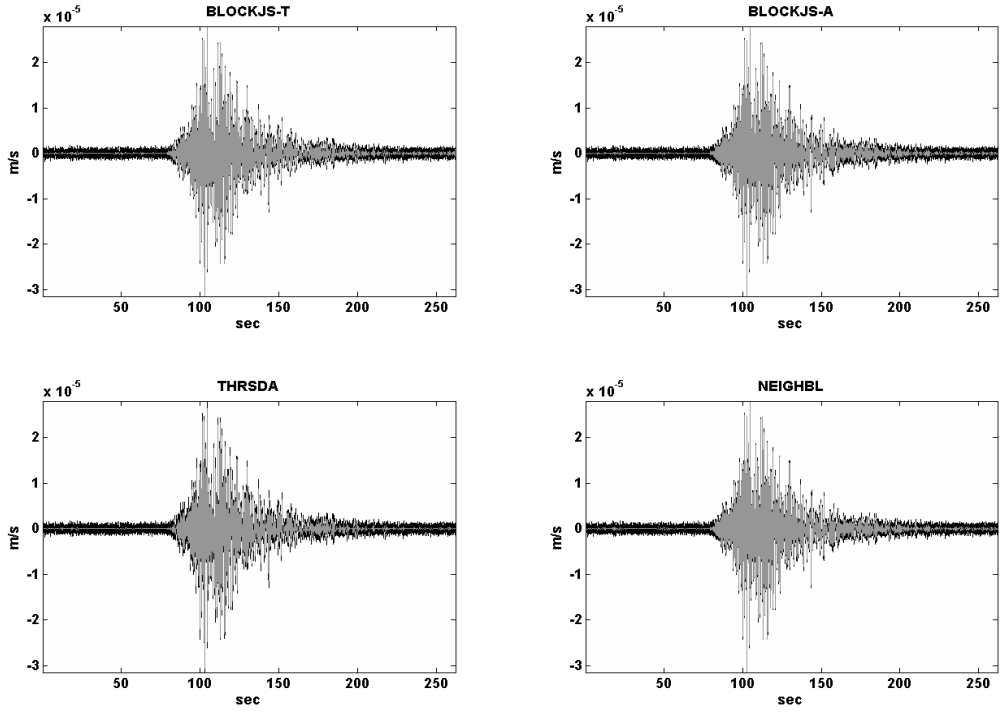


Figure 5.22 Original (black line) and denoised (grey line) recording from B category using TBLOCKJS-T (top left), BLOCKJS-A (top right), THRSDA (bottom left) and NEIGHBL (bottom right).

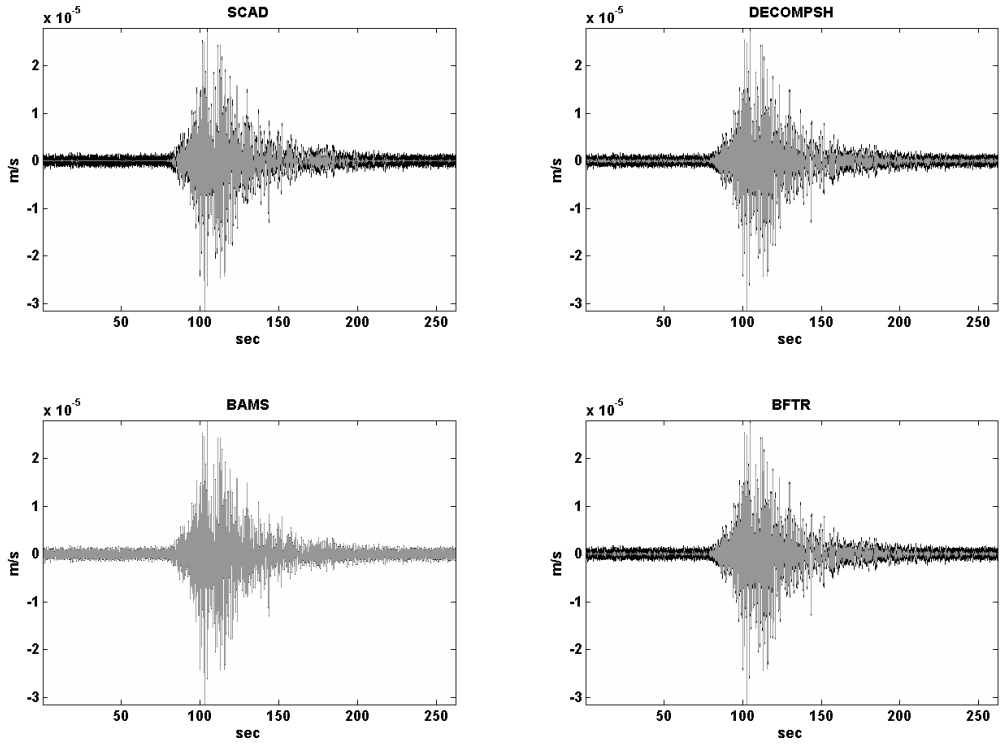


Figure 5.23 Original (black line) and denoised (grey line) recording from B category using SCAD (top left), DECOMPSH (top right), BAMS (bottom left) and BFTR (bottom right).

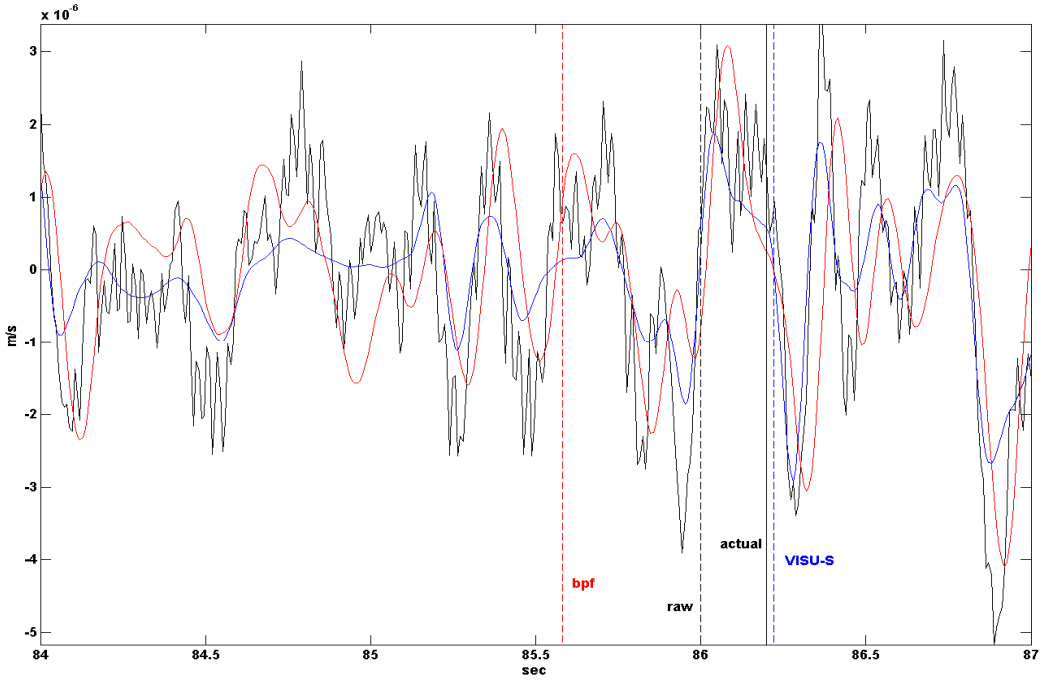


Figure 5.24 P-wave arrival of the original (black line), wavelet (VISU-S) denoised (blue line) and bandpass filtered (red line) category B recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, unfiltered recording – black dashed line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

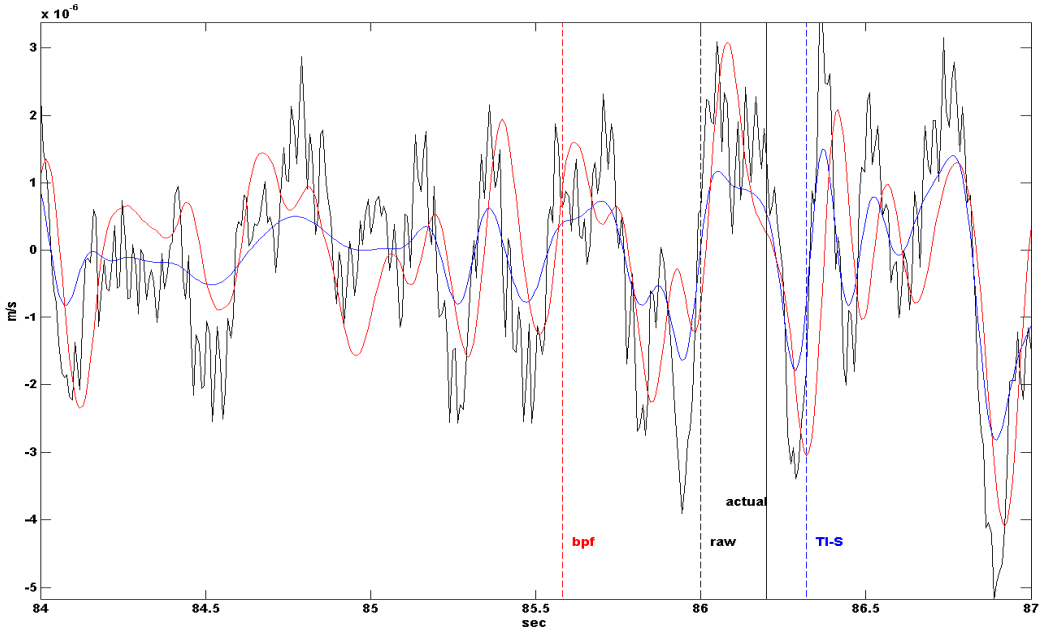


Figure 5.25 P-wave arrival of the original (black line), wavelet (TI-S) denoised (blue line) and bandpass filtered (red line) category B recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, unfiltered recording – black dashed line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

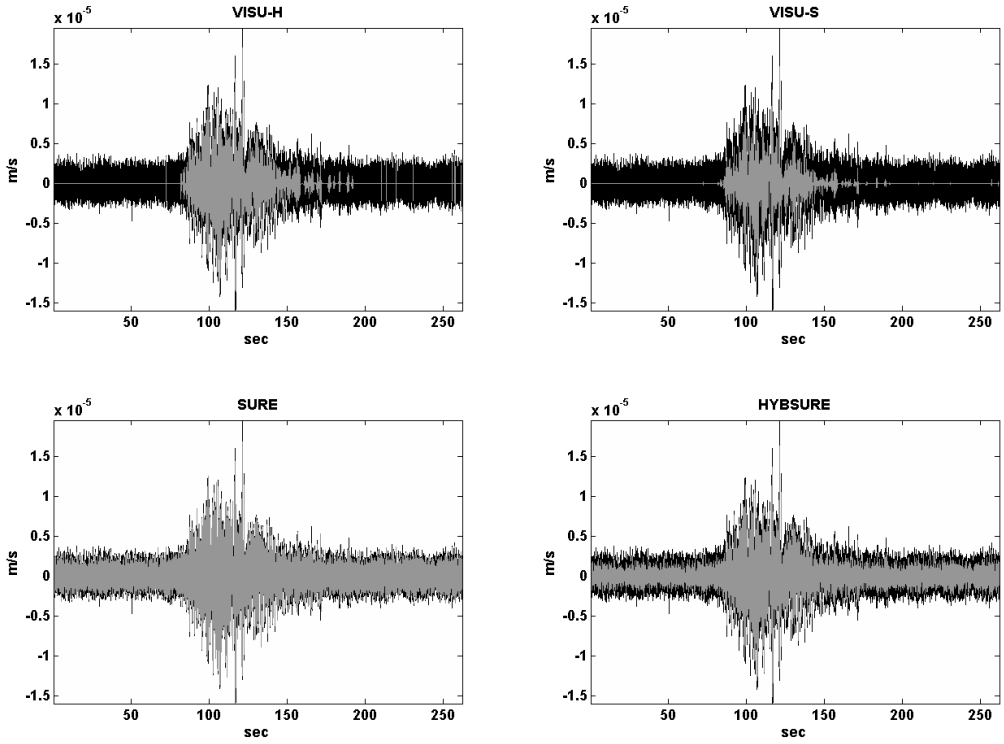


Figure 5.26 Original (black line) and denoised (grey line) recording from C category using VISU-H (top left), VISU-S (top right), SURE (bottom left) and HYBSURE (bottom right).

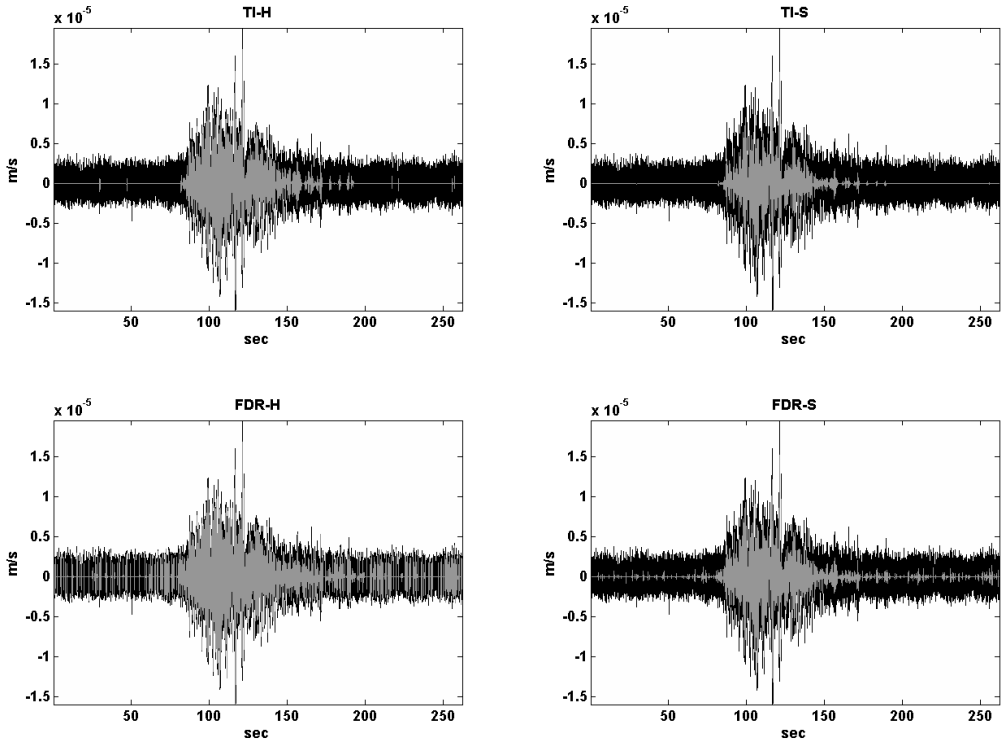


Figure 5.27 Original (black line) and denoised (grey line) recording from C category using TI-H (top left), TI-S (top right), FDR-H (bottom left) and FDR-S (bottom right).

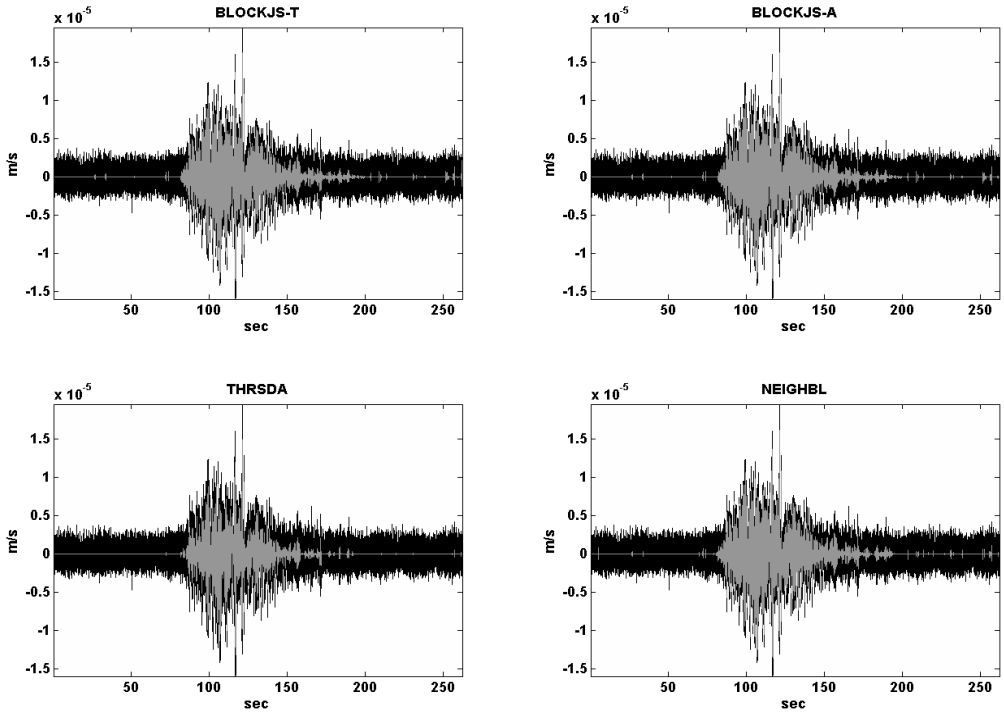


Figure 5.28 Original (black line) and denoised (grey line) recording from C category using TBLOCKJS-T (top left), BLOCKJS-A (top right), THRSDA (bottom left) and NEIGHBL (bottom right).

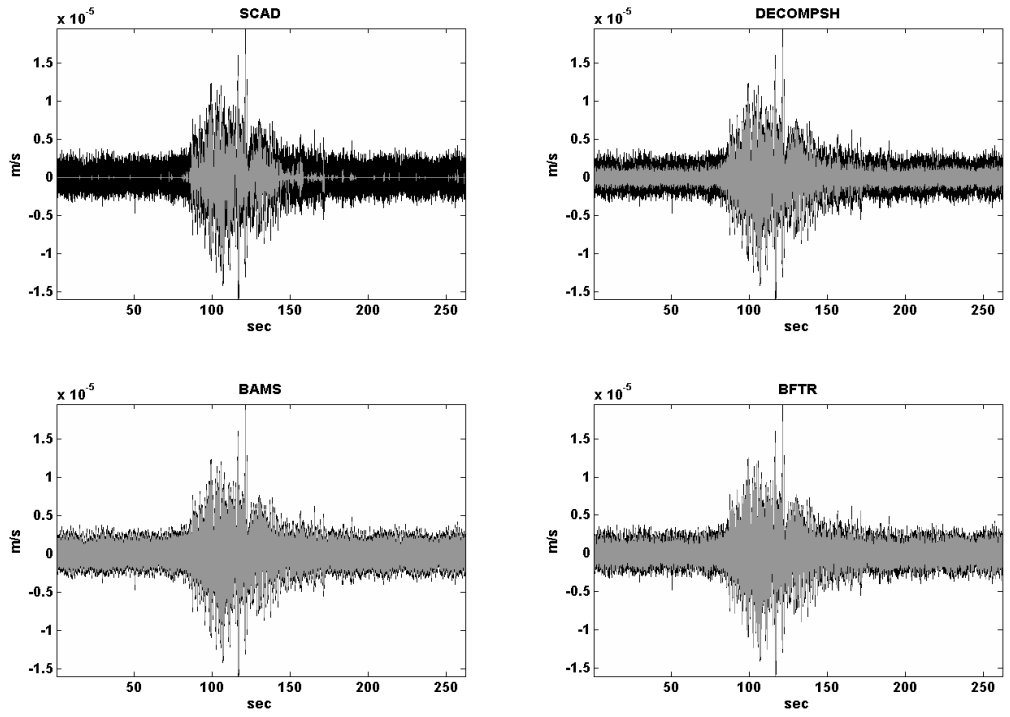


Figure 5.29 Original (black line) and denoised (grey line) recording from C category using SCAD (top left), DECOMPSH (top right), BAMS (bottom left) and BFTR (bottom right).

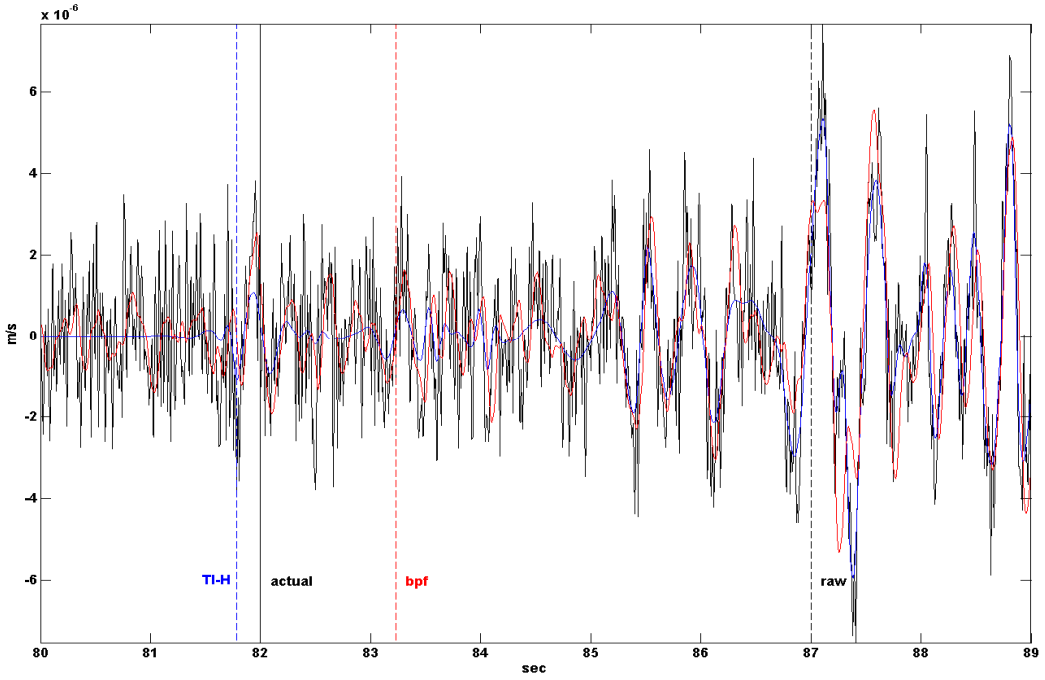


Figure 5.30 P-wave arrival of the original (black line), wavelet (TI-H) denoised (blue line) and bandpass filtered (red line) category C recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, unfiltered recording – black dashed line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

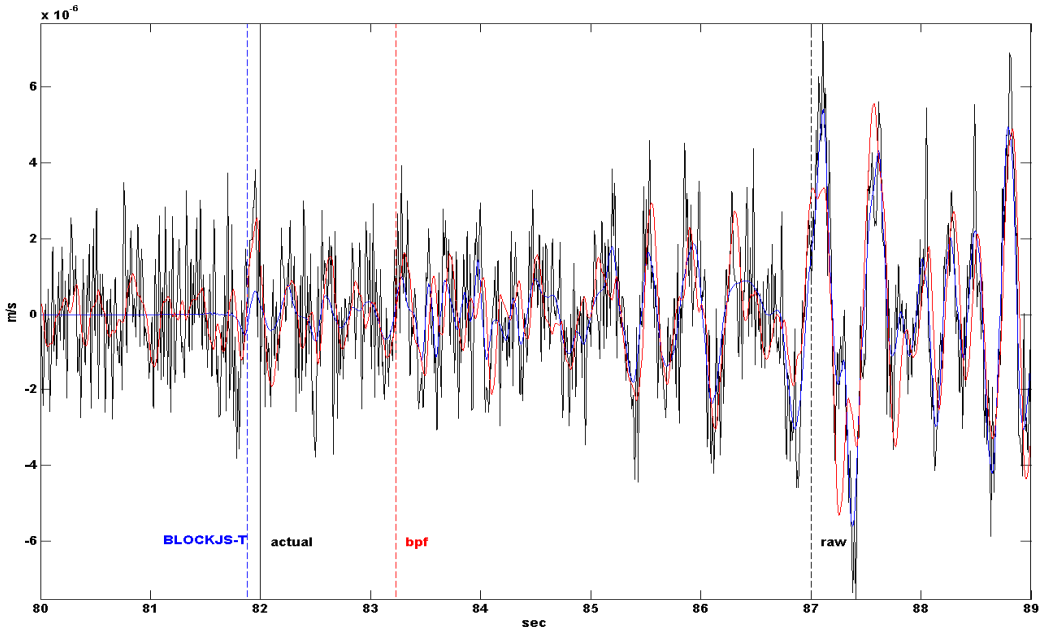


Figure 5.31 P-wave arrival of the original (black line), wavelet (BLOCKJST) denoised (blue line) and bandpass filtered (red line) category C recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, unfiltered recording – black dashed line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

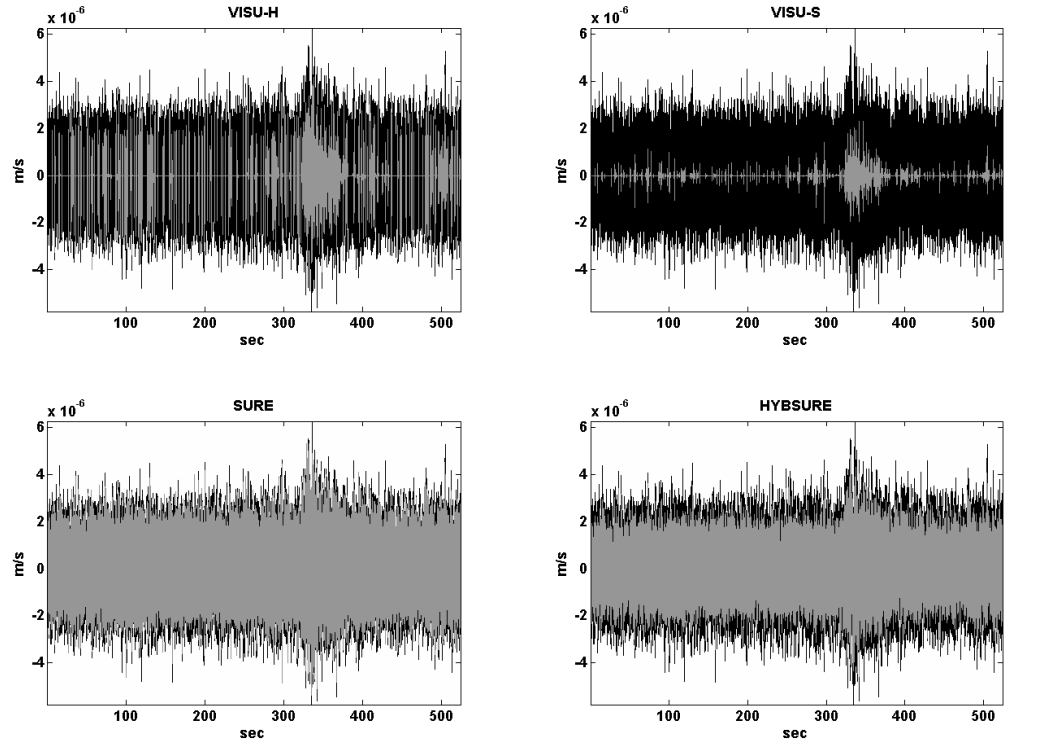


Figure 5.32 Original (black line) and denoised (grey line) recording from D category using VISU-H (top left), VISU-S (top right), SURE (bottom left) and HYBSURE (bottom right).

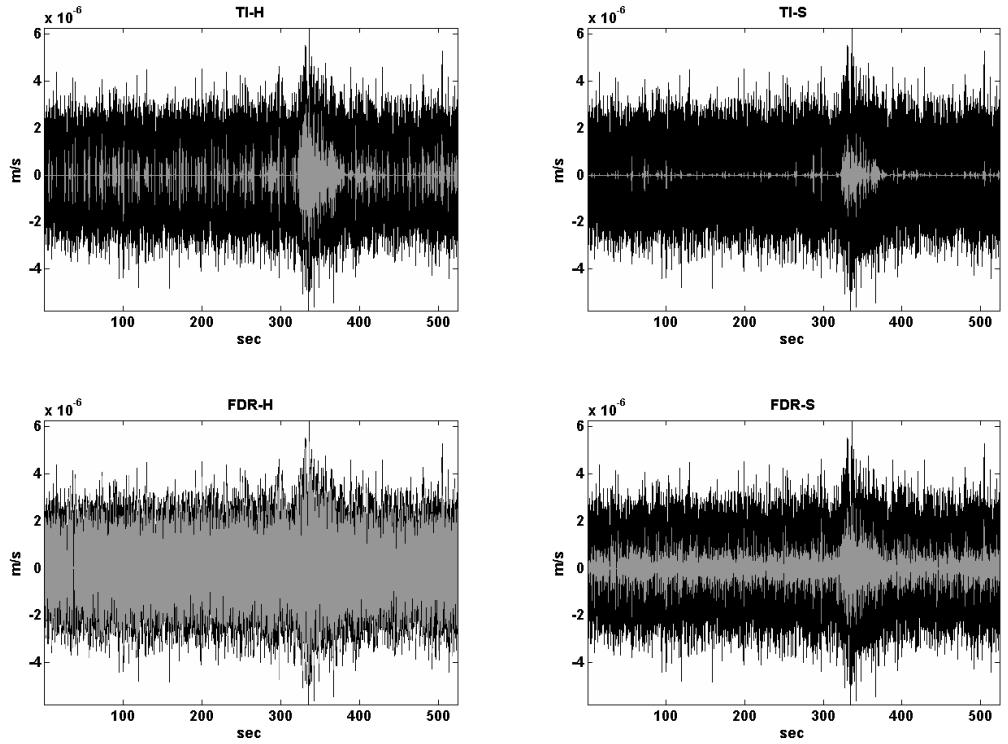


Figure 5.33 Original (black line) and denoised (grey line) recording from D category using TI-H (top left), TI-S (top right), FDR-H (bottom left) and FDR-S (bottom right).

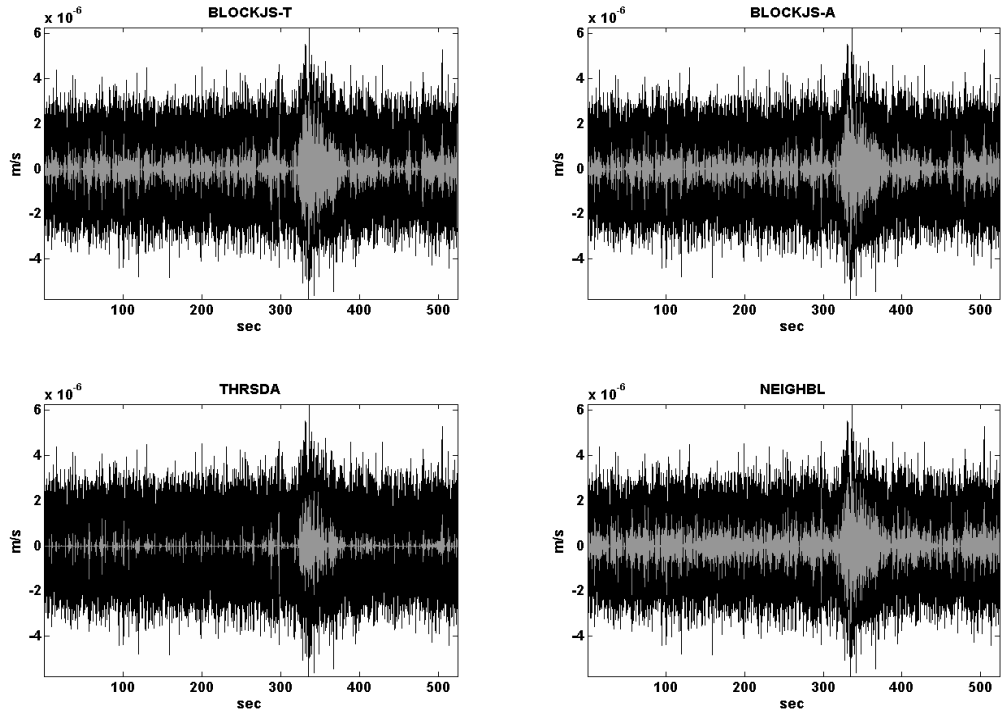


Figure 5.34 Original (black line) and denoised (grey line) recording from D category using TBLOCKJS-T (top left), BLOCKJS-A (top right), THRSDA (bottom left) and NEIGHBL (bottom right).

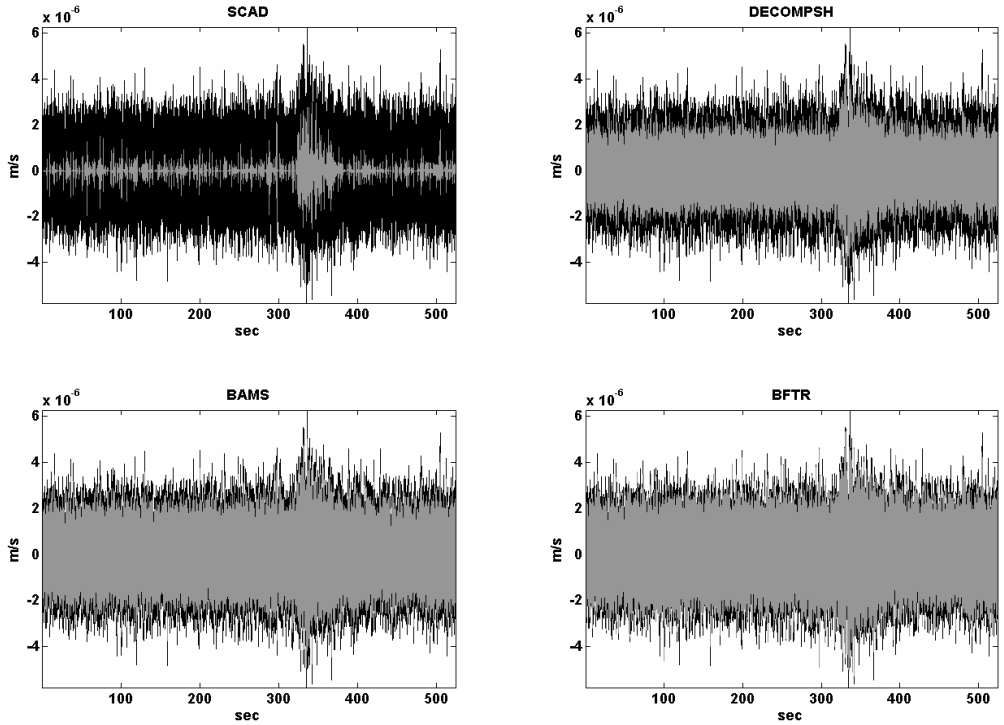


Figure 5.35 Original (black line) and denoised (grey line) recording from D category using SCAD (top left), DECOMPSH (top right), BAMS (bottom left) and BFTR (bottom right).

