UNIVERSITÀ DI PISA

Department of Earth Science

PhD in Earth Sciences

XXVIII Cycle

Dynamic characterization of soils by means seismic passive measures and their utility in Seismic Microzonation studies

Author:

Enrico Paolucci

Supervisor:

Prof. Dario Albarello

Acknowledgements

First of all I want to thank Prof. Dario Albarello, my thesis supervisor, for his guidance and patient support throughout these years.
I would like to express my sincere gratitude to Dr. Pauline Galea, Dr. Sebastiano D'Amico, Daniela Farrugia and Dr. Stefano Parolai for their collaboration, suggestions and their warm hospitality at the University of Malta and GeoForschungsZentrum (GFZ) of Potsdam (Germany).
Many thanks to my colleagues and friends Enrico Lunedei, Domenico Pileggi and Giacomo Peruzzi for the moral and scientific support.
My special thanks to my friends and all the persons who were close to me during these years of study.
I dedicate this thesis to my mother, my uncle and in particular to my father.

ABSTRACT

Ambient vibrations contain information on the local S-wave velocity structure, which can be obtained from the Rayleigh wave phase velocity dispersion curve by means of seismic array measurements. The horizontal-to-vertical (HVSR or H/V) spectral ratio from single stations provides a direct estimate of the soil resonance frequency and also contains information on the average S-wave velocity and the total thickness of the sedimentary cover. Both curves can be used in a joint inversion procedure in order to obtain a more detailed V_S profile of the subsoil. The use of these passive seismic techniques for dynamic characterization of soils is significantly growing in the last few years and, in particular, their application in the framework of Seismic Microzonation studies is becoming of paramount importance.

This PhD thesis is characterized by an applicative nature. Its main goals were:

- to verify and evaluate the reliability of the passive seismic techniques in presence of a significant S-wave velocity inversion in the subsoil;
- to check the informative capability of the extensive use of the ambient vibration prospecting, also identifying the lateral geological heterogeneities at the scale ranging from hundreds of meters to tens of kilometers.

The objectives and the research lines dealt in this dissertation are devoted to improve the informative potential of the passive seismic measurements for purposes related to Seismic Microzonation studies.

The first three Chapters of the thesis were devoted to review the state of the art concerning the origin and the nature of the ambient vibrations and the different analysis procedures adopted for the single station and seismic array measurements. Moreover, in the Chapter 4 a review of the methods aimed to estimate the V_S profile was given, both concerning the simplified approaches and the most common inversion procedures.

In order to accomplish the first purpose of the dissertation, a study is carried out in the Maltese Archipelago, where the stratigraphy of the investigated sites is characterized by the presence of a thick buried low S-wave velocity layer. This geological setting is clearly visible in outcropping and easily deduced by using the geological map. This study is showed in Chapter 5 and it demonstrated the optimal capability of the passive seismic techniques and the joint inversion procedure based on Genetic Algorithms in resolving both the presence and the characteristics of the thick buried low-velocity layer in the stratigraphy. Moreover, it has been shown that the effective Rayleigh wave dispersion curves obtained using the ESAC method represents an excellent marker to identify this particular subsoil characteristic.

As concerns the second objective of this thesis, the results of an extensive use of the ambient vibration prospecting carried out in the area damaged by May-June 2012 seismic sequence in Emilia Romagna (Northern Italy) are summarized and analyzed in Chapter 6. The visual inspection of the HVSR curves allowed to identify three zones where analogous seismic response is expected. Despite the fact these general trends are the effect of the subsoil configuration, H/V spectral ratios show a significant dependence on meteo-climatic conditions: in particular, HVSR peak

amplitudes in the low frequency range (< 0.5 Hz) correlate significantly with the sea wave activity in the Central Mediterranean. Anyway, resonance frequencies estimated from HVSR peaks appear persistent and have been used to estimate the local depth of impedance contrasts responsible for seismic resonance phenomena. To this purpose, average V_S values were assessed using a simplified procedure constrained by Rayleigh dispersion curves deduced from the seismic arrays. In this way, two main interfaces have been detected: at depths of about 50–100 and 500–600 m.
Finally, in order to perform an automatic identification of the large scale areas detected by similar HVSR curves, a procedure based on Principal Component Analysis (PCA) was proposed in Chapter 7. After defining the grouping criteria, It was observed that this technique is able to group together similar HVSR curves, obtaining in this way a first identification of the large scale heterogeneities of the study area. Another significant characteristic of PCA is the immediate recognition of the main patterns representing the overall variance of the original dataset. Moreover, the analyzed study cases demonstrate that this technique and the adopted grouping procedure discriminate especially well the large scale heterogeneities in contexts where significant geological changes occur.

RIASSUNTO

Le vibrazioni ambientali contengono informazioni sulla struttura di velocità delle onde di taglio nel sottosuolo che può essere stimata dalla curva di dispersione delle velocità di fase delle onde di Rayleigh ottenuta mediante le misure di array sismico. La curva dei rapporti spettrali H/V (o HVSR) ottenuta dalle misure a stazione singola fornisce una stima diretta della frequenza di risonanza del sito e contiene informazioni sulla velocità media delle onde S e sullo spessore totale della copertura sedimentaria. Entrambe le curve possono essere utilizzate in una procedura di inversione congiunta con lo scopo di ottenere un profilo di V_S del sottosuolo più dettagliato. L'uso di queste tecniche di sismica passiva finalizzate alla caratterizzazione dinamica dei terreni sta crescendo significativamente negli ultimi anni e, in particolare, la loro applicazione nel contesto degli studi di Microzonazione Sismica sta diventando di fondamentale importanza.

Questa tesi di dottorato è caratterizzata da una natura applicativa. I suoi scopi principali sono stati:

- verificare e valutare l'affidabilità delle tecniche di sismica passiva in presenza di una marcata inversione di velocità delle onde S nel sottosuolo;
- verificare la capacità informativa dell'uso estensivo di tali metodi, identificando anche le eterogeneità geologiche laterali alla scala che va dalle centinaia di metri alle decine di chilometri.

Gli obiettivi e le linee di ricerca trattati in questa tesi sono finalizzati a migliorare il potenziale informativo delle misure di sismica passiva per scopi connessi agli studi di Microzonazione Sismica.

I primi tre Capitoli sono dedicati alla rassegna dello stato dell'arte riguardante l'origine e la natura delle vibrazioni ambientali e le differenti procedure di analisi adottate per le misure a stazione singola e per quelle di array sismico. Inoltre, nel Capitolo 4 è stata fornita una rassegna dei metodi finalizzati alla stima del profilo di V_S, comprendente sia gli approcci semplificati, sia le più comuni procedure di inversione.

Per conseguire il primo scopo del lavoro, è stato realizzato uno studio nell'Arcipelago Maltese, dove la stratigrafia dei siti indagati è caratterizzata dalla presenza di uno spesso strato sepolto a bassa velocità delle onde S. Questa configurazione geologica è chiaramente visibile in affioramento e facilmente deducibile consultando la carta geologica. Questo studio ha dimostrato l'ottima capacità delle tecniche di sismica passiva e della procedura di inversione congiunta basata sugli Algoritmi Genetici nel determinare sia la presenza che le caratteristiche dello strato a bassa velocità nel sottosuolo. Inoltre, è stato mostrato che la curva di dispersione effettiva delle velocità delle onde di Rayleigh ottenuta attraverso il metodo ESAC rappresenta un eccellente indicatore per identificare questa particolare configurazione di sottosuolo.

Per quanto riguarda il secondo obiettivo della tesi, nel Capitolo 6 sono stati mostrati e analizzati i risultati di un utilizzo estensivo delle prospezioni di simica passiva effettuate nell'area danneggiata dalla sequenza sismica di Maggio-Giugno 2012 in Emilia Romagna (Italia Settentrionale). L'ispezione visuale delle curve HVSR ha permesso di identificare tre zone dove è attesa la stessa risposta sismica. Nonostante gli andamenti generali delle curve siano effetto della configurazione del sottosuolo, i rapporti spettrali H/V mostrano una significativa dipendenza alle condizioni

meteo-climatiche: in particolare, le ampiezze dei picchi in bassa frequenza (< 0.5 Hz) sono significativamente correlate con l'attività delle onde del mare nel Mediterraneo Centrale. Tuttavia, le frequenze di risonanza stimate dai picchi HVSR appaiono persistenti e sono state usate per stimare la profondità dei contrasti di impedenza responsabili dei fenomeni di risonanza sismica individuati. A questo scopo, sono stati stimati i valori di V_S media usando una procedura semplificata vincolata dalle curve di dispersione delle onde di Rayleigh ricavate dagli array sismici. In questo modo, sono state individuate due interfacce principali, una alla profondità di 50-100 m e l'altra a 500-600 m.

Infine, con lo scopo di effettuare un'identificazione automatica delle aree a larga scala individuate da curve HVSR simili, nel Capitolo 7 è stata proposta una procedura basata sull'Analisi delle Componenti Principali (PCA). Dopo aver definito i criteri di raggruppamento, è stato osservato che questa tecnica è in grado di raggruppare insieme simili curve H/V, ottenendo in questo modo una prima identificazione delle eterogeneità a larga scala dell'area di studio. Un'altra caratteristica significativa della PCA è l'immediato riconoscimento dei pattern principali che rappresentano la varianza complessiva del dataset originale. Inoltre, i casi studio analizzati dimostrano che questa tecnica e le procedure di raggruppamento adottate discriminano particolarmente bene le eterogeneità a larga scala in contesti caratterizzati da variazioni geologiche significative.

CONTENTS

INTRODUCTION

After past and recent earthquakes, it was possible to observe how the spatial distribution of damage maximum effects was very heterogeneous on equal vulnerability conditions of the buildings. Since 1970s, several studies conducted in areas stricken by strong earthquakes (Seed and Idriss, 1969; Stone *et al.*, 1987; Seed *et al.*, 1990) witnessed that the local geological-geomorphological conditions may contribute to generate significant local amplification effects of the ground shaking.

When an earthquake occurs, the ground motion detected at two nearby sites characterized by different outcropping materials can show very different features (Fig. I), concerning the amplitude values, the overall duration and the frequency content. This latter characteristic might play an important role considering that the buildings can react to the input ground motion emphasizing shaking depending on their frequency of oscillation (Parolai, 2012).

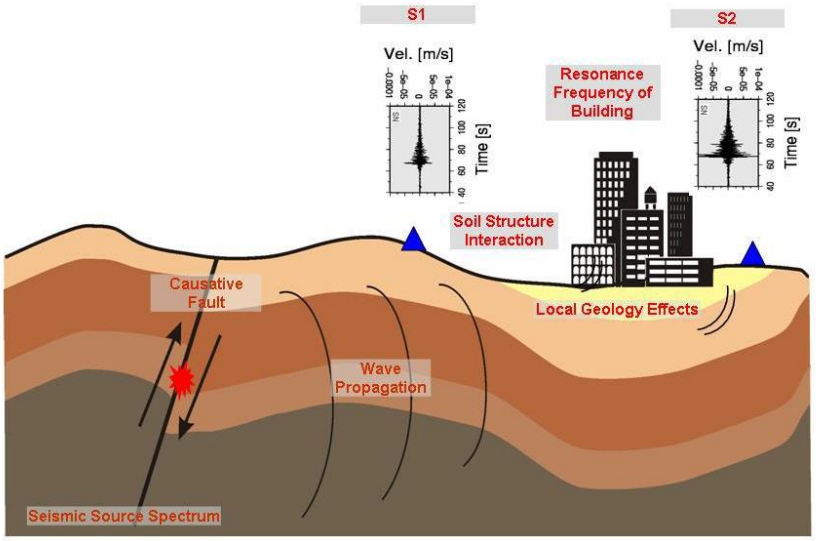

Fig. I - Simplified scheme showing seismic wave propagation after a seismic event. Blue triangles show the location of the two seismic stations: S1 is located on rock site and S2 on more recent soft sediments (Parolai, 2012).

In general, the main factor that can concur to the modification of the ground motion within a sedimentary basin is the impedance contrast between the soft sedimentary cover and the bedrock: the first one is characterized by low density ρ_{Sed} and S-wave velocity V_{Sed}, the second one by high density ρ_{Bed} and S-wave velocity V_{Bed}. The impedance contrast c, defined by the relation

$$c = \frac{\rho_{Bed}\ V_{Bed}}{\rho_{Sed}\ V_{Sed}} \qquad (1)$$

establishes the relevance of the resonance phenomenon, derived by the trapping of the seismic wave energy within the sedimentary cover due to multi-reflection of the waves at particular frequencies. Considering an 1-D configuration and the simplified assumption of vertical

propagation of S waves, starting from the results of Takahashi and Hirano (1941), the modulus of the site amplification function *H(f)* can be described by the following relation (Parolai, 2012):

$$|H(f)| = \left(\frac{(1+r)^2}{1+2rcos(4\pi f\tau)+r^2}\right)^{\frac{1}{2}} \quad (2)$$

where τ is the travel time of the S waves in the soft sedimentary layer and *r* is the reflection coefficient that is related to the impedance contrast *c*. The maximum value of *H(f)* is obtained when cos($4\pi f\tau$) is equal to -1 that is when $f = 1/(4\tau)$. This value of *f* is defined as the fundamental resonance frequency f_0. From these considerations, it derives the well know equation (Kramer, 1996; Parolai, 2012):

$$f_0 \approx \frac{V_{Sed}}{4H_{Sed}} \quad (3)$$

that relates the f_0 value with the average shear wave velocity (V_{Sed}) and the thickness (H_{sed}) of the soft sedimentary layer.

Other significant phenomena can concur to the ground motion amplification. These effects, related with 2-D or 3-D geological-geomorphological variations, mainly concern focusing, diffraction and scattering of body waves by irregular surface geometry, both within the sedimentary basins and in sites located on pronounced topographic relief (Aki, 1988; Safak, 2001).

Moreover, the occurrence of a seismic event can induce phenomena of ground instability that can produce significant permanent deformations (e.g., seismic induced landslide, liquefaction and near surface faulting; Kramer, 1996).

All these effects, more commonly defined like “site effects”, play an important role in the seismic risk assessment. Furthermore, an analysis aiming to the identification and the evaluation of site effects at the investigated area scale assumes a paramount importance for the reconstruction of damaged structures in areas stricken by an earthquake and for the planning of a possible urban development. A study of this kind is called Seismic Microzonation. Basically, this study is focused to assess the local seismic hazard identifying portions of the study area characterized by homogenous seismic behavior. In particular three different situations are identified (Albarello *et al.*, 2015):

- *the stable areas*, i.e. where no amplification phenomena are expected respect to the reference ground motion at a flat rock site;
- *the stable areas susceptible to local amplification*, i.e. where amplification phenomena induced by the local stratigraphic and geomorphological configuration are expected respect to the reference ground motion at a flat rock site;
- *the areas susceptible to instability*, i.e. where permanent deformation phenomena induced by earthquake occurrence (slope instability, liquefaction or fault rupture) are expected.

Seismic Microzonation studies provide this kind of information at the municipality scale (about hundreds of squared kilometers) on the basis of a "holistic" approach (Mucciarelli, 2008), that is using a interdisciplinary criterion taking in account jointly geology, geophysics, seismology geotechnics and structural engineering.
Specific guidelines were released in 2008 ("Indirizzi e Criteri per la Microzonazione Sismica" or ICMS; italian version: GdL MS, 2008; english version: SM Working Group, 2015), tested in field after the earthquake occurred in L'Aquila (Central Italy) in 2009 (GdL MS-AQ, 2010) and upgraded in 2011 (AA.VV., 2011). Nowadays, these studies constitute in Italy a basic element of civil protection strategy and the drafting of guidelines has been necessary to harmonize the Seismic Microzonation projects distributed over the whole Italian territory and performed by different local technical bodies and practitioners with the support of research and academic institutions (Albarello *et al.*, 2015).
The realization of this kind of analysis has different costs and complexities as a function of the level of detail. In order to make them achievable for municipalities having different economical resources, the ICMS are developed into three levels:

- *level 1* is a semi-qualitative level aiming to reconstruct the reference geologic model identifying the areas where an homogenous seismic behavior is expected. In this first phase the collection of existing data and the geological survey play an important role;
- *level 2* is a more quantitative level aiming at quantifying amplification phenomena by using simplified approaches based on 1-D scheme;
- *level 3* includes detailed measurements with advanced techniques and concerns more complex situations (2-D, 3-D effects, liquefaction, etc..) where simplified approaches are inappropriate.

In view of what has been described previously about the nature of site effects, it is clear that Seismic Microzonation studies, the seismic design and a more general seismic response analysis are based on the characterization of the dynamic behavior of soils under the influence of a cyclic disturbance such as that induced by the occurrence of an earthquake. The relevant parameters that effectively represent the physical-mechanical properties of the soil in such conditions are the shear-wave velocity (V_S) profile and strain dependent dynamic factors like the shear modulus and the damping factor. Particularly, the first parameter is essential to investigate the shear wave resonance effects within a sedimentary basin.
These information can be obtained both from laboratory measurements conducted on samples taken by perforations and from "in situ" measurements with appropriate techniques of seismic prospecting. Although the first ones present the advantage of operating in large deformations conditions (such as those that occur during an important seismic events), they are affected by the problem of the effective representation of the data obtained from a small size sample and often in disrupted conditions due the drilling. Seismic measurements can overcome these problems and provide important information about the soils dynamic properties though limited to the small deformations field.

Several seismic prospecting techniques exist to retrieve the local V_S profile. There are two main families: the invasive methods, essentially based on borehole seismic techniques (down-hole, cross-hole, up-hole) and non invasive methods, based on surface seismic techniques. This latter family is divided in:

- V_S refraction and reflection seismic prospections;
- methods based on surface-wave analysis. These ones can be carried out by using:
  - active techniques;
  - passive techniques by means ambient vibrations analysis.

The borehole methods provide a V_S profile estimate more reliable than the other ones but in addition to the invasive nature, they are characterized by relatively higher costs than surface seismic techniques.

The surface-wave methods, used also at large scale to investigate the Earth structure, exploits the geometric dispersion of surface waves to infer the properties of the medium: in particular surface waves having different frequency (f) propagate interesting portion of subsoil up to a depth proportional to their wavelength (λ). The propagation velocity (V) is strictly dependent on the physical properties of the layers involved. Since f and λ are related according to the equation $\lambda = V/f$, the surface waves at high frequency are characterized by lower wavelengths and they propagate through the shallower layers of the subsoil, conversely low frequency waves affect the deeper layers. These information allow to reconstruct the surface wave dispersion curve, that is the function that relates the surface wave velocity values with the frequency: this curve is strictly correlated with the subsoil configuration.

Despite the cheapness, these techniques are characterized by greater interpretation ambiguity compared to borehole seismic tests. In fact, the estimate of V_S profile by using the surface-wave applications is accomplished by adopting a strategy based on estimating the experimental dispersion curve from field data and subsequently solving an inverse problem (Socco *et al.*, 2010). Because this latter procedure is mathematically illposed, the non uniqueness of the solution plays an important uncertainty role.

In particular, the standard procedure for surface-wave analysis can be divided into three main steps (Socco *et al.*, 2010; Foti *et al.*, 2011; Fig. II):

1. acquire the experimental data;
2. process the signal to obtain the experimental dispersion curve;
3. solve the inverse problem to estimate the V_S profile at the investigated site.

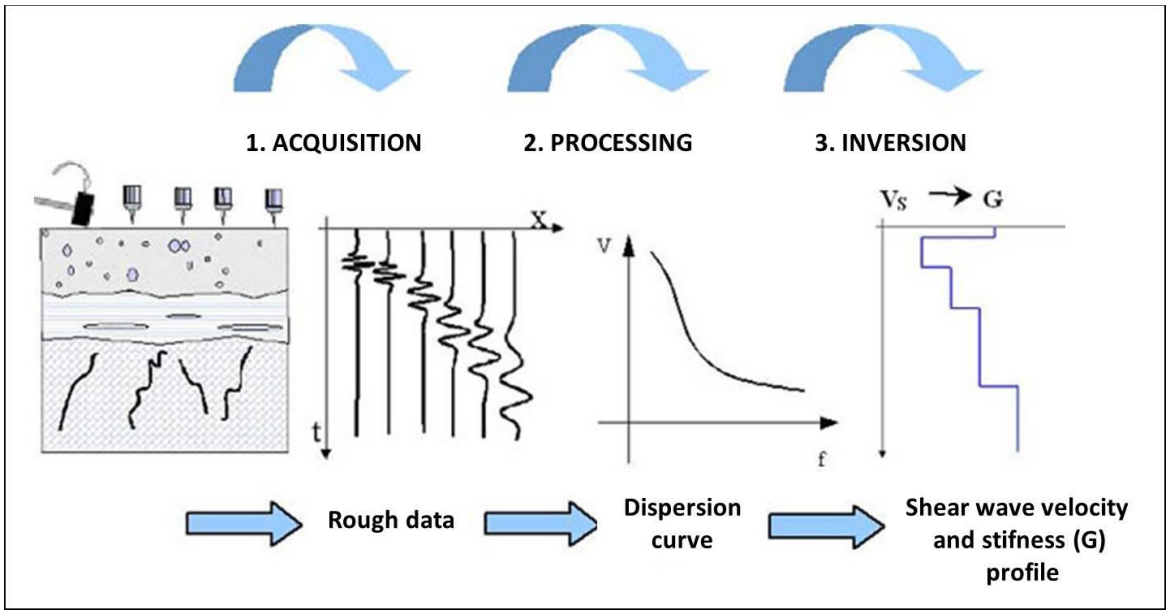

Fig. II - Standard procedure scheme adopted to infer the V_S profile by using surface-wave techniques.

An important limitation that characterizes the active techniques (both surface-wave and body-wave based) is their sensitivity to an eventual presence of high levels of environmental noise, such as those ones that occur close to industrial plants or in urban areas. In such situations the application of these methods requires intense energization sources, which may be incompatible with the application context. Moreover, the active surface seismic methods are characterized by limited investigation depth, essentially due to their high frequency energization.

In the last decade, the interest on surface-wave passive techniques is considerably increased. These methods, commonly called passive seismic prospections, are essentially founded on the study of ambient vibrations produced by natural and anthropic sources randomly distributed around the measuring points. These measurements can be performed in multi-station configurations (seismic arrays) or using a single station acquisition.
In the first configuration, the ambient vibration field is recorded through a group of synchronized sensors deployed in the investigated area with appropriate geometries (essentially 2D linear or random geometries). Usually, only the vertical component of the ground motion is acquired: in this case, the goal of this technique is the retrieving of Rayleigh wave velocity (V_R) dispersion curve, which can be obtained by using suitable different analyses. In particular, two main approaches are used: one is based on spatial autocorrelation analysis (or SPAC; e.g., Aki, 1957; Okada, 2003), the other one on frequency-wavenumber analysis (or *f-k*; e.g., Lacoss et al., 1969; Capon, 1969).
In the single station measurements the ambient vibration are recorded with a three-directional sensor and analyzed with HVSR (or H/V) technique (e.g., Nakamura, 1989). The aim of this method is the direct estimate of the fundamental resonance frequency of the soil from the spectral ratio between the horizontal (H) and the vertical (V) components of the ground motion. Despite the

debates (e.g., Fäh *et al.*, 2001; Lunedei and Albarello, 2010) that afflict the physical interpretation of the H/V ratios in function of the frequency (HVSR curve), the frequency corresponding to the maximum value of the H/V function has a strict correspondence with the fundamental resonance frequency of the sedimentary cover (e.g., Bonnefoy-Claudet *et al.*, 2006a).
The main limitations of the seismic passive techniques are linked to the presence of uncontrolled sources that make more complex the interpretation and the analysis of the data compared to the active techniques. In particular, it is possible to obtain reliable results for geognostic aims only exploiting the statistical properties of the ambient vibration wavefield. This feature clearly differentiates these techniques from the active ones which are strongly characterized by a deterministic theoretical approach.
On the other hand, the exploitation of the ambient vibration wavefield provides different benefits. In fact, seismic passive methods have energization sources virtually limitless and costless; moreover they are particularly suitable in contexts where the active methods can fail due to low signal-noise ratio (e.g., in urban areas). Another important benefit is related to the spectral structure of the ambient vibrations: in fact, thanks to their higher low-frequency content it is possible to retrieve information about deeper structures as compared to the active prospecting. Furthermore, the possible use of an inversion procedure based on the joint provision of the surface-wave dispersion curve and the HVSR curve (e.g., Arai and Tokimatsu, 2005; Parolai *et al.*, 2005; Picozzi and Albarello, 2007; Albarello *at al.*, 2011) allows to obtain reliable V_S profile estimates up to hundreds of meters depths.

All these characteristics make the seismic passive measurements an excellent tool for the above described Seismic Microzonation studies. In fact, the main targets of such projects are the urban areas, where the active methods suffer of important drawbacks. Moreover, taking into account the frequency range of engineering interest (0.5-5 Hz) and the shear-wave velocity values typical of shallow sedimentary bodies (of the order of hundreds of meters per second), the resulting wavelength of interest for Seismic Microzonation is of the order of tens to hundreds of meters (Albarello *et al.*, 2015): in view of this, considering the higher penetration depth with respect to the active measurements, the ambient vibration prospecting is definitely more suitable to individuate the major impedance contrasts depth.
Actually, the use of passive seismic measurements is common in this framework, in fact it provides an important support to the development of the three investigation level described previously (Albarello *et al.*, 2015). In particular:

- in the first level the single station acquisitions analyzed with HVSR technique represent an important tool to individuate the effective presence of resonance phenomena and then to locate the areas where these effects occur. Moreover, they are particularly useful to support the geological survey aiming at identifying lithological transitions where major seismic impedance contrasts occur in the subsoil (Albarello *et al.*, 2011; Gallipoli et al., 2011). Furthermore, exploiting the direct estimate of the resonant frequency value, it is possible to roughly assess the resonant interface depth in terms of “shallow” (few meters),

"deep" (tens of meters) and "very deep" (hundreds of meters) (Albarello *et al.*, 2015). Finally, the quickness and the ease of use make HVSR an optimal analysis tool in this first Seismic Microzonation phase, where the needed information must have an extensive nature;

- in the second level, a first order quantification of the amplification effects by using tabulated parameters is required. These tables (called "abacuses"; Peruzzi et al., 2016) are based on the use of integral parameters like the average S-wave velocity in the first 30 meters of subsoil (V_{S30}) and resonance period of S waves. In this context, seismic array techniques are a very useful tool to estimate V_{S30} value exploiting simplified interpretative procedures which don't need detailed interpretations based on inversion procedures of experimental dispersion curves (D'Amico *et al.*, 2008; Albarello and Gargani, 2010; Albarello *et al.*, 2011);
- In the third level, it is necessary to obtain very detailed quantitative information about the dynamic characteristics of the subsoil in presence of more complex geological settings. In this context, where adequate inversion approaches are expected to be used, joint inversion techniques that consider both H/V and Rayleigh waves dispersion curves are very useful to constrain the local V_S profile up to depths of the order of hundreds of meters (Picozzi and Albarello, 2007; Albarello *et al.*, 2011). Moreover, beyond V_S profile estimate, adequate passive seismic array analysis can provide an assessment of the damping factor at the small strain level (Albarello and Baliva, 2009; Parolai, 2014).

