Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
GEOFISICA

Ciclo XXIX

Settore Concorsuale di afferenza: 04/A4
Settore Scientifico disciplinare: GEO/10

THREE COMPONENT SEISMIC ARRAY ANALYSIS IN THE DISCRETE WAVELET DOMAIN

Presentata da: Andrea Angelillo

Coordinatrice Dottorato:
Prof.ssa Nadia Pinardi

Relatore:
Dott. Luca D'Auria

Esame finale anno 2017

INDEX

RINGRAZIAMENTI

Nessun obiettivo importante si raggiunge mai in piena autonomia.

La stesura di questa tesi e il lavoro di ricerca che è stato svolto negli ultimi 3 anni, affrontato parallelamente ad altri importanti ambiti obiettivi personali, hanno richiesto un impegno imponente, con il superamento di tante difficoltà e un grande spirito di abnegazione.

Tutto questo sarebbe riduttivo attribuirlo solo ai miei sforzi. Voglio quindi ringraziare la mia compagna, Lisa, che con grande pazienza, discrezione e dedizione mi è stata sempre vicino in questo tempo, aiutandomi e sUpportandomi e sOpportandomi come poche persone hanno la capacità di fare.

Ovviamente voglio ringraziare mia Madre e mio Padre perché sempre con i loro consigli, ma anche con le loro diversità, sanno essere fonte di ispirazione e di esempio e sono stati capaci di rendermi la persona testarda, determinata, e ambiziosa che sono, dandomi serenità nell'operare le scelte che più mi potessero rendere felice. Ovviamente per ultimo, ma non in ordine di importanza vi è un enorme ringraziamento a Luca, sincero amico e preparatissimo relatore. Con lui ho imparato tanto, e non ho difficoltà a riconoscere in lui uno dei professionisti più preparati che conosco.

Oltre ad avermi trasmesso tanto professionalmente parlando, devo ringraziarlo per l'opportunità che mi offrì 3 anni, con cui ho potuto operare un cambio radicale nella mia vita professionale passata, dandomi anche la possibilità di lavorare questi ultimi 3 nella mia città, vicino alla mia famiglia e alla mia compagna.

INTRODUCTION

The main purpose of this PhD thesis was the development of an innovative methodology for the seismic array data analysis, named DWT-MuSiC (Discrete Wavelet Transform -Multiple Signals Classification).

Seismic arrays are nowadays widely used in geophysics and many methodologies have been developed to obtain as much information is possible for their utilization.

DWT-MuSiC is a new proposed method intended to be able to perform near-real time analysis relating to the detection of different seismic wave field and their characterization, starting from raw seismic array data.

The innovative point of DWT-MuSiC is that it is thought to combine the resolution of the MuSiC (Multiple Signals Classification) algorithm, methodology used for frequency estimation and source location, proposed by Schmidt (1986), and the potentialities of the discrete wavelet domain analysis. The DWT-MuSiC, whose algorithm structure is further discussed in the chapter 3, in fact other than to distinguish the presence of different wavefronts, provides both the direction of arrival of the front themselves (back azimuth) and their apparent speed of advancement (the inverse of the slowness), returning even information about the polarization of each identified phases, preserving furthermore spatial information of the original signals, as well as their amplitude spectrum.

More precisely, the program starts performing a preliminary transformation of the signals that are going to be processed in the wavelet domain by means of discrete wavelet transform. After getting the wavelet coefficients relative to the several frequency bands, and different temporal positions in which the original signals are decomposed, DWT-MuSiC uses the coefficients themselves in order to analyze their covariance between the different sensors that compose the array and obtains the desiderate results.

The analysis of the covariance is made by MuSiC method that, applied in a proper way with a multi dimensional grid search method, permit to distinguish the useful information related to coherent the seismic sources, from the incoherent seismic

noise. In this way, for each time/frequency band interval, it is possible to recognize a certain number of seismic phases related to different seismic sources.

After having distinguish the different seismic phases, and after having characterized them with all the researched parameters, DWT-MuSiC revise the results with the use of a non linear optimization function in order to overcome the limitation of the grid search method resolution.

This thesis is structured in different chapters in order to give an extensive overview about the topic. In particular, in the first chapter is presented a general overview about the seismic arrays, their application in seismology and the basic principle related to their utilization. In the second chapter are introduced the main theoretical concepts involved in the presented methodology, in chapter 3 is explained in detail the structure of the DWT-MuSiC and its operation, while in chapter 4 are presented some synthetic test used to validate the methodology, as well as the comparison with others analysis like the beamforming and the MuSiC method used in the Fourier domain.

In chapter 4 are moreover presented 2 applications to real cases that show some of the potential applications of the methodology in different geophysical contests. The first is an analysis of a volcano-tectonic event registered at Mount Vesuvius, Italy, and the second is the characterization of array data acquired during an active seismic survey at Krafla caldera, Island.

CHAPTER 1: SEISMIC ARRAY BACKGROUND

1.1 SEISMIC ARRAYS OVERVIEW

A seismic array consists of a certain number of seismometers arranged in a well-defined geometric configuration. They are used in geophysics in many fields of application. The installation of the firsts seismic arrays started at the beginning of the 1960s when, for military purposes, it was necessary to have an analytic methodology capable of improve the threshold of detection of underground nuclear tests made worldwide, discriminating at the same time between them and global natural earthquakes [e.g. Douglas et al, 1999].

After this first purpose, since then, seismic arrays started to be used also for civil scientific purposes thanks to its powerful potentiality.

Some example of the nowadays application consist of estimation of the seismic phases slowness vectors [e.g. Shyh-Jeng et al., 1993] as well as locating and tracking volcanic tremor [e.g. Almendros et al., 1997], signals extractions and polarization analysis [e.g. Meersman et al., 2006], characterization of a rupture propagation during an earthquake [e.g. Goldstain et al., 1991] refining small-scale structures in the Earth's interior [e.g. Weber et al., 1996] or high resolution tomographic images on regional scales [e.g. Arlitt et al., 1999].

Another aspect that contributes to the appeal of arrays in geophysics, has been the advent of digital data loggers allowing a fast development of portable arrays, which actually can be easily deployed for experiments of short time duration like the study of seismic noise [e.g. Saccarotti et al., 2001] or time limited experiments aimed at the detection and tracking of the tremor or long period events on active volcanoes [e.g. Saccarotti et al., 2008]. The physical principles behind the use of the seismic array analysis, acting like a directional receivers, is the same as in other field of applications where arrays are used since more time, like telecommunications or radio astronomy or radar science.

This means that in seismology, they can measures the directional properties of the wave-field (slowness or wave number vector) radiated by one or more seismic sources.

A seismic array differs from a local network of seismic stations mainly by the techniques used for data analysis, by which the detectable threshold of seismic signals, with respect to the ambient seismic noise within the Earth, can be significantly enhanced [e.g. Schimmel and Paulssen, 1997].

The main advantage of seismic arrays, in fact, if compared to single seismological stations, is the improvement of the signal-to-noise ratio (SNR) due to the summation of the individual recordings of the array stations.

In addition, the seismic arrays can determine directional information of seismic signals by which is possible to locate the source of a seismic signal by a single array measurement.

This enables to study phases that normally do not show up in seismograms of single stations with amplitudes large enough to study travel times and/or waveforms. This is the primary reason why seismic arrays are very useful in studies of the small-scale surveys.

Besides the large-scale acquisitions that are only possible with traditional seismic stations, many regional and local studies have been made possible with seismic arrays act like a directional receivers [e.g. Inza et al., 2011]. Arrays helped to resolve fine-scale structure well below the resolution level of global seismology in many different places in the Earth, from the crust using body waves and surface waves, the upper mantle, the lower mantle, the core-mantle boundary, and the inner core [e.g. Rothert et al., 2001 ; Vidale et al., 2000].

Different arrays can have different characteristic considering the considering the purpose for which they are used.

One of the most important characteristic of a seismic array is its geometrical configuration, i.e. array aperture, sensors interspacing, number of array stations, etc. A careful choice of these parameters is essential for obtaining good results for the chosen analytical purposes.

Different array designs have been tested in literature, and, depending on their application, their optimal characteristics has been discussed significantly [Schweitzer et al., 2002].

The aimed objective is to obtain, as a response function used to detect the seismic source, a one that present a sharp maximum with a rapid suppression of the energy next to it.

Additionally, a configuration able to minimize spatial aliasing should be taken in account, for example avoiding distances between stations similar to the wave number window of interest. These prerequisites, obviously are dependent on the wavelengths of the seismic phases that are going to be studied.

Therefore seismological arrays are deployed taking into account also the expected wave field frequency content. To fulfill these criteria, the number of array stations, the inter-station spacing, and the configuration of the array can be varied.

To make some examples it is possible to say that the aperture of the array affects the array response in terms of the ability of the array to separate the wave numbers of two incoming wavefronts, and the number of stations controls the quality of the array response, optimizing the signal to noise ratio making possible to detect weaker signals.

The inter-station spacing of the array stations defines the position of the side lobes in the response function and the largest resolvable wave number. The smaller is the inter-station spacing, the larger the wavelength of a resolvable seismic phase will be.

Finally, the geometry of the array controls the azimuthal dependency of the resolution and the quality and the position of the side lobes.

Seismic arrays are designed to exploit the coherence of seismic signals between sensors in order to detect and characterize the impinging wave field. However, signal coherence decreases with increasing distance between measurement locations due to effects that include, but are not limited to, signal multi-pathing, dispersion, and wavefront distortion. Therefore, the design of an array is a balance between ensuring the sensor separations are small enough to guarantee acceptable

signal coherence, yet large enough to provide the required resolution when estimating signal azimuth and velocity.

All these conditions, as explained play a role in the array response quality and as it is possible to imagine, it is not always simple to work with real data acquired in the perfect theoretical condition.

Anyway the most important aspect is that a registered wavefronts has to produce a significant phase change in relation with the array aperture.

The condition is satisfied imposing that the array aperture mast be at least four times greater than the wave length of the seismic phase that we want to analyze. This is expressed by the following relation:

$$\lambda = \frac{V}{f} \leq 4 \cdot Array\ aperture \qquad \textbf{(Eq. 1.1)}$$

Where V is the seismic velocity of the medium and f is the frequency of the seismic wave

In order to avoid spatial aliasing, moreover, the wavelength of a wavefront has to be at least comparable with the array interspacing. This condition is described by the equation:

$$\lambda = \frac{V}{f} \geq Array\ interspacing \qquad \textbf{(Eq. 1.2)}$$

These conditions are considered to the prerequisites for being able to perform an array analysis, and are valid for all the methodologies involved in the array analysis and further, more strictly conditions can be adopted depending on the methodology involved.

1.2 THREE COMPONENT SENSORS AND SEISMIC WAVE POLARIZATION

Because of the elastic properties of Earth materials and the presence of surface boundary, different types of seismic waves propagate within it.

Compressional (P) and Shear (S) waves propagate through the Earth's interior and are known as body waves. Love and Rayleigh waves propagate primarily at and near the Earth's surface and are called surface waves.

Different wave types produce different oscillations of the medium within them are travelling and this schematically as represented in figure 1.1 .

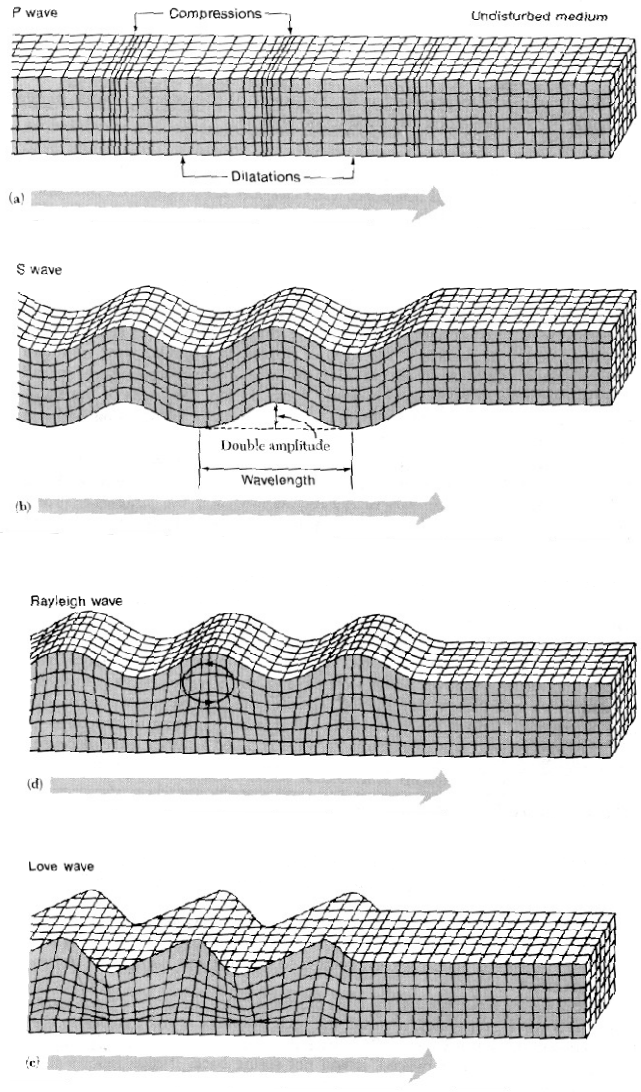


FIGURE 1.1 *Representation of the disturbance caused by the passage of a P-wave(a), S-Wave(b), a Rayleigh wave(c) and a Love-wave(d)*

The P-waves, figure 1.1(a), create a disturbance that is propagated as a compression and dilatation of the material, generated in the direction of propagation. The particle motion related to an S-wave propagation, figure 1.1 (b), on the contrary, show shear motion directed horizontally to the direction of propagation.

Regarding the surface waves, a Rayleigh-waves propagation through a volume of elastic material, Fig.1.1(c), cause a disturbance that generates an elliptical motion caused by the combination of both a vertical, (perpendicular to the direction of propagation but in the plane of the ray path), and horizontal (in the direction of propagation) particle motion. The last type of wave, Love-waves, Fig. 1.1(c) cause a disturbance that is propagated horizontally and perpendicularly to the direction of propagation. Considered what written before, it is clear that the use of P-waves seismic data alone may not be enough to characterize completely a wavefront that impinge a seismic array.

Having the possibility to identify different seismic components is really important to distinguish different structure present in the subsurface.

When a P-wave arrives at subsurface rock interfaces at non-normal angles of incidence, in fact, a conversion of P-wave takes place, generating an S wave,.

S waves can be composed by a horizontal (SH) and a vertical (SV) waves causing the rock particles to oscillate perpendicular to the direction of the propagating wavefront and orthogonal to each other as well. These three different components of the seismic reflected wavefront can be recorded with sensors that recognize the full particle motions, and are called multi component or three component sensors

Geophones used for conventional seismic data acquisition are constrained to respond to just one component, i.e. the vertical component, but the multi-component sensors have the motion sensing elements arranged in a single casing and are used for recording the complete seismic wave field. Orientation of sensors is also an important aspect in multi component seismic acquisition, because the processing depends on reliable information about geophone direction and polarity. Current techniques of acquisition generally keep all the geophones in the same predefined orientation such that all the axial and transverse components maintain the same polarity and direction.

Most modern seismometers include a configuration to measure three orthogonal separate elements that allow the determination of the simultaneous movement in three different directions; two horizontal, aligned east-west and north-south, and the third vertical. This configuration is also takes in consideration in this thesis.

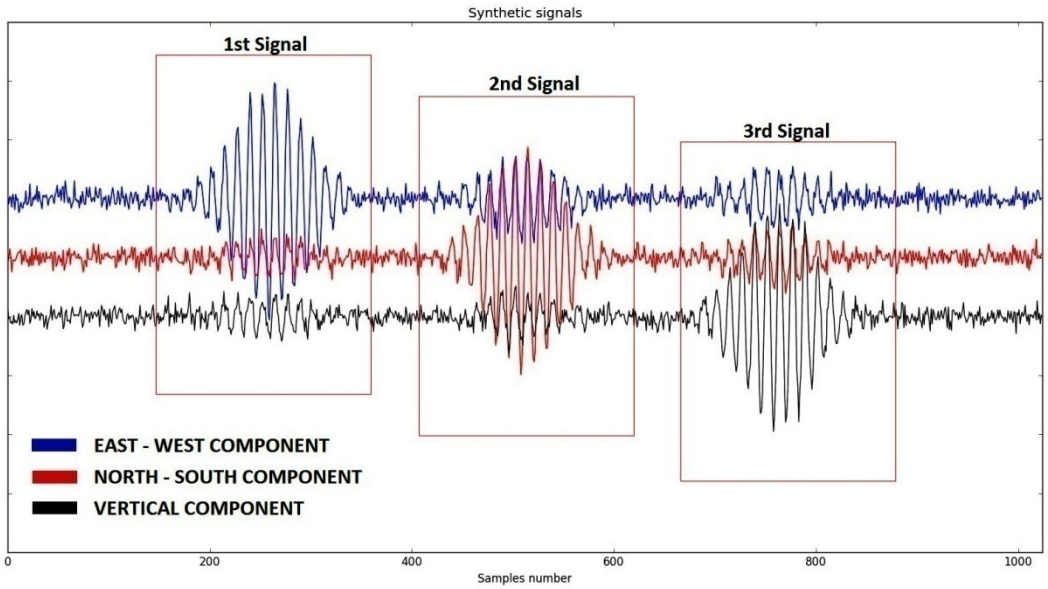


FIGURE 1.2 *The image show the 3 different traces registered by a 3 component seismic sensor, with different signals amplitudes dependent on the wavefront polarization*

In Figure 1.2 is presented for example a synthetic seismic signal that well show the registration of the 3 different components and the relative amplitude produced caused by the orientation of the wavefront in respect to the station, and its polarization pattern.

1.2 DIRECTION OF ARRIVAL AND SLOWNESS CHARACTERIZATION

The majority of array seismology methods assume a plane wave arriving at the array. This assumption is valid for distances from the source much larger than about 4 wavelengths [e.g. Rost et. al.,2002].

$$Distance\ source \gg Array\ aperture \qquad \textbf{(Eq. 1.3)}$$

The directions of approach and propagation of the wavefront projected onto the horizontal plane are defined by the following parameters:

1) **Φ Backazimuth** = angle of wavefront approach, measured clockwise from the North to the direction pointing towards the source.

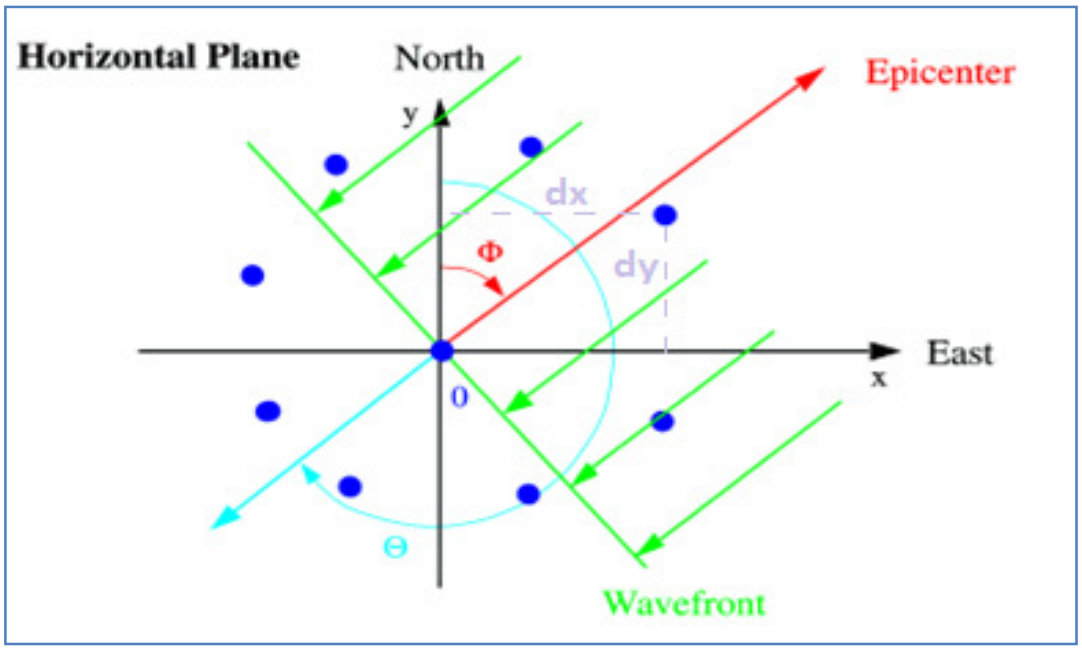


FIGURE 1.3 *Schematic representation of the back azimuth angle of an impinging wavefront.*

2) **i vertical incidence angle** = vertical angle of wavefront approach, measured from the vertical 0°<i< 90°

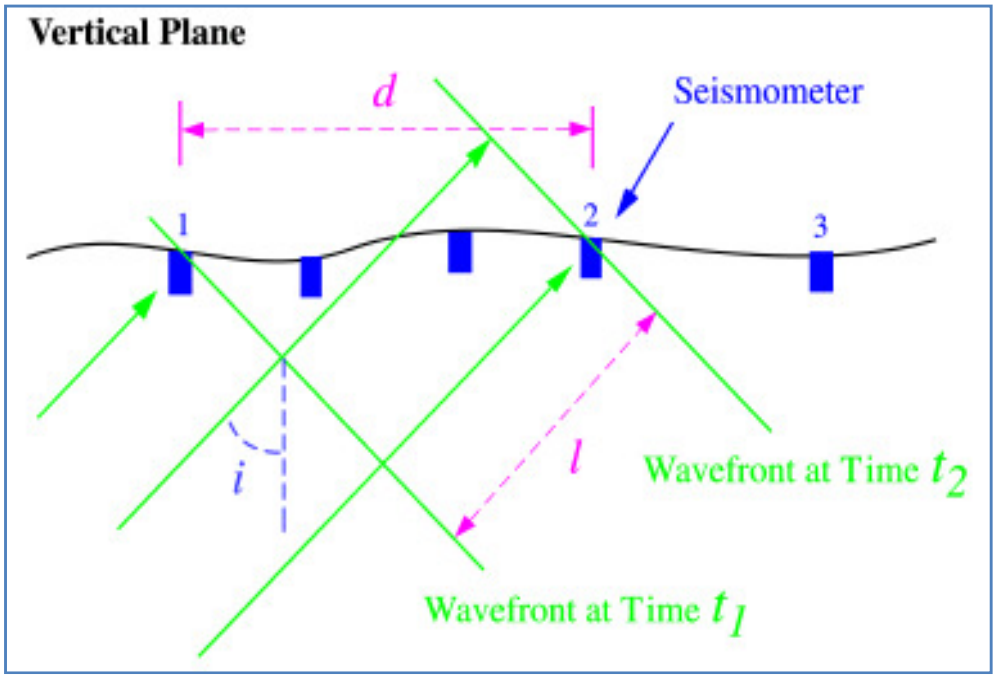


FIGURE 1.4 *Schematic representation of the vertical incidence angle of an impinging wavefront.*

3) **Slowness u**= the inverse of the propagation velocity of the wavefront across the array $1/v_0$.

$$\mathbf{u}= (ux, uy, uz) = \left(\frac{\sin(\theta)}{v_o}, \frac{\cos(\theta)}{v_o}, \frac{1}{v_o \tan i}\right) =$$

$$= \frac{1}{v_0}(\sin i \sin\theta\,, \sin i \cos\emptyset, \cos i) \qquad \textbf{(Eq. 1.4)}$$

The slowness vector **u** points toward the direction of the wave source, and its modulus is the reciprocal of the wave speed.

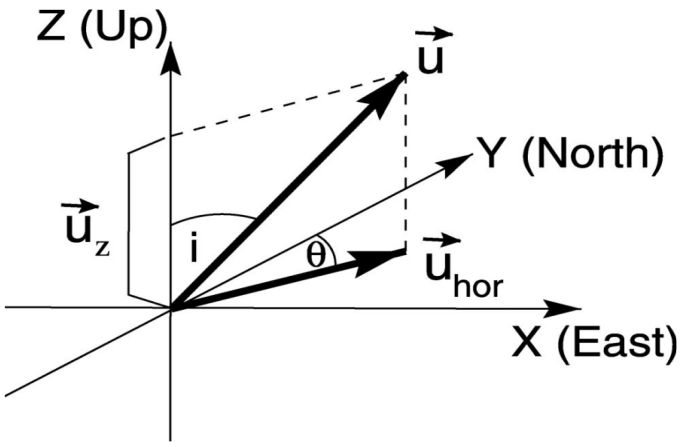


FIGURE 1.5 *Representation of the slowness vector and its 3 component.*

Although it is possible to install seismic array sensors at different depths (i.e. in wells), generally a seismic array stations are all placed on the ground level, in a flat area, to form a two-dimensional planar geometry. That is the difference between 2-dimensional and 3-dimensional arrays. Roughly speaking, in array analysis we measure the difference of the arrival times or the phase difference at various stations. It is then enough clear that in a 2-dimensional array it is possible to experience the same delay in the arrival time at different stations for various configuration of slowness and incidence angle. From this is clear that the biggest limitation of a two–dimensional array is that is not possible to reconstruct the real slowness of the incident wavefront but only a component of it called apparent slowness. Considering that the apparent speed is defined as follow:

$$V_{app} = \frac{V_0}{\sin(i)} \qquad \textbf{(Eq. 1.5)}$$

where V_0 is the real wave velocity beneath the array.

Apparent slowness vector **u** is defined as follows:

$$\mathbf{u} = (ux, uy) = \left(\frac{\sin(\theta)}{v_{app}}, \frac{\cos(\theta)}{v_{app}}\right) = \frac{1}{v_0}(\sin i \sin\theta, \sin i \cos\theta) \qquad \textbf{(Eq. 1.6)}$$

1.4 BEAMFORMING

Beamforming [Barlett,1948], is one of the most basic and common methodologies used in array data processing, aimed at determining the slowness and the back azimuth of a seismic phase.

Beamforming technique is not only used in seismology, but being based on a relative simple principle, has a wide application also in radar analysis, wireless communications, astronomy or acoustic problems [Brooks et al., 2006 ; Steyskal, 1987].

The principle at the basis of the beamforming analysis is to detect the slowness and the back azimuth of a certain seismic phase by improving the signal-to-noise ratio (SNR) of the stacked traces registered by the array seismic sensors. After shifting in time the traces of the different sensors till obtaining a perfect alignment, summing them all together, is possible to maximize the constructive interference related to the coherent signals, reducing at the same time the incoherent noise.

In order to do that, the operation to align the seismic traces to make them comparable before summation, is only possible after having found the right values of delay time, specific for each single sensor.

The delay time is defined as the extra time needed by a seismic wavefront to reach the different seismic sensors that compose an array. For a certain number of stations, with known coordinates, the delay time is only dependent on the apparent slowness and back azimuth angle of the wavefront.

Considering a generic array, with a geometric configuration like the one represented in Fig.1.6, it is possible to express the position of all the sensors, relative to a reference origin point.

Generally the reference system is placed in correspondence of a specific station having a central position in the array, or associated to the array geometric center.

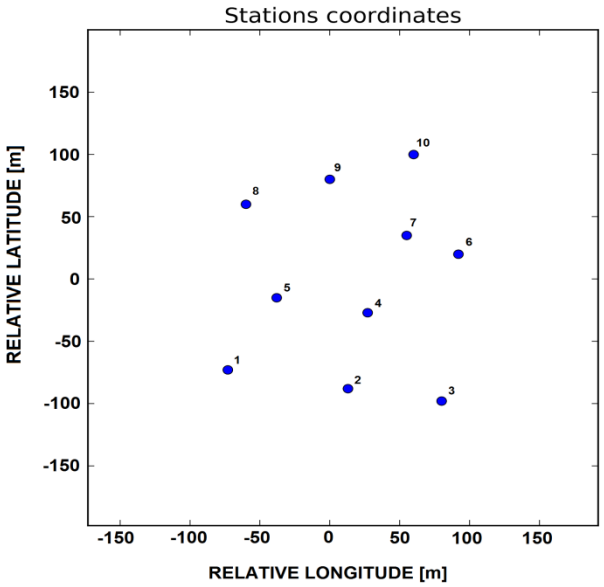


FIGURE 1.6 *Map view of the spatial distribution of the station composing a computer created seismic.*

Assuming a plane wave approximation, being the distance of the source much larger than the array aperture, If we define O (x_o,y_o,z_o) as the center of the reference system, where (x,y,z) are the Cartesian coordinates in [m] with positive axes towards East (x), towards North (y), and vertically (z), a seismic signal registered at the station o is defined as:

$$x_o(t) = f(t) + n_o(t) \qquad \textbf{(Eq. 1.7)}$$

Where *f(t)* is the signal n_o is the seismic noise at station *j-th*(x_j,y_j,z_j) having a distance $\boldsymbol{r_j}$ from the center *O*, the registered trace can be expressed as:

$$x_j(t) = f\left(t - \mathbf{r_j} \cdot \mathbf{u}\right) + n_j(t) \qquad \textbf{(Eq. 1.8)}$$

The same signal, after removing the delay time is expressed by the following relation:

$$\tilde{x}_j(t) = f(t) + n_j(t + \mathbf{r_j} \cdot \mathbf{u}) \qquad \textbf{(Eq. 1.9)}$$

The beamforming function is then:

$$Beam(t) = f(t) + \frac{1}{M}\sum_{j=1}^{M} n_j(t + \mathbf{r_j} \cdot \mathbf{u}) \qquad \textbf{(Eq. 1.10)}$$

where *M* is the number of the stations.

An example of the beamforming methodology is applied to some synthetic signals showed in Fig 1.6. The synthetic signals are related to a P wave crossing the array indicated in figure Fig.1.5 and the wavefront characteristics are: frequency 15 Hz, speed of 350 m/s and back azimuth angle of 45°.