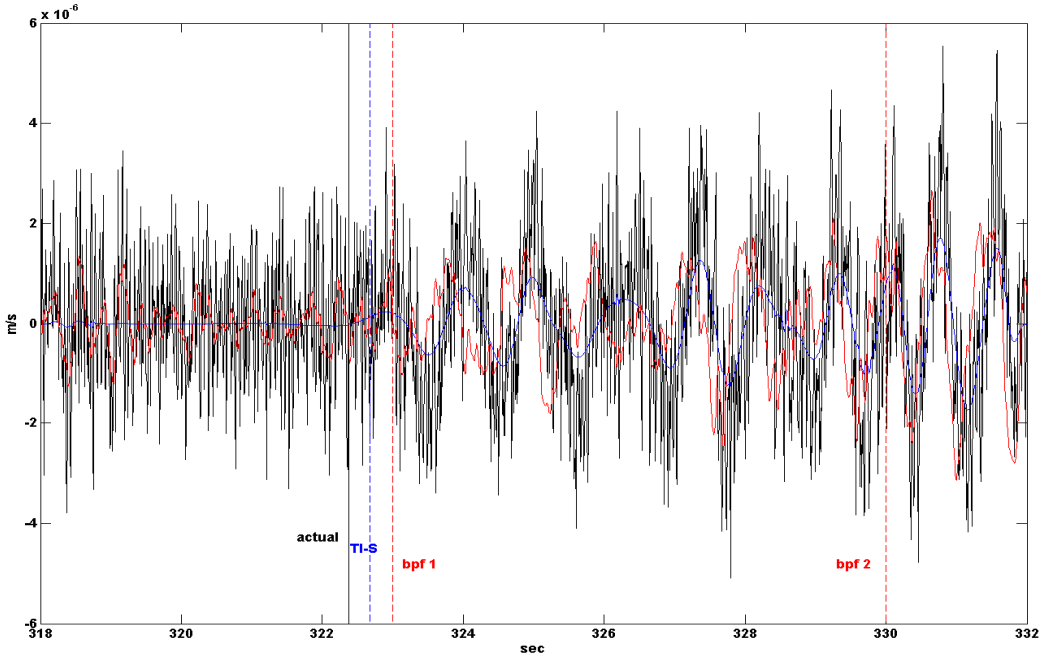


Figure 5.36 P-wave arrival of the original (black line), wavelet (TI-S) denoised (blue line) and bandpass filtered (red line) category D recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

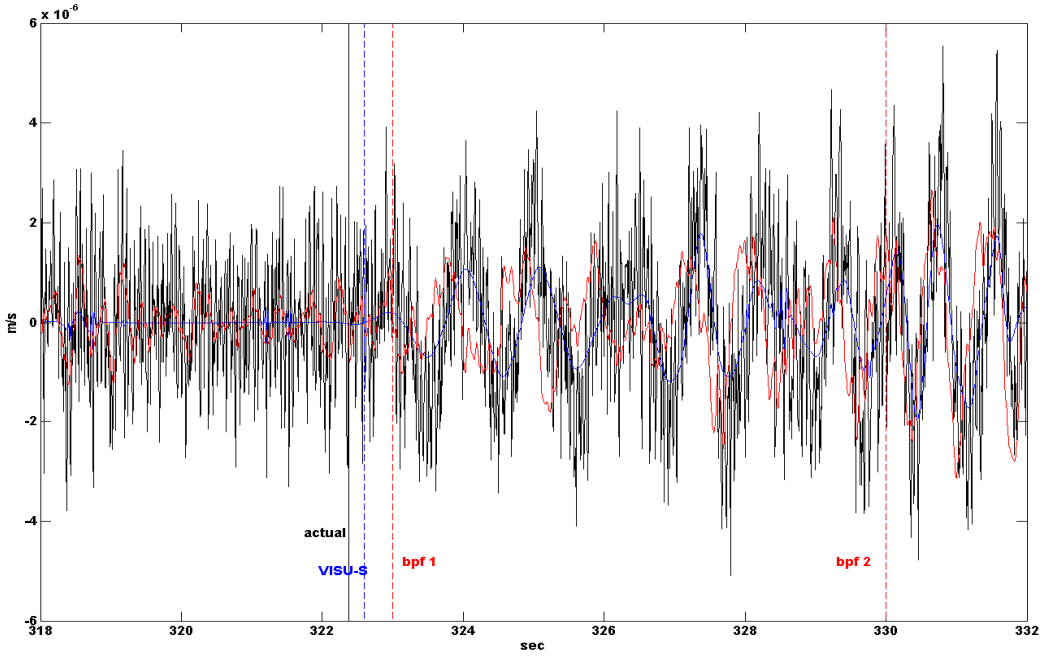


Figure 5.37 P-wave arrival of the original (black line), wavelet (VISU-S) denoised (blue line) and bandpass filtered (red line) category D recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

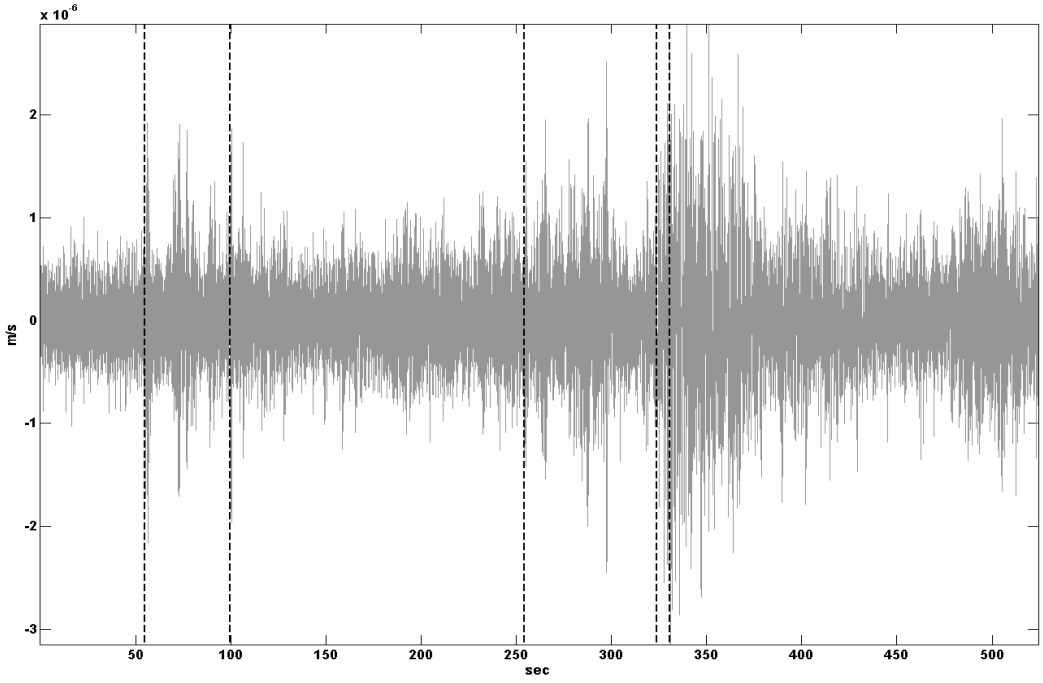


Figure 5.38 Bandpassed category D recording which triggers (dashed vertical lines) STA/LTA multiple times (dummy events) Performance evaluation of denoising using LA(8) basis for GN3 (top left), CL6

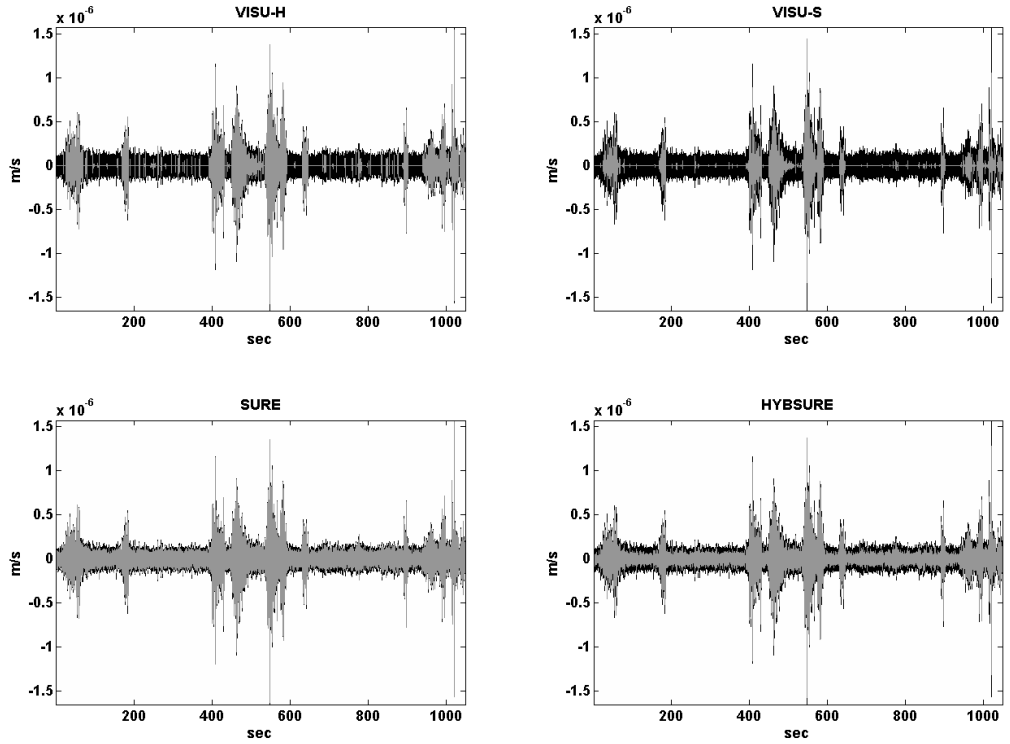


Figure 5.39 Original (black line) and denoised (grey line) recording from E category using VISU-H (top left), VISU-S (top right), SURE (bottom left) and HYBSURE (bottom right).

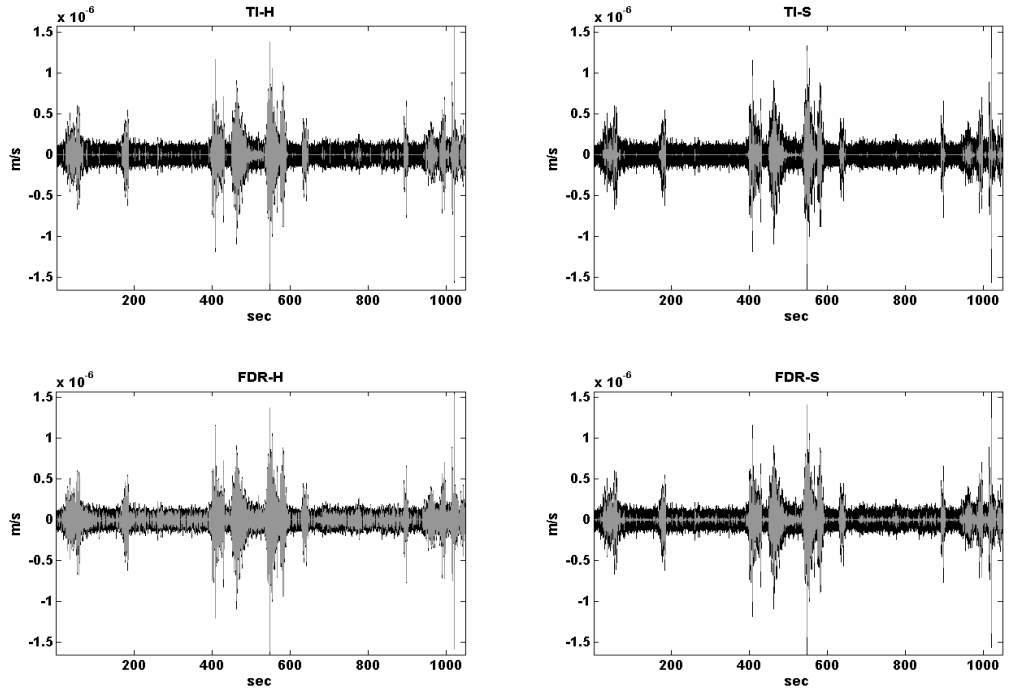


Figure 5.40 Original (black line) and denoised (grey line) recording from E category using TI-H (top left), TI-S (top right), FDR-H (bottom left) and FDR-S (bottom right).

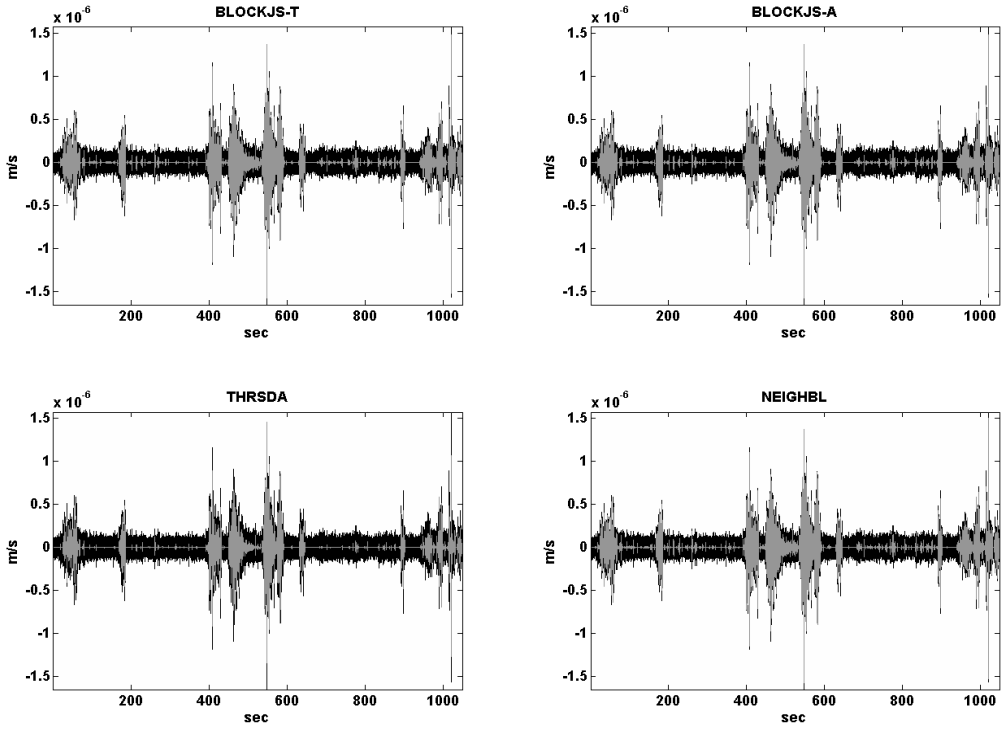


Figure 5.41 Original (black line) and denoised (grey line) recording from E category using TBLOCKJS-T (top left), BLOCKJS-A (top right), THRSDA (bottom left) and NEIGHBL (bottom right).

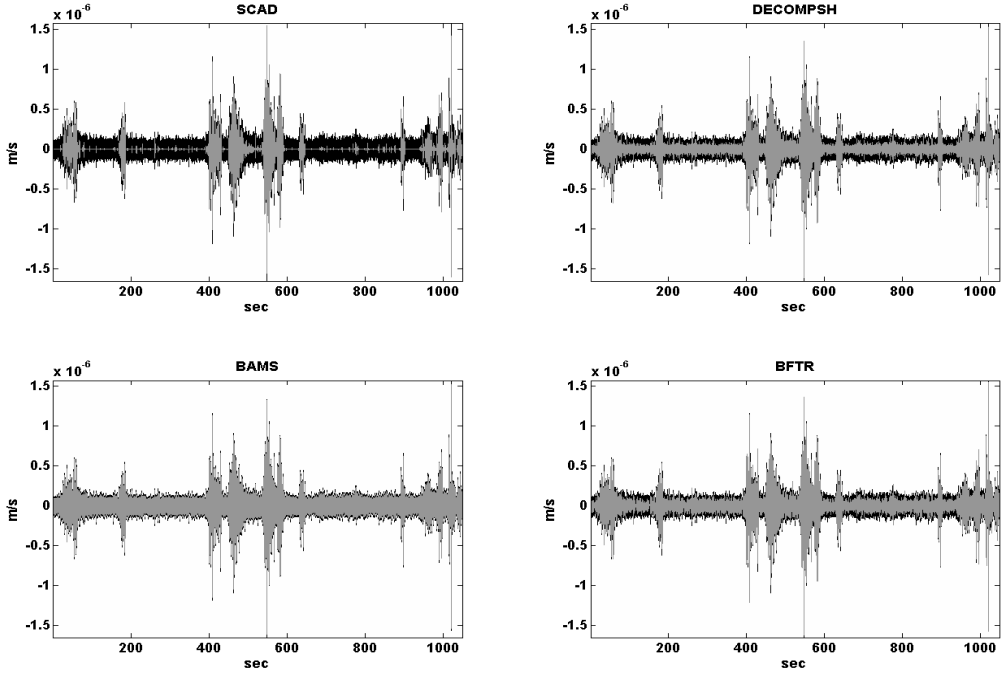


Figure 5.42 Original (black line) and denoised (grey line) recording from E category using SCAD (top left), DECOMPSH (top right), BAMS (bottom left) and BFTR (bottom right).

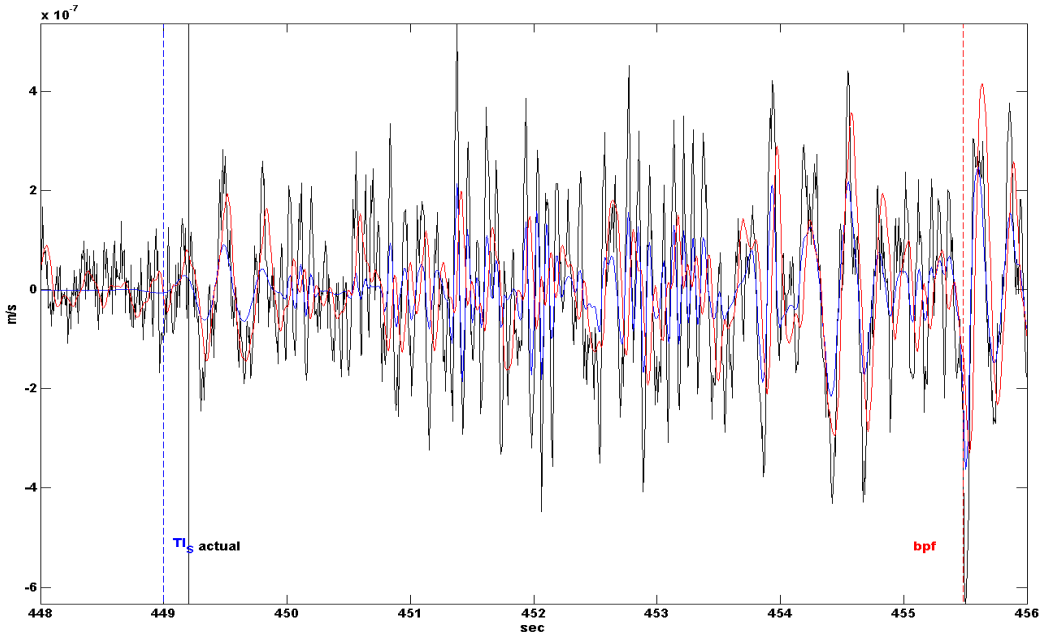


Figure 5.43 P-wave arrival of the original (black line), wavelet (TI-S) denoised (blue line) and bandpass filtered (red line) category E recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

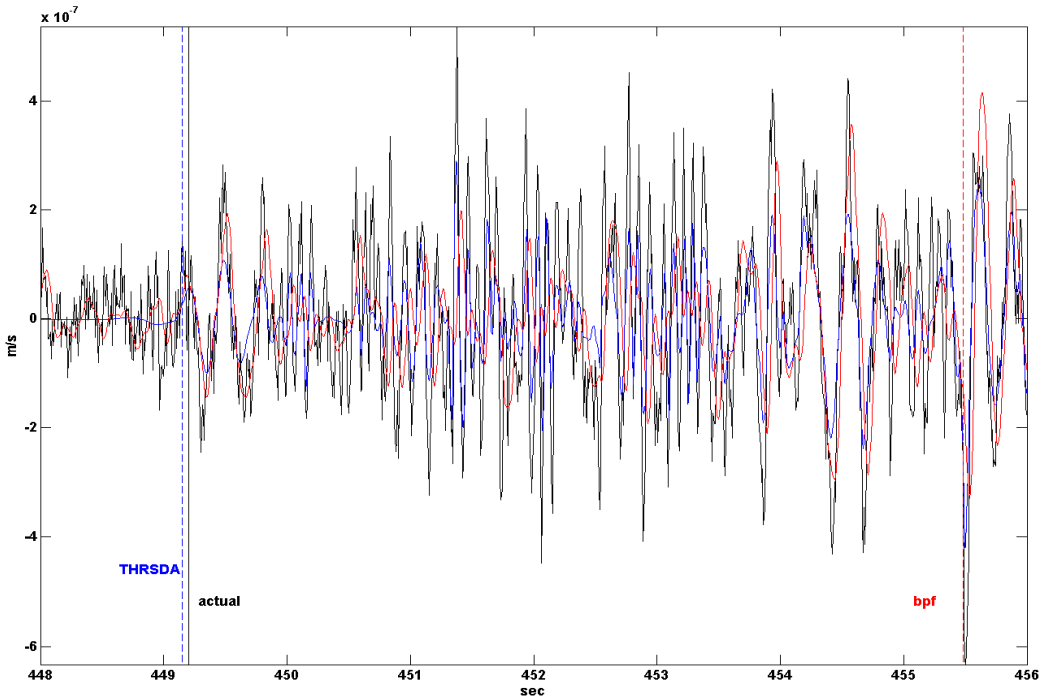


Figure 5.44 P-wave arrival of the original (black line), wavelet (THRSDA) denoised (blue line) and bandpass filtered (red line) category D recording . Trigger times using STA/LTA presented as vertical lines (event's reported actual time – black solid line, wavelet denoised – blue dashed line, bandpass filtered – red dashed line).

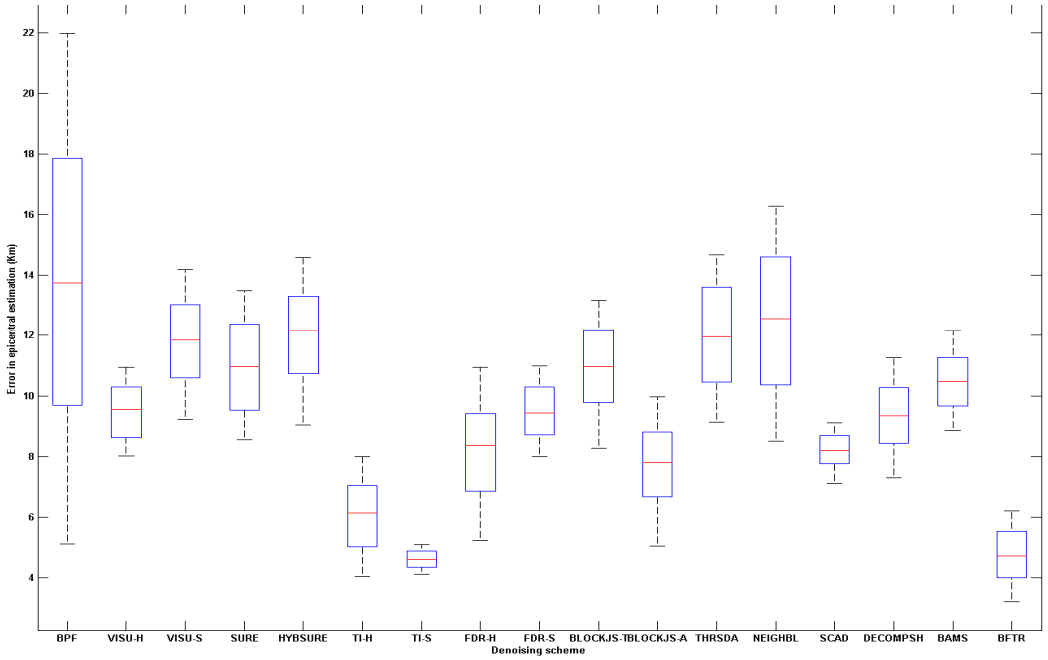


Figure 5.45 Distance errors in epicentral estimation for each denoising scheme using 250 events from category A

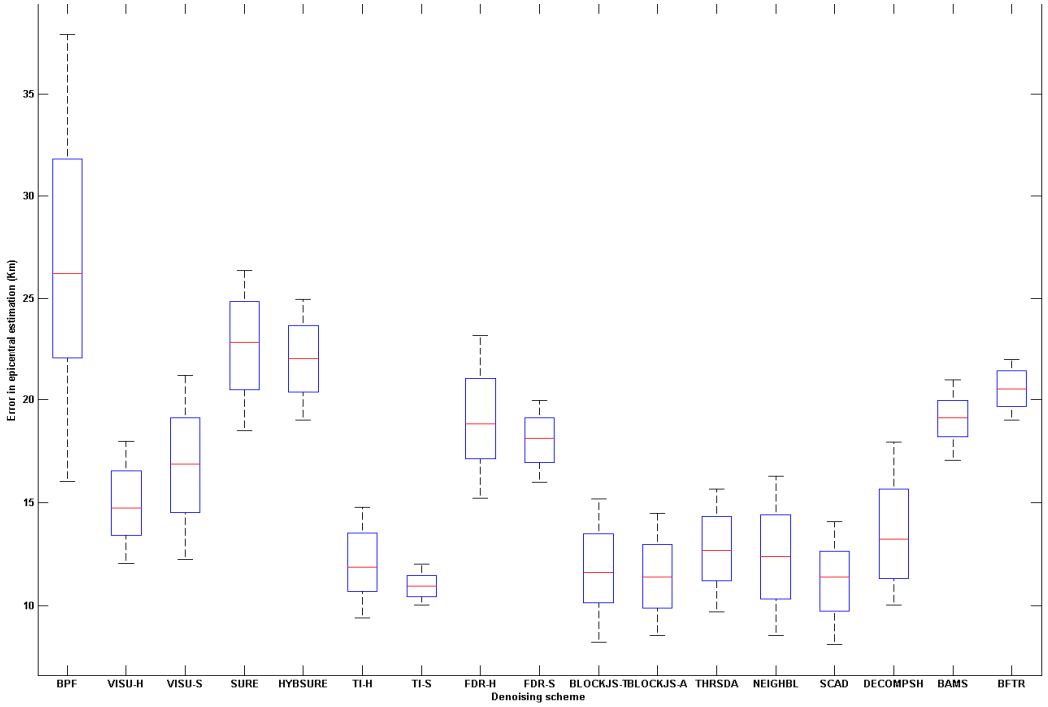


Figure 5.46 Distance errors in epicentral estimation for each denoising scheme using 250 events from category B

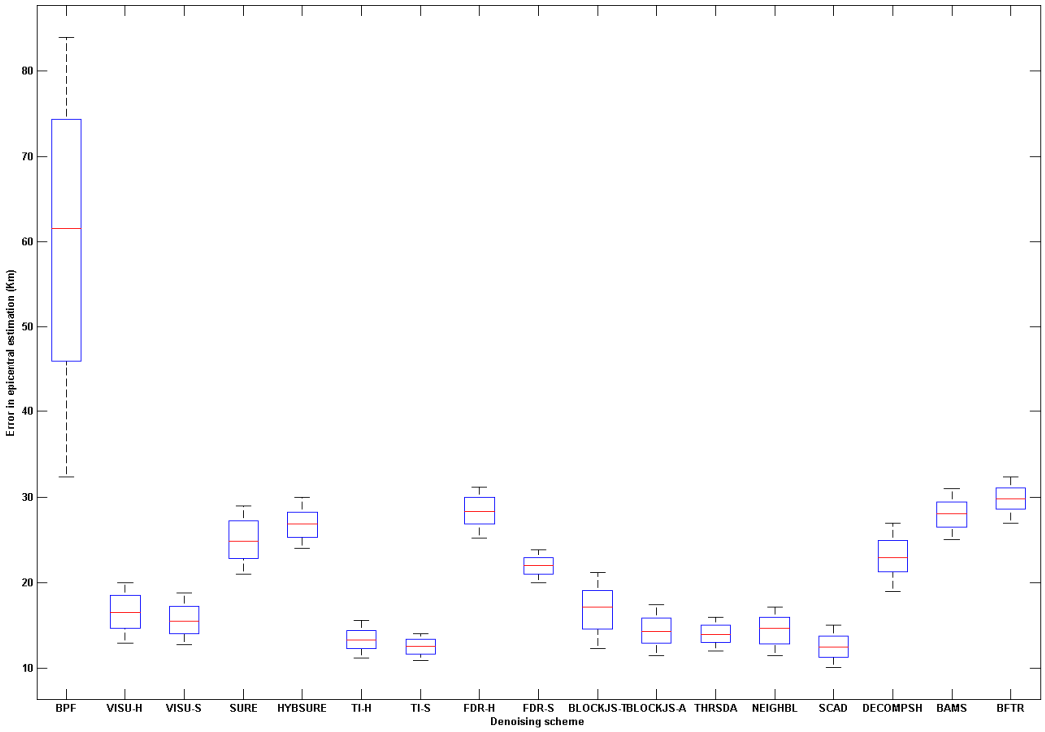


Figure 5.47 Distance errors in epicentral estimation for each denoising scheme using 250 events from category C

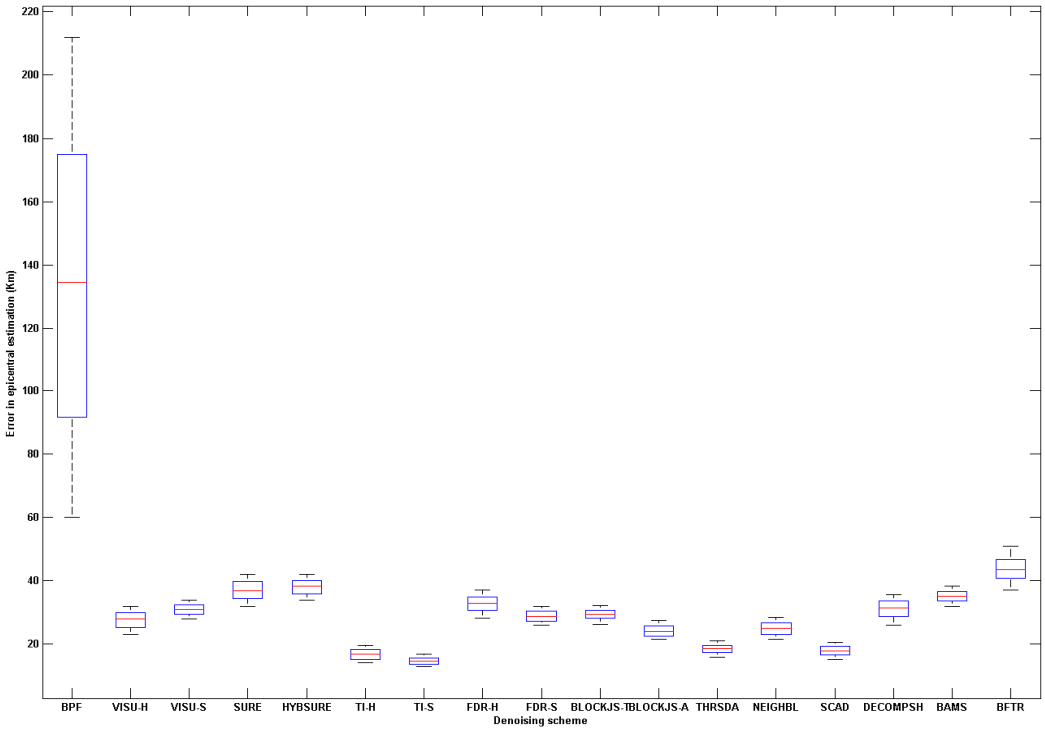


Figure 5.48 Distance errors in epicentral estimation for each denoising scheme using 250 events from category D

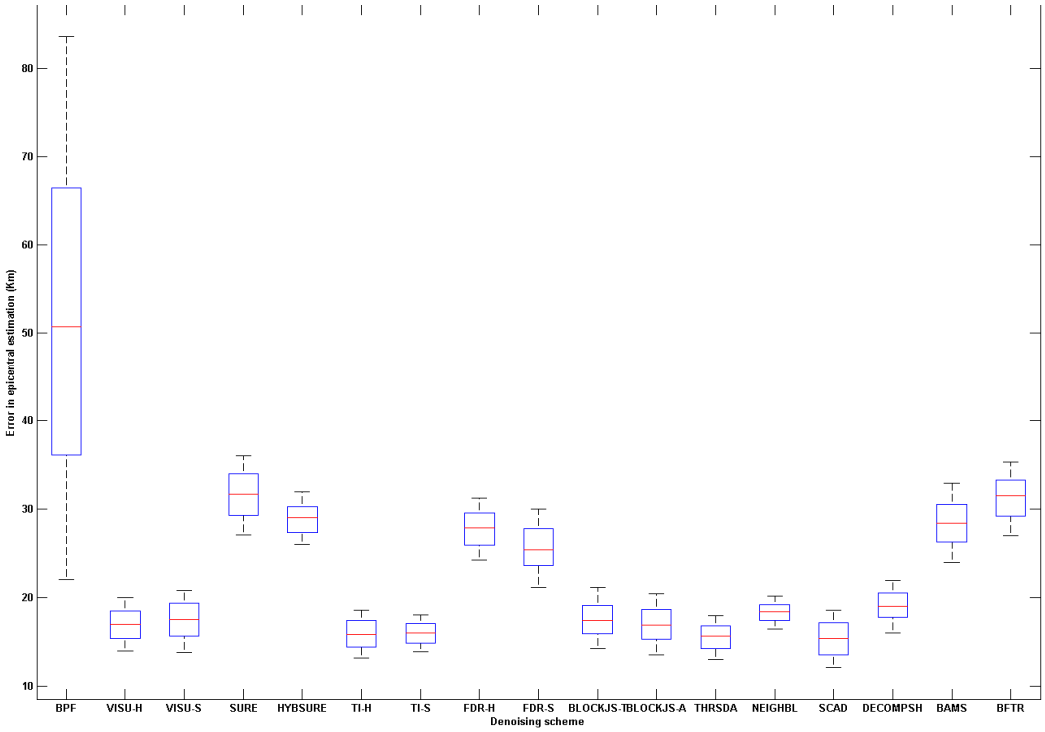


Figure 5.49 Distance errors in epicentral estimation for each denoising scheme using 250 events from category E

5.10 Discussion

Two basic conclusions derived from the examination of the results:

- Every wavelet denoising method provides similar or better results from the commonly used bandpass filtering
- There is no specific WDM that can be used a-priori.

Regarding the superiority of WDM over bandpass filtering is not surprisingly since the wavelet methods provide better time localization. This superiority is not expressed only as an improvement to SNR but also because the wavelet denoised seismograms did not suffer from delays at P-wave pulses. The latter is very important for the validity of results provided by a seismological network because an error in P wave's arrival time could produce wrong estimations for earthquake's epicenter. For seismograms that belong to category A this delay produced errors with an average value of 14 km. This value probably seems negligible for the estimation of a big earthquake but is unacceptable for modern, high resolution, seismic tomography methods which use the big earthquakes as triggering sources. As expected, as the

quality of the seismograms decreased the delay increases. One important point that must clarify is that for each seismogram the values in STA/LTA detector that used are the optimum ones in order to provide the better result regarding P wave estimation. This is very important since because for each seismogram the bandpassed filtered seismograms need some investigation in order to produce the better result. On the other hand, for the wavelet denoising methods the commonly used values for STA/LTA produce the above results. This is also very important since it relaxes the analyst from finding the better combination of STA/LTA values in order to detect every event but at the same time to reject other events (explosions, artificial noises). The poor performance of bandpass filtering is verified at categories C, D and E where the produced results are unacceptable for a seismological network.

For each of the five categories several WDM produce very good results. In general, even the worst WDM produce better results from bandpass filtering. This is a useful conclusion since it can lead to the confirmation of WDM as a defacto denoising method to modern seismological networks. Except the computational complexity (which is negligible when using today's high performance PCs) it seems that there is no other reason that can stop using wavelets for denoising seismograms.

As already mentioned, there is no unique WDM that produce the better results. From our results, it seems that TI-H, Bayesians and NEIGHBL produce best overall results but we can qualify them as the only appropriate. A suggestion that we can make is the use of the hybrid system that will select the most appropriate WDM according to seismogram's SNR. This is not new idea because a system with this type selection scheme is proposed by Pazos et al (2003). Our proposal differs because we propose the use of different WDM according to SNR where Pazos et al approach selects different thresholds according to SNR. In contrary, they didn't deal with seismograms with SNR<3. These seismograms (category D) are included to our analysis not to primarily show that WDM can handle them also but to show the ability of wavelets to detect non regional earthquakes or microearthquakes that will remain undetected using conventional bandpass filtering.

Category E deals with the disturbances that a seismometer could acquire. These are not unusual, even if they exist over a short period of time (e.g. car passing, construction works near seismological station). The ability to have the same information quality from this type of stations remains high since WDM are not so

sensitive to these disturbances and they can handle the real seismic signal accordingly.

WDM cannot by themselves provide a complete solution to earthquake detection but they can be used to denoise with high performance the seismic signals which then must be passed to appropriate earthquake attribute estimators (hypocentral estimation, magnitude calculation, phase detectors e.t.c) for further processing.

5.11 Summary

Wavelet denoising schemes have proved to be well adapted to several types of signals. For non-stationary signals, such as seismograms, the use of linear and non-linear wavelet denoising methods seems promising. The contribution of this chapter is a comparison study for wavelet denoising methods suitable for seismic signals, which proved from previous studies their superiority against appropriate conventional filtering techniques. The importance of wavelet denoising methods relies on two facts: they recovered the seismic signals having fewer artifacts than conventional filters (for high SNR seismograms) and at the same time they can provide satisfactory representations (for detecting the earthquake's primary arrival) for low SNR seismograms or microearthquakes. The latter is very important for a possible development of an automatic procedure for the regular daily detection of small or non-regional earthquakes especially when the number of the stations is quite big. Initially, their performance is measured over a database of synthetic seismic signals in order to evaluate the better wavelet approach. As expected, there is no wavelet denoising method that outperforms the others but best overall results gained from TI-S, Bayesian and NEIGHBL.

6. The Potential for a Wavelet Earthquake Early Warning System in South Aegean

6.1 Introduction

Since the complexity involved in earthquake process does not allow the use of a practical earthquake prediction method, the only approach that can reduce the seismic hazard mitigation is by using systems that can generate alerts before the destructive S waves arrive to the area of interest. Such a system is called Earthquake Early Warning System (EEWS) and its objective is to provide a few to tens of seconds of warning time for the oncoming ground motions allowing for mitigation actions in short term. EEWS that estimate the severity of ground shaking few seconds after its onset time are in operation in Japan (Nakamura 1984, 1988, 1989), Taiwan (Teng et al., 1997, Wu et al., 1998, 1999, Wu and Teng, 2002) and Mexico (Espinosa et al. 1995). The information produced by the above systems can be used by authorities in order to minimize the property damage and the loss in urban areas as well as to provide real time loss estimation for emergency response and recovery (Wu et al., 2002).

The measures needed for EEWS are the earthquake's magnitude and its epicentral location. Their rapid and accurate determination is the main goal for the successful operation of an EEWS. An earthquake generates two main types of waves: the P ans S waves. The direct P waves have smaller amplitude and travel with faster velocity than S waves which have lower velocity but higher amplitude and thus hold the destructive attributes. The traditional way of estimating a local magnitude (M_L) requires the whole seismogram. This is the most accurate method but from the seismic mitigation point is useless since the magnitude can only be estimated after the arrival of the destructive part of the earthquake. During last years, research has focused on whether an accurate earthquake magnitude can be rapidly estimated using only tht first few second of the P wave recordings acquired close to the epicenter, stations. Several studies on EEWS (Nakamura, 1988, Allen and Kanamori, 2003, Lockman and Allen, 2005, Olson and Allen, 2005, Kanamori, 2005, Wu and Kanamori, 2005a,b) used the dominant period of the first few seconds of the P wave as the observable that correlates with M_L. They found that the predominant period

tends to increase with magnitude even where the rupture is not complete (large earthquakes) within the first 3-4 secs. This hopeful estimator is not without shortcomings. As Wolfe (2006) and Simmons et al. (2006) pointed out the predominant period estimator has significant scattering mainly caused by its recursive calculation based on a spectral domain relation.