This dissertation concerns the dynamic characterization of soils by using passive seismic measurements, both through single station and seismic array configurations. The aim of the work is twofold:

- ✓ to verify and evaluate the reliability of this kind of techniques in the contexts characterized by complex geological settings, that is situations not related with simple 1-D scheme where the S-wave values increase with depth;
- ✓ to check the informative capability of the extensive use of the ambient vibration prospecting.

In particular the research activities are developed in three main directions:

1. exploration of the effects due to a significant S-wave velocity inversion in the V_S profile on the HVSR curves and Rayleigh wave phase velocity dispersion curves;
2. deduction of the dynamic characteristics of the subsoil from passive seismic measurements at the scale of about thousands of square kilometers;
3. detection of lateral geological heterogeneities at scale of tens of kilometers by means the information inferred from extensive HVSR surveys.

As concerns the first research line, in recent years some studies, both theoretical and experimental, were carried out on HVSR curve (e.g., Di Giacomo *et al.*, 2005; Castellaro and Mulargia, 2009) and on Rayleigh wave dispersion curve (Tokimatsu *et al.*, 1992a; Tokimatsu, 1997; Zhang and Lu, 2003). The study carried out in this dissertation has an experimental nature and it is

focalized in both analyses pointing out the characteristics induced in the two curves by the presence of a thick buried low S-wave velocity layer. Moreover, the reliability of the joint inversion procedures in the estimating of V_S profile in this particular context was explored. The investigated sites were chosen in order to have a reliable comparison with the achieved data and the real geological configuration: in fact these sites, located in the Malta's Archipelago, are characterized by a geological setting clearly visible in outcropping and easily deduced by using the geological map. The location of the low velocity layer, characterized by an overall thickness in the order of tens of meters, is always included in the first hundred meters depth: this characteristic makes this study suitable for the Seismic Microzonation purposes.

Regarding the second aspect of this dissertation, an extensive application of the passive seismic technique was attempted in the area damaged by May-June 2012 seismic sequence in Emilia Romagna region (Northern Italy) (Paolucci *et al.*, 2015). The amount of data collected by different research institutes, universities and geological regional services allowed the experimentation of suitable simplified procedures aiming to retrieve the subsoil dynamic characteristics and to estimate the depths of the major impedance contrasts in the whole area (about 2500 Km^2).

As concerns the third research line, in order to identify the large scale lateral heterogeneities in this study area, the single station measurements (more spatially spread and numerous than the multiple station ones) and their respective HVSR curves were taken into account: in particular, H/V curves with similar trend were grouped in order to attempt to individuate the areas with analogous geological configurations. To do this, two procedures were used: the first one is based on a simple visual inspection of the curves, the second one is based on the statistic technique of the Principal Component Analysis (PCA). This latter procedure was investigated with the aim to automatically group the HVSR curves after the retrieving their predominant trends. In order to explore the reliability, the effectiveness and the drawbacks of the PCA application for these purposes, this technique was attempted in different geological contexts where several single station measurements were collected.

All the research lines dealt in this dissertation are devoted to improve the informative potential of the passive seismic measurements for purposes related to Seismic Microzonation studies. In particular, the first research activity concerns mainly the third level of investigation, where detailed and accurate measurements are required. Instead, the second and the third activity concern mainly the use of the huge amount of data collected in the framework of the first level of investigation, where information with extensive character are required.

This thesis is structured in seven chapters:

- Chapter 1 is devoted to the ambient vibrations, highlighting their nature and the main characteristics;
- in Chapter 2 the HVSR method is treated, describing its physical basis, the experimental and the computational procedures to obtain the H/V curve and its informative capability; moreover, the effect of S-wave velocity inversion on HVSR is pointed out, by reference to the main works present in the literature;

- the Chapter 3 deals the principal aspects of the seismic array analysis, describing the most common techniques directed to obtain the Rayleigh wave dispersion curve and the techniques able to retrieve the V_R value between a pair of sensors; moreover, the effect of S-wave velocity inversion is reviewed also for the dispersion curve;
- Chapter 4 illustrates the procedures aimed to estimate the V_S profile, both using simplified approach and inversion procedures;
- Chapter 5 describes the experimental study carried out in the Malta's Archipelago to assess the influence of a sharp V_S inversion in the subsoil on the passive seismic measurements;
- in Chapter 6, the extensive use of the ambient vibration monitoring carried out in Emilia Romagna is illustrated, describing the procedures used to individuate the large scale heterogeneities and to estimate the depths of the resonant interfaces;
- finally, Chapter 7 is devoted to the Principal Component Analysis and its application in the automatic identification of the large scale heterogeneities. In particular, three experimental cases are described: Emilia Romagna (Italy), Collesalvetti (Livorno, Italy) and Montecatini Terme (Pistoia, Italy).

CHAPTER 1
The ambient vibrations

1.1 Introduction

In absence of seismic events, the Earth's surface is constantly affected by small vibrations characterized by amplitudes on the order of 10^{-4} - 10^{-2} mm, below human sensing. This kind of ground shakings is called *seismic noise*, *microtremors* or *ambient vibrations*.

In seismology this background tremor is commonly referred to as seismic noise because, despite its weak intensity, it can significantly disturb the earthquake recordings. Seismologists are used to define as *microseisms* the natural vibrations with frequencies lower than 1 Hz; instead, the term *microtremors* (most used in earthquake engineering) is referred to the signal components with frequencies greater than about 1 Hz. Recently, the term *ambient vibrations* has been adopted in several disciplines (e.g., geology, applied geophysics, oil exploration, earthquake engineering) where this phenomenon can be exploited as a useful signal. Moreover, this latter expression includes both natural and anthropic small vibrations, without distinctions about their spectral content.

Since the second part of the 19th century, the ambient vibrations were the subject of several studies (Ferrari *et al.*, 2000). These early researches aspired mainly to explain the geographic distribution and the origin of the ambient vibrations sources; due to the limitations about the instruments and the processing techniques, basically these studies had a qualitative character.

Starting from 1950s, the technical improvements in equipment and the increase of the knowledge in seismology field allowed significant advances in the understanding of the ambient vibrations nature and of their wavefield. In this period, some authors showed the effectiveness of the ambient vibrations in achieving information about the subsoil (e.g., Kanai, 1954; Aki, 1957; Akamatsu, 1961). These pioneering works helped the development of geophysical prospecting techniques essentially based on dispersive character of surface wave contained in seismic noise wavefield.

From 1970s up to now the number of publications dealing this argument has considerably increased: most of these works treat the use of ambient vibration analyses aimed to the dynamic characterization of the subsoil and, starting from 1990s, their application in the Seismic Microzonation studies.

1.2 Origin of ambient vibrations

From several research works (e.g., Gutenberg, 1958; Asten, 1978; Bonnefoy-Claudet, 2006b) it is well known that the ambient vibrations are generated by natural sources (e.g., ocean currents, tides, waves striking the coasts, wind, atmospheric pressure variations, water flow in rivers) and

anthropic sources (power plants, factories, cars, human footsteps, trains). They contribute in different spectral domains (Tab. 1.1):

- at frequencies below 0.5 Hz, the sources are essentially natural. Some authors termed this phenomena *microseisms*;
- in the frequency range between 0.5 and 1 Hz, the sources are both natural and anthropic;
- at frequencies higher than 1 Hz, the sources are essentially anthropic. In this case the ambient vibrations are also referred like *microtremors* or *cultural noise*.

Sources	Gutenberg (1958)	Asten (1978, 1984)
Oceanic waves striking along the coasts	0.05–0.1 Hz	0.5–1.2 Hz
Monsoon/Large scale meteorological Perturbations	0.1–0.25 Hz	0.16–0.5 Hz
Cyclones over the oceans	0.3–1 Hz	0.5–3 Hz
Local scale meteorological conditions	1.4–5 Hz	
Volcanic tremor	2–10 Hz	
Urban	1–100 Hz	1.4–30 Hz

Tab. 1.1 - Summary of ambient vibrations sources according to frequency after the studies of Gutenberg (1958), Asten (1978), Asten and Henstridge (1984) (from Bennefoy-Claudet, 2006b).

It is important to note that the frequency of 1 Hz it is not an absolute border between natural and anthropic origin, in fact the seismic noise wavefield can be affected by local meteorological conditions up to 5 Hz. Moreover, some authors (Young *et al.*, 1996) showed that the wind-generated noise can be broadband, ranging from about 0.5 Hz up to about 15 to 60 Hz: this influence in high frequency band is probably related with the direct action of the wind on trees, bushes, buildings and other structures.

Fig. 1.1 is a summary of noise levels observed in a worldwide network of 75 permanent seismic stations at a time of no seismic event related to earthquakes of nuclear explosions (Peterson, 1993). The graph shows the upper (New High Noise Model, NHNM in Fig.1.1) and lower (New Low Noise Model, NLNM in Fig.1.1) envelopes (solid lines) of a large collection of curves representing the power spectral density of ground acceleration as a function of the period. These power spectra envelopes, calculated for the vertical component of background noise, demonstrate that the spectral characteristics of ambient vibration wavefield are largely similar worldwide.
The two clear peaks at around 5 s and 11 s are originated to the ocean activity. In particular, the smaller one is related with the primary ocean microseisms, which are generated only in shallow

waters in coastal regions, where the wave energy can be converted directly into seismic energy either through vertical pressure variations or by the smashing surf on the shores (Bormann and Wielandt, 2013). As concerns the main peak at around 5 s, it is associated with the secondary ocean microseisms activity: this kind of phenomenon is generated by the superposition of ocean waves propagating towards the continent and the waves of equal period travelling in the opposite direction due the reflection from the coast (Bormann and Wielandt, 2013).

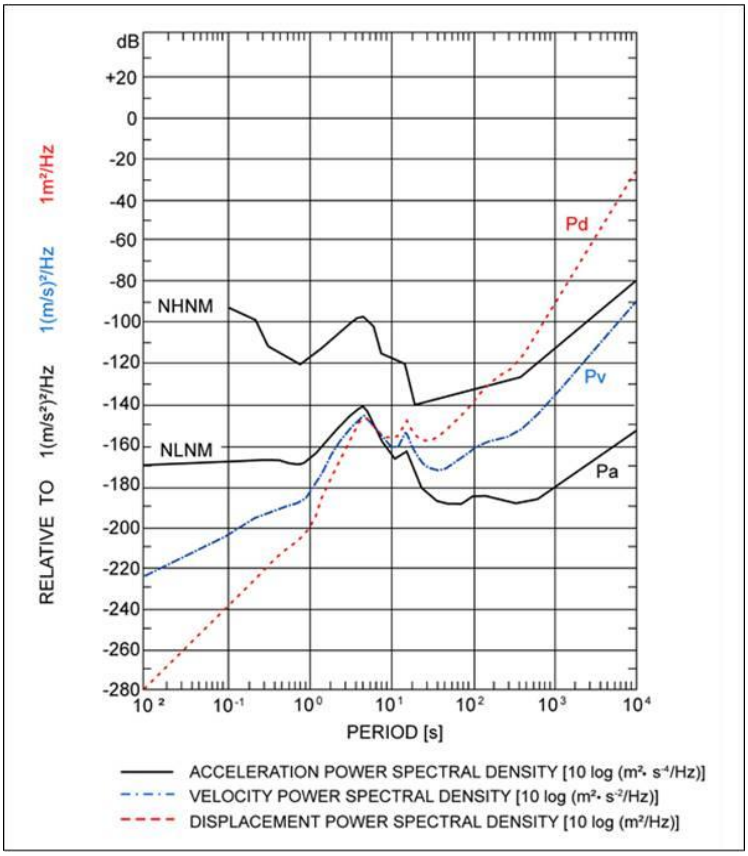

Fig. 1.1 - Black solid lines represent the envelope curves of acceleration power spectral density of the ambient vibrations vertical component (Pa**,** expressed in units of dB related to 1 $(m/s^2)^2/Hz$) as a function of period (according to Peterson, 1993). The upper envelope is the new global high noise model (NHNM) and the lower one corresponds to the new low noise model (NLNM). Dashed lines represent the displacement and velocity power spectral density (Pd in red and Pv in blue) calculated for the NLNM (from Bormann and Wielandt, 2013).

Despite these global analogies, the ambient vibrations are characterized by a large temporal and spatial variability.

Fig. 1.2-A show some records of the seismic noise vertical components acquired in the same day (7 December 1991) in the city of Mianeh (Iran) at different hours. Such records show different trends and amplitudes after some hours and similarity during the record length of about 10 seconds. It is possible to note that the main differences exist between day and night, where the most important variations in human activity (traffic, factories, etc…) occur. Fig. 1.2-B shows that the power

spectral density values of these records are characterized by important dissimilarities at frequency higher than 1 Hz, confirming their relation with cultural noise.

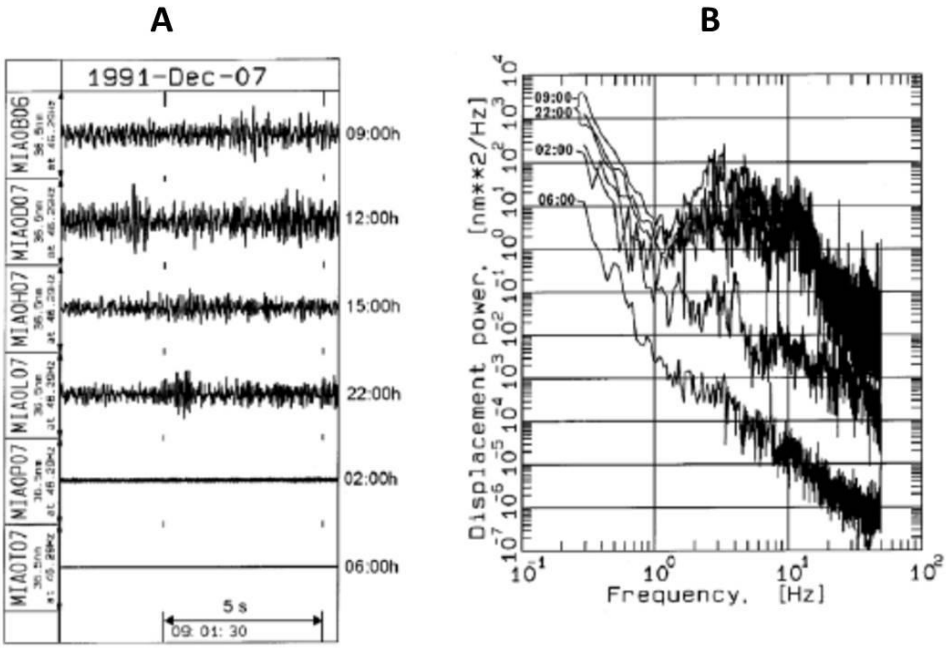

Fig. 1.2 - Comparison of vertical components of some ambient vibration time series (A) and related displacement power spectra (B) obtained in the town of Mianeh (Iran). The measurements were made at different times of the day (from Bormann and Wielandt, 2013).

Other interesting results about seismic noise time variability were obtained by Yamanaka *et al.* (1993). In this case, ambient vibrations were recorded for 10 minutes every two hours between 17 and 25 September 1990.

Fig. 1.3 shows the variations in amplitude of the signal harmonic components with period 0.3 and 6.5 seconds, along with swell heights detected in the oceanographic station of Begg-Rock, located 100 km south-west of Los Angeles. It is possible to note that the component of 0.3 seconds has a regular and daily amplitude variation, with a maximum at midday and a minimum at midnight; moreover, the relative amplitudes are smaller on Saturdays and Sundays. As concerns the component with period of 6.5 seconds, it manifests a clear correlation with the oceanic motions despite a slight phase shift.

These outcomes emphasize the unequal behaviour of ambient vibrations in the spectral domain, witnessing that the amplitude variations between different days are mainly linked with low frequency ocean activity.

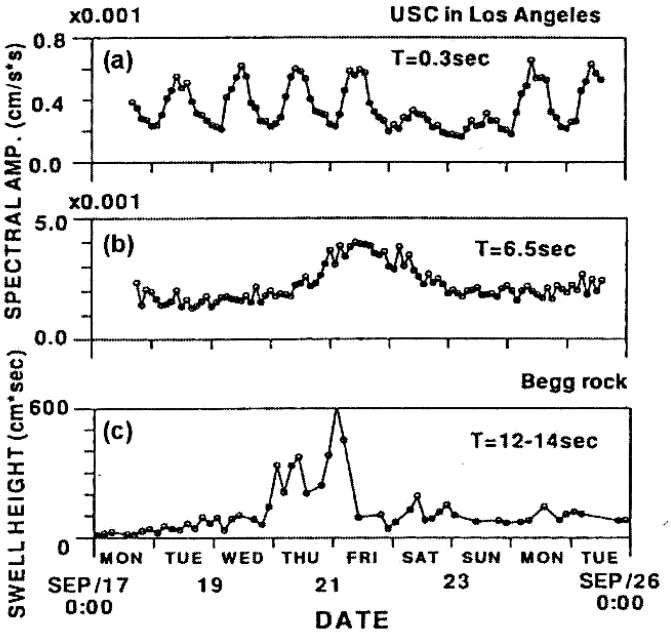

Fig. 1.3 - Spectral amplitude variations of ambient vibrations over time at 0.3 seconds (a) and 6.5 seconds (b). These spectral amplitudes were obtained by the geometrical mean of two horizontal sensors installed in Los Angeles. (c) Recordings of swell height variations over time about a hundred kilometres from Los Angeles (from Yamanaka *et al.*, 1993).

In Fig. 1.4 the power spectra of seismic noise acquired at the same time in 10 seismic stations around the city of Obihiro (Japan) are showed (Okada, 2003). The spectra are obtained by records of 45 minutes length and are compared in 6 different times of 21 and 22 July 1989.

It is clear that the overall spectral structure is similar but differences are evident, especially for frequencies higher than 1.5 Hz, both in the different times and for the different stations. Considering this frequency range, two important characteristics arise:

- how showed previously, the temporal variation can be attributed to the daily change of anthropic sources energy;
- the spatial variation between the seismic stations is related to the microtremors response respect to the local geological configuration.

On the contrary, the power spectra for the frequencies below 1.5 Hz are not characterized by significant differences, appearing more stable especially for the different measure sites.

This example allows realizing that the ambient vibrations contain significant information about the soil response, in that they vary spatially due to the local characteristics of the subsoil for the same energization conditions (type and distribution of the sources).

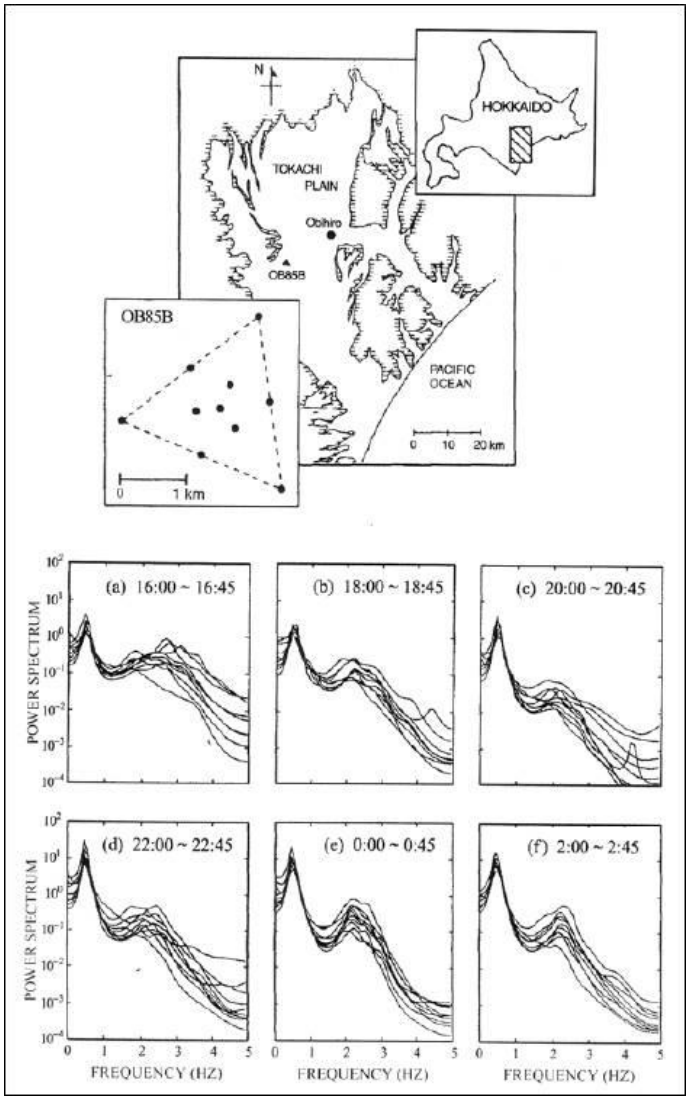

Fig. 1.4 - Power spectra of 10 seismic stations located around the city of Obihiro (Japan). The spectra were obtained by ambient vibration recordings with 45 min length at 6 different times. The upper panel shows the spatial configuration of the stations and their relative distance (from Okada, 2003)

1.3 Nature and composition of the ambient vibration wavefield

Considering the kind of phenomena that generate the ambient vibration wavefield, the ground shaking associated is characterized by considerable complexity and irregularity.

Several studies demonstrated that the seismic noise is constituted by a combination of body and surface seismic waves, propagating in different directions (SESAME, 2004a). This set of waves varies spatially and over time, depending on the physico-mechanical properties of the means they cross, as well as on the characteristics and the position of their sources.

Despite the conflicting outcomes resulting from the main works in literature, the understanding of the ambient vibrations nature quantifying the proportion of each type of waves is a paramount key to develop the forward model to be implemented in the inversion procedures.

The early studies carried out in this topic were devoted to explore the ratio of body to surface wave energy in the seismic noise wavefield. Douze (1964) performed seismic noise acquisitions at the surface and at various depths: he noted a clear exponential decay with depth of the signal amplitude, as well as a decay rate decreasing respect to the period (Fig. 1.5). This characteristic suggests the significant presence of surface waves.

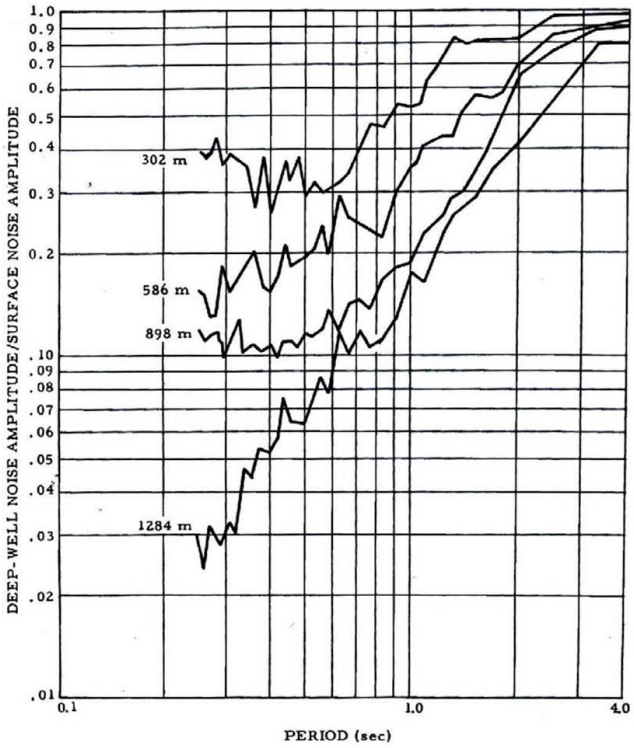

Fig. 1.5 - Ratio of spectral amplitudes (vertical component) of seismic noise at different depths over surface noise amplitude as a function of period (from Douze, 1964).

Table 1.2 summarizes the results of vary authors. Despite the outcomes are quite different (since the microtremors can exhibit remarkable differences site by site) it is noticeable the important role of the surface waves, especially at frequencies lower than 1 Hz. Instead, at higher frequencies the seismic noise wavefield is suspected to be a mix of body waves and Rayleigh waves (fundamental mode and/or higher modes).

	0					0.5					1					1.5					2			Hz
Douze						R1/P					R1/P										R3/P			
Toksoz	Ro		R+	P/R+		P																		
Li 1984														R+	and /	or	P						=>	20hz
Horike 1985								Ro								Ro /	R+						=>	3 hz
Yamanaka 1994				Ro																				

P = Body waves
R*n* = Rayleigh waves
n=0 : fundamental mode
n=1,2 ... : ***n*** higher mode
n=+ : higher modes (no order precision)

Tab. 1.2 - Synthesis of different authors studies (Douze, 1967; Toksoz, 1964, 1968; Li, 1984; Horike, 1985; Yamanaka, 1994) about nature of ambient vibration wavefield (from SESAME, 2004a).