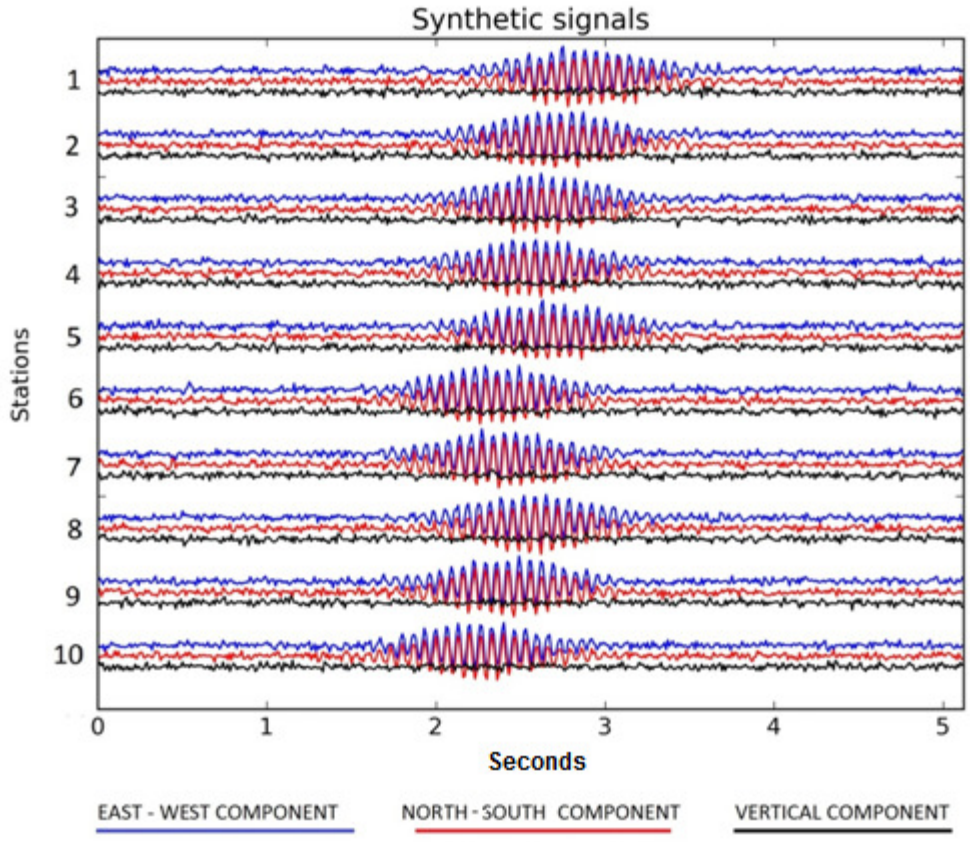


FIGURE 1.7 *Synthetic seismic signals showing a time delay due to the different registration time at the different sensors.*

The delay of the signals through the stations, clearly visible in Figure 1.7, is corrected for the wavefront slowness and back azimuth, obtaining a perfect alignment of the seismic traces as shown in Figure 1.8.

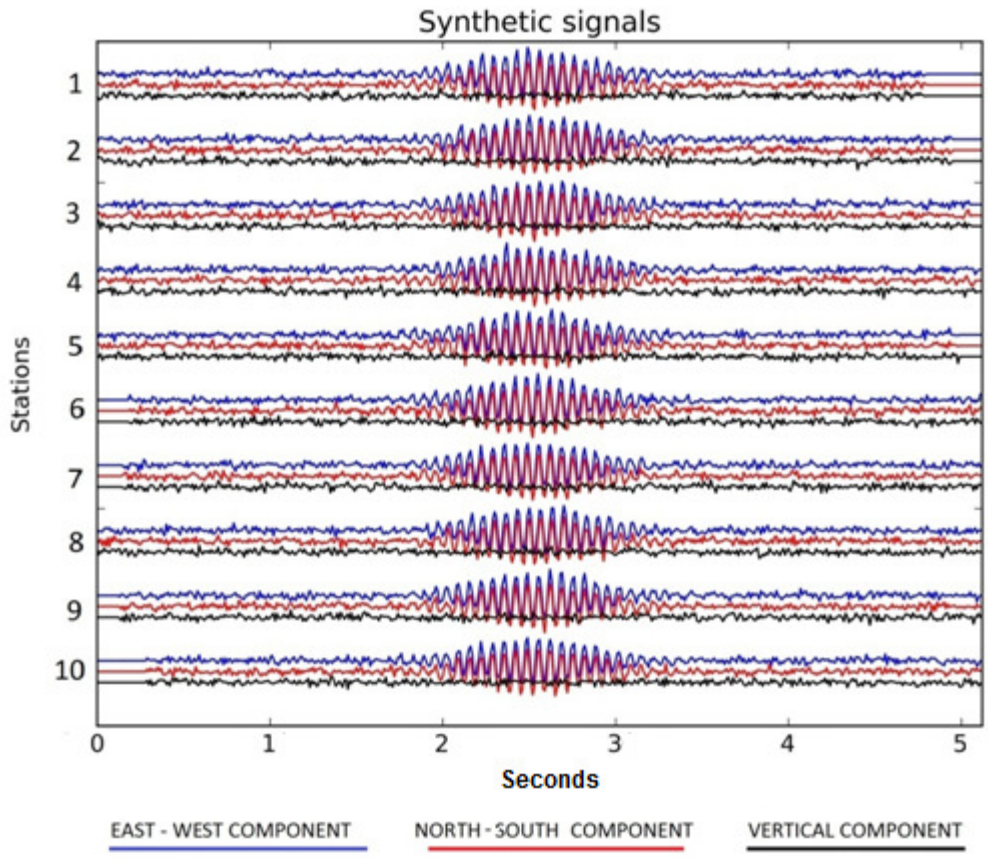


FIGURE1.8 *Synthetic seismic signals after the time alignment*

If we assume that the noise present in the traces has a zero mean value and a variance σ^2, it is possible to sum the aligned signals obtaining , for an array with M stations, an improvement of the signal-to-noise in comparison with the signal-noise ration of a single station (s) specified by:

$$SNR = \sqrt{M}\, s \qquad \textbf{(Eq. 1.11)}$$

where M is the number of the stations.

This is done assuming perfectly coherent signals f(t) at every array station, and completely uncorrelated noise $n_j(t)$.

The results can be observed in Fig. 1.9 showing the stacked signals.

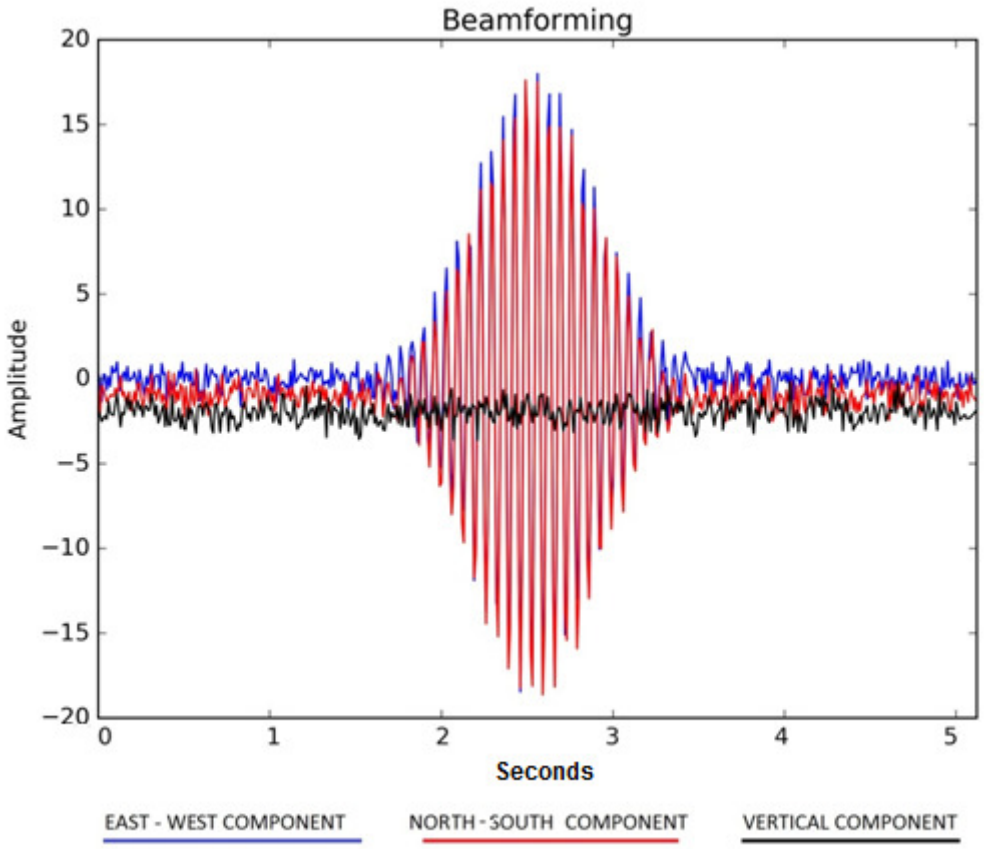


FIGURE 1.9 *Beamforming of the signals obtained summing all the seismic aligned previously aligned.*

To obtain a great accuracy of the results, it is important to reconstruct as precisely as possible the complete slowness vector of a particular seismic phase. Beamforming is a reliable methodology even if the theoretical results expected can be influenced negatively by other factors like local variation of wave speeds beneath the array stations. Considering the simplicity on implementation and the few computational resources needed to perform the analysis, Beamforming continues to be one of the most used techniques for array analysis

Some examples of application of the beamforming methodology, compared to the DWT-MuSiC results, are presented in chapter 4, while in the following chapter we will introduce the main theoretical background of the latter, to better understand the principle on which the latter is based.

CHAPTER 2 : THEORETICAL BACKGROUND

2.1 INTRODUCTION TO THE RELEVANT THEORIES

Seismic arrays have shown their potentiality in application in a variety of sectors. Consequently much attention has been given to arrays in geophysics especially in studies detecting plane-wave signals and their relative slowness vectors and backazimuth [e.g. Kennett et al., 2003].

In the scientific literature there are numerous different techniques which, using different approaches, are used to analyze arrays seismic data. Each of them has different advantages and limitations.

One of the first and basic approach used to characterize the wave field in terms of backazimuth and slowness, consisted to carrying out space-time processing of the array data, estimating the parameters using a cross correlation method between signals.

One of the simplest techniques using this approach is the beamforming method, already explained in paragraph 1.4.

A numerous of alternative methods for enhancing the results obtained by the traditional beamforming have also been proposed to improve the resolution in detecting closely spaced signal sources [Van Veen, Buckley, 1998; Frost,1972].

The main advantage of the time domain methods is the relative simplicity of analysis and the small computational resources that in general they needs. The main limitation is that, analyzing all in once the entire bandwidth of the data, it is impossible to separate the single source contribution, eventually giving inaccurate results caused by reciprocal interference of the sources.

The limitations of the time domain approaches brought to the development of new different methodologies based on the analysis of the frequency content of the signals. Also in this case numerous methods have been proposed. Among the most commonly used is the Capon maximum likelihood technique [Capon, 1969] .

Regarding the methods based on frequency domain analysis, it is possible to affirm that even if in theory they are capable to detect multiple signals sources having different frequency contribution, in practice in seismology that not always can be effectively applied due to a strong non-stationary and transient nature of the signals. In fact these methods were designed to face with nearly monochromatic signals.

Another solution consist in the introduction of subspace-based estimation techniques that offered new possibilities in the sensor array signal processing.

The subspace-based approach relies on certain geometrical properties of the assumed data model, resulting in a high resolution capability providing that experimental data accurately reflects the model assumptions.

Pisarenko (1973) was one of the first to exploit the structure of the data model and from his approach derives the MuSiC method developed by Smith. MuSiC is one of the most efficient techniques applied to array analysis and used to estimate the backazimuth of signals sources and it is widely applied in literature, not only in seismology but also in other field of application .

In its original form, as proposed by Smith in (1977), MuSiC methodology was applied to electromagnetic wave time series data and it is shown its capabilities to resolve multiple closely spaced sources, being selectively sensitive to the strongest ones.

The MuSiC method, in its original form, assumes that the observed signals were stationary in time, and its constituent sources uncorrelated. While these assumptions may apply to different fields of applications, as for example in telecommunication, in seismology that assumptions are generally no more valid. The majorities of the seismic signals, are transient and arrive simultaneously from different directions (e.g. scattered waves).

The innovation introduced in this thesis has been to provide a methodological improvement that made the characterization of the wavefront more precise and complete adopting a generalization of the MuSiC approach.

The creation of the DWT-MuSiC is in fact an adaptation of the original MuSiC with the basic idea to decompose the signals in different time/frequency intervals using the discrete wavelet transform, and then analyze them one interval per time. In this way it is possible to consider each single interval, as the signal is stationary which was otherwise not feasible by only using MuSiC as proposed in the original version.

The use of wavelet transform in fact is particularly useful for the analysis of transients, aperiodic and non-stationary signal features where, through its use, subtle changes in signals morphology may be highlighted over the scale of interest.

Wavelet analysis is moreover particularly valuable because of its ability to preserve simultaneously the spectral and temporal information from an original signal by employing a window of variable width.

Thus wavelet transforms produce a time–frequency decomposition of the signal which separates individual signal components more effectively comparing to more traditional methods like short time Fourier transform (STFT). This flexible temporal–spectral aspect of the transform allows a local scale-dependent spectral analysis of individual signal features. In this way both short duration / high frequency and longer duration / lower frequency information, can be captured simultaneously. The use of the DWT-MuSiC, other than to distinguish the arrival of different wavefronts, uses also the MuSiC method to provide information about the polarization of each identified phases.

The main limitation of the methodology is that for each analysis performed for a single time-frequency interval, all the signal information is contained in one single wavelet coefficient and this makes difficult to discriminate among more than one seismic phase for each interval.

A detailed explanation of the DWT-MuSiC algorithm structure is in chapter 3, while in the next paragraphs we introduce the basic theoretical elements of the MuSiC method and an overview of the Wavelet analysis theory.

2.2 MUSIC ANALYSIS

To describe the MuSiC theory, we start supposing to receive at N sensors, a number of q plane waves with the same angular frequency ω. The signal received at the station $\mathbf{x}_i$ can be written as:

$$\psi(\mathbf{x_i}, t) = \sum_{m=1}^{q} A_m \, e^{i(\mathbf{k_m} \cdot x_i - \omega t + \phi_m(t))} + \eta(\mathbf{x_i}, t) \qquad \textbf{(Eq. 2.1)}$$

where $\eta(\boldsymbol{x}_{i,}t)$ is the noise, $\boldsymbol{k}_m$ is the wave vector of the m-th signal and φ_m is the phase information of the m-th signal. The covariance R_{ij} relative to the stations x_i and x_j can be then be defined as:

$$\mathrm{R_{ij}} = \langle \psi(x_i, t) \psi^{\dagger}(x_i, t) \rangle_t \qquad \textbf{(Eq. 2.2)}$$

where$\langle\rangle_{\mathbf{t}}$ is the time average and †is the Hermitian conjugate;
When the signals are stationary, the correlation of the signals received at the station x_i and x_j, is expressed by the element of the covariance matrix as:

$$\mathrm{R_{ij}} = \sum_{m=1}^{q} |A_m|^2 \, e^{i\mathbf{k_m} \cdot (x_i - x_j)} + \sigma^2 \delta_{\mathrm{ij}} \qquad \textbf{(Eq. 2.3)}$$

Where σ^2 is the noise intensity and A is the amplitude of the m-th signal.
We define the total signals vector as:

$$\mathbf{\Psi(t)} = [\psi(\mathbf{x}_1, t), \psi(\mathbf{x}_2, t), \psi(\mathbf{x}_N, t)]^T \qquad \textbf{(Eq. 2.4)}$$

where T means the transpose operator, and considering also the **(2.2)** and (**2.3),** we have that the signal vector can be expressed as follows:

$$\Psi(t) = \sum_{m=1}^{q} A_m \, \mathbf{u}(\mathbf{k}_m) e^{-i(\omega t + \phi_m(t))} \quad \textbf{(Eq. 2.5)}$$

where the directional information is contained in the term $\mathbf{u}(\mathbf{k}_m)$ and is defined as:

$$\mathbf{u}(\mathbf{k}_m) = [e^{i\mathbf{k}_m \cdot \mathbf{x}_1}, \ldots, e^{i\mathbf{k}_m \cdot \mathbf{x}_N}]^{\mathbf{T}} \quad \textbf{(Eq. 2.6)}$$

Considering the **(2.2), (2.3), (2.4), (2.5) and (2.6),** the covariance matrix can be then written as:

$$\boldsymbol{R} = \sum_{m=1}^{q} |A_m|^2 \, \mathbf{u}(\mathbf{k}_m) \otimes \mathbf{u}^{\dagger}(\mathbf{k}_m) + \sigma^2 \mathbf{I} \quad \textbf{(Eq. 2.7)}$$

Or in matrix notation:

$$\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{U}^{\dagger} + \sigma^2 \mathbf{I} \quad \textbf{(Eq. 2.8)}$$

Where **U** is the matrix of the spatial signal signature formed using the concatenation of the N directional vectors of the q sources and having dimension [N x q], **I** is the identity matrix and **S** is the matrix containing the intensities $|Am|^2$ of the vectors $\mathbf{u}(\mathbf{k}_m)$. $\sigma^2\mathbf{I}$ is the noise power matrix.

Now presuming that the condition q < N is verified, the covariance matrix **R** has rank q. We can proceed decomposing it into eigenvalues and eigenvectors.

After the decomposition we obtain a series of N eigenvalues and the associated eigenvectors that must be separated into two orthogonal subspaces: the noise subspace and signal subspace. The signal space will be identified by the largest eigenvalues (q) while the other eigenvalues N-q will be relative to the noise subspace. It is fundamental to highlight that the N-q eigenvectors of **R** are orthogonal to the spatial signals vectors $\mathbf{u}(\mathbf{k}_m)$.

Defining $\mathbf{E_s}$ the matrix [N x q], collinear to the directional vectors and associated with the q largest eigenvalues, and the $\mathbf{\Lambda}_s$ the diagonal matrix [q x q] of the relative eigenvectors, and defining $\mathbf{E_s}$[N x N-q] the eigenvalues matrix associated with the N-q eigenvalues (subspace noise) orthogonal to the spatial signals vectors $\mathbf{u}(\mathbf{k}_m)$, with the associated eigenvectors diagonal matrix [q x q] $\mathbf{\Lambda}_n$, then the covariance matrix can be rewritten as:

$$\mathbf{R} = \mathbf{E_s}\mathbf{\Lambda_s}\mathbf{E_s^\dagger} + \mathbf{E_n}\mathbf{\Lambda_n}\mathbf{E_n^\dagger} \qquad \textbf{(Eq. 2.9)}$$

Where $\mathbf{E_s}\mathbf{\Lambda_s}\mathbf{E_s^\dagger}$ is the contribution to the covariance matrix relative to the signals and $\mathbf{E_n}\mathbf{\Lambda_n}\mathbf{E_n^\dagger}$ is the contribution relative to the noise.
Afterwards, it is necessary to find a set of steering vectors to project them on the noise space:

$$\mathbf{a(k)} = [e^{i\mathbf{k}_m \cdot \mathbf{x}_1}, \dots, e^{i\mathbf{k}_m \cdot \mathbf{x}_N}] \qquad \textbf{(Eq. 2.10)}$$

To estimate the signal direction vectors $\mathbf{u}(\mathbf{k}_m)$ we have to find the steering vectors that give the minimum projection in the noise subspace.
To do this operation, we define the MuSiC estimator function as follows:

$$\mathbf{D(k)} = \frac{\mathbf{1}}{|\mathbf{a(k)} \cdot \mathbf{E_n}|^2} \qquad \textbf{(Eq. 2.11)}$$

Once the spectrum of the function D(k) has been correctly determined using all the set of steering vectors, than the research of the maximum value of the function can be performed with a grid search method.

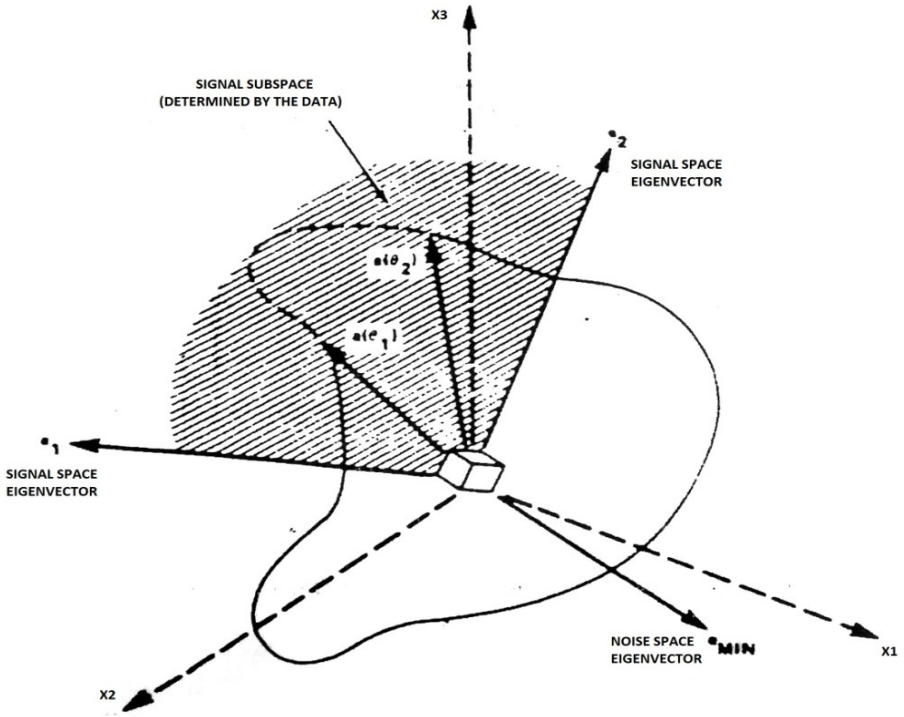


FIGURE 2.1 *Schematic representation of the signal and noise subspace spanned by some steering vector projecting their image on noise subspace.*

2.3 WAVELET ANALYSIS

In signal analysis Fourier transform is only able to provide information about the frequency composition of a given signal but is not useful to reveal the temporal localization of a signal. Another limitation of Fourier transform is that the sin(x)and cos(x) in which the original function is transformed are infinite and periodic functions. These functions are smooth and do not adapt to the sharp changes of the input signal.

This limitation can be partially solved dividing the original signal with a sliding time windows of fixed dimension to keep the information about the temporal localization. This solution is known as short time Fourier transform and provides a certain temporal resolution by highlighting the spectral response for each time interval.

Wavelet analysis represents a much powerful solution, to analyze data maintaining the time and spectral information and for its potentiality is widely used in several scientific fields like, engineering, pattern recognition and geophysics.

Wavelets are oscillatory functions of short durations and that is the reason of their name. There are many wavelets to choose from; however, by far the most popular are the Mexican hat wavelet, the Morlet wavelet or the Coiflet wavelet. (Fig. 2.2)

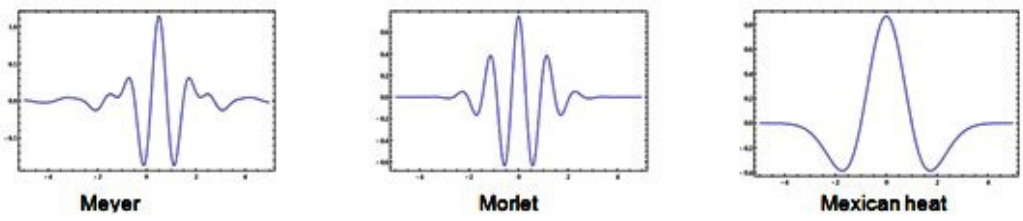


FIGURE 2.2 *Representation of 3 of the most common used wavelet form.*

In wavelet analysis, we use an integration mechanism similar to that of Fourier Transform to compute the wavelet coefficients:

$$C(scale, shift) = \int_{-\infty}^{\infty} x(t)\psi(scale, shift)dt \qquad \textbf{(Eq. 2.12)}$$

The idea is to take a wavelet ψ, that we call mother wavelet, and scale and shift it to create a new wavelet. Then we measure how well a similar size segment on the original signal fits the wavelet shape. This results in some coefficients that represent the level of adaptation. The larger the absolute value the better is the match.

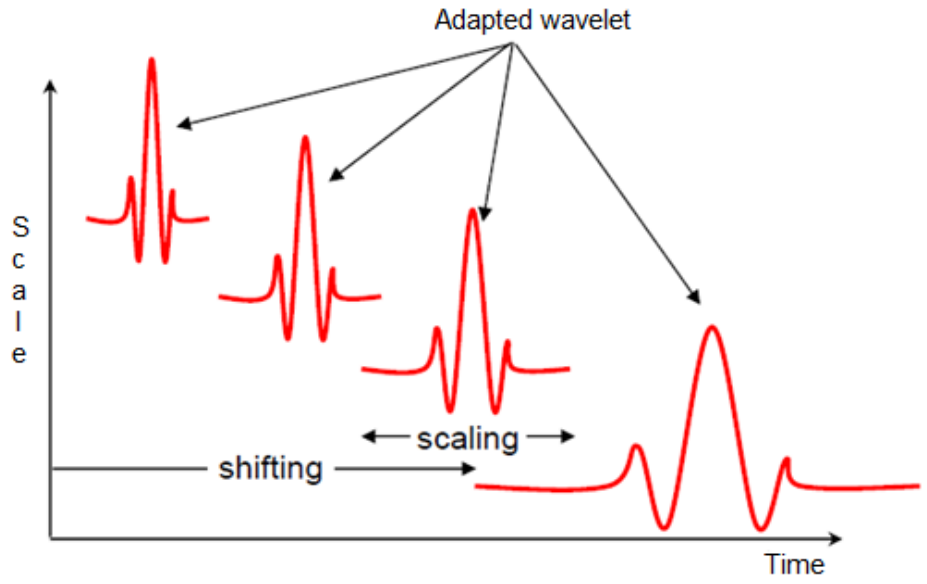


FIGURE 2.3 *Representation of a mother wavelet scaled and shifted in order to perform the wavelet transform analysis.*

Wavelet analysis is particularly useful for the analysis of transients, aperiodic and non-stationary signals because it allows generating a time–frequency decomposition of the signal which separates individual signals more effectively than the traditional short time Fourier transform. In this way both short duration, high frequency and longer duration, lower frequency information can be appreciated simultaneously.

Wavelet transforms as they are defined come in two distinct types: the continuous wavelet transform and the discrete wavelet transform.

The continuous wavelet transform of a time signal, **x(t)**, is defined as:

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi * \left(\frac{t-b}{a}\right) \mathrm{d}t \qquad \textbf{(Eq. 2.13)}$$

where ψ*(t) is the complex conjugate of the analyzing wavelet function ψ(t), *a* is the dilation parameter of the wavelet and *b* is the location parameter of the wavelet.

In order to be classified as a wavelet, a function must satisfy certain mathematical criteria. The conditions are:

Wavelets must have finite energy:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 \mathrm{d}t \leq \infty \qquad \textbf{(Eq. 2.14)}$$

If $\hat{\boldsymbol{\psi}}(\boldsymbol{f})$ is the Fourier transform of $\boldsymbol{\psi}$ (*t*), i.e.

$$\hat{\psi} = \int_{-\infty}^{\infty} \psi(t) e^{-i(2\pi f)t} \mathrm{d}t \qquad \textbf{(Eq. 2.15)}$$

Then the following condition must hold:

$$C_g = \int_0^{\infty} \frac{\left|\hat{\psi}(f)\right|^2}{f} \mathrm{d}f < \infty \qquad \textbf{(Eq. 2.16)}$$

This implies that the wavelet must have a zero mean.

The contribution to the signal energy at the specific *a* scale and *b* location is given by the two-dimensional wavelet energy density function defined as:

$$E(a,b) = |T(a,b)|^2 \qquad \textbf{(Eq. 2.17)}$$

Peaks in E(a) highlight the dominant energetic intervals within the signal.

For a certain transformed signal, the original signal may be reconstructed using an inverse wavelet transform, defined as follows:

$$x(t) = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_{0}^{\infty} T(a,b) \psi_{a,b}(t) \frac{\mathrm{d}a\mathrm{d}b}{a^2} \quad \textbf{(Eq. 2.18)}$$

The continuous wavelet transform can be calculated over an arbitrarily fine time–frequency grid. This characteristic involves a great disadvantage because the CWT involves a lot of computational resources. This is limiting and moreover unnecessary because a vast amount of repeated information is contained within this redundant representation of the continuous wavelet transform *T (a, b)*.

For this reason to save computational resources, often it is used the discrete wavelet transform (DWT), in which the wavelets are discretely sampled over a dyadic grid.

In its most common form, the DWT employs a dyadic and orthonormal wavelet basis functions and exhibits zero redundancy. The transform integral in fact is determined only on a discrete grid of *a* scales and *b* locations. In practice, the input signal is treated as an initial wavelet approximation to the underlying continuous signal from which, using a multi resolution algorithm, the wavelet transform and inverse transform can be computed discretely, quickly and without loss of signal information.

A natural way to sample the parameters *a* and *b* is to use a logarithmic discretization of the *a* scale and link this, in turn, to the size of steps taken between *b* locations. To link *b* to *a* we move in discrete steps to each location *b*, which are proportional to the *a* scale. This kind of operation can be expressed as:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} \psi\left(\frac{t - n b_0 a_0^m}{a_0^m}\right) \quad \textbf{(Eq. 2.19)}$$

where the integers *m* and *n* control the wavelet dilation and translation respectively; a_0 is a specified fixed dilation step parameter and b_0 is the location parameter.. This power-of-two logarithmic scaling of both the dilation and translation steps is known as the dyadic grid arrangement. The dyadic grid is the

simplest and most efficient discretization for practical purposes and lends itself to the construction of an orthonormal wavelet basis.

Discrete dyadic grid wavelets are orthogonal to each other and are normalized to have unit energy. This is expressed as:

$$\int_{-\infty}^{\infty} \psi_{m,n}(t)_{m',n'}(t)\mathrm{d}t = \begin{cases} 1 & if\ \ m = m' and\ n = n' \\ 0 & otherwhise \end{cases} \quad \textbf{(Eq. 2.20)}$$

With the DWT, you always end up with the same number of coefficients as the original signal samples, but many of the coefficients may be close to zero in value. As a result, you can often throw away those coefficients and still maintain a high-quality signal approximation. Using the dyadic grid derived from equation **(2.18),** the discrete wavelet transform (DWT) can be written as:

$$T_{m,n} = \int_{-\infty}^{\infty} x(t)\psi_{m,n}(\mathrm{t})\mathrm{d}t \quad \textbf{(Eq. 2.21)}$$

Where $\boldsymbol{T_{m,n}}$ is known as the coefficient at scale and location indices ***(m, n)***.