The purpose of this chapter is to demonstrate the use of wavelets as the common processing tool for earthquake's rapid magnitude determination and epicentral estimation in order to move from time-frequency domain estimators to time-scale domain estimators. The goal is to use the same set of wavelet coefficients that characterize the seismogram (and especially its P wave portion) in order to use one technique (WT) for double use (magnitude and location estimation). Wavelet Magnitude Estimation (WME) is used to derive a scaling relation between earthquake's magnitude and wavelet coefficients. In addition, the epicentral estimation is achieved by a new method (Wavelet Epicentral Estimation – WEpE) which is based on the combination of wavelet azimuth estimation (Galliana, 2007) and two stations method (Rydelek and Pujol, 2004). WME and WHE running in parallel, used as the core of a regional WEES in South Aegean.

6.2 Fundamentals of EEWS

The physical attributes that involved in an earthquake early warning system depend on: (1) strong ground shaking from a damaging earthquake is caused by shear (S) waves and the following surface waves, (2) typical crustal P-wave velocity is about 6-8 km/sec, whereas S- and surface waves travel at about half the speed of the P-waves, and (3) the seismic wave velocities are much slower than electromagnetic signals transmitted by telegraph, telephone, or radio at about 300,000 km/sec.

Cooper (1868) was the first who propose an earthquake early warning system more than one hundred years ago. He proposed setting up seismic detectors 120km away California and when an earthquake triggered them, an electric signal would be sent by telegraph to San Francisco. This signal would then ring a big bell in the City Hall to warn citizens that an earthquake had occurred. Cooper's scheme was never implemented. Heaton (1985) proposed a seismic computerized alert network for southern California. Nakamura (1988) implemented the single-station urgent earthquake detection and alarm system (UrEDAS) for the high speed train system in

Japan. Espinosa-Aranda et al. (1995) established the first earthquake early warning system for large events in order to produce alerts to the public in the Mexico City. Subsequently, EEWS have been implemented, experimented, and commented upon by several different groups

An EEWS warns an area of interest (e.g. a urban area) of forthcoming strong shaking, normally with a few seconds to a few tens of seconds of early warning time, i.e., before the arrival of the destructive S- and surface waves of a damaging earthquake. These few seconds of advanced warning time will be useful for the pre-programmed procedures for various critical facilities**,** such as stopping underground wagons and high-speed trains to avoid potential derailment, the controlled shutoff of gas pipelines to minimize fire hazards, the safe-guarding of computer facilities to avoid the loss of vital data and the preparation of auxiliary power sources in case of power failure in critical facilities (e.g. hospitals)

An EEWS is basically a network of seismic recording systems, consisting of seismological stations (local or central) and a transmission infrastructure spreading the alarm to end users (Heaton, 1985) to initiate personal or automatic security measures. The success of an EEWS is directly related to the reduction in total losses produced in a region of interest for very critical facilities.

EEWS may be categorized by the configuration of their seismic network as regional or site-specific (Kanamori, 2005). Regional EEWS consist of wide seismic networks covering a portion of the area which is likely to be the source of a catastrophic earthquake. The recorded data are sending to a processing centre in order to calculate earthquake parameters such as magnitude, location, faulting mechanism and peak ground velocities. Then, these values used to estimate the level of shaking in the affected area. The whole process requires a significant amount of time and for this reason in a possibly large portion of the region, there is no way the alarm be produced before the arrival of destructive waves. This area called the *blind zone* (Kanamori, 2005). These systems are guided to applications such as shake maps (Wald et al., 1999), which are territorial distributions of ground shaking available immediately after the event to the local authorities for emergency management, in order to direct their efforts to areas that expected to have the highest losses. These systems called Rapid Response Systems (Wieland, 2001).

When the system can spread the alarm during the event, before the ground motion hits some sites or areas of interest then it can classified as real-time EEWS. Even then, the system does not have enough time to spread the alert, The only known case where the EEWS has a more relaxed time margin for spreading the alert is the system that installed in Mexico city. This is achieved because seismic active zone is known and is sufficiently far away. The used EEWS is focused in this area (the subduction zone of the Pacific coast) by using 12 digital strong motion field stations located along a 300 km stretch of the coast. Each field station includes a computer that continually processes seismic activity within a 100 km radial area around each station. Two seconds are required for the information on an event to reach Mexico City.

While regional systems directly improve the resiliency of communities to earthquakes, site-specific EEWS are designed using as primary objective the increase of safety margin for critical systems such as nuclear power plants, lifelines or transportation infrastructures. The seismological networks that used in these cases are smaller than those of the regional type because they cover only the surroundings of the system like protection armour for the seismic waves. The topology of the network depends on the time needed to activate the safety procedures before the S wave's arrival. In these Seismic Alert System*s* the alarm is generally produced when the first station triggered by an S wave. The difference from the aforementioned systems is that the source parameters of the event are not essential. This is because the calculation of source parameters is time-consuming. Also there is no significant uncertainty about the propagation of seismic waves since the path between the network and the site is limited. The alarm thresholds of these systems are configured in away that can produce alarm for every possible event even if it proved false. This is done because false alarms are preferable than lost alarms; the cost for issuing an alarm for an inexistent event is lower from the cost of a lost alarm. Among site-specific systems a paradigmatic example is that of the Ignalina nuclear power plant in Lithuania (Wieland et al. 2000) which is designed to produce an alert 4 to 8 seconds before the destructive ground motion reaches the reactor. By knowing that the time needed to secure the reactor is 2 secs, this EEWS is an example of successful design.

6.3 Wavelet Magnitude Estimation

6.3.1 Introduction

The use of predominant period (τ_c) of P-wave as a potential indicator of earthquake's magnitude produced signignificant results in several cases (Japan – Nakamura, 1988, Horiuchi et al., 2006a; Taiwan, Wu & Teng, 2002; Italy, Olivieri et al., 2008). τ_c is calculated as :

$$\tau_c = \sqrt{4\pi^2 \frac{\int_0^{\tau_0} u^2(t)dt}{\int_0^{\tau_0} \dot{u}^2(t)dt}} = \sqrt{\frac{\int_0^{\infty} |\hat{u}(f)|^2 df}{\int_0^{\infty} f^2 |\hat{u}(f)|^2 df}} \qquad (6.1)$$

Where *u(t)* and $\dot{u}(t)$ are the ground motion displacement and velocity as a function of time, respectively. τ_0 is the duration of the P wave that used for the calculation and usually taken between 3 and 4 secs. For the calculation an iterative algorithm is used which can be implemented for real time operation (Allen & Kanamori, 2003).

The τ_c is obvious in time domain but Eq.6.1 shows that it must be calculated by means of a spectral domain expression too. That means that τ_c is a weighted measure of the spectral energy density in the seismogram, derived from its FT $\hat{u}(f)$ as a function of frequency *f*. The result is that there is strong possibility that the information contained in time frequency representation of a seismogram could not be estimated accurately. A typical example is shown in Fig.6.1 where the spectrograms of two seismograms are computed (Hann tapering, 50% overlapping) and superimposed with the calculated τ_c for 4secs. What is suprising is that τ_c has non zero values several seconds before the first arrival of P wave (which indicated by left vertical line). As Simmons et al,. (2006) indicated this is the main reason for high scattering of magnitude estimation that observed in several studies (Allen & Kanamori, 2003, Olson & Allen, 2005, Abercrombie, 2005)

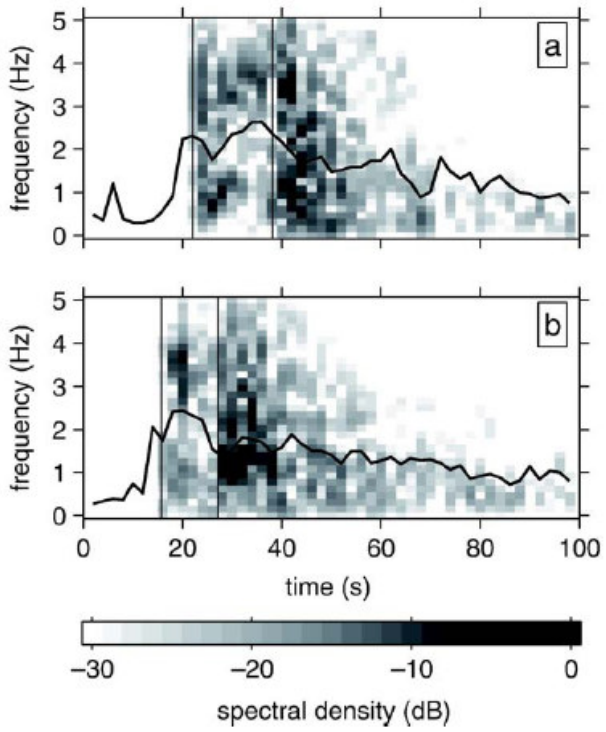


Figure 6.1 Spectrograms of two different seismograms from different earthquakes. Thick line indicates the values of predominant frequency; vertical lines indicate the P wave arrival (left line) and S wave arrival (right line) (Simmons et. al.,,2006).

The above shortcomings can be avoided if the magnitude estimator is investigated in time-scale domain by using WT. This idea is introduced by Simmons et al., (2006) who concluded that there is correlation between the station averaged thresholded wavelet coefficients and local magnitude at higher wavelet scale.

6.3.2 Data analysis and method's details

The area of interest that the possible EEWS will protect defined at the South Aegean sea, focusing on the island of Crete. Since the island lies over the back arc of Hellenic subduction zone it is expected that strong and devastating earthquakes might occur. This verifed the last decades from the several over M_L 6 earthquakes. The primary installed seismological network from National Observatory of Athens (NOA) had an interstation average distance round 220km. Recent installations from Geofon group and SNC-LGS all over Crete and surrounding islands decrase the average distance to 60km. This combined network cannot be characterized as dense;

for this reason is important to investigate the performance of the WME which has only be used in dense seismological networks (California – Simmons et. al., 2006; Japan, Mcguire et. al.,, 2008)

For the current research data from 181 earthquakes used. Data were collected from the seismological stations of SNC-LGS (from 2003 until 2008) and GEOFON stations (from 1999 to 2008, public available at www.orfeus-knmi.nl/data/webdc). It is important to clarify that in the current research, the earthquakes that used have $M_L \geq 3.8$ and their depth were up to 35km. we note that for each earthquake the used seismograms are those that recorded from stations up to 120km away from the epicentral (twice the average interstation distance in HSNC). A graphical representation of the above is shown in Fig.6.2

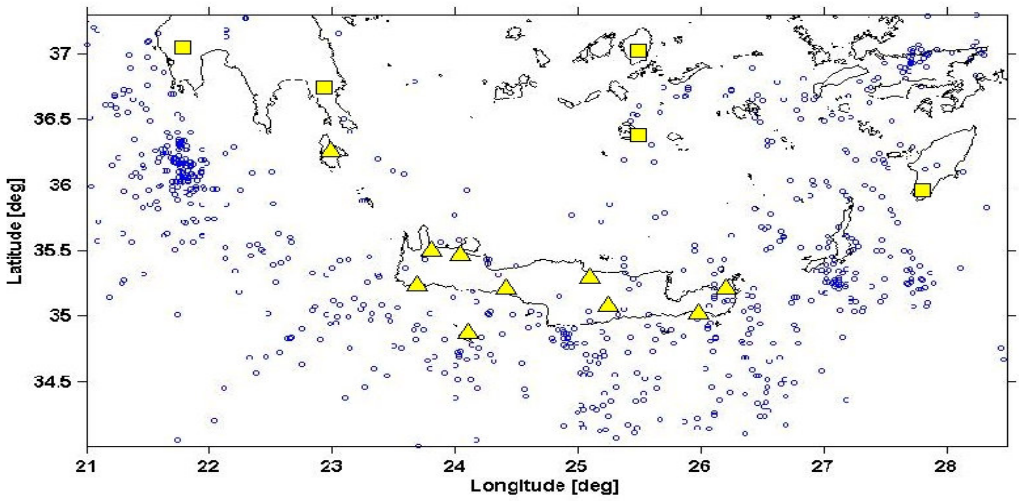


Figure 6.2 : Earthquakes (blue dots) and seismological stations (triangle:SNC-LGS, square: Geofon) used for this study

From each station only the vertical seismograms were used. Each seismogram is denoised using an appropriate (according to SNR) the wavelet denoising scheme as discussed in chapter 3. From each denoised seismogram 12sec around indicated onset time (as this derived from published earthquake bulletins, www.iaspei.org and www.emsc-csem.org) is wavelet transformed using MODWT. Unlike Simmons et. al, (2006) who used the lifting approach of WT here the use of a redundant WT is

dictated in order to avoid boundary effects; a fact the previous authors detected but didn't take into account. The MODWT used the LA(8) basis and run over 7 scales.

6.3.3 Results

A typical set of earthquakes recorded from the aforementioned stations is depicted in Fig.6.3.

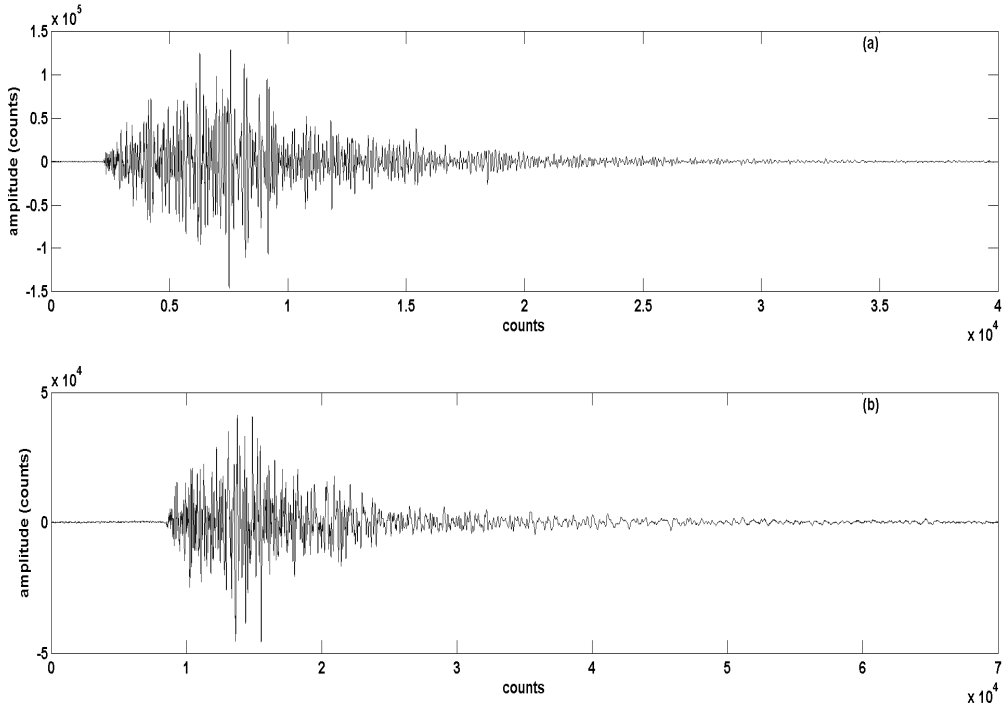


Figure 6.3: Vertical component seismograms from two earthquakes (a) M_L=5.5, depth 9km, epicentral distance 72km and (b) M_L=5, depth 11km, epicentral distance 111km

Transforming the above seismograms using MODWT with settings mentioned earlier produces the following results, as presented in Fig. 6.4 and Fig.6.5

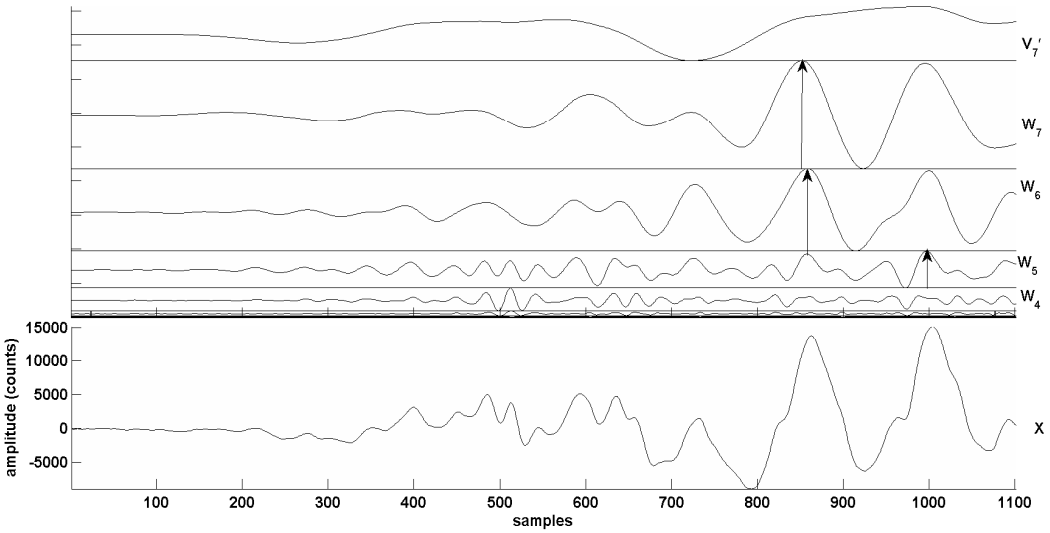


Figure 6.4: MODWT of fig.6.3.a seismogram. Vertical arrow indicates the presence of maximum amplitude coefficient at corresponding scale.

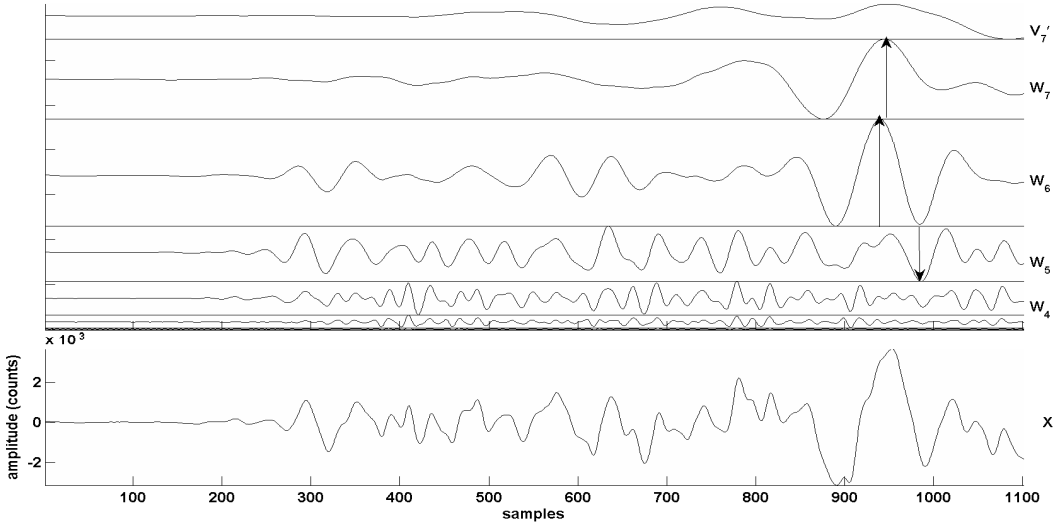


Figure 6.5: MODWT of Fig.6.3.b seismogram. Vertical arrow indicates the presence of maximum amplitude coefficient at corresponding scale.

The above results are characteristic for all the examined seismograms. A preliminary result that can be extracted from Fig.6.4 and Fig.6.5 is that the maximum amplitude coefficients are presented at higher scales. Also for the stronger earthquake the maximum coefficient is closer to the P wave arrival. These two observations

agree with the observations in predominant period studies where it is observed that stronger earthquakes present lower predominant frequencies in the first few seconds after P wave arrival.

The goal is to translate the wavelet transformed seismogram to information related to earthquake's magnitude. For this reason we keep first significant wavelet coefficient at each scale for 4secs time window (t=0 denotes P-wave arrival). For each distinct magnitude the average of the absolute wavelets coefficients value is calculated over all stations that satisfy the epicentral and depth criteria. From these averages, the best-fit regression line is derived. The total number of vertical component seismograms that satisfy the epicentral and depth criteria were 962. Results are presented at Fig.6.6 for scales 4-7 (scales 1 to 3 are not present any significant correlation).

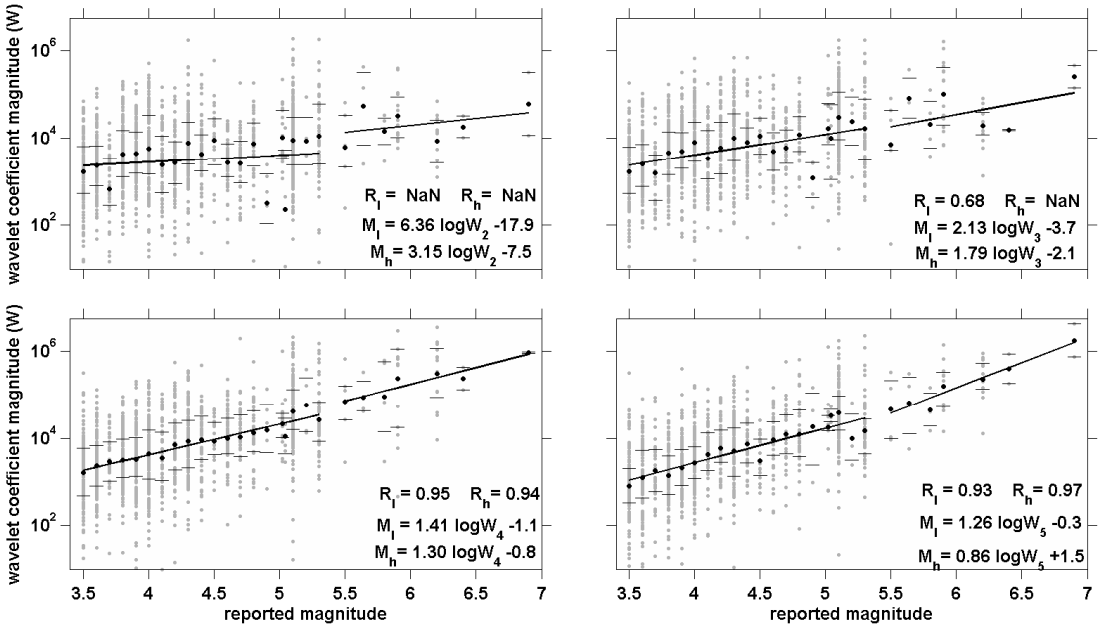


Figure 6.6: The correlation of wavelet amplitudes at scale 4 (top left), 5 (top right), 6 (bottom left) and 7 (bottom right) with earthquake magnitude. Results for individual seismograms are shown as grey dots (some share the same amplitude values), the 25th and 75th percentile as thin black dashes and the averages at every distinct magnitude by black squares. The thick line is the least-squares best fit line to the average values for high (M_h) and low (M_l) magnitude values.

The results from Fig.6.6 show that there are significant correlations between station averaged wavelet coefficients at last two scales (6 and 7). More specific, best-fit regression is achieved by low earthquake magnitude range in scale 6 and high magnitude range in scale 7. By using the corresponding equations to predict the magnitude, the average errors are below 0.7 at maximum. These facts is presented in Fig.6.7

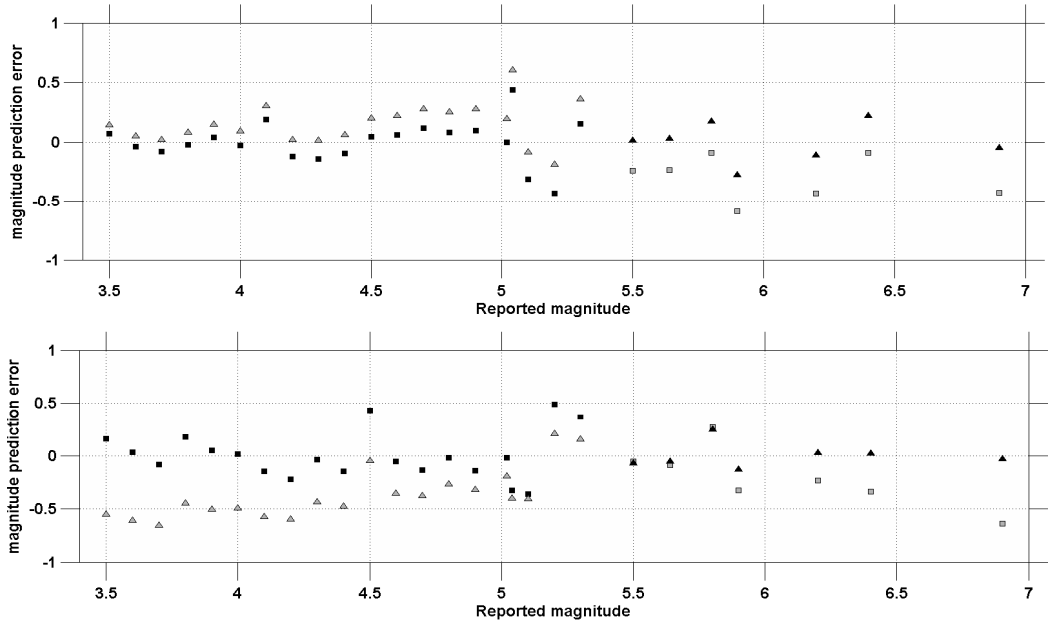


Figure 6.7: Magnitude errors of the prediction using equations derived at scales 6 (top) and 7 (bottom). Errors shown as : black squares when low regression used for low magnitude prediction, grey triangles when high regression used for low magnitude prediction, black triangles when high regression used for high magnitude prediction and grey squares when low regression used for high magnitude prediction. The reported magnitudes extracted from EMSC catalogue ($M_{L,EMSC}$)

From the above results (see Fig.6.6 and 6.7), the minimum errors produced when using low regression for low magnitude prediction from scale 6 and high regression for high magnitude prediction from scale 7. Thus the two equations characterizing the magnitude prediction using wavelets coefficients for South Aegean are:

$$M_l = 1.41 * \log(W_6) - 1.1 \qquad (6.2)$$

For reported magnitude $M_{L,EMSC} \leq 5.2$

And

$$M_h = 0.86 * \log(W_7) + 1.5 \qquad (6.3)$$

For reported magnitude $M_{L,EMSC} \geq 5.2$

In order to compare the WME with predominant frequency estimator for the South Aegean seismological station, analysis using the latter estimator was performed for the same set of earthquakes and for the same distance and depth criteria. The errors between predicted magnitude using predominant frequency estimator shown in Fig.6.8.

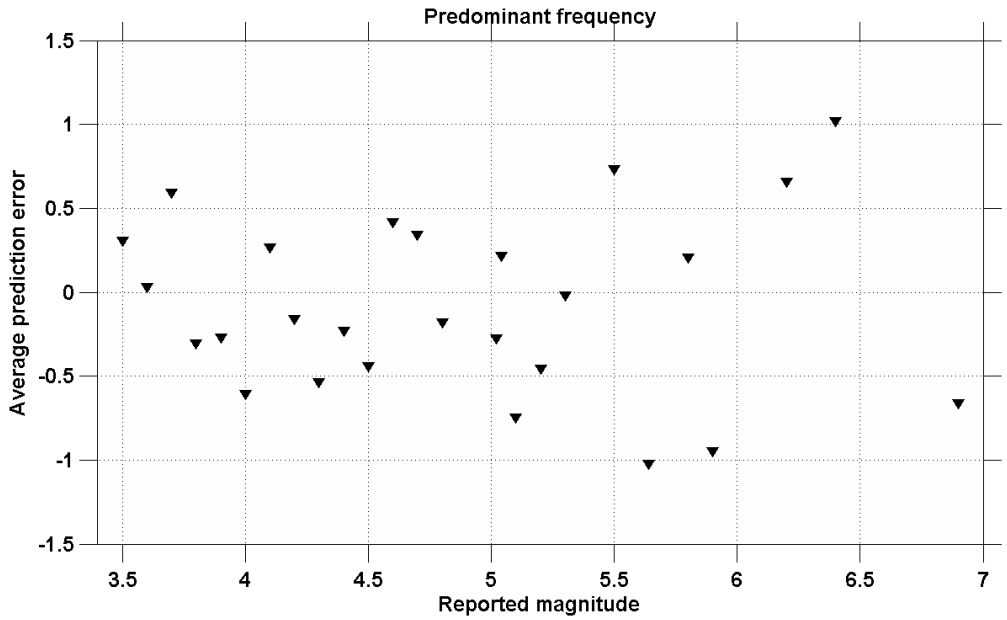


Figure 6.8: Magnitude prediction errors when predominant frequency used as estimator.

The bigger scattering in results in comparison with WME, agrees with the same observations from Wu & Teng (2002) and Wu et al. (2006a) and shows that WME is more suitable estimator for South Aegean area.

6.4 Wavelet Epicentral Estimation

6.4.1 Introduction

There are many approaches to standard earthquake location, which is performed when all the phase arrival times (P and S mainly) for an event are available. The goal for an EEWS is the use of a reliable method which can extract result as early is possible. For this reason the majority of the modern epicentral location methods focus on a solution that can be derived using only the P arrivals and

if possible without waiting for triggering from three stations (Front et al., 2004, Lomax, 2005, Rydelek and Pujol, 2004, Horiouchi, 2005, Satriano et al., 2008). The aforementioned methods use the principle of equal differential time formulation (EDT) of Font et al. (2004) and Lomax (2005) for standard earthquake location. EDT is a generalization of the master-station method (Zhou, 1994) and the "method of hyperbolas" cited by Milne (1886). Horiuchi et al. (2005) combine standard L_2–norm event location, EDT location on quasi-hyperbolic surfaces, and the information from not yet arrived data to constrain the event location beginning when there are triggered arrivals from two stations. The two arrivals define a hyperbolic surface on which the event can be located.

The above studies showed that it is possible to estimate epicentral location with accepted accuracy for an EEWS by using two stations and waiting the third in order to refine the result. Two problems that still are under investigation from the above methods are: (a) what is the time difference that the EEWS surcharged when one or more stations of the seismological network are not operational and (b) is it possible to increase the accuracy of the estimated by two stations epicentral ? For answering these problems we propose the use of WEpE which is a new method based on the combination of wavelet azimuth estimation (Galliana, 2007) and two stations method (Rydelek and Pujol, 2004)

.

6.4.2 Wavelet Azimuth Estimation (WAE)

For a regional earthquake, the P wave is linearly polarized in the direction of propagation. The N, E and Z components of a seismogram can be projected onto the propagation direction. This projection, along the direction of incoming seismic signal L, presents the maximum variance (Fig 6.9).

If the incoming seismic signal is noise free then this projection is zero in any other direction than L. Obviously this is not the case of real seismograms. In practice we expect the variance in direction L to be dominant and having extremely low values at any other direction. Folowing this fact, Magotra et al., (1989) provided the necessary formulation to calculate the azimuth (θ angle) and elevation (φ angle).

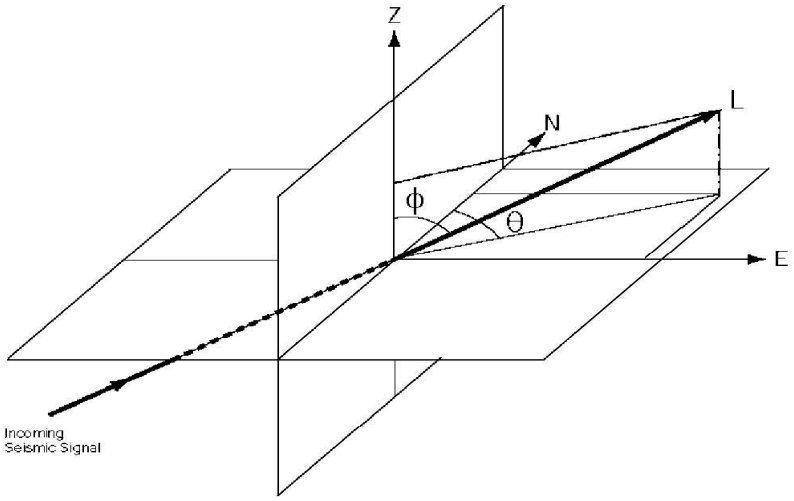


Figure 6.9 : Projection of the incoming seismic signal (Galliana, 2007)

For the automatic azimuth estimation the method of Galliana et al., (2007) is used. This method uses three steps that can be summarized as follows:

a) Denoising. It is very important that the incoming P wave to be as much noise free is possible; otherwise the difference between the maximum variance value and the minimum ones (as described earlier in the previous paragraph) could not be clearly dominant. Also a noisy seismogram (as discussed in chapter 3) usually provides false information about P wave onset time.

b) P wave arrival detection. A modified STA/LTA detector is used for this purpose. The modification consists in the use of a LTA threshold along with the threshold value of ordinary STA/LTA. Under this approach, the first arrival time is defined when the value of STA/LTA ratio exceeds its threshold as well as the associated LTA also exceeded. In the present research the signal that used is not the denoised signal as proposed by Galliana et. al., (2007) but the wavelets coefficients of the denoised seismogram associated with scale 3. We note that the selection of scale 3 is consistent with the frequency proposed by Galliana et. al., (2007) since for 125Hz sampling rate this scale corresponds to frequency band 15.6 – 7.8Hz). After experimenting using the data of HSNC, the most appropriate values that must be used for Galliana's detector are STA/LTA = 1.5 and LTA = 0.002 although as it will explained later we can select a value up to 0.006. In addition the window length of the detector is investigated below.

c) Azimuth estimation. This is achieved by using an adaptive length window in order to include only two periods of each component of denoised seismogram. Then, the union of three component windows defines the final time window where the covariance matrix and the eigenvector associated with the maximum eigenvalue is estimated. By using this eigenvector, the θ and φ angles (thus azimuth and elevation) are calculated (Magotra, 1989).

In order to investigate the applicapibility of WAE method in installed seismological network we used a set of 20 local shallow earthquakes with magnitude from 3 to 6. We used the HYPO2000 (Klein, 2000) software to estimate accurately the location and the azimuth of each event. The specific software is used as the *defacto* package from the vast majority of the seismological networks all over the world for the manual estimations of epicentral location. It accuracy is ensured when an event is recorded at least from three stations. The events that used for the current comparison are recorded from at least four stations. HYPO2000 can also produce the errors associated with each angle giving us the ability to compare its accurate results with the results obtained by the WAE method. Table 6.1 shows the events used. Then these events are used by the WAE in order to estimate the azimuth as explained before.

Event	*Longitude*	*Latitude*	*Magnitude (M_L)*	*Depth (km)*
1	21.11	35.89	3.3	3
2	23.74	36.18	3.1	12
3	23.63	36.21	4.2	11
4	25.8	36.23	4.5	6
5	22.41	35.26	4.5	12
6	21.64	36.28	5.0	9
7	26.79	34.25	5.3	10
8	22.64	35.22	4.8	11
9	21.81	36.32	6.0	23
10	21.80	36.17	5.6	7
11	23.70	36.29	3.4	3
12	23.87	34.32	3.6	6
13	21.90	36.36	3.9	7
14	26.77	35.33	4.9	9
15	24.74	36.06	3.5	8
16	23.52	35.25	5.1	11
17	21.69	34.27	4.8	11
18	23.84	36.37	4.4	9
19	24.76	34.36	5.4	7
20	26.82	35.27	5.5	11

Table 6.1.: Events source parameters used for azimuth test study

The results from HYPO2000 estimation and discussed method are presented in Table 6.2 and their difference is depicted in Fig.6.10. The results from HYPO2000 used as reference values and difference between them and the calculated value of WAE method denoted as WAE error.