More recently, some authors attempted to quantify the relative proportion of Rayleigh and Love waves in ambient vibration wavefield. To achieve this purpose, they performed seismic arrays measurements using three-directional sensors. Table 1.3 synthesizes the results obtained by these works.

	Frequency range	Rayleigh waves (%)	Love waves (%)
Chouet et al. 1998 (volcanic tremor)	*> 2 Hz*	*30 %*	*70 %*
Yamamoto 2000	3 – 10 Hz	< 50 %	> 50 %
Arai et al. 1998	1 – 12 Hz	30 %	70 %
Cornou 2002	< 1 Hz	60 %	40 %

Tab. 1.3 - Synthesis of the conclusions after different studies (Chouet et al., 1998; Yamamoto, 2000; Arai and Tokimatsu, 1998; Cornou, 2002) about the proportion of Rayleigh and Love waves in seismic noise (from SESAME, 2004a).

The outcomes, obtained through different array techniques, are quite ambiguous. Considering the poor state of art in this argument, it is impossible to deduce general conclusions about this proportion. For frequencies higher than 1 Hz, the results pointed out by Yamamoto and Arai *et al.* are quite in agreement; for frequencies lower than 1 Hz it is difficult to attempt reliable hypothesis since only one study has been done (Cornou in 2002).

An interesting study was performed by Tokimatsu (1997). He showed that the surface-wave propagation through stratified medium can be characterized by a modal nature.
Modeling synthetic seismic noise for three different shear wave velocity profiles, he computed the Rayleigh waves dispersion curves and the relative dispersion curves of the higher modes (Fig. 1.6).
It is possible to note that in the case 1, where V_S increasing monotonically with depth, the computed dispersion curve follows the theoretical dispersion curve of fundamental Rayleigh mode. Instead, in the case 2 and 3, where the V_S values varies irregularly with depth, the higher

modes can increase their energy content at higher frequencies affecting significantly the Rayleigh wave velocity.

(a)

Layer No.	Thick-ness H (m)	Density ρ (Mg/m³)	V_p (m/s)	V_s (m/s)		
				Case 1	Case 2	Case 3
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	2	1.8	360	80	180	80
2	4	1.8	1000	120	120	180
3	8	1.8	1400	180	180	120
4		1.8	1400	360	360	360

(b)

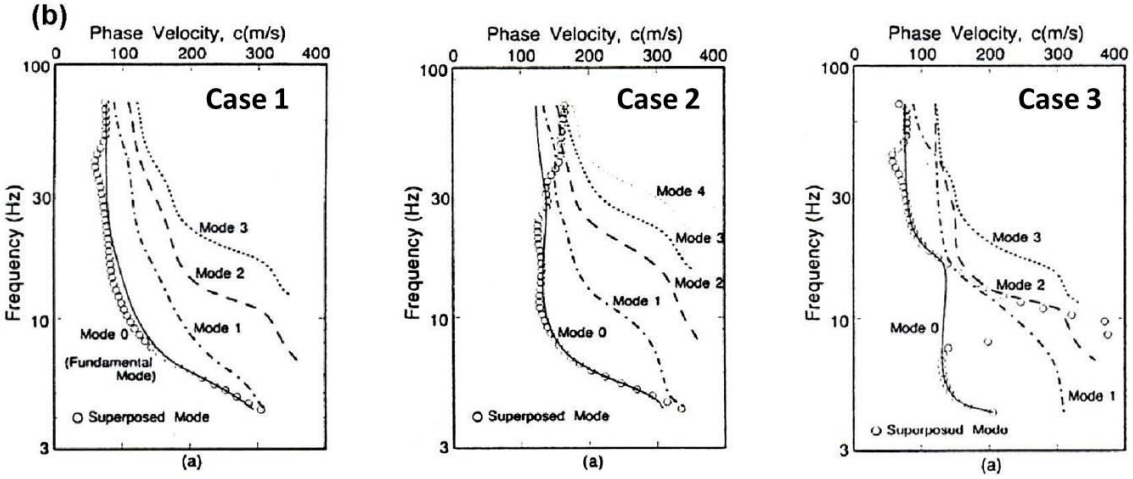

Fig. 1.6 - Velocity profiles (a) used by Tokimatsu (1997) to compute numerically Rayleigh waves dispersion curves (b). Computed dispersion cures (dots) from synthetic seismic noise are compared with theoretical higher modes dispersion curves of Rayleigh waves (from SESAME, 2004a).

1.4 Ambient vibrations as a stochastic and stationary process

The body and surface waves forming the ambient vibration wavefield are generated by different kinds of sources, randomly distributed both in space and in time. These characteristics and the various geological conditions in which the waves propagate (Fig.1.7), make complex the form of the recorded signal, which shows an evident randomness in terms of amplitude. In view of this, the seismic noise amplitude can be defined as a stochastic phenomenon, that is a process characterized by undetermined and unrepeatable features at a certain place and time.

The study of this kind of processes requires statistical approaches able to reduce the complexity of the analysis: in particular, if the amplitude of the ambient vibrations is considered a stochastic variable, it is possible to define a probability density function to describe it (Okada, 2003).

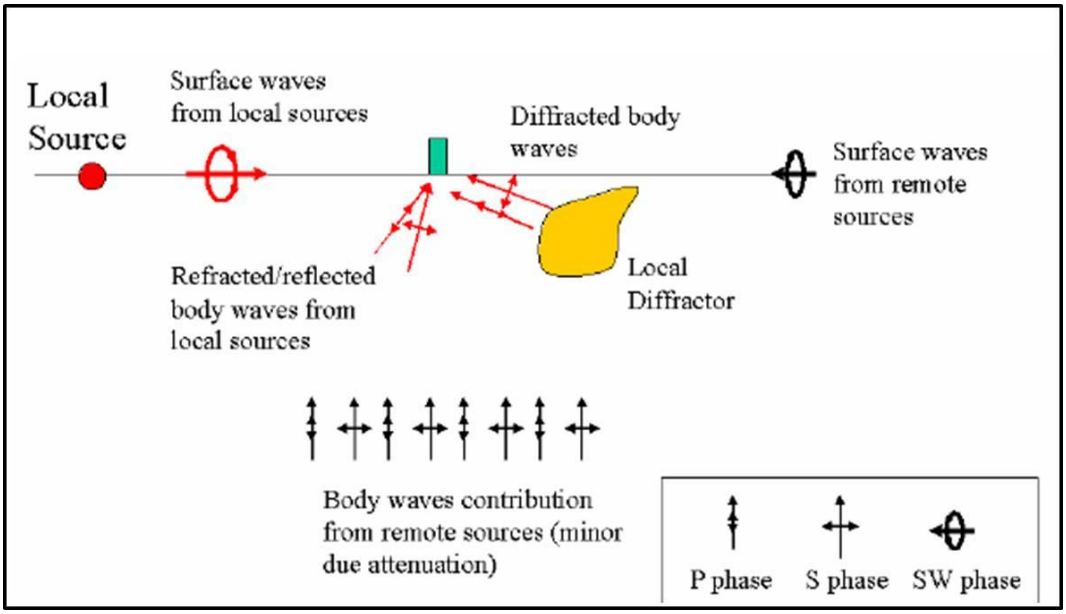

Fig. 1.7 - Qualitative scheme of the ambient vibration wavefield.

Nogoshi and Igarashi (1970) and Sakaji (1998) showed that the probability density function of the amplitude of the ambient vibrations approximates a normal distribution. In particular, this latter author, analyzing records of 10 minutes maximum length, demonstrated that the frequency distribution of the amplitude is well approached by a Gaussian distribution when data length exceeds 4 minutes (Fig. 1.8). Furthermore, he showed that the average, the standard deviation and the autocorrelation function of the microtremors amplitude are characterized by very little variations (Fig 1.8).
These results prove that the ambient vibrations satisfy the properties of a stochastic and temporally stationary process. However, this is not always true when the acquisition exceed the length of few hours (Okada, 2003). Moreover, as demonstrated by Groos and Ritter (2009), considerable deviations from a Gaussian distribution can be due to large amplitude transient and sinusoidal signals (cars, footsteps, industrial machinery, etc...) that can interfere during the recording time (Fig. 1.9).

In view of what described above and considering that the spatial stationarity of the microtremors is strictly related with the geological characteristics of the subsoil, it is evident that the ambient vibrations can be considered a stochastic and stationary process for a suitable spatial extent and for a time length of few hours. This assumption is of paramount relevance for the analysis techniques based on the seismic noise wavefield acquisition.

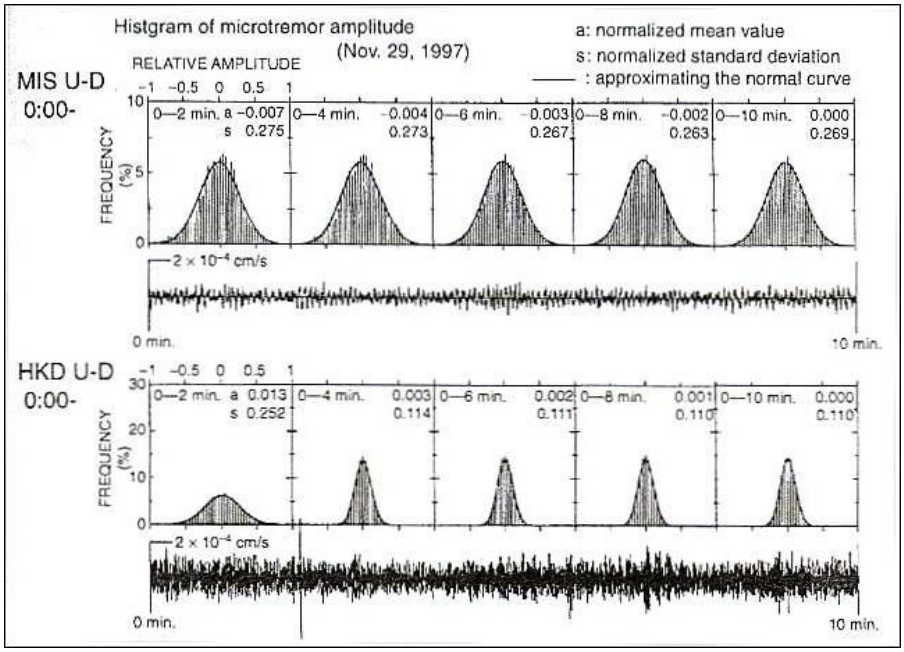

Fig. 1.8 - Probability density functions for two seismic stations (MIS and HKD) of the vertical component amplitudes of the microtremors, estimated progressively expanding the sampling window from 2 to 10 minutes. The curves represent the probability density function determined by assuming the histogram follows normal distribution (from Okada, 2003).

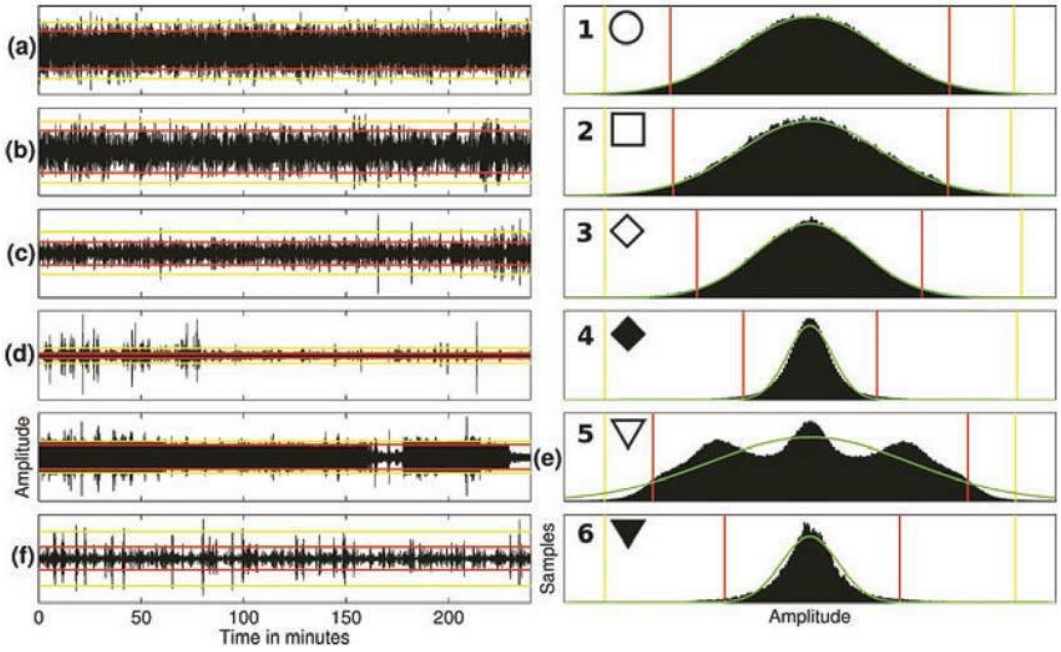

Fig. 1.9 - Time series (on the left) and their histograms with the Gaussian distributions (on the right) of seismic noise vertical component recorded in Bucharest according to Groos and Ritter (2009). The red and yellow lines on the right delimit the 2-sigma and 3-sigma range of the Gaussian distributed time series. Note that more the time series are dominated by transient signals, more the frequency distributions deviate from the respective Gaussian ones. The non-bell shaped multimodal distribution under (e) is due to time series dominated by sinusoidal signals and the asymmetric distribution under (f) is due to the dominance of asymmetric signals (from Bormann and Wielandt, 2013).

CHAPTER 2
The HVSR technique

2.1 Introduction

The Horizontal-to-Vertical Spectral Ratio (HVSR or H/V) technique is widely used tool for a quick detection and evaluation of seismic-amplification effects in terms of S-wave resonance frequency by using a single-station measurements carried out on the Earth's surface. This technique requires the acquisition of the three components of the ground motion and consists in performing the ratio between the horizontal and vertical Fourier spectrum. This ratio, which is expressed as a function of the frequency, is called the H/V (or HVSR) curve (or function).
This method was first introduced by Nogoshi and Igarashi (1971), and revised by Nakamura (1989, 1996, 2000). These authors have highlighted the correlation between the frequency value of the H/V peak and the fundamental resonance frequency of the site (f_0), thus suggesting the capability of this technique to provide information on the dynamic characteristics of the subsoil (see Eq. 3). In the last twenty years, many studies based on experimental and theoretical approach (e.g., Field and Jacob, 1993; Lermo and Chavez-Garcia, 1993; Lachet and Bard, 1994; Fäh, 1997) have shown that the HVSR method can be successful applied for identifying the fundamental S-wave resonance frequency of sedimentary deposits when a sharp impedance contrast occurs between the soft materials and the underlying stiffer formations. However, some disputes concerning the applicability of this technique to evaluating the site amplification are ongoing. In particular, many authors consider the H/V peak as a good estimate of the fundamental resonance frequency of the site, but unable to estimate the site amplification factor (e.g., Kudo, 1995; Bard, 1999; Konno and Omachi, 1998; Mucciarelli, 1998; Fäh *et al.*, 2001): in this case, the HVSR peak amplitude and the site amplification would be connected only by a qualitative relationship.

2.2 HVSR theory

The HVSR theory is characterized by some controversial aspects. These issues concern the exact physical interpretation of the H/V curve and most of them are related with the nature of the ambient vibration wavefield and of its sources. In fact, in order to exploit the H/V curve to infer the dynamic characteristics of the subsoil, theoretical models are needed to relate this curve with the mechanical and geological properties beneath the investigated site. Nowadays, several theoretical models of the ambient vibration wavefield exist. In particular, they are mainly divided in:

- models based on body waves theory, where it is assumed that body waves play a dominant role in the ambient vibration wavefield;
- models where it is assumed a dominant role of the surface waves and where the H/V ratio is interpreted in terms of ellipticity of Rayleigh waves;

- models where no ex ante assumption is required about the seismic phases present in the expected wavefield. For this, the HVSR curve is provided by full wavefield simulations.

The differences in the H/V curve modelling can affect significantly the outcomes of inversion procedures used to retrieve the V_S profile from the experimental data.

2.2.1 HVSR ratio by body waves theory

In the first theoretical explanation proposed by Nakamura (1989), the author assumes that the ambient vibrations wavefield is composed by S and Rayleigh waves. He computes the H/V spectral ratio by using this relationship:

$$HVSR(\omega) \equiv \frac{A^{FW}_{H,surface}(\omega)}{A^{FW}_{V,surface}(\omega)}, \tag{2.1}$$

where $A^{FW}_{H,surface}$ and $A^{FW}_{V,surface}$ are the total (*FW* means full wavefield) spectral amplitudes of the horizontal and, respectively, vertical ground-motion at the Earth's surface, and ω is the angular frequency.

The Author shows that the ratio in Eq. 2.1 "removes" the effect of surface waves and, at the fundamental frequency f_0, equals the horizontal soil transfer-function normalized by the vertical one considering two main assumptions:

1. the analogous H/V spectral ratio computed at the bedrock is approximately unitary;
2. the vertical component of the ground motion is not amplified.

In this view, the H/V ratio directly represents the site amplification function interesting the horizontal component of the ground-motion.

Nakamura (2000) revised this theory by considering the ambient-vibration wavefield composed by just vertically incident P and S waves along with Rayleigh waves. Then, he separates these different contributions by considering the horizontal and vertical spectral amplitudes of the ground-motion at the Earth's surface in this way:

$$\begin{aligned} A^{FW}_{\delta,surface}(\omega) &= A^{BW}_{\delta,surface}(\omega) + A^{SW}_{\delta,surface}(\omega) \\ &= T_\delta(\omega) \cdot A^{BW}_{\delta,bedrock}(\omega) + A^{SW}_{\delta,surface}(\omega), \end{aligned} \tag{2.2}$$

where δ indicates horizontal (*H*) and vertical (*V*) component, *BW* are the body waves, *SW* the surface waves and T_δ are the horizontal and vertical transfer functions. Taking in account this expression, the H/V ratio at the surface can be defined as (Lunedei and Malischewsky, 2015):

$$HVSR(\omega) = \frac{A^{BW}_{H,bedrock}(\omega)}{A^{BW}_{V,bedrock}(\omega)} \cdot \frac{T_H(\omega) + \frac{A^{SW}_{H,surface}(\omega)}{A^{BW}_{H,bedrock}(\omega)}}{T_V(\omega) + \frac{A^{SW}_{V,surface}(\omega)}{A^{BW}_{V,bedrock}(\omega)}}. \quad (2.3)$$

Assuming that the H/V spectral ratio at the bedrock is approximately unitary, from Eq. (2.3) it is possible infer that the HVSR function is close to T_H/T_V when the surface-wave contribution is not significant.

In this theory, amplification in P-waves is not expected near the S-wave resonance frequency, since P-wave velocity is supposed to be many times greater that the S-wave one, so the H/V function should approximate the S-wave transfer function, around its peak frequency (Lunedei and Malischewsky, 2015). If instead Rayleigh waves play a more important role, the HVSR curve approximates the ratio between the horizontal spectral amplitude and the vertical spectral amplitude of the surface waves, whose peak frequency should approximate the S-wave site resonance frequency. Finally, the author states that the maximum of the H/V function represents the S-wave fundamental resonance frequency and the relative amplification factor, regardless of the considered seismic phases.

In agreement with Nakamura (2000), Herak (2008) proposed a theory to compute the H/V curve which only involves P and S waves vertically incident on horizontal, homogeneous and isotropic layers. Differently by the previous body waves theory, no a priori hypothesis is made about the P-wave amplification. Assuming further that horizontal and vertical motions are approximately equal on the bedrock, the HVSR function at the surface is:

$$HVSR(\omega) \equiv \frac{A^{BW}_{H,surface}(\omega)}{A^{BW}_{V,surface}(\omega)} = \frac{AMP_S(\omega)}{AMP_P(\omega)}, \quad (2.4)$$

where $A^{BW}_{H,surface}$is the S-wave (horizontal) spectral amplitude and $A^{BW}_{V,surface}$ is the P-wave vertical) spectral amplitude, both computed on the surface. AMP_P and AMP_S are the P-wave and S-wave amplification spectra.

Unlike the Nakamura's theory, Eq. (2.4) shows that a direct estimation of the S-wave transfer function by the H/V ratio is not always possible, because it is clear that $HVSR(\omega) \simeq AMP_S(\omega)$ only if $AMP_P(\omega) \simeq 1$ for the frequencies of interest. This approximation is valid only when P-waves have much higher resonant frequency value than S-waves one.

2.2.2 Interpretation of HVSR ratio in terms of ellipticity of Rayleigh waves

As described in Chapter 1, the composition of ambient vibrations in term of the different seismic phases is not yet clearly understood. In particular, the most controversial aspect is related with the estimation of the relative contributions of all the various phases that travel through the subsoil.
In fact, from the results derived by in field experiments and numerical simulations, it is possible to infer that significant variations of the different seismic phases content in the ambient vibrations wavefield can occur and that they depend on the subsoil stratigraphy and on the sources characteristics as well as on the different frequency ranges.
In view of this, beside the models based on body waves theory, models that consider the surface-wave dominance were developed. In particular, several authors (e.g., Lachet and Bard, 1994; Tokimatsu 1997; Konno and Ohmachi, 1998; Fäh *et al.*, 2001; Wathelet *et al*., 2004) agree on the close relation existing between the H/V spectral ratio and the ellipticity of Rayleigh waves. This relation is justified by the simple consideration that Rayleigh waves are generally energetically predominant in the vertical component of ambient vibration wavefield.
Fig. 2.1 shows the theoretical ellipticity curves for the first 5 modes of Rayleigh waves. One of the most important characteristics of these curves is the frequency dependence: in fact, the infinite peaks and troughs occurring at various frequencies in the ellipticity curves correspond to the vanishing of the vertical and horizontal component, respectively.

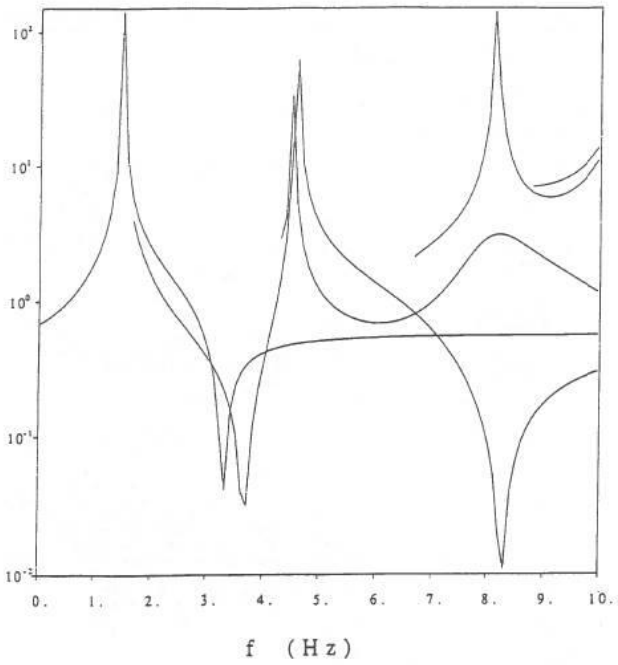

Fig. 2.1 - Example of ellipticity curves for Rayleigh waves in a stratified medium. The ellipticity curves (log scale) are displayed as a function of the frequency for the 5 first higher modes (from Bard, 1999).

Konno and Ohmachi (1998), point out that the H/V peak by ambient vibrations could be explained by the ellipticity of the fundamental Raylegh mode as well as by the Airy phase of the fundamental

Love mode. In particular, these authors investigate the particle motion related to fundamental Rayleigh mode as a function of the fundamental period T of the subsoil profile and, in particular, of the velocity contrast between the single layer and the half-space. They showed that only in presence of very high impedance contrasts, for increasing period the H/V curve corresponds to the presence of horizontal motion only (i.e., vanishing of the vertical one). It is worth noting how this frequency dependent particle motion simply can explain the presence of peaks and troughs in the H/V curves.
Moreover, in numerical simulations performed by these authors, the H/V-peak amplitude roughly approximates the S-wave amplification factor, providing that a specific proportion between Rayleigh and Love waves exists; this mimics the Nakamura's statement, but in terms of surface waves instead that of body waves (Lunedei and Malischewsky, 2015).

Fäh *et al.* (2001) explore two ways to generate H/V synthetic curves. The first one is a numerical simulation made by a finite difference technique, where both superficial and buried ambient vibration sources are considered. These sources, with positions, depths and time-dependences chosen randomly, are distributed around a receiver. The second technique focuses on Rayleigh wave ellipticity of fundamental and higher modes, to explain the H/V-peak frequency: in fact, only this element is considered reliable in the HVSR curve by the authors, in that the peak amplitude and other features of the curve also depend on other parameters besides than the S-wave velocity profile.
Moreover, they identify stable parts of the H/V ratio that are independent on the sources distance and are dominated by the ellipticity of the fundamental Rayleigh mode, in the frequency band between the H/V-peak frequency (which they check to be close to the site S-wave fundamental resonance-frequency) and the first minimum of the H/V curve (Lunedei and Malischewsky, 2015).

More recently, Bonnefoy-Claudet *et al.* (2004, 2006a, 2008) carry out a systematic study of the H/V curve by numerical simulations, in which ambient vibrations are generated by a multitude of point-like forces, randomly oriented in the space and located relatively near to the observation point. By using several simple stratigraphical profiles, Bonnefoy-Claudet *et al.* (2008) check the good correspondence between the H/V-peak frequency and the S-wave resonance one. Moreover, they shows that the interpretation of the H/V main peak in terms of Rayleigh ellipticity seems limited to profiles characterized by high impedance contrast (more than 4).
Overall, this work highlighted the importance of all surface wave phases propagating in the subsoil in constructing a suitable H/V model, as well as the paramount role of the impedance contrast in computing the HVSR peak.

2.2.3 HVSR by full wavefield simulations

The models described previously confirm that all seismic phases should be taken into account to provide a reliable interpretation of the H/V curves. The recent theories based on full wavefield simulations assume that both body and surface waves play a significant role in the ambient vibration wavefield.