For the DWT the transform integral remains continuous but is determined only on a discretized grid of *a* scales and *b* locations.

Orthonormal dyadic discrete wavelets are associated with scaling functions and their dilation equations.

The scaling function is associated with the smoothing of the signal and has the same form as the wavelet, given by

$$\phi_{m,n}(t) = 2^{-\frac{m}{2}}\phi(2^{-m}t - n) \quad \textbf{(Eq. 2.22)}$$

They have the property:

$$\int_{-\infty}^{\infty} \phi_{0,0}(t)\, dt = 1 \quad \textbf{(Eq. 2.23)}$$

where $\varphi_{0,0}(t)=\varphi(t)$ is referred to as the father wavelet or scaling function. The father wavelet is orthogonal to translations of itself, but not to dilations of itself. The scaling function can be convolved with the signal to produce approximation coefficients as follows:

$$S_{m.n} = \int_{-\infty}^{\infty} x(t)\phi_{m,n}(t)\,\mathrm{d}t \quad \textbf{(Eq. 2.24)}$$

A signal **x(t)** can then be represented using a combined series expansion using both the approximation coefficients and the detail coefficients as follows:

$$x(t) = \sum_{n=-\infty}^{\infty} S_{m_0.n}\phi_{m_0.n}(t) + \sum_{n=-\infty}^{m_0}\sum_{n=-\infty}^{\infty} T_{m.n}\psi_{m,n}(t) \quad \textbf{(Eq. 2.25)}$$

We can see from this equation that the original continuous signal is expressed as a combination of an approximation of itself, at arbitrary scale index m_0, added to a succession of signal details from scales m_0 down to negative infinity.

The signal detail at scale m is defined as:

$$d_m(t) = \sum_{n=-\infty}^{\infty} T_{m.n}\psi_{m,n}(t) \quad \textbf{(Eq. 2.26)}$$

Hence we can write equation (2.25) as:

$$x(t) = x_{m_0}(t) + \sum_{n=-\infty}^{\infty} d_m(t) \quad \textbf{(Eq. 2.27)}$$

From the equation (2.27) is possible to show that:

$$x_{m-1}(t) = x_m(t) + d_m(t) \quad \textbf{(Eq. 2.28)}$$

which tells us that if we add the signal detail at an arbitrary scale (index m-1) to the approximation at that scale we get the signal approximation at an increased resolution. This is called a multi-resolution representation.

The numerical computation of the DWT can be made very rapidly by using an algorithm (fast DWT) which determines the DWT coefficients at each scale through repeated iterations.

The algorithm, starting from an input signal, produces two different sets of coefficients: approximation coefficients cA_1, and detail coefficients cD_1. These vectors are obtained by convolving the input signal with the low-pass filter function for approximation, and with the high-pass filter function for detail. This operation is repeated continuing splitting the approximation coefficients cA_1 in two parts using the same scheme, replacing the input signal by cA_1, and producing new set of coefficients, $cA_2$and cD_2 relative to the second scale of decomposition, and so on.

Here follows a schematic representation of the algorithm after which utilization, the signal will be represented by the following structure having a tree shape:

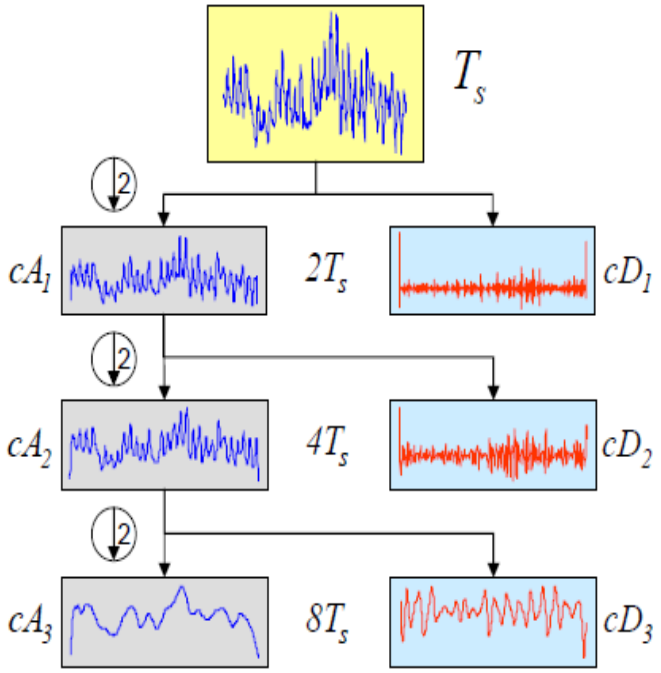


FIGURE 2.4 *The image show the multi resolution wavelet decomposition performed at different scales and shows for each scale the relative approximation and detail coefficients obtained*

After having introduced in this chapter the fundamental theory involved for the development of the DWT-MuSiC methodology, in the next chapter is described in detail the structure of the algorithm and how is the working method for each step of the analysis it performs.

CHAPTER 3: DWT-MUSIC ALGORITHM

3.1 ALGORITHM OVERVIEW

The main purpose of this PhD thesis was to develop an innovative methodology for the seismic array data analysis, named DWT-MuSiC (Discrete Wavelet Transform-Multiple Signal Classification)

DWT-MuSiC, is based on a combination of MuSiC and the potentialities of the discrete wavelet analysis. The use of the DWT-MuSiC, other than to distinguish the arrival of different wavefronts, providing both their backazimuth and their apparent speed of advancement (the inverse of the apparent slowness), returns even information about the polarization of each identified phase. With the analysis results, moreover, temporal information as well as amplitude spectrum content related to the original signals are preserved.

The aim of this thesis, however, was not only to develop such a methodology, but also to develop a software package, potentially capable to perform near-real time analysis in different geophysical contests, as monitoring, where fast arrays analysis is a critical task. For this purpose the DWT-MuSiC algorithm was implemented using *Python*, a modern and innovative programming language.

Python is a widely used high-level programming language used for general purpose programming, created by Guido van Rossum and first released in 1991.It has a design philosophy which emphasizes code readability and a syntax which allow writing calculation in fewer lines of code. Moreover it is widely used and interpreters are available for many operating systems, allowing its code to run on a wide variety of systems, is open source and has a community-based development model, as do nearly all of its variant implementations.

Developing the methodology, particular emphasis was placed in finding a compromise in producing high quality results, keeping computing time within few minutes, using current desktop computers. A large possibility of customizable settings are also implemented in DWT-MuSiC algorithm, helping the user to prioritize the quality of the analysis rather than the time needed to perform it,

leaving also the option to perform the complete algorithm workflow characterizing signals polarization, slowness and backazimuth, or limiting the results just to some specific analysis.

For better understanding how the algorithm of DWT-MuSiC works, here it is presented the completed algorithm workflow that is schematized into 4 different steps, which are subsequently discussed in separate paragraphs. The steps are:

1) Wavelet decomposition
2) Preliminary backazimuth and slowness wavefront estimation
3) Wavefront polarization characterization
4) Non-linear optimization of the results

To overview the whole process, the algorithm starts making a preliminary transformation of the seismic signals by means of the discrete wavelet transform (DWT), obtaining the wavelet coefficients that are going to be processed.

After getting the wavelet coefficients, DWT-MuSiC performs the backazimuth and slowness estimation, as well as the polarization characterization, on each single wavelet frequency/time interval of the decomposed signal.

After distinguishing the seismic phases, and after having characterized them for all the explored parameters, DWT-MuSiC revises the results using a non-linear optimization technique in order to overcome the problem of the resolution of grid search methods.

The application of DWT-MuSiC of course is subject to some assumptions related to the limits of the MuSiC methodology and to the range of application of seismic arrays.

The main assumptions are listed below:

1) The methodology is applicable to the case of a planar array
2) The seismic signals are assumed to be under the far-field assumption (seismic source >> array aperture)

3) Seismic waves are assumed to propagate in a homogeneous medium in the surroundings of the array.
4) The signals come from an undefined numbers of seismic sources that are anyway assumed to be less than the number of array sensors.
5) Seismic noise is assumed to be spatially incoherent

In the next paragraph it is explained in detail the algorithm workflow here introduced.

3.2 DISCRETE WAVELET TRANSFORM ALGORITHM

The first operation that the algorithm performs is the decomposition of the seismic signals by means of the DWT. Before executing this step, a preliminary operation is needed in order to obtain the analytic representation of the seismic signals. This lets the MuSiC algorithm being able to extrapolate signals phases information. The use of the analytic representation of the signal is in fact essential in order to obtaining at the end of the process complex wavelets coefficients that preserve the phase information.

In signal processing, the analytic representation of a real data sequence is defined as:

$$u_a(t) = u(t) + i \cdot H(t) \qquad \textbf{(Eq. 3.1)}$$

Where *u(t)* is the original data, *i* is the imaginary unit and *H(t)* the Hilbert transform. The Hilbert transform is a linear operator that takes a function, $u(t)$, and produces a function, $H(t)$, in the same domain:

$$H(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u(s)}{t-s} ds \qquad \textbf{(Eq. 3.2)}$$

The Hilbert transformed signals have the same amplitude and frequency content as the original sequence but phases of individual components are shifted of $\pi/2$.

After having calculated the analytical representation of the input signal, the algorithm proceeds performing the discrete wavelet decomposition calculating the complex wavelet coefficients.

The input signal S is then decomposed as indicated in Eq. 3.3 according to the DWT theory.

$$S=[cD_1,cA_1;cA_2,, cD_2 ; \ldots cA_n, cD_n] \qquad \textbf{(Eq. 3.3)}$$

where n is the last level of decomposition and cD and cA are the detail and approximate Coefficient of each level.

For this thesis a Coiflet wavelet was chosen as mother wavelet. The choice was supported by the motivation that the Coiflet wavelet has a symmetrical shape that well adapts to the seismic signals and is computationally cheap.

The levels of the DWT decomposition was automatically selected each time by the used algorithm, aiming to the maximum level it was able to compute on the basis of the input signals. Once the algorithm completes the analysis, for each input signal, it returns a matrix of values:

$$\boldsymbol{\Sigma} = \begin{Bmatrix} \sigma_{\omega_1,t_1};\ \sigma_{\omega_1,t_2};\sigma_{\omega_1,t_3} \dots \sigma_{\omega_1,t_m} \\ \sigma_{\omega_2,t_1};\ \sigma_{\omega_2,t_2};\sigma_{\omega_2,t_3} \dots \sigma_{\omega_2,t_m} \\ \dots \\ \sigma_{\omega_n,t_1};\ \sigma_{\omega_n,t_2};\sigma_{\omega_n,t_3} \dots \sigma_{\omega_n,t_m} \end{Bmatrix} \quad \textbf{(Eq.3.4)}$$

Where ω_n is the n-th decomposition level and t_m is the m-th wavelet time interval.

In the case of the DWT-MuSiC, the algorithm performs, once a time, the wavelet decomposition of all the signals registered at the array stations, taking the coefficients of only one seismic component of each station a time. It is moreover important to highlight that the coefficients coming out after all DWT decomposition, are stored in 3 different datasets, each one containing the coefficients of a single sensor component registered by all the array stations :

$$\Sigma_{\mathbf{x}}=[C_1, C_2 \dots C_i] \ , \ \Sigma_{\mathbf{y}}=[C_1, C_2 \dots C_i] \ , \ \Sigma_{\mathbf{z}}=[C_1, C_2 \dots C_i] \quad \textbf{(Eq.3.5)}$$

where i is the i-th array station, and x,y,z are the seismic components.

The reason for separating the coefficients, is because for each time/frequency wavelet interval, MuSiC analysis is performed first separately on each component and then recombined together only at a final stage.

To better visualize and understand the structure of wavelet coefficients obtained as outputs, each dataset Σ of coefficients is presented with a particular plot, called

scalogram. This plot has 3 axis: x=time, y=scale of decomposition and a color legend that represents the coefficient amplitude. A schematic representation of a scalogram is shown in figure 3.1. It is possible to note how the time windows length is adapted automatically to best fit each DWT level content that is analyzed.

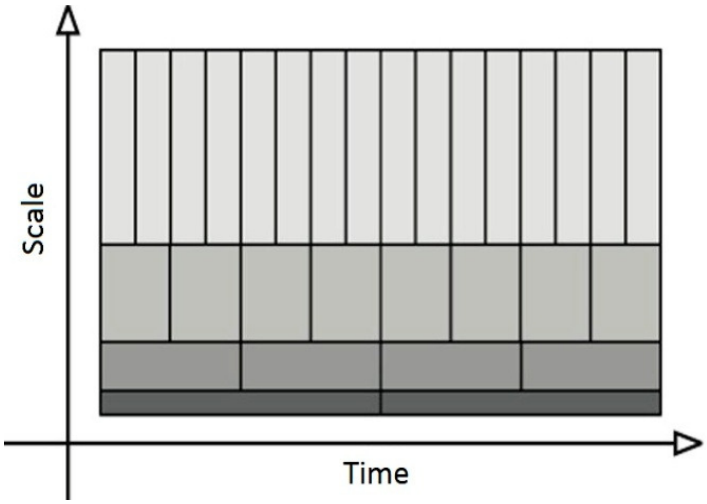


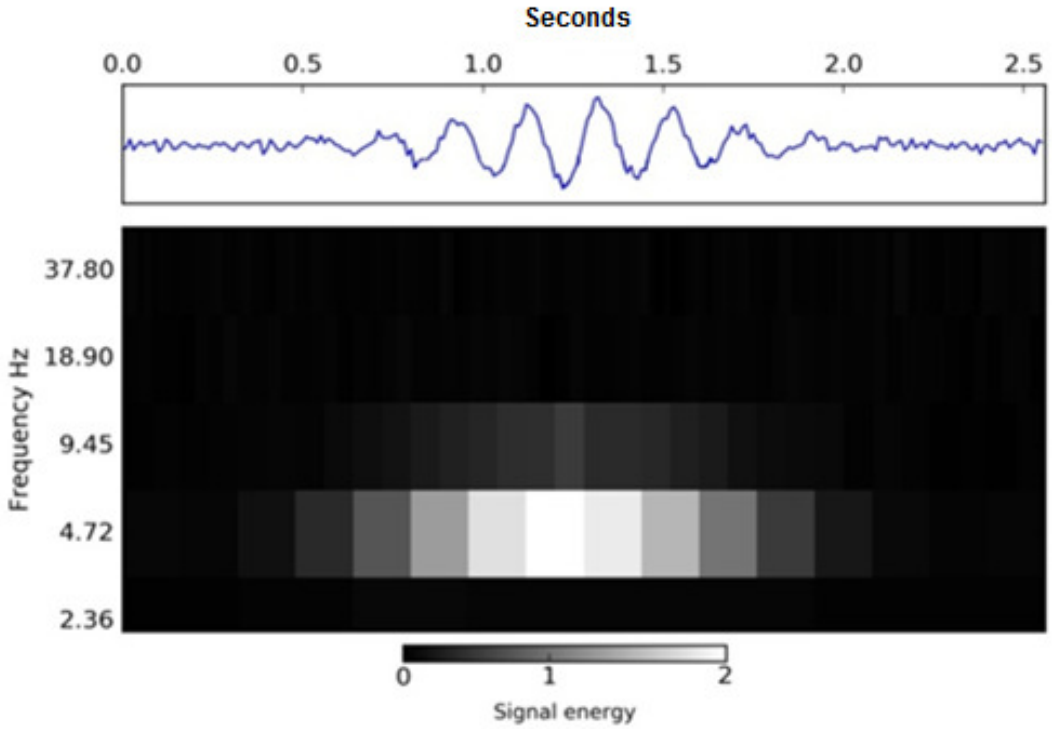
FIGURE 3.1 *Representation of a scalogram with the DWT intervals*

For better understanding the outputs returned by the DWT-MuSiC algorithm, in the figure 3.2(a)(b)(c) are shown 3 different scalograms, each one relative to a single component (East-West, North-South, and Vertical) obtained from a synthetic signal shown above each of them.

In figure 3.3, moreover, it is represented the scalogram of the total signal components, obtained combining the contribution of all the component together. The total coefficient amplitude has been reconstructed starting from the single component contributions with the following relation:

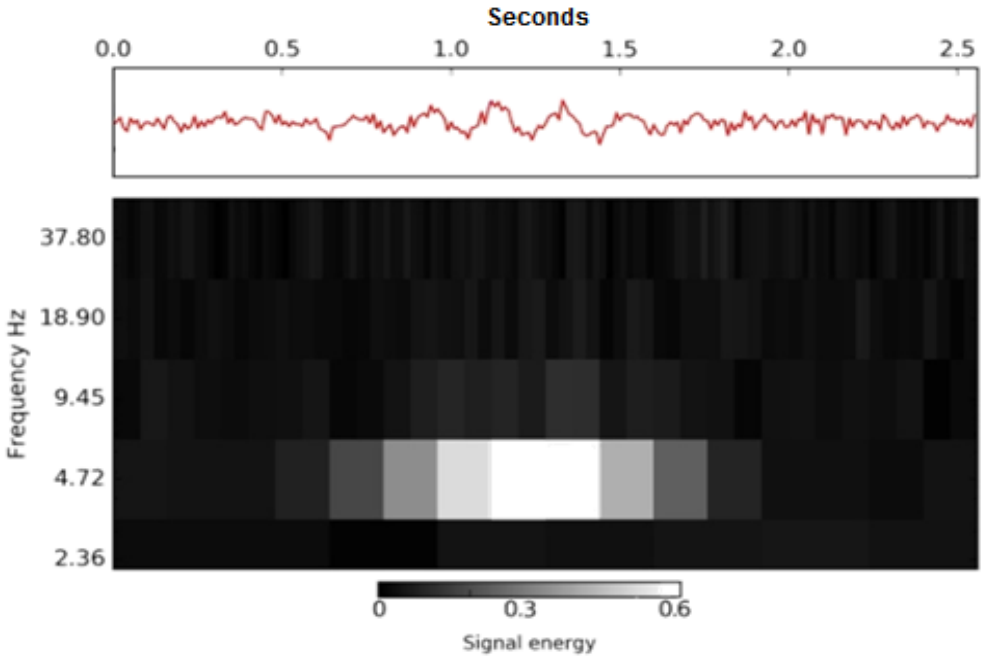
$$\sigma_{\text{total}}(\omega, t) = \sqrt{\sum_{j=1}^{3} \sigma_j(\omega, t)^2} \qquad \textbf{(Eq. 3. 6)}$$

where ω is the DWT decomposition level, t is the temporal reference and j is the number of seismic component



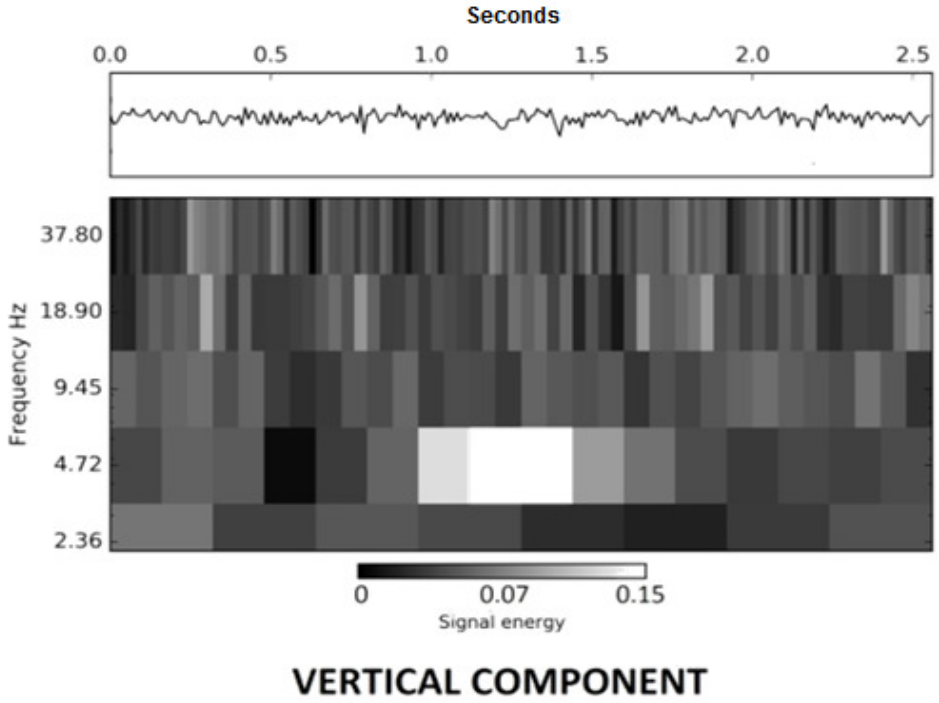
EAST- WEST COMPONENT

(a)



NORTH - SOUTH COMPONENT

(b)



(c)

FIGURE 3.2(a)(b)(c) *Single component scalograms of a synthetic seismic signal decomposed by means of the DWT*

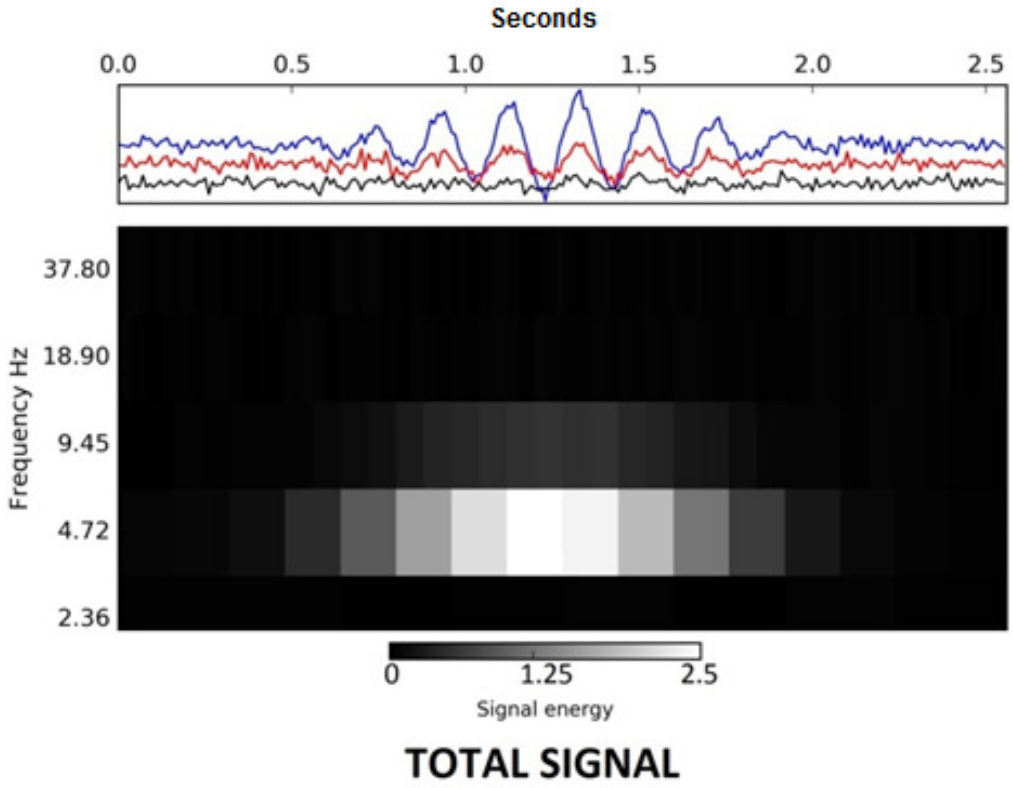
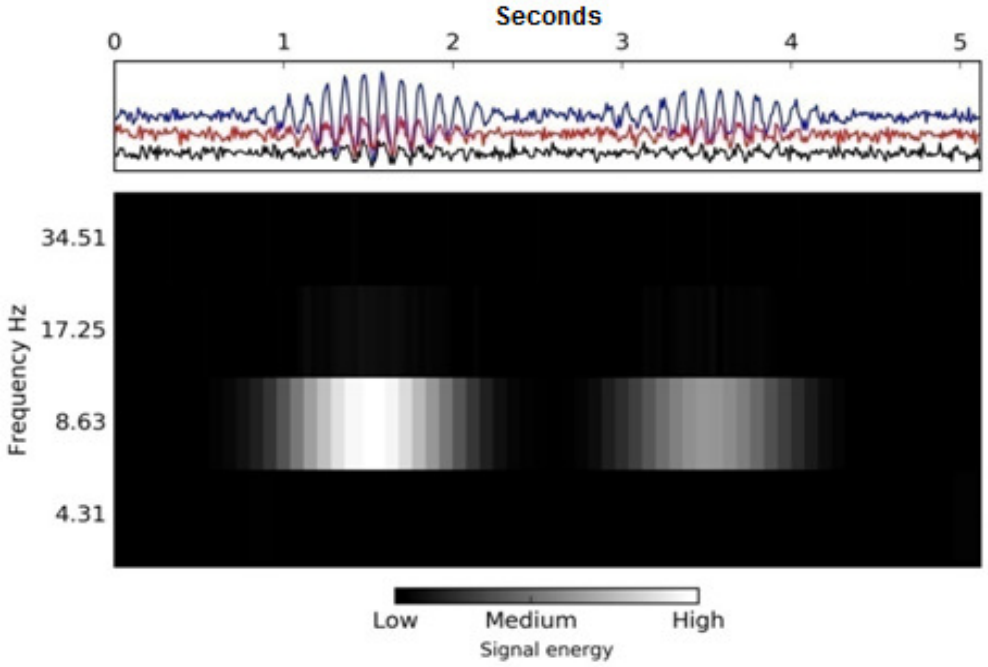


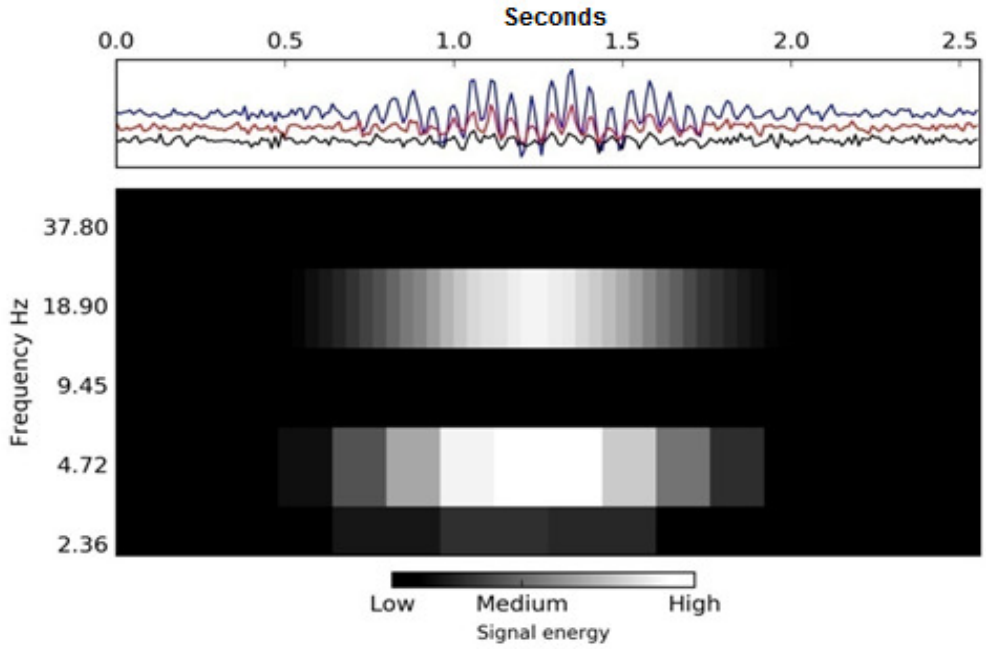
FIGURE 3.3 *Scalogram showing the coefficient amplitude of the 3 component wavelet coefficients recombined together.*

From the analysis of the figure 3.2(a)(b)(c) it has to be noted that instead of indicating the DWT decomposition level on the y axis, we show the central frequency for an easier interpretation.

To be able to attribute a central value of frequency to each level, we used as reference the associated mother wavelet, checking its amplitude spectrum with a Fourier transform. Although a wavelet is not a monochrome function, it is possible to check its amplitude spectrum taking the value of the maximum amplitude as reference. Continuing observing the signal wavelet decomposition and its representation through scalograms, it is interesting to see how the wavelet coefficient amplitude is a parameter that easily indicates where the energy of the signal is located in terms of temporal position and frequency content. That is especially true in the case where the analyzed seismic signal has the contribution of more than one seismic source. In this case it is extremely important to immediately recognize different seismic phases. The seismogram represented in figure 3.4(a) for example represents 2 signals with similar frequency but arriving at the station at different time. The figure 3.4(b) represents, on the contrary, a scalogram of 2 signals with different frequencies that arrive at the same time.



(a)



(b)

FIGURE 3.4 *Scalogram showing a seismic signal with 2 different wavefronts. In figure (a) the 2 waves have similar frequencies and different arrival time, in figure (b) waves have different frequencies and overlaps.*

The information coming from coefficient amplitudes of the signals is very important because gives a quick overview of where the seismic energy of the different phases are located and subsequently where maximum signal coherence in the following MuSiC analysis has to be expected.

The DWT-MuSiC uses the wavelet coefficient amplitudes as a threshold value, in order to perform the subsequent MuSiC analysis just within the intervals where the amplitude overpass a chosen arbitrary value and the signals information are expected to be higher. This is made to save computational time.

3.3 BACKAZIMUTH AND SLOWNESS ESTIMATION

Starting from the transformed signal by means of DWT algorithm, the next step that DWT-MuSiC does, is to perform the analysis aimed at detecting different seismic phases present in the seismograms, characterizing them in term of backazimuth and slowness.

MuSiC analysis is designed to be performed on nearly monochromatic signals, so to be adapted to a seismograms, that for their nature are broadband, the analysis has to be carried out several times, each time considering only the wavelet coefficients relative to a particular narrow frequency band. In order to analyze separately the different DWT time intervals, moreover, for each frequency band, only the coefficients related to a specific time window are selected. In this way the DWT-MuSiC will perform multiple analysis , for each different frequency bands, each one having a certain numbers of DWT time intervals.