Event	*Azimuth by Hypo2000 (deg)*	*Azimuth by WAE (deg)*
1	91.5	87.1
2	56.7	61.5
3	134.2	132.1
4	89.1	90.9
5	23.5	26.7
6	45.1	48.7
7	16.9	16.1
8	22.4	24.5
9	123.5	121.9
10	78.8	78.2
11	144.6	139.8
12	56.6	58.2
13	55.7	54.2
14	45.6	46.9
15	61.6	58.7
16	45.7	47.3
17	44.2	41.8
18	34.6	33.3
19	49.2	49.1
20	112.8	110.1

Table 6.2: Azimuth estimations using WAE and HYPO2000

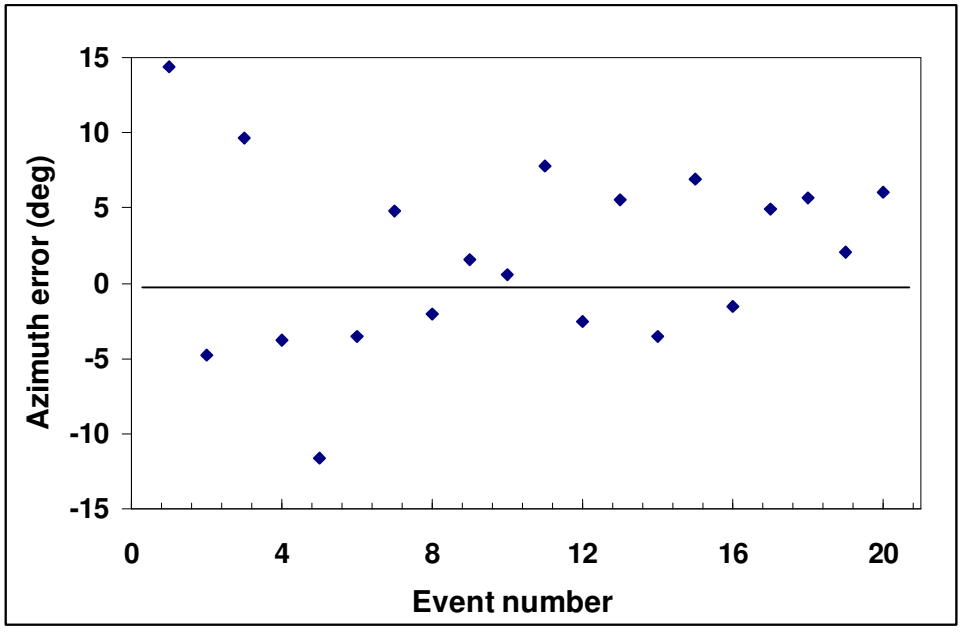


Figure 6.10. Difference between the azimuth estimated by WAE method and the azimuth obtained by HYPO2000 (WAE uses STA/LTA = 1.5 and LTA=0.002)

As mentioned earlier, the LTA threshold can vary from 0.002 to 0.006. Low value of LTA gives a sensitive detector which can produce arrival estimations prior to the real. On the contrary, high values of LTA produce arrival estimations close to the real but there is big possibility to include also other phases than P, leading in this way to wrong azimuth estimations. In order to investigate how this variation affects the error produced between WAE and HYPO2000 solution for each event we have used an LTA threshold ranging from 0.001 to 0.006 with 0.001 steps. For each case and event the errors calculated and the results are depicted in Fig 6.11

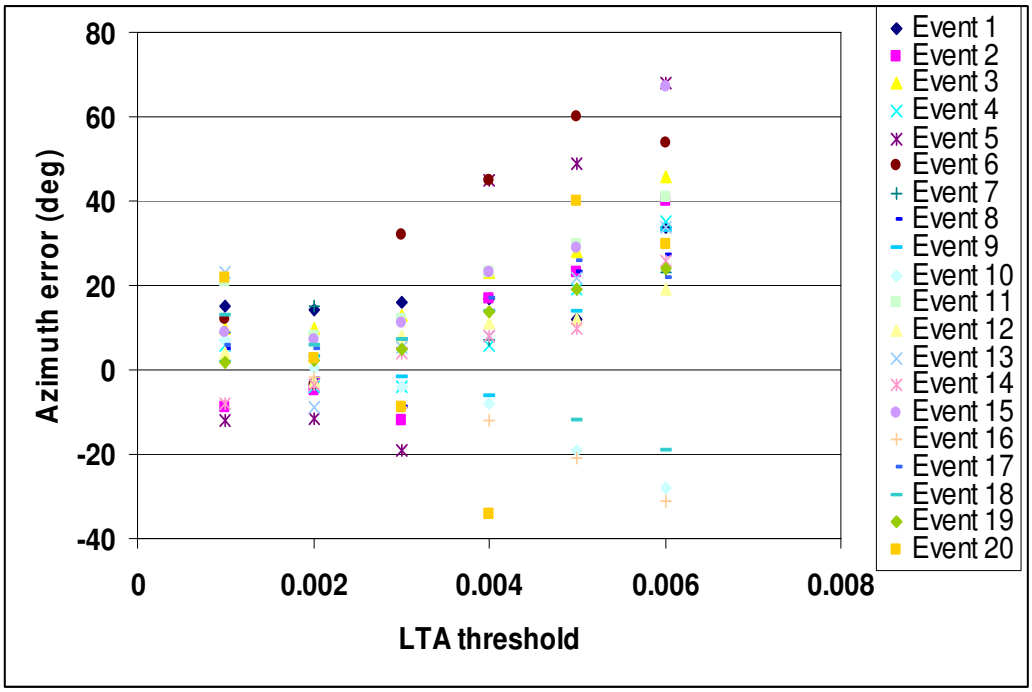


Figure 6.11 : Azimuth errors for WAE method for different LTA thresholds

The results indicate that WAE can determine the azimuth with a standard deviation around 5.9°. The LTA threshold that seems suitable to the seismological stations of SNC – LGS is 0.002 since this is the value that produces the minimum azimuth errors in the majority of examined events (17 from 20). By keeping in mind that the whole procedure will need around 1.2secs in a typical PC we can conclude that WAE is a suitable method for an EEWS.

6.4.3 Epicentral estimation using Two-Station Subarray (TSS) method

TSS method is based on the concept of "not yet arrived data" proposed by Horiouch et al., (2005). A typical configuration for illustrating the method is shown in Fig 6.12

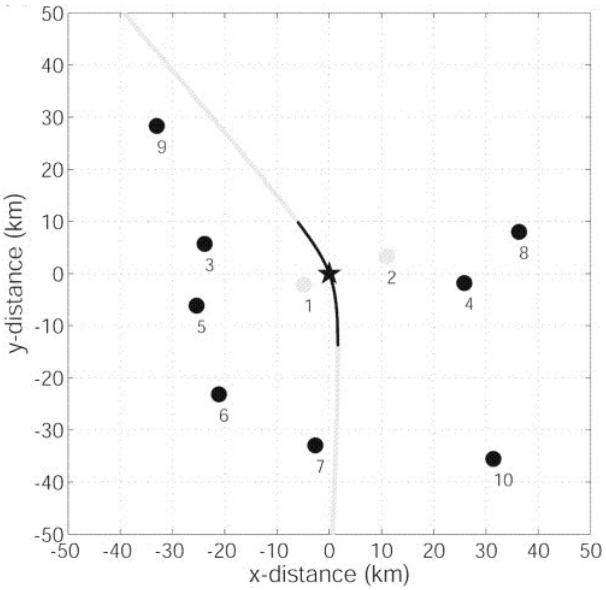


Figure 6.12: Illustration of TSS method. Star denotes the earthquake, grey dots are the triggered stations and black dots are the non triggered stations (Rydelek & Pujol, 2004)

Ideally, the seismic waves generated by this earthquake will eventually be recorded at all seismic stations in the figure. In case where only stations 1 and 2 (grey circles) at distances d_1 and d_2 are triggered from the first two *P*-wave arrivals at times t1 and t_2, respectively. It can be easily shown (in the good approximation of a half-space Earth that $d_2 - d_1 = V_p(t_2 - t_1)$, where V_p is the local seismic P-wave velocity. Since this velocity has known value , it is obtained that the earthquake occurs along a curve that satisfies a simple condition: the locus of points such that the difference in the distances from two fixed stations (or points) to any point on the curve is constant. This is the definition of a hyperbola, which is drawn as the grey curve in Fig.6.12. If we waited for the P-wave arrival at a third station, two other hyperbolas could be constructed, and the area of common intersection would thus provide the approximate location of the epicentre of the earthquake. The hyperbolic method of earthquake location was first used by Mohorovicic (1915) and remained in use over the years until the advent of computer-based methods.

The two-station subarray method proposed here is based on the information that the epicentral location along the hyperbola can be constrained due to the fact that the first two arrivals were recorded by two particular stations, as illustrated in Fig.6.12. Because station 3 did not record the first P-wave arrival, the distance along the constrained (black) segment of the hyperbola from the epicenter to station 3 must be larger than to station 1. If the epicenter were on the hyperbola slightly north of the black segment, then station 3 would have recorded the first arrival and not station 1, and the hyperbola would have been different. The same argument is used to constrain the southern extent of the black segment; going slightly to the south of it means that station 7, instead of station 1, would have recorded the first P arrival. The length of the black segment of the hyperbola in Fig.6.12 is about 25 km, and in general this length will depend on station geometry and interstation spacing in the seismic array, as shown later for actual earthquakes.

6.4.4 Wavelet Epicentral Estimation

The TSS method discussed above needs two stations in order to estimate the hyperbola where the epicentre possibly lies. If the distance between the first two stations is quite high (~100km) then the hyperbola could extent up to 80 km. This is not accepted as epicentral estimation. In the case of SNC-LGS network there are cases where the interstation distance between two subsequent stations is quite high. For this reason there is a need for eliminating the limits where the hyperbola could extent. In order to achieve this, the use of WAE at the first triggered station, is proposed. The idea is that if the azimuth is known from the incoming seismic signal then we can define an area of interest as an isosceles triangle; with vertex angle the standard deviation of 5.9^{o} (as calculated earlier) pointing to the direction of azimuth and with equal sides calculated by the distance of the closest operating station (Fig. 6.13)

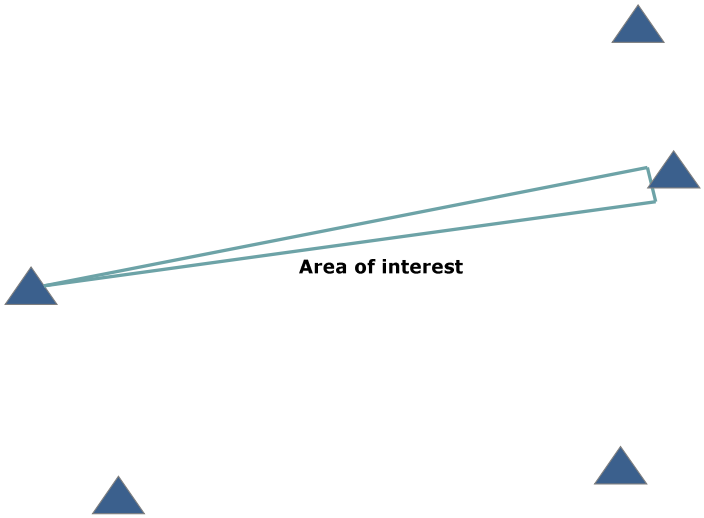
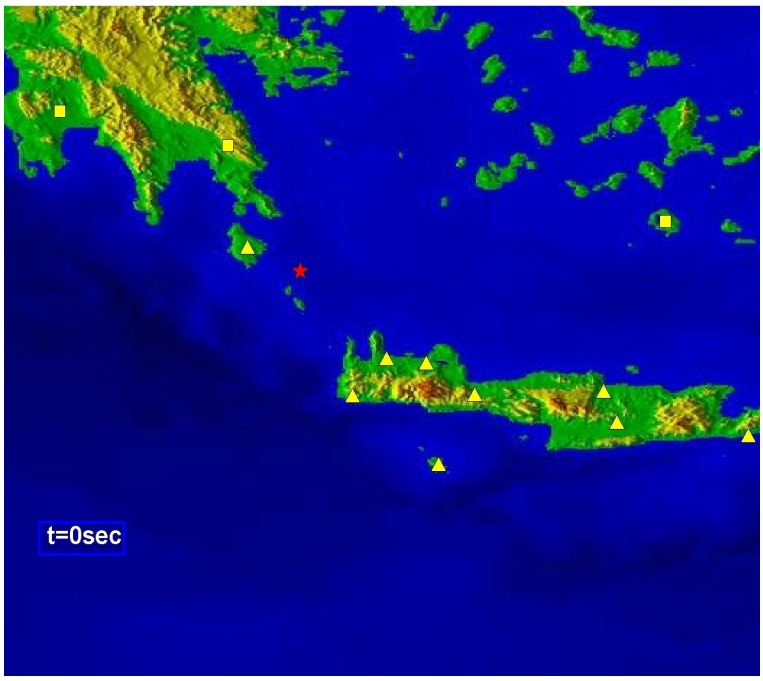
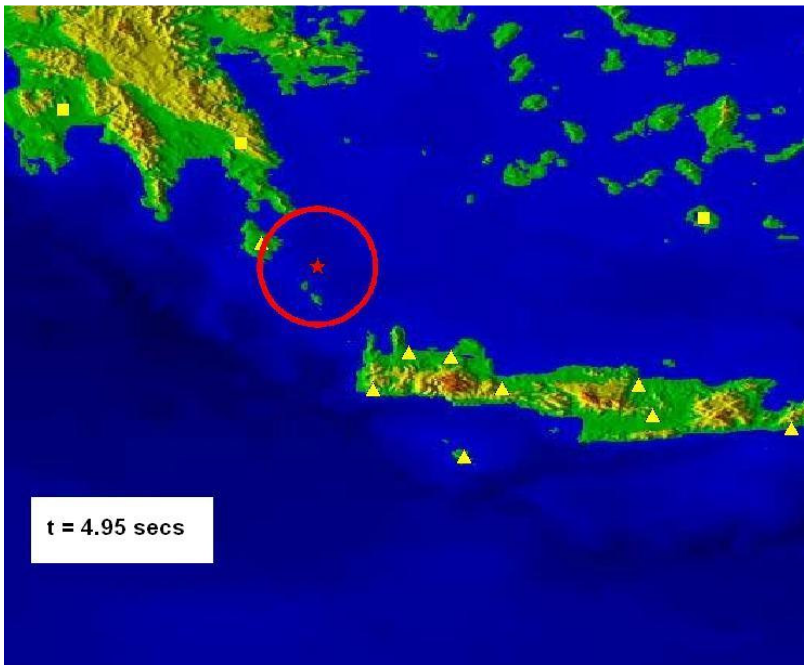


Figure 6.13: Illustration of the construction of "area of interest" using azimuth estimation

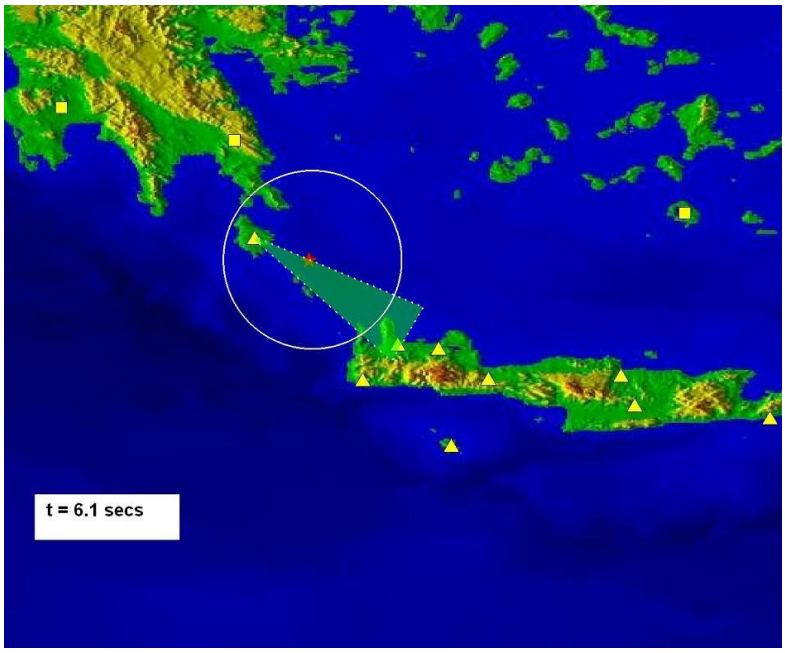
Since this area of interest can be defined using the first triggered station when the second station triggers and thus the hyperbola can be estimated, its limits can be shrink immediately. This is achieved by assuming as valid part of the hyperbola the part that lies down in the area of interest which is calculated before the second station triggers. This is depicted as a series of snapshot in Fig 6.14 where the M_L=6.5 earthquake occurred at January 8, 2006 used.



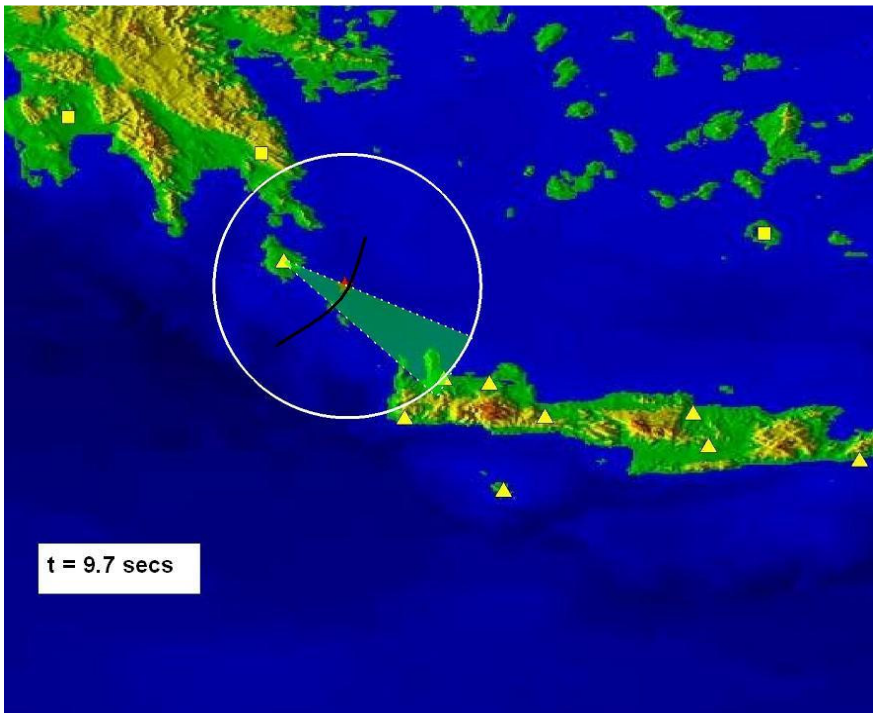
(a)



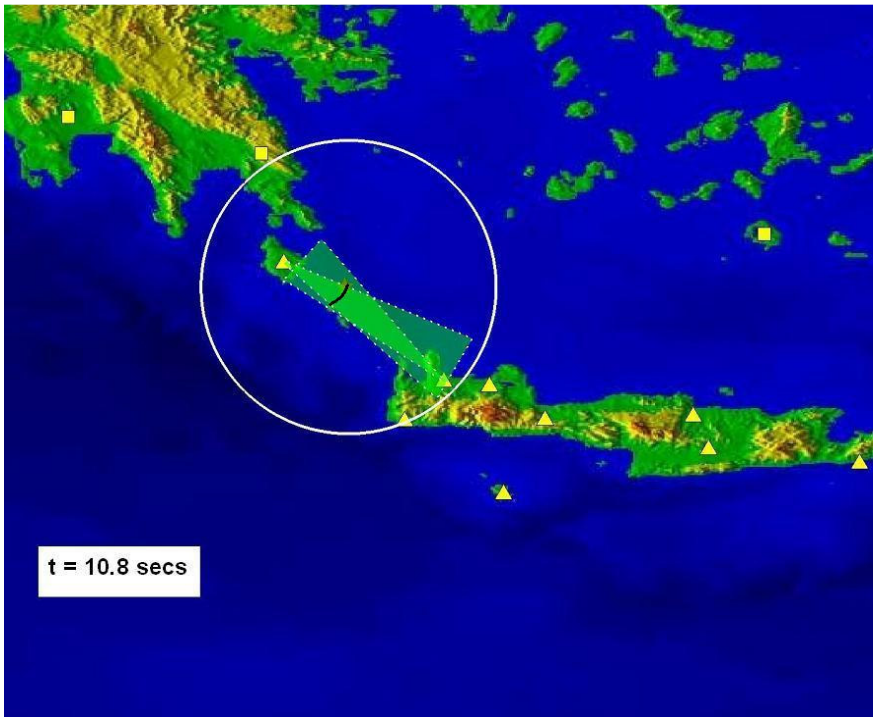
(b)



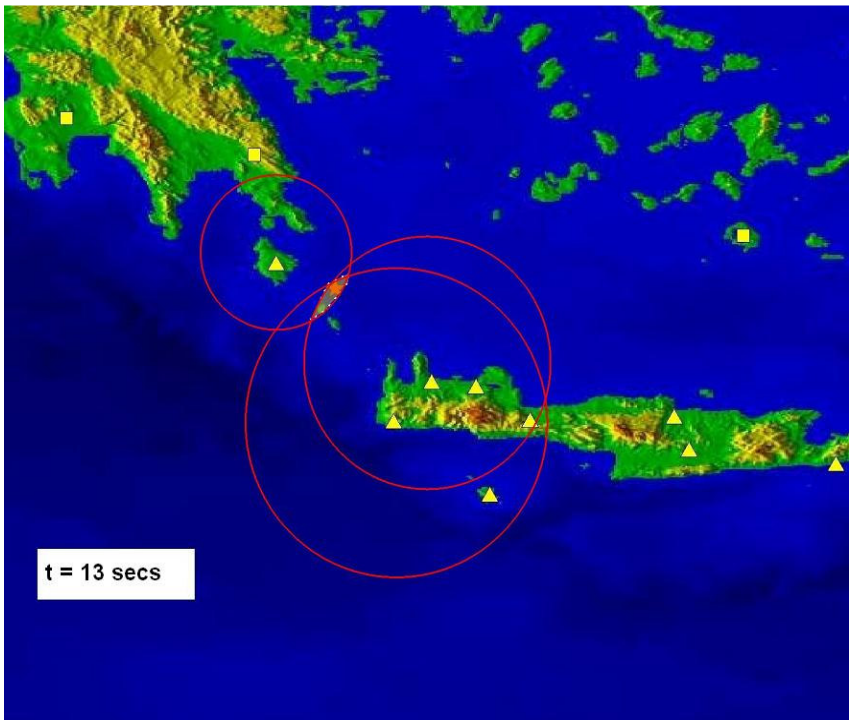
(c)



(d)



(e)



(f)

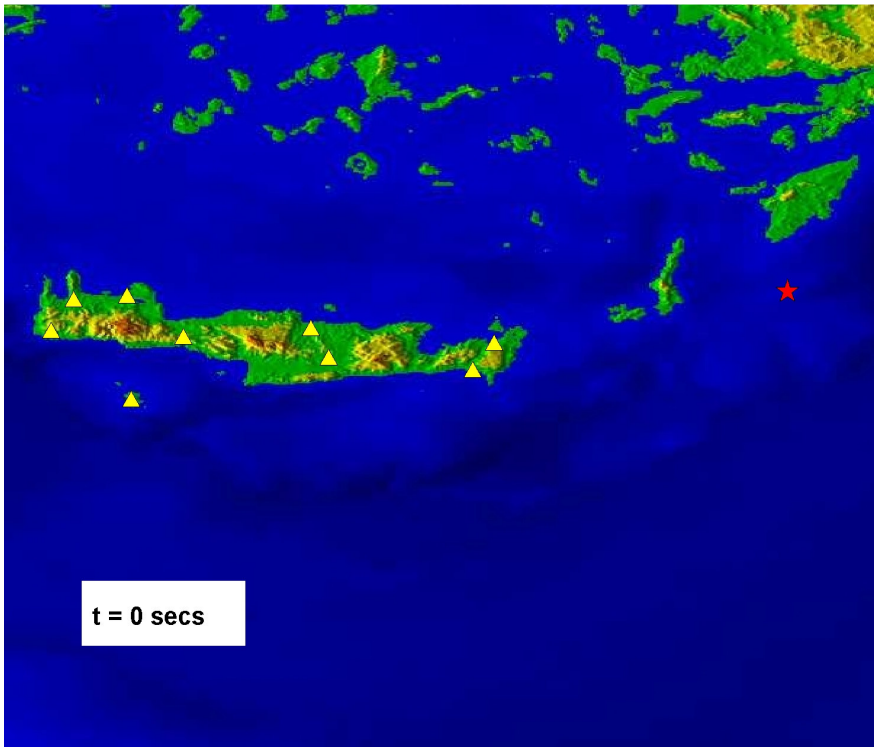
Figure 6.14: Snapshots of evolutionary detection of the epicentral location using WAE method

The time begins at earthquake's occurrence time (Fig.6.14.a). The first station triggers after 4.95secs (Fig.6.14.b) and from the P-wave is able to estimate the azimuth using WAE method. This is completed after 1.15secs and an area of interest can be defined (Fig.6.14.c). The area of interest (green shaded area) defines the area where the epicentrel expected to be. At t=9.7secs the second nearest station triggers and a hyperbola (black solid line) can be defined now using the TSS method (Fig.6.14.d). At t=10.8secs the second area of interest can be defined along with hyperbola (Fig.6.14.e). At this stage the hyperbola estimated from TSS method extends up to 90km. By keeping as valid only the portion that lies down inside the common part of two areas of interest, the possible hyperbola length decreased down to 12km which can be qualified as acceptable.

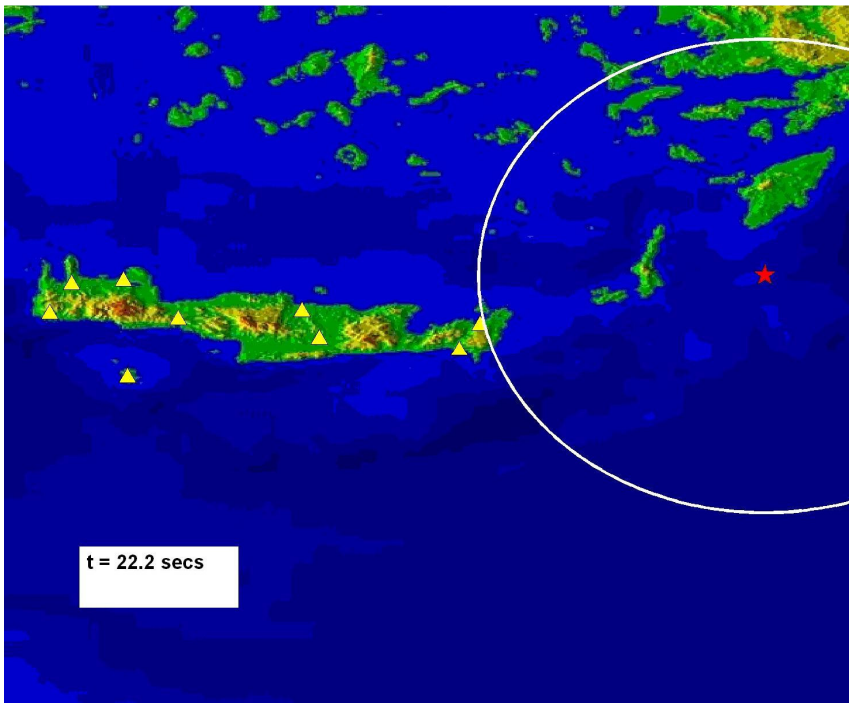
Another case of interest is when the earthquake occurs outside the network. In this case the WAE method proves valuable because it shrinks the intrinsic inability of TSS method to estimate a reliable hyperbola. This case is demonstrated in Fig. 6.15 where an earthquake occurs (red star in Fig 6.15.a). When the first nearest station triggers, it is possible to estimate the area of interest (Fig.6.15.b) after 1.1 secs

(Fig.6.15.c). During the calculation time the second station of the network triggers and it is possible to estimate the hyperbola. Since the event is outside the network the hyperbola will extent to east increasing the error in estimation (black solid line in Fig.6.15.c).

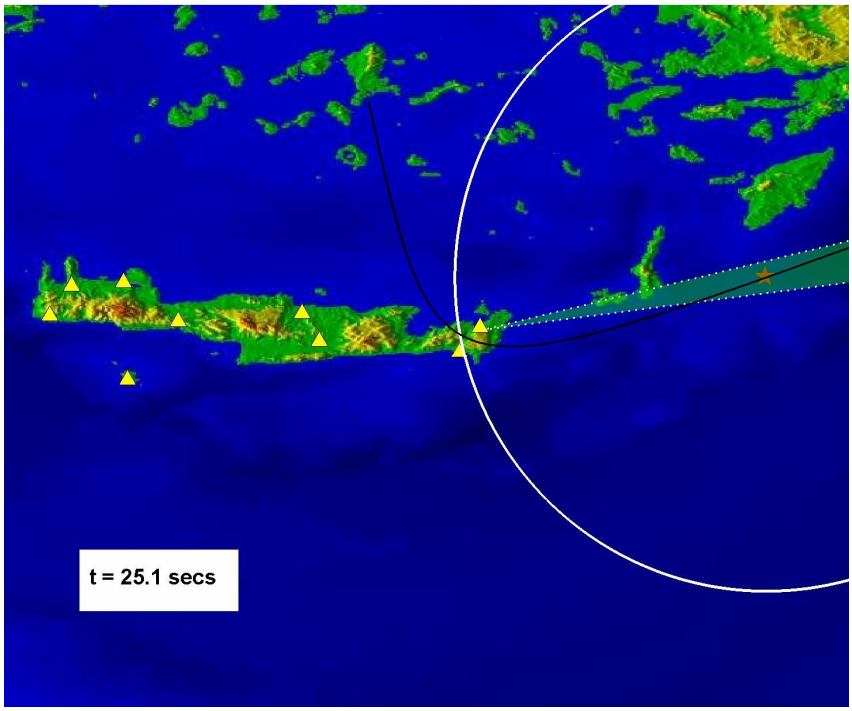
When the second triggered station estimates the area of interest then the limits of hyperbola reduced and the epicentral estimation lies between the union of two areas of interest and corresponding hyperbola points (black solid line in Fig. 6.15.d)



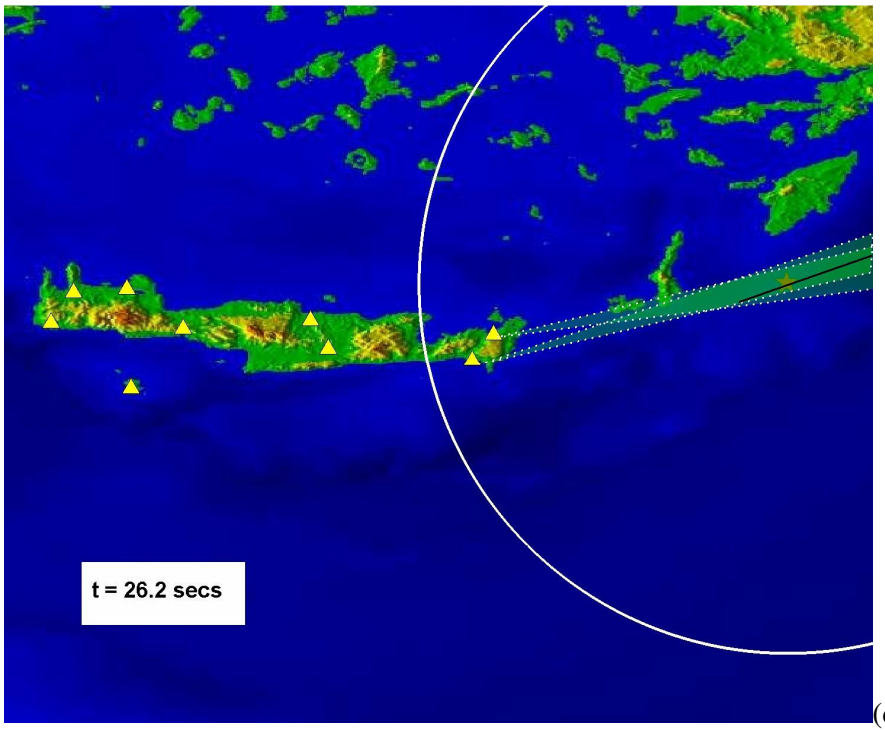
(a)



(b)



(c)



(d)

Figure 6.15: Snapshots of evolutionary detection of the epicentral location using WAE method with earthquake occurred outside network

In EEWS accurate earthquake location is of lesser concern compared to minimizing the amount of time needed to locate it, albeit imprecisely. Therefore the trade off with the WEpE method is between the uncertainty in the epicentral location and the time spent waiting for more stations to trigger. The usefulness of WEpE depends on the distance between stations. A dense network (~10km between stations) can provide location estimations using only the TSS method since the first three stations will trigger the first 3 secs (assuming a P wave velocity of 7km/sec) so there are no obvious benefits. For the National network where the interstation distance varies from 60km to 100km, WepE could provide epicentral estimations enough time before the third station triggers. As a general rule one can conclude that the case when unfortunately the distance between the stations is large then the greater the time saved compared to the time waiting the third station to trigger but at the cost of uncertainty in epicentral location. For large earthquakes this uncertainty becomes side issue since the main concern for an EEWS is a rapid estimation of location and magnitude.

6.5 Summary

In this chapter, the ability of WT as a common processing tool for the purposes of EEWS is demonstrated. More specific, it is shown that is possible to use the same set of wavelet coefficients from a transformed seismogram for performing two actions : rapid magnitude estimation and epicentral location. For magnitude estimation, it is shown that the absolute value of the wavelet coefficients at lower scales of a WT (here 6 and 7) correlate with the local magnitude; for this two empirical relations are calculated: one for low magnitude prediction (up to 5.2) and one for high magnitude prediction (higher than 5.2). In general, the errors produced by WME are no more than ±0.6. Compared with predominant frequency estimator, WME shows better performance for the South Aegean area. For epicentral location purpose it is shown that a new proposed technique, WHE, which combines two different known approaches from seismology (azimuth estimation and epicentral estimation using two stations), can be used to estimate in real time the epicental with improved accuracy and especially when epicentral location is outside the area defined by installed seismological network. This fact makes this approach very important for the purposes of a regional EEWS.

7. Multiresolution Wavelet Analysis of Seismicity

7.1 Introduction

The time distribution of seismic events was the subject of several studies, dealing with their time-scaling properties (Wilson & Domnic, 1990, Godano, 1996, Oncle et.al.1996). A new wavelet based technique, has been recently developed for the purpose of dynamic variability of seismicity investigation. It is based on scaling analysis of the temporal properties of sequences of earthquakes. This is not a new approach since many studies used set of scale-independent measures in order to examine the time fluctuations of a series of earthquakes (e.g., periodogram power-law exponent, Telesca et.al, 1998, the Allan–Factor power-law coefficient, Telesca et.al. 2002, the detrended fluctuation analysis power-law exponent, Telesca et.al 2002). The objective is to estimate the least-squares-fit slope of the investigated measure in a log-log (double logarithmic) plot coordinate system. These approaches provide the researcher with a quick picture and understanding of the correlation properties of earthquake series independently on the scale. Their disadvantage is that they are vulnerable to nonstationarities; a situation that is not very rare when characterizing the time dynamics of an earthquake sequence..

Multiresolution wavelet analysis has proved a valuable and accurate tool for the analysis of signals at multiple scales, especially in presence of nonstationarities which often contaminates such signals (Akay, 1987, Albroudi, 1996, Arneodo et.al., 1998). This method has been successfully used to discriminate healthy patients from those with cardiac pathology (Thurner et.al, 1998, Ashkenaz et.al., 1998). The above method produces a scale-dependent measure, the wavelet-coefficient standard deviation, which has been demonstrated to perform better than the scale-independent measures (Teich et.al, 2001). Recently, Telesca et al. (2004) have analyzed the time variation of the wavelet-coefficient standard deviation for two seismic sequences occurred in Italy, distinguishing between aftershock-related dynamics and background behaviour.

The wavelet-coefficient standard deviation is used for the analysis of South Aegean seismicity. Initially, the seismic sequence from 1970 until 2003 is examined and the strongest event can easily identified at lower wavelet scales. The procedure is repeated with a more complete catalogue (from 2003 to 2008) where the a set of double strong earthquakes is identified at lower scales

7.2 Formulation

As already mentioned in detail at Chapter 2, one of the basic advantages of wavelets is that they project a signal in the frequency domain as well as in the time domain simultaneously, unlike the usual Fourier transform where an event is accurately represented in the frequency or in the time domain only. This characteristic permits the successful derive of data with different behaviour through several scales and thus cannot provide a consistent multiresolution analysis.

In this chapter the analysis of time intervals between successive seismic events (IS) $\{\tau_i\}$, $i = 1, \ldots, L$, where L is the length of the series, is taking place The initial step is to transform the IS from time space into a space of wavelet coefficients in order to derive the scale dependant measures.

Mathematically speaking, the coefficients are obtained by carrying applying a DWT

$$W_{a,b}^{wav} = 2^{-a/2} \sum_{i=1}^{L} t_i \psi(2^{-a} i - b) \tag{7.1}$$

where the scale variable a and the translation variable b are integers, L represents the total number of IS intervals analyzed, and ψ is the wavelet function. The DWT is evaluated at the points (a, b) in the scale-interval-number plane. Smaller scales correspond to more rapid variations and therefore to higher frequencies. Orthogonality in DWT ensures that the information deduced at a certain scale a is disjoint from the information at other scales (Thurner, 1998).

The standard deviation of the wavelet coefficients, as a function of scale is defined as:

$$\sigma_{wav}(a) = \sqrt{\left[\frac{1}{N-1} \sum_{b=1}^{N} \left(W_{a,b}^{wav} - \left\langle W_{a,b}^{wav} \right\rangle \right)^2 \right]} \tag{7.2}$$

where N is the number of wavelet coefficients at a given scale a and the indicate the average among the coefficients at a scale a (Telesca et.al.,2004)

7.3 Geological and seismological settings

The SW segment of the Hellenic Arc (34ºN - 37.5º N, 20º E - 26º E) is the most active plate margin of the Mediterranean area, with correspondingly high seismicity and relatively

frequent occurrences of large earthquakes. The most dominant geotectonic feature is the Hellenic Trench, where the eastern Mediterranean oceanic lithosphere (front part of the African plate) is sub ducted under the Aegean microplate. The Hellenic Trench is parallel to the Hellenic Arc which consists of the outer sedimentary arc and the inner volcanic arc. The average distance between them is 120 km. The Sea of Crete appears a maximum depth around 2 km and defined between between the sedimentary arc and the volcanic arc. The front part of the African oceanic lithosphere is subducting under the continental Aegean Sea lithosphere as part of the collision process of Africa–Eurasia plates. (McKenzie, 1978; Jackson, 1994; Kiratzi and Papazachos, 1995 among many others, Benetatos et. Al, 2004)

The SW Hellenic Arc and Trench system generates large shallow- and intermediate-depth earthquakes many of which have been are reported since early historic times. Magnitudes of up to 8.0 have been reported in the literature, (e.g., Papazachos, 1990; Papazachos and Papazachou, 1997). Although such figures may have been overestimated, (some of them were drawn on the basis of ancient and medieval archival data), they still signify the great seismogenetic potential of the area.