Lunedei and Albarello (2010), following the work of Field and Jacob (1993), developed a model considering the case of a continuous distribution of random independent point sources, which are active at the surface of a flat, viscoelastic layered Earth. A limit of this model, later denominated as DSS (Distributed Surface Sources), is that it does not regard the presence of unavoidable deep scattering, because sources are only superficial, but it has to be said that the hypothesis of a stratified Earth presumes that the discontinuities due to layer-interfaces are more relevant that any other inhomogeneity (included the deep scattering generators). In addition to the full wavefield model, a surface wave version of DSS was developed (Lunedei and Albarello, 2009): in both models, the H/V curve is computed as square root of the ratio between the average spectral powers on the horizontal plane and along the vertical direction.

Albarello and Lunedei (2011) obtain some insights about this full wavefield model of the ambient vibration structure. For a stratigraphical profile-set equal to the group M2* in Table 2.1, three frequency ranges are identified (Lunedei and Malischewsky, 2015):

- low-frequency domain (below the S-wave resonance frequency value, f_S), where the spectral powers of the ambient vibrations are relatively low. In this domain, the soft shallow layer acts as a high-pass filter and both near sources and body-wave phases dominate the wavefield;
- high-frequency domain (above $\max\{f_P, 2f_S\}$, where f_P is the P-wave resonance frequency), where the surface waves (both Love and Rayleigh, in their fundamental and higher modes) dominate the wavefield. Spectral powers smoothly decrease with frequency as an effect of material damping;
- intermediate frequency domain (at and in-between f_S and $\max\{f_P, 2f_S\}$), where most of the seismic energy of the ambient vibrations is concentrated. In this range, sharp peaks in the horizontal and vertical spectral powers are revealed around its bounds f_S and $\max\{f_P, 2f_S\}$. Irrespective of the subsoil structure and source-free area considered, horizontal ground motion is dominated by surface waves, with a varying combination of Love (in the fundamental mode) and Rayleigh waves, that depends on the Poisson's ratio. In the vertical component, Rayleigh and other phases play different roles, both depending on the dimension of the source-free area and of *VP* and *VS* depth profiles.

In synthetic H/V curves produced by this model, the peak frequency is generally very near to f_S, regardless to the Poisson's ratio and the dimension of the source-free area. A weak sensitivity is revealed with respect to the impedance contrast only, and these findings enforce the common idea that the H/V peak-frequency is a good estimate of f_S (Lunedei and Malischewsky, 2015). In general, the H/V peak-amplitude value increases with the impedance contrast, but no linear

relationship exists between these two parameters. Moreover, a significant dependence of the H/V-peak amplitude on the source-free area dimension (the amplitude tendentially increases with its dimension), on the shallow-layer Poisson's ratio (the amplitude increases with it) and thickness (the amplitude decreases when the thickness increases) was obtained.

M2					
***h* (m)**	***V_S* (m/s)**	***v***	***ρ* (g/cm^3)**	***D_P***	***D_S***
25	200	0.333	1.9	0.001	0.001
5000	1000	0.333	2.5	0.001	0.001
∞	2000	0.257	2.5	0.001	0.001
M2*					
***h* (m)**	***V_S* (m/s)**	***v***	***ρ* (g/cm^3)**	***D_P***	***D_S***
25	200	0.01-0.49	1.9	0.001	0.001
5000	228-1520	0.333	2.5	0.001	0.001
∞	2000	0.257	2.5	0.001	0.001
M3					
***h* (m)**	***V_S* (m/s)**	***V_P* (m/s)**	***ρ/ρ_4***	***D_P***	***D_S***
5	30	500	1	0.001	0.001
25	100	500	1	0.001	0.001
50	150	500	1	0.001	0.001
∞	500	1500	1	0.001	0.001

Tab. 2.1 - Stratigraphical profiles used in the numerical testing. The symbols h, V_S, V_P, v, ρ, D_P and D_S represent, respectively, the values of thickness, S-wave velocity, P-wave velocity, Poisson's ratio, density, damping for P-wave and damping for S-wave (from Lunedei and Malischewsky, 2015).

More recently, some authors (Sánchez-Sesma and Campillo, 2006; Sánchez-Sesma *et al.*, 2011; Kawase *et al.*, 2011) proposed another kind of model named DFA (Diffuse Field Approach), where it is assumed that ambient vibrations constitute a diffuse wavefield. In this assumption, the seismic waves propagate in every (three-dimensional) spatial direction in a uniform and isotropic way. Moreover, a specific energetic proportion between P and S waves exists, which is always the same, both spatially and temporally. The link between the H/V curve and the subsoil configuration is simply given by the Green's function, computed for source and receiver located in the same position: its imaginary part, in the spectral domain, is proportional to the average spectral-power of ambient-vibration ground-motion. A key element in the DFA theory, which is implied in the diffuse character of the wavefield, is the loss of any trace of the sources' characteristics, so no link between displacement and its sources is involved in this theory, *ergo*, no description of ambient-vibration sources is needed (Lunedei and Malischewsky, 2015). This model can be use to describe the ambient vibrations considering the presence of all the seismic phases (full wavefield) or their surface-wave component only.

In order to study the differences and similarities between DSS and DFA models, a set of synthetic tests was performed considering the set of profiles showed in Tab. 2.1. The group of profiles M2*

is generated by varying the profile M2, and both sets basically consist of a layer overlying a half-space. Instead, the profile M3 is characterized by two main and a weak impedance contrasts.
As concerns the profile M2, DSS model shows a significant contributions of body waves for frequencies around and below the S-wave resonance frequency f_S (2 Hz in this case), and a clear surface-wave dominance for frequencies larger than the P-wave one f_P (4 Hz in this case). This characteristic causes the amplitude differences between the blue HVSR curve (full wavefield, FW) and the green one (surface waves, SW) in Fig. 2.2-a around the peak frequency. Instead, when a source-free area with a radius of 10 m is inserted in the simulations (Fig. 2.2-b), the FW produces an H/V peak very similar to the one of the SW. Regarding the DFA model, the H/V curves generated by its simulations show no difference between FW and SW; moreover, the H/V peak amplitudes are less that the DSS ones.

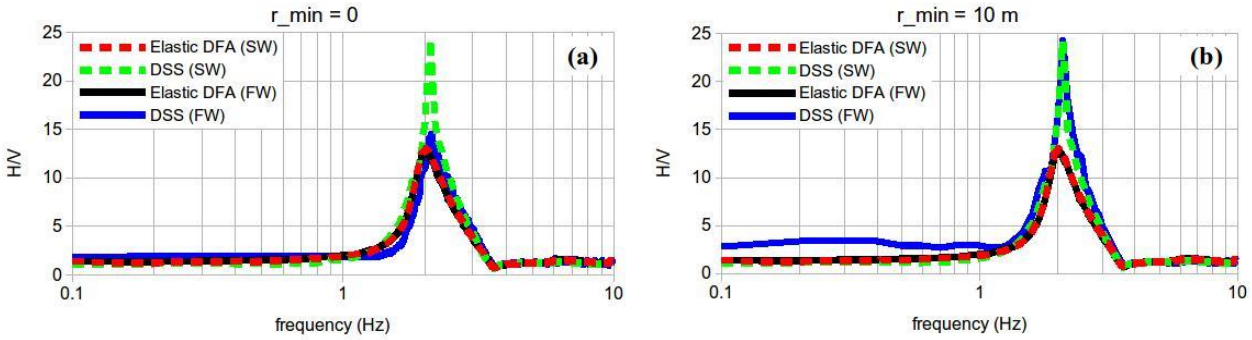

Fig. 2.2 - H/V curves for the stratigraphy M2 obtained by the DSS for the full-wavefield (blue) and the surface-wave component (green), as well as by the DFA for the full-wavefield (black) and the surface-wave component (red); a: no source-free area is considered; b: a source-free area with radius 10 m is set in the DSS (From Lunedei and Malischewsky, 2015).

Considering the profile M3, which presents more complicated configuration with main and secondary impedance contrasts, the H/V trends computed with DFA and DSS are characterized by more remarkable differences. The elastic DFA computation presents two maxima at about 0.4 and 1.5 Hz (probably associated with the two subsoil principal interfaces), both for the FW (black lines in Fig. 2.3) and the SW (red lines in Fig. 2.3). For the DSS-FW model (blue line in Fig. 2.3), it is possible to note that a more strictly correspondence with the DFA-FW curve exists only removing the close sources from the simulations (Fig. 2.3-b): in this way, the surface waves play a predominant role respect to the body waves. Differently, the SW models show very similar results in all investigated cases.

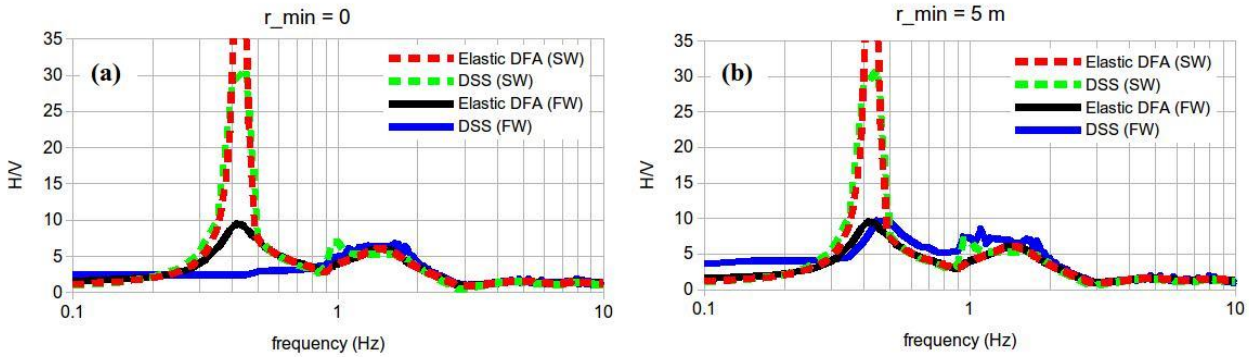

Fig. 2.3 - H/V curves from DFA and DSS method for the profile M3; a: sources are allowed on the whole Earth's surface; b: near sources are removed from around the receiver up to the distance of 5 m (from Lunedei and Malischewsky, 2015).

In view of this comparison, it is possible to conclude that both the DSS and the DFA provide reasonable full-wavefield and surface-wave synthetics of H/V spectral ratios, in particular for stratigraphic profiles with a dominant impedance contrast (profile M2). Moreover, it is clear that the results relative to DSS can depend on the source distribution around the receiver.

Lunedei and Albarello (2015) proposed a new version of the full-wavefiled DSS model. This new theory describes the ambient-vibration ground-motion displacement and its generating force fields as three-dimensional stochastic processes stationary both in time and space. In this frame, the displacement power can be linked with the source filed power *via* the Green's function, which, in turn, depends on the subsoil configuration (Lunedei and Malischewsky, 2015).
This formalization reduces computational efforts with respect to the previous version of the DSS model based on distributed surface sources and may provide synthetic HVSR-curve patterns that are in line with those given by that computationally more troublesome version, as well as with those deduced under the assumption that the ambient vibrations constitute a diffuse wavefield (Lunedei and Albarello, 2015).

2.3 HVSR curve computation

Several protocols of numerical analysis for the H/V function determination starting from experimental measurements exist (Mucciarelli, 1998; Bard, 1999; SESAME, 2004b). A significant element of the selected procedure is the ability to provide a stable HVSR curve as well as an estimate of its confidence interval.
Picozzi *et al*. (2005a) analyzed the statistical properties of the H/V spectral ratios in order to define optimal strategies for numerical processing and identification of possible artefacts. All the HVSR curves shown in this thesis are retrieved from ambient vibration recordings following this study and, in particular, using this procedure:

1. the recording is divided into windows of 20 to 60 seconds length (suitable time interval to ensure a good spectral resolution), each of which is analyzed separately. The number of temporal windows required, i.e. the recording duration, must be such as to ensure the statistical stabilization of the signal (e.g. about 20 minutes of recordings should be enough);

2. within each window, for each of the three ground motion components these analyses are performed:
   - linear trend removal (detrend) and tapering of the signal ends with a Hann cosine function (to avoid leakage phenomena);
   - spectral amplitudes computation by using Fast Fourier Transform (Cooley and Turkey, 1965);
   - the smoothing of spectra with triangular moving window with frequency dependent half-width (5% of central frequency). This is a mandatory operation to avoid the presence of spurious peaks and excessive irregularities in the H/V curve (e.g., Bindi *et al.*, 2000; Picozzi *et al.*, 2005a; Fig. 2.4);

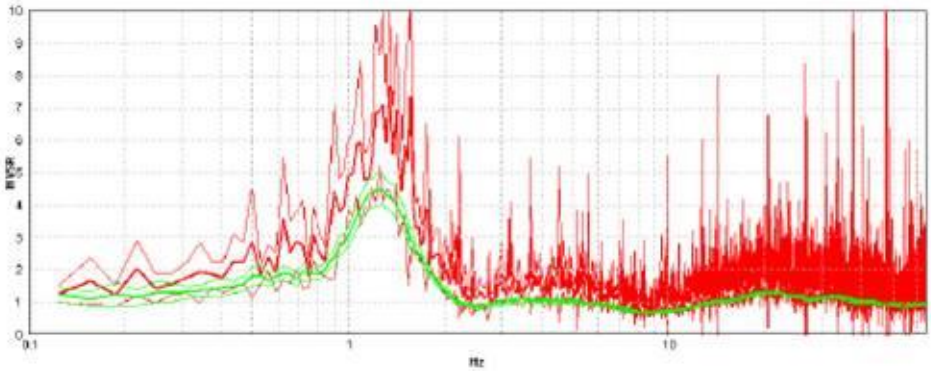

Fig. 2.4 - Experimental H/V curves smoothed with a moving triangular window with half-width equal to 5% of the central frequency (thick green line) and without smoothing (thick red line).

3. the geometrical mean of the two horizontal components (H_N and H_E) spectral amplitudes are computed through the relationship:

$$\overline{H} = \sqrt{H_N \cdot H_E} \tag{2.5}$$

where $\overline{H}$ represents the mean of the horizontal components amplitudes. After this, the H/V curves in each time window are obtained by the spectral ratio between $\overline{H}$ and the vertical component spectral amplitude;

4. the final H/V function and the corresponding 95% confidence interval (Fig. 2.5) are achieved by averaging the H/V ratios obtained for each time window, thus identifying the frequency value for which the H/V ratio is maximum (resonance frequency f_0).

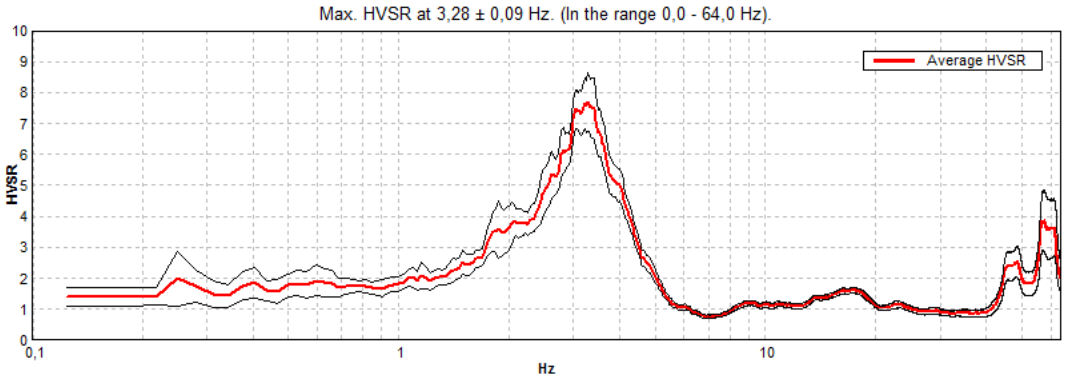

Fig. 2.5 - Example of experimental average H/V curve (thick red line) and the 95% confidence interval around average HVSR values (thin black lines).

As concerns the point 3, Albarello and Lunedei (2013) compared different definitions for the merging of the horizontal components (arithmetic mean, geometric mean, vector summation, quadratic mean and maximum horizontal value) for two different methods to define the average H/V spectral ratio:

a) the square root of the ratio between the arithmetic mean of the spectral powers on the L time-windows,
b) The arithmetic mean of the H/V ratios computed in each of the L time-windows (the method used in this thesis and described at point 4).

From this comparison, it results that the H/V estimates are biased of 46 % to more than 100 % and that, while the definition (a) quickly reduces its bias-size (for all cases 1–6) as L increases, this does not happen for the definition (b) (Lunedei and Malischewsy, 2015).
In conclusion, the results suggest that experimental H/V spectral ratio values could strongly depend on the processing strategy adopted for their estimate and that eventual inversion and testing procedures should account for these differences (Albarello and Lunedei, 2013).

2.4 Quality criteria of the HVSR curve

To estimate the quality of HVSR measure it is fundamental to assess the time stability of the H/V function during the acquisition (temporal stationarity) and the possible presence of directional phenomena that can produce spatial heterogeneities in the ambient vibration wavefield.
To perform these inspections, two analyses are carried out:

- to assess the stationarity assumption, the H/V spectral ratios obtained in the different time windows as a function of the recording time are represented (Fig. 2.6-A), thus identifying and removing the possible strong transient signals (e.g., cars or people passing too near the

sensor) that can significantly affect the outcomes of the whole analysis (Parolai et al., 2004);

- to check the directionality of the recorded signal, the horizontal components are rotated at regular intervals in the 0°-180° degree range in order to highlight the presence of significant spatial heterogeneity in the ambient vibration wavefield (Fig. 2.6-B). A wavefield characterized by a single dominant source may excessively affect the spectral characteristics of the signal, so that the relevant H/V ratio would no longer dependent only by the site physical properties.

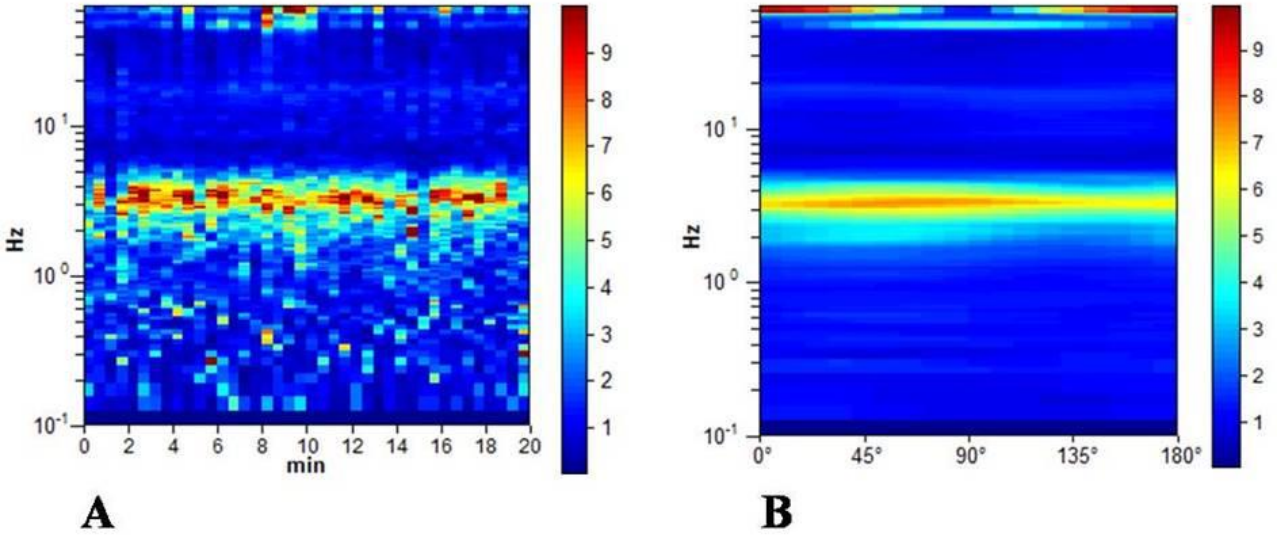

Fig. 2.6 - Contour plot of the H/V curves computed for each time window. The scale-bar on the right indicates H/V amplitudes (A). Contour plot of H/V curves obtained by the rotation and combination of the horizontal components of motions. The scale-bar on the right indicates H/V amplitudes (B).

To warrant optimal applications of HVSR prospecting techniques and to provide comparable measurements among themselves, a definition of quality standards to select meaningful measurements is required. An important aspect concerns the reliability and the clearness of the H/V function maxima: in fact, an unsatisfactory statistical quality of the measure can be produced spurious peaks that can hamper a correct identification of resonance effects. Some criteria in that regard are established in the framework of the European research project SESAME (Site Effects Assessment Using Ambient Excitations), though these statistical norms can not be considered exhaustive.

Recently, Albarello *et al.* (2011) proposed some classification criteria including a larger number of elements, turns out to be more complete than that proposed by the SESAME group (SESAME, 2004b). In order to provide an immediate indication about the quality of single HVSR measure and preventing over-interpretation of bad experimental results, the authors have developed the following classification scheme with three classes (Albarello *et al.*, 2011):

- *class A*: trustworthy and interpretable HVSR curve, which represents a reference measurement, that can be considered representative of the dynamical behavior of the subsoil at the site of concern by itself;
- *class B*: suspicious HVSR curve, which should be used with caution and only if it is coherent with other measurements performed nearby;
- *class C*: bad HVSR curve (it is hardly interpretable), to be discarded.

Criteria used to classify a single measurement as of class A are:

1. stationarity, i.e., the HVSR curve included in the frequency range of interest shows a persistent shape for at least the 30% of the measurement duration;
2. isotropy , i.e., the azimuthal amplitude variations do not exceed 30% of the maximum;
3. absence of artifacts, i.e., there are not symptoms of electromagnetic noise or peaks of industrial origin into the frequency range of interest;
4. physical plausibility, i.e., HVSR maxima are characterized by a localized lowering of the vertical amplitude spectral component, forming an “eye-shaped” structure (Fig. 2.7);

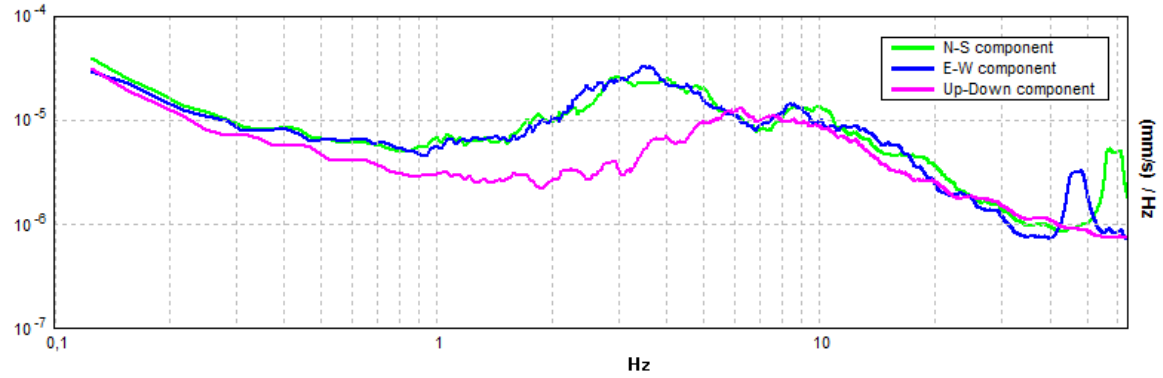

Fig. 2.7 - Amplitude spectra of the three components of the ground motion referred to the H/V curve in Fig. 2.5: it is possible to note a localized lowering of the vertical amplitude spectral component located in correspondence of the HVSR curve peak frequency.

5. statistical robustness, i.e., SESAME criteria for a reliable H/V curve are fulfilled;
6. representative sampling, i.e., the measurement took place for at least 15 minutes.

A measurement is in class B if one or more of the previous conditions are not fulfilled.

Measurements of class B become of class C if:

- a rising drift exists from low to high frequencies, that indicates a movement of the instrument during the acquisition (Forbriger, 2006), or
- electromagnetic or industrial disturbances affect several frequencies in the frequency range of interest (Fig. 2.8).

Actually, these criteria aim at the “first-glance” identification of good and unreliable measurements (A and C classes respectively) and at identifying doubtful results (B class), that require careful inspections, and they do not concern the possibility to provide a physical interpretation of the curve in terms of “absence/presence” of resonance phenomena. For this

purpose, the SESAME conditions for "peak clearness" are also taken into account. On this basis, two sub-classes (type) were introduced:

- *type 1:* the HVSR curve presents at least one clear peak in the frequency range of interest (possible resonance);
- *type 2:* the HVSR curve does not present any clear peak in the frequency range of interest (absence of resonance).

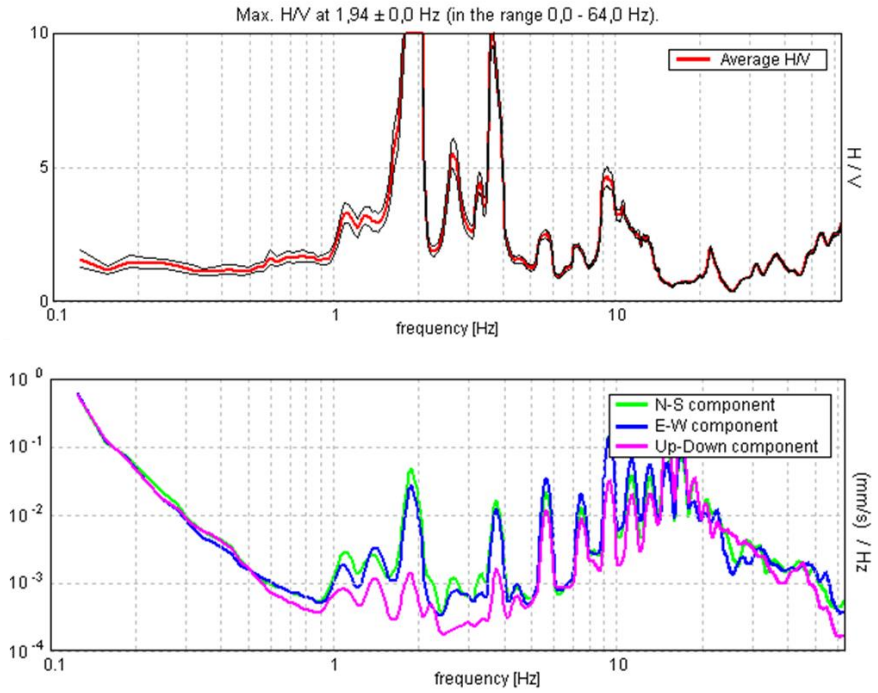

Fig. 2.8 - Effect of strong Industrial disturbances in the H/V curve and in the spectra of the three components of the ground motion.