Each single analysis is performed on a dataset of wavelet coefficients that has as many elements as the number of the stations composing the array:

$$C_j=[\sigma_{1,j}(\omega_k,t_k),\ \sigma_{2,j}(\omega_k,t_k)\dots\ \sigma_{i,j}(\omega_k,t_k)] \qquad \textbf{(Eq. 3.7)}$$

Where *i* is the i-th array station, $1<j<3$ is the seismic component and *k* is the k-th interval.

To create the input covariance matrix, according to the MuSiC theory, the algorithm multiplies the coefficient dataset with the conjugate transpose of itself:

$$R = C \cdot C^{\dagger} \qquad \textbf{(Eq . 3.8)}$$

where the symbol † is the conjugate transpose.

The covariance matrix is then decomposed in terms of eigenvalues and eigenvectors in order to separate the noise subspace to the signals subspace.

As explained in the paragraph 2.2 the MuSiC theory says that a threshold has to be chosen in order to divide the two subspaces, on the basis of the reduction in the eigenvalues amplitude.

Despite it is possible to use different to face to this problem, e.g. Akaike's Information Criteria, for what concern for this thesis, the signal subspaces is always associated only to the first eigenvalue having the greatest value, attributing the noise information to all others.

The reason of this choice is that experimentally, displaying the value of all the eigenvalues on a plot, it has been verified that only the first one had a high value associable to the seismic signals, while all the others showed a near null value (figure 3.5). This is true also when multiple synthetic seismic sources were simulated.

This characteristic is related to the fact that the covariance matrix has been created starting from only one wavelet coefficient per seismic station instead of the signal itself, the original version of MuSiC algorithm says.

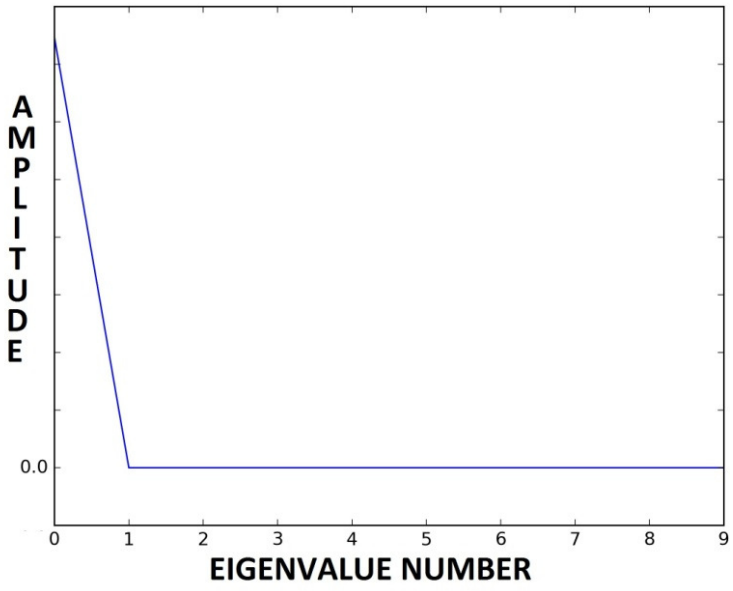


FIGURE 3.5 *Example of eigenvalues and their related amplitudes obtained decomposing the covariance coefficients matrix*

After having selected and isolated the eigenvalues and the relative eigenvectors associated to the noise subspace, the algorithm continues its analysis, determining

the backazimuth and the slowness of coherent seismic phases presented in each considered signal interval.

To make this, following the MuSiC method, the eigenvalues of the noise subspace E_n are used in conjunction by a set of steering vectors, to check where the projection of these on the noise subspace is low.

The Eq. 3.9 recalls the formula of the MuSiC estimator function D(k):

$$\mathbf{D(k)} = \frac{\mathbf{1}}{\mathbf{|a(k)\cdot E_n|^2}} \qquad \textbf{(Eq. 3.9)}$$

The exploration of the noise subspace, is performed with a grid search method, defining different steering vectors with different couples of $\mathbf{u_x}$ and $\mathbf{u_y}$ slowness values, taken at regular steps between a minimum and a maximum preselected range. A Steering vector is an L-dimensional complex vector containing responses of all L elements of the array to a narrowband source of unit power.

For each node of the grid then, we obtain a value of D(k) function that is representative of how the used steering vector, associated to a defined couple of $\mathbf{u_x}$ and $\mathbf{u_y}$ values represents the signal subspace.

With a certain steering vector, the higher is the value of the D(k) function, the much representative its associated slowness values are.

For each analysis performed the $\mathbf{u_x}$ and $\mathbf{u_y}$ grid limits are not fixed, but vary depending on the frequency band of the signal is being analyzed.

The necessity to set different slowness limits is aimed to eliminate artifacts due to spatial aliasing and to assure that the analysis is performed within its validity range. There are some constrains that were considered structuring the algorithm analysis, some of them related to the use of the wavelet, some other already introduced in chapter 1 and strictly related to the use of the array.

The first condition is based on the consideration that the wavelength of a wave has to be at least comparable with the array interspacing:

$$\lambda = \frac{v}{f} \geq Array\ interspacing \qquad \textbf{(Eq. 3.10)}$$

To satisfy this, for an analyzed frequency band, the analysis is high limited in the explored slowness. This is done in order to consider only signals having the wavelength longer than the array interspacing, thus reducing the aliasing phenomenon.

Another condition is related to the possibility of resolving the phase difference at the stations. To make it possible the wavelength of the signals must be comparable with the array aperture:

$$\lambda = \frac{v}{f} \leq 2 \cdot Array\ aperture \qquad \textbf{(Eq. 3.11)}$$

To satisfy this condition, again, assuming we decided to analyze a certain interval of speed based on local geology and velocity medium, it is enough to set a lower limit in the wavefront slowness (1/v) to be analyzed.

The last condition is justified by the necessity that a seismic phase, to be analyzed, have to travel through the entire array. Considering the length of the considered wavelet at a given scale (t), the condition is expressed as**:**

$$t \geq \frac{Arrayaperture}{v} \qquad \textbf{(Eq. 3.12)}$$

Considering that each DWT time intervals *t* has different length on the basis of the explored frequencies, also in this case, it is necessary to set a maximum value of slowness (1/v) to be analyzed.

At the end of the process, when for each single node of the grid, the complete spectrum of the function D(k) function is obtained, the program return as output, the $\mathbf{u_x}$ and $\mathbf{u_y}$ related to the maximum value of the D(k) spectrum.

With these values, we can obtain backazimuth and apparent speed values using the following relations:

$$\boldsymbol{\Phi} = \tan^{-1}(u_y/u_x)$$

$$\boldsymbol{V_{app}} = \sqrt{{u_x}^2 + {u_y}^2} \qquad \textbf{(Eq. 3.13)}$$

The D(k) spectrum can also presented on a 2D polar plot (figure3.6).

The diagram is structured in order to have on the X-axis the interval of analyzed East-West slowness interval, while the Y-axis the slowness component on the North-South direction. The advantage of using this plot is based on the possibility to read it as a compass diagram. Since the Y-axis coincides with the North-South direction and the X-axis with the East-West one, the position of the maximums of the D(k) function is easily be associate to a particular backazimuth. Different circular lines of equal apparent propagation speed, with increasing values toward the center, are also represented on the plot for better reference. Below there is a schematic representation of the diagram just introduced:

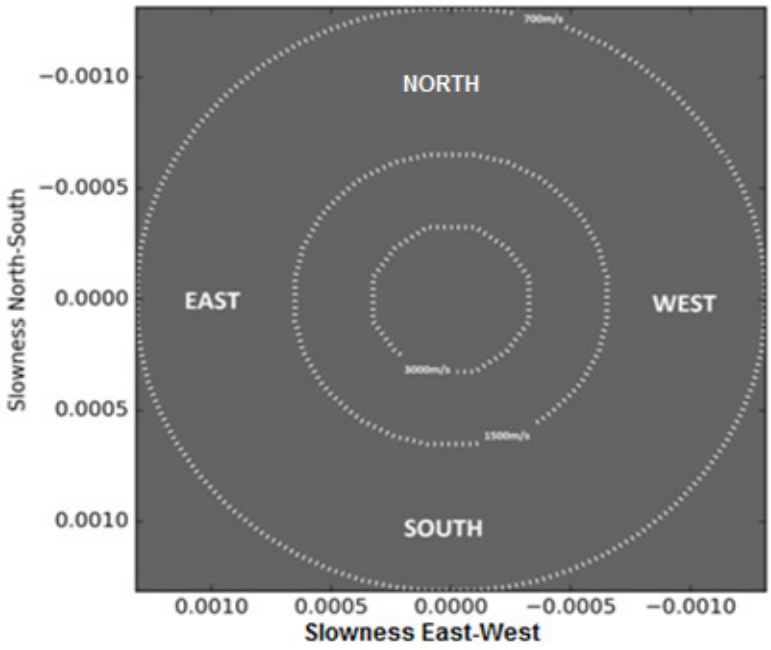


FIGURE 3.6 *Representation of the circular diagram used to plot the D(k) MuSiC spectrum.*

As explain before, the function D(k) is calculated independently for each single seismic component, but in order to be representative of the total seismic signal, we need combining the single seismic components spectrum D(k) all together. In order to do this, the DWT-MuSiC can follows two different strategies.

The first, used to maximizing the coherence between the three components combines the components spectrum using the following relationship:

$$D(k) = \sqrt{\sum_{n=1}^{3} (D_n(k)/\widetilde{D_n(k)})^2} \qquad \textbf{(Eq. 3.14)}$$

where the symbol ~ is the median of the function.

The second method, that can be selected in place of the previous one, consists incomparing the value of D(k) for each component, creating a new spectrum taking for each position the maximum value among the three spectrum.

$$D(k) = Max[D_1(k), D_2(k), D_3(k)] \qquad \textbf{(Eq. 3.15)}$$

The results in general are similar, but the choice of one instead of the other has to be evaluated in any single case. For each combination method in fact, the outputs can be slightly different depending if there are more different seismic phases and their energy contribution on each seismic component. In figure 3.8 there is an example of a backazimuth and slowness estimation relative to a simulated seismic phase coming from 45° N and having an apparent speed of 600m/s. The analysis has been performed with a simulated array having an aperture of 200 meters and composed by10 three-component sensors (figure 3.7).

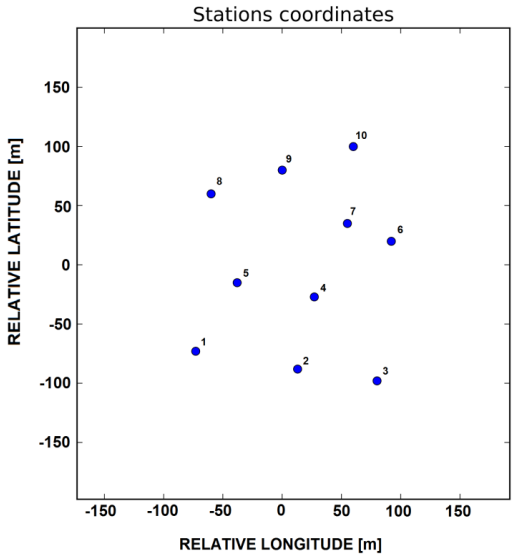


FIGURE 3.7 *Plan view of the synthetic array geometry*

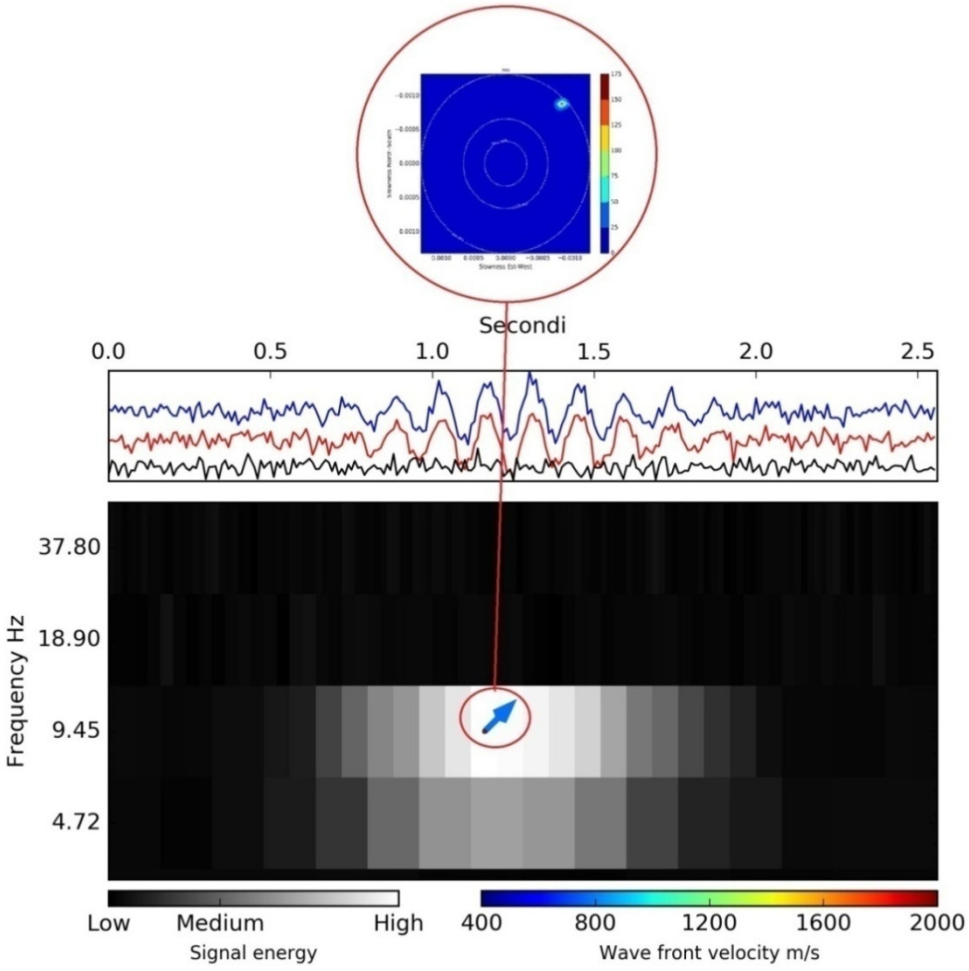


FIGURE 3.8 *MuSiC analysis of a synthetic signal coming from 45° N and having an apparent speed of 600m/s.*

In the figure3.8 we can see the scalogram showing the wavelet coefficient of the signal (the 3 component together) and in correspondence to the interval with the higher energy, the results of the DWT-MuSiC analysis. The backazimuth and slowness results are presented by means of arrow. The direction where the arrow points, correspond to the backazimuth, whereas the speed of propagation is indicated by the color of the arrow itself.

Figure3.9 shows a detail of the three seismic component spectrum from which the total D(k) function is reconstructed.

As we can see from the amplitude of the signals in figure 3.9, the majority of the energy signal is contained in the horizontal component. While the vertical component is almost all composed of noise.

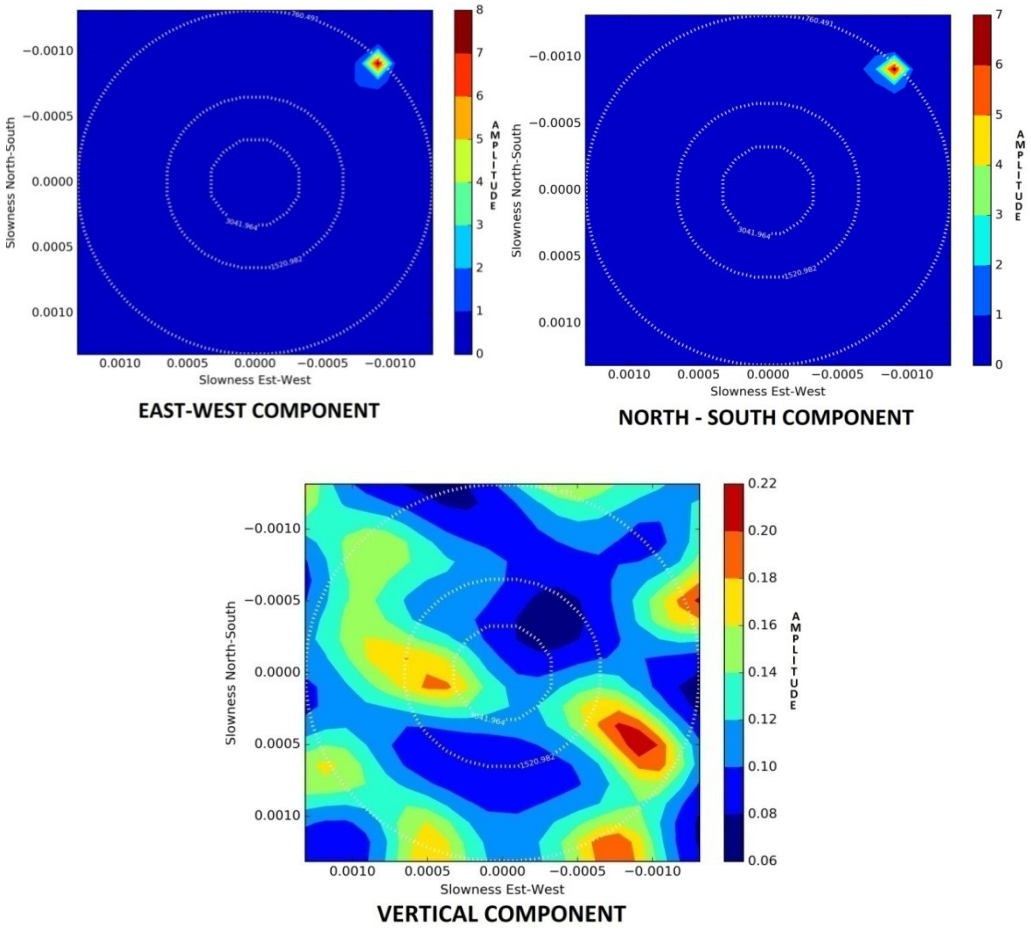


FIGURE 3.9 *MuSiC analysis of a synthetic signal coming from 45° N and having an apparent speed of 600m/s.*

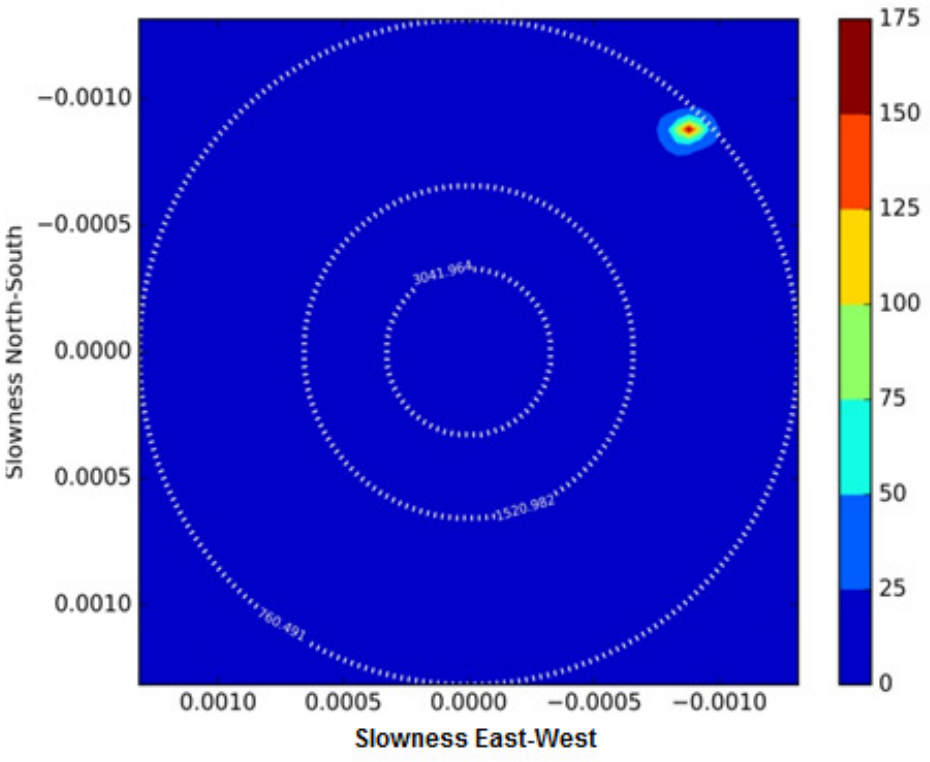


FIGURE 3.10 *MuSiC analysis of a synthetic signal coming from 45° N and having an apparent speed of 600m/s.*

As it was previously described, it is possible to select the intervals where to execute the DWT-MuSiC analysis by setting a threshold value based on the maximum amplitude of the wavelet coefficients.

In figure 3.11 we perform DWT-MuSiC, setting a lower threshold value. What is possible to notice is that the MuSiC analysis is performed on a greater number of coefficients. Where the signal energy is less, the coherence of the analysis decreases, showing only random orientations where the signal/noise ratio is too low. This show the importance of choosing a valid threshold value for saving computational time and to improve the readability of the results.

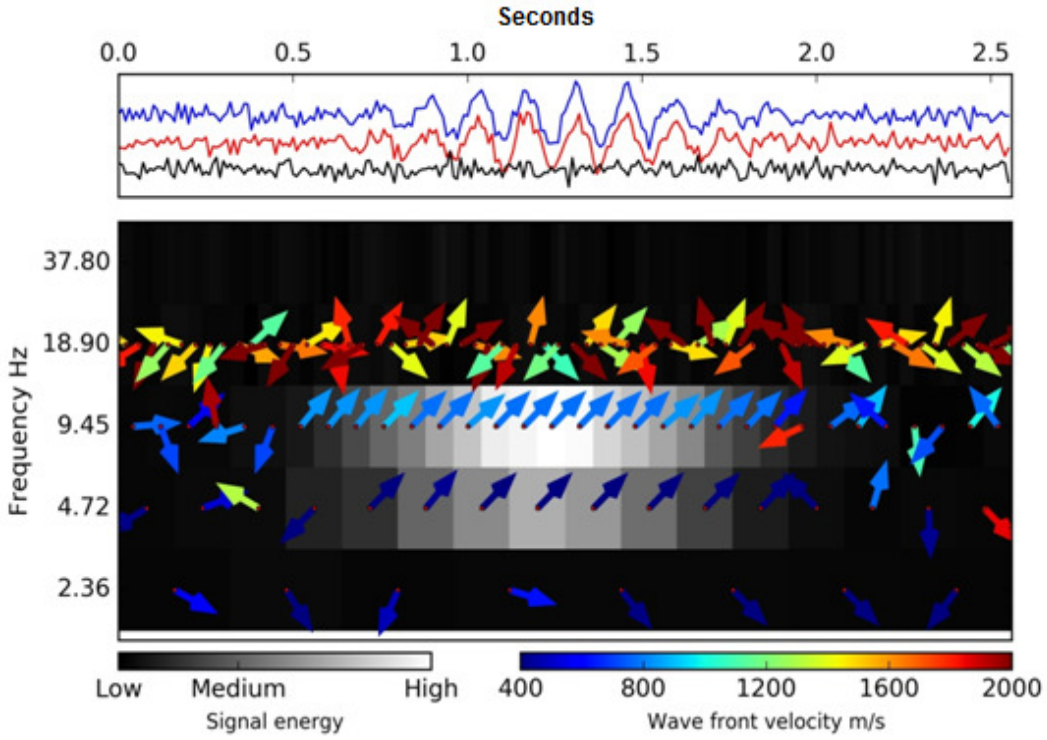


FIGURE 3.11*MuSiC analysis of a synthetic signal coming from 45° N and having an apparent speed of 600m/s.*

3.4 POLARIZATION ESTIMATION

This paragraph describes the second step of the DWT-MuSiC algorithm workflow, during which the polarization of the previously identified seismic phases are characterized.

The seismic wave field is a combination of polarized waves in the three-dimensional space. The polarization is a characteristic of the wave related to the displacement of particles that shows a preferred direction of motion depending on the source properties and the Earth structure. Body waves for example, usually generate a linear particle motion with different directions depending on the wave type (P or S). Rayleigh waves, on the other hand, generate retrograde elliptical polarization. The characterization of the polarization is useful to improve the understanding of the source process and the Earth structures crossed by the recorded waves. The objective of the polarization analysis is the determination of the type of polarization (linear, elliptical, transversal or longitudinal) and its orientation in the space, by means of the following parameters:

4) **Polarization azimuth α** = angle of the main polarization axis, measured clockwise from the North. $0° < \alpha < 360°$

5) **Polarization inclination β** = vertical angle of the main polarization axis, measured from the vertical $0° < \beta < 90°$

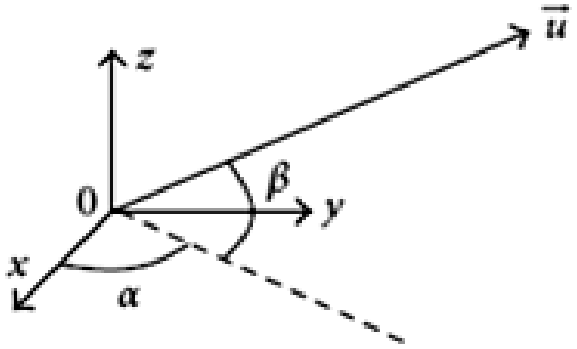


FIGURE 3.12 *Diagram showing polarization azimuth α and inclination β in a Cartesian reference system.*

6) **Polarization ellipticity ε**= a parameter indicating how much a wavefront polarization is close to circular pattern. 0<ε< 1. A value of 0 identifies a linear polarization typical of body wave. Greater value, usually ε>0,5 identify progressively a near circular polarization typical of surface waves.

If the components oscillation have the same phase in fact, will produce a linear polarization, but if it is presented a phase shift in between components, these will generate an elliptical polarization as is shown in the figure 3.13. When the phase shift is exactly ±π/2°, then circular polarization case is verified.

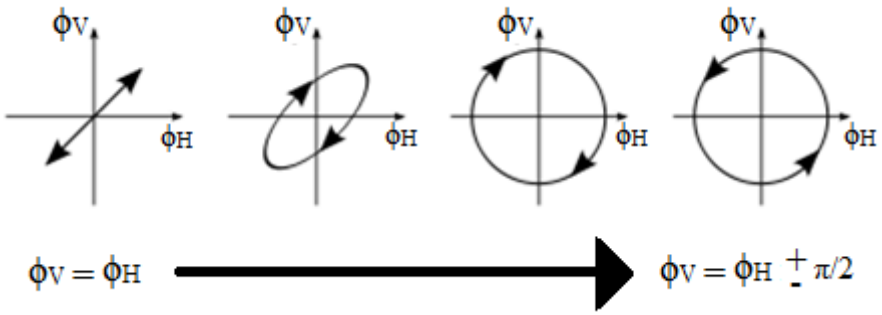


FIGURE 3.13 *Various examples of polarization diagrams. From the left to the right the ellipticity value is increasing passing from a linear polarization to a circular one. ϕV and ϕH represent the horizontal and vertical phases.*

Considering the coordinate system defined by the 3 seismic component registered by the seismic sensors, the ellipticity can be described by the relative phase variation of the horizontal component in relation with the vertical one.

$$\varepsilon = \frac{\sqrt{(\phi_z - \phi_x)^2 + (\phi_z - \phi_y)^2}}{a} \qquad \textbf{(Eq. 3.16)}$$

where ϕ_x and ϕ_y are the phases along the East-West and North-South directions, while ϕ_z is the phase value on the vertical plane. a=4.44 and is a normalizing constant considering the maximum phase difference of π between the vertical and each horizontal component.

Before describing the algorithm in detail, we need to clarify the reason why DWT-MuSiC is structured in such a way to perform the backazimuth and slowness estimation and from polarization in 2 separate steps.

The reason to structure the algorithm in this way is because, even if theoretically the slowness, the backazimuth and the polarization estimation could be done simultaneously, the computational time required to perform MuSiC calculation dealing with a 5-D space parameter (backazimuth, slowness, polarization azimuth, polarization inclination and ellipticity) is not acceptable to the scope for which DWT-MuSiC is thought.

Considering so, it resulted practical to divide the process of estimation into two different steps reducing drastically the time needed for obtaining the results, without moreover having a negative dependence between the 2 steps. It has been experimented in fact that the polarization characterization of the algorithm doesn't depends much on the success of previous analysis, even if its results would be not so accurate. For example if we decide to perform the analysis with a coarse grid search algorithm for time saving purpose, DWT-MuSiC is still able to find the signals polarization domain with good accuracy.

The way how the polarization analysis is performed, is by starting a new MuSiC analysis searching also in this case the maximum D(k) estimator function's with a grid search method. This time the function spans a 3-D space defined by the polarization parameters (α,$\beta$and ε).

Despite the similarity of the previous step aimed to backazimuth and slowness estimation, there are some differences in the way how the polarization characterization is performed.

Polarization in fact is calculated considering the interaction of all the 3 seismic components signals and for this reason the MuSiC estimator function D(k) has to be calculated taking into account the wavelet coefficients of the all 3 seismic component at the same time. This is done computing, for each DWT interval, the covariance matrix of the long concatenated vector coefficient C(s,t) defined as follows:

$$C(s,t) = [Cx(s;t), Cy(s;t), Cz(s;t)] \qquad \textbf{(Eq. 3.17)}$$

where Cx, Cy, Cz are the 3 component coefficients, *s* is the DWT decomposition level and *t* is the time interval analyzed.

After having analyzed the eigenstructure of the matrix, selecting the eigenvalues related to the noise and signal subspace as theory says, a new set of steering vectors are created to span those subspaces.