Crete lies in the forearc of the Hellenic subduction zone, indicating the transition zone between the African and Eurasian plates. Active subduction is indicated by a north-dipping Wadati-Benioff seismic zone extending beneath Crete to a depth of about 200 Km (Le Pichon et. al.,, 1979; Knapmeyer and Harjes, 2000). Nowadays the Hellenic arc is associated with moderate arc-parallel extension and strong compression perpendicular to it. Subduction at the Hellenic subduction zone appears to have been operating continuously since at 26 Ma and likely back to 40 Ma (Spakman et al., 1988; Meulenkamp et al., 1988)

The seismicity of the region provides valuable information on the geometry of the subduction zone, the deformation and the stress field in the region. The Aegean Wadati-Benioff zone has been extensively studied (Knapmeyer, 1997). Benioff zone of Hellenic subbuction zone defined between 60 and 160 km depth by means of intermediate depth earthquakes. Inside the borders of Hellenic arc and around west and central Crete, shallow events clusters form two NE–SW striking lineaments, one close to the coast of Crete and the other at least 150km long which crosses central Crete (Telesca et. Al, 2004, 2007)

7.4 Application to South Aegean seismicity until 2003

The data that used in this study are the interevent times between successive seismic events of magnitude $M \geq 3.7$, from 1970 to 2003, extracted from the GI-NOA catalogue (http://www.gein.noa.gr/services/cat.html). $M_c = 3.7$ represents the minimum magnitude of completeness of both catalogues, evaluated by the Gutenberg–Richter analysis (Gutenberg & Richter, 1956). Fig.7.1 shows the epicenter distribution of the earthquakes analyzed.

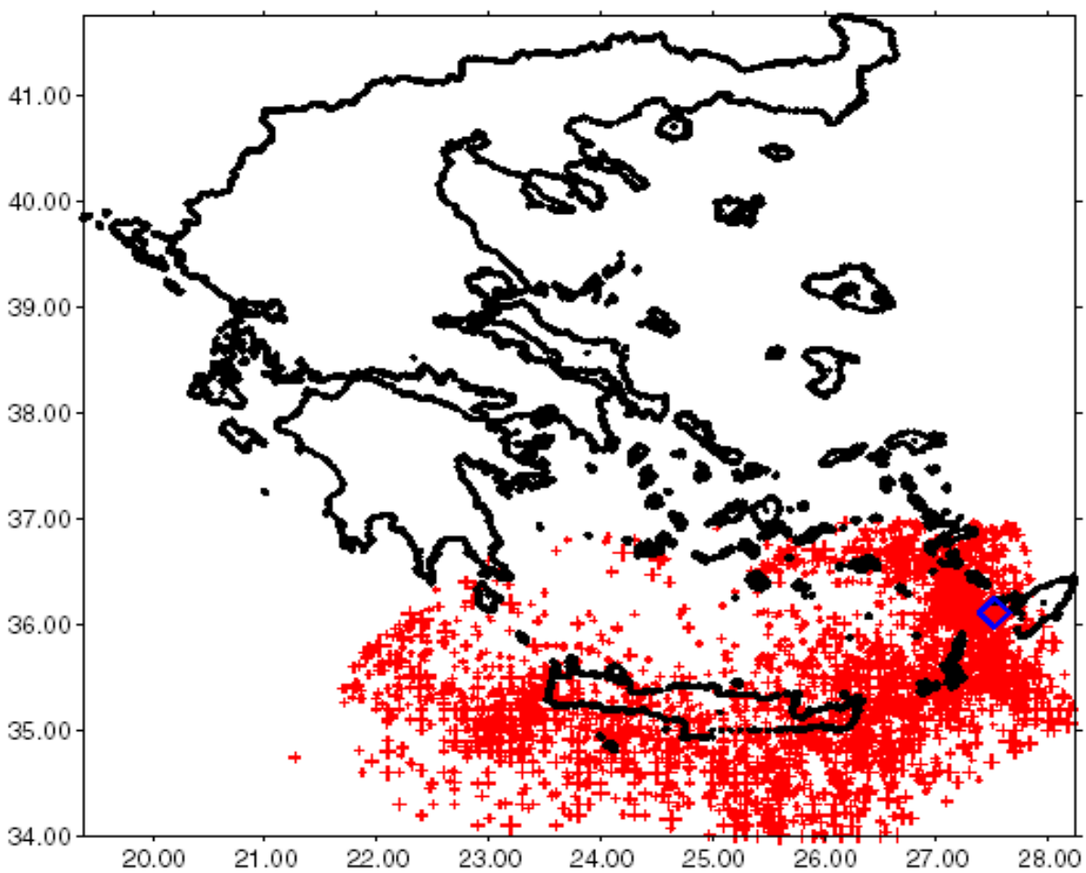


Figure 7.1 Epicentral distribution of earthquakes ($M \geq 3.7$) occurred in southern Aegean area over the period 1970–2003. The diamond indicates the largest earthquake of the sequence ($M = 6.1$) (Telesca et.al, 2007).

The 1970–2003 seismic sequence is composed by $L = 3214$ interevent intervals. The average and the standard deviation of the interevent times are approximately 5×10^3 min and 7.7×10^3 min.

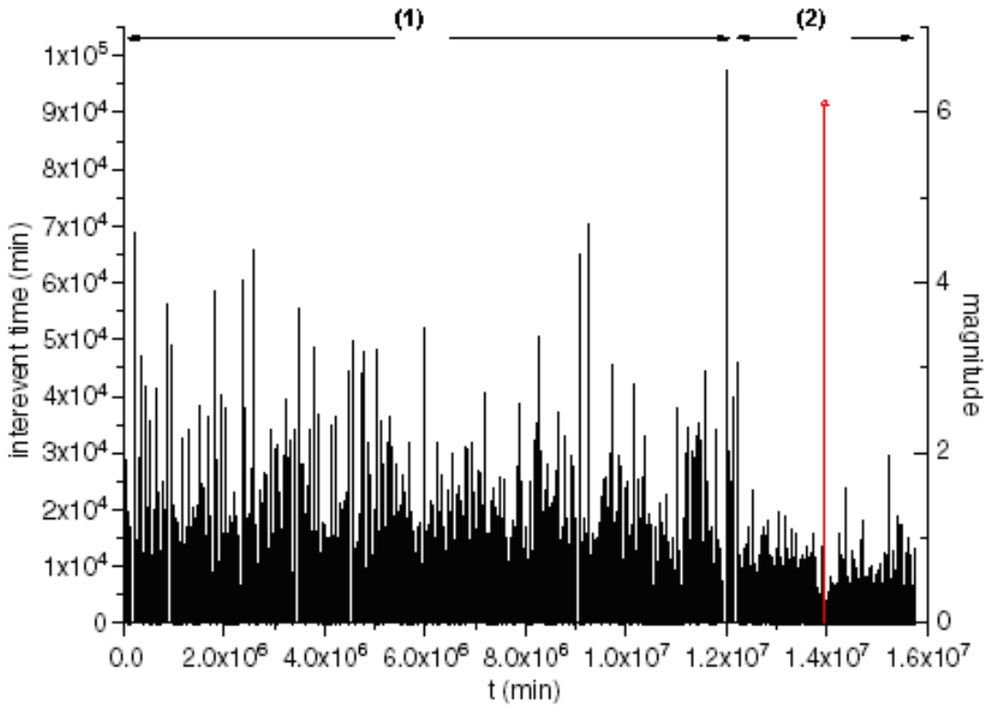


Figure 7.2. Interevent times between two successive events versus the time occurrence of the second event. Red arrow indicates the strongest event. (Telesca et. al., 2007)

Fig.7.2 shows the interevent times between two successive events versus the time occurrence of the second event. The vertical arrow indicates the occurrence time of the largest event of the sequence ($M = 6.1$). Two temporal ranges [(1) & (2)] of the IS can be identified: in the region (1) the distribution of the interevent times seems rather varying with relatively high mean interevent time, in the region (2) the distribution appears more concentrated around a lower value for the mean interevent time. Multiresolution wavelet analysis was selected as the tool for the investigation of time fluctuations of the earthquake IS. Based on the theoretical definitions on section 7.2 and because there is a relation between the length L of the series and the number of scales m (Thurner et.al, 1998) we could investigate up to 11 scales for the southern Aegean sequence. This is derived from the relation is $L = 2^a$

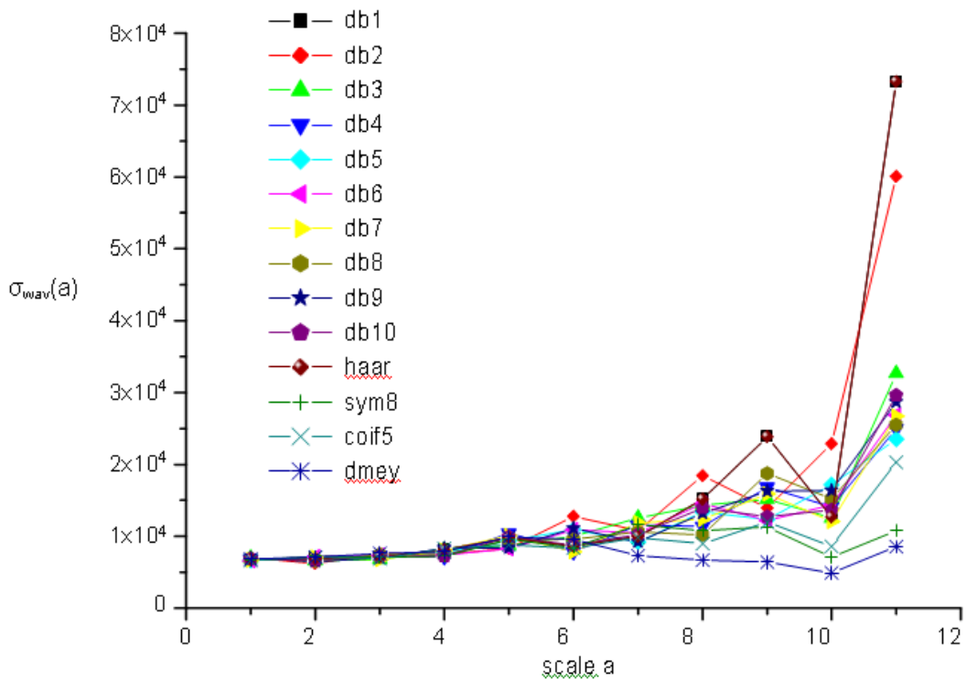


Figure 7.3. $\sigma_{wav}(a)$ for the seismic catalogue of southern Aegean area, calculated by means of different wavelet functions. (Telesca et. al., 2007)

Fig. 7.3 shows the standard deviation of the wavelet coefficients $\sigma_{wav}(m)$, calculated by means of different wavelets. Initially, one can easily identify that there is strong variability of the wavelet coefficients. This is depicted by the high value of $\sigma_{wav}(a)$ as the scale increases; this is an indication of the strong fluctuations displayed by the IS. In addition, $\sigma_{wav}(a)$ increases with the scale a. Lower scales are associated with higher frequency oscillations, thus an increase of $\sigma_{wav}(a)$ with the scale indicates that the fluctuations of higher frequency are less strong than those of lower frequency. Another observation is that all the wavelet bases produce similar results at small scales up to $a = 5–6$. This can be explained by keeping in mind that the wavelet coefficients is given by the relation $L/2^a$ which means that at high scales a, the number of wavelet coefficients becomes decreases. This situation produce spreading among the standard deviations of different wavelet bases.

The time evolution of the $\sigma_{wav}(a)$ is investigated by using a fixed event number window (500 events), shifting through the entire series. The same window legth is used for the analysis of time variation of the $\sigma_{wav}(a)$ for $a = 1, \ldots, 8$. Different wavelets bases have been used and produced similar results. Fig. 7.4, in the present section shows only the results obtained by means of the Daubechies 10 and Haar wavelets. Fig.7.4 through Fig.7.8 shows the variation of $\sigma_{wav,m}(t)$ with *wav* = *coif5*, *db5*, *db10*, *dmey*, *haar* for the southern Aegean catalogue. In each plot, the strongest earthquake (indicated by the vertical arrow) that occurred in the area ($M = 6.1$) during the period of observation, is reported. The first common feature in all the plots is the variability of $\sigma_{wav,a}(t)$. There is significant variability which means that there is temporal variability of the multiscale properties of the IS. The second, most striking feature is a pattern that is observed in all the plots: mainly at lower scales, the major shock of the seismic sequence is associated with a decrease of the temporal evolution of the $\sigma_{wav,a}(t)$. Regarding the last scales (a= 7 and $a = 8$) which correspond to lower fluctuations of IS, a more regular and cyclical structure is assumed since there is no such sharp decrease as before..

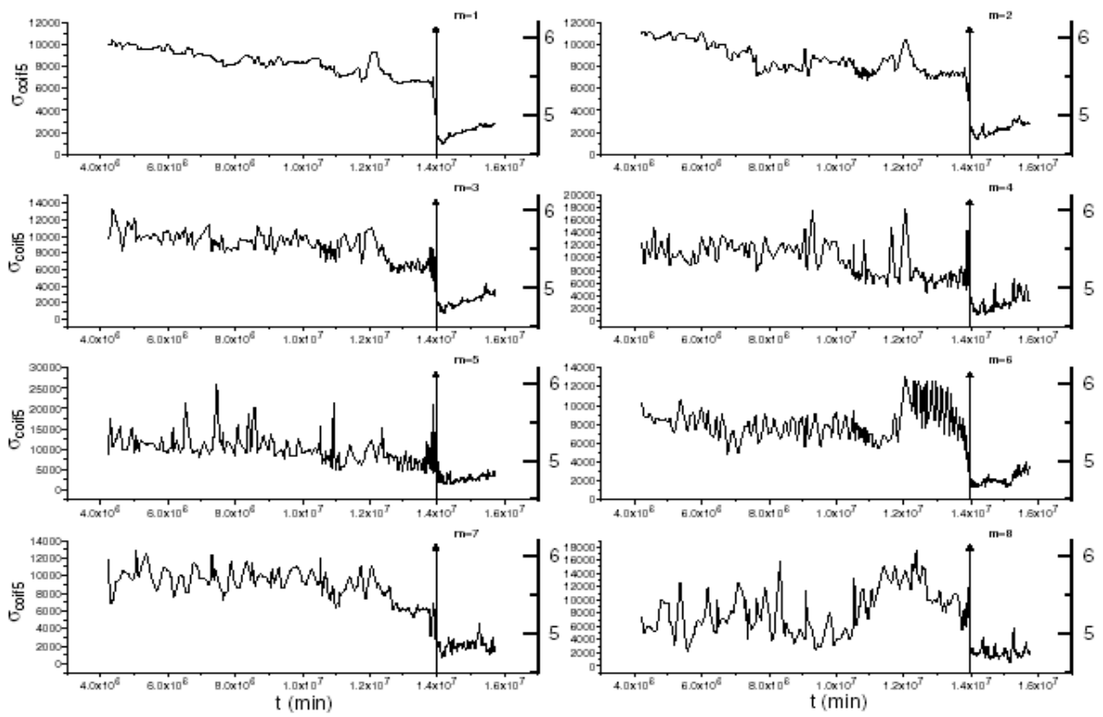


Figure 7.4. Time variation of $\sigma_{coif5,a}$ with a ranging from 1 to 8. (Telesca et. al., 2007)

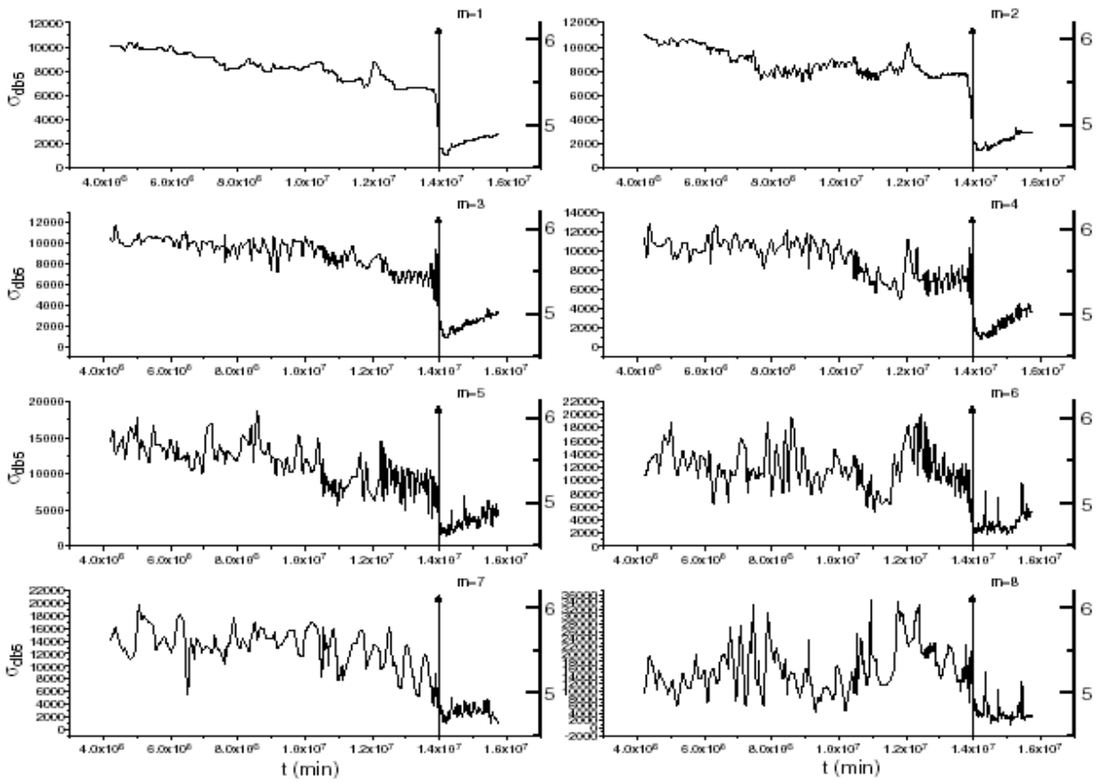


Figure 7.5. Time variation of $\sigma_{db5,a}$ with a ranging from 1 to 8. (Telesca et. al., 2007)

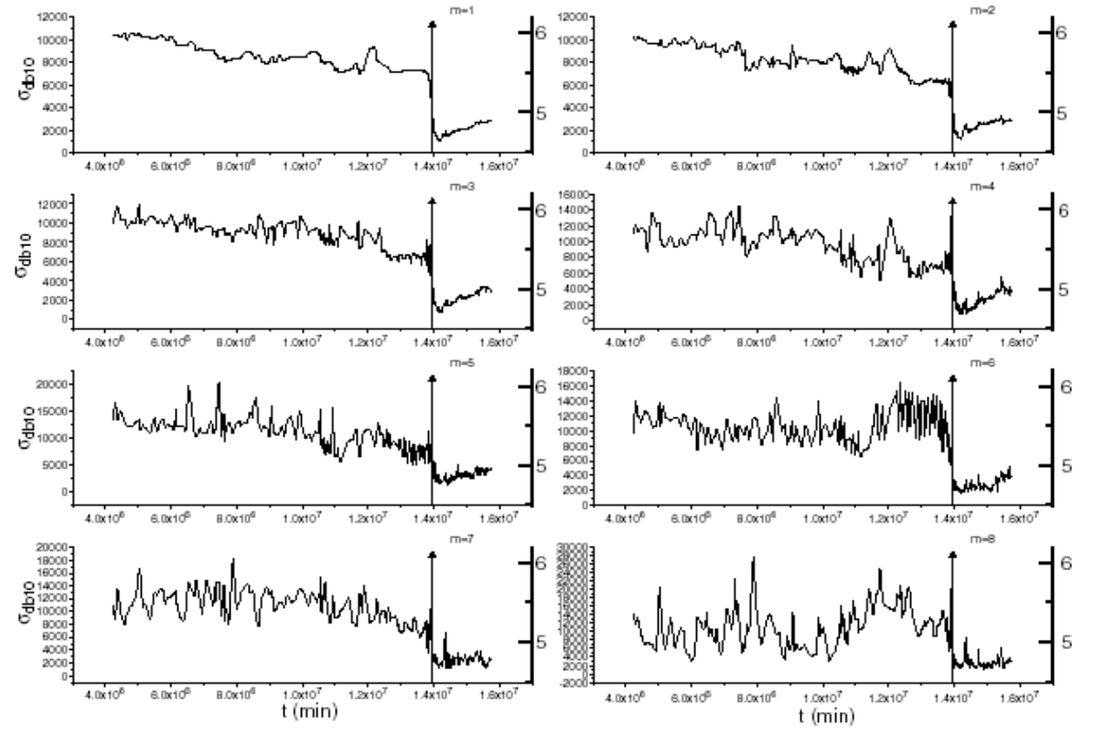


Figure 7.6. Time variation of $\sigma_{db10,a}$ with a ranging from 1 to 8. (Telesca et. al., 2007)

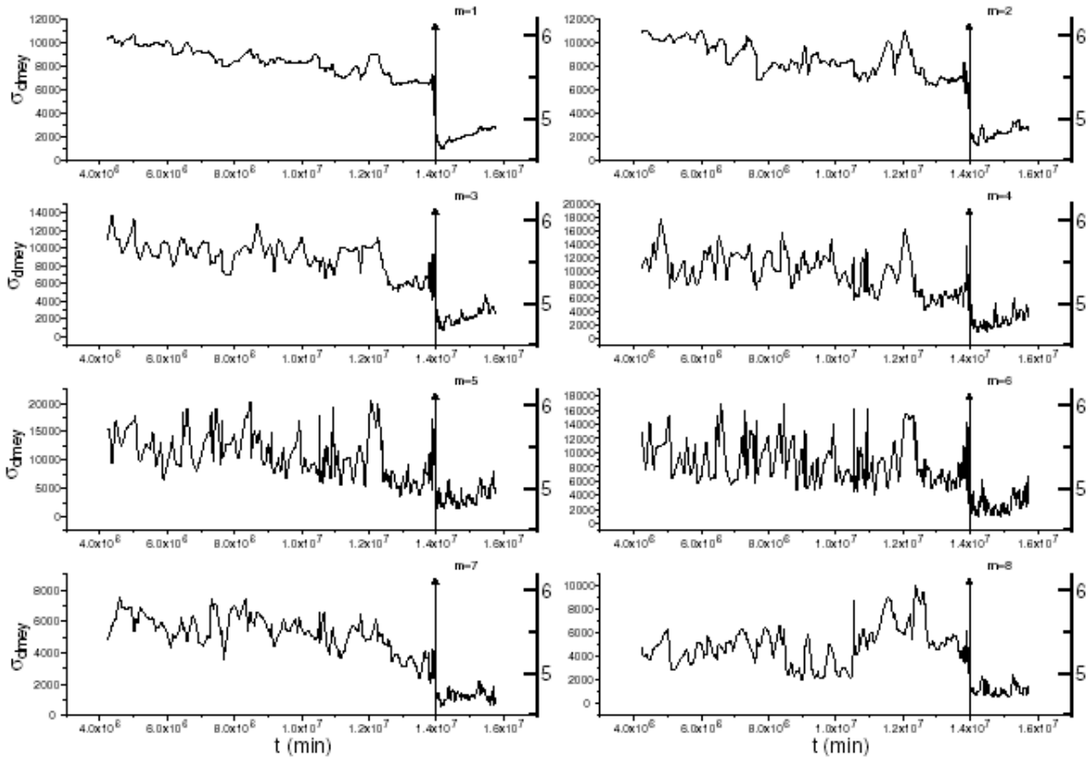


Figure 7.8. Time variation of $\sigma_{dmey,a}$ with a ranging from 1 to 8. (Telesca et. al., 2007)

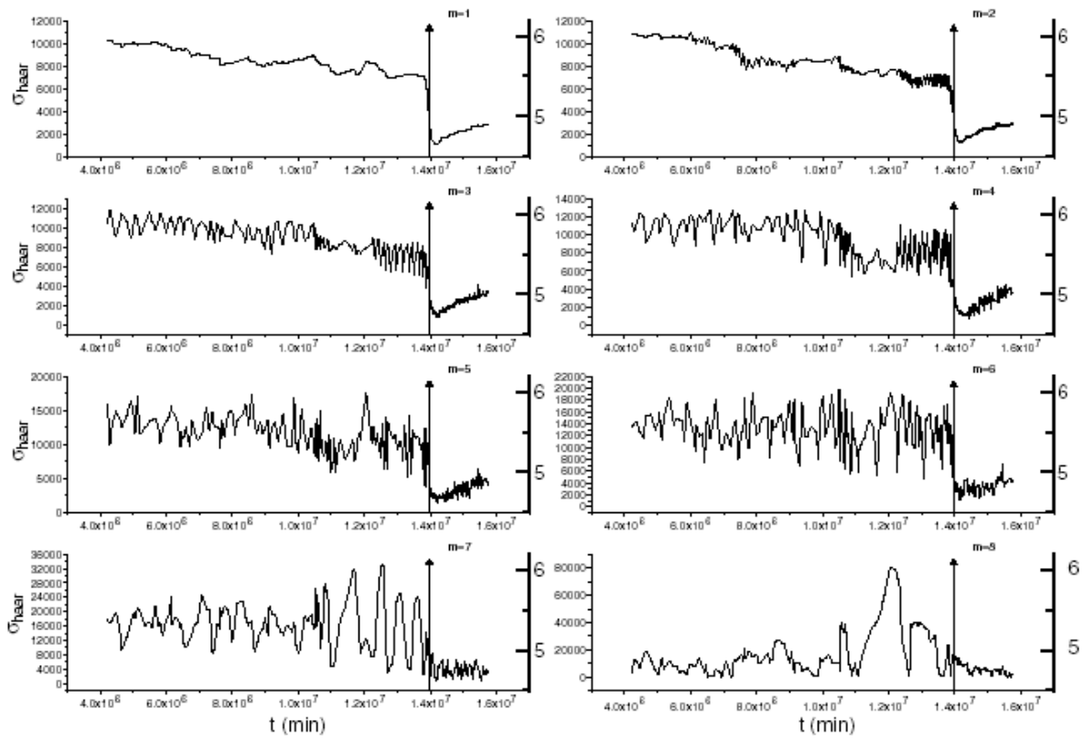


Figure 7.7. Time variation of $\sigma_{haar,a}$ with a ranging from 1 to 8. (Telesca et. al., 2007)

7.5 Evaluation to South Aegean seismicity after 2003

In this section we investigate if the results from the previous section hold, for the forthcoming seismic sequence (from 2003 up to July 2008) at the same region. A new interevent time series between two successive events versus the time occurrence of the second event, is generated. This is analyzed by means of the same wavelet bases used before in order to investigate the similarity of extracted results.

The data that used are derived combined from the GI-NOA earthquake catalogue (http://www.gein.noa.gr/services/Noa_cat/CAT2003.TXT) and HSNC's catalogue, for the period from Jan 2003 until July 2008. The epicentre distribution of these events is shown in Fig.7.9. The histogram of the number of events versus magnitudes shows a simple distribution (Fig.7.10). The maximum in number of earthquakes occurs in the magnitude range 3 to 3.5, which is a first indication of the value of M_c of the catalogue. A Gutenberg-Richter analysis is performed in order to estimate the minimum magnitude of completeness (M_c). As shown in Figure 7.11 the examined catalogue proofs complete for magnitudes higher than 3.2. The M_c value is lower than the respective value for the 1970 – 2003 catalogue. This can be explained by considering that the last years, the number of seismographs installed in the examined area increased, comparing to the number of seismographs were installed during the period 1970 – 2003. A denser seismological network has the ability to detect earthquakes with lower magnitude produced with this way, a catalogue with lower minimum value of completeness.

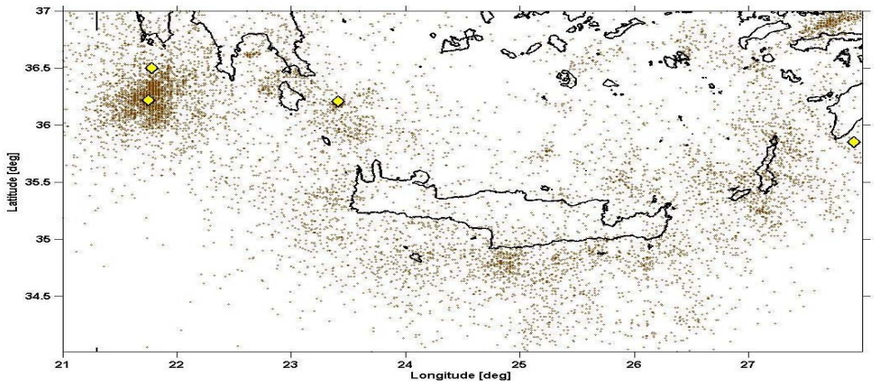


Figure 7.9 Epicentre distribution of earthquakes ($M \geq 3.2$) occurred in southern Aegean area over the period 2003–2008. Diamonds indicate the largest earthquakes of the sequence.

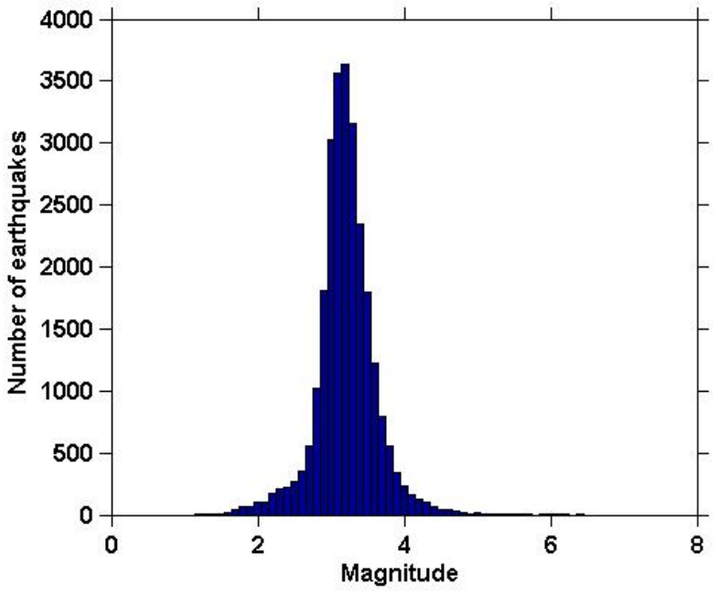


Figure 7.10 Histogram of the number of events versus the magnitude for the 2003-2008 seismic catalogue. Magnitude bins are equal to 0.1

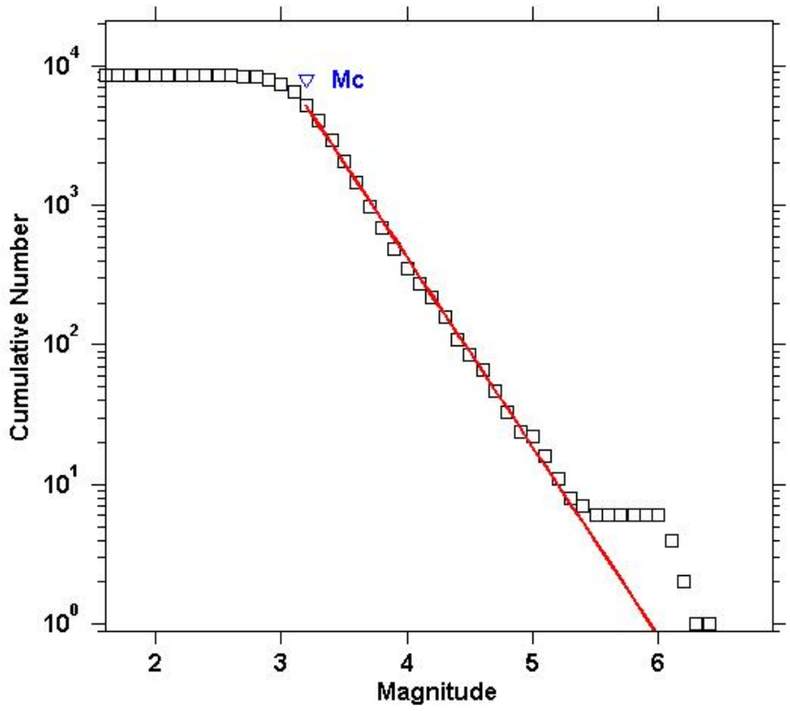


Figure 7.11 The frequency – magnitude distribution of earthquakes for 2003-2008 catalogue. Histogram of the number of events versus the magnitude for the 2003-2008 seismic catalogue.

Due to this fact, the number L of interevent intervals is also increased. The 2003-2008 sequence consists of L=3914 intervals as shown in Fig. 7.12 where the three largest earthquakes (8/1/2006 M=6.4, 14/2/2008 M=6.2, 14/2/2008 M=6.1 and 15/7/2008 M=6.4) indicated by red vertical lines. Instead of examine only the bigger earthquake as before, we select the three biggest. This is done in order to check the validity of the proposed method to larger data sets which may include more than one major event. This could be useful for the case of investigation of temporal patterns in declustered catalogues.

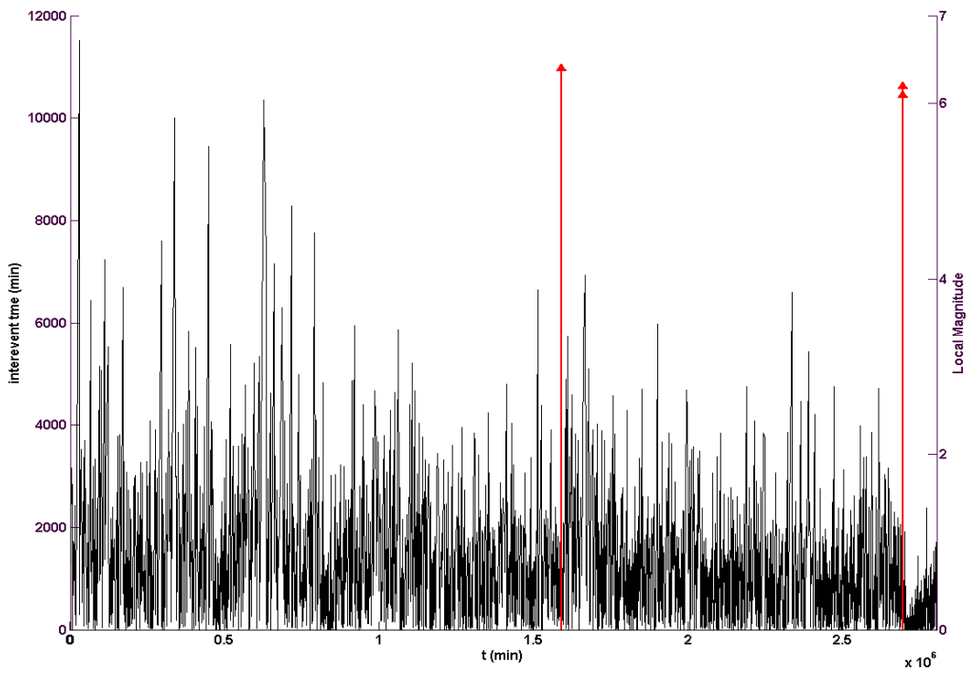


Figure 7.12. Interevent times between two successive events versus the time occurrence of the second event for the period Jan 2003 until May 2008. Red vertical arrow lines indicates occurrence time of strongest events (M > 6.1).

Following the guidelines derived from Fig.7.3 the results from first 8 scales will be presented. The wavelet analysis performed using the same wavelets and parameters as in previous section. The results are shown in Fig.7.13 through Fig.7.17

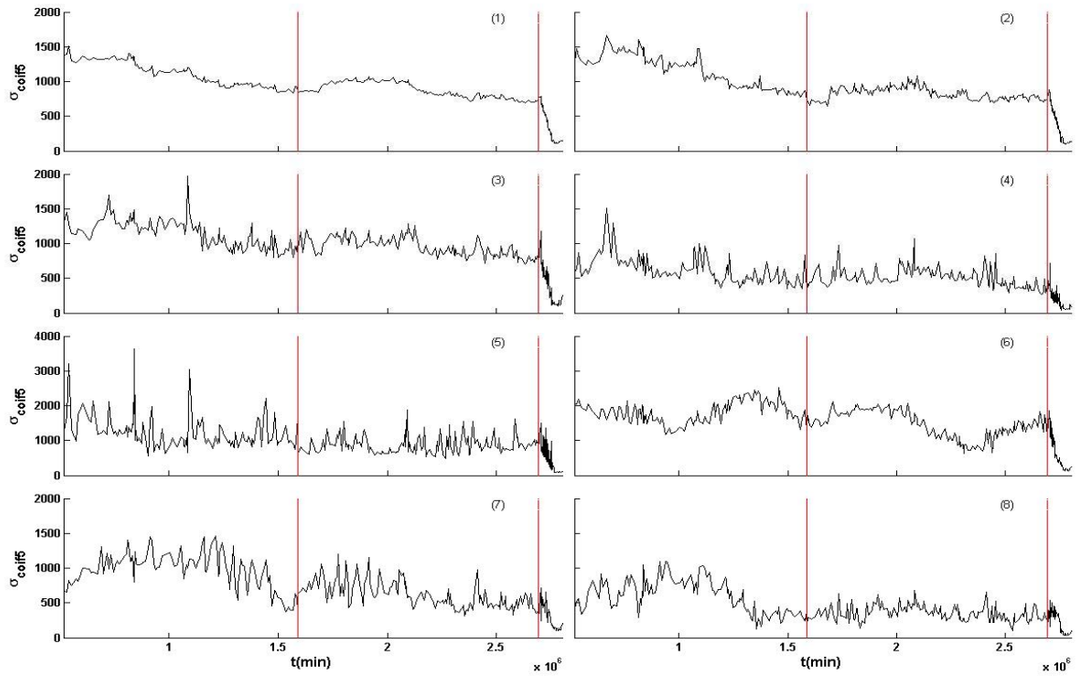
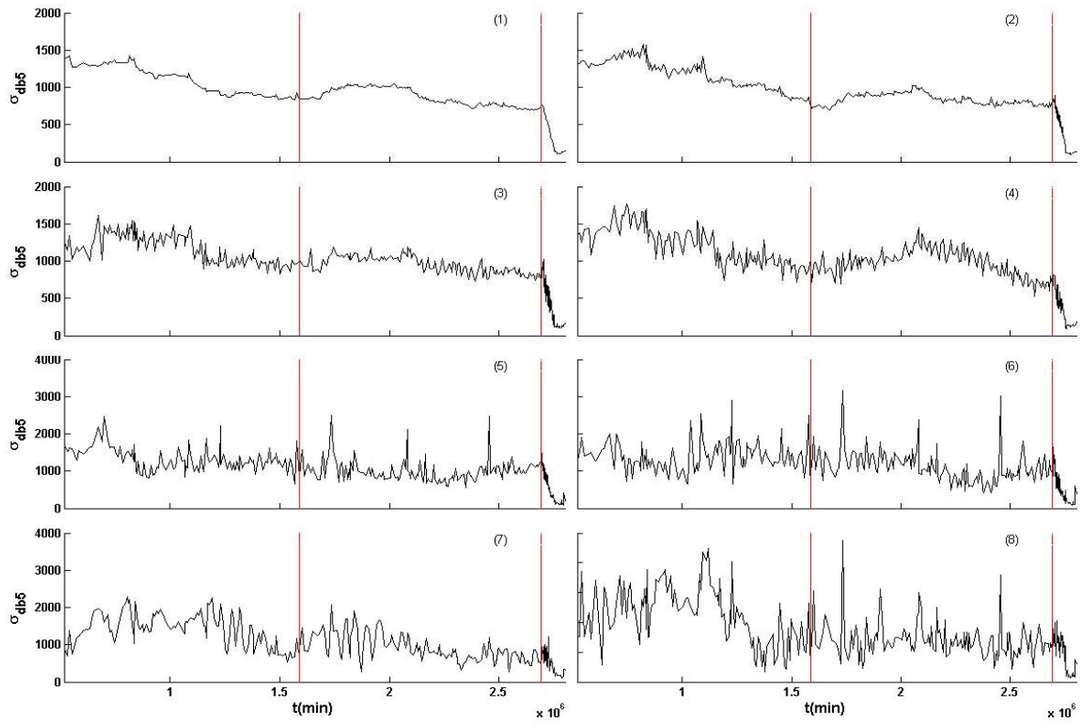


Figure 7.13. Time variation of $\sigma_{coif5,a}$ with a ranging from 1 to 8



F**igure 7.14**. Time variation of $\sigma_{db5,a}$ with a ranging from 1 to 8.