2.5 Stratigraphic use and role in Seismic Microzonation studies of the HVSR curve

An important aspect related to the HVSR curve is its use for stratigraphic and exploratory purposes. By the direct estimate of the resonance frequency f_0 and exploiting the approximate relationship (Eq. 3) between this parameter and the thickness (H_{Sed}) and the average S-wave velocity (V_{Sed}) of the sedimentary cover, it is possible to retrieve preliminary information on the shallow subsoil structure. In particular, it is easy to deduce that once estimated the f_0 value identifying the H/V peak and known the V_{Sed} value, it is possible to estimate the resonant interface depth. Albarello *et al*. (2011) proposed a simple abacus (Tab. 2.2) in order to obtain a very rough estimate of the latter parameter in absence of S-wave information for the shallow materials and, therefore, exploiting only the knowledge of f_0 value.

f_0 (Hz)	h (m)
< 1	> 100
1 ÷ 2	50 ÷ 100
2 ÷ 3	30 ÷ 50
3 ÷ 5	20 ÷ 30
5 ÷ 8	10 ÷ 20
8 ÷ 20	5 ÷ 10
> 20	< 5

Tab. 2.2 - Approximate summary relationship between HVSR peak frequency (f_0) and resonant interface depth (h) (from Albarello *et al.*, 2011).

The cheapness, the robustness and the quickness that characterize the Horizontal-to-Vertical Spectral Ratios deduced by single station measurements (specially using very portable instruments) makes this analysis a very attractive tool in the first level of Seismic Microzonation studies and in the phases immediately following major damaging earthquakes (e.g, Albarello *et al.*, 2011; Gallipoli *et al.*, 2011; Priolo *et al.*, 2012). Considering the fact that areas of the order of hundreds of square Km are involved, in these preliminary frameworks extensive approaches are required than more detailed ones, often characterized by strong economic efforts and relatively long field procedures. Furthermore, in the emergency post-earthquake phases, it is of paramount importance to quickly identify the sites where provisory recovery area could be safely emplaced due to lack of significant local amplification effects during the seismic sequence.
The detection of these phenomena is one of the basic information (along with the estimate of f_0 and resonance interface depth) that the HVSR technique can provide: in fact, the presence of at least one maximum in the H/V curve is assumed as an indication of possible seismic resonance phenomena at the investigated site (Fig. 2.9-A), while a flat curve (Fig. 2.9-B) denotes their absence in the considered frequency range.
As well as the information described above, the use of the HVSR analysis in Seimic Microzonation and emergency contexts allows (Gallipoli *et al.*, 2011):

- to provide an estimate, at least qualitatively and comparatively, of the possible seismic impedance contrast taking into account the peak amplitude;
- to define the lithologies and the geological formations that can constitute the sedimentary cover and the bedrock;
- to provide a map of the resonance frequency of sedimentary covers, helping geologists to refine on formation limits and for identification of the areas with a danger of soil-structure resonance.

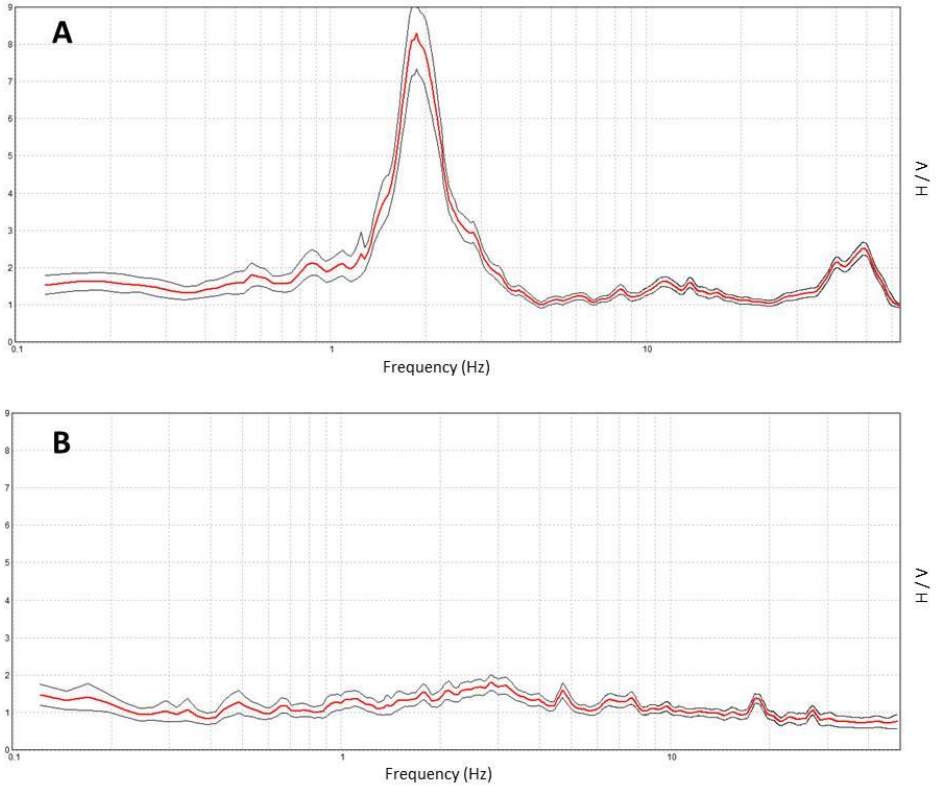

Fig. 2.9 - Comparison of different HVSR curve trends. A: an example of HVSR curve related to an amplifying site (the curve exhibits a clear sharp peak). B: flat HVSR curve that denotes the absence of seismic resonance phenomena in the whole investigated frequency range (0.1-64 Hz). Red lines represent the average HVSR curves and black lines their 95% confidence interval.

2.6 The effect of S-wave velocity inversion on HVSR curve

One of the purposes of this dissertation is the study of the effect of a sharp shear wave velocity inversion in the subsoil on the passive seismic measurements and, moreover, to verify the reliability and capability of these techniques in estimating the V_S profile in this unusual geological setting.

The aim of this paragraph is a review of the main works present in literature dealing this topic, in particular focusing on the influence of the S-wave velocity inversion on the HVSR curve. The literature about this issue is minimal and the main studies, based on experimental measurements and numerical testing, are performed in the last decade by Di Giacomo *et al*. (2005) and Castellaro and Mulargia (2009).

The first work was devoted to check the stability of the H/V ratio calculated at the Venosa seismic station site (Southern Italy), which lies on a geological setting characterized by a shallow velocity inversion (at about 20 m depth) due to the presence of a rigid conglomerates layer over clayey material. The authors analyzed nearly two years of data, composed of 244 triggered ambient vibration records (with relatively high amplitude) and 44 earthquakes, with the aim to compare

the resulting HVSR curves obtained by the two dataset. For the two years taken into account, the station recorded local, regional (including the San Giuliano Di Puglia 2002 sequence) and teleseismic earthquakes, covering a wide range of epicentral distances and azimuths.
The seismic noise HVSR curve, which is not affected by evident seasonal changes, is characterized by an high frequency peak (related with the presence of anthropogenic fill) and one at low frequency (due to the resonant interface between the clays and the bedrock) as well as a clear trend of the curve below 1 in the frequency range 1-8 Hz (Fig. 2.10). As concerns the frequency peaks (especially with the low frequency one), the H/V spectral ratios obtained by the seismic events records show a good correlation (Fig. 2.10), but conversely of the curve described previously, they does not exhibit the same pattern below amplitude value 1 in the frequency range 1-8 Hz, even though the amplitudes are not very high (generally about a value 2).

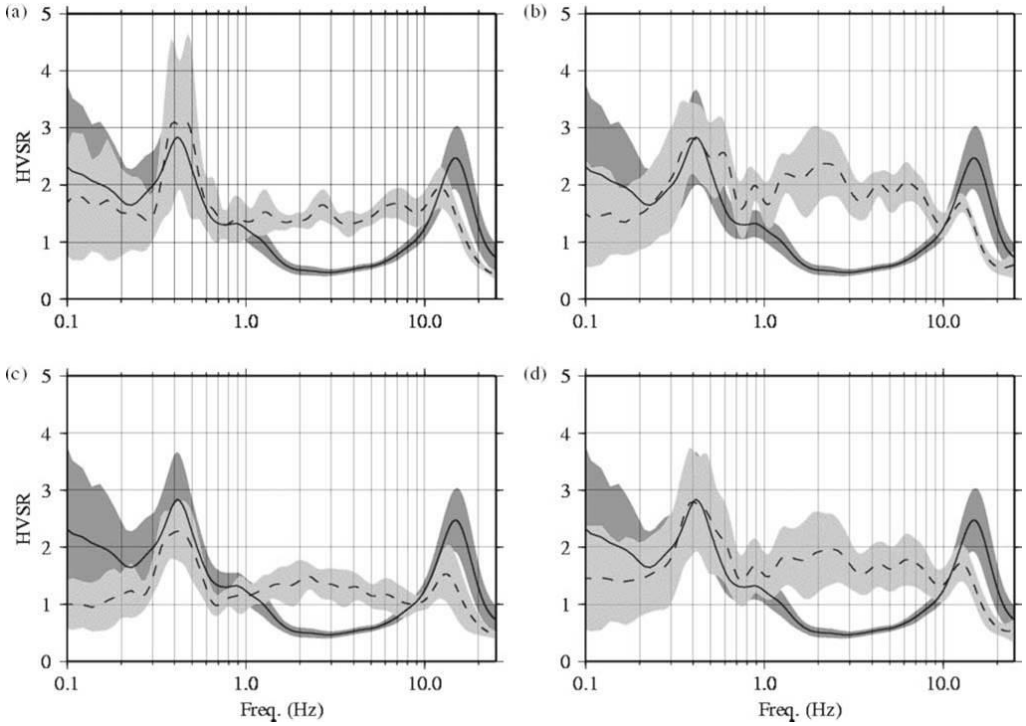

Fig. 2.10 - Comparison between the mean of the H/V spectral ratio of earthquakes (dashed lines) and ambient vibrations (solid lines). Dark-gray area represents ± 1 S.D. of the seismic noise HVSR. Light-gray area represents ± 1 S.D. of (a) the local earthquakes; (b) San Giuliano di Puglia 2002 sequence; (c) regional and teleseismic events; (d) all the analyzed earthquakes (from Di Giacomo *et al.*, 2005).

Very similar HVSR trends were found by the authors using a numerical modeling in a layered 1D medium (exploiting of an improved Thomson-Haskell propagator matrix method; Wang, 1999) both for ambient vibrations and for earthquakes (Fig. 2.11). In particular, they used the same parameter values in both cases (except for the shallow anthropic materials) inferred by a trial and error procedure and inserting a V_S inversion between the conglomerates and the clay (Fig. 2.11).

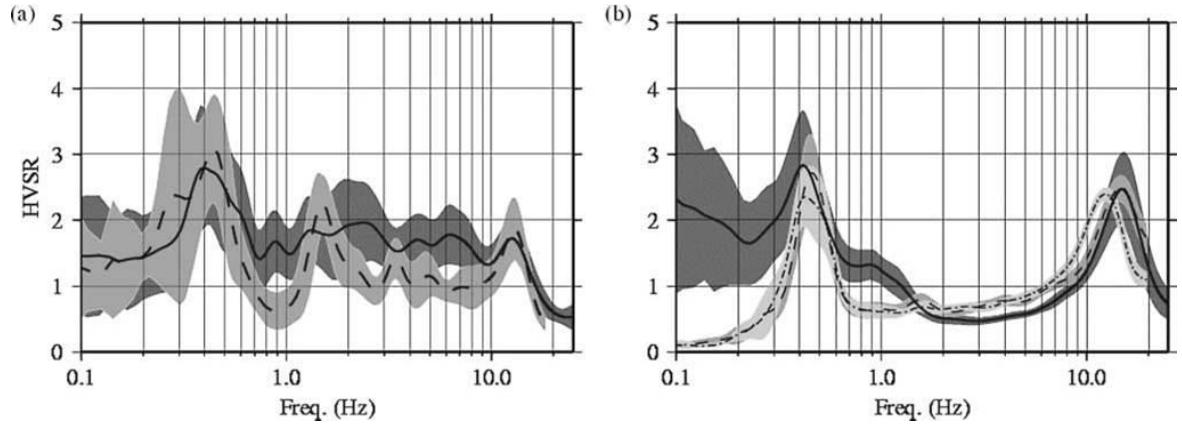

Fig. 2.11 - (a) Comparison between the mean HVSR spectral ratio of real earthquakes (solid line) and synthetic ones (dashed line). Dark-gray area represents ± 1 S.D. of the real earthquakes. Light-gray area represents ± 1 S.D. of the synthetic earthquakes. (b) Comparison between the mean HVSR spectral ratio of ambient vibrations (solid line), and synthetic ones (dashed line and dashed-dotted line) with different values for the anthropogenic fill. Dark-gray area represents ± 1 S.D. of the seismic noise HVSR. Light-gray areas represent ± 1 S.D. of the two HVSRs of the synthetic seismic noise (from Di Giacomo *et al.*, 2005).

As concerns the amplitude differences between the HVSR ambient noise and the earthquake HVSRs in the frequency range 1-8 Hz, the authors suppose that they could stem from the different composition and propagation of the seismic waves: in particular, the velocity inversion might determine the incidence of the seismic waves at the site, and the source position. In fact, because the sources of seismic noise are located at the surface (i.e., above the velocity inversion), the propagation is mainly lateral and the wavefield is dominated by surface waves. In contrast, seismic waves from earthquakes propagate nearly vertically through the layers underneath the site because these sources are located at 8–10 km depth.

From the above described work, it is possible to deduce that trends in the HVSR curve characterized by spectral ratio values lower than 1 could be a mark to identify the presence of an S-wave velocity inversion in the subsoil. A work more focused about this feature, both using experimental data and numerical modeling, it was realized by Castellaro and Mulargia (2009). In order to better discriminate the H/V phenomenology under shear-wave velocity inversions, the authors analyzed several experimental cases where the V_S values increase monotonically with depth and vice versa.

In the first case, as described in the paragraph 2.3, the H/V peak is due to a relative detachment of the horizontal components from the vertical one: this detachment has an “eye-shape” structure in the single spectra plot (Fig. 2.7) and it is always the clear signature of an H/V peak with stratigraphic origin. Moreover, the authors observe that in the absence of V_S inversions with depth, the vertical spectral component lays at the same amplitude level or below that of the horizontal ones, except maxima (usually with a narrow bandwidth, as showed in Fig. 2.7 for the frequency range 6-7 Hz) due to the ellipticity minimum of the fundamental mode Rayleigh wave occurring at $2f_0$, where f_0 is the frequency of the ellipticity maximum (Fäh *et al.*, 2001). Therefore,

in the absence of velocity inversions, the H/V curve has an amplitude equal to or higher than 1 (H/V ≥ 1) in the entire frequency domain except around $2f_0$ (as showed in Fig. 2.5).
In the cases of velocity inversion, the authors analyzed several different situations (inversions due to natural stratigraphy, cavities and thin artificial soils like concrete and pavements) where the subsoil characteristics are known by boreholes and geophysical information. In all these cases, a key feature distinguishes these spectral ratios from those obtained on positive velocity gradients: the H/V curve remains below 1 for a large frequency interval (Fig. 2.12 and Fig. 2.13). On the single component spectra, this implies that the vertical component of motion has higher amplitude values than the horizontal ones for the same wide frequency interval (Fig. 2.12).

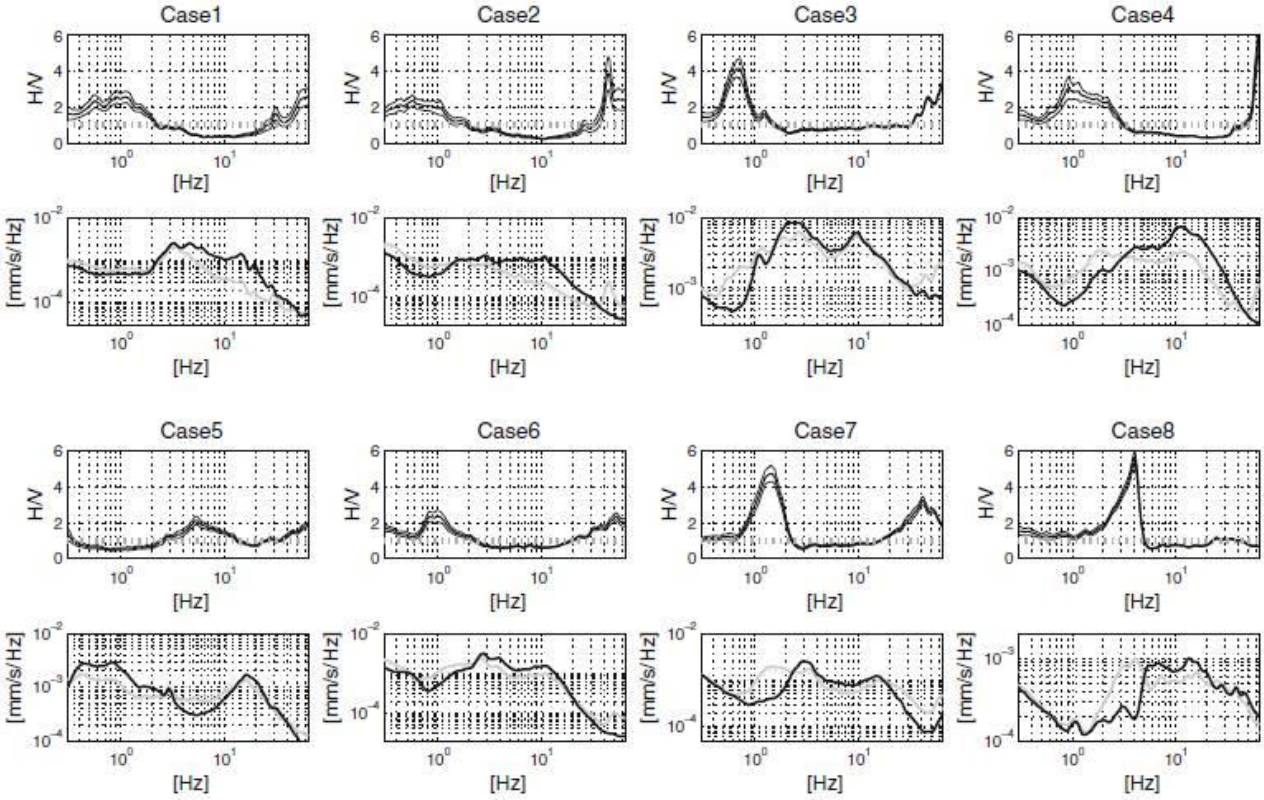

Fig. 2.12 - HVSR obtained velocity inversion cases due to natural stratigraphy. First and third row show average H/V curves with 2σ confidence intervals. Second and fourth row show average spectra: vertical component is represented in black, EW horizontal component is the gray solid line, NS horizontal component is the gray dotted line (from Castellaro and Mulargia, 2009).

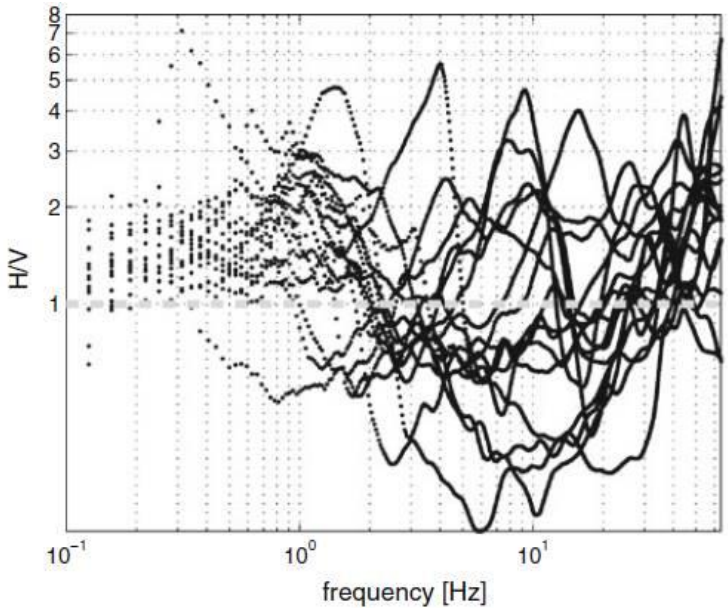

Fig. 2.13 - Summary of the H/V curves of Fig. 2.12 and 8 more cases. It is possible to note the H/V < 1 wide ranges (from Castellaro and Mulargia, 2009).

After that, Castellaro and Mulargia focus on the effect generated by shallow velocity inversions due by stiff artificial soils in the HVSR curve.

In order to parametrize this effect and to observe the single spectra behavior, the authors used numerical models to produce synthetic HVSRs. In particular, they adopt 1-D layered viscoelastic Haskell-Thomson solid in Herrmann formulation (2002) assuming that the ambient vibrations are composed of Rayleigh waves in the fundamental mode (Fäh *et al.*, 2001) and also evaluating the addition of Love waves (Bonnefoy-Claudet *et al.*, 2008). In the modeling section they do not consider the absolute H/V values but the ratio H/V / $(H/V)_0$, where H/V is measured on stiff artificial soil and $(H/V)_0$ on natural soil; hence, the spectral ratio is defined as P/P_0, where P is the spectrum of the horizontal or vertical component on stiff artificial soil and P_0 on natural soil. Considering a simple model (a single soft layer above the half-space) and adding a superficial stiffer layer of thickness varying from 0.1 to 1 m (Fig. 2.14-A), they note that the H/V curve is lowered in amplitude from 40 to about 1 Hz (peak included, Fig. 2.14-B), the horizontal Rayleigh and Love components are decreased in the entire frequency domain (Fig. 2.14-D,F) and that the vertical component increases from about 3 to 20 Hz (Fig. 2.14-E). This synthetic example shows that even a very thin stiff artificial soil can strongly affect the H/V curve down to low frequencies.

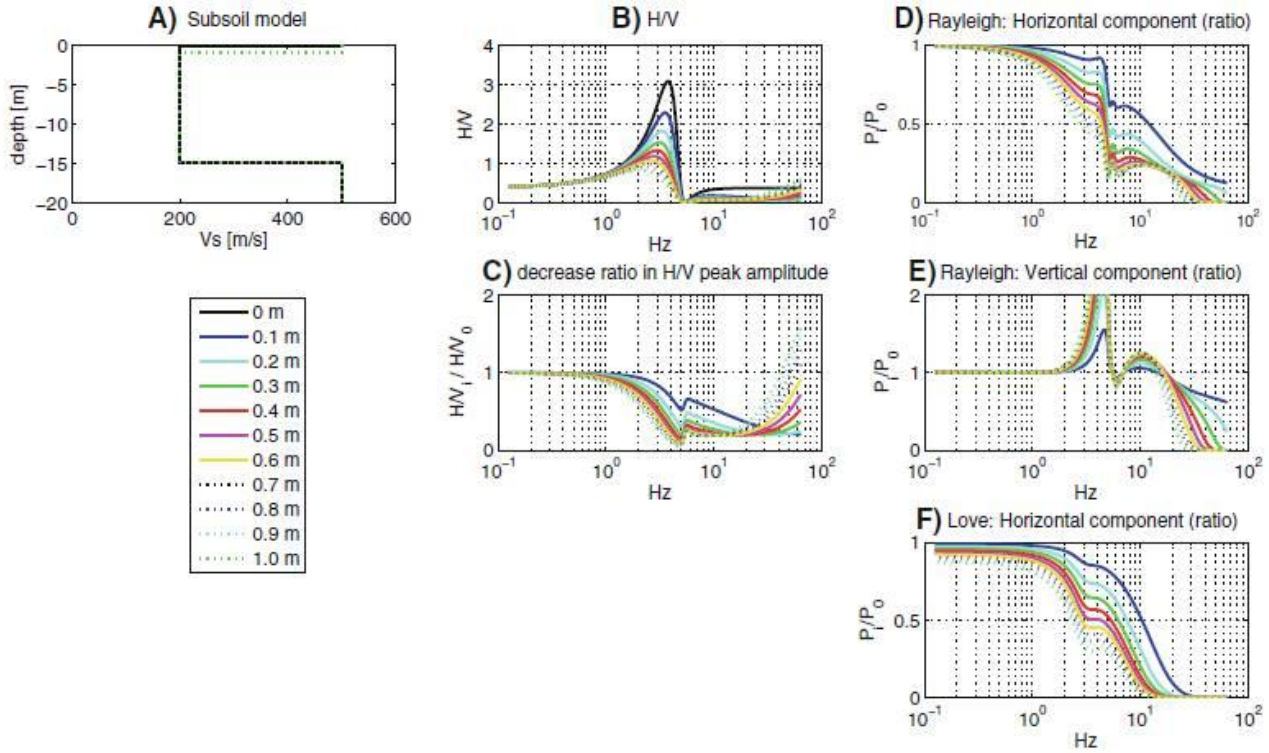

Fig. 2.14 - Results from modelling. A) subsoil model (black solid line: 15 m soft layer + bedrock, blue solid line: 0.1 m stiff artificial layer, 14.9 m soft layer + bedrock, cyan solid line: 0.2 m stiff artificial layer, 14.8 m soft layer + bedrock etc.), B) modelled Rayleigh H/V ratio for each subsoil model, C) Rayleigh H/V depression ratio = $(H/V)_i$ / $(H/V)_0$, where $(H/V)_i$ = H/V ratio in presence of the i-th stiff artificial layer and $(H/V)_0$ = H/V ratio on natural soil, D) P_i/P_0, where P_i = horizontal spectrum of the horizontal Rayleigh wave in presence of the i-th stiff artificial layer and P_0 = horizontal Rayleigh wave on natural soil; E) as D) for the vertical Rayleigh wave, F) as D) for the horizontal Love wave (from Castellaro and Mulargia, 2009).