This time the steering vectors are defined differently, and vary only according to different values of α, βand ε, maintaining a fixed values of $\mathbf{u_x}$ and $\mathbf{u_y}$ slowness, as found in the first step of the analysis. In this case also the relative amplitudes of the steering vectors among the seismic components, is variable according the polarization. If we consider in fact the relation between the coordinates of the propagation vector in the sensor coordinates system,

$$a = \left[\cos\alpha \cdot \sin\beta, \sin\alpha \cdot \sin\beta, \cos\beta\right]^T \qquad \textbf{(Eq. 3.18)}$$

an incoming wave s(t) generates three different amplitudes as output on a multi component sensor defined as $s_x(t)$, $s_y(t)$, and $s_z(t)$ .The response model to s(t) can thus be defined as:

$$\begin{pmatrix} s_x(t) \\ s_y(t) \\ s_z(t) \end{pmatrix} = aw(t) \qquad \textbf{(Eq. 3.19)}$$

where *w(t)* is the signal source function.

The steering vectors a(k), considering how the covariance matrix is structured, span 3 independent seismic components and so, results from the concatenation of the 3 steering vectors related to the 3 components:

$$a(k) = [a(k)_x , a(k)_y , a(k)_z] \quad \textbf{(Eq. 3.20)}$$

After performing the analysis for all the DWT coefficients, the results are determined by checking for which position of the grid search, the function has its maximum, returning as output, α, β and ε related to that position.

The spectrum of the estimator function is displayed on a specific plot having on the x-axis the polarization azimuth, and on y-axis the inclination.

To show the variations of the polarization of the phases present in the whole seismic signal analyzed, the results are presented on a separate scalogram, with superimposed the polarization information.

An example is shown in figure 3.14 where both the polarization and slowness-backazimuth relative to two synthetic simulated seismic phases.

For that simulation we considered two seismic phases: the first is an S wave having a backazimuth of 50° N a speed of 900m/s, an incidence of 30° and a frequency of 5 Hz.

The second is a Rayleigh wave having a backazimuth of 80° N, an apparent speed of 600m/s, a horizontal incidence and a frequency of 9 Hz.

The analysis have been performed with the same simulated array shown in figure 3.9

The upper scalogram shows backazimuth and slowness results, while the lower one is associated to polarization characterization.

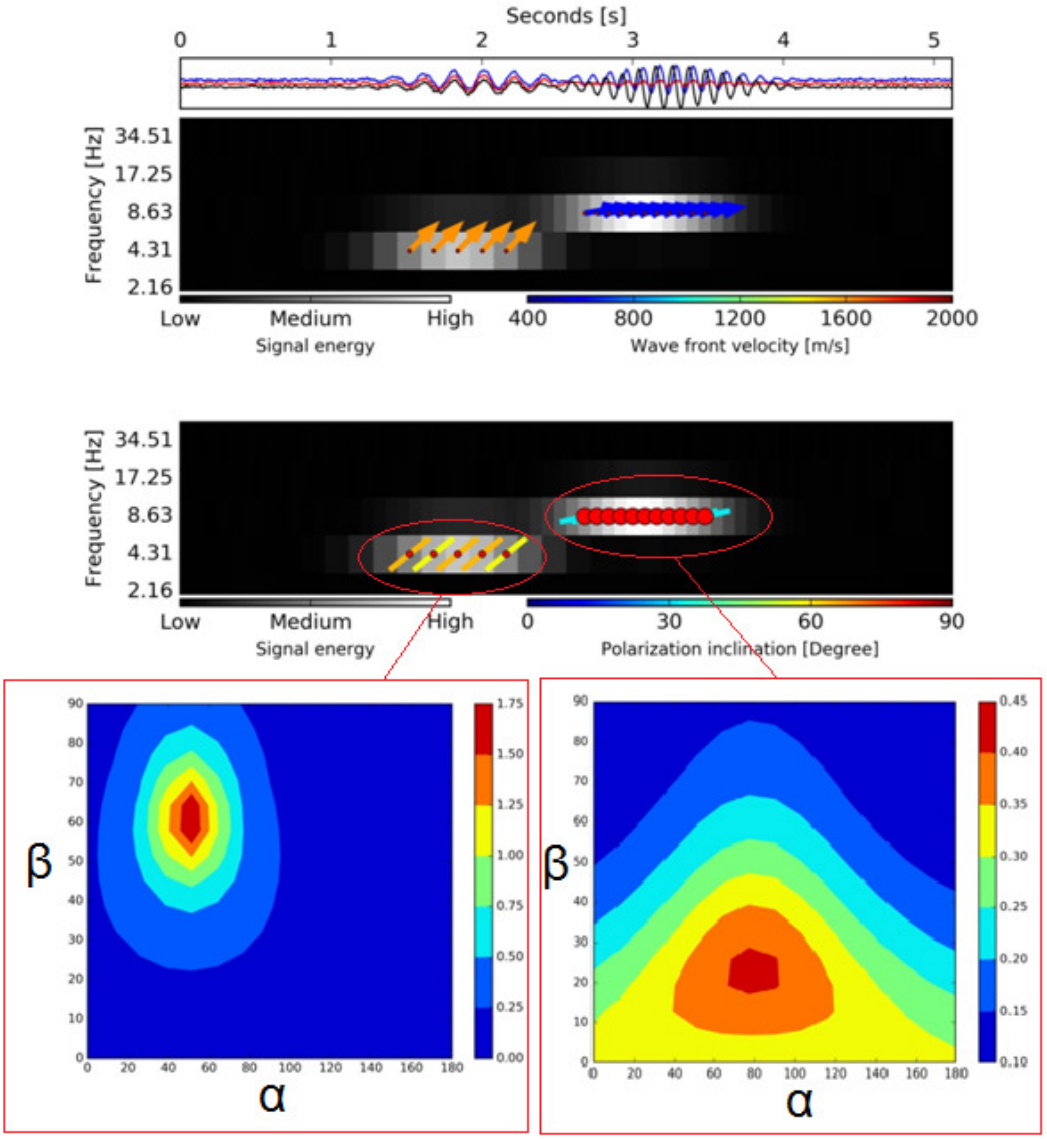


FIGURE 3.14 *Completed analysis performed on a synthetic signal composed by two wavefronts (S wave and Rayleigh wave).*
In the figure are presented the 2 scalograms relative to backazimuth and slowness outputs, and polarization outputs. In the lower part of the figure are reported also 2 representative plots showing the MuSiC polarization spectrum relative to both the simulated waves.

The backazimuth and slowness results are presented in the upper scalogram by means of arrows, as explained in the previous paragraph.

In the lower scalogram, the direction where the bar points corresponds to the azimuth of polarization, whereas the color represents its angle of inclination. The red circle, over imposed on the scalogram, gives a representation of the ellipticity of the associated main polarization. If the polarization is purely linear, the circle is

represented as a single small red dot. The larger is the circle the greater is the ellipticity.

Presenting the results in such a way is useful to easily interpret the information about the wave field, for example comparing the backazimuth of the wavefronts with the polarization azimuth, helps understanding the type of the seismic wave crossing the array.

The interpretations of the results as well as the presentation of some case studies are explained in the following chapter 4.

3.5 COMPUTATIONAL SPEED-UP

After having completed, for all the DWT coefficients, the analysis to obtain all the wavefront characteristics, the next and last step of the algorithm is to increase the precision of the already obtained output.

The parameters estimation using a grid search methods, in fact are strongly dependent on the numbers of nodes present in the search grid itself. Although the algorithm was designed in a way the user can select a fine grid search rather than a coarse one, it is not convenient in terms the computational resources to increments the precisions of results everywhere on the grid.

What is more convenient in fact is just to have a higher precision in the surrounding of the model space where the MuSiC estimator function has its maximum value.

The best solution, to save time and computational resources, is to calculate the MuSiC function D(k) starting with a relatively coarse grid search, and then increasing the accuracy of the results searching with a non linear optimization a better result.

The way used to perform this optimization is the Nelder-Mead simplex algorithm, [J.A. Nelder, R. Mead, 1965]. This method is commonly used to find the minimum, or eventually the maximum of a function in a multidimensional space and it is generally applied to nonlinear optimization problems for which derivatives may not be known.

The algorithm starts from an initial set of points representing solution estimates relative to an objective function $F(x_1,x_2,x_3,x_n)$.

The number of points supplied is one more than the spatial dimension of the model space. They form a “simplex” of $n + 1$ points within an-dimensional vector space. Examples of “simplices” include a line segment on a line, a triangle on a plane, a tetrahedron in a three-dimensional space and so forth.

The algorithm then evaluates the function at each point on the simplex and tries to extrapolate the values of the objective function measured at each test point, in order to find a new test point and to replace the old with the new one, and so on.
It considers various ways of seeking a better estimate, including replacing one vertex of the simplex by its reflected image, or by shrinking or expanding the simplex.
The basic approach is to replace the worst point with a point reflected through the centroid of the remaining *n* points. If this point is better than the best current point, then it can try stretching exponentially out along this line. On the other hand, if this new point isn't better than the previous value, then we are stepping across a valley, so it shrinks the simplex towards a better point.
Although the user specifies an initial simplex of starting values, the algorithm is not constrained to search only within that simplex. This means that the user cannot force the algorithm to search only within a restricted region. The method is well suited to solve optimization problems where the objective function varies smoothly and is unimodal.
The way how Nelder–Mead method is used by DWT-MuSiC algorithm, starting from the input results, is to define new values of the D(k) function, changing the following parameters:

1) East-West slowness vector $\mathbf{u}_x$
2) North-South slowness vector $\mathbf{u}_y$
3) Polarization azimuth α
4) Polarization inclination β
5) Ellipticity ε

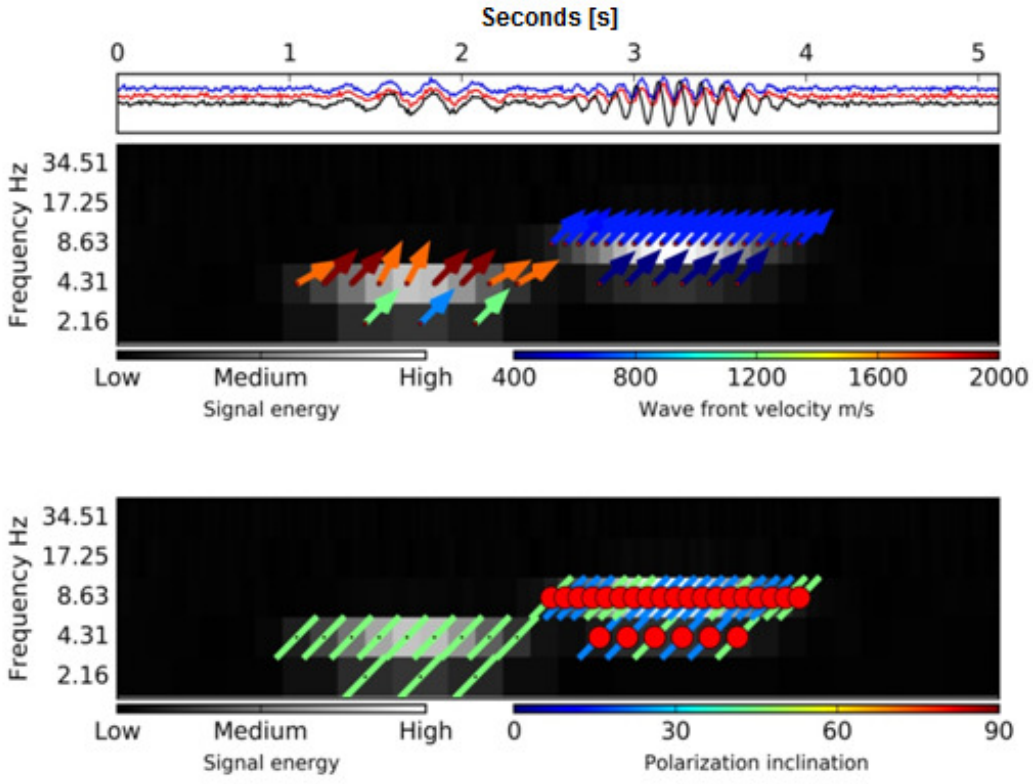
In figure 3.15 is reported an example of an array signal analyzed first with the MuSiC grid search algorithm (a), and then optimized with the use of the Nelder-Mead simplex algorithm (b).

For that simulation we considered two different wavefronts. The first is a P wave having a backazimuth of 45° N, an incidence of 90°,a speed of 1100m/s and a frequency of 4 Hz.

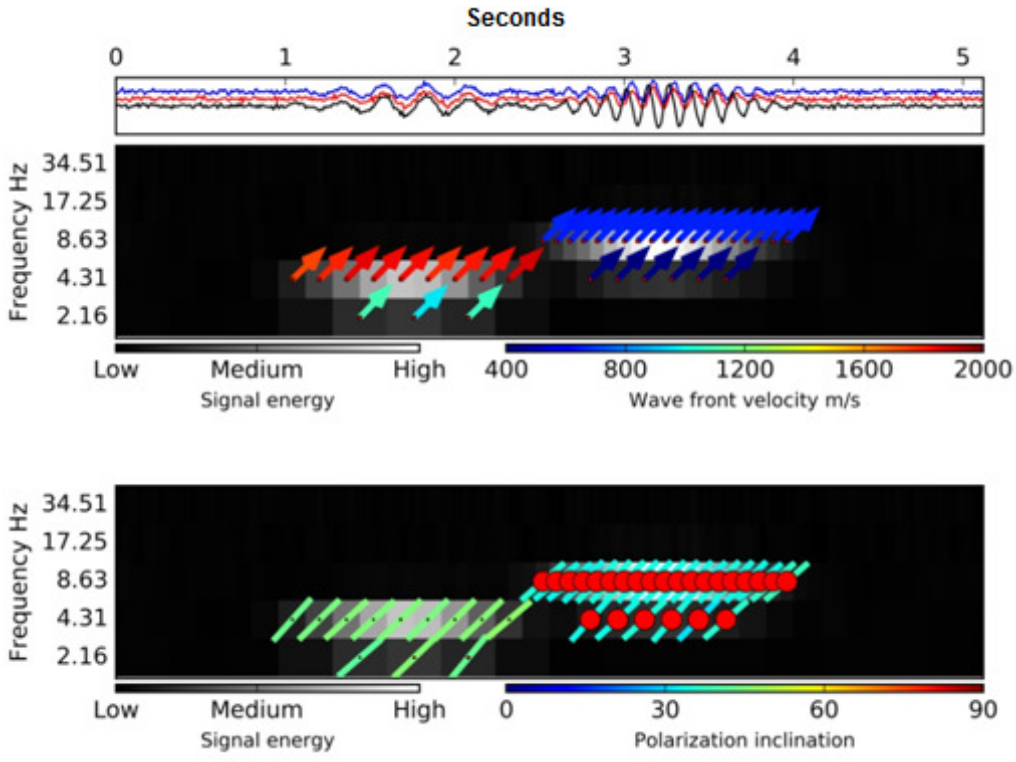
The second is a Rayleigh wave having a backazimuth of 45° N with an apparent speed of 800m/s, incidence of 90° and a frequency of 8 Hz.

As can be seen from the results, the quality of the results is greatly improved

.



(a)



(b)

FIGURE 3.15 *Comparison between outputs before (a) and after(b) non-linear optimization. It is possible to notice the greater accuracy in the results and a better reliability in the identified seismic phases.*

After having explained in detail how the algorithm works, in the next chapter are presented some tests that show the results obtained in different scenarios.

CHAPTER 4: APPLICATION OF DWT-MUSIC TO SYNTHETIC AND ACTUAL DATASETS

In this chapter are shown some indicative application of the DWT-MuSiC algorithm on synthetic datasets, performed to validate the methodology in different seismological scenarios. Together with the presentation of the synthetic tests, in paragraph 4.3 we show the results obtained applying the DWT-MuSiC to two different real cases: a volcano-tectonic event registered at Mount Vesuvius (Italy) and the analysis of array data acquired during an active seismic survey made at the Krafla caldera (Island).

4.1 TESTS ON SYNTHETIC DATA

During this work, for the development of the methodology a great number of synthetic tests have been executed in order to evaluate the capabilities and limitations of the DWT-MuSiC algorithm, checking the reliability of the obtained results. In this chapter we selected four significant tests among all those performed, to show in particular the applicability of the methodology in discriminating and characterizing:

1) A single wavefront in low signal / noise ratio condition.
2) Multiple wavefronts partially overlapped, uncorrelated, having a similar frequency content but impinging on the array at different times.
3) Multiple wavefronts, partially overlapped, uncorrelated, having different frequency content but impinging on array at the same time.
4) A wavefront composed by different transient phases. In particular we simulated a seismogram with the first arrival corresponding to the P wave, followed by the S wave and a Rayleigh wave having an elliptical polarization pattern.

For each performed test we present the results obtained with the DWT-MuSiC and we also compare the results obtained with other methodologies. In particular we considered the comparison of the DWT-MuSiC, to the standard beamforming analysis, and to the MuSiC analysis in the Fourier's domain. The aim of the comparison is to highlight the differences in executing array analysis in the time domain, in the frequency domain and in the wavelet domain.

The beamforming method was already described in Chapter 1 and we briefly introduce here the MuSiC applied to the Fourier domain. Instead of being carried on wavelet coefficients obtained through DWT, with this approach, the analysis is simply performed on the Fourier coefficients of the input signals. In this way it is possible to divide a broad band input signal into several monochromatic functions and perform an analysis similar to the MuSiC originally proposed by Smith. For each frequency, and its associated coefficients related to the signals registered at the array stations, we performed a MuSiC analysis in a similar way as for the DWT coefficients (see chapter 3).The main difference in the results, comparing the DWT-MuSiC algorithm is the complete loss of temporal information that especially in presence of transient signals, which are very common in seismology, represents a big limitation.

For the execution of synthetic tests we used a synthetic small-aperture array composed of 10 stations and having an array aperture of 200m (Figure 4.1). This geometry mimics the typical setup of arrays used in different seismological applications.

The position of the stations has been imagined to have a random distribution, equally distributed, in order to balance the resolution in all the directions, limiting at the same time the occurrence of spatial aliasing phenomena which can arise from a regular geometry of the array.

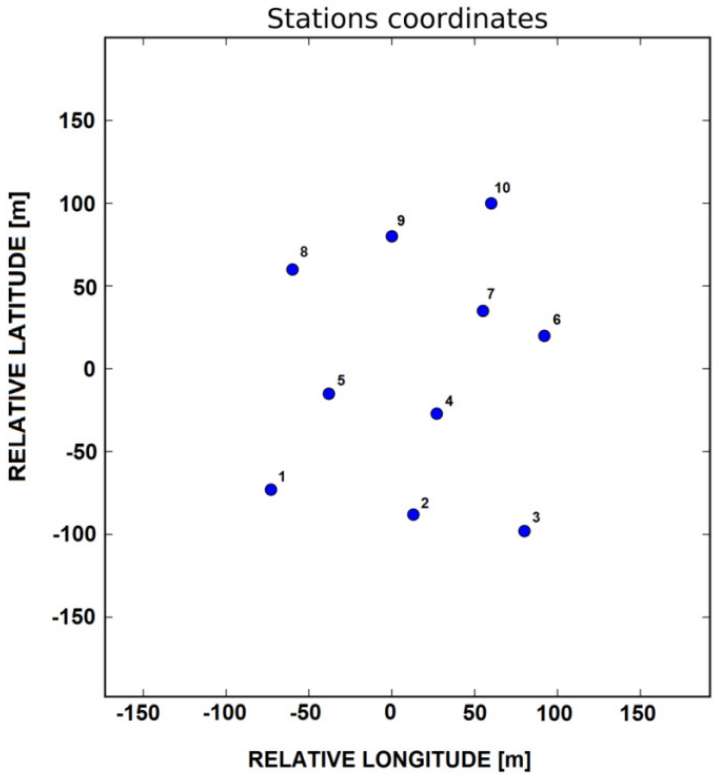


FIGURE 4.1 *Plan view of the simulated array used to perform the synthetic tests.*

The synthetic signals were obtained simulating each seismic phase with a damped sine waves having a certain frequency.

To perform each synthetic test we choose a time window long enough to permit the entire wave to cross all array stations, leaving a padding time of 2 seconds before the first appearance of the signal at the first station and 2 seconds after the wavefront leaves the array.

In this way it was assured that at all the stations, the entire signal was analyzed without loss of information.

For each simulated wavefront the input parameters used to generate the signals were:

1) The backazimuth angle
2) The wavefront velocity
3) The angle of inclination of the wavefront, respect to the vertical
4) The type of the seismic wave

5) The signal/noise ratio
6) The relative arrival time of wavefronts referenced to the center of the array

The noise simulated for the tests is assumed to be uncorrelated between the array stations and possessing a white spectrum.

- ***Test 1: Detection of a low signal to noise ratio signal***

In this first test it is shown the capability of the DWT-MuSiC algorithm in detecting and characterizing a single wavefront having a low signal/noise ratio.
The aim of the test is to show how the DWT is able to detect coherent phases even if their energy is low compared to the uncorrelated background noise. The simulated wavefront has the following characteristics:

1) P wave
2) Speed of propagation of 900m/s
3) backazimuth of 240°
4) 40° of inclination
5) Frequency of 10Hz.

The minimum signal to noise ratio that is presented here was the lowest one that produced coherent results, and had a value of 1.5
In figure 4.2 are reported the seismograms of the synthetic signals.

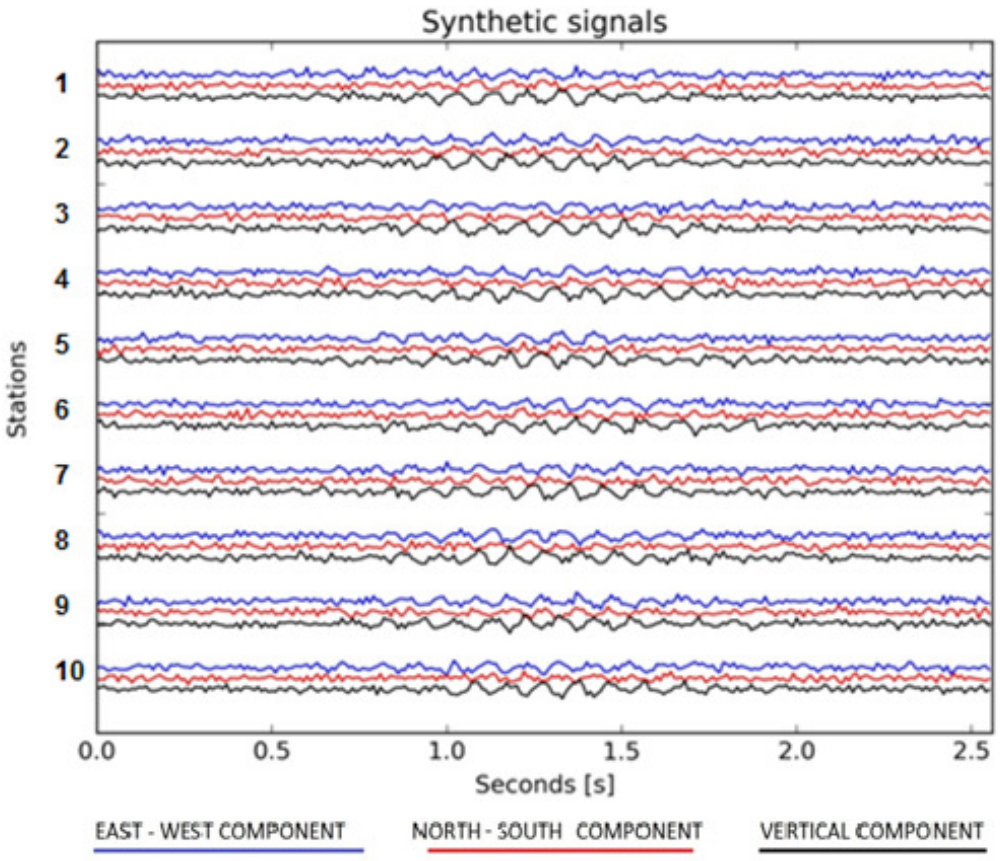


FIGURE 4.2 *Synthetic seismograms used to perform the test1.*

The test was performed setting a threshold value of 30% of the maximum wavelet coefficients amplitude. This means that wavelet coefficients having lower amplitude have been ignored because they were deemed uninformative.

The estimation of the backazimuth and slowness was performed on a grid of 15x15 nodes while the polarization inclination and azimuth was calculated on a grid of 10x10 nodes. The results were obtained in about 2 minutes of computation with a dual core CPU @ 2.5 GHz laptop computer.

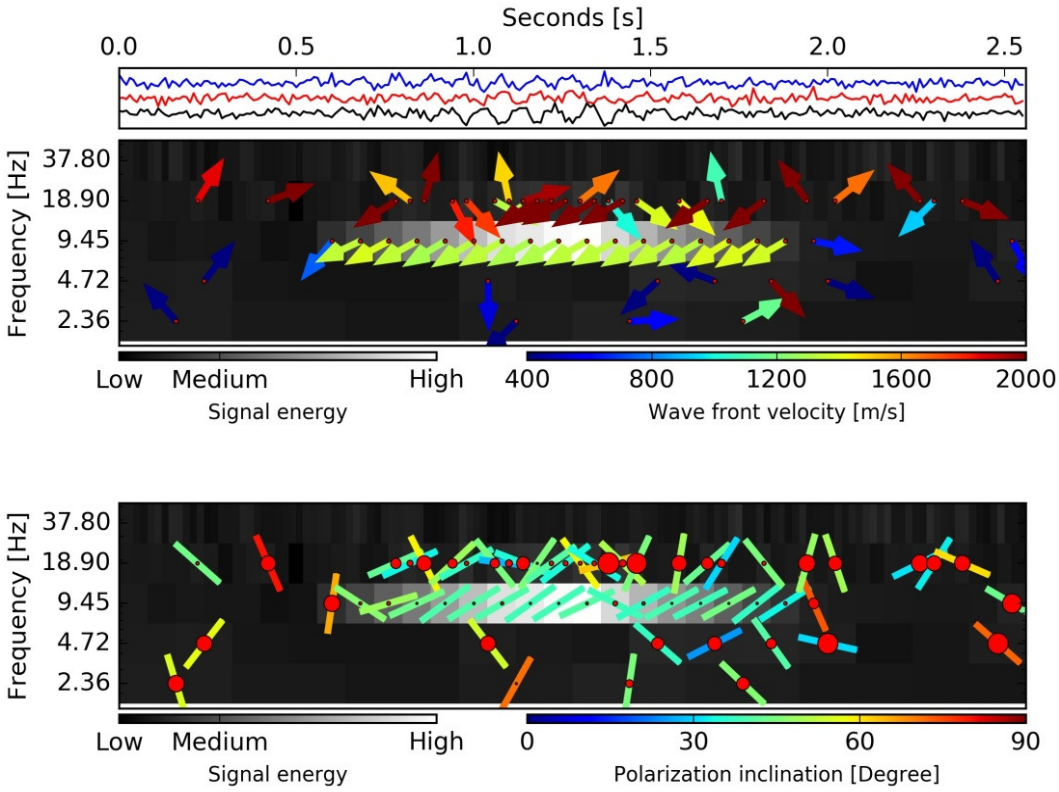


FIGURE 4.3 *Results of the first test showing the correct characterization of a weak signal with a signal to noise ratio of 1.5. The results in correspondence with the maximum energy of the signal show the presence of a P-wave having a S-W backazimuth and parallel polarization. Where the signal no noise amplitude ratio decrease the results loose coherence.*

	Frequency [Hz]	Time interval [s]	Backazimuth	Apparent Speed [m/s]	Polarization azimuth	Polarization inclination	Ellipticity
WAVEFRONT	9.4	1.3	240°	1310	239°	43°	0

TABLE 4.1 *In this table are reported the numerical results returned by the DWT-Analysis. The results refer to the interval identified by the 9.45 Hz frequency band and the 1.3 s time interval, where the signal presented its maximum energy.*

The results of the analysis are shown in figure 4.3 and in table 4.1. It is possible to see how DWT-MuSiC was able to recognize the presence of the wavefront characterizing it correctly in terms of backazimuth, slowness and polarization.

Taking a look at the results reported in figure 4.3 and table 4,1, within the interval where the wavefront energy was maximum, we can see that the analysis correctly returned a backazimuth value of 240° and an apparent speed of the wavefront of

~ 1300m/s. The apparent speed obtained in output is correctly greater than the real speed chosen to simulate the wavefront because its inclination angle of 40°.
In the lower scalogram, we can see from the red circle marker, that the results indicate not only almost linear polarization, typical of a P wave, but also a concordance in the polarization azimuth and the wavefront backazimuth.

- ***Test 2: characterization of 2 wavefronts partially overlapped having different frequency content and impinging at the array at different time.***

In this test we verified the capability of the algorithm in detecting and characterizing the arrival of two different wavefronts impinging the array at almost the same time but having different frequencies. The aim is to show how the DWT–MuSiC algorithm discriminate both the wavefronts, while the same operation is more difficult using the beamforming analysis and with the MuSiC analysis applied in the Fourier domain.
The synthetic seismograms used for this test are made by the overlap of:

1) The first wavefront associated to a P wave, speed of 900m/s, backazimuth of 240°, 45° of inclination and a frequency of 10Hz
2) The second P wavefront, speed of 900m/s, backazimuth of 150°, 35° of inclination and a frequency of 4Hz.

In figure 4.4 are reported the seismograms of the synthetic signal.