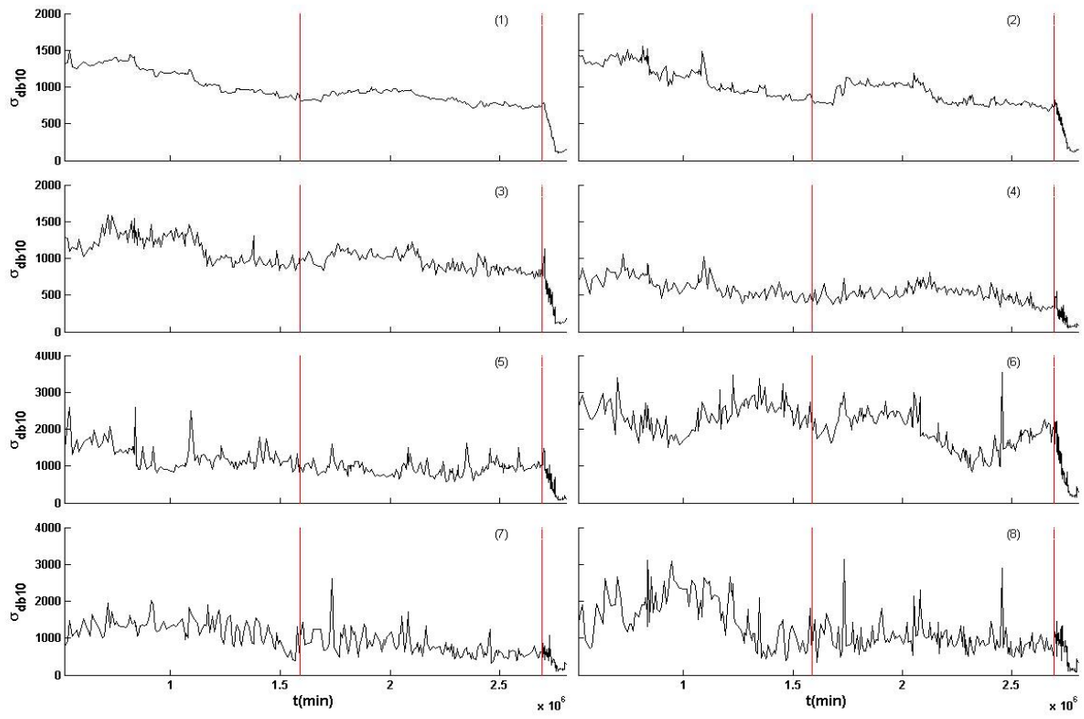


Figure 7.15. Time variation of $\sigma_{db10,a}$ with a ranging from 1 to 8

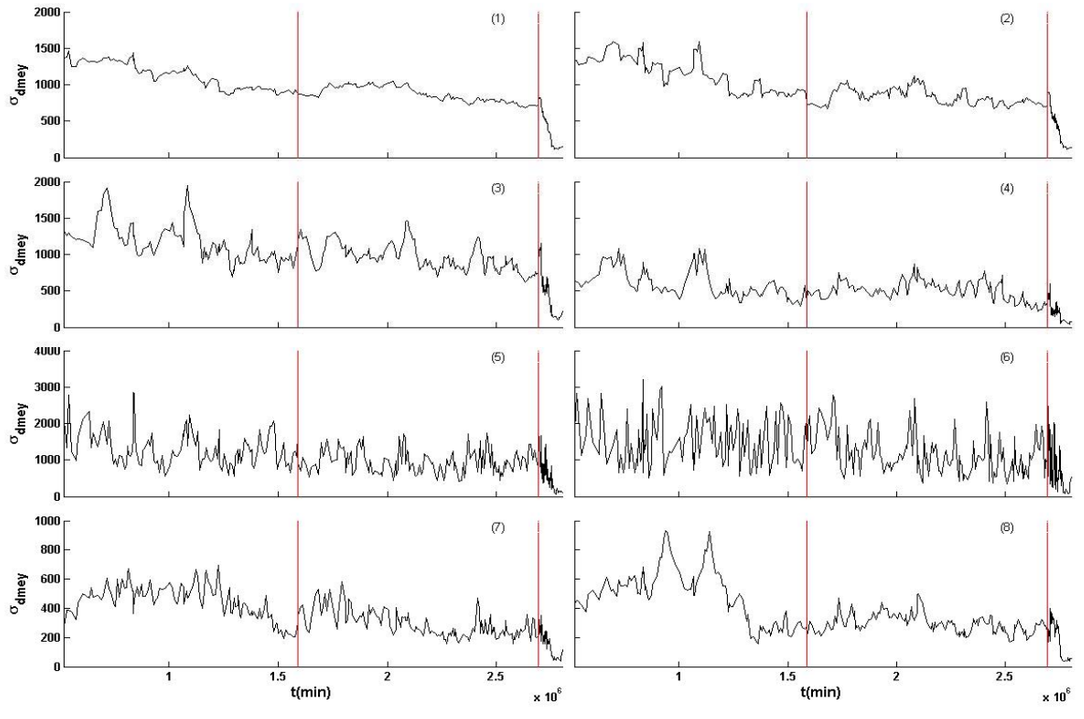


Figure 7.16. Time variation of $\sigma_{dmey,a}$ with a ranging from 1 to 8.

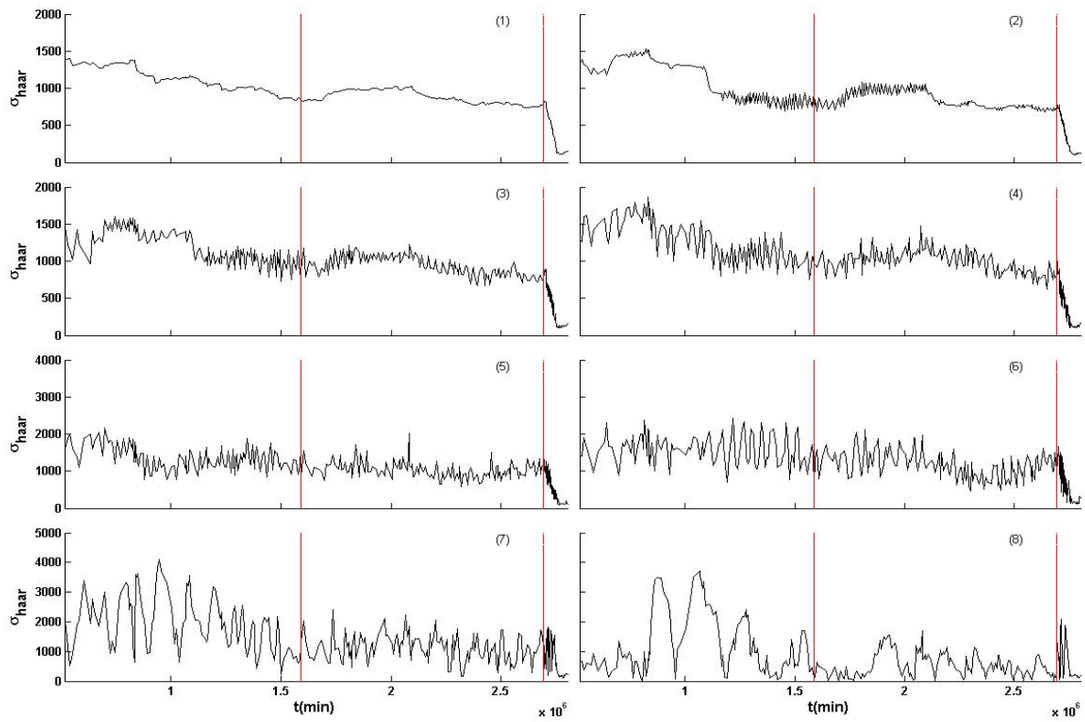


Figure 7.17. Time variation of $\sigma_{haar,a}$ with a ranging from 1 to 8.

The results derived from the previous catalogue (1970-2003) also hold here. More specific, for all wavelet bases, at scales from 1 to 6 (except *dmey* wavelet) one can associate again the occurrence of the major shock with a decrease in temporal evolution of the $\sigma_{wav,a}(t)$. The interesting point in the examined catalogue is that it contains a set of "twin" earthquakes (earthquakes that can be associated with the same epicentral area and share very close occurrence times). The major decrease of standard deviation of wavelet coefficients exist not at the time of the strongest earthquake but at the time where "twin" earthquakes exist. This leads to the possible conclusion that multiresolution wavelet analysis could provide an *aposteri* indicator of the maximum emitted earthquake energy which in some cases could be presented by a set of strong earthquakes. This result need further elaboration and is a part of a future research plan.

7.7 Summary

In this chapter, the time dynamics of seismicity by means of the temporal variation of the wavelet-coefficient standard deviation $\sigma_{wav,a}$ was examined. Initially the seismicity of southern Aegean area, during 1970-2003 as a point process, was analyzed. The strongest event in the sequence seems to discriminate two different temporal regimes, characterized by different dynamics displayed by the wavelet-coefficient standard deviation measure. The change from one regime to the other is very sharp at lower scales, corresponding to higher frequency fluctuations, due to background seismicity or the high number of aftershock events, while smoother at higher scales, corresponding to lower frequency fluctuations which correspond to long geodynamic processes. The same analysis provided with the seismicity from January 2003 until July 2008 which reveals the same results for background seismicity but the sharp change in lower scales occurred when the couple of the strongest earthquakes events appeared. Therefore, this method can possibly used to indicate the time occurrence of the maximum emitted earthquake energy.

8. Conclusions and Recommendations for Future Work

Conclusions

This research focused in the installation of a regional seismological network and the applications of wavelets in problems that arise in several stages. Apart from the engineering aspects involved for the design, installation, deployment, operation and maintenance of seismological network, there are also four original contributions to knowledge using wavelets approaches:

- Identification of disturbing transients in HVSR used for site effect characterization during design stage of seismological network.
- Investigation of best wavelet denoising approaches for seismological signals.
- Application of a wavelet based EEWS in South Aegean capable of producing reliable magnitude and epicentral estimates after 5 secs from the closest to the event station P-wave arrival.
- Characterization on the dynamics of South Aegean seismicity catalogue.

The South Aegean area and especially the area around Crete island is one of the most seismically active zones of the world and the most active in East Mediterranean. At the end of 2002 the limited number seismological stations that covered the area operated by two institutes: GI-NOA and GEOFON. Their combined network has an average interstation distance around 180km and only a limited number of stations was available for online data transmission. In addition, real time processing (event's source parameter estimation and dissemination of results) was done only at GEOFON's site, in Germany and only for stations that were capable to send their data in near real-time. These circumstances produced a fuzzy framework concerning the use of seismological network for local interest purposes capable of providing accurate in real time estimations. The design of SNC at LGS came to overcome, at least, the above limitations. The main goals that described during the design stage of SNC were:

- Dense network coverage
- Real time data transmission using all possible communication combinations
- Real time data processing at CRC
- Rapid dissemination of results in different forms

- Ability for easy upgrade and concurrent real time processing for the evaluation of new algorithms

The installation of SNC begun at 2003. Each seismological station comprised of a high performance DAS with 24bit resolution, capable of storing 40 days of data, providing real time data transmission and telemetry, consuming only 2.4W at full operation (with 3 channel active sensor, disk writing and continuous data transmission and telemetry).

For the communication to the CRC, a TCP/UDP based protocol, the RTP, is selected as proved to be the most appropriate for seismological network purposes. It provides the basic functionality as TCP/UDP but it is optimized for seismological data transmission aspects such as datagram implementation, expected delays from satellite links, recovery from loss of connection without data loss and relatively small implementation. All the data links that used compose a VPN fully isolated from non certified users but fully accessible using internet to registered users. The data links that used were wired and wireless. Wired links installed at locations that were closer to 2km from the local PTO's point of presence. This is dictated because the technology that used, ADSL, which in Crete Island performs satisfactory for locations up to 2km. The availability survey that performed for year 2007 shows that for 5 from 6 examined stations we have at least 92% availability. Wireless link using dedicated satellite connections used for stations where wired links could not provide the expected availability. The parameters of satellite link extracted using as test case the worst case scenario which is a future installation at western part of F2 beam of used satellite (Hellas-Sat II). The near 100% availability of these links verifies the validity of selected installation parameters of satellite link.

Data collection, event and monitoring, real time processing and dissemination of results based on commercial packages that modified to cooperate with additional software written by the author. More specific, data collection based on the implementation of RTP protocol and provided by the DAS manufacturer (Reftek inc.). It is modified by the author to adapt to the specific needs of SNC. A complementary (to the one provided by DAS manufacturer) web application is created for event and monitoring. This new application has the advantage that is adapted to the type of web browser that a certified user runs: other options are available for the certified user that uses a desktop or laptop computer to access the SNC while fewer options are available for a certified user that uses his cellular phone or palmtop through GPRS.

Real time processing is based on the SNDP software. Among other advantages (that described in chapter 3), this software has the ability to run processes that produce the same result (e.g. magnitude estimation) in parallel providing a unique way to test and evaluate different algorithms for the same problem and thus selecting the best one. The algorithms of real time estimation of magnitude and epicentral that discussed in Chapter 6 are implemented in this way. Finally, the dissemination of results required a new set of software modules, written from scratch, since there were no available at the SNDP package. These modules are responsible for the creation (after an event estimation) of appropriate messages (emails, notification to online bulletin, sms, visible and auditive alerts), the update of the corresponding database, the production of static as well as dynamic mapping images, the selection of appropriate data portions for further analysis and the update of the corresponding web page.

The second part of this thesis consist of the applications of wavelets for specific problems that arised during the operation of SNC. The first problem deals with the use of redundant WT for the identification of small amplitude transients on HVSR. Since HVSR method is used to estimate site's characteristics it is obvious that was used for seismological's stations location characterization. HVSR method assumes that the ambient noise that uses is stationary because it uses FT to calculate the spectra of horizontal and vertical component of ambient noise. Generally, it is advised to avoid long duration disturbances (e.g. near source man made or artificial noise) since it is expected to contaminate the spectral content of the estimated spectra. High amplitude disturbances usually can be detected by STA/LTA algorithm but when these disturbances have amplitude comparable to ambient noise it is very difficult to identify them (since STA/LTA is an amplitude comparison algorithm). The method that proposed is to project the components of ambient noise to wavelet domain where the frequency content and its time attributes (start – stop time) can easily be identified by an appropriate selection of scales. This process can run as a preliminary test to every HVSR survey in order to give the analyst the option to initially identified time windows that contain disturbances and then to exclude these portions of signal from further analysis. An algorithm implementing the above procedure and capable of handling the two common file formats used in HVRS surveys (SAF and ASCII) is also presented.

The next application that investigated is the denoising of seismological signals using wavelets. The use of WDM in seismology is shown in Chapter 5, were some

successful cases presented. The need for WDM arised from the fact that the commonly used filtering schemes in seismological packages introduced unacceptable distortions to P wave arrival leading to at least wrong epicentral estimations. Since today, there are many promising WDM. The contribution of this Chapter is to investigate the applicability of the latest WDM on the seismological signals. More specific, a comparison study, for 20 WDM using 1250 seismograms with different SNR is carried out. The result is that every WDM outperforms conventional filtering but there is no unique WDM that can perform satisfactory for all the cases.

For real time data processing a potential for EEWS based on wavelets, is proposed. Rapid (4-6 secs after event occurs) determinations are the keystone of a successful and reliable EEWS. The use of WT for this application is presented only in a couple of studies and for magnitude determination only. In chapter 6, a wavelet based EEWS for South Aegean is introduced. The proposed method consists of combining wavelet based methods for magnitude and epicentral estimation. More specific, the wavelet magnitude estimation by Simmons (2006) is modified and used for South Aegean in order to derive a correlation between the amplitude of wavelet coefficients and local magnitude. Such a correlation is observed in last two WT scales and the appropriate magnitude prediction equations are estimated. For epicentral estimation, the proposed method based on the composition of two different approaches: the epicentral estimation using two station subbaray (TSS by Rydelek & Pujol, 2006) and wavelet azimuth estimation (WAE by Galliana et al., 2004). By combining these two methods, as never used before, a epicentral estimation with two triggering stations is available with a 5.9^{o} uncertainty. If the event happens outside the network, the proposed method has the advantage that uncertainty remains at 5.9^{o} while in the TSS method the uncertainty increased dramatically. WME and WHE are capable to produce reliable results after 5secs which means that are suitable for implementation in an EEWS at South Aegean area

The last application is the use of WT in multiresolution wavelet analysis of seismicity. Since the beginning of SNC around 5500 events were recorded and identified. This number provides a satisfactory material for the investigation of time dynamics of seismicity. The measure that used is the temporal variation of the wavelet-coefficient standard deviation $\sigma_{wav,m}$. Initially the seismicity of southern Aegean area, during 1970-2003 as a point process, was analyzed. The strongest event in the sequence seems to discriminate two different temporal regimes, characterized

by different dynamics displayed by the wavelet-coefficient standard deviation measure. The change from one regime to the other is very sharp at lower scales, corresponding to higher frequency fluctuations, while smoother at higher scales, corresponding to lower frequency fluctuations. The same analysis provided with the seismicity from January 2003 until July 2008 which reveals the same results for background seismicity but the sharp change in lower scales occurred when the couple of the strongest earthquakes events appeared. Therefore, this method can possibly used to indicate the time occurrence of the maximum emitted earthquake energy.

Recommendations for Further Work

Reviewing the research described in this thesis a number of recommendations can be made regarding the further development of the proposed methods as well as a suggestion for new research programmes that can now be accomplished thanks to the achievement of reliable results from the application of WT to seismological signals:

- Expansion of the seismological network. Even with the present topology the interstation distance in South Aegean combining all operating networks is not comparable with networks located in other seismogenic zones (e.g. California with 35 km interstation distance, South Italy with 40 km interstation distance). Additional stations installations are inevitable if we want to test and develop new methodologies for sophisticated EEWS. LGS planning 8 new, satellite linked installations until 2010.
- Implementation of selection algorithm for WDM. As shown in Chapter 5 there is no unique WDM that fits to all the data. Thus an algorithm capable of identifying the quality of incoming data (e.g. SNR, transient detection) could be useful for the selection of appropriate WDM. The future involvement of Artificial Neural Networks in this procedure, is promising
- Evaluation of WME. The continuous evaluation of WME is mandatory in order to derive the best predicted magnitude equations. For this reason, the algorithm of recalculation of best fit equations after every final magnitude estimation (by human analyst) must be implemented in the real time system. Under this approach, after a human analyst estimates the final local magnitude, this value will inserted to the database and then the amplitude of wavelet coefficients will be estimated. A new best fit

regression will be calculated and the new equations will be used for the magnitude prediction of the next event. The future expansion of HSNC requires a distance correction factor to the expressions of WME.

- Investigation of locally applicability of WME. Since the WME produced by averages over all stations, it will be interesting to investigate if the same relations hold for local (one station) estimations. If this will verified then the process of WME could be on site and only the estimated magnitude will be sent to CRC. This will decrease the time needed for producing an alert since it is easier to transfer an alert through a link rather than the data needed to estimate at CRC this alert.

References

Abramovich, F., Bailey, T.C. & Sapatinas, T. (2002). Wavelet analysis and its statistical applications. The Statistician, 49, 1–29.

Abramovich, F. & Benjamini, Y. (1995). Thresholding of wavelet coefficients as multiple hypotheses testing procedure. In Wavelets and Statistics, Antoniadis, A. & Oppenheim, G. (Eds.), Lecture Notes in Statistics 103, pp. 5–14, New York: Springer-Verlag.

Abramovich, F. & Benjamini, Y. (1996). Adaptive thresholding of wavelet coefficients. Comput. Statist. Data Anal., 22, 351–361.

Abramovich, F., Sapatinas, T. & Silverman, B.W. (1998).Wavelet thresholding via a Bayesian approach. J. R. Statist. Soc. B, 60, 725–749.

Abramovich, F., Benjamini, Y., Donoho, D.L. & Johnstone, I.M. (2000). Adapting to unknown sparsity by controlling the false discovery rate. Technical Report 00-19, Department of Statistics, Stanford University, USA

Abercrombie, R., (2005), The start of something big?, Nature 438, 171–173.

Akay M, (editor), (1997), Time frequency and wavelets in biomedical signal processing. Piscataway (NJ): IEEE Press.

Allen, R. M., and H. Kanamori (2003). The potential for earthquake early warning in southern California, Science 300, 786–789.

Allen, R. M. (2004). Rapid magnitude determination for earthquake early warning, in The Many Facets of Seismic Risk, G. Manfredi, M. R. Pecce, and A. Zollo (Editors), Proceedings of the Workshop on Multidisciplinary Approach to Seismic Risk Problems, Sant'Angelo dei Lombardi, 22 September 2003, Universita degli Studi di Napoli "Federico II", Napoli, Italy, 15–24.

Albroubi A, Unser M, (editors) (1996), Wavelet in medicine and biology. Boca Raton (FL): CRC Press.

Ansary, M. A., Yamazaki, F., and Katayama, T. (1995). Statistical analysis of peaks and directivity of earthquake ground motion, International Journal of Earthquake Engineering and Structural Dynamics, Vol. 24 (11), pp. 1527-1539

Antoniadis, A. & Fan, J. (2001) Regularization of wavelets approximations (with discussion). J. Am. Statist. Ass., 96

Antoniadis , A., (2000), Wavelet Estimators in Nonparametric Regression, Technical report, IMAG-LMC, Grenoble, France

Argoul, F., Arneodo, A., Grasseau, G., Gagne, Y., Hopfinger, E.J., Frisch, U.: (1989), Nature, 338, 51.

Arneodo, A., G. Grasseau, and M. Holschneider, (1988) Wavelet transform of multifractals, Phys. Rev. Lett., 61, 2281–2284.

Arneodo, A., Bacry, E., Muzy, J.F.: (1995), Physica A, 213, 231-75.

Arneodo, A., d'Aubenton-Carafa, Y., Bacry, E., Graves, P.V., Muzy, J.F., Thermes, C.: (1996), Physica D, 96, 291.

Arneodo, A., Manneville, S., Muzy, J.F.(1998), Eur. Phys. J. B, 1, 129.

Ashkenazy Y, Lewkowicz M, Levitan J, Mψlgaard H, Thomsen PEB, Srmark K.(1998), Discrimination of the healthy and sick cardiacautonomic nervous system by a new wavelet analysis of heartbeat intervals. Fractals;6:197–203.

Athichanagorn, S., (1999), Development of an interpretation methodology for long-term pressure data from permanent downhole gauges, PhD Thesis, Stanford University

Athanassopoulos, G.A., (2000), Site Effects on the Seismic Ground Response of the City of Chania, Greece, Technical Report, Deliverable D2, research project SEISMOCARE.

Bard, P.Y., (1999) Microtremor measurement: a tool for site effect estimation. in The Effects of Surface Geology on Seismic Motion. Irikura K, Kudo K, Okada H and Sasatami T (Editors), Balkema, Rotterdam,1251–1279.

Bardet J.P, Ichii, K. and Lin, C., (2000) EERA: a computer program for equivalent-linear earthquake site response analyses of layered soil deposits. University of Southern California, 40 pp.

Borcherdt, R. and Glassmoyer, G., (1992), On the characteristics of local geology and their influence on ground motions generated by the Loma Prieta earthquake in the San Francisco Bay region, California, Bulletin of the Seismological Society of America; v. 82; no. 2; p. 603-641

Baumjohann, W., Treumann, R.A., Georgescu, E., Haerendel, G., Fornacon, K.H.,Auster, U.,(1999), Ann. Geophys., 17, 1528.

Bendjoya, Ph., Slezak, E., Froeschle: (1991), Astron. Astrophys., 251, 312.

Benetatos, C., Kiratzi, A., Papazachos, K., Karakaisis, G., (2004), "Focal mechanisms of shallow and intermediate depth earthquakes along the Hellenic Arc", Journal of Geodynamics, vol 37, issue 2, pp.253-296

Benzi, R., Biferale, L., Crisanti, A., Paladin, G., Vergassola, M., Vulpiani, A.: (1993), Physica D, 65, 352.

Beylkin G., (1992), On the Representation of Operators in Bases of Compactly Supported Wavelets. SIAM Journal of Numerical analysis, 29, 1716-40

Bijaoui, A., Slezak, E., Mars, G.: (1993), Wavelets, Fractals and Fourier Transforms, Oxford University Press, Oxford., p. 213.

Breiman, L., Peters, S. (2000). Comparing automatic smoothers (a public service enterprise). Int. Statist. Rev., 60, 271–290.

Breiman L (1995). Better subset regression using the nonnegative garrotte. Technometrics 37, 373-38

Bruce A.G. and Gao H.Y (1996). Applied Wavelet Analysis with S-Plus. New York, Springer

Bouchon, M., Barker, J., (1996), Seismic response of a hill: The example of Tarzana, California, Bulletin of the Seismological Society of America; February 1996; v. 86; no. 1A; p. 66-72

Cai, T.T. (1999). Adaptive wavelet estimation: a block thresholding and oracle inequality approach. Ann. Statist., 27, 898–924.

Cai,T.T. & Silverman, B.W. (2001). Incorporating information on neighboring coefficients into wavelet estimation. Sankhy¯ a, Series A, 63

Carniel, R., Malisan, P., Barazza, F., Grimaz S., (2008), Improvement of HVSR by wavelet analysis, Soil Dynamics and Earthquake Engineering, Vol. 28, Issue 4, Pages 321-327

Chipman, H.A., Kolaczyk, E.D. & McCulloch, R.E. (1997). Adaptive Bayesian Wavelet Shrinkage. J. Am. Statist. Ass., 92, 1413–1421.

Chui, C. K., (1992a), Wavelet Analysis and Its Applications, vol. 1, An Introduction to Wavelets, Academic, San Diego, Calif..

Chui, C. K., (1992b), Wavelet Analysis and Its Applications, vol. 2, Wavelets— A Tutorial in Theory and Applications, Academic, San Diego, Calif.

Clyde, M. & George, E.I. (2000). Flexible empirical Bayes estimation for wavelets. J. R. Statist. Soc. B, 62, 681–698.

Cohen, A., Kovacevic, J., (1996), "Wavelets: the mathematical background," Proc. IEEE, 84:514–522

Coifman R. and Donoho D.L. (1995). Translation Invariant De-Noising. In Wavelets and Statistics (Lecture Notes in Statistics, Vol 103) edited by Antoniadis A and Openheim G, New York, Springer-Verlag,125-50

Coifman R and Wickerhauser MV (1992). Entropy based algorithms for best basis selection. IEEE Trans. on Inf. Theory, 32, 712-718

Cornish, C., (2004), Maximal overlap wavelet statistical analysis with applications to atmospheric turbulence. Boundary-Layer Meteorology, v119 i2. 339-374.

Cooper, J.,D., (1868) Letter to Editor. San Francisco Daily Evening Bulletin, Nov. 3, 1868

Donoho, D. L. & Johnstone, I. M. (1994) Ideal spatial adaptation by wavelet shrinkage, Biometrika, 81, pp. 425± 455.

Donoho, D. L. & Johnstone, I. M., Krkyacharian, G. & Picard, D., (1995) Wavelet shrinkage -asymptopia, Journal of the Royal Statistical Society, Series B, 57, pp.301-337.

Daubechies, I., (1988), Orthonormal bases of compactly supported wavelets, Commun. Pure Appl. Math., XLI, 909–996.

Daubechies I., (1990), The wavelet transform: time-frequency localization and signal analysis. IEEE Transactions on Information Theory, vol. 36, no. 5, pp. 961-1005.

Daubechies, I., (1992), Ten Lectures on Wavelets, 357 pp., Soc. for Ind. and Appl. Math., Philadelphia, Pa.

Donoho, D.L. & Johnstone, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81, 425–455.

Donoho, D.L. & Johnstone, I.M. (1995). Adapting to unknown smoothness via wavelet shrinkage. J. Am. Statist. Ass., 90, 1200–1224.

Donoho, D.L. & Johnstone, I.M. (1998). Minimax estimation via wavelet shrinkage. Ann. Statist., 26, 879–921.

Donoho, D.L., Johnstone, I.M., Kerkyacharian, G. & Picard, D. (1995). Wavelet shrinkage: asymptopia? (with discussion). J. R. Statist. Soc. B, 57, 301–337.

Douglas, A., (1997), Bandpass filtering to reduce noise on seismograms: Is there a better way ?, Bull. Seism. Soc. 87, 770-777

Duffin, R. J. and Schaeffer, A. C., (1952)A Class of Nonharmonic Fourier Series, Transactions of the American Mathematical Society,72, 341-366.

Efron, B., Tibshirani, R., (1993), an introduction to the bootstrap, Chapman & Hall / CRC, Florida, US

EMERIC I – CRINNO Project (2006). http://www.crete-region.gr/greek/programs /CRINNO/

Espinosa-Aranda J, Jiménez A, Ibarrola G, Alcantar F, Aguilar A, Inostroza M, Maldonado S (1995) Mexico City seismic alert system. Seism Res Lett 66: 42-53

Faccioli E., (1991), Seismic amplification in the presence of geological and topographic irregularities. In: Proceedings of the second international conference on recent advances in geotechnical earthquake engineering and soil dynamics. St Louis, Missouri, vol. II, p. 1779–97.

Field, E., Jacob, K., (1993), The theoretical response of sedimentary layers to ambient seismic noise, Geophys. Res Lett., vol. 20, n^{o} 24, pp. 2925-2928

Fligge, M., Solanki, S.K.: (1997), Astron. Astroph. Supl., 124, 579.

Fligge, M., Solank, S. K., Beer, J., (1999). Determination of solar cycle length variations using the continuous wavelet transform. Astronomy and Astrophysics 346, 313–321.

Frick, P., Galyagin, D., Hoyt, D.V., Nesme-Ribes, E., Schatten, K.H., Sokoloff, D.,Zakharov, V.: (1997), Astron. Astroph., 328, 670.

Foster, G.: (1996), Astron. J., 112 (4), 1709.

Gabor, D., (1946), Theory of Communications, Journal of the Institute of Electrical Engineering, London III, 93, 429-457.

Galiana-Merino, J. J., J. Rosa-Herranz, J. Giner, S. Molina, and F. Botella (2003). De-noising of short period seismograms by wavelet packet transform, Bull. Seism. Soc. Am. 93, no. 6, 2554–2562.

Galiana-Merino, J. J., J. Rosa-Herranz, J. Giner, S. Molina, and F. Botella (2004). Regularized Deconvolution of Local Short-Period Seismograms in the Wavelet Packet Domain ,Bull. Seism. Soc. Am. 94, no. 4, 1467–1475.

Galiana-Merino, J. J., J. Rosa-Herranz, J. Giner, S. Molina, and F. Botella (2007), Wavelet Transform Methods for Azimuth Estimation in Local Three-Component Seismograms, Bull. Seism. Soc. Am. 97, no. 3, 793 -805

Girardi, M., Fadda, D., Escalera, E.; Giuricin, G., Mardirossian, F.; Mezzetti, M.: (1997), Astroph. J., 490, 56.

Gitterman, Y., Zaslavsky,Y., Shapira, A., Shtivelman, (1996) Empirical site response evaluations : case studies in Israel, Soil Dynamics and Earthquake Engineering Vol. 15, Issue 7, pp 447-463

Godano C, Caruso V. (1995),Multifractal analysis of earthquake catalogues. Geophys J Int; 121:385–92.

Goupillaud, P., A. Grossmann, and Morlet, J., (1984), Cycle-octaves and related transforms

in seismic signal analysis, Geoexploration, 23, 85–102.

Graps, A., (1995), "An introduction to wavelets," IEEE Computational Science and Engineering, 2(2).

Grossman, A., Morlet, J., (1984). Decomposition of Hardy function into square integrable wavelets of constant shape. SIAM J. Appl. Math., pp. 723-773

Gutenberg B, Richter C.(1956), Earthquake magnitude, intensity, energy, and acceleration. Bull Seism Soc Am; 46:105–45.

Haar, A., (1910), Zur Theorie der Orthoganalen Funktionen-Systeme, Annals of Mathematics, 69, 331-371.

Hartzell, S.,et. al., (1994) Variability in nonlinear sediment response during the 1994 Northridge, California, earthquake, Bull. Seism. Soc. Am.,v.88; no. 6; p. 1426-1437

Haynes, P.H., Norton, W.A.: (1993), Wavelets, Fractals and Fourier Transforms, Oxford University Press, Oxford, p. 229.

Heaton, T.,H., (1985), A model for a seismic computerized alert network. Science 228, pp. 987-990

Heisenberg, W., (1929), "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik", Zeitschrift für Physik, 43 1927, pp. 172-198.

Holschneider, M., (1995), Wavelets: An Analysis Tool, 423 pp., Oxford Sci., New York.

Horike, M., (1996), Geophysical exploration using microtremor measurements, Xth World conference earthquake engineering, Acapulco, 2033, Elsevier Science Ltd

Horike M., Zhao B and Kawase H (2001). Comparison of site response characteristics inferred from microtremors and earthquake shear waves. Bull. Seism. Soc. Am. 81, 1526–1536.

Horiuchi S, Negishi H, Abe K, Kamimura A, Fujinawa Y (2005) An Automatic Processing System for Broadcasting Earthquake Alarms. Bull. Seism. Soc. Am., no.95, p. 708-718

Hough, S., E., Seeber, L., Lerner-Lam, A., Armbruster, J., C., Guo, H., (1990), Empirical Green's function analysis of Loma Prieta aftershocks Bull. Seism. Soc. Am., 81: 1737 - 1753

Hough, S.E., Seeber, L., Rovelli, M., Malagnini, L., (1992), Ambient noise and weak-motion excitation of sediment resonances: Results from the Tiber Valley, Italy, Bulletin of the Seismological Society of America; v. 82; no. 3; p. 1186-1205

Huang, S.Y. & Lu, H.H.-S. (2000). Bayesian wavelet shrinkage for nonparametric mixedeffects models. Statist. Sinica, 10, 1021–1040.

Kanamori, H. (2005). Real-time seismology and earthquake damage mitigation, Annu. Rev. Earth Planet. Sci. 33, 5.1–5.20, doi 10.1145/ annurev.earth.33.092203.122626.

Kareem, A., Gurley, K. and Kantor, J.C. (1993), "Time-Scale Analysis of Nonstationary ProcessesUtilizing Wavelet Transforms", Proceedings of the 6th International Conference on Structural Safety and Reliability, Innsbruck, Austria, Balkema Publishers, Amsterdam, Netherlands.

Kiratzi, A., and Louvari E., (2003), Focal mechanisms of shallow earthquakes in the Aegean

sea and the surrounding lands determined by waveform modelling: a new database, Journal of Geodynamics, 36, 251-274.

Kiratzi, A.A., Papazachos, C.B., (1995), Active deformation of the shallow part of the subducting lithospheric slab in the southern Aegean, Journal of Geodynamics 19 (1), pp. 65-78

Klein, F. W. (2002). User's guide to HYPO-2000, a Fortran program to solve for earthquake locations and magnitudes, U.S. Geol. Surv. Open-File Rep. 02-171,123

Komm, R.W., Gu, Y., Hill, F., Stark, P.B., Fodor, I.K.: 1999, Astroph. J., 519, 407.

Konno, K., Ohmachi, T., (1998), Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor, Bull. Seis. Soc. Am., v. 88; no. 1, pp. 228-241

Knapmeyer, M., Harjes, H.-P. (2000), Imaging crustal discontinuities and the downgoing slab beneath western Crete, Geophysical Journal International 143 (1), pp. 1-21

Knapmeyer M. (1999), Geometry of the Aegean Benioff zones. Annali Geofis.;42:27–38.