Moreover, several experimental cases are analyzed by the authors to verify additional effects due the presence of still artificial soils. Fig. 2.15 shows an example: they performed several ambient vibration recordings in a very narrow area (20 x 1 m), some of them on a stiff artificial soil (composed of cement layer about 30 cm thick) and the other ones on natural soil. The site stratigraphy is well known from borehole information and is characterized by a shallow impedance contrast (about 5 m depth) and one at more higher depth. It is possible to note that the HVSRs on natural soil show clear peaks at 11–17 Hz (Fig. 2.15-C, thin lines): these maxima totally disappear on the H/V curve obtained on the artificial stiff soil (Fig. 2.15-C, thick lines) though the subsoil stratigraphy remains identical excluding the first 30 cm of cement (Fig. 2.15-A,B). Similar H/V patterns were obtained by numerical modeling (50% Rayleigh and 50% Love waves in the fundamental mode) for the two stratigraphic configurations (Fig. 2.15-D). Observing the single component spectra, an eye-shaped pattern with a local minimum of the vertical component is present on natural soil in correspondence of the high frequency peak (black arrow in Fig. 2.15-E), while on artificial soil the horizontal components are clearly lowered. An interesting feature is that the vertical component appears to be much less affected and still shows a clear local minimum in

correspondence of the expected H/V peak (black arrow in Fig. 2.15-F): this suggests that even in the presence of velocity inversions which force the H/V curve below 1, the signature of the "hidden" layer can still be recognized in the single spectra.

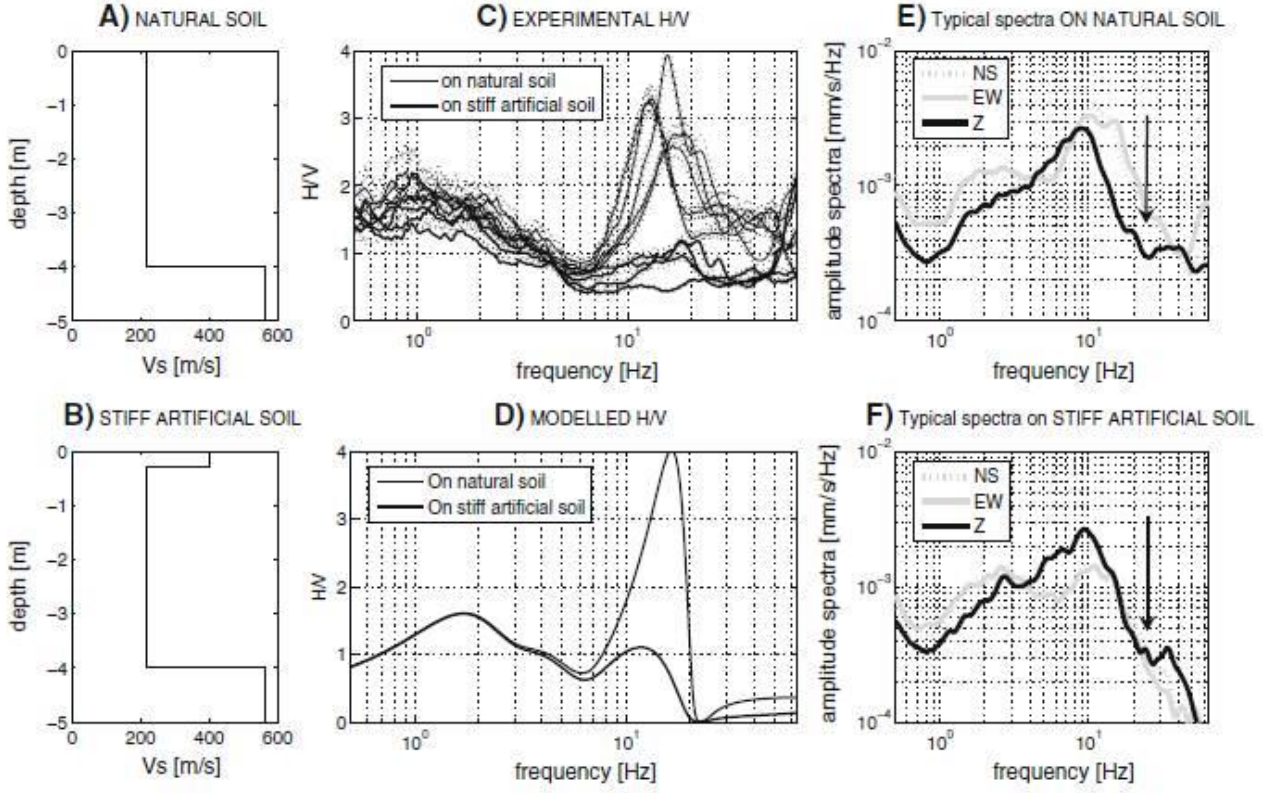

Fig. 2.15 - A) stratigraphy on natural soil; B) stratigraphy on stiff artificial soil; C) experimental H/V curves on natural (thin lines) and on stiff artificial (thick lines) soil, D) modelled H/V for the case on natural soil A) (thin line) and on stiff artificial soil B) (thick line); E) typical spectra on natural soil; F) typical spectra on stiff artificial soil. The arrow indicates the "eye-shaped" spectral pattern responsible for the H/V peak on natural soil (panel E) and the same pattern on stiff artificial soil (panel F) which, due to the H/V amplitude reversal, cannot result in a H/V larger than 1 (from Castellaro and Mulargia, 2009).

From this study, it is clear that the phenomenological signature of S-wave velocity inversions in the H/V curves consists of a spectral ratio below 1 for a wide range of frequencies, due to the decrease of the horizontal spectral components (compared to the case of no velocity inversion) and to a possible rise of the vertical component (this latter component is generally less affected). The H/V < 1 pattern can have different origins (e.g., Rayleigh wave ellipticity, strong transients on the vertical components, etc.) but in these cases it is usually present over a restricted frequency range.

Moreover, this effect can be observed not only in the presence of natural velocity inversions (e.g., stiff gravel banks overlying soft clays or cavities) but also on very thin stiff artificial soils (e.g., asphalt, etc.) above sediments: in this latter case, the decay of the horizontal components can affect dramatically the resulting HVSR curve hiding also strong peaks, which can be rescued by analyzing the single spectra.

CHAPTER 3
Seismic array techniques

3.1 Introduction

As described in Chapter 1, although the ambient vibrations are characterized by stochastic and irregular behaviour, they show significant features of statistical stationarity such as to allow their application for geophysical exploration purposes.
In the last decades, many researchers focused their attention on the analysis of the surface wave component of the ambient vibration wavefield. The identification of these seismic phases is due to two principal reasons:

1. the surface waves are characterized by a lower geometric attenuation and higher energy content respect to the body wave component of the ambient vibration wavefield;
2. from a statistical point of view, the surface waves show characteristics of persistence which bring out the signal coherence associated to them.

The methods based on the surface wave analysis are defined as seismic array or multi-station measurements. These methods, also used for active seismic sources (MASW technique, Park *et al.*, 1999), consist in the placing of a definite number of synchronized seismic sensors in the investigated area according to suitable configurations.
As explained in the Introduction, the aim of seismic array methods is the surface wave dispersion curve estimate, namely the function that relates the Rayleigh or Love wave (group or phase) velocity with the frequency (or period). Through suitable inversion techniques, this curve can provide significant information about the subsoil mechanical properties.
Usually, the commonly available equipment allows the recording of the vertical component of ambient vibrations only: for this reason, this dissertation and a large part of works are focused on studying the propagation of Rayleigh waves. Considering that Rayleigh waves sample portions of the subsoil proportional to their wavelengths and that their phase velocities are strongly conditioned by the S-wave velocities of the sampled layers, the Rayleigh wave phase dispersion curve is used to infer information on the subsoil structure, and in particular on the local S-wave velocity profile.
Furthermore, the study of Rayleigh wave propagation represents a necessary prerequisite to infer information about Love wave propagation. In fact, although the horizontal component of ground motion is the only source of information about Love wave propagation, to estimate the Love wave dispersion curve three-component sensor recordings are required.

In the seismic array measurements the ambient vibrations are acquired concurrently in several points on the ground and for a suitable time length with a dual purpose:

- to strengthen the informative content than to the spurious signals, thus allowing a more accurate investigative capacity;
- to study the characteristics of the seismic waves propagation: velocity and direction.

The spatial configuration of the sensors depends on the analysis method to be used (and/or vice versa). In the ambient vibration study, the seismic arrays can be deployed using two geometries:

1. linear geometry: the receivers are deployed according to a specific direction with an equal or irregular spacing (e.g., ReMi technique; Louie, 2001);
2. 2-D geometry: the receivers are deployed on an area according to a suitable configuration with an equal or irregular spacing. The sensor configuration can be characterized by circular, linear (e.g., L or cross shape configuration) and random (without any preferential direction) geometries.

The studies carried out in this dissertation are based on 2-D geometry seismic arrays. Therefore in the following paragraph the commonly used 2-D seismic array techniques will be treated.

3.2 2-D seismic array techniques

2-D seismic arrays were originally proposed at the beginning of the 1960s as a new type of seismological tool for the detection and identification of nuclear explosion (e.g., Bormann, 1966; Frosch and Green, 1966). Since then, seismic arrays have been applied at various scales for many geophysical purposes. At the seismological scale, they were used to obtain refined velocity models of the Earth's interior (e.g., Birtill and Whiteway, 1965; Kværna, 1989; Rost and Thomas, 2002). At smaller scales, seismic arrays have been used for the characterization of surface wave propagation, and the extraction of information on the shallow subsoil structure (i.e., the estimation of the local S-wave velocity profile). Especially in the last decades, due to the focus of seismologists and engineers on estimating the amplification of earthquake ground motion as a function of local geology, the analysis of ambient vibrations recorded by 2D arrays offered the opportunity to characterize a site down to great depths (hundred meters) avoiding the need of large and heavy active sources. Due to this significant characteristic, the interest in analyzing seismic noise recorded by arrays has hugely grown (e.g., Horike, 1985; Tokimatsu, 1997; Ohori *et al.*, 2002; Okada, 2003; Scherbaum *et al.*, 2003; Parolai *et al.*, 2005; Asten, 2006; Parolai *et al.*, 2006; Köhler *et al.*, 2007).

Regardless the used technique, the outcomes of the 2-D seismic array analysis also depend on how the ambient vibration wavefield is spatially sampled by the seismic sensors configuration. Although several geometries exist, it is possible to give some general rules:

- the maximum distance between the sensors must be similar to the desired investigation depth;
- the minimum distance should be at least half of the minimum wavelength of interest, also taking into account the sampling frequency and the expected velocity values;
- the number of sensors must be enough to provide a good sampling of the distances between the minimum and the maximum one, thus avoiding spatial undersampling problems;

- the orientation of all the pairs of sensors should allow an azimuthal coverage as wide as possible.

In the following sub-sections, a detailed overview of the 2-D seismic arrays techniques used in this dissertation is done. In particular, two main methodologies are treated:

- ✓ the methods based on the Spatial AutoCorrelation analysis;
- ✓ the methods based on the frequency-wavenumber (*f-k*) analysis.

Finally, a short review concerning the analysis based on cross-correlation in time domain is done, with a particular focus on its application in small scale studies.

3.2.1 Methods based on the Spatial AutoCorrelation analysis

3.2.1.1 SPAC method

The SPatial AutoCorrelation (SPAC) method was proposed by Aki (1957, 1965). This author analysed the ambient vibration recordings to estimate the Rayleigh wave phase velocity using a statistical approach: in particular, considering the ambient vibration wavefield as a stochastic and stationary process in time and space, he assumed that this signal represents the sum of waves propagating without attenuation in a horizontal plane in different directions with different powers, but with the same phase velocity for a given frequency. Moreover, he supposed that waves with different propagation directions and different frequencies are statistically independent.

The spatial correlation function between two receivers can be defined as:

$$\phi(r,\varphi) = \langle u(x,y,t)u(x+r\cos\varphi, y+r\sin\varphi\,, t)\rangle, \tag{3.1}$$

where $u(x,y,t)$ and $u(x+r\cos\varphi, y+r\sin\varphi\,, t)$ is the signal measured at the first receiver (with coordinates (x,y)) and at the second receiver (with coordinates $(x+r\cos\varphi, y+r\sin\varphi)$) at time *t*; *r* is the inter-station distance, φ is the azimuth of the two receivers measured from the direction of the *x* axis and the angle brackets $\langle\ \rangle$ denote the temporal averaging.

Using sensors disposed on a circumference of radius *r* and regularly spaced among themselves, it is possible to compute the azimuthal average of the spatial correlation function $\bar{\phi}(r)$ between each of them and a sensor placed in the circumference center. This function is defined as:

$$\bar{\phi}(r) = \frac{1}{\pi}\int_0^{\pi} \phi(r,\varphi)d\varphi\,. \tag{3.2}$$

For single-mode scalar wave with Rayleigh phase velocity $c(\omega)$, the azimuthally averaged spatial correlation function $\bar{\phi}(r)$is related to the temporal power spectrum $\Phi(\omega)$ by:

$$\bar{\phi}(r) = \frac{1}{\pi}\int_0^{\infty} \Phi(\omega) J_0\left[\frac{\omega}{c(\omega)} r\right] d\omega, \tag{3.3}$$

where J_0 is the Bessel function of zeroth order. Filtering the waves through a narrow-band filter centered at the frequency ω_0, the power spectrum can be expressed in the form:

$$\Phi(\omega) = P(\omega_0)\delta(\omega - \omega_0), \quad \omega > 0, \tag{3.4}$$

where $P(\omega_0)$ is the spectral power density of the signal at the frequency ω_0 and $\delta(\omega)$ is the Dirac delta function. Inserting this expression into Eq. 3.3, the azimuthally averaged spatial correlation function is obtained as:

$$\bar{\phi}(r, \omega_0) = \frac{1}{\pi}\, \mathrm{P}(\omega_0) J_0\left[\frac{\omega_0}{c(\omega_0)} r\right]. \tag{3.5}$$

Defining the normalized spatial correlation coefficient $\rho(r, \varphi, \omega_0)$ in this form

$$\rho(r, \varphi, \omega_0) = \frac{\phi(r,\varphi,\omega_0)}{\phi(0,\varphi,\omega_0)}, \tag{3.6}$$

the Eq. 3.5 becomes

$$\bar{\rho}(r, \omega_0) = J_0\left[\frac{\omega_0}{c(\omega_0)} r\right]. \tag{3.7}$$

This equation shows that the phase velocity $c(\omega_0)$ of the waves can be derived by fitting the experimental values of $\bar{\rho}(r, \omega_0)$ (calculated for a series of receivers set in a circular pattern around a central reference receiver) with the theoretical ones computed by the Bessel function J_0 at different tentative combinations of parameters *r*, ω_0 and $c(\omega_0)$ (Fig. 3.1).
To perform this, Aki (1957) proposed the use of a least-squares fitting method. In particular, since *r* in a circular array is constant and $\omega_0 = 2\pi f$, the coefficient $A = 2\pi r_0$ is a constant value. Therefore, the Eq. 3.7 can be rewritten as:

$$\bar{\rho}(r_0, f) = J_0\left[\frac{Af}{c(f)}\right] = J_0(x), \tag{3.8}$$

where $x = \frac{Af}{C(f)}$ and $c(f)$ are as a function of the frequency *f*. Due to the simplicity of the Eq. 3.8, it is possible to infer that the optimum Bessel function $J_0(x)$ can be obtained from a set of observed values of $\bar{\rho}(r_0, f)$ and f by the use of the least-squares fitting method (Okada, 2003).

In practice, the estimation of the Rayleigh wave phase velocity *c* is carried out for each frequency *f* by

$$C(f) = \frac{Af}{x}. \tag{3.9}$$

This generates a solution for the argument *x* of the Bessel function which is equal to the experimental spatial correlation coefficient.

The Rayleigh wave dispersion curve obtained according to this approach is relevant only in a narrow band of frequencies that depends on the radius *r* of the adopted circular array. Therefore, in order to extend the frequency domain of the dispersion curve it is necessary to repeat the ambient vibrations acquisition with different radius arrays (Fig. 3.1).

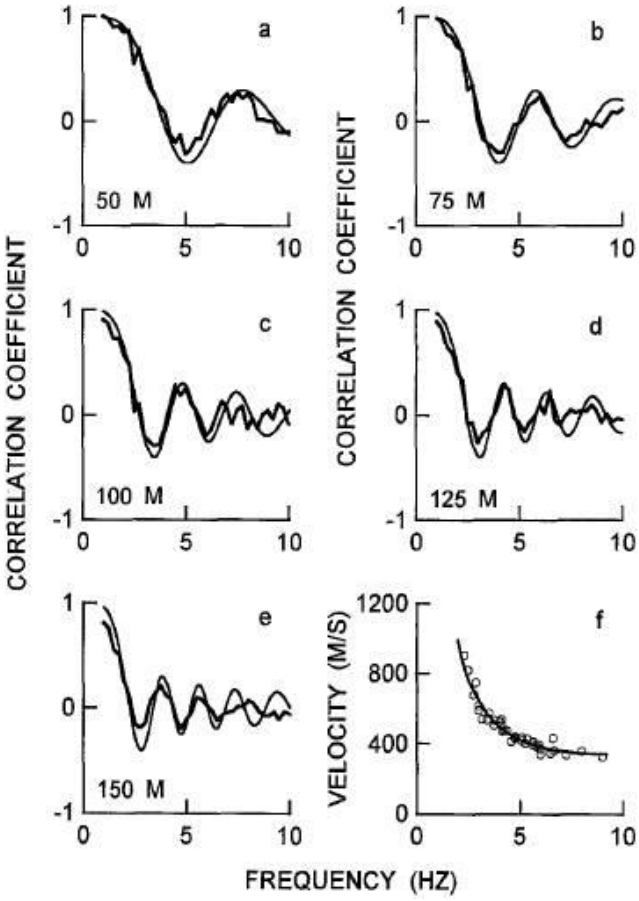

Fig. 3.1 - Azimuthal averages of the spatial correlation coefficients (bold lines) obtained as a function of bandpass-filter frequency for semi-circular arrays (deployed in the same location) with radius (a) r = 50 m, (b) r = 75 m, (c) r = 100 m, (d) r = 125 m, and (e) r = 150 m. The smooth thin lines superimposed on the data represent the theoretical spatial correlation coefficients derived from Eq. 3.7 taking into account the Rayleigh phase velocity values (open circles) showed in the dispersion curve in (f) (from Chouet *et al.*, 1998).

3.2.1.2 ESAC method

As described previously, the maximum length of the array must be similar to the desired investigation depth. When this depth value is in the order of hundred meters, could be not easy to deploy a seismic array with circular or semi-circular geometry, especially in urban areas. To overcome this problem, Ling and Okada (1993) and after Ohori *et al*. (2002), proposed a modified version of the SPAC approach, named Extended Spatial AutoCorrelation (ESAC) method. Unlike SPAC technique, this approach is suitable for any kind of array geometry both linear (Fig. 3.2-A) and random: this characteristic makes the ESAC method a very attractive tool to estimate the Rayleigh wave dispersion curve in the framework of Seismic Microzonation projects.
In this technique, utilized in the studies carried out for this dissertation, the spatial correlation function for each couple of sensors can be computed in the frequency domain in this way (Malagnini *et al.*, 1993; Ohori *et al.*, 2002, Foti *et al.*, 2011):

$$\rho_{jn}(\omega) = \frac{\frac{1}{M}\sum_{m=1}^{M} Re\left({}_mS_{jn}(\omega)\right)}{\sqrt{\frac{1}{M}\sum_{m=1}^{M} {}_mS_{jj}(\omega)\sum_{m=1}^{M} {}_mS_{nn}(\omega)}} = Re\left[Coh_{jn}(\omega_0)\right], \tag{3.10}$$

where ${}_mS_{jn}$ is the cross-spectrum for the m^{th} segment of data between the j^{th} and the n^{th} sensor and *Re* indicates its real part; ${}_mS_{jj}$ and ${}_mS_{nn}$ are the power spectra (or autospectra) of the m^{th} segments at sensor *j* and *n* respectively. *M* is the total number of used segments. The quantity on the right of the above equation is also defined as coherence matrix (Coh_{jn}).
Therefore, Eq. 3.7 can be re-written as:

$$\rho\left(r_{jn}, \omega_0\right) = J_0\left[\frac{\omega_0}{c(\omega_0)} r_{jn}\right], \tag{3.11}$$

where r_{jn} is the distance between the station *j* and *n*.
Unlike the SPAC method, defined as frequency approach, in the ESAC method the Rayleigh phase velocity $c(\omega_0)$ is computed from Eq. 3.11 in the distance domain. In particular, for a fixed frequency value, the spatial correlation function is computed for each inter-station distances of the array and the phase velocity is obtained using Eq. 3.11 in order to best fitting the observed pattern of $\rho_{jn}(\omega)$ as a function of the inter-station distance (Fig. 3.2-B).
At each interested frequency ω_n, the optimal value of $c(\omega_n)$ is determined through an iterative grid-search procedure (Parolai *et al.*, 2006), where the tentative phase velocity is generally varied over large intervals (e.g., between 100 and 3000 m/s) in small steps (e.g. 1 m/s). The best-fit between the experimental spatial correlation values $\rho_{exp}(r_i, \omega_n)$ and the theoretical ones is obtained for the $c(\omega_n)$ value that minimizes the following misfit function (Baliva, 2006):

$$Mis(\omega_n) = \sqrt{\frac{\sum_{i=1}^{N}\left[\rho_{exp}(r_i,\omega_n) - J_0\left(\frac{\omega_n}{c(\omega_n)}r_i\right)\right]^2}{N}} = \sqrt{\frac{\sum_{i=1}^{N}\left[Coh_i(\omega_n) - J_0\left(\frac{\omega_n}{c(\omega_n)}r_i\right)\right]^2}{N}}, \qquad (3.12)$$

where *N* is the total number of the stations and r_i is the inter-station distance.
Experimental spatial correlation values which differ by more than two standard deviations from the theoretical values obtained with the minimum misfit velocity can be removed before the next iteration of the grid-search.

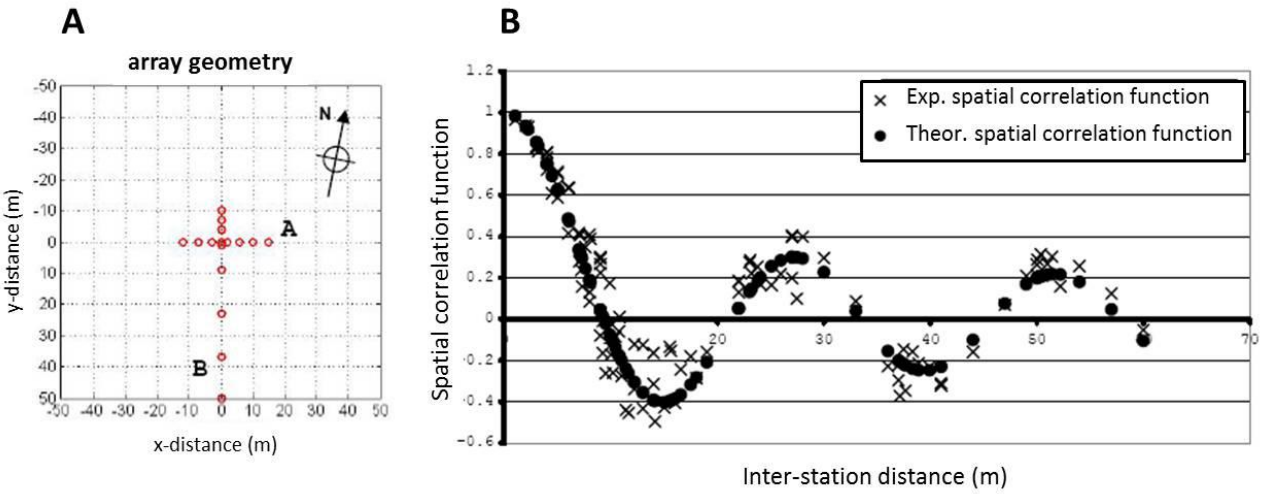

Fig. 3.2 - Panel A shows a linear array geometry that can be exploited using the ESAC method; red circles represent the position of the seismic sensors. Panel B shows the distribution of the spatial correlation function values as a function of the distance between the sensors at a fixed frequency value. The crosses denote the experimental values, the black circles are the theoretical ones.

Considering again the Eq. 3.10, it is possible to note that since the function $\rho_{jn}(\omega)$ is related with the cross-spectrum, this expression points out the proportional relationship between the spatial correlation function and the phase difference between the signals recorded at station *j* and *n*. Therefore, the accuracy of the ESAC method in determining the Rayleigh wave phase velocity ultimately depends on the estimate of the phase differences between the station acquisitions. Zhang *et al.* (2004) show that the uncertainty $\delta c(\omega_0)$ on the phase velocity estimation from Eq. 3.11 is:

$$\delta c(\omega_0) \approx \frac{1}{\tau}\frac{c(\omega_0)^2}{\Delta r}, \qquad (3.13)$$

where τ is the sampling frequency of the signal and Δr is the inter-station distance. Therefore, to equal sampling frequency, in order to ensure a good accuracy of the measure for a defined velocity range, it is needed to adopt an appropriate distance between the sensors. More the

expected velocity values are high, more the inter-station distance able to determine them is greater.