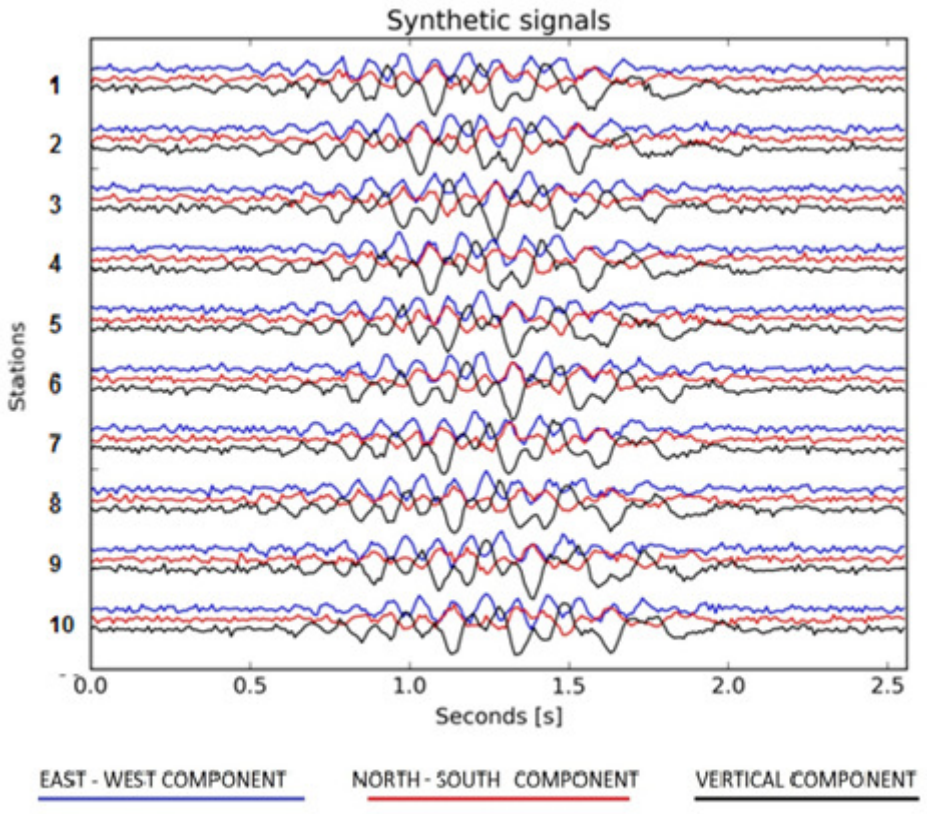


FIGURE 4.4 *Synthetic seismograms used to perform the test 2.*

The test was performed setting a threshold value where to perform the analysis, based on the 40% of the maximum wavelet coefficients amplitude. The analysis on a dual core CPU @ 2.5GHz laptop took about 2 minutes to be performed. The estimation of the backazimuth and slowness was performed on a grid of 15x15 nodes while the polarization inclination and azimuth was calculated on a grid of 10x10 nodes. The results of the DWT-MuSiC analysis is reported in figure 4.5 and in table 4.2:

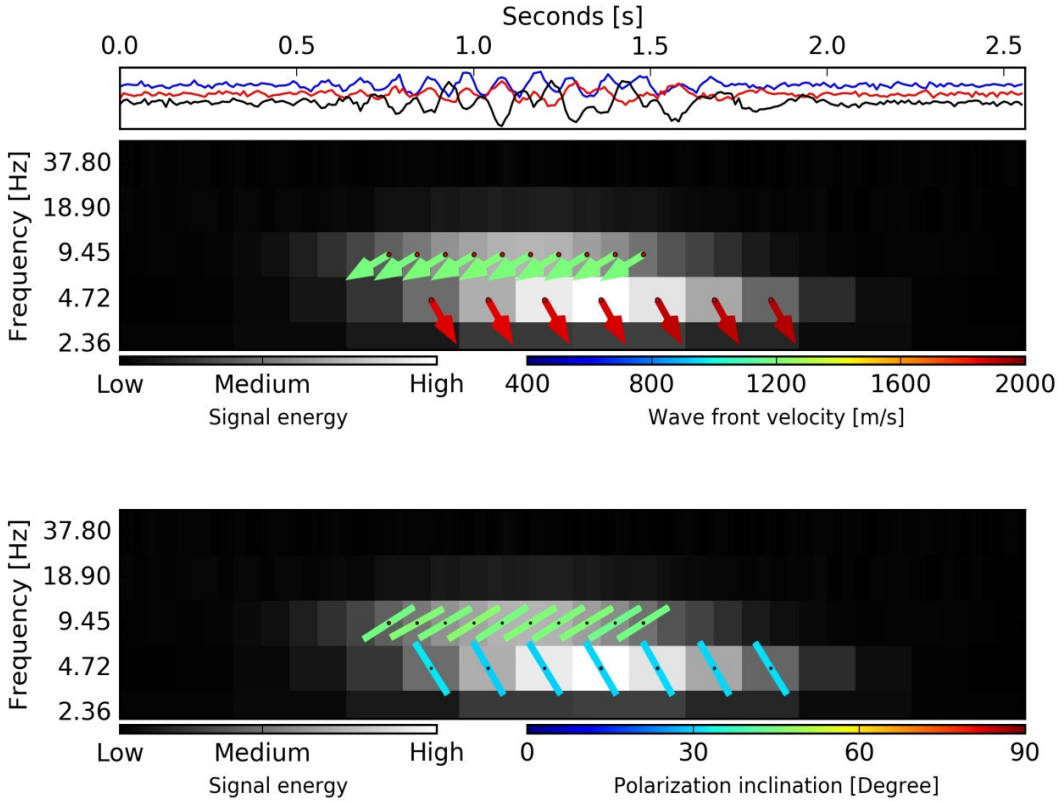


FIGURE 4.5 *Results of the first test showing the correct characterization of the 2 wavefront that partially overlap in time. The simulated P waved can be distinguish thanks to their different frequency content, permitting to isolate their own characteristics correctly.*

	Frequency [Hz]	Time interval [s]	Backazimuth	Apparent Speed[m/s]	Polarization azimuth	Polarization inclination	Ellipticity
Wavefront1	9.4	1.3	240°	1200	239	45°	0
Wavefront 2	4.7	1.4	150°	1870	151°	32°	0

TABLE 4.2 *In this table are reported the numerical results returned by the DWT-Analysis. For the 1st wavefront the results refers to the interval identified by the 9.45 Hz frequency band and the 1.2 s time interval . For the 2nd wavefront the results refers to the interval identified by the 4.72 Hz frequency band and the 1.2 s time interval. where the signal presented its maximum energy.*

Taking a look at the results reported in figure 4.5 and table 4.2, we can see that performing the analysis both the wavefronts are correctly characterized in terms of backazimuth an apparent speed.

In the lower scalogram, the polarization results indicate for both the wavefronts a linear polarization, typical of a P wave, and a concordance in the polarization azimuth with the wavefronts backazimuth. Reading the results as well as the scalogram colorbar it is possible to check the value in degrees of the polarization inclination.

Considering the same input signals, it is interesting to make a comparison with the other results obtained with a beamforming analysis, shown in figure 4.6 and the MuSiC analysis applied to the Fourier domain shown in figure 4.7.

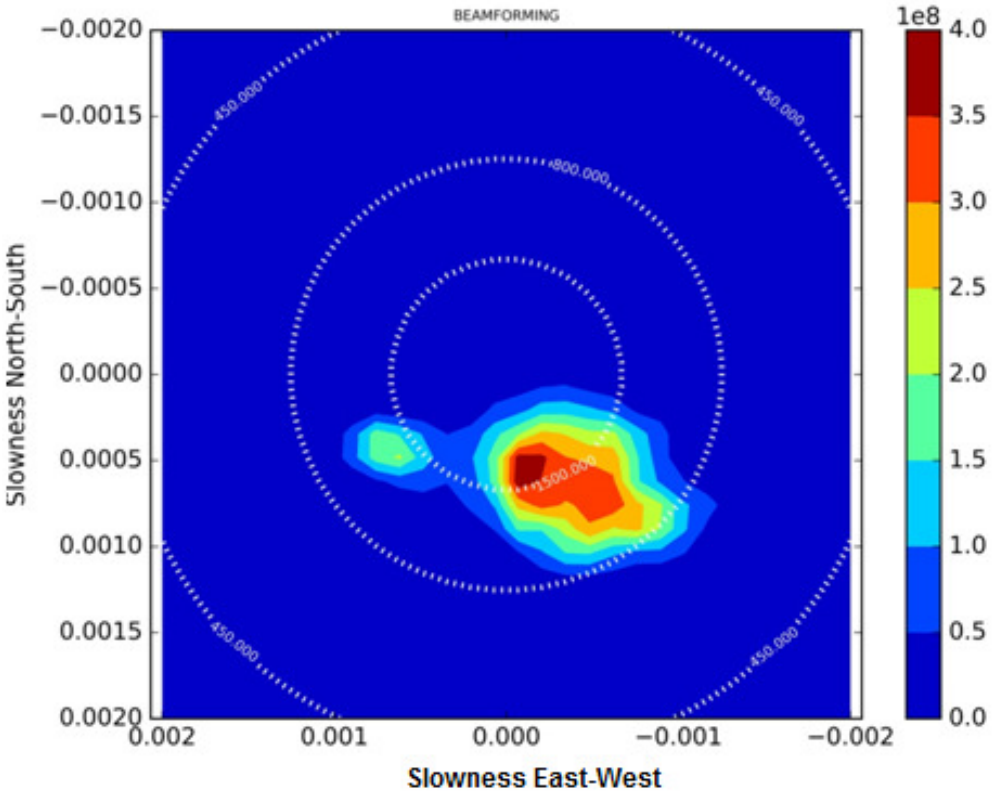


FIGURE 4.6 *Results of the beamforming analysis showing the peak in the response spectrum produced by the two simulated wavefronts. The peak of the Beamforming spectrum is in correspondence of a backazimuth of 173° and an apparent speed of 1600m/s*

As is possible to note the beamforming analysis show some limitation in returning the exact information of the two simulated wavefronts.

The analysis in fact returns a significant peak positioned in a wrong location in correspondence of the values indicated in figure 4.6. A second peak of lower

amplitude is also present, indicating the existence of another wavefront, but results shows that the presence of the wavefront having the lower frequency, caused only a smaller relative maximum in the response function. This is clearly an effect of the overlap of the 2 wavefronts that in this case can't be isolated properly to evaluate singularly their contributions.

The results moreover do not provide any information about the frequency of the wavefronts as well as about their arrival time.

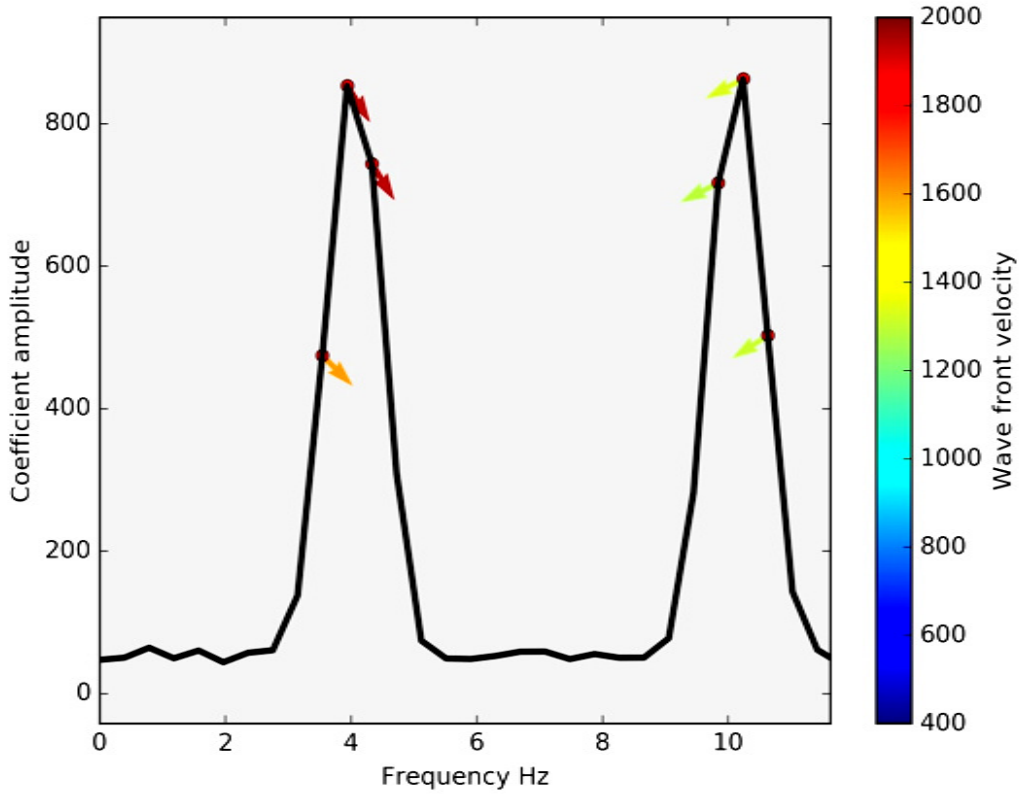


FIGURE 4.7*Results of the MuSiC analysis in the Fourier domain. It is possible to see the 2 peaks in the amplitude spectrum associated to the 2 wavefronts. The 1st wavefront, in correspondence of 3.9 Hz, is characterized with a backazimuth of 148° and an apparent speed of 1935 m/s. The with their relative characterization. The 2nd wavefront, in correspondence of 10 Hz, is characterized with a backazimuth of 242° and an apparent speed of 1310 m/s.*

The results obtained with the MuSiC applied to Fourier domain are shown in figure 4.7. The analysis identified two clear peaks in the amplitude spectrum of the signals, in correspondence of the proper frequencies of the simulated wavefronts. The analysis performed at these frequencies, as can be seen from the results, correctly characterized the wavefronts backazimuth and apparent speed even if with a lower precision than the analysis performed with the DWT-MuSiC method. This technique was able to recognize two different wavefronts having different frequencies, but cannot locate them in the time domain. This makes impossible to perform any evaluation on transient signals that changes gradually their proprieties.

- ***Test 3: Recognition of 2 wavefronts partially overlapped having the same frequency content but impinging on the array at different times.***

In this test it is highlighted the capability of the algorithm in detecting and characterizing the arrival of two different wavefronts impinging the center of the array with a difference in time of 1 second but having the same frequency. The interesting of this test is to observe how of the DWT–MuSiC algorithm recognizes correctly both the wavefronts. Also in this case it is presented the comparison between the DWT-MuSiC methodology and both the beamforming technique and the MuSiC analysis applied to the Fourier domain.

The simulated wavefronts used in this test has the following characteristics:

1) The first wavefront is a P wave, speed of 900m/s, backazimuth of 240°, 45° of inclination and a frequency of 5Hz.
2) The second wavefront is an S wave having a traversal polarization, speed of 800m/s, backazimuth of 150°, 30° of inclination and a frequency of 5Hz.

In figure 4.8 we report the seismograms of the synthetic signal.

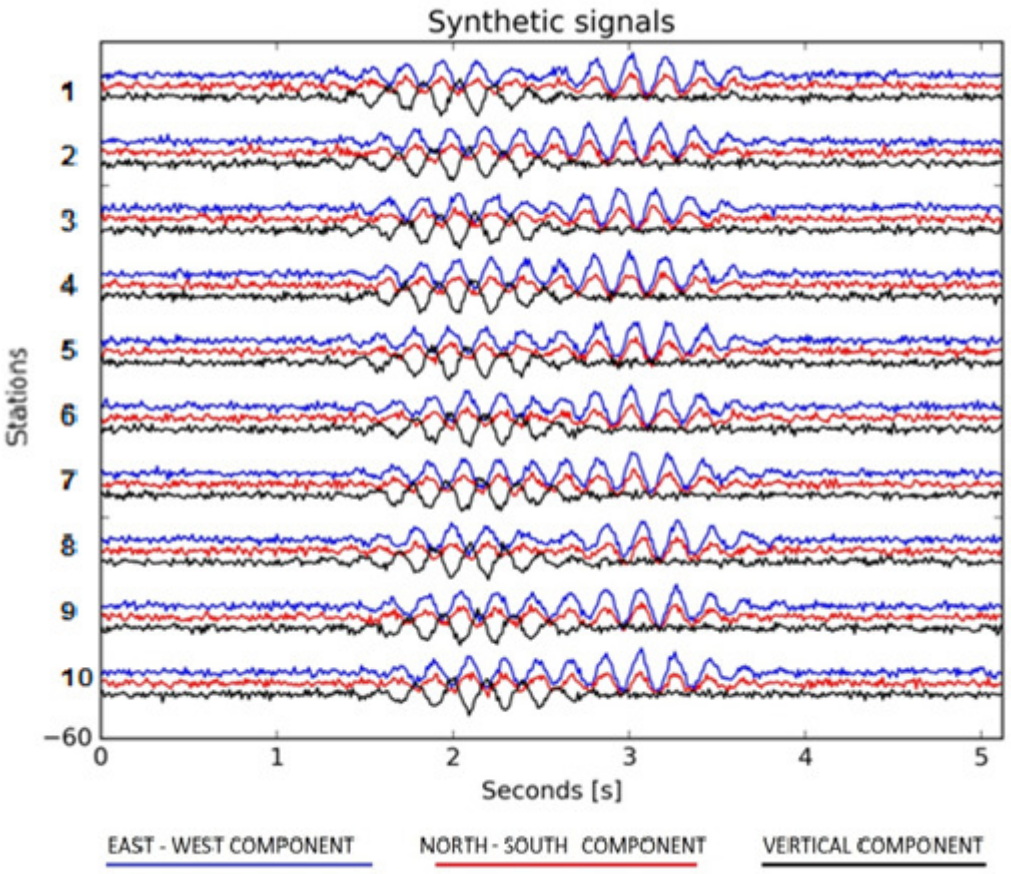


FIGURE 4.8 *Synthetic seismograms used to perform the test 3.*

The DWT-MuSiC analysis was performed choosing a threshold value of the wavelet coefficient amplitude of the 40% of the maximum value. The analysis performed on a dual core CPU @ 2.5GHz laptop took about 3 minutes to be performed. The estimation of the backazimuth and slowness was performed on a grid of 15x15 nodes while the polarization inclination and azimuth was calculated on a grid of 10x10 nodes.

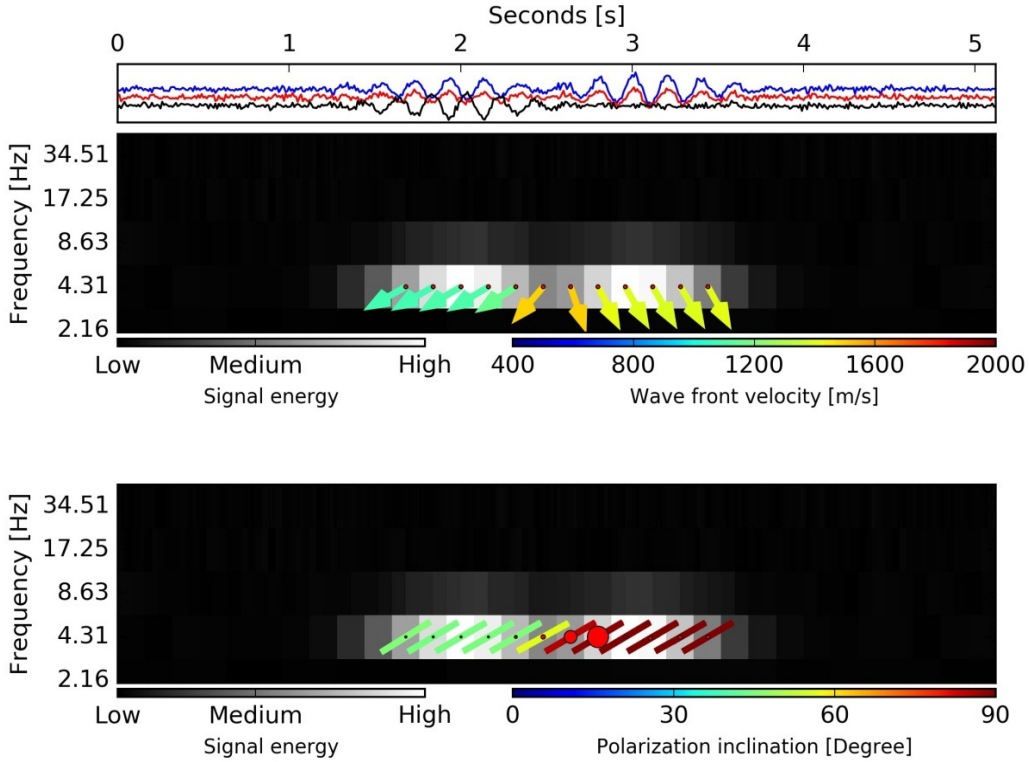


FIGURE 4.9 *Results of the first test showing the correct characterization of the 2 wavefronts having the same frequency but impinging at the array with 1s. of difference. The simulated P waves can be clearly distinguish by their difference in backazimuths, speeds and polarizations. content, permitting to isolate their own characteristics correctly.*

	Frequency [Hz]	Time interval [s]	Backazimuth	Apparent Speed[m/s]	Polarization azimuth	Polarization inclination	Ellipticity
Wavefront 1	4.3	2.0	239°	1110	240°	44°	0
Wavefront 2	4.3	2.9	150°	1390	239°	89°	0

TABLE 4.3 *In this table are reported the numerical results returned by the DWT-Analysis. For the 1st wavefront the results refers to the interval identified by the 4.3 Hz frequency band and the 2.0 s time interval . For the 2nd wavefront the results refers again to the interval identified by 4.3 Hz frequency band but at 2.9 s time interval where the latter signal presented its maximum energy.*

The results of the DWT-MuSiC analysis is reported in figure 4.9 and table 4.3. Taking a look at the results we can see that also on this test the analysis correctly returned the backazimuth and the apparent speed values relative to the simulated wavefronts.

In the lower scalogram, we can see that the polarization results indicate for both the wavefronts a linear polarization, typical of the body wave. For the first wave the results show a concordance between the polarization azimuth and the backazimuth, but for the second one we see that the polarization azimuth is orthogonal to the wavefront backazimuth. From these observations it is possible to deduce that the first wavefront is related to a P wave while the second one is an S wave. Considering the same input signals, the results obtained with a beamforming analysis, are shown in figure 4.10 while the MuSiC analysis applied to the Fourier domain are shown in figure 4.11.

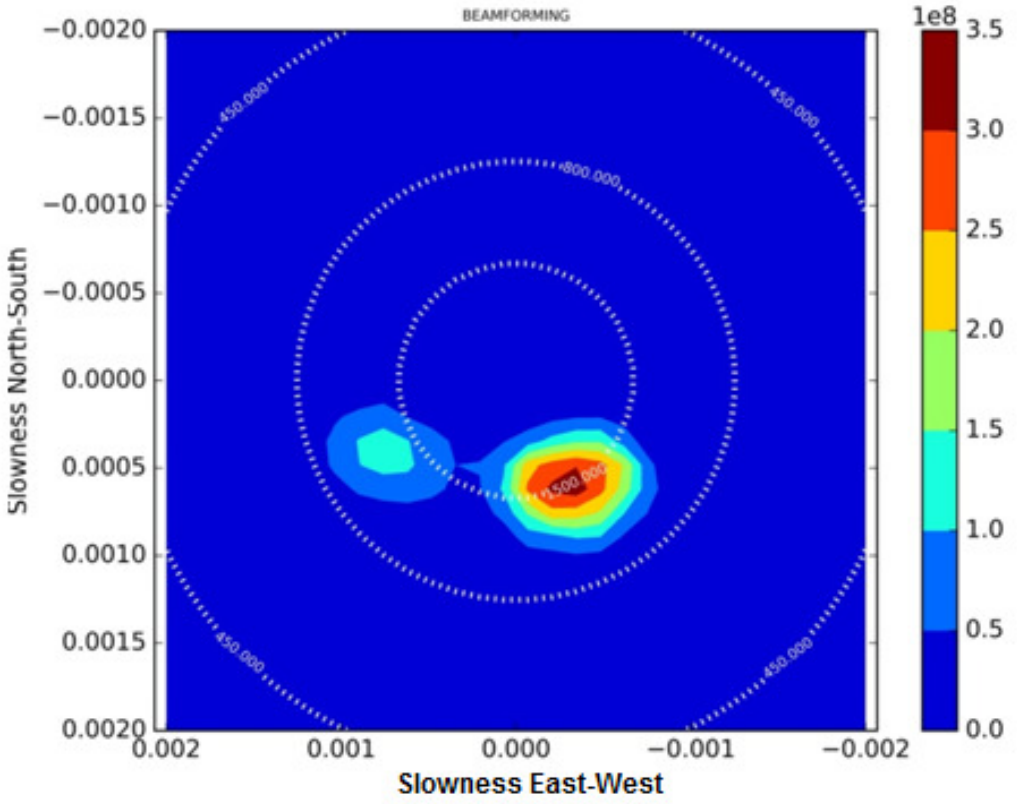


FIGURE 4.10 *Results of the Beamforming analysis applied to the synthetic signals. The peak of the response function is in correspondence of a backazimuth of 150° and an apparent speed of 1408 m/s*

As is possible to note the Beamforming analysis show also in this case some limits in returning the exact information of the 2 simulated wavefronts.

The analysis in fact, as in the previous test, returned a peak positioned in an imprecise position. The results in this case show that the analysis prioritized the

detection of the wavefront having the higher amplitude attributing only a smaller relative peak in correspondence to the other one.

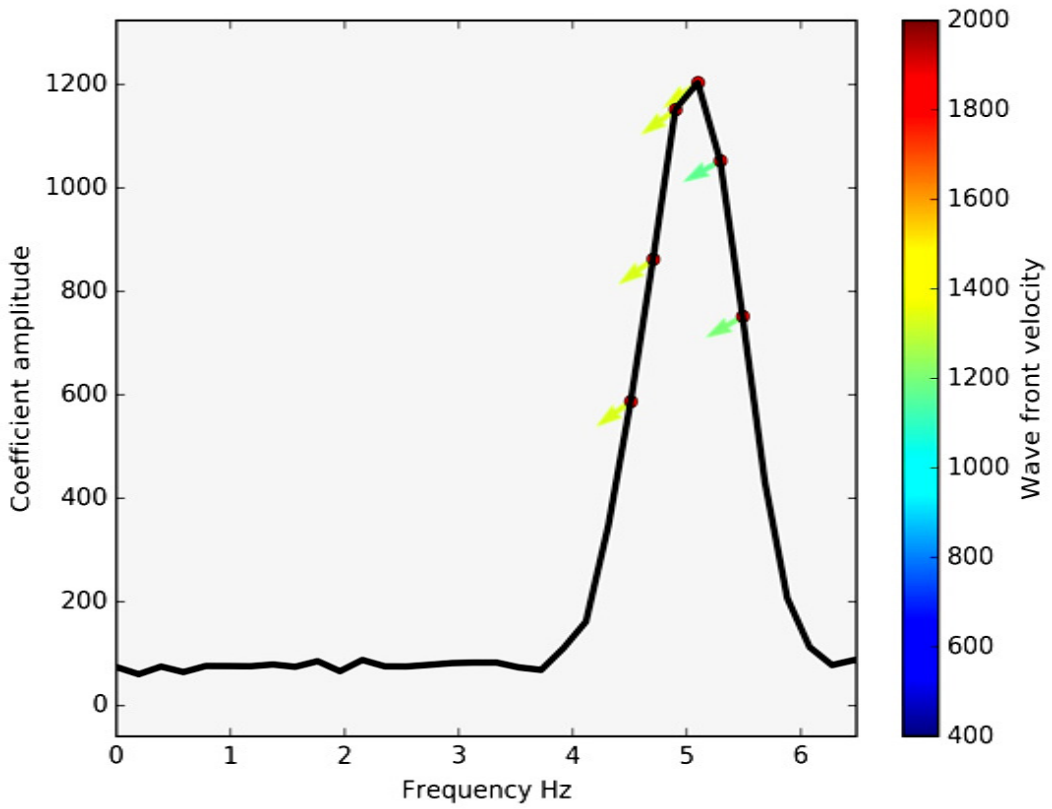


FIGURE 4.11 *Results of the MuSiC analysis in the Fourier domain. It is possible to see the peak in the amplitude spectrum associated to the 1st wavefront in correspondence of 5 Hz. The MuSiC analysis returned a backazimuth of 234° and an apparent speed of 1343 m/s. The results in this case are not precise and the second wavefront remain undetected. This is cause by the overlap in the frequency domain of the 2 wavefronts, having both 5 Hz..*

The results obtained with the MuSiC applied in the Fourier domain show only one clear pick in the amplitude spectrum in correspondence of the frequency of 5 Hz. The analysis performed at this frequency, as can be seen from figure 4.11, returned only a single value of backazimuth and slowness leaving undetected the wavefront coming from 240°. This shows that with this methodology approach, when more than one wavefront with the same frequency is present, the results are only partially correct, showing the presence in detecting one of the 2 seismic phases leaving the other ones undetected, having also the possibility to generate some artifacts in the MuSiC response function, due to the reciprocal interference.

• ***Test 4: Recognition of a transient seismogram composed by different wavefronts having different characteristics***

In this last test we simulated different kind of waves coming from the same seismic source. In particular it is first simulated the arrival of the P wave, followed by the S wave, again followed by the arrival of the Rayleigh tail. The scope of the test is to show how the DWT is able show the variation present in the seismogram due to these different waves. The simulated wavefront has the following characteristics:

1) The P wave has a speed of 900m/s, backazimuth of 130°, 60° of inclination and a frequency of 10 Hz.
2) The S wave has a speed of 700m/s, backazimuth of 130°, 30° of inclination and a frequency of 7 Hz.
3) The Rayleigh wave has a speed of 500m/s, backazimuth of 130°, 90° of inclination and a frequency of 3 Hz.

In figure 4.12 are reported the seismograms of the synthetic signal.

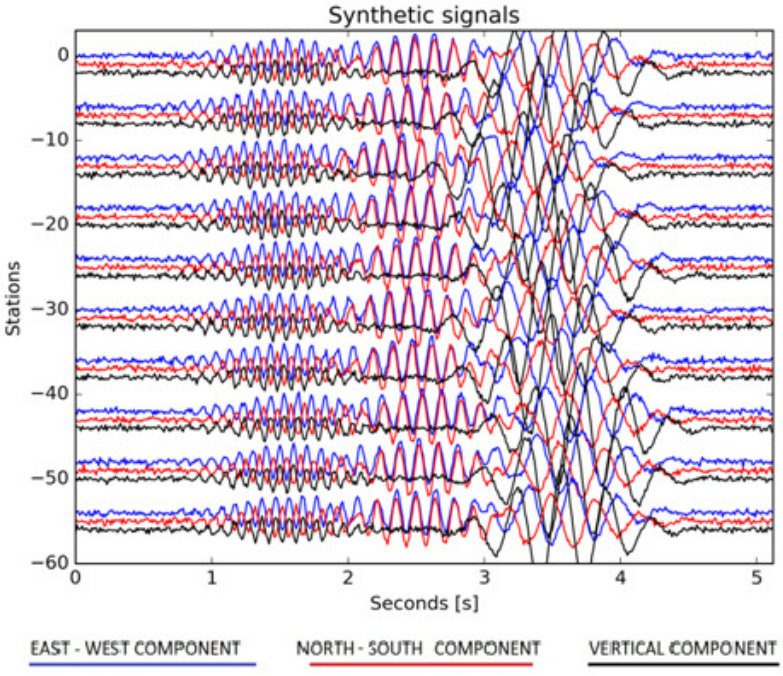


FIGURE 4.12 *Synthetic seismograms used to perform the test*

The analysis was performed choosing a threshold value of the wavelet coefficient amplitude of the 30% of its maximum value. The analysis on a dual core CPU @ 2.5GHz laptop typical laptop took 2 minutes to be performed. The estimation of the backazimuth and slowness was performed on a grid of 15x15 nodes while the polarization inclination and azimuth was calculated on a grid of 10x10 nodes. The results of the analysis are shown in figure 4.13 and in table 4.4.