Lachet, C., Bard, y., (1994), Numericl and theoretical investigations on the possibilities of Nakamura technique, J. Phys. Earth, 42,377-397

Le Pichon, X., Angelier, J., Aubouin, J., Lyberis, N., Monti, S., Renard, V., Got, H., Chronis, G., (1975), From subduction to transform motion: a seabeam survey of the Hellenic trench system, Earth and Planetary Science Letters 44 (3), pp. 441-450

Le Pichon X, Chamot-Rooke N, Lallemant S, Noomen R, Veis G.(1995), Geodetic determination of the kinematics of central Greece with respect to Europe: implications for eastern Mediterranean tectonics. J Geophys Res;100:12675–90.

Lay, T., Wallace, T.C., 1995. Modern Global Seismology. Academic Press, San Diego, CA, 521pp.

Lermo, J., Francisco, S., Chavez-Garcia, J., (1994), Are microtremors useful in site response evaluation ?, Bull. Seism. Soc. Am. 84, no. 5, 1350–1364.

Lima Neto, G. B., Pislar, V., Durret, F., Gerbal, D.; Slezak, E.: (1997), Astron. Astroph., 327, 81-89.

Lockman, A., and R. M. Allen (2007). Magnitude-period scaling relations for Japan and the Pacific Northwest: Implications for earthquake early warning, Bull. Seismol. Soc. Am. 97, no. 1, 140–150, doi 10.1785/0120040091.

Lockman, A. B., and R. M. Allen (2005). Single-station earthquake characterization for early warning, Bull. Seism. Soc. Am. 95, 2029–2039.

Lomax, A., (2005) A Reanalysis of the Hypocentral Location and Related Observations for the Great 1906 California Earthquake. Bull Seism Soc Am 95:861-877

Lucek, E.A., Balogh, A.: (1997), Geophys. Res. Lett., 24, 18, 2387.

Magotra, N., N. Ahmed, and E. Chael (1989). Single-station seismic event detection and location, IEEE Trans. Geosci. Remote Sensing 27, 15–23.

Magri, L., Mucciarelli, M., Albarello, D., (1994), Estimates of site seismicity rates using ill-defined macroseismic data, Pure and Applied Geophysics PAGEOPH 143 (4), pp. 617-

632

Mallat, S., (1989a). A theory for multiresolution decomposition -the wavelet representation, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. II, No. 7, pp. 674-692

Mallat, S., (1989b) Multifrequency channel decomposition of images and wavelet models, IEEE Trans. Acoust. Speech Signal Anal., 37(12), 2091–2110.

Mallat, S., (1989c), Multiresolution Approximations and Wavelet Orthonormal Bases of L2(R), Transactions of the American Mathematical Society, 315, No. 1, 69-87.

Mallat S.and Hwang L.(1992). Singularity detection and processing with wavelets. IEEE Trans. Information Theory, vol. 38, no. 2, pp. 617-643

Mallat, S.,(1998), A Wavelet Tour of Signal Processing, Academic Press.

Mastrolorentzo, G., (2004), Profili Di Velocita' Vs Nell'abitato Di Chania'(Creta), Msc Thesis, Universita degli Studi della Basilicata.

Meulenkamp, J.E., Wortel, M.J.R., van Wamel, W.A., Spakman, W., Strating, E.(1988), On the Hellenic subduction zone and the geodynamic evolution of Crete since the late Middle Miocene, Tectonophysics 146 (1-4), pp. 203-215

McClusky S. Global positioning system constraints on plate kinematics and dynamics in the eastern Mediterranean and Caucasus. J. Geophys Res 2000;105:5695–720.

McGuire, J., Simons, F., J., Collins, J., A., (2008), Analysis of seafloor seismograms of the 2003 Tokachi-Oki earthquake sequence for earthquake early warning, Geoph. Res. Lett., vol 35, L14310

McKenzie D.(1972),Active tectonics in the Mediterranean region. Geoph.Res.Astr;30:109–85.

McKenzie D. (1978). "Active tectonics of the Alpine-Himalayan belt: the Aegean Sea and surrounding regions" Geophys. J. R. Astr. Soc., vol. 55, pp. 217-254.

McNamara, D.E., Buland, R.P., (2004), Ambient noise levels in the continental United States, Bull. Seism. Soc. Am., 94, (4), pp. 1517-1527

Meyer, Y., (1992), Wavelets and Operators, Cambridge Univ. Press,New York.

Meyer, Y., (1993), Wavelets: Algorithms and Applications, SIAM.

Morlet, J., G. Arens, E. Fourgeau, and D. Giard, (1982a), Wave propogation and sampling theory, 1, Complex signal and scattering in multilayered media, Geophysics, 47(2), 203–221.

Morlet, J., G. Arens, E. Fourgeau, and D. Giard, (1982b), Wave propagation and sampling theory, 2, Sampling theory and complex waves, Geophysics, 47(2), 222– 236.

Morretin, P., (1997), Wavelets in statistics, 3rd International Conference on Statistical Data Analysis and Related Methods, Neuchatel, Switzerland

Mouri, H., Takaoka, M., Kubotani, H.: (1999), Phys. Lett. A, 461, 82.

Mucciarelli, M., and Monachesi G., (1997), A quick survey of local amplifications and their correlation with damage observed during the Umbro–Marchesan (Italy) earthquake of

September 26, 1997. J Earthquake Eng. .2, 2, pp. 325–337.

Mucciarelli M (1998). Reliability and applicability of Nakamura's technique using microtremors: an experimental approach, J. Earthquake Eng., 2, 625–638.

Mucciarelli M, Gallipoli MR and Arcieri M (2003). The stability of the horizontal-tovertical spectral ratio of triggered noise and earthquake recordings. Bull. Seism. Soc. Am., 93, 1407–1412.

Muhlmann, W., Hanslmeier, A.: (1996), Solar Phys., 166, 445.

Muzy, J.F., Bacry, E., Arneodo, A.: (1991), Phys. Rev. Lett., 67, 25, 3515.

Muzy, J.F., Bacry, E., Arneodo, A.: (1993), Phys. Rev. E, 47, 2, 875.

Nakamura Y. (1989). A method for dynamic characteristics estimations of subsurface using microtremors on the ground surface, Q. Rept. Railway Technical Research Institute Japan 30, 25–33.

Nakamura Y., (2000). Clear identification of fundamental idea of Nakamura's technique and its applications, Proc. 12 World Conf. on Earthq. Ength, pp. 2656-2664

Nakamura, Y. (1988). On the urgent earthquake detection and alarm system (UrEDAS), in Proc. 9th World Conf. Earthquake Eng. 7, 673–678.

Nason, G.P. (1994). Wavelet regression by cross-validation. Technical Report 447, Department of Statistics, Standord University, USA.

Nason G and Silverman B (1995), The Stationary Wavelet Transorm and some Statistical Applications. In Wavelets and Statistics (Lecture Notes in Statistics, Vol 103) edited by Antoniadis A and Openheim G, New York, Springer-Verlag, 281-99

Nogoshi M., and Igarashi T (1970). On the propagation characteristics estimations of subsurface using microtremors on the ground surface, J. Seism. Soc. Japan 23, 264–280.

Nogoshi M. and Igarashi T (1971). On the amplitude characteristics of microtremor (part 2), J. Seism. Soc. Japan 24, 26–40.

Odgen, T., Parzen, O. (1996). Change-point approach to data analytic thresholding. Statist. Comput. 6, 93-99.

Olivieri, M., Schweitzer, J., (2007), An Empirical Procedure for Rapid Magnitude Estimation in Italy, Bull. Seism. Soc. Am., Vol. 97, No. 5, pp. 1750–1755

Olivieri, M., Allen, R., Wurman, G., (2008), The Potential for Earthquake Early Warning in Italy Using ElarmS, Bull. Seism. Soc. Am., Vol. 98, No. 1, pp. 495–503

Olson, E. L., and R. M. Allen (2005). The deterministic nature of earthquake rupture, Nature 438, 212–215.

Oncel AO, Main IG, Alptekin O., Cowie PA.(1996), Spatial variation in the fractal properties of seismicity in the north Anatolian fault zone. Tectonophysics;257:189–202.

Otazu, X., Ribo, M., Peracaula, M., Paredes, J.M., Nunez, J.: (2002), Mon. Not. R. Astron. Soc., 333, 365-372.

Pazos, A.,, Gonzalez, M.J., Alguacil, G., (2003) Non-linear filter, using the wavelet transform, applied to seismological records, Journal of Seismology 7: 413–429.

Pantin, E., Starck, J.L.: (1996), Astron. Astroph. Suppl., 118, 575.

Parolai S and Galliana-Merino J (2006), Effects of transient seismic noise of H/V spectral ratio, Bull. Seism. Soc. Am. 96, 228–236.

Parolai, S., Picozzi, M., Strollo, A., Pilz, M., Di Giacomo, D., Liss, B., Bindi, D.,(2009), Are Transients carrying Useful Information for estimating H/V spectral ratios, In Increasing seismic safety by combining enginnering technologies and seismological data, Mucciarelli, M., Herak, M., and Cassidy, J., (editors), Springer Science, NL

Papadopoulos I. Th., Savvaidis A., Theodoulidis N. P., Papazachos C. B., Vallianatos F.,(2006), Shallow Geological structure properties using HVSR and comparison with other geophysical information, 1st ECEES, Geneva

Papazachos, B. C.(1990) Seismicity of the Aegean and surrounding areas, Tectonophysics, 178, 287-308.

Papazachos CB, Kiratzi AA. (1996) A detailed study of the active crystal deformation in the Aegean and surrounding area. Tectonophysics, 253:129–53.

Papazachos, B., Papazachou, C., (1997). The Earthquakes of Greece. Ziti Editions, Thessaloniki. 304 pp.

Papazachos BC, Karakostas VG, Papazachos CB, Scordilis EM. (2000) The geometry of the Wadati-Benioff zone and lithospheric kinematics in the Hellenic arc. Tectonophysics;319:275–300.

Percival D. and Walden A. (2000), Wavelet methods for time series analysis, Cambridge University Press, Cambridge, UK

Press, W., Teukolsky, S. A., Vetterling, W. T., and Flannery, B., (1995) Numerical Recipes in C, Cambridge (1995).

Provaznic, I., (2001), Wavelet analysis for signal detection-applications to experimental cardiology research. PhD thesis, Brno University of Technology.

Pinsky, M.A., (2001), "Introduction to Fourier Analysis and Wavelets", Brooks/Cole.

Roux, S., Muzy, J.F., Arneodo, A.: (1999), Eur. Phys. J. B, 8, 301.

Rydelek P, Pujol J (2004) Real-time seismic warning with a 2-station subarray. Bull Seism Soc Am 94:1546-1550

Rydelek, P., and S. Horiuchi (2006), Is earthquake rupture deterministic?, Nature, 442, E5–E6, doi:10.1038/nature04963.

Satoh, T., Kawase, H., Iwata, T., Higashi, S., Sato, T., Irikura, K., Huang, H.-C., (2001), S-wave velocity structure of the Taichung Basin, Taiwan, estimated from array and single-station records of microtremors, Bull.Seism.Soc.Am,vol 91,n.5,pp1267-1282

Satriano, C., Zollo, A., Lomax, A., (2008), Real-Time Evolutionary Earthquake Location for Seismic Early Warning, Bull Seism Soc Am, vol 98, no 3.

Shi Y, Bolt BA.(1982) The standard error of the magnitude-frequency b value. Bull Seism Soc Am;72:1677–87.

Schukla , P., (2003), Complex wavelet transforms and their applications, MPhil thesis, University of Strathclyde, Scotland

SE-RISK (Advanced techniques for SEismic RISK reduction in Mediterranean Archipelago Regions) project (2008). http://www.ims.forth.gr/joint_projects/se_risk/se_risk.html

SESAME (Site EffectS assessment using AMbient Excitations) project (2004), EVG1CT-2000-00026. http://sesame-fp5.obs.ujf-grenoble.fr/

Scherbaum, F., (1996). Of poles and zeros: fundamentals of digital seismology, Kluwer Academic publishers, p.292

Schneider, K., Kevlahan, N.K.-R., Farge, M.: (1997), Theor. Comp. Fluid Dyn., 9. 191 -206.

Shensea J (1992), The Discrete Wavelet Transform: Wedding the a Trous and Mallat Algorithms, IEEE Transactions on Signal Processing, 40, 2464-2482

Spakman, W., Wortel, M.J.R., Vlaar, N.J., (1988), The Hellenic subduction zone: a tomographic image and its geodynamic implications. Geophysical Research Letters 15 (1), pp. 60-63

Spudich, P., Hellweg, M., Lee, W.H.K., (1996), Directional topographic site response at Tarzana observed in aftershocks of the 1994 Northridge, Bull. Seism. Soc. Am., v. 86; no. 1B; p. S193-S208

Simmons, F., J., Dando, B., Allen, R., (2006), Automatic detection and rapid determination of earthquake magnitude by wavelet multiscale analysis of the primary arrival, Earth & Planetary Science Letters, 250, p.214-223

Strang, G., (1994), Wavelets, American Scientist, Vol. 82, Issue 3, p.250-255.

Stewart SW (1977). Real time detection and location of local seismic events in central California. Bull. Seism. Soc. Am. 67, 433–452.

Teich MC, Lowen SB, Jost BM, Vibe-Rheymer K, Heneghan C.(2001) Heart rate variability: measures and models. In: Akay M, editor. Nonlinear biomedical signal processing. Dynamic analysis and modeling, vol. II. New York: IEEE Press; p.159–213.

Telesca L, Cuomo V, Lapenna V, Macchiato M., (2001), Identifying space-time clustering properties of the 1983–1997 Irpinia-Basilicata (Southern Italy) seismicity. Tectonophysics, 330:93–102.

Telesca L, Cuomo V, Lapenna V, Macchiato M. (2001), Depth-dependent time clustering behaviour in seismicity of southern California. Geophys Res Lett;28:4323–6.

Telesca L, Lapenna V, Vallianatos F. (2002), Monofractal and multifractal approaches in investigating scaling properties in temporal patterns of the 1983–2000 seismicity in the western Corinth graben, Greece, Phys Earth Planet Int., 131:63–79.

Telesca L, Lapenna V, Alexis N. Multiresolution wavelet analysis of earthquakes (2004). Chaos Soliton Fract.,;22:741–8.

Telesca, L., Nikolintaga, I., Vallianatos, F, (2006), "time-scaling analysis of southern Aegean seismicity", Chaos, solitons & Fractals, 28, (2), pp.361-366

Telesca, L., Hloupis, G., Nikolintaga, I., Vallianatos, F., (2007), Temporal patterns in southern Aegean seismicity revealed by the multiresolution wavelet analysis, Communications in Nonlinear Science and Numerical Simulation 12 (8), pp. 14181426

Teng TL, Wu YM, Shin TC, Tsai YB, Lee WHK (1997) One minute after: strong motion

map, effective epicenter, and effective magnitude. Bull Seism Soc Am, no. 87:12091219

Thurner S, Feurstein MC, Teich MC. Multiresolution wavelet analysis of heartbeat intervals discriminates healthy patients from those with cardiac pathology. Phys Rev Lett 1998;80:1544–7.

Torrence C and Compo G (1999), A practical guide to wavelet analysis, Bull. Amer. Met. Soc. 79, 61—78.

Tokimatsu, K, (1997), Geotechnical site characterizations using surface waves, Proc. 1st Intl. Conf. Earthquake Geotechnical Engineering

Vannucci, M. & Corradi, F. (1999). Covariance structure of wavelet coefficients: theory and models in a Bayesian perspective. J. R. Statist. Soc. B, 61, 971–986.

Vetterli, V. and Kovacevic, J., (1995),Wavelets and Subband Coding, Prentice-Hall.

Vidakovic, B. (1998). Wavelet based nonparametric Bayes methods. In Practical Nonparametric and Semiparametric Bayesian Statistics, Dey, D.D., M¨ uller, P. & Sinha, D. (Eds.), Lecture Notes in Statistics 133, pp. 133-155, New York: Springer-Verlag.

Vidakovic, B. & Ruggeri, F. (2000). BAMS method: theory and simulations. Discussion Paper, Institute of Statistics and Decision Sciences, Duke University, USA.

Vidakovic, B. (2000). Statistical Modeling by Wavelets. New York: John Wiley & Sons.

Wald, D., Quitoriano, T.H.,Heaton, H., Kanamori, C.W., Scrivner, C,Worden C.,(1999). TriNet “ShakeMaps”: Rapid Generation of Peak Ground Motion and Intensity Maps for Earthquakes in Southern California, Earthq. Spectra, Vol.15, No. 3,pp 537-556

Wieland, M.,(2001) Earthquake Alarm, Rapid Response, and Early Warning Systems: Low Cost Systems for Seismic Risk Reduction. International Workshop on Disaster Reduction, Reston, VA.

Wieland, M., Griesser, M., Kuendig, C., (2000) Seismic Early Warning System for a Nuclear Power Plant. Proc. of 12° World Conference on Earthquake Engineering. Auckland, New Zealand

Wilson TH, Dominic J. Fractal interrelationships between topography and structure. Earth Surf Process Land 1990;23:509–25.

Wolfe, C. J. (2006). On the properties of predominant-period estimators for earthquake early warning, Bull. Seism. Soc. Am. 96, no. 5, 1961–1965,doi 10.1785/0120060017.

Wornell, G. W., (1995), Signal Processing With Fractals: A Wavelet-Based Approach, Prentice-Hall, Englewood Cliffs, N. J.

Wu YM, Zhao L (2006a) Magnitude estimation using the first three seconds P wave amplitude in earthquake early warning. Geophys Res Lett 33:L16312

Wu YM, Yen HY, Zhao L, Huang BS, Liang WT (2006b) Magnitude determination using initial P waves: A single-station approach. Geophys Res Lett 33:L05306

Wu, Y.-M., and H. Kanamori (2005a). Experiment on an onsite early warning method for Taiwan early warning system, Bull. Seism. Soc. Am. 95, 347–353.

Wu,Y.-M., and H. Kanamori (2005b). Rapid assessment of damage potential of earthquakes

in Taiwan from the beginning of P waves, Bull. Seism. Soc. Am. 95, 1181–1185.

Wu YM, Teng TL, (2002), A virtual sub-network approach to earthquake early warning. Bull Seism Soc Am 92:2008-2018

Wu YM, Teng TL, Shin TC, Hsiao NC (2003) Relationship between peak ground acceleration, peak ground velocity, and intensity in Taiwan. Bull Seism Soc Am 93:386-396

Wu YM, Hsiao NC, Teng TL, Shin TC (2002) Near real-time seismic damage assessment of the rapid reporting system. TAO 13; p.313-324

Wu YM, Chung JK, Shin TC, Hsiao NC, Tsai YB, Lee WHK, Teng TL (1999) Development of an integrated seismic early warning system in Taiwan – case for the Hualien area earthquakes. TAO 10:719-736

Wu YM, Shin TC, Tsai YB (1998) Quick and reliable determination of magnitude for seismic early warning. Bull Seism Soc Am 88:1254-1259

Wu YM, Chen CC, Shin TC, Tsai YB, Lee WHK, Teng TL (1997) Taiwan Rapid Earthquake Information Release System. Seism Res Lett 68:931-943

Yamanaka, K., Kolosov, O., V., Nagata, Koda, K., Nishino, H., Tsukahara, Y., (1993), Analysis of excitation and coherent amplitude enhancement of surface acoustic waves by the phase velocity scanning method, J. Appl. Phys. 74, 6511 ; DOI:10.1063/1.355140

Zhang, Y., Paulson, K.V.: (1997), Pure Appl. Geophys., 149, 405.

Zhou H (1994) Rapid 3-D hypocentral determination using a master station method. J.Geophys Res 99:15439-15455

Zollo, A., M. Lancieri, and S. Nielsen (2006), Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records, Geophys. Res. Lett., 33, L23312,

Appendix A - Publications derived from this research

i) F. Vallianatos, **G. Hloupis, «**HVSR technique improvement using Redundant Wavelet Transform», in Increasing Seismic Safety by combining Engineering Technologies and Seismological Data, Mucciarelli M., Herak M., (editors), NATO ARW series, Springer, 2009

ii) F. Vallianatos, **G. Hloupis**, J.P.Makris, Challenges in Computational Solid Earth System science in the frame of CYCLOPS project, Geophysical Research Abstracts, Vol. 10, EGU2008-A-07118, 2008

iii) **G. Hloupis,** F. Vallianatos, Rapid Earthquake Magnitude Estimation using Wavelets: Application for an Early Warning System in South Aegean, Geophysical Research Abstracts, Vol. 10, EGU2008-A-11710, 2008

iv) **G.Hloupis,** Wavelets in the Analysis of seismoelectromagnetic signals, IUGG conference, 2007, Italy

v) F. Vallianatos, **G. Hloupis**, J. Stonham, Rapid Wavelet Magnitude estimation for Seismic Early Warning, Geophysical Research Abstracts, Vol. 9, 09796, 2007

vi) F. Vallianatos, **G. Hloupis**, J. Stonham, Wavelet Based processing of Microtremors Signals, Geophysical Research Abstracts, Vol. 9, 5-2-2007

vii) F. Vallianatos, **G. Hloupis**, M. Moisidi, I Papadopoulos, J. Makris, «Site Effect Studies using the 8th of January 2006 Kythira Earthquake data recorded in Crete (Southern Greece)», Geophysical Research Abstracts, Vol. 9, 5-2-2007

viii) Luciano Telesca, **George Hloupis**, Irini Nikolintaga and Filippos Vallianatos. Temporal patterns in southern Aegean seismicity revealed by the multiresolution wavelet analysis, Communications in Nonlinear Science and Numerical Simulation, March 2006

ix) Vallianatos, F.; Makris, J.; Soupios, P.; Saltas, V.; **Hloupis, G**.; Nikolintaga, I.; Kokinou, E.; Papadopoulos, I.; Moisidi, M.; Kalisperi,D. A new telemetry seismological network in the front part of the Hellenic arc: its contribution to the study and understanding of the seismic behavior of the area. EGU, Vienna, 2-7 April 2006

x) Vallianatos, F.; **Hloupis, G**.; Papadopoulos, I.; Moisidi, M. Site Response Evaluations following the January 8, 2006 Kithira Earthquake *(solicited)*, Geophysical Research Abstracts, Vol. 8, 08527, 2006

xi) F. Vallianatos, J. Makris, P. Soupios, V. Saltas, **G. Hloupis**, I. Nikolintaga, E. Kokinou, I. Papadopoulos, M. Moisidi, D. Kalisperi, A new telemetry seismological network in the front part of the Hellenic arc: its contribution to the study and understanding of the seismic behaviour of the area, Geophysical Research Abstracts, Vol. 8, 08365, 2006

xii) F. Vallianatos, P. Soupios, J. P. Makris, V. Saltas, I. Papadopoulos and **G. Hloupis,** 2006, Construed Geotechnical Characteristics of Foundation Beds by Geophysical Measurements, Proceedings of 2nd International Conference Advances in Mineral Resources Management and Environmental Geotechnology, Hania, Crete, Greece, 25-27 September 2006

xiii) F. Vallianatos, J. P. Makris, P. Soupios, V. Saltas, I. Papadopoulos, **G. Hloupis**, M. Kouli, D. Alexakis, E. Kokkinou, I. Nikolintaga, D. Kalisperi, M. Moisidi, 2005, Coupling GeoEnvironmental Research & Education: Examples from the Technological Educational Institute of Crete, WSEAS, Engineering Education, 2005, 12-14 July, Vouliagmeni, Athens, ISBN 960-8457-28-9, pp. 503-510

xiv) Kalisperi D., **Hloupis G.**, Makris J.,D. Rust, Vallianatos F., Saltas, V., Soupios P., Vardiampasis I., Educational Software Bundle for Studying Magnetotelluric Theory and Specific Geoelectric Structure Models, WSEAS, Engineering Education, 2005, 12-14 July, Vouliagmeni, Athens, ISBN 960-8457-28-9, pp. 477-486.

xv) **G. Hloupis**, I. Stavrakas, V. Saltas, D. Triantis, F. Vallianatos, J. Stonham, "Identification of contamination in sandstone by means of dielectric and conductivity measurements", WSEAS transactions on Circuits and Systems, Issue 3, vol. 4, March 2005, 148 – 156.

xvi) **G.Hloupis,** F.Vallianatos, J.Stonham, Identification of Short-period transients in microtremors using wavelet transform, Geophysical Research Abstracts, Vol. 7, 09927, 2005

xvii) F. Vallianatos, **G. Hloupis**, V. Saltas, J. Stonham, D. Triantis, "Wavelet analysis of time-series dielectric measurements as a tool for identification of soil contamination", 2nd Int. conf. at Applied Geophysics for Engineering, Messina, Italy, 13-16 October 2005.

xviii) **G. Hloupis**, D. Triantis, F. Vallianatos, J. Makris, J. Stonham, "Detection of short period transients in geophysical signals using wavelet transform", WSEAS transactions on Communications, Issue 3, vol. 3, July 2004, 904 – 909.

xix) **G.Hloupis,** F.Vallianatos, J.Stonham, Stationarities vs. Non-stationarities in HVSR technique, 1st International Conference Advances in Mineral Resources Management and Environmental Geotechnology, Chania, Crete, Greece June 2004

xx) **G.Hloupis,** F.Vallianatos, J.Stonham, A Wavelet representation of HVSR technique**,** Proceedings of 10th International Congress Geological society of Greece, Thessaloniki, April 2004.

xxi) F. Vallianatos**, G. Hloupis**, J. Makris, J. Stonham, "Wavelet analysis of a ULF geomagnetic field measured in Crete, Greece", EGU, 1st general assembly, 25 – 30 April 2004, Nice, France.

xxii) F. Vallianatos, P. Soupios, J. P. Makris, V. Saltas**, G. Hloupis,** "Study of the tectono-karstic voids using electrical tomography and microtremor measurements", IV International Workshop on Magnetic, Electric and Electromagnetic Methods in Seismology and Volcanology (MEEMSV-2004), 5 – 9 September 2004, La Londe les Maures, France.

Appendix B
Results from ambient noise surveys at stations' locations

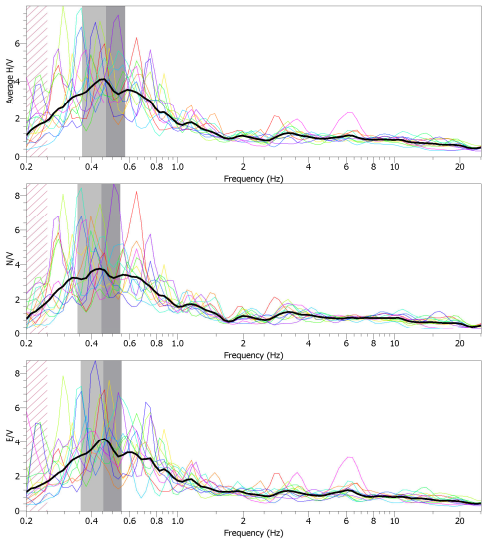


Figure B1: Station CHAN

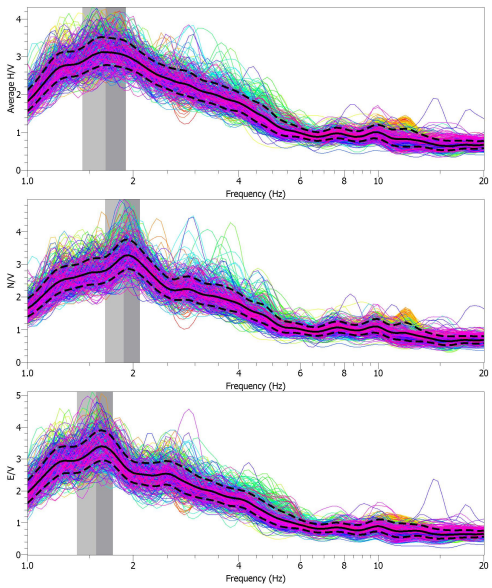


Figure B2: Station KTHR

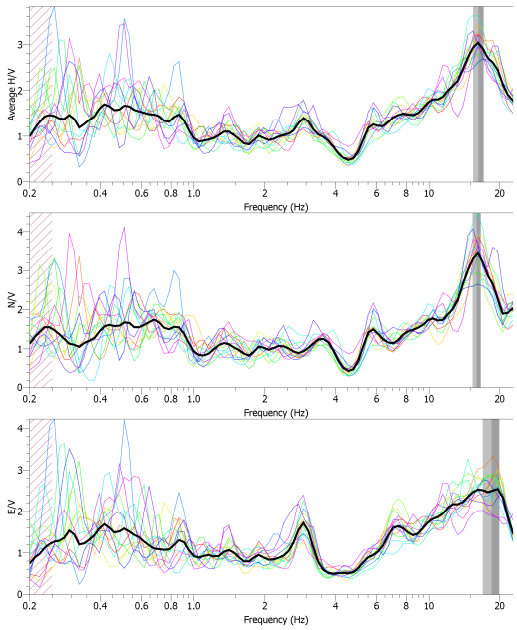


Figure B3: Station KNDR

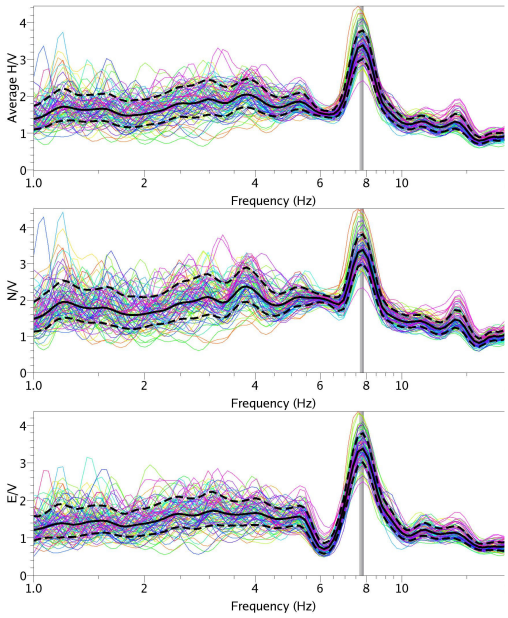


Figure B4 :Station STIA

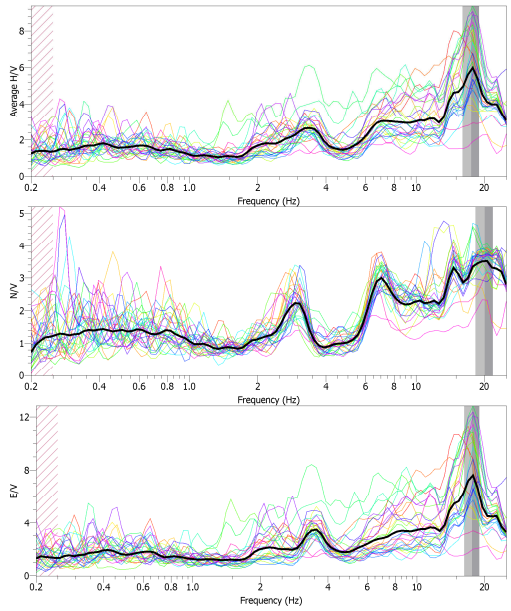


Figure B5: Station HERA

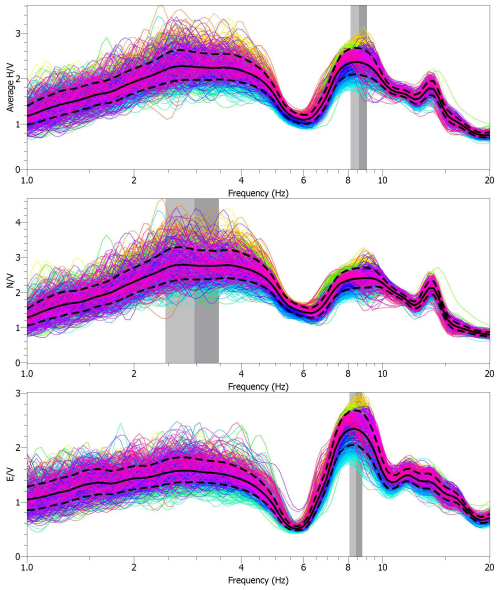


Figure B6: Station GVDS

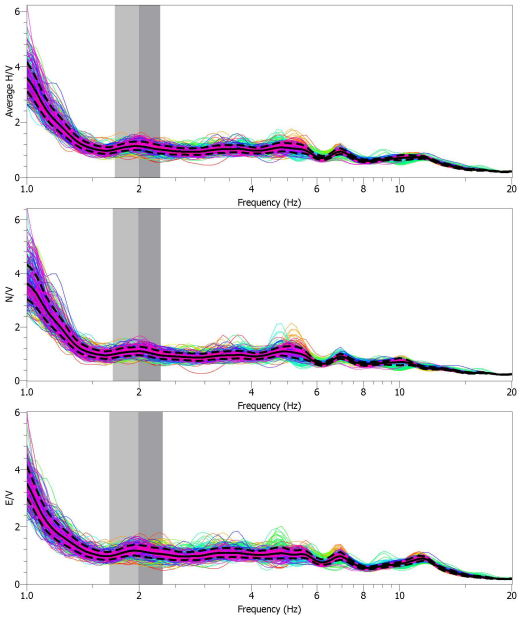


Figure B7: Station PRNS

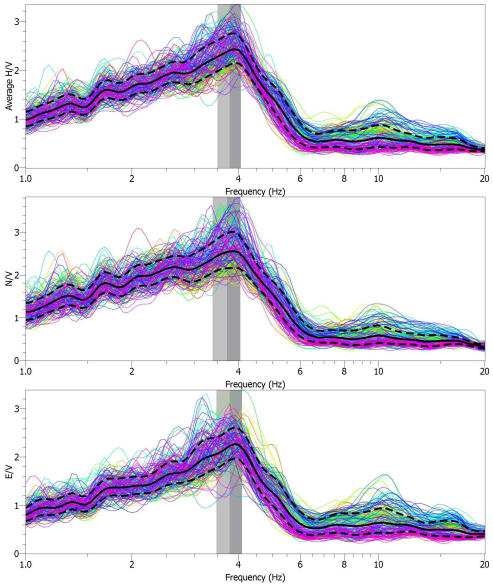


Figure B8: Station KLDN

Appendix C: Satellite Link Specifications

Digital Link Budget

Thursday, July 03, 2006

Service Name	*HUB to Remotes Channel*
Coverage	*South Greece*
Uplink earth station	*Chania, Greece*
Downlink earth station	*Rodhos, Greece*
Satellite name	*Hellas Sat 2*

Link Input Parameters	Up	Down	Units
Site latitude	35.37N	36.40N	degrees
Site longitude	24.48E	28.25E	degrees
Site altitude	0.5	0.5	km
Frequency	14	12	GHz
Polarization	Vertical	Horizontal	
Rain model	ITU (29.5)	ITU (40.3)	(mm/h or zone)
Availability (average year)	99.89	99.91	%
Antenna aperture	2.4	1.2	metres
Antenna efficiency / gain	65	65	% (+ prefix dBi)
Coupling loss	1	0.5	dB
Antenna tracking / mispoint error	1	0.2	dB
LNB noise figure / temp		1	dB (+ prefix K)
Antenna noise		30	K
Adjacent carrier interference	28	28	dB
Adjacent satellite interference	26	26	dB
Cross polarization interference	28	28	dB
Uplink station HPA output back-off	3		dB
Number of carriers / HPA	1		
HPA C/IM (up)	30		dB
Uplink power control	0		dB

Uplink filter truncation loss	0	dB

Calculations at Saturation	Value	Units
Gain 1m^2	44.38	dB/m2
Uplink C/No	103.72	dB.Hz
Downlink C/No	94.50	dB.Hz
Total C/No	94.01	dB.Hz
Uplink EIRP for saturation	77.34	dBW

General Calculations	Up	Down	Units
Elevation	46.12	46.26	degrees
True azimuth	155.90	162.26	degrees
Compass bearing	152.70	158.62	degrees
Path distance to satellite	37333.23	37323.27	km
Propagation time delay	0.124530	0.124497	seconds
Antenna efficiency	65.00	65.00	%
Antenna gain	49.06	41.70	dBi
Availability (average year)	99.89	99.91	%
Link downtime (average year)	9.643	7.889	hours
Availability (worst month)	99.582	99.649	%
Link downtime (worst month)	3.052	2.563	hours
Spectral power density	-56.57	-27.85	dBW/Hz

Uplink Calculation	Clear	Rain Up	Rain Dn	Units
Uplink transmit EIRP	41.67	41.67	41.67	dBW
Transponder input back-off (total)	9.00	9.00	9.00	dB
Input back-off per carrier	35.67	36.85	35.67	dB
Mispoint loss	1.00	1.00	1.00	dB
Free space loss	206.81	206.81	206.81	dB
Atmospheric absorption	0.11	0.11	0.11	dB
Tropospheric scintillation fading	0.29	0.29	0.29	dB
Atmospheric losses total	0.41	0.41	0.41	dB
Total path loss (excluding rain)	207.22	207.22	207.22	dB

Rain attenuation	0.00	1.19	0.00	dB
Uplink power control	0.00	0.00	0.00	dB
Uncompensated rain fade	0.00	1.19	0.00	dB
C/No (thermal)	68.05	66.87	68.05	dB.Hz
C/N (thermal)	18.87	17.69	18.87	dB
C/ACI	28.00	26.81	28.00	dB
C/ASI	26.00	24.81	26.00	dB
C/XPI	28.00	26.81	28.00	dB
C/IM	30.00	30.00	30.00	dB
Eb/(No+Io)	15.07	13.93	15.07	dB