As mentioned previously, the final outcome of this procedure is the dispersion curve estimate. If it is assumed that the ambient vibration wavefield is dominated by surface waves, the measurements of the ground motion vertical component allow to estimate the Rayleigh wave dispersion curve related to the fundamental mode. In the presence of higher modes, the Eq. 3.11 should be modified in this form:

$$\rho\left(r_{jn},\omega_0\right) = \sum_{n=1}^{M} \frac{P_n(\omega_0)}{P(\omega_0)} J_0\left[\frac{\omega_0}{c_n(\omega_0)} r_{jn}\right], \tag{3.14}$$

where $c_n(\omega_0)$ is the phase velocity of the n^{th} mode that at the frequency ω_0 is characterized by a spectral power $P_n(\omega_0)$; $P(\omega_0)$ is the overall spectral power at the same frequency. Considering the simple case where a planar wave propagates along the direction connecting the j^{th} and the n^{th} sensor (Arai and Tokimatsu, 2004), the Eq. 3.11 becomes

$$\rho\left(r_{jn},\omega_0\right) = cos\left[\frac{\omega_0}{c(\omega_0)} r_{jn}\right]. \tag{3.15}$$

Juxtaposing the Eq. 3.14 and 3.15, it is possible to obtain an "effective" phase velocity $c_e(\omega_0)$ from this relationship

$$c_e\left(r_{jn},\omega_0\right) = \frac{\omega_0 r_{jn}}{\cos^{-1}\left[\sum_{n=1}^{M} \frac{P_n(\omega_0)}{P(\omega_0)} J_0\left[\frac{\omega_0}{c_n(\omega_0)} r_{jn}\right]\right]}, \tag{3.16}$$

which represents a sort of mean of the phase velocities associated to the different modes, weighted to the correspondent spectral powers. This means that the ESAC method (and SPAC as well) allows to estimate the Rayleigh wave effective phase velocity that corresponds to the phase velocity of the fundamental mode only if the higher modes are characterized by negligible spectral powers at the interested frequencies.

3.2.1.3 SPAC zero crossing method

This method constitutes one of the possible ways to retrieve the phase velocity values from the spatial correlation spectrum and it is generally adopted in circular arrays analyzed by using the SPAC method (e.g., Metaxian *et al.*, 1997; Chouet *et al.*, 1998). An interesting application of this approach was proposed by Ekstrom *et al*. (2009) with the aim to achieve a dispersion curve estimate between each pair of sensors belonging to a large scale array with random spatial

geometry. The final purpose of this study was the accomplishment of a Rayleigh wave phase velocity map of western United States by using the recordings of the ambient vibration vertical component acquired on more than 600 stations of the USArray during 2006-2008.

Ekstrom *at al.* (2009) justify the application of the SPAC formulation considering separately each station pairs relying on the Part 3 of Aki (1957) work: here the author argues that the azimuthally averaged spatial correlation spectrum $\bar{\rho}(r, \omega_0)$ (Eq. 3.7) can be replaced in the analysis by the spectrum obtained for a single station pair. Moreover, Aki's (1957) states that this assumption can be exploited only if that the stochastic ambient vibration wavefield is sufficiently isotropic.

Fig. 3.3 shows the spatial correlation spectra for the two station pairs belonging USArray, located in Washington State and separated by 282 km (Fig. 3.3-A) and 145 km (Fig. 3.3-B), respectively. The real parts of the spectra (dark blue lines) resemble a Bessel function in their oscillatory character, but the amplitudes of the peaks do not decrease monotonically with frequency as for J_0. This amplitude behavior depends on the background noise and non-linear effects of the data processing as well as the azimuthally non-uniform power of the ambient vibration witnessed by the non-vanishing imaginary part (Fig. 3.3) of the spatial correlation spectrum (Cox, 1973).

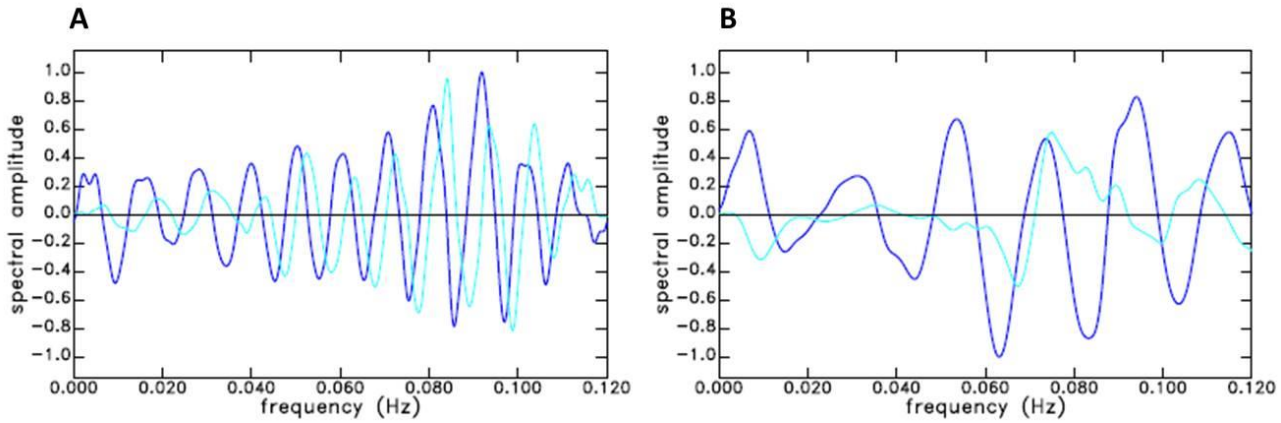

Fig. 3.3 - Spatial correlation spectra for two station pairs belonging to USArray; the stations are separated by 282 km (A) and 145 Km (B). Dark blue lines show the real parts of the spectra and light blue lines show the imaginary parts (from Ekstrom *et al.*, 2009).

Therefore, it is evident that dispersion information cannot readily be deciphered from the detailed shape of the spectrum. In view of this, the authors decided to use the locations of the zero crossings as the dispersion observables, assuming that these points in the spectrum should be insensitive to variations in the spectral power of the background noise. In particular, if ω_n denotes the frequency of the n^{th} observed zero crossing and z_n denotes the n^{th} zero of J_0 (namely the argument value that makes null the J_0 value), it is possible to determine the corresponding phase velocity $c(\omega_n)$ as

$$c(\omega_n) = \frac{\omega_n r}{z_n}, \tag{3.17}$$

following Aki's (1957) approach. In observed spatial correlation spectra, association of a given zero with a particular zero crossing of J_0 may be difficult because background noise in the spectrum can cause missed or extra zero crossings. To overcome this problem, the authors developed a set of phase velocity estimates $c_m(\omega_n)$, where *m* takes the values 0, ±1, ±2, etc., indicating the number of missed or extra zero crossings. Following the Eq. 3.17, this relationship can be expressed as

$$c_m(\omega_n) = \frac{\omega_n r}{z_n + 2m}. \tag{3.18}$$

Fig. 3.4 displays the dispersion curves obtained by the spectra showed in Fig. 3.3. The connection of the $c_m(\omega_n)$ values computed by Eq. 3.18 by considering the positive-to-negative zero crossings (downward triangles in Fig. 3.4) and of the ones computed by the negative-to-positive zero crossings (upward triangles in Fig. 3.4) generates two dispersion curves that can be assessed for consistency.

Moreover, from Fig. 3.4 it is evident that several dispersion curves are derived from the phase-velocity measurements resulting from different possible choices of *m* for each zero. At low frequencies, the authors discard all the curves that do not fall within a realistic range (dashed lines in Fig. 3.4), while at higher frequencies they do not consider the dispersion curves characterized by step-like velocity changes (namely evaluating the curve smoothness).

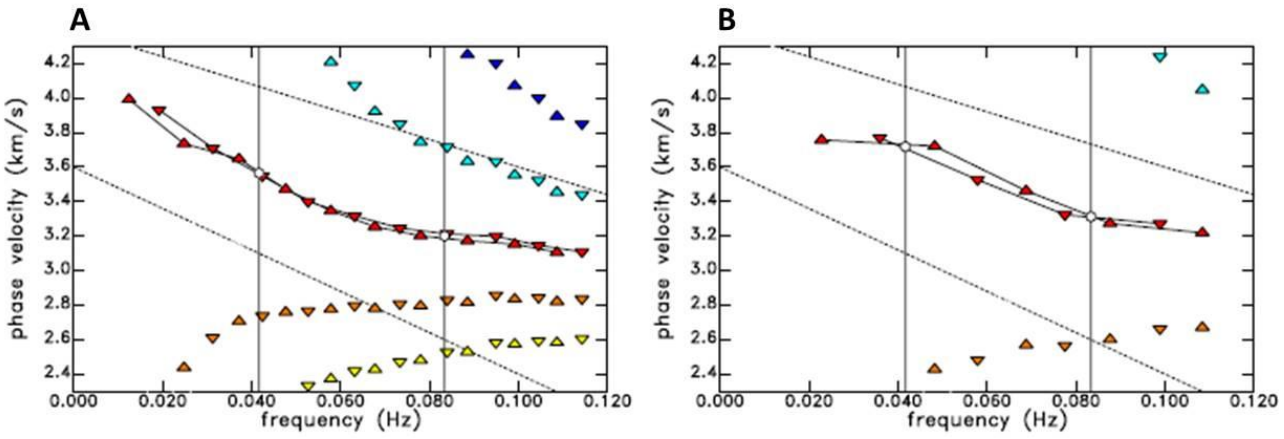

Fig. 3.4 - Dispersion diagrams showing the phase-velocity values derived from the zero crossings of the spectra in Fig. 3.3. Downward triangles show zero crossings from positive to negative, upward triangles show crossings from negative to positive. Red triangles correspond to m = 0, with no missing or extra zeros. Orange triangles correspond to m = 1; yellow to m = 2; light blue to m = -1; and dark blue to m = -2. Dashed lines outline the region used to define acceptable phase velocities at each frequency. Black lines connect upward- and downward-pointing triangles to define the dispersion curves that are sampled at discrete frequencies for the subsequent analysis. Thin vertical lines and white dots show the frequencies and phase velocities values taken into account to accomplish the tomographic Rayleigh wave velocity maps at 12 s and 24 s (from Ekstrom *et al.*, 2009).

After that, the remaining dispersion curve is evaluated by considering the two sub-dispersion curves (one connecting up-crossing zeros and one connecting down-crossing zeros) as a quality criterion: in particular, the dispersion curve is not included in the subsequent analysis if the phase velocity values of the two sub-dispersion curves at considered frequency differ than more a defined threshold.
In order to assess the quality of the estimated dispersion curves and to validate the above described approach, the authors collected the phase-velocity measurements at 12 s period (0.083 Hz) and used them in an inversion to obtain a Rayleigh phase velocity map for the Western part of the United States. The resulting map is characterized by consistent velocity values with the crustal geological settings of the investigated area; moreover, this map presents a strong similarities with that one proposed by Lin *et al*. (2008) analyzing the correlation of the ambient vibration recordings in time domain (*vide infra*).
Ekstrom (2014) used the same methodology and dataset with the aim to infer Rayleigh and Love phase velocity maps of Western and Central United States through least-square tomographic method.

Recently, Pilz *et al*. (2012) and Pilz *et al*. (2013) explored the use of the SPAC zero crossing method to retrieve the Rayleigh wave phase velocity values between each pair of sensors belonging to a small scale array, with maximum inter-station distance of the order of 100 m. The authors follow the procedure described above and proposed by Ekstrom *et al.* (2009) taking into account only the case where m=0 (no extra or missing zero crossing) in the Eq. 3.18; the final purpose was the accomplishment of a 3-D shear wave ambient vibration tomography at local scale. After a reliability validation of the method on a synthetic example with a known modeled subsurface structure, the authors retrieved a shear wave tomography at different frequencies at Nauen test site (Germany) and near the city of Bishkek (Kyrgyzstan). In particular, in the first case Pilz *et al*. (2012) obtain a subsoil imaging very similar at that one achieved exploiting the same dataset by Picozzi *et al*. (2009) using the ambient vibrations cross correlation in the time domain (*vide infra*).
These studies witness that the SPAC zero crossing method can be a useful tool to point out the Rayleigh wave phase velocity (and consequently the S-wave velocity) differences within of a local scale seismic array. Moreover, since the authors used a limited number of seismic stations (about 20) and recording times of several tens of minutes, the use of this technique could be effective in urban framework and therefore for Seismic Microzonation studies.

The reliability of the SPAC zero crossing method in retrieving the Rayleigh phase velocity (V_R) dispersion curve between two sensors was tested in the framework of this dissertation. This experiment was carried out in a test site located near the city of Rabat (Malta). This site is characterized by a simple 1-D geological setting, with outcropping soft clayey material overlying a stiffer limestone formation (the geology of the Maltese archipelago is described in detail in Chapter 5). The ambient vibrations were acquired for 20 minutes with two synchronized three-

directional sensors with a sampling frequency of 256 Hz and an acquisition time of 20 minutes. Three acquisitions with different inter-sensor distances *D* were performed: 60 m, 80 m and 110 m. The SPAC zero crossing analysis was carried out for the vertical component of the acquired signal considering in the Eq. 3.18 *m* values equal to -2, -1, 0, 1, 2. Fig. 3.5 shows the five resulting Rayleigh wave dispersion curves for the inter-station distance of 110 m: as showed in Fig. 3.4, the curves related with *m*=1 and *m*=2 are characterized by an anomalous trend (the V_R values decrease almost linearly towards the lower frequencies) and therefore they can be easily discarded. Also the curve with *m*=-2, despite it presents a normally dispersive trend, lets some doubts about its reliability, since V_R values of about 2000 m/s are not much credible for the subsoil materials in this site. Conversely, the dispersion curves with *m*=0 and *m*=-1 are characterized by compatible values with the geological features of the investigated area. The dispersion curve plots for the inter-station distances of 60 m and 80 m show very similar features but for more narrow frequency ranges.

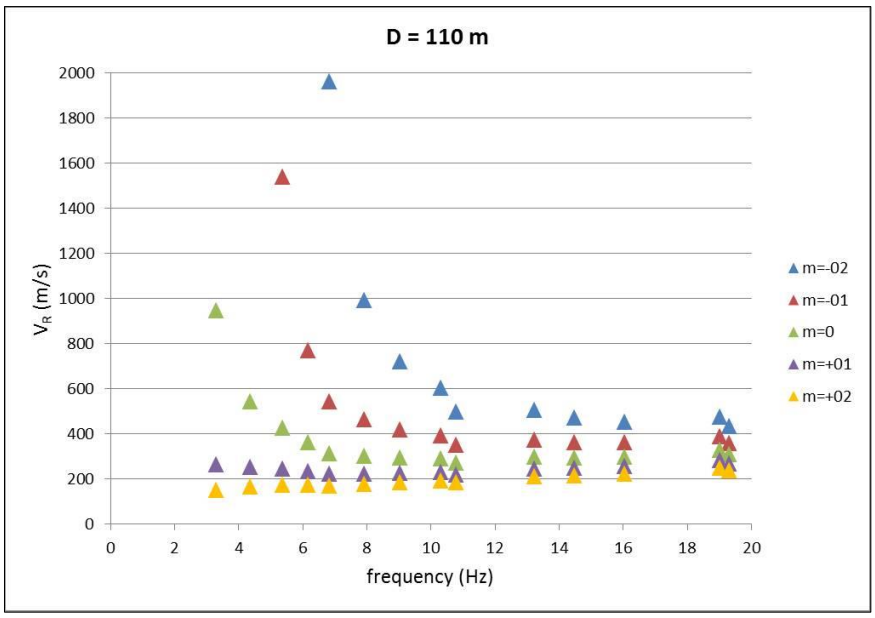

Fig. 3.5 - Rayleigh phase velocity curves obtained by the SPAC zero crossing method for a sensor pair separated by a distance of 110 m. The curves are obtained considering different *m* values in Eq. 3.18.

This test site was also investigated by using an L shape seismic array (Farrugia *et al.*, 2015) formed by 42 vertical geophones regularly spaced with distance of 5 m and maximum considered inter-station distance of about 150 m. The ambient vibrations were acquired for 20 minutes and the dispersion curve was obtained by the ESAC method.

Plotting this curve with the other ones achieved by SPAC zero crossing method (Fig. 3.6), it is possible to realize that, for each of the three considered inter-station distances, the "real" dispersion curve corresponds with that one obtained by considering *m*=0 (Fig. 3.6-A). Moreover, it

is evident that the SPAC zero crossing curves achieved for D=80 m and D=110 m cover a wider frequency range respect to the ESAC dispersion curve.

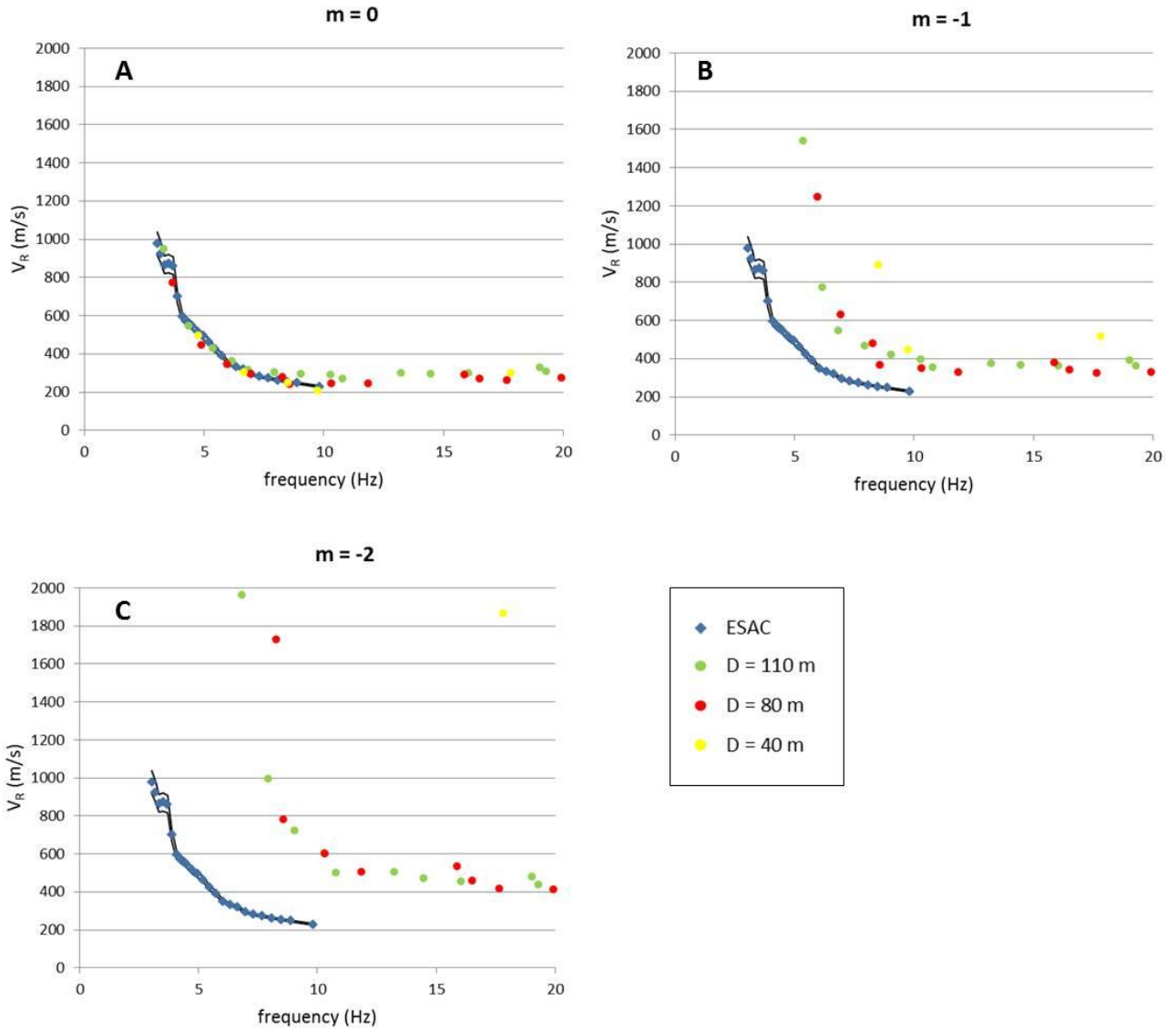

Fig. 3.6 - Rayleigh phase velocity curves obtained by the SPAC zero crossing method for sensor pairs separated by a distance of 110 m (green dots), 80 m (red dots) and 40 m (yellow dots). Panel A, B and C shows the curves with *m* values equal to 0, -1 and -2 respectively. Blue dots and the thin black lines represent the ESAC dispersion curve and its confidence interval.

The outcomes obtained by this small scale test confirm the capability of the SPAC zero crossing technique in retrieving a reliable Rayleigh phase velocity dispersion curve using only two sensors. On the other hand, this technique provide more than one reasonable curve and the identification of the "real" dispersion curve can not be trivial without the presence of other geophysical information. Moreover, in order to assess the lateral heterogeneities of the subsoil, the use as a comparison of a dispersion curve retrieved by 2-D seismic arrays formed by several sensors has

not sense: in fact, SPAC, ESAC and *f-k* curves represent a sort of mean of the surface velocity values characterizing the investigated subsoil portion within the array. Furthermore, in presence of sharp lateral heterogeneities, these techniques suffer severe drawbacks and they can provide a strongly biased dispersion curve estimate.

3.2.2 Methods based on the frequency-wavenumber (f-k) analysis

3.2.2.1 General notions

As showed in Chapter 1, each stationary random process can be characterized by a power spectral density function, or power spectrum, which provides information concerning the frequency composition of the power of the phenomenon under study. The *f-k* techniques are aimed to study the ambient vibrations acquired by a 2-D array by means the estimate of the power spectral density function (or power spectrum) in the frequency-wavenumber domain ($P(f,k)$). This function describes the phase and amplitude of planar waves (with frequency *f*) that propagate in the *x-y* plane of the array with velocity and directions defined by the two-dimensional horizontal wavenumber vector $k(k_x, k_y)$. Moreover, the $P(f,k)$ function provides the energy distribution imaging among the different harmonic components of the wavefield, in relation to their velocity and propagation direction.
In particular, for a defined frequency, the absolute maximum and the possible secondary ones are individuated in the k_x-k_y plane (Fig. 3.7): these maxima are related to the more energetic seismic phases at the considered frequency.

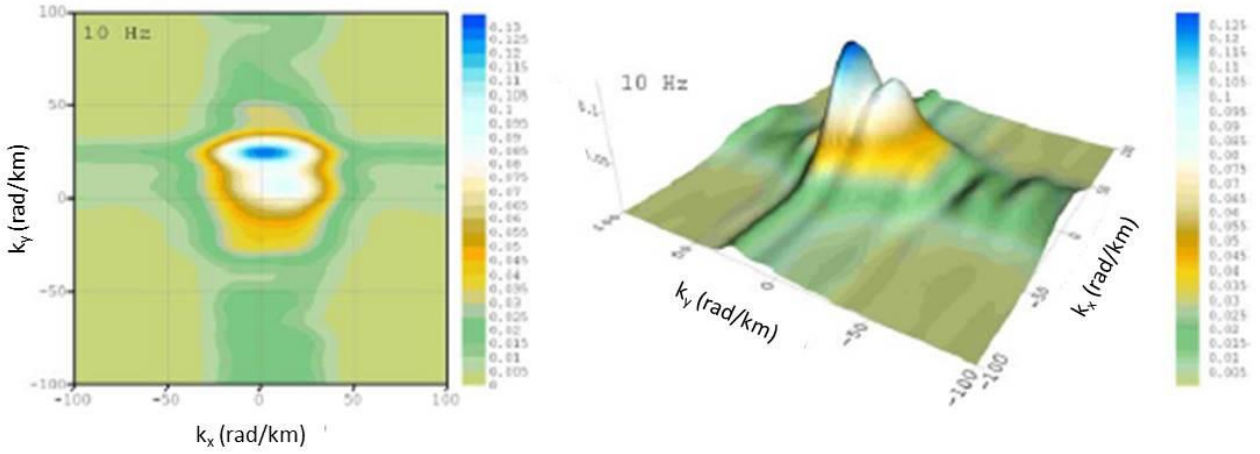

Fig. 3.7 - Example of the power spectral density function values (indicated in the colored bars) obtained by *f-k* Beam Forming method (*vide infra*) for the 10 Hz frequency as a function of the wavenumber vector horizontal components (k_x and k_y). The 2-D and 3-D representations are showed respectively on the left and on the right panel (from Baliva, 2006).

Considering a certain frequency f_0, the wavenumber vector $k_0(k_{x0,}k_{y0})$ is related with the phase velocity $c(f_0)$ and the propagation angle β of the signal by these simple relationships:

$$k_0 = \frac{2\pi f_0}{c(f_0)}, \tag{3.19}$$

$$c(f_0) = \frac{2\pi f_0}{k_0} = \frac{2\pi f_0}{\sqrt{{k_{x0}}^2 + {k_{y0}}^2}}, \tag{3.20}$$

$$k_{x0} = \frac{2\pi f_0 . \cos\beta}{c(f_0)} \text{ and } k_{y0} = \frac{2\pi f_0 . \sin\beta}{c(f_0)}, \tag{3.21}$$

$$\beta = arctg\left(\frac{k_{y0}}{k_{x0}}\right). \tag{3.22}$$

It is clear that, after the identification of the power spectral density maxima and its coordinates k_{x0} and k_{y0} in the *f-k* spectrum at the frequency f_0, the phase velocity value computation is immediate.

It is worth remembering that the *f-k* techniques do not investigate the nature of the waves detected, but rather they analyze the energetic content of the wavefield. If a higher mode surface wave is dominant, such a higher mode wave is detected, and if a body wave is dominant, the body wave is detected. Therefore, the *f-k* methods are not aimed to individuate surface waves alone. These techniques are used to retrieve the surface wave dispersion curve for the passive seismic prospecting method because it is assumed that surface waves are relatively dominant over the other seismic phases mixed into the ambient vibration wavefield (Okada, 2003).