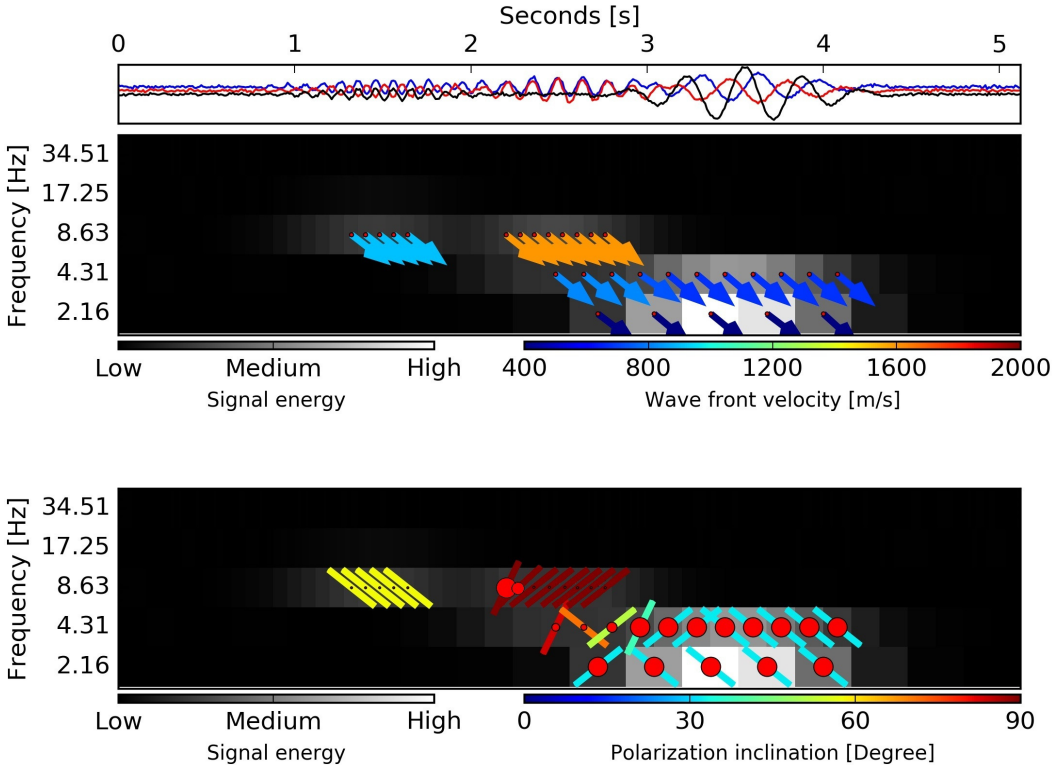


FIGURE 4.13 *Results of the DWT-MuSiC analysis showing the characteristics of the 3 different simulated wavefronts. From the left to the right is it possible to identify the P wave front*

	Frequency [Hz]	Time interval [s]	Backazimuth	Apparent speed [m/s]	Polarization azimuth	Polarization inclination	Ellipticity
Wavefront 1	8.6	1.5	239°	970	234°	60°	0
Wavefront 2	8.6	2.5	238°	1480	58°	89°	0
Wavefront 3	4.4	3.5	238°	510	238°	35°	0.85

TABLE 4.4 *In this table are reported the numerical results returned by the DWT-Analysis. The characterization data of the wavefronts is referred to the DWT intervals considered where the signals presented their maximum energy.*

As is possible to see, DWT-MuSiC is able to detect the variations in time present in the seismogram. If we take a look at the results, we can see that the analysis correctly recognized the presence of three wavefronts coming from 130° and having different characteristics. Considering together the lower and the upper scalograms it is possible to see how the firsts two wavefronts have an almost linear polarization, typical of the body waves while the third one present an high level of ellipticity indicated by the big red circle. The concordance between the polarization azimuth and the wavefront backazimuth make it possible to distinguish the P wave with the SH wave that, conversely, possess a polarization angle orthogonal to the backazimuth. All these information deduced from the results would not have been available from beamforming analysis as shown in figure 4.14. The results in fact only show one peak in the estimator function making the different phases undetectable.

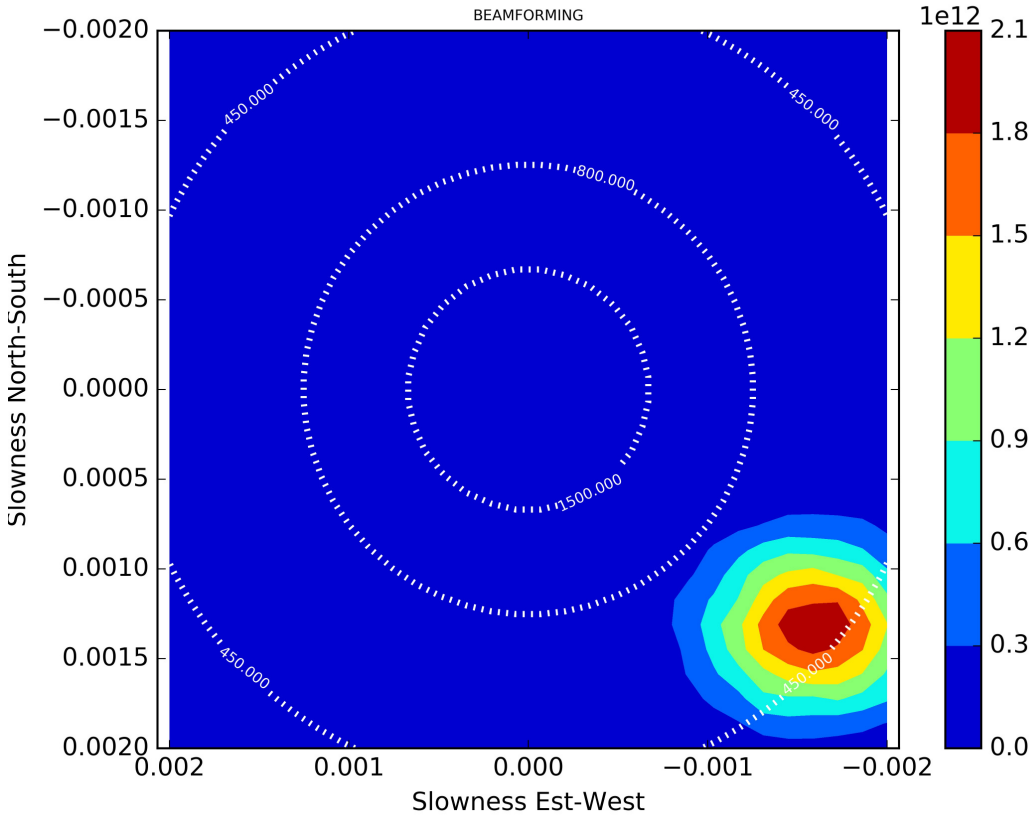


FIGURE 4.14*Results of the beamforming analysis performed on the 3 wavefronts. The Beamforming analysis returned only one peak in correspondence of a backazimuth of 129° and an apparent speed of 486 m/s*

With the MuSiC analysis applied in the Fourier domain (figure 4.15) on the contrary is possible to distinguish the three different wavefronts, but again, without the possibility to collocate it in the time domain it is impossible to deduce even the basic feature of the seismic signal.

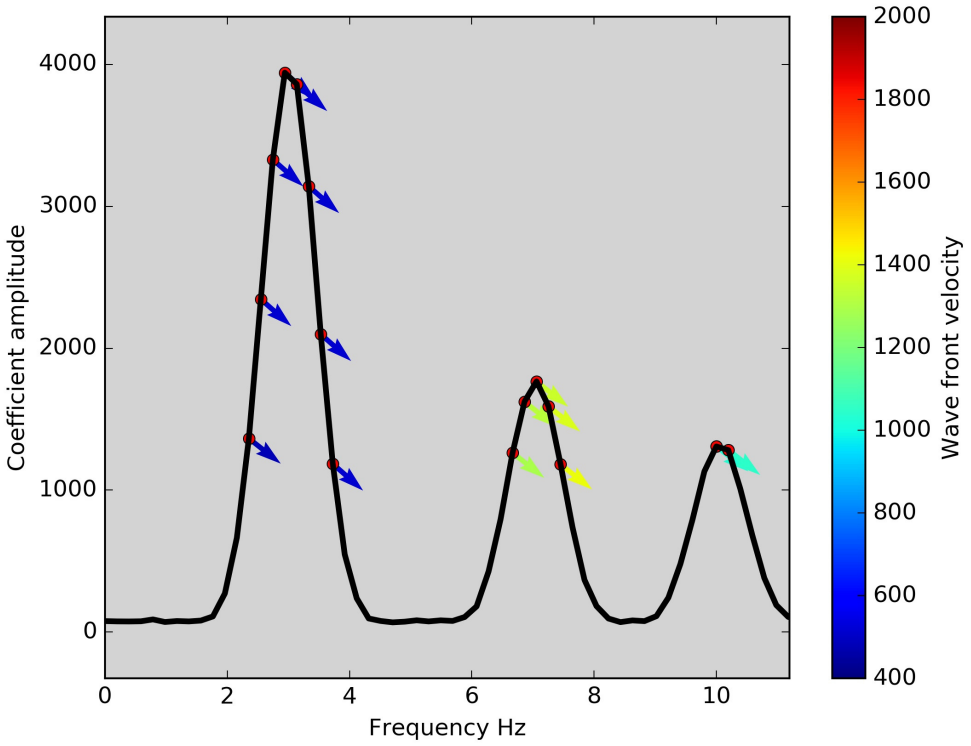


FIGURE 4.15 *Outputs obtained performing the MuSiC analysis in the Fourier domain. It is possible to see the three peaks associated to the 3 wavefronts with the returned characterization. The first peak at 3 Hz returned a backazimuth of 129° and an apparent speed of 507 m/s. The second peak at 7 Hz returned a backazimuth of 129° and an apparent speed of 1352, while the 3rd one at 10 Hz returned a backazimuth of 130° and an apparent speed of 1020 m/s*

4.2 APPLICATION TO A VOLCANO-TECTONIC EVENT RECORDED AT MOUNT VESUVIUS

Mt. Vesuvius is one of the most dangerous volcanos in the world due to its proximity to the urban area of Naples and its volcanic activity is characterized by explosive eruptions.

Mount Vesuvius background seismicity has been described in various studies [Castellano et al., 2002; Del Pezzo et al., 2004, and references therein], and to monitor its activity, Vesuvius Observatory, operates a continuous surveillance for the detection of possible precursors of eruptions.

Together with different geophysical and geochemical instruments and a seismic network also array are used for monitoring purposes, especially to detect non impulsive sources generating signals, like volcanic tremor and LP events where the evaluation of the wavefronts characteristics offer a powerful tool to estimate the temporal and spatial evolution of seismic source [Bianco et al., 2005].

Examples of application of these methods to volcanic signals have been reported, for example, for Kilauea [Saccorotti et al., 2001], Stromboli [Chouet et al., 1997, Deception Island [Ibanez et al., 2000], and Mt. Vesuvius [Saccorotti et al., 2001].

The array analysis also helps in discriminating artificial seismic from natural events simply determining if the backazimuth of a seismic phase point or not toward the volcano.

The application of the DWT-MuSiC to Mount Vesuvius array data consisted in the analysis of seismograms registered with a small-aperture array that has been installed from 2006 to 2010 on the SW side of Mt. Vesuvius, to improve the seismic monitoring of the volcano. It was composed of 14 multi component Lennartz LE-3D seismometers.

The figure 4.16 shows the configuration of the array and its location respect to the crater.

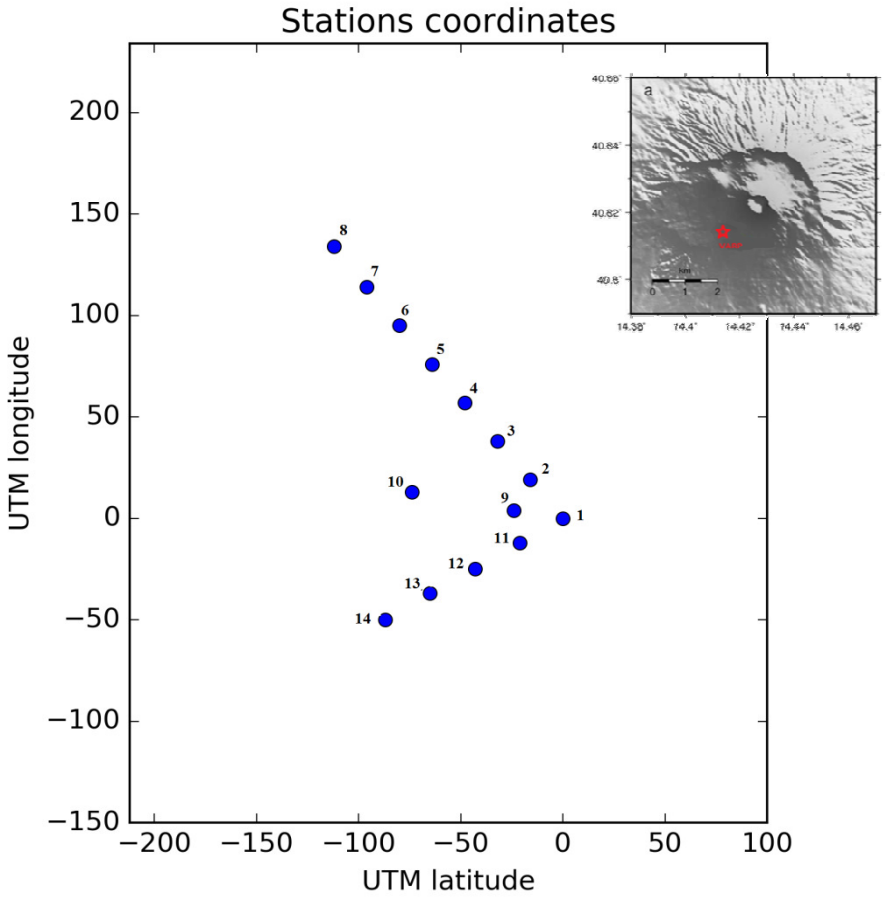


FIGURE 4.16 *Plan view of the array used to acquire the seismic data. In the upper right corner the position relative to the Mont Vesuvius crater.*

The 11/09/2009 at 21:06 UTC, the array registered an event of magnitude Md=1.0 that is reported in figure 4.17.

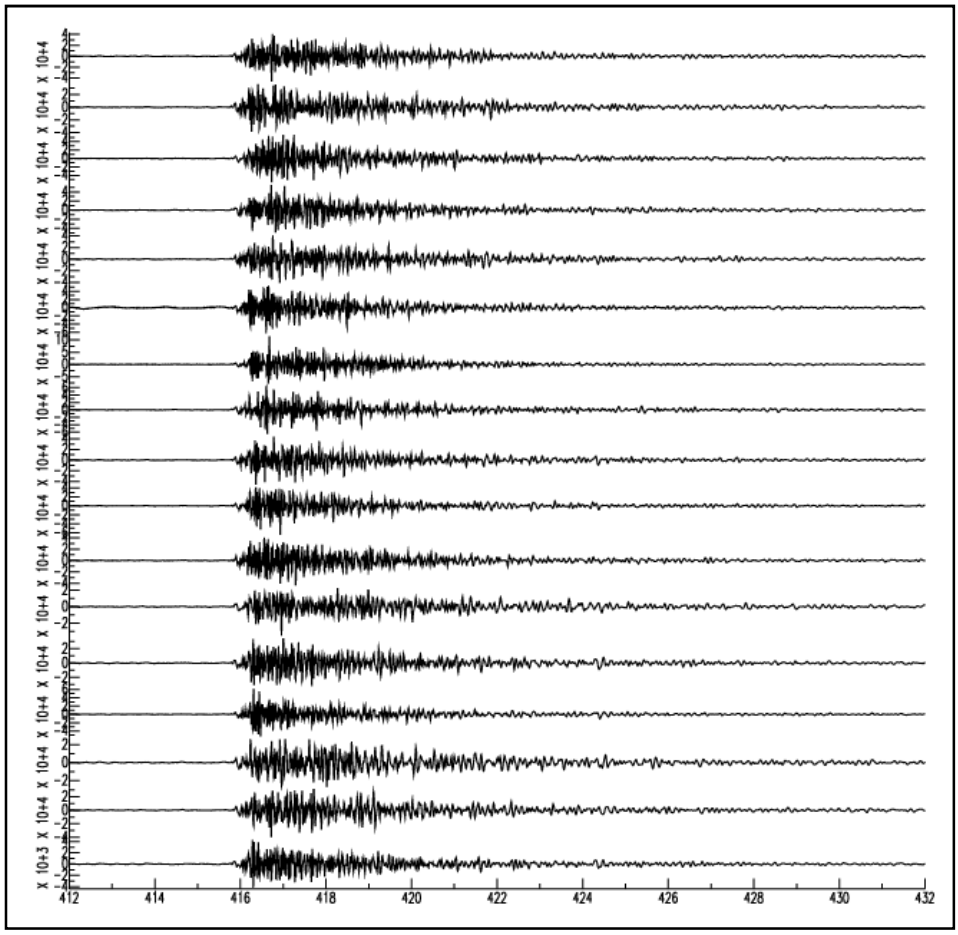


FIGURE 4.17 *Seismograms of the event recorded the 11/09/2009 at 21:06 UTC.*

The analysis was performed on a 2.5 s window on the seismograms pre-processed to remove the mean value. The input seismograms used for the analysis are shown in figure 4.18.

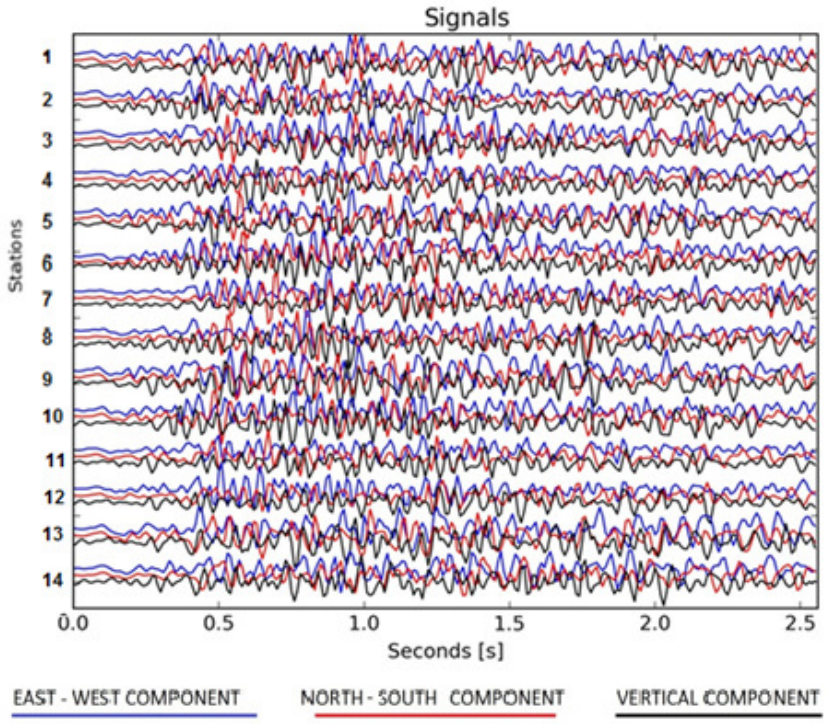


FIGURE 4.18 *Seismogram after pre-processing*

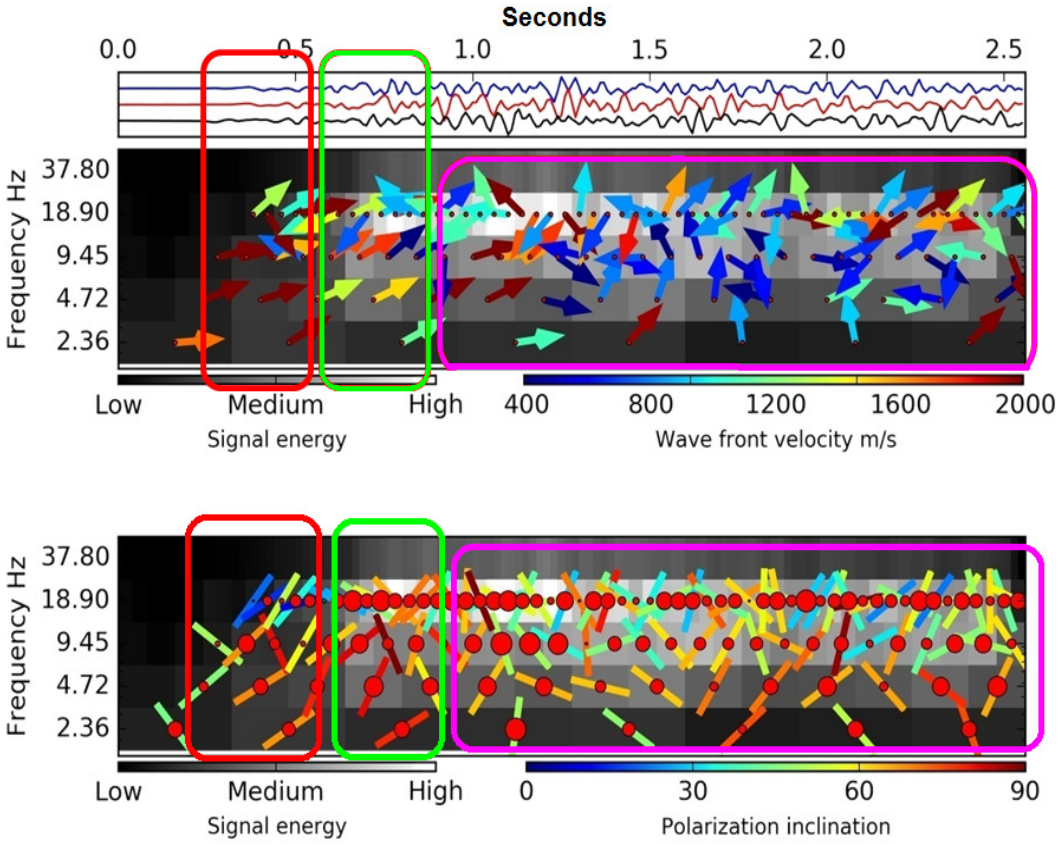


FIGURE 4.19 *Results of the DWT-MuSiC analysis showing three different identified phases. In red the P waves, in green the S waves and in magenta the coda waves.*

The results of the analysis indicate that the first phases identified, highlighted with a red box in figure 4.19 is associated to P waves, which propagates with an apparent velocity greater than 2000 m/s, with a backazimuth of ~80°, or in other words pointing toward the crater area. This is in agreement with the locations obtained for the same events using the ordinary location technique applied to the data recorded by the permanent seismic network. We can observe the increment of the coefficient energy interpreted as the arrival of the S wave and having a greater amplitude (Green box in figure 4.19)

The S waves shares the same backazimuth of the P waves but have a lower apparent speed. The polarization angle moreover is deeper and its azimuth is orthogonal to the backazimuth indication, as expected for S waves.

The coda part of the seismogram presents highly scattered backazimuth values, revealing the absence of a predominant direction of propagation of this wavefield (Magenta box in figure 4.19). This is in accordance with previous studies concerning the wavefield of coda waves at Mt. Vesuvius. [e.g. Saccorotti et al.,2001 ; Tramelli et al., 2009]. However they show that DWT-MuSiC has the ability to identify and discriminate seismic phases, proving a set of important parameters associated with them. This suggests that results of DWT-MuSiC can be used as a starting point for the development of methods for the automatic detection of P and S phases from seismic array recording of earthquakes.

4.3 APPLICATION TO AN ACTIVE SEISMIC EXPERIMENT AT KRAFLA CALDERA (ICELAND)

Krafla Caldera in northern Iceland is a well-monitored and extensively drilled caldera system that underwent a major rifting and eruption episode in 1975 to 1984. In 2009, during the execution of a well, for the Iceland Deep Drilling Project (IDDP), the borehole IDDP-1, aimed to reach supercritical fluids at 4 km depth in Krafla Caldera, unexpectedly encountered rhyolite melt at 2.1 km depth, that was not hypnotized to be there [e.g. Hólmgeirssona et al., 2014]. The drill site in fact was chosen paradoxically because magma was not expected at shallow depth, based on the occurrence of seismicity to twice that depth beneath the site during the last eruption, and on 3-D resistivity structure.

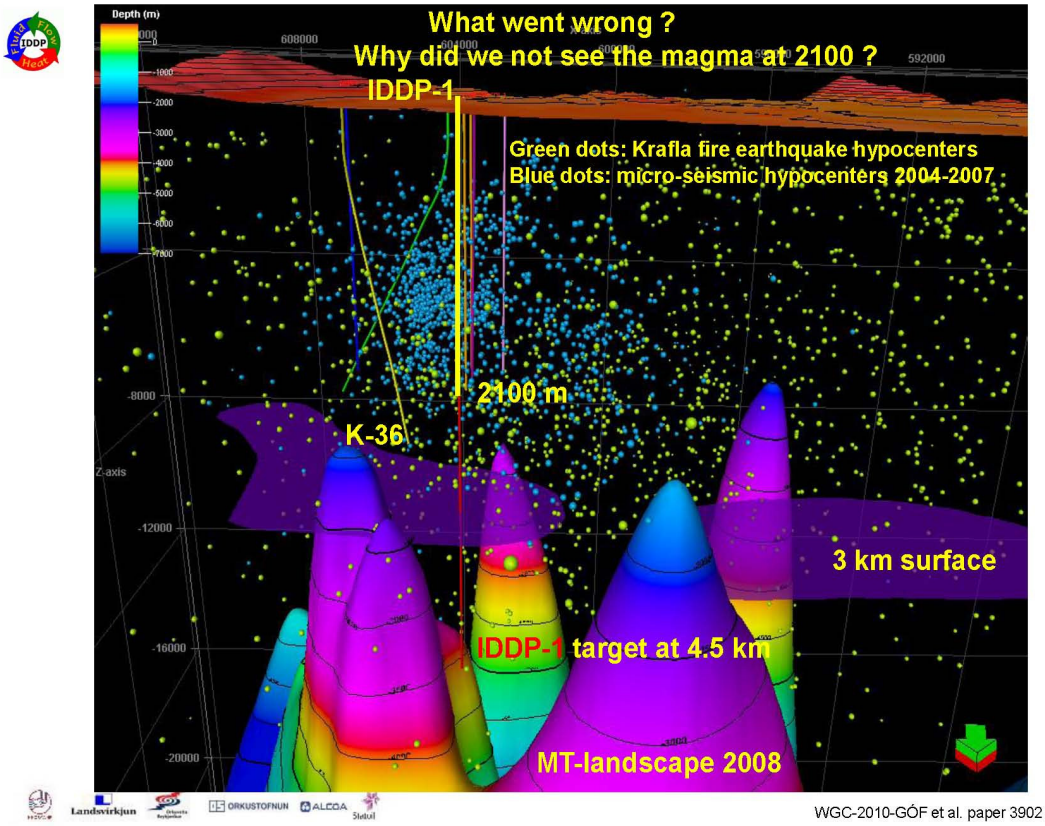


FIGURE 4.20 *3D representation of the IDDP-1 borehole trajectory and the magmatic chamber encountered.*

This discovery opened the way to the project KMDP finalized to understand the thermal, chemical, and mechanical behavior of the active magma system discovered by IDDP-1 and the ductile rocks enclosing it. [e.g. Elders et al. ,2014; Friðleifsson et al., 2014]

Further drillings have been conducted in order to improve the understanding of volcanic hazards in calderas and better interpretation of precursory deformation and seismic signals that may herald eruption.

Part of the experiment was carried out by INGV in collaboration with University of Iceland and ISOR on 6-16 September 2015, and consisted in a seismic survey aimed to determine the geometry and the physical properties of the shallow (2.1 km) magma chamber through the analysis of its reflected phases.

During the experiment were positioned 30 seismic stations, 2 seismic arrays and were exploded59 artificial shots in order to collect enough data.

We checked the capability of the DWT-MuSiC to detect some P-to-S converted seismic phases, which can be associate to the top of the magma chamber. We used the seismogram acquired by one of the arrays, called the array “B”.

The array had an aperture of ~140m and the stations were located as in figure 4.21.