Downlink Calculation	Clear	Rain Up	Rain Dn	Units
Satellite EIRP total	52.00	52.00	52.00	dBW
Transponder output back-off (total)	4.00	4.00	4.00	dB
Output back-off per carrier	30.67	31.85	30.67	dB
Satellite EIRP per carrier	21.33	20.15	21.33	dBW
Mispoint loss	0.20	0.20	0.20	dB
Free space loss	205.47	205.47	205.47	dB
Atmospheric absorption	0.09	0.09	0.09	dB
Tropospheric scintillation fading	0.29	0.29	0.29	dB
Atmospheric losses total	0.38	0.38	0.38	dB
Total path loss (excluding rain)	205.85	205.85	205.85	dB
Rain attenuation	0.00	0.00	1.39	dB
Noise increase due to precipitation	0.00	0.00	1.78	dB
Downlink degradation (DND)	0.00	0.00	3.16	dB
Total system noise	133.36	133.36	200.73	K
Figure of merit (G/T)	19.75	19.75	17.98	dB/K
C/No (thermal)	63.84	62.65	60.67	dB.Hz
C/N (thermal)	14.65	13.47	11.49	dB
C/ACI	28.00	26.81	28.00	dB
C/ASI	26.00	24.81	26.00	dB
C/XPI	28.00	26.81	28.00	dB

C/IM	18.00	16.81	18.00	dB
Eb/(No+Io)	10.53	9.35	8.34	dB
Totals per Carrier (End-to-End)	Clear	Rain Up	Rain Dn	Units
C/No (thermal)	62.44	61.25	59.94	dB.Hz
C/N (thermal)	13.26	12.07	10.76	dB
C/ACI	24.99	23.80	24.99	dB
C/ASI	22.99	21.80	22.99	dB
C/XPI	24.99	23.80	24.99	dB
C/IM	17.73	16.61	17.73	dB
C/(No+Io)	60.41	59.23	58.68	dB.Hz
C/(N+I)	11.23	10.05	9.50	dB
Eb/(No+Io)	9.22	8.05	7.50	dB
System margin	2.00	2.00	2.00	dB
Net Eb/(No+Io)	7.22	6.05	5.50	dB
Required Eb/(No+Io)	5.50	5.50	5.50	dB
Excess margin	1.72	0.55	0.00	dB

Earth Station Power Requirements	Value	Units
EIRP per carrier	41.67	dBW
Antenna gain	49.06	dBi
Antenna feed flange power per carrier	-7.39	dBW
Uplink power control	0.00	dB
HPA output back off	3.00	dB
Waveguide loss	1	dB
Filter truncation loss	0	dB
Number of HPA carriers	1	
Total HPA power required	-3.3878	dBW
Required HPA power capability	0.4584	W
Spectral power density	-56.57	dBW/Hz

Space Segment Utilization	Value	Units
Overall link availability	99.800	%
Information rate (inc overhead)	0.1313	Mbps

Transmit rate	0.1656	Mbps
Symbol rate	0.0828	Mbaud
Occupied bandwidth	0.0994	MHz
Noise bandwidth	49.18	dB.Hz
Minimum allocated bandwidth required	0.1076	MHz
Allocated transponder bandwidth	0.1100	MHz
Percentage transponder bandwidth used	0.31	%
Used transponder power	21.33	dBW
Percentage transponder power used	0.22	%
Max carriers by transponder bandwidth	327.27	
Max carriers by transponder power	464.14	
Max transponder carriers limited by:-	Bandwidth	[327.27]

Digital Link Budget

Thursday, July 03, 2006

Service Name	*Remotes to HUB Channel*
Coverage	*South Greece*
Uplink earth station	*Rodhos, Greece*
Downlink earth station	*Chania, Greece*
Satellite name	*Hellas Sat 2*

Link Input Parameters	Up	Down	Units
Site latitude	36.40N	35.37N	degrees
Site longitude	28.25E	24.48E	degrees
Site altitude	0.4	0.5	km
Frequency	14	12	GHz
Polarization	Horizontal	Vertical	
Rain model	ITU (40.3)	ITU (29.5)	(mm/h or zone)
Availability (average year)	99.88	99.92	%
Antenna aperture	1.2	2.4	metres

Antenna efficiency / gain	65	65	% (+ prefix dBi)
Coupling loss	0.2	0.5	dB
Antenna tracking / mispoint error	0.5	0.5	dB
LNB noise figure / temp		1	dB (+ prefix K)
Antenna noise		30	K
Adjacent carrier interference	28	28	dB
Adjacent satellite interference	26	26	dB
Cross polarization interference	28	28	dB
Uplink station HPA output back-off	2		dB
Number of carriers / HPA	1		
HPA C/IM (up)	60		dB
Uplink power control	0		dB
Uplink filter truncation loss	.2		dB
Required HPA power capability	MIN		W

Satellite Input Parameters	Value	Units
Satellite longitude	39.00E	degrees
Transponder type	TWTA	
Receive G/T	6	dB/K
Saturation flux density	-86.5	dBW/m2
Satellite attenuator pad	0	dB
Satellite ALC	0	dB
EIRP (saturation)	52	dBW
Transponder bandwidth	36	MHz
Input back off total	9	dB
Output back off total	4	dB
Intermodulation interference	18	dB
Number of transponder carriers	AUTO	

Carrier/Link Input Parameters	Value	Units
Modulation	4-PSK	

Required bit error rate performance	10^-8	
Required Eb/No without FEC coding	11.97	dB
Required Eb/No with FEC coding	5.5	dB
Information rate	0.128	Mbps
Overhead	35	%
FEC code rate	0.793	
Spreading gain	0	dB
Reed Solomon code	1	
(1 + Roll off factor)	1.2	
Carrier spacing factor	1.3	
Bandwidth allocation step size	0.01	MHz
System margin	2	dB

Calculations at Saturation	Value	Units
Gain 1m^2	44.38	dB/m2
Uplink C/No	103.72	dB.Hz
Downlink C/No	100.22	dB.Hz
Total C/No	98.62	dB.Hz
Uplink EIRP for saturation	76.84	dBW

General Calculations	Up	Down	Units
Elevation	46.26	46.12	degrees
True azimuth	162.26	155.90	degrees
Compass bearing	158.62	152.70	degrees
Path distance to satellite	37323.27	37333.23	km
Propagation time delay	0.124497	0.124530	seconds
Antenna efficiency	65.00	65.00	%
Antenna gain	43.04	47.72	dBi
Availability (average year)	99.88	99.92	%
Link downtime (average year)	10.519	7.013	hours
Availability (worst month)	99.549	99.683	%

Link downtime (worst month)	3.292	2.313	hours
Spectral power density	-54.45	-31.25	dBW/Hz

Uplink Calculation	Clear	Rain Up	Rain Dn	Units
Uplink transmit EIRP	38.96	38.96	38.96	dBW
Transponder input back-off (total)	9.00	9.00	9.00	dB
Input back-off per carrier	37.88	39.59	37.88	dB
Mispoint loss	0.50	0.50	0.50	dB
Free space loss	206.81	206.81	206.81	dB
Atmospheric absorption	0.11	0.11	0.11	dB
Tropospheric scintillation fading	0.30	0.30	0.30	dB
Atmospheric losses total	0.41	0.41	0.41	dB
Total path loss (excluding rain)	207.22	207.22	207.22	dB
Rain attenuation	0.00	1.71	0.00	dB
Uplink power control	0.00	0.00	0.00	dB
Uncompensated rain fade	0.00	1.71	0.00	dB
C/No (thermal)	65.84	64.14	65.84	dB.Hz
C/N (thermal)	15.47	13.76	15.47	dB
C/ACI	28.00	26.29	28.00	dB
C/ASI	26.00	24.29	26.00	dB
C/XPI	28.00	26.29	28.00	dB
C/IM	60.00	60.00	60.00	dB
Eb/(No+Io)	12.67	10.97	12.67	dB

Downlink Calculation	Clear	Rain Up	Rain Dn	Units
Satellite EIRP total	52.00	52.00	52.00	dBW
Transponder output back-off (total)	4.00	4.00	4.00	dB
Output back-off per carrier	32.88	34.59	32.88	dB
Satellite EIRP per carrier	19.12	17.41	19.12	dBW
Mispoint loss	0.50	0.50	0.50	dB
Free space loss	205.47	205.47	205.47	dB
Atmospheric absorption	0.09	0.09	0.09	dB

Tropospheric scintillation fading	0.29	0.29	0.29	dB
Atmospheric losses total	0.38	0.38	0.38	dB
Total path loss (excluding rain)	205.85	205.85	205.85	dB
Rain attenuation	0.00	0.00	1.00	dB
Noise increase due to precipitation	0.00	0.00	1.40	dB
Downlink degradation (DND)	0.00	0.00	2.41	dB
Total system noise	133.36	133.36	184.18	K
Figure of merit (G/T)	25.47	25.47	24.07	dB/K
C/No (thermal)	67.34	65.64	64.94	dB.Hz
C/N (thermal)	16.97	15.26	14.57	dB
C/ACI	28.00	26.29	28.00	dB
C/ASI	26.00	24.29	26.00	dB
C/XPI	28.00	26.29	28.00	dB
C/IM	18.00	16.29	18.00	dB
Eb/(No+Io)	11.80	10.10	10.48	dB

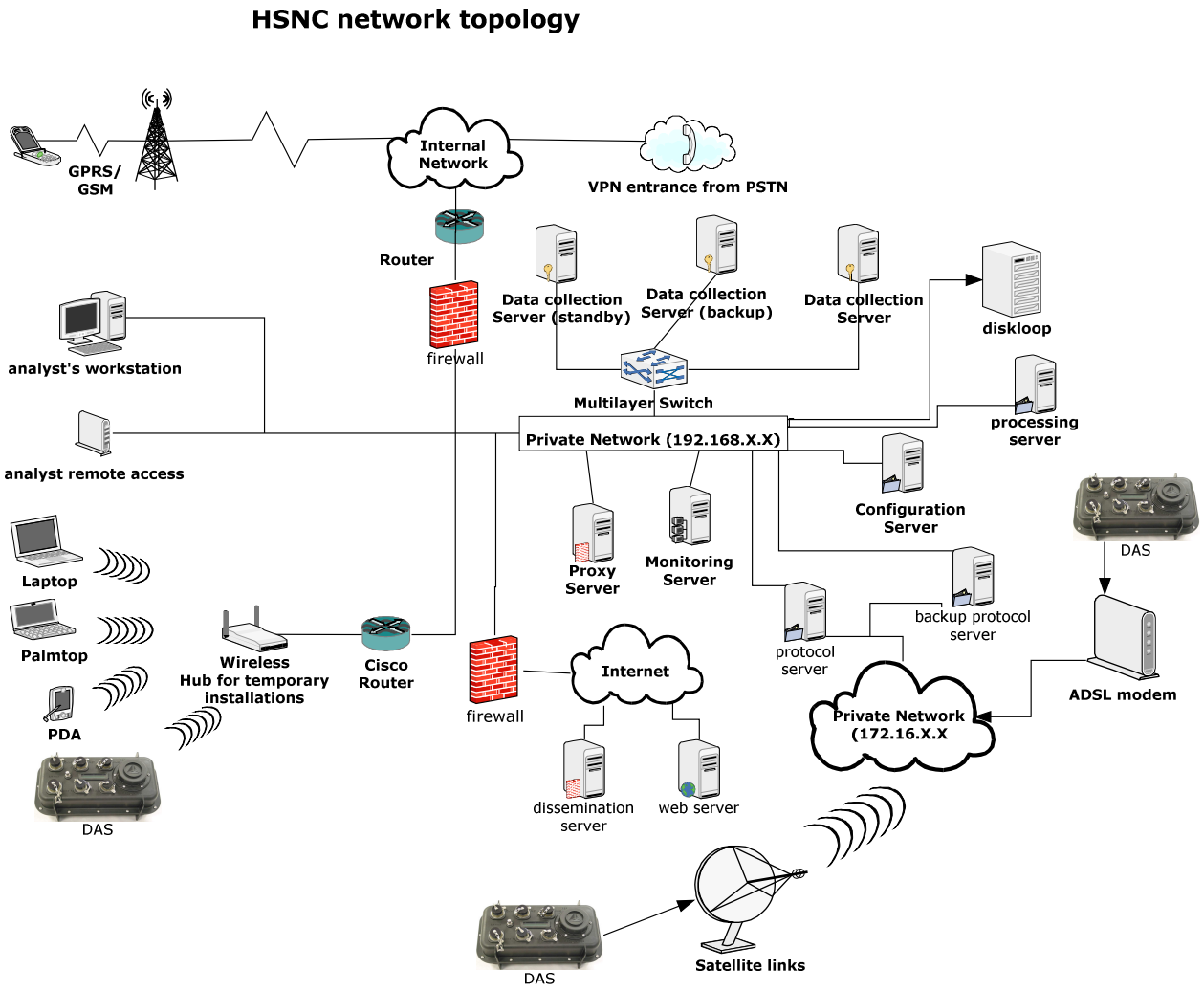
Totals per Carrier (End-to-End)	Clear	Rain Up	Rain Dn	Units
C/No (thermal)	63.52	61.81	62.36	dB.Hz
C/N (thermal)	13.15	11.44	11.98	dB
C/ACI	24.99	23.28	24.99	dB
C/ASI	22.99	21.28	22.99	dB
C/XPI	24.99	23.28	24.99	dB
C/IM	18.00	16.29	18.00	dB
C/(No+Io)	61.58	59.88	60.80	dB.Hz
C/(N+I)	11.21	9.50	10.43	dB
Eb/(No+Io)	9.21	7.50	8.43	dB
System margin	2.00	2.00	2.00	dB
Net Eb/(No+Io)	7.21	5.50	6.43	dB
Required Eb/(No+Io)	5.50	5.50	5.50	dB
Excess margin	1.71	0.00	0.93	dB

Earth Station Power Requirements	Value	Units
EIRP per carrier	38.96	dBW

Antenna gain	43.04	dBi
Antenna feed flange power per carrier	-4.08	dBW
Uplink power control	0.00	dB
HPA output back off	2.00	dB
Waveguide loss	0.2	dB
Filter truncation loss	.2	dB
Number of HPA carriers	1	
Total HPA power required	-1.6772	dBW
Required HPA power capability	0.6796	W
Spectral power density	-54.45	dBW/Hz

Space Segment Utilization	Value	Units
Overall link availability	99.800	%
Information rate (inc overhead)	0.1728	Mbps
Transmit rate	0.2179	Mbps
Symbol rate	0.1090	Mbaud
Occupied bandwidth	0.1307	MHz
Noise bandwidth	50.37	dB.Hz
Minimum allocated bandwidth required	0.1416	MHz
Allocated transponder bandwidth	0.1500	MHz
Percentage transponder bandwidth used	0.42	%
Used transponder power	19.12	dBW
Percentage transponder power used	0.13	%
Max carriers by transponder bandwidth	240.00	
Max carriers by transponder power	772.30	
Max transponder carriers limited by:-	Bandwidth	[240.00]

Appendix D: Network Topology of Central Recording Centre



Appendix E: Installations

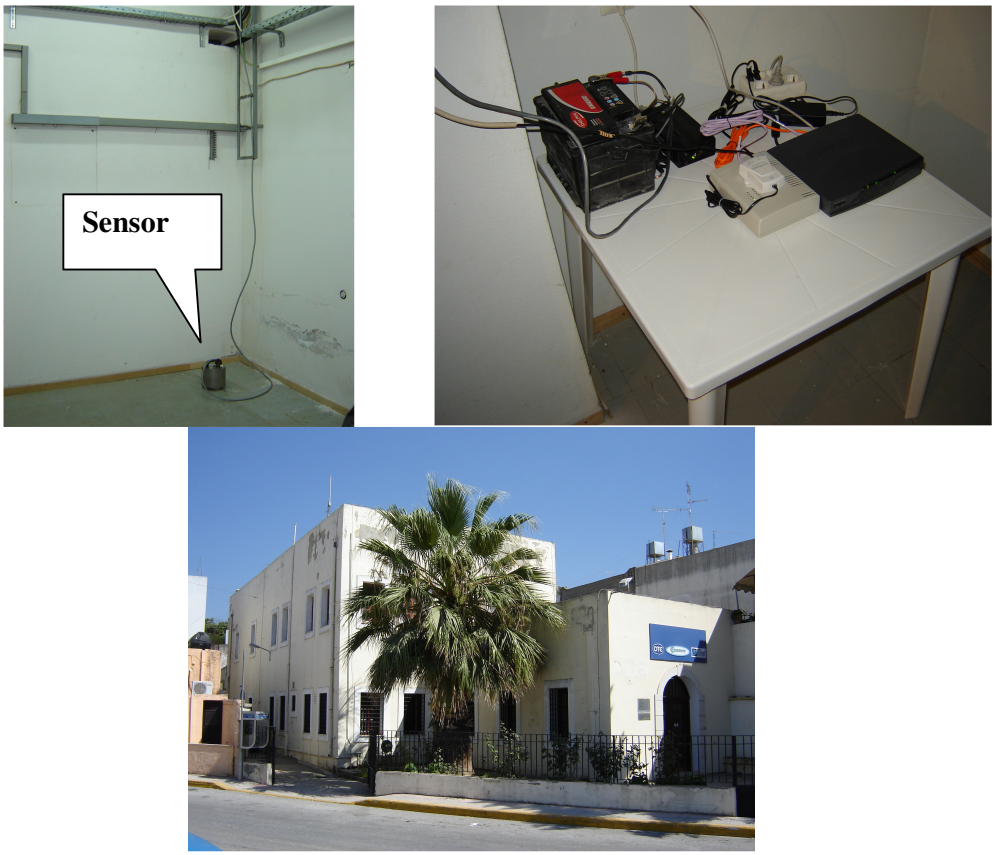


Figure E.1: KSTL seismological station

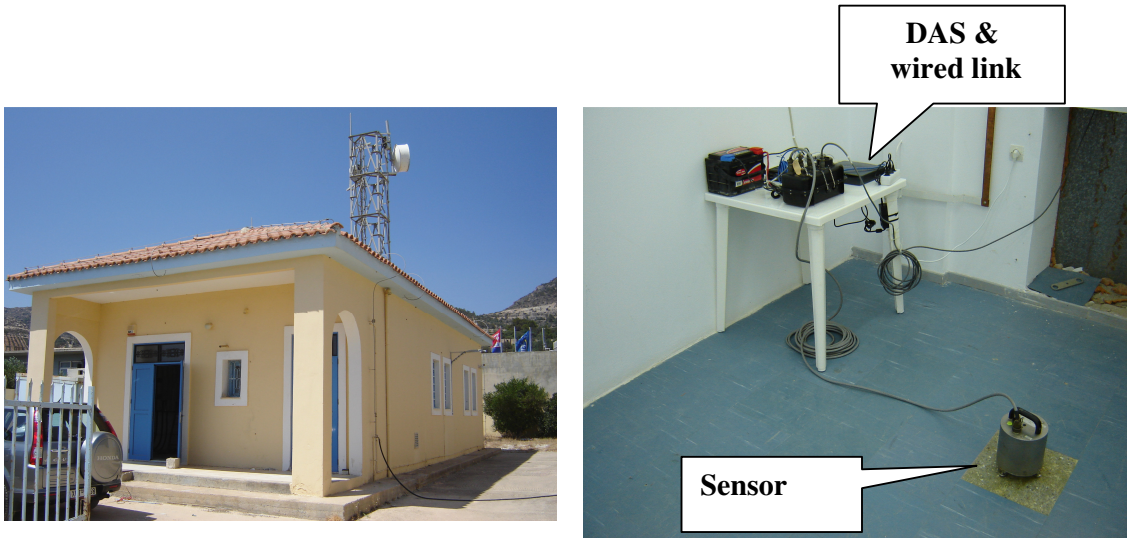


Figure E.2: FRMA seismological station

Figure E.3: GVDS seismological station

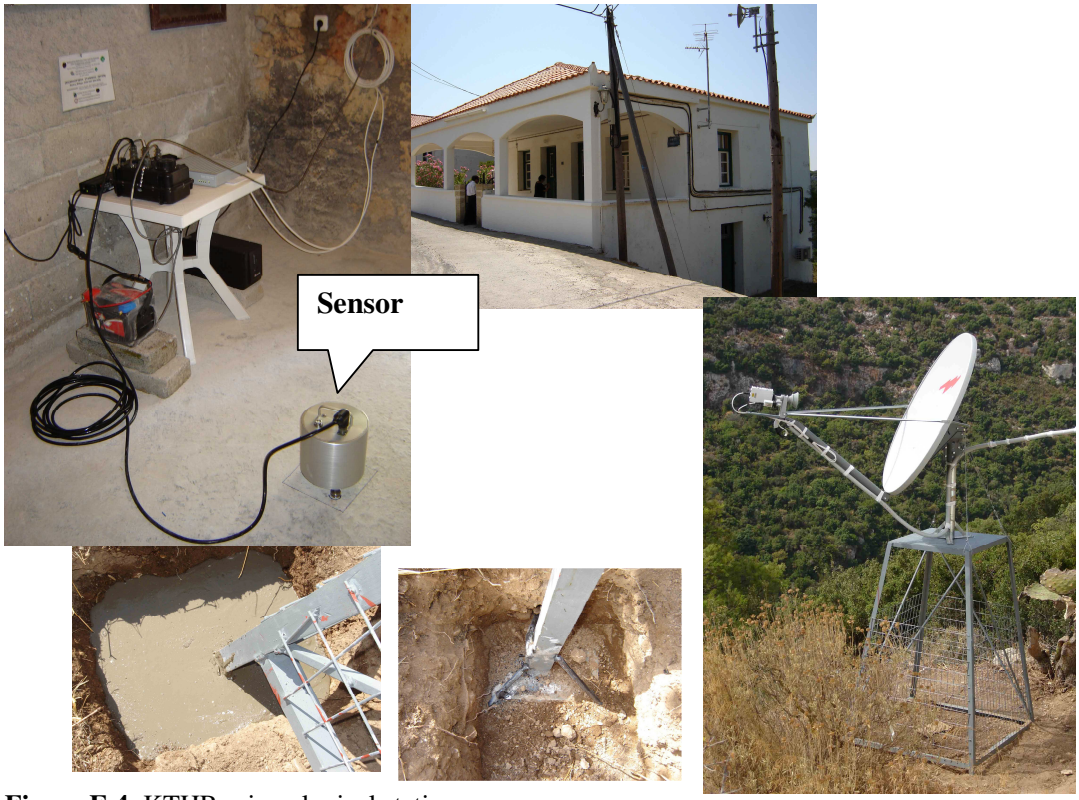
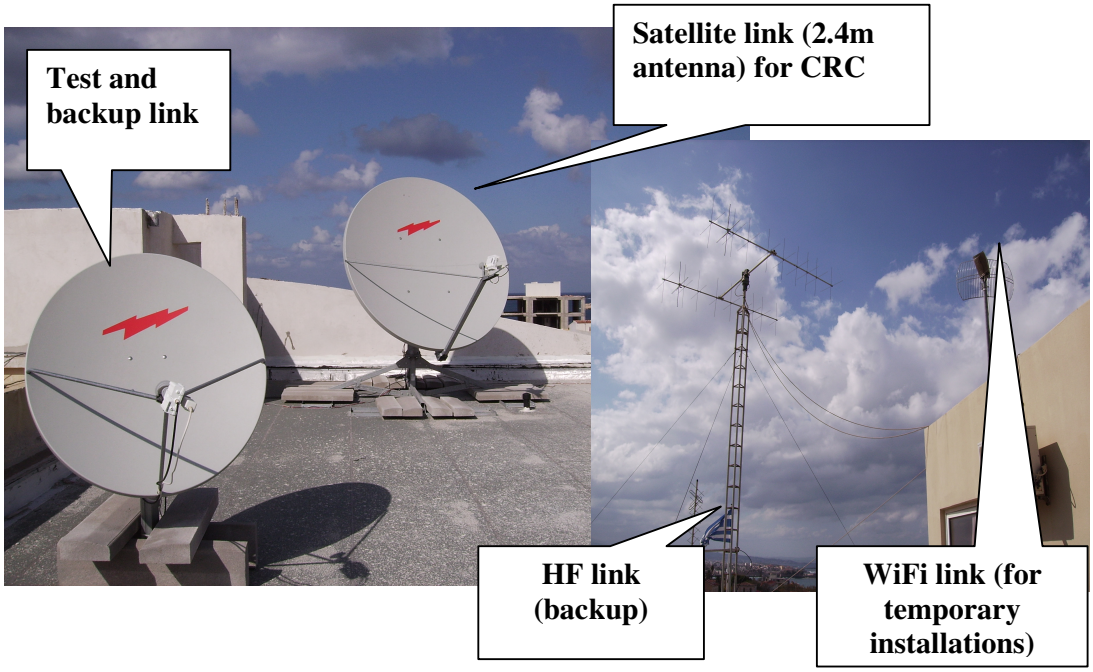
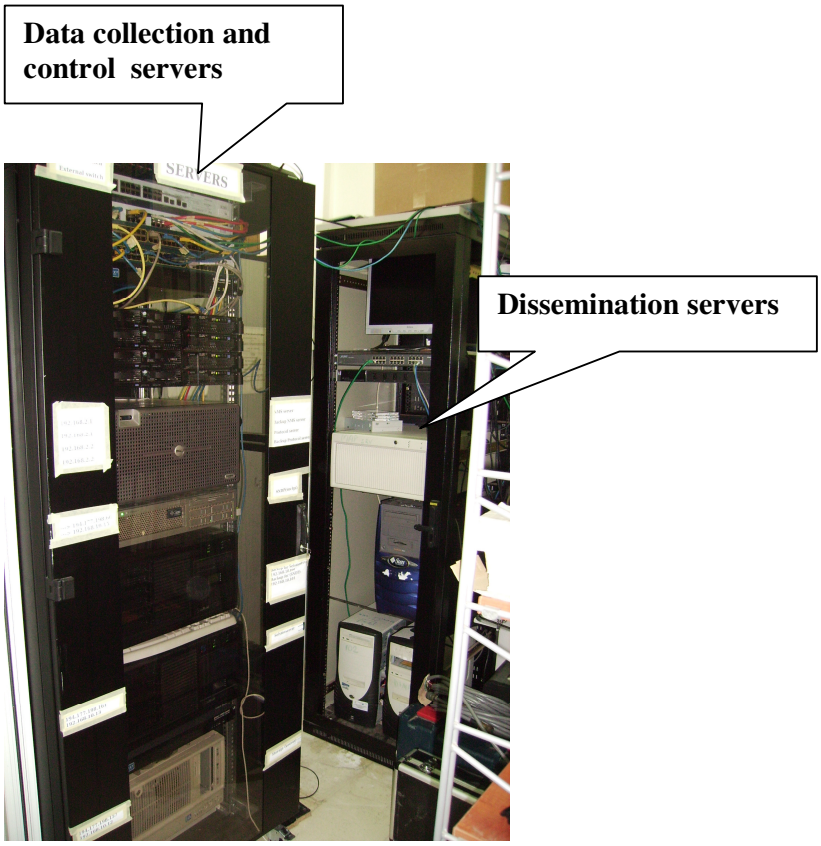


Figure E.4: KTHR seismological station


Test and backup link
Satellite link (2.4m antenna) for CRC
HF link (backup)
WiFi link (for temporary installations)


Data collection and control servers
Dissemination servers

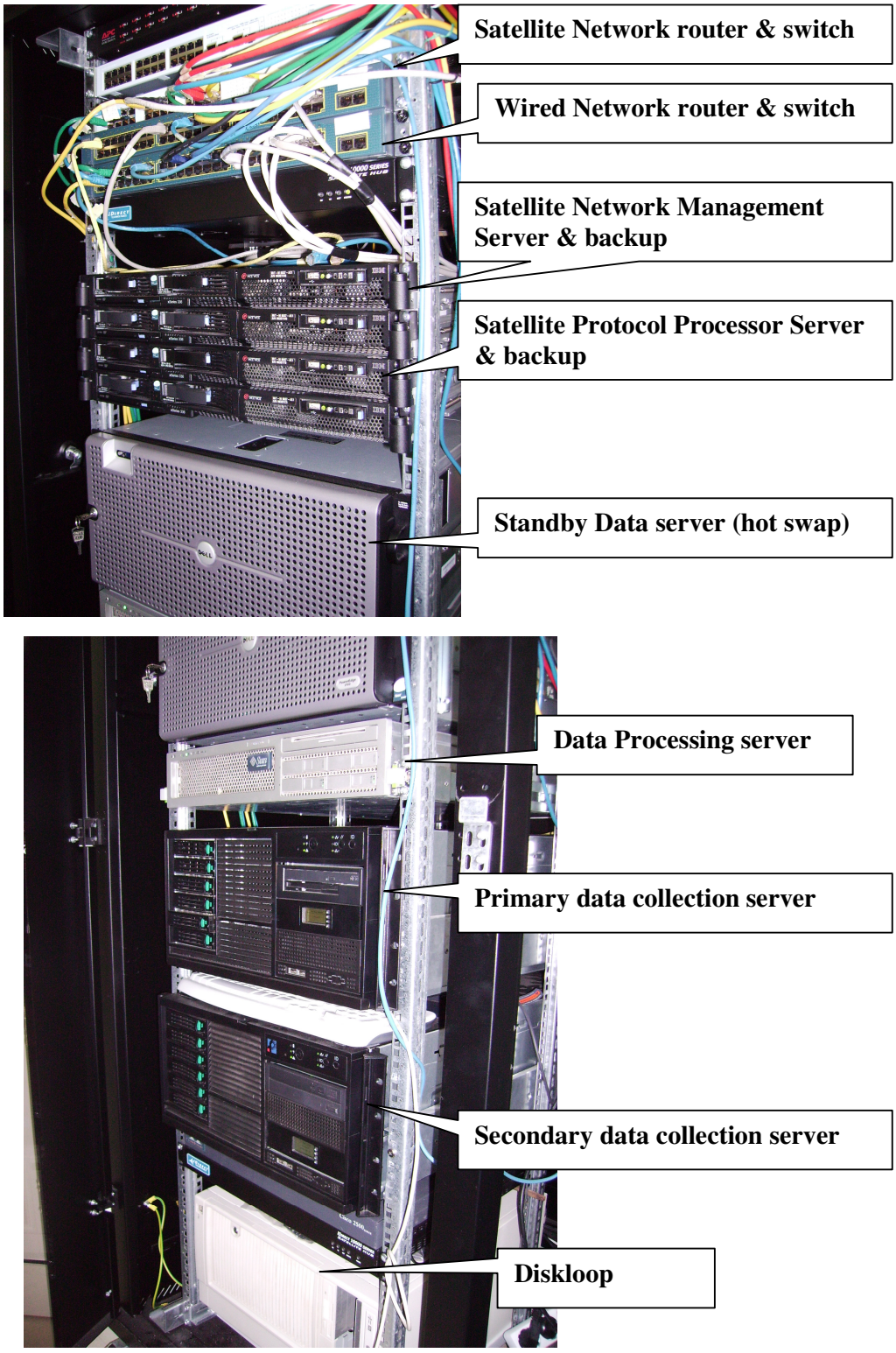


Figure E.5: Central Recording Centre

Appendix F : Result from synthetic signals' denoising

	Index	NL-DTWT	SCAD	HS	CV-H	CV-S	RHT-S	TI-H	TI-S
1	RS1	-2.5	-2.00	-3.00	-3.38	-2.94	-3.16	-4.00	-3.50
2	CL1	5.47	5.23	6.07	5.56	5.51	6.11	7.12	6.59
3	RQ1	25.81	26.00	24.67	29.90	27.86	27.60	34.45	29.56
4	GN1	8.79	8.90	10.89	10.21	9.50	10.56	12.87	11.88
5	RS2	6.78	7.20	6.81	6.87	6.82	6.71	7.87	7.34
6	CL2	8.39	6.78	9.36	12.35	10.37	9.84	12.10	10.73
7	RQ2	9.3	9.10	8.76	9.05	9.18	9.57	11.76	10.26
8	GN2	6.18	6.00	7.23	5.90	5.75	6.87	8.45	7.84
9	RS3	6.21	6.10	6.60	7.34	6.45	6.60	7.98	7.29
10	CL3	14.55	15.20	15.56	13.26	13.24	15.24	16.02	15.79
11	RQ3	18.49	17.80	17.28	21.42	19.00	18.72	21.91	19.60
12	GN3	16.54	16.00	17.94	19.35	17.09	17.64	19.87	18.90
13	RS4	9.51	8.65	10.41	9.09	9.54	10.88	12.87	11.64
14	CL4	24.56	24.58	25.55	22.10	23.93	26.24	28.87	27.21
15	RQ4	11.56	12.45	11.43	14.12	13.17	12.79	14.67	13.05
16	GN4	28.98	29.02	30.20	29.44	29.96	30.08	31.38	29.32
17	RS5	11.35	12.45	12.05	11.53	11.73	12.97	14.56	12.67
18	CL5	13.99	15.06	14.45	11.11	13.21	14.48	17.88	15.39
19	RQ5	12.98	14.67	13.17	14.55	14.49	14.04	15.76	13.77
20	GN5	12.89	13.87	14.20	11.08	12.61	13.53	14.52	13.67
21	RS6	16.85	15.68	16.68	14.47	16.48	17.74	19.01	16.99
22	CL6	17.54	18.98	19.18	19.94	19.73	19.21	19.95	18.63
23	RQ6	38.15	22.45	28.86	33.62	37.77	36.00	42.54	34.00
24	GN6	18.85	18.07	19.71	20.78	20.86	20.52	21.34	19.55
25	RS7	22.15	23.45	22.24	20.40	23.00	23.58	25.67	22.82
26	CL7	37.89	38.02	39.52	32.99	38.31	39.21	41.40	38.53
27	RQ7	29.87	30.06	28.54	27.04	30.76	32.07	32.45	30.49
28	GN7	20.14	20.45	22.49	21.82	22.68	23.55	24.53	23.51
29	RS8	38.78	32.56	36.60	37.44	41.20	38.80	39.91	38.26
30	CL8	41.56	38.90	41.87	44.74	46.65	44.71	44.60	43.23
31	RQ8	40.25	42.66	39.48	49.06	48.28	42.78	45.01	42.25
32	GN8	45.68	43.54	46.26	40.87	46.78	46.67	48.98	46.89
33	RS9	34.25	34.60	32.79	28.53	33.93	35.30	34.02	33.40
34	CL9	38.14	36.06	38.42	41.47	43.03	40.20	39.12	36.77
35	RQ9	32.54	33.55	33.55	28.62	33.06	33.25	33.55	33.00

Table A.1. SNR Results

	Index	NL-DTWT	BL-A	BL-T	NBL	FDR-H	FDR-S	BHT
1	RS1	-2.5	-2.85	-3.72	-3.98	-2.82	-3.27	-4.10
2	CL1	5.47	5.83	6.45	6.78	5.72	5.80	6.98
3	RQ1	25.81	27.14	29.96	32.32	27.04	28.24	32.12
4	GN1	8.79	9.28	10.22	12.01	9.38	10.53	11.98
5	RS2	6.78	6.89	7.41	6.54	6.58	6.76	7.07
6	CL2	8.39	8.98	12.39	11.76	9.27	10.92	11.98
7	RQ2	9.3	9.37	10.05	10.65	9.15	9.20	11.93
8	GN2	6.18	6.06	6.34	7.20	6.47	6.58	8.01
9	RS3	6.21	6.48	7.57	6.99	6.28	6.95	7.40
10	CL3	14.55	15.12	15.17	15.60	14.26	13.86	15.97
11	RQ3	18.49	18.29	20.44	20.40	18.27	19.44	20.91
12	GN3	16.54	15.93	17.89	18.43	16.87	18.16	19.02
13	RS4	9.51	9.81	10.76	11.08	10.07	9.91	11.51
14	CL4	24.56	25.05	24.46	26.00	24.17	23.55	22.45
15	RQ4	11.56	12.08	13.65	13.45	11.77	12.88	13.65
16	GN4	28.98	26.89	28.52	28.04	29.01	27.91	29.02
17	RS5	11.35	12.54	13.48	13.56	11.92	11.52	14.50
18	CL5	13.99	14.20	14.55	13.55	14.32	12.70	16.42
19	RQ5	12.98	13.39	14.67	13.86	13.51	13.44	14.98
20	GN5	12.89	12.58	12.16	13.51	13.07	11.82	14.01
21	RS6	16.85	16.63	17.84	17.02	17.09	15.01	18.99
22	CL6	17.54	18.50	21.26	18.99	18.74	18.29	19.85
23	RQ6	38.15	32.77	38.16	41.00	35.27	32.13	43.09
24	GN6	18.85	17.77	20.84	19.93	19.29	19.12	22.95
25	RS7	22.15	23.36	24.43	24.00	23.27	20.59	25.01
26	CL7	37.89	38.71	39.87	40.23	38.61	34.11	40.34
27	RQ7	29.87	29.97	30.79	32.98	30.55	27.42	34.56
28	GN7	20.14	20.71	21.61	25.67	21.50	21.57	22.49
29	RS8	38.78	37.70	42.38	41.01	39.56	35.98	40.01
30	CL8	41.56	41.49	48.73	44.00	42.89	41.75	45.06
31	RQ8	40.25	41.05	47.43	44.34	43.35	43.20	44.96
32	GN8	45.68	42.91	44.36	48.08	45.41	41.84	49.88
33	RS9	34.25	34.37	34.38	33.56	34.70	29.54	33.80
34	CL9	38.14	38.07	44.45	40.02	39.42	37.05	39.97
35	RQ9	32.54	32.18	31.90	33.67	33.67	29.38	34.01

Table A.2. SNR_p Results