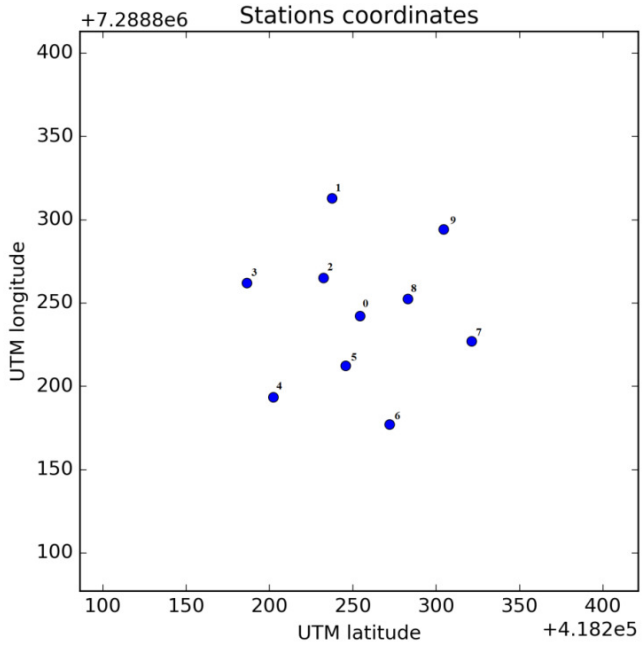


FIGURE 4.21*Plan view of the array B used for the active seismic survey at Krafla caldera.*

At this array were registered different shots identified by the letter T and a progressive number, having the source location as shown in figure 4.22

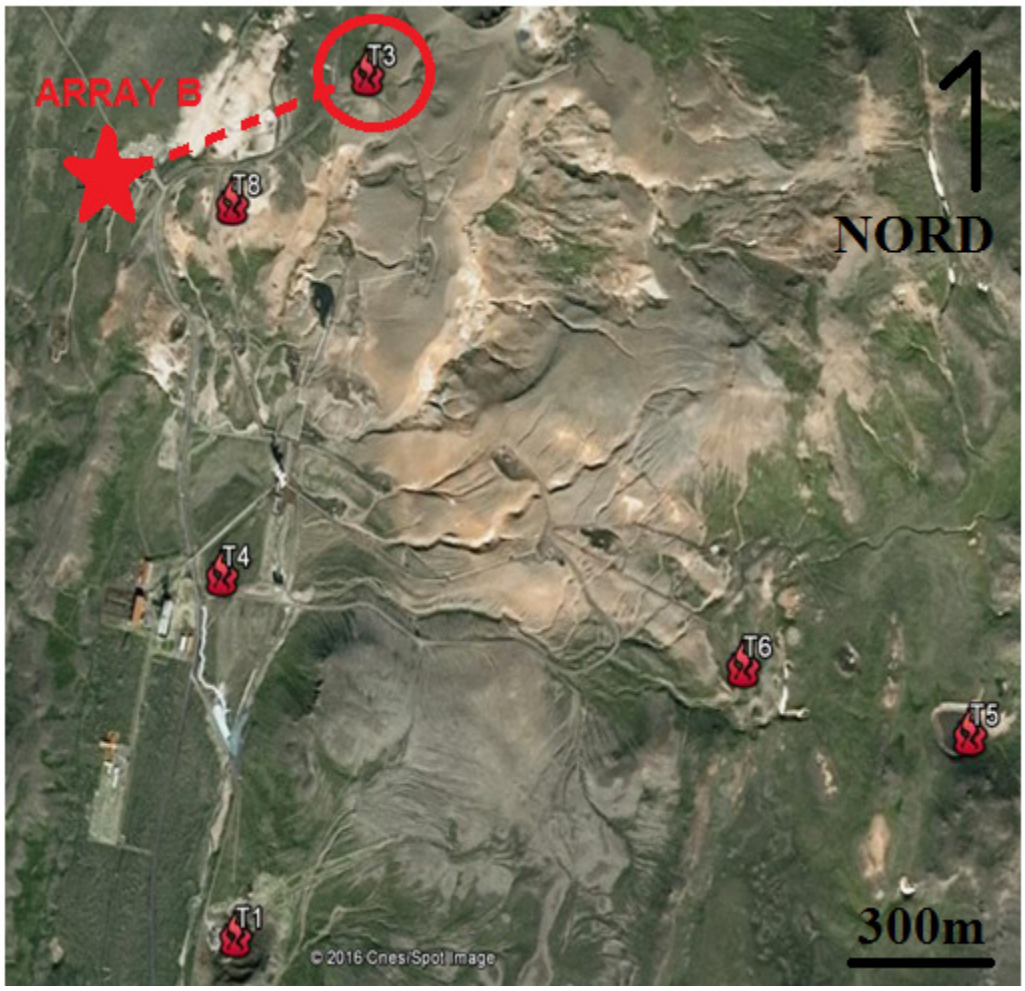


FIGURE 4.22 *Map of the array B location and the position of the shots.*

The case presented in this thesis is the seismic acquisition associated with the T03 shot. The motivation in choosing this shot was based on the better quality of the signal and the bigger distance from the shot to the array, that better approximates the far field condition necessary for the array analysis.

We show in figure 4.23 the seismograms registered at the array stations:

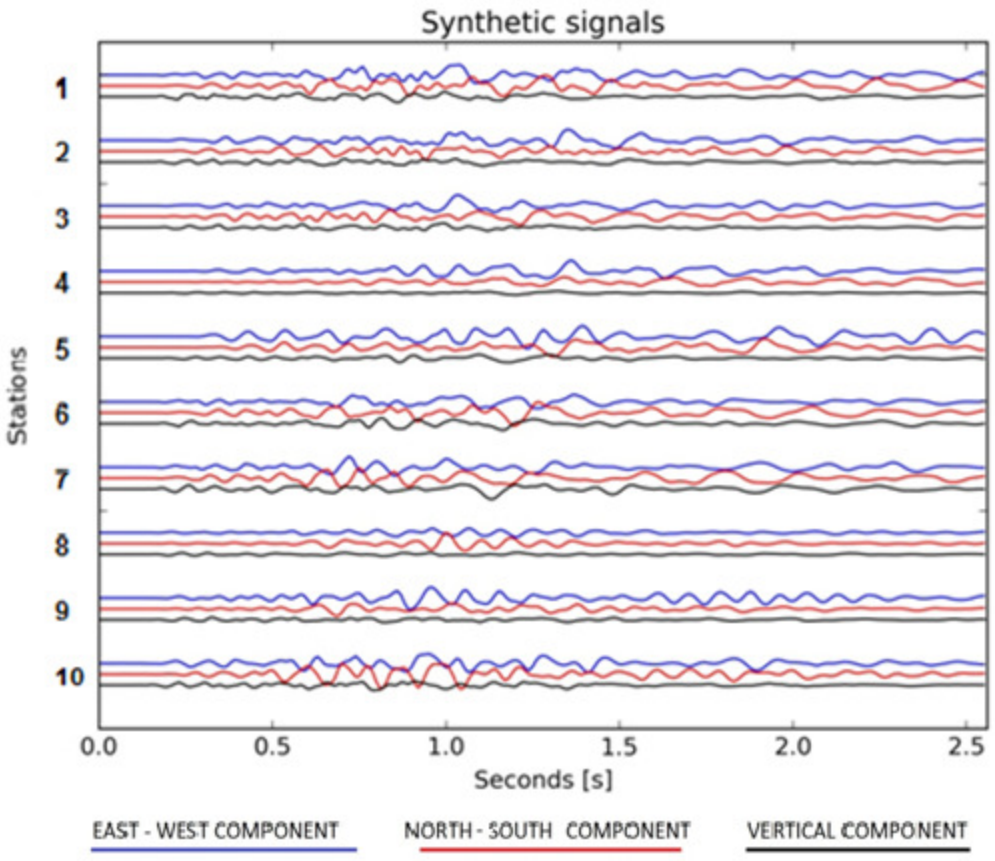


FIGURE 4.23 *Seismograms relative to the shot T3 , used to perform the test. The used seismograms are relative to a time window of 2.5 seconds. The raw signals before being used has been optimized to remove the mean error on the tracks and normalize to level out the relative amplitudes.*

We selected a 2.5 s window at the beginning of the signal. After having removed the mean on the signals we performed the DWT-MuSiC analysis obtaining the following results (figure 4.24):

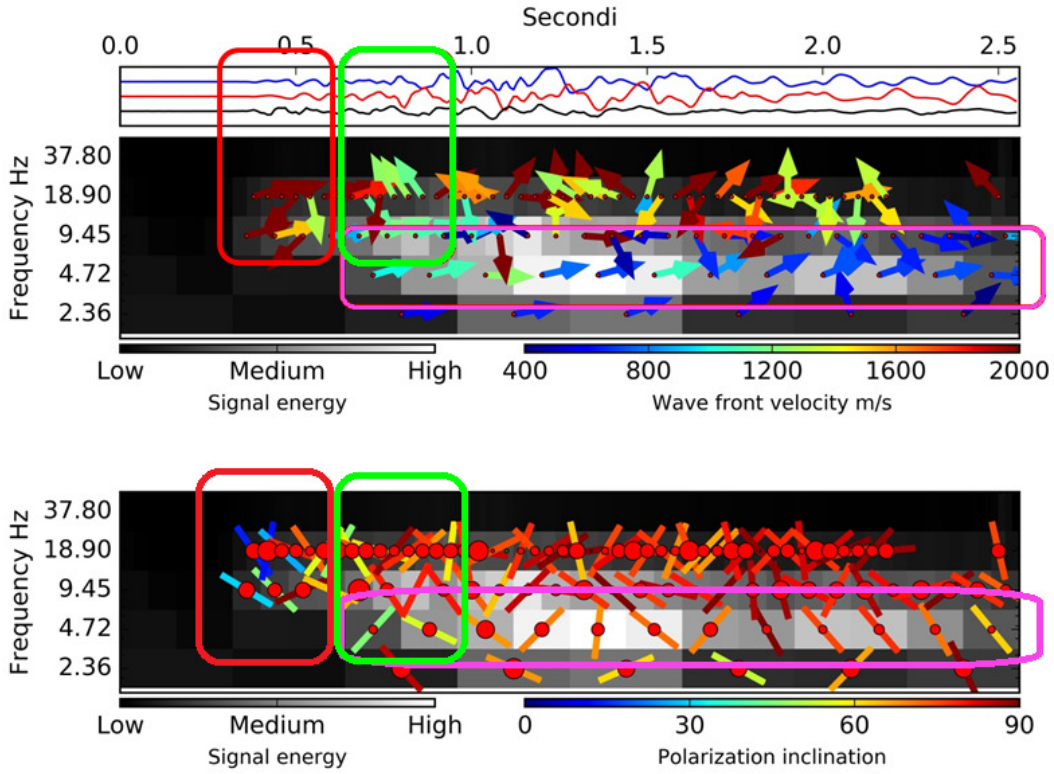


FIGURE 4.24 *Results of the DWT-MuSiC analysis showing three different identified phases. In red the P waves, in green the S waves and in magenta the coda wave.*

From the analysis results it is possible to note a general trend of the wavefronts that show a coherent backazimuth direction of 70°-80° in accord to the relative position of the shot T03 respect to the array. We can also observe that the greatest part of the signal energy is localized on the frequency band between 3 and 10 Hz, in the time intervals comprise to 1.0 and 1.5 s. Looking at the results more in details it is possible moreover to distinguish different seismic phases. In particular is possible to highlight some coherent phases located at the beginning of the seismogram (highlighted in red in the figure 4.24) that can be associated to the first arrival of the P waves. This can also be observed in the same position in the polarization results where, it is possible to observe a low polarization inclination angle and a concordance in polarization azimuth and wavefront backazimuth.

Proceeding the analysis of the seismogram it is possible to highlight also another group of seismic phases (green circled in the figure 4.24) that may represent the reflected waves from the magmatic chamber.

These phases have a Nord-West backazimuth in contraposition to the general East-North-East trend. The resulting speed is lower and also the polarization azimuth, shows a direction which differs to the wavefront backazimuth. The third part of the seismogram that is possible to identify, are the coda waves (circled in blue) that present a lower frequency and a lower speed of propagation. They are likely to be associated to surface waves generated by the shot.

We can affirm that the S wave evidenced by the analysis were likely to be generated at the top of the magmatic chamber because the artificial source used for this survey, in absence of a discontinuity is only capable producing P and Rayleigh waves. In conclusion it is possible to affirm that even if the condition to perform this kind of analysis was not optimal, due to the complex wavefield and the low signal/noise ratio, it is possible to use the DWT-MuSiC as a tool to highlight the presence of some specific phase within signals generated by artificial shots. In this case the use of DWT-MuSiC clearly evidenced the presence of a converted P-to-S seismic phase. The analysis of all the dataset would provide important information about the position of the magmatic chamber and in delineating its shape.

CONCLUSIONS

The main purpose of this PhD project was develop an innovative methodology for the seismic array data analysis able to return a complete characterization of the seismic wavefield and useful to be used in different geophysical contests like volcanic monitoring, identification of underground structures, characterization of ambient noise, identification of wake seismic signal etc.

The implemented methodology, named DWT-MuSiC (Discrete Wavelet Transform - Multiple Signals Classification), has shown to be able to perform near-real time analysis providing, in different contests and even in presence of multiple seismic sources, the detection of the seismic phases and the characterization of their parameters, in terms of direction of arrival (backazimuth), apparent speed of propagation and polarization. The method, moreover, being based on the discrete wavelet decomposition of the seismogram, is also capable to reference the results in terms of time and frequency content, making in this way possible the analysis of transient seismic signals.

The development of the DWT-MuSiC algorithm was made completely from scratch, creating a tool with *Python*, a modern and versatile open-source programming code that, and the time of writing, this tool is intended to be distributable in the next future to the scientific community through a stand-alone package to be used in various geophysical contests.

The realization of the thesis was split in three different phases:

1) The design and development of the chosen methodology in *Python* language
2) Its validation by means of editable synthetic signals, specifically generated by new Python language scripts
3) The execution of some comparison tests between the DWT-MuSiC algorithm and different array analysis techniques, also in this case implemented in *Python* language from scratch
4) The application of the DWT-MUSIC to 2 different real cases; the first to volcano-tectonic array data registered at Mount Vesuvius, Italy, and the

second to data acquired during an active seismic survey at Krafla caldera, Island.

The synthetic tests performed with the DWT-MuSiC analysis have shown the ability of the algorithm to detect wake signals also on signals that present to noise ratio of 1.5. The tests moreover verified that the method returned correct information in those cases where the analysis is performed on multiple seismic phases that overlap in terms of time of registration or / and frequency content of the signals themselves. The comparison between other methodology like the beamforming and MuSiC itself applied in the Fourier domain, shown how the information returned by the DWT-MuSiC was much more complete.

The use of discrete wavelets and the optimization algorithm implemented in the analysis permit moreover to save computational time analyzing only the important data, allowing the recovery of hidden information when more than one wavefronts overlap.

The application to the real cases was important to test the applicability of the methodology in different contests. The analysis performed on the data collected at Mount Vesuvius has shown that the registered event was confirmed to be a natural volcano-tectonic event, excluding the possibility of an artificial explosion. This was possible collocating the direction of source from the backazimuth analysis and identifying different seismic phases in the seismogram typical of the seismicity of the area. The analysis of the array data collected at the Krafla caldera, confirmed the presence of some reflected S phase in the seismogram, caused by the presence of a shallow magmatic chamber. The evidence of the reflected S phase suggests the possibility to use the DWT-MuSiC as a support methodology in acquiring useful information to map the magmatic chamber itself.

To conclude, the DWT MuSiC methodology gave promising results and it will be interesting to test possible further application of the methodology that could involve also other field of application like detection of local micro-earthquakes, real-time seismological volcano monitoring , industrial applications (e.g. seismic-

while-drilling, wide-angle exploration), or also engineering applications like detection and characterization of nuclear detonations.

BIBLIOGRAFY

T. D. Abhayapala, R. A. Kennedy (2000). Nearfield broadband array design using a radially invariant modal expansion. J. Acoust. Soc. Am. 107 (1), January 2000.

P. S. Addison (2005). Wavelet transforms and the ECG: a review. Physiol. Meas. 26 (2005), R155-R199

D. Alfsmann, Heinz G. Göckler, Stephen J. Sangwine, Todd A. Ell (2007). Hypercomplex algebras in digital signal processing: benefits and drawbacks. Signal Processing Conference, 2007 15th European

J. Almendros, J.M. Ibanez, G. Alguacil, E. Del Pezzo, R. Ortiz, Array tracking of the volcanic tremor source at Deception Island, Antarctica (1997). Geophysical Research letters, vol.24, no.23, 3069-3072, December 1, 1997.

J. Almendros, J.M. Ibanez, G. Alguacil, E. Del Pezzo (1999). Array analysis using circular-wave-front-geometry: an application to locate the nearby seismo-volcanic source. Geophys. J. Int. 1999, 136, 159-170.

S. Argentieri, P. Danès (2007). Broadband Variations of the MUSIC High-Resolution Method for Sound Source Localization in Robotics. IEEE/RSJ International Conference on Intelligent Robots and Systems San Diego, CA, USA, Oct 29 - Nov 2, 2007.

R. Arlitt, Kissling E., Ansorge J., TOR working group (1999). 3-D crostal structure beneath the TOR array and effects on teleseismic wavefronts. Tectonophysics 314, 309-319, 1999.

N. M. Astaf'eva (1996). Wavelet analysis: basic theory and some applications. Physics-Uspekhi 39(11), 1085-1108.

M. S. Bartlett (1948). Smoothing periodograms from time series with continuous spectra. Nature, *161*(4096), 686-687.

T.Bartosch, J.Wassermann (2004). Wavelet Coherence Analysis of Broadband Array Data Recorded at Stromboli Volcano, Italy. Bulletin of the Seismological Society of America, Vol. 94, No. 1, pp. 44–52, February 2004.

F.Bianco, P. Cusano, S. Petrosino, M. Castellano, C. Buonocunto, M. Capello, E. Del Pezzo (2005). Small-aperture Array for Seismic Monitoring of Mt. Vesuvius. Seismological Research Letters, Volume 76, Number3, May/June 2005

A. Bottero, Y. Cansi, B. Massinon (1994). Automatic processing of seismic events recorded on a mini-array. Signal analysis combined with neural networks. Annali di geofisica, Vol. XXXVII, N.5, September 1994.

T.F. Brooks, , W. M. Humphreys (2006). A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays . Journal of Sound and Vibration. Volume 294, Issues 4–5, July 2006, Pages 856–879

Y. Cansi (1995). An automatic seismic event processing for detection and location: The P.M.C.C. method. Geophysical research letters, vol.22, no.9, 1021-1024, May 1, 1995

M. Castellano, Buonocunto, C., Capello, M., & La Rocca, M. (2002). Seismic surveillance of active volcanoes: the Osservatorio Vesuviano Seismic Network (OVSN, southern Italy). Seismological Research Letters, 73(2), 177-184.

J. Capon. 1969. high-Resolution Frequency-Wavenumber Spectrum analysis. Proc. IEEE,57(8), 2408-2418

F.J. Chavez-Garcıa, F. Luzon (2005). On The Correlation Of Seismic Microtremors. Journal of Geophysical Research, Vol. 110, B11313,2005

S-J. Chiou, B. A. Bolt (1993). Seismic wave slowness-vector estimation from broad-band array data. Geophys. J. Int. 114, 234-248

B. Chouet, G. Saccorotti, M. Martini, P. Dawson, G. De Luca, G. Milana, R. Scarpa (1997). Source and path effects in the wave fields of tremor and explosions at Stromboli Volcano, Italy. Journal of geophysical research-all series-, 102, 15-129.

P.-J. Chung (2007). Array processing method introduction. Academic Press Library in Signal Processing: Array and Statistical Signal processing, Chapter 14.

L. D'Auria, F. Giudicepietro, M. Martini, M. Orazi, R.Peluso, G.Scarpato (2010). Polarization Analysis in the Discrete Wavelet Domain: An Application to Volcano Seismology. Bulletin of the Seismological Society of America, Vol.100, No.2, pp.670-683,April 2010.

B. Darren, Ward, Z. Ding, R. A. Kennedy (1998). Broadband DOA Estimation Using Frequency Invariant Beamforming. IEEE Transactions On Signal Processing, Vol. 46, No. 5, May 1998.

E. Del Pezzo, F. Bianco, G. Saccorotti. (2004). Seismic source dynamics at Vesuvius volcano, Italy. *Journal of volcanology and geothermal research*, *133*(1), 23-39

A. Douglas, Bowers D., Marshall P.D., Young J. B., Porter D., Wallis N. J.(1999). Putting nuclear-test monitoring to the test. Nature, 398,1999, 474 – 475.

K. De Meersman, M. van der Baan, J.-M. Kendall (2006). Signal Extraction and Automated Polarization Analysis of Multicomponent Array Data. Bulletin of the Seismological Society of America. Vol. 96. No.6, 2415-2430, December 2006.

W.A. Elders, G.Ó.Friðleifsson, A.Albertssonb (2014). Drilling into magma and the implications of the Iceland Deep Drilling Project (IDDP) for high-temperature geothermal systems worldwide.Geothermics 49 (2014) 111–118

G. Ó. Friðleifsson, Ármannsson, H., Guðmundsson, Á., Árnason, K., Mortensen, A.K., Pálsson, B., Einarsson, G.M. (2014), Site selection for the well IDDP-1 at Krafla, Geothermics, 49, 9-15.

O.L. Frost (1972). An algorithm for linear constrained Adaptive Array Processing. Proc. of IEEE,60(8), 926-935, august 1972

M. Furumoto, T. Kunimoto, H. Inoue, I. Yamada, K. Yamaoka, A. Ikami, Y. Fukao (1990). Twin sources of high-frequency volcanic tremor of Izu-Oshima volcano, Japan. Geophysical research letters, vol.17, no.1, 25-27, January 1990

J. J. Galiana-Merino, J. Rosa-Herranz, P. Jauregui, S. Molina, J. Giner (2007). Wavelet Transform Methods for Azimuth Estimation in Local Three-Component Seismograms. Bulletin of the Seismological Society of America, Vol. 97, No. 3, pp. 793–803, June 2007.

P. Goldstein, R. J. Archuleta (1987). Array analysis of seismic. Geophysical research letters, vol.14, no.1, 13-16, January 1987

P. Goldstein, R. J. Archuleta (1991). Deterministic Frequency-Wavenumber Methods and Direct Measurements of Rupture Propagation During Earthquakes Using a Dense Array: Theory and Methods. Journal of geophysical research, vol. 96, no.B4, 6173-6185, April 10,1991.

J. Guilbert, J. Vergoz, E. Schissele´, A. Roueff, and Y. Cansi (2005) . Use of hydroacoustic and seismic arrays to observe rupture propagation and source extent of the Mw = 9.0 Sumatra earthquake. Geophysical research letters, vol. 32, l15310.

X. Gong, Y. Xu, Z. Liu (2008). Quaternion ESPRIT for Direction Finding with a Polarization Sentive Array. IEEE Xplore Conference: Signal Processing, 2008. ICSP 2008.

M. Hobiger, N. Le Bihan(2), C. Cornou, P.-Y. Bard (2009), Rayleigh Wave Ellipticity Estimation From Ambient Seismic Noise Using Single And Multiple Vector-Sensor Techniques. 17th European Signal Processing Conference (EUSIPCO 2009) 2037-2041.

S. Hólmgeirssona, Guðmundssona, Á., Bóassonb, H.Á., Ingasonb, K., Sverrissonb, H., Thórhallssonc, S.(2014), Drilling of the well IDDP-1, Geothermics, 49, 23-30.

J. M. Ibanez, e. Del Pezzo, J. Almendros, M. La Rocca, G. Alguacil, Ramon Ortiz, Alicia Garcia (2000). Seismovolcanic signals at Deception Island volcano, Antarctica: Wave field analysis and source modeling. Journal of geophysical research, vol. 105, no.B6, 13,905-13,931, June 10,2000.

L.A. Inza, J.I. Mars, J.P. Metaxian , G.S. O'Brien, O. Macedo (2011). Localization with multicomponent seismic array. Author manuscript published in "The Fourth International Workshop on Computational Advances in Multi-Sensor Adaptive Processing. San Juan : Puerto Rico (2011)

M. La Rocca, E. Del Pezzo, M. Simini, R. Scarpa, G. De Luca (2001) Array Analysis of Seismograms from Explosive Sources: Evidence for Surface Waves Scattered at the Main Topographical Features. Bulletin of the Seismological Society of America, 91, 2, pp. 219–231, April 2001.

R. Jeffers, K. L. Bell, Harry L. Van Trees (2002). Broadband Passive range estimation using music. IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), 0-7803-7402-9/02.

C. A. Langston (2007). Spatial Gradient Analysis for Linear Seismic Arrays. Bulletin of the Seismological Society of America, Vol. 97, No. 1B, pp. 265–280, February 2007.

H. Mack, E. Smart (1972). Frequency Domain Processing of Digital Microbarograph Array Data. Journal of Geophysical research, Vol.77, NO.3, 20 January 1972.

F. Marvasti, A. Amini, F. Haddadi, M. Soltanolkotabi, B. H. Khalaj, A. Aldroubi, S. Holm, S. Sanei, J. Chambers (2012). A Unified Approach to Sparse Signal Processing. EURASIP Journal on Advances in Signal Processing December 2012, 2012:44

S. Miron. N. Le Bihan, J. I. Mars (2005). Vector-Sensor MUSIC for Polarized Seismic Sources Localization. EURASIP Journal on Applied Signal Processing 2005:1, 74–84.

S. Miron, N. Le Bihan, J. I. Mars (2004). Polarized source characterization using vector-music. Signal Processing conference, 2004 12th European.

S. Mykkeltveit, k. Stebol, d. J. Doornbos, e. S. Husebye(1983). Seismic array configuration optimization bulletin of the seismological society of america, vol. 73, no. 1, pp. 173-186, February 1983.

J. A. Nelder, R. Mead (1965). A simplex method for function minimization, in Computer Journal, vol. 7, 1965, pp. 308–313

E. A. Okal, Y. Cansi (1998) Detection of PKJKP at intermediate periods by progressive multi-channel correlation. Earth and Planetary Science Letters 164 (1998) 23–30

C. Paulus and J. I.Mars (2010).Vector-Sensor Array Processing for Polarization Parameters and DOA Estimation.Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2010, Article ID 850265, 13 pages.

R. A. Kennedy, T. D. Abhayapala, D. B. Ward (2002). Broadband Nearfield Beamforming Using a Radial Beampattern Transformation. IEEE Transactions on signal processing, vol. 46, no. 8, January 2002.

R. A. Kennedy, D. B. Ward, T. D. Abhayapala (1999). Nearfield Beamforming Using Radial Reciprocity. IEEE Transactions on signal processing, vol. 47, no. 1, January 1999.

B. L. N. Kennett, D. J. Brown, M. Sambridge, and C. Tarlowski (2003). Signal Parameter Estimation for Sparse Arrays. Bulletin of the Seismological Society of America, Vol. 93, No. 4, pp. 1765–1772, August 2003

H. Krim, M. Viberg (1996). Two decades of array signal processing research: the parametric approach. IEEE Signal Processing Magazine Volume: 13, Issue: 4, Jul 1996

P. Kumar, E. Foufoula-Georgou (1997). Wavelet analysis for geophysical applications. Review of geophysics,35, 4 / 385-412, November 1997.

A. F. Kushnir, v. M. Lapshin, V. I. Pinsky, J. Fyen (1990). Statistically optimal event detection using small array data. Bulletin of the seismological society of America, vol. 80, no. 6, pp. 1934-1950, December 1990.

S. Rost, C. Thomas (2002). Array Seismology: Methods and applications. Reviews of geophysics, 40, 3 / September 2002.

G. Saccorotti, B. Chouet, M. Martini, R. Scarpa (1998). Bayesian Statistics Applied to the Location of the Source of Explosions at Stromboli Volcano, Italy. Bulletin of the Seismological Society of America, Vol. 88, No. 5, pp. 1099-1111, October 1998.

G.Saccarotti, V.Nisii, E. Del Pezzo (2008). Coherent-subspace array processing based on wavelet covariance: an application to broad-band, seismo-volcanic signals. Geophys. J Int. 174, 435-450

G.Saccarotti, R. Maresca, E. Del Pezzo (2001). Array analyses of seismic noise t Mt. Vesuvius Volcano, Italy. Journal of Volcanology and Geothermal Research 110 (2001) 79-100

A. Saiga, K. Yamaoka, T. Kunitomo, T. Watanabe (2006) . Continuous observation of seismic wave velocity and apparent velocity using a precise seismic array and ACROSS seismic source. Earth Planets Space, 58, 993-1005, 2006.

M. Schimmel, Paulssen, H. (1997). Noise reduction and detection of weak, coherent signals through phase-weighted stacks. Geophysical Journal International, 130(2), 497-505.

E. Schisselé, J. Guilbert, S. Gaffet, Y. Cansi (2004). Accurate-Time-frequency-wavenumber nalysis to study coda waves. Geophys. J. Int.158, 577-591.

R. O. Schmidt (1986). Multiple Emitter Location and Signal Parameter Estimation. IEEE transactions on antennas and propagation. Vol. AP-34, No.3, March 1986.

J. Schweitzer, J. Fyen, S. Mykkeltveit, S. J. Gibbons, M. Pirli, D. Kühn, T. Kværna (2012). Seismic Arrays. - In: Bormann, P. (Ed.), *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, Potsdam : Deutsches GeoForschungsZentrum GFZ, pp. 1—80.

H. Steyskal(1987). Digital beamforming antennas - An introduction. Microwave Journal (ISSN 0026-2897), vol. 30, Jan. 1987, p. 107, 108, 110 (7 ff.).

B. Stump, M.-S. Jun, C. Hayward, J.-S. Jeon, Il-Y. Che, K. Thomason, S. M. House, J. McKenna (2004). Small-Aperture Seismo-Acoustic Arrays: Design, Implementation, and Utilization. Bulletin of the Seismological Society of America, Vol. 94, No. 1, pp. 220–236, February 2004.

C. Torrence, G. P. Compo (1998). A Practical Guide to the Wavelet Analysis. Bulletin of the American Meteorological Society: Vol.79, No.1, 61-78

A. Tramelli, E. Del Pezzo, M. C. Fehler (2009) 3D Scattering Image of Mt. Vesuvius. Bulletin of the Seismological Society of America June 2009 vol. 99 no. 3 1962-1972

B.D. Van Veen, K.M. Buckley (1988). Beamforming: A versatile Approach to Spatial Filtering. IEEE ASSP Magazine, pp 4-24, April 1988

M. Weber, C.W.Wicks Jr. (1996). Reflections from a distant subduction zone, Geophysical research letters 23 1996, pp.1453-1456.

M. H. Wirth, r. R. Blandford, and R. H. Shumway (1976). Automatic seismic array and network detection bulletin of the seismological society of america. Vol. 66, no. 4, pp. 1375-13811. August 1976.

K.. Yamaoka, T. Kunitomo, K. Miyakawa, K. Kobayashi, M. Kumazawa (2001). A trial of monitoring temporal variation of seismic velocity using an ACROSS system. The Island Arc (2001) 10, 336-347.