3.2.2.2 Beam-Forming method (BFM) and Maximum Likelihood Method (MLM)

The most popular *f-k* techniques used for geophysical exploration purposes are the Beam-Forming method (BFM) (Lacoss *et al.*, 1969) and the Maximum Likelihood Method (MLM) (Capon, 1969).

The estimate of the *f-k* spectrum $EP_{BFM}(f,k)$ by the BFM is given by (Picozzi *et al.*, 2010):

$$EP_{BFM}(f,k) = \sum_{l,m=1}^{n} \Phi_{lm}\, exp\,[ik(X_l - X_m)], \tag{3.23}$$

where *n* is the number of sensors, Φ_{lm} is the cross power spectrum between the l^{th} and m^{th} data, *i* is the imaginary unit and X_l and X_m are the coordinates of the l^{th} and m^{th} sensors, respectively.

Instead, the estimate of the *f-k* spectrum $EP_{MLM}(f,k)$ for the MLM is expressed as (Picozzi *et al.*, 2010):

$$EP_{MLM}(f,k) = \left(\sum_{l,m=1}^{n} {\Phi_{lm}}^{-1}\, exp\,[ik(X_l - X_m)]\right)^{-1}. \quad (3.24)$$

With respect to the BFM, the inversion of the cross-power spectral complex matrix (Φ_{lm}) is the only additional processing step. The inverse of Φ_{lm} can be computed following standard numerical approaches, like performing the singular value decomposition (SVD; e.g. Press *et al.*, 1992).
The estimate EP_{BFM} and EP_{MLM} of the true P_{BFM} and P_{MLM} f –k spectra may be considered the convolution of the true functions with a frequency window function W_f and the wavenumber window functions W_{BFM} and W_{MLM} for the BFM and MLM, respectively (Lacoss *et al.* 1969). The first window function W_f is the transfer function of the tapering function applied to the signal time windows (Kind *et al.* 2005). The function W_{BFM} is defined as the array response function (or spatial window function by Lacoss *et al.*, 1969; or beam-forming array response function by Capon, 1969) and it only depends on the distribution of the sensors in the array, namely on the spatial sampling of the ambient vibration wavefield. This function can be represented as the power spectral density function $P_{BFM}(f,k)$ of a planar wave striking the seismic array vertically from below (namely a signal corresponding to a two-dimensional Dirac delta function). For the wavenumber vector k_0, the array response function can be described as (Horike, 1985):

$$W_{BFM}(k,k_0) = \frac{1}{n^2}\sum_{l,m=1}^{n} exp\,[i(k-k_0)(X_l - X_m)], \quad (3.25)$$

In practice, this function represents a kind of spatial filter for the wavefield.
The array response function exhibits a central peak (or lobe) of unitary amplitude (k_x and k_y = 0) and several secondary lateral peaks with amplitude lower than one due to aliasing phenomena (Fig. 3.8). The main peak width and the number and amplitude of the secondary lobes are the principal features of the resulting spectrum.
For a complex wavefield, aliasing is likely to occur for lower wavenumbers due to the summation of the lateral peaks of W_{BFM}: hence, the estimation of this function is of paramount importance to define the potential aliasing limits (k_{max}) of the chosen array geometry. Moreover, it is clear that the thinner the central peak is, the more capable is the array to distinguish two waves travelling at close wavenumbers (Whatelet, 2005). The resolution limit (k_{min}) is controlled by the width of the central peak. For simple and regular array geometries, k_{min} and k_{max} are related to the minimum and maximum distance between sensors; instead, for irregular array geometries, the estimation of W_{BFM} is necessary for the definition of objective wavenumber limits (Whatelet, 2005). For instance, a rule of thumb to define the aliasing and resolution limits from W_{BFM} is to set k_{max} at the first secondary peak exceeding the W_{BFM} value equal to 0.5 and k_{min} being measured at the mid-height of the central peak (Woods and Lintz, 1973; Asten and Henstridge, 1984).

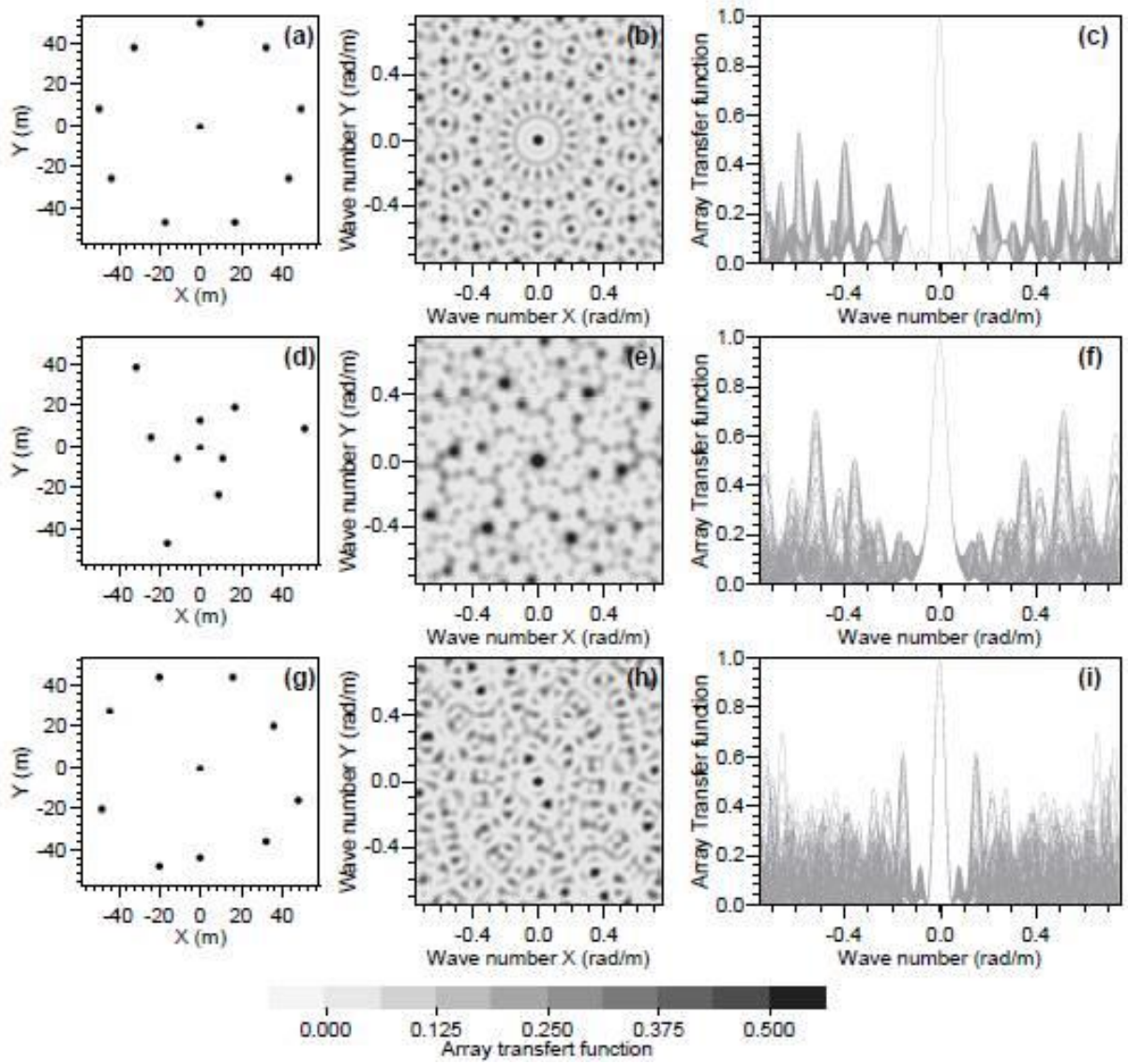

Fig. 3.8 - Theoretical array responses for 10 sensors. Panels (a), (d), and (g) show the array geometries. Panels (b), (e), and (h) show the theoretical array responses in the plane (k_x, k_y). Panels (c), (f), and (i) show the sections across theoretical array responses (from Wathelet, 2005).

As concerns the wavenumber window function W_{MLM}, for a wavenumber vector k_0, it can be expressed in the form (Horike, 1985)

$$W_{MLM}(f, k, k_0) = \left|\sum_{j=1}^{n} A_j(f, k_0)\right| W_{BFM}(k, k_0), \tag{3.26}$$

where

$$A_j(f, k_0) = \frac{\sum_{l=1}^{n} q_{jl}(f, k_0)}{\sum_{j,l=1}^{n} q_{jl}(f, k_0)}, \tag{3.27}$$

and where q_{jl} represents the elements of the cross power spectral matrix. From these relationship, It is evident that W_{MLM} depends not only on the array configuration through the function W_{BFM}, but also on the quality (i.e. the signal to-noise ratio) of the data (Horike, 1985). In fact, the wavenumber response is modified by using the weights $A_j(f, k_0)$, which depend directly on the elements $q_{jl}(f)$. In particular, coherent signals are associated with large weights of A_j and their energy is emphasized in the *f-k* spectrum, reducing the amplitude and the presence of the secondary aliasing lobes; on the contrary, if the coherency is low, the weights A_j are small and the energy in the *f -k* spectrum is damped (Kind *et al.*, 2005).

Experimental outcomes (Wood and Lintz, 1973; Asten, 1976) show that the resolving power in *f-k* spectrum estimating of the MLM is higher than of the BFM (Fig. 3.9) mainly due the adaptive nature of its wavenumber window function. However, being the MLM more sensitive to the signal to-noise ratio, when the analyzed signals contain strong transients, its higher resolving power with respect to BFM cannot be considered a general property (Okada 2003).

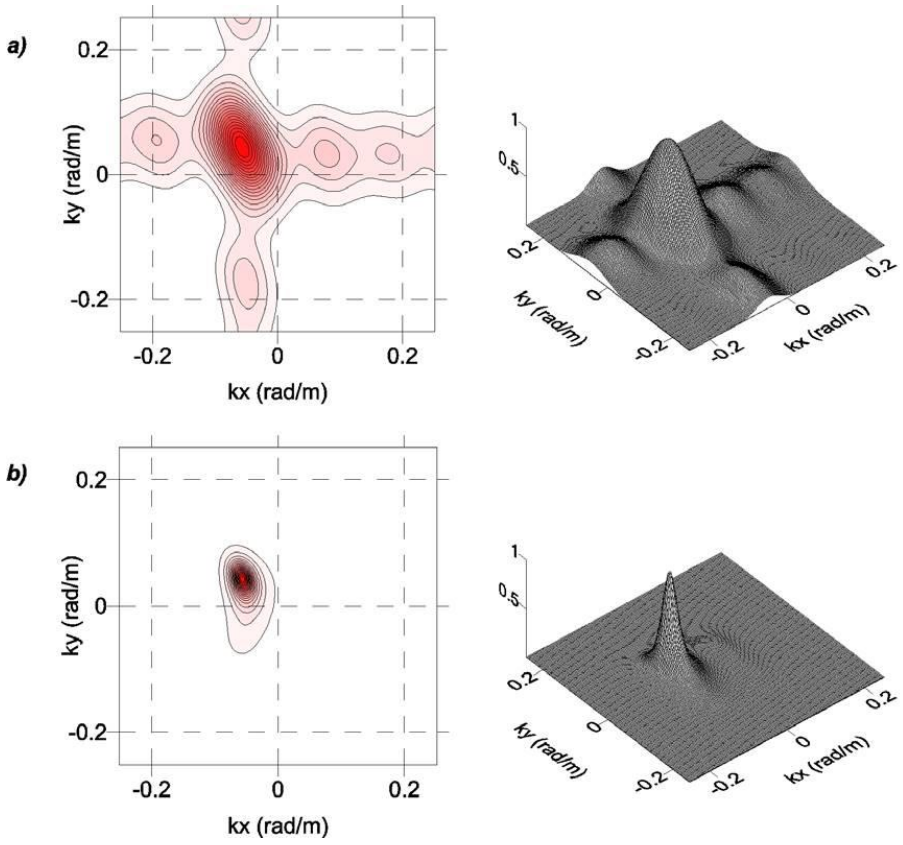

Fig. 3.9 - Contour and mesh plots of the frequency-wavenumber power spectral density function. a) $EP_{BFM}(f, k)$ estimate from beam-forming method. b) $EP_{MLM}(f, k)$ estimate from maximum likelihood method (from Picozzi, 2005).

Considering the dependence of W_{MLM} on W_{BFM}, it is clear that the array geometry is a factor having a strong influence on both EP_{BFM} and EP_{MLM}. In fact, several large side lobes located around the main central peak can remain in the *f-k* spectrum, hampering the phase velocity estimate.

For this reason, it is considered good practice to undertake a preliminary evaluation of the array response when the survey is planned, taking into account various array apertures and geometries with overlapping wavenumber ranges. These characteristics can be based on a first guess of the dispersion curve inferred considering the expected geology: the wavenumber ranges must cover the whole dispersion curve down to the expected resonance frequency (Whatelet, 2005). Moreover, irregular configurations of even only a few sensors should be preferred, because they allow one to obtain a good compromise between a large aperture, which is necessary for sharp main peaks in the EP_{BFM} and EP_{MLM}, and small inter-station distances, which are needed for large aliasing periods (Kind *et al.* 2005).

3.2.2.3 Comparison with ESAC

The dispersion curves obtained from the *f-k* BMF and MLM analyses are compared with that one obtained from the previously described ESAC method in Fig. 3.10. It is possible to note that all the showed curves look normally dispersive and are in good agreement with each others over a wide frequency range. However, at lower frequencies (below 5 Hz) the *f-k* methods provide larger estimates of phase velocity than ESAC: this bias in the phase velocity estimate is defined by Okada (2003) as "degeneration phenomenon of *f-k* spectra". He states that *f-k* methods are able to use wavelengths up two to three times the largest interstation distance, whereas with the ESAC method one may investigate wavelengths up to 10 to 20 times the largest interstation distance, being therefore more reliable in the low-frequency range (Foti *et al.*, 2011). This effect is a serious drawback of these procedures, because it will affect the deeper portion of the S-wave velocity profile (e.g. bedrock velocity) that can be retrieved inverting the *f-k* dispersion curves. Okada (2003) concludes stating that to avoid the degeneration phenomenon, the method requires a large array as well as a large number of sensors. However, a large number of sensors will increase the field effort and decrease field efficiency, while a large array size adversely affects the assumption that layers are subparallel under the array (Okada, 2003).

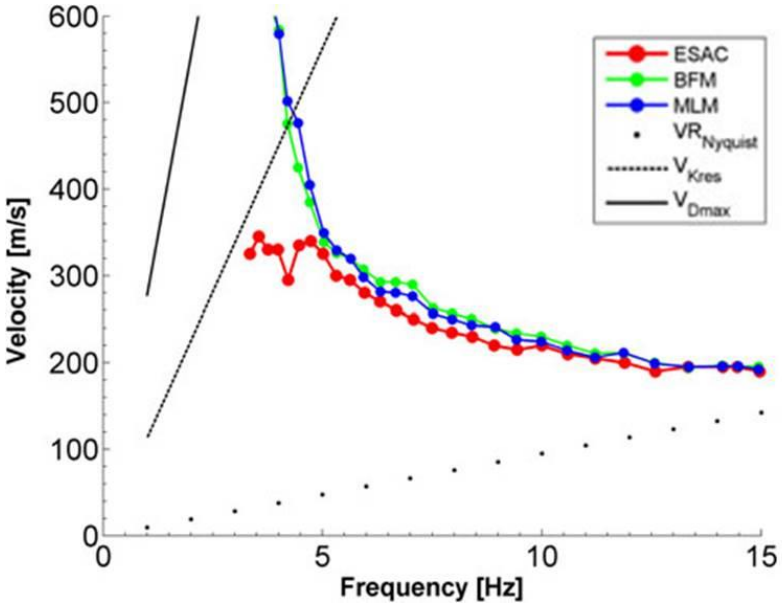

Fig. 3.10 - Phase velocity dispersion curves obtained by ESAC (red line), MLM (blue line) and BFM (green line) analysis. The dotted black line indicates the theoretical aliasing limit calculated as $4f \cdot d_{min}$, where d_{min} is the minimum interstation distance. The dashed gray line indicates the lower frequency threshold of the analysis based on $2\pi f \cdot \Delta k$, where Δk is calculated as the half-width of the main peak in the array response function. The continuous grey line indicates the lower frequency threshold of the analysis based on the criterion $f \cdot d_{max}$, where d_{max} is the maximum interstation distance in the array (from Foti *et al.*, 2011).

Finally, from the qualitative point of view it is possible to state that (Picozzi, 2005):

- The ESAC computation is simpler and faster;
- By checking the fit between the spatial correlation coefficient and the Bessel function for various frequencies, the ESAC method is able to tell if inappropriate phase velocity values related to others seismic waves are detected;
- The *f-k* estimate, even if it requires more computational efforts, allows detecting significant information about the direction of the main sources of the wave (that is especially important in the case that noise is not randomly distributed) and might help separating the contribution of higher modes of surface waves to the wavefield (Okada, 2003).

3.2.3 Cross-correlation in time domain

It is well know that the ground motion generated at a site on the Earth surface by any radiating energy source can be represented as the convolution of a source function of any suitable distribution of volume forces and the Green's function, that represents the response of the Earth to a sudden point load (Aki and Richards, 2002). The form of this function is determined by the dynamic characteristics (seismic wave velocities and damping) of the subsoil materials. Since the second half of the last century, several studies witnessed that the Green's function relative to a portion of the crust can be retrieved by cross-correlating seismographic registrations of the same

event at two sites located along the direction of wave propagation (e.g. Dziewonski and Hales, 1972).

More recently, several studies showed the possibility to retrieve the Green's function from cross-correlation in time domain of ambient vibration measurements (see Larose *et al.*, 2006 and references therein): this procedure relies on the assumption that the ambient vibration wavefield can be considered as a random diffuse wavefield (see Subparagraph 2.2.3) generated by a number of independent stochastic sources or by the presence of a scattering medium (e.g., Weaver and Lobkis, 2001, 2004; Snieder, 2004). Shapiro and Campillo (2004) confirmed the possibility in estimating the Rayleigh wave component of Green's functions and its dispersion characteristics between two stations by the cross-correlation in time domain of simultaneous long seismic noise sequence recordings. In particular, these authors retrieved the Rayleigh group velocity dispersion curve between each considered pair of stations computing a broadband (0.008–0.07 Hz) cross-correlation and applying to it a frequency-time analysis (FTAN) of Levshin *et al.* (1989) (Fig. 3.11).

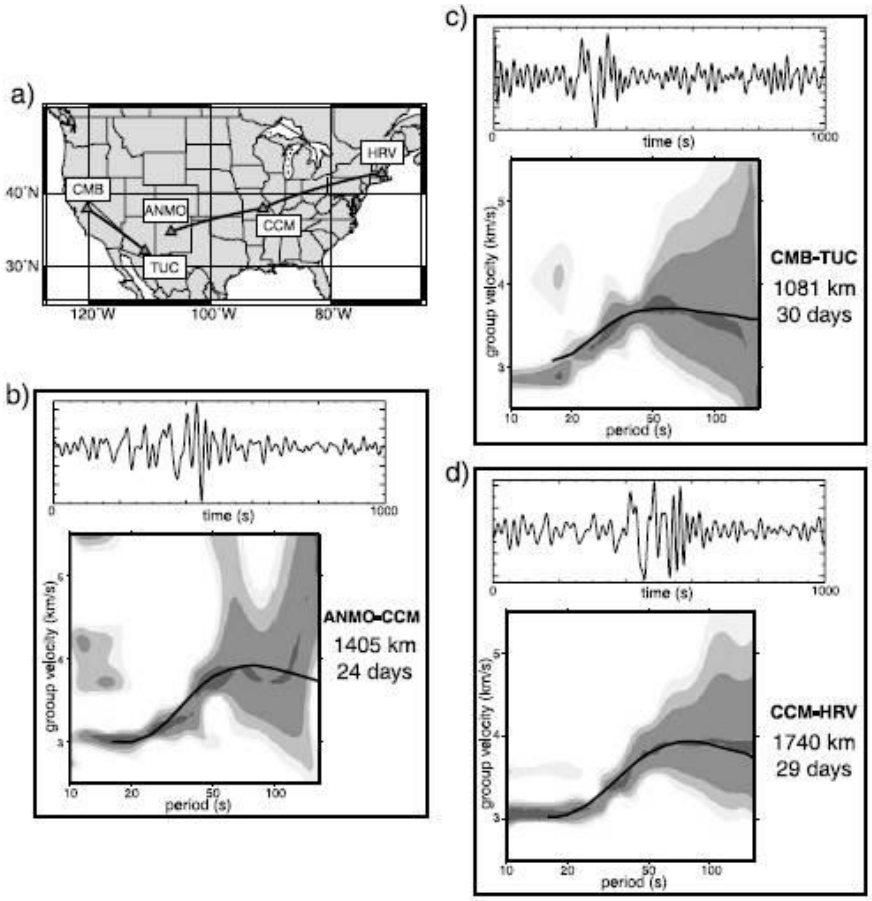

Fig. 3.11 - Analysis of broadband (0.008–0.07 Hz) cross-correlations computed for three pairs of stations from the vertical component of the ambient vibrations recorded between January 10 and February 8, 2002. (a) Maps showing locations of stations in US. (b)-(d) results for three pairs of stations in US. For each pair of stations, the upper frame shows the cross-correlation in time domain; the lower frame shows the comparison of the Rayleigh wave group velocity dispersion curve computed from this cross-correlation with a dispersion curve (solid black line) predicted for the corresponding inter-station path from global Rayleigh wave group velocity tomographic maps (Ritzwoller *et al.*, 2002) computed from surface waves generated by seismic events (modified from Shapiro and Campillo, 2004).

This procedure, also generally defined as seismic interferometry, is commonly used to accomplish surface wave tomography maps at continental scales in order to infer an image of the crust and the upper-mantle structure (e.g., Shapiro and Campillo, 2004; Shapiro *et al.*, 2005; Gerstoft *et al.*, 2006; Yao *et al.*, 2006; Yang *et al.*, 2007; Lin *et al.*, 2008). Hence, for these kinds of studies, ambient vibration recordings from national seismic networks and frequencies well below 1 Hz were considered.

Several studies witness the reliability of the ambient vibration cross-correlation for frequencies greater than 1 Hz in order to infer the subsurface structure from depths of a few meters to several hundred meters (e.g., Schuster, 2001; Dong *et al.*, 2006; Halliday *et al.*, 2007, 2008). An application of seismic high frequency seismic interferometry to constrain damping profiles in the shallow subsoil was also provided by Albarello and Baliva (2009). Moreover, several authors (e.g., Chavez-Garcia and Luzon, 2005; Chavez-Garcia and Rodriguez, 2007; Yokoi and Margaryan, 2008) showed, for a small scale experiment at a site with a homogeneous subsoil structure, the equivalence between the results obtained by cross-correlation in the time domain and the SPAC method.

An interesting application of this approach for frequency greater than 1 Hz was proposed by Picozzi *et al.* (2009). In view of what was obtained over regional scales, the authors explored the cross-correlation in time domain in order to retrieve the Rayleigh wave group velocity values from the Green's function estimate between each pair of sensors belonging to a small scale array, with maximum inter-station distance of the order of 100 m. The final purpose of this study was the accomplishment of an ambient vibration tomography at local scale aimed to identify the lateral heterogeneities in terms of Rayleigh wave group velocity within the subsoil portion sampled by the array. The seismic array (the same one analyzed by Pilz *et al.*, 2012; see part 3.2.1.2) was located in a well know site (previous geophysical investigations were present, Fig. 3.12) in Nauen (Germany) and it was composed by 21 stations deployed with irregular geometry and with minimum and maximum inter-sensors distance of about 5 m and 100 m respectively (Fig. 3.12). The vertical component ambient vibrations were acquired for a week.

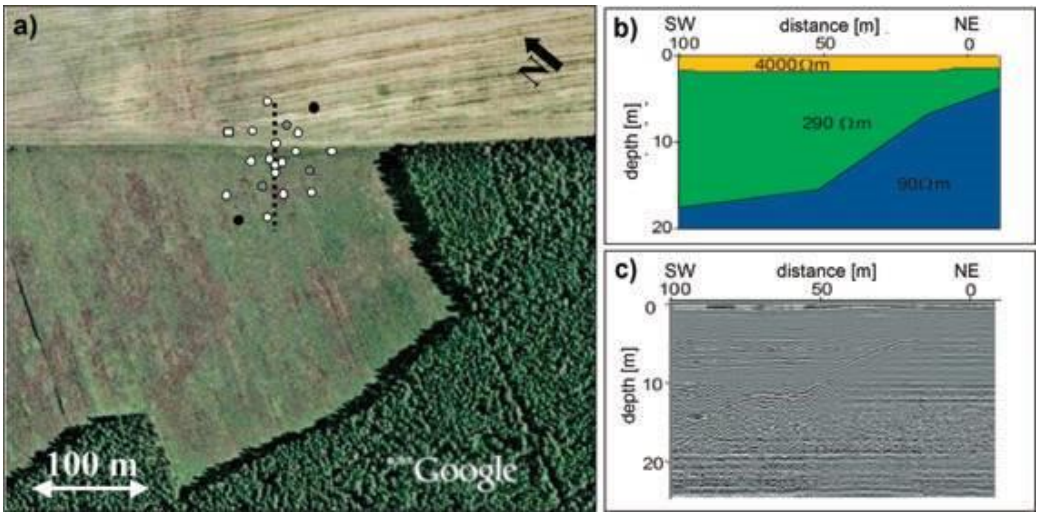

Fig. 3.12 - (a) Nauen test site: location of the sensors (white dots, black dots and grey dots) and of the geoelectric and GP radar profiles (dotted line). (b) geoelectric profile. (c) GP radar profile (from Picozzi *et al.*, 2009).

From the continuous data streams, 480 non-consecutive noise windows 30 s wide were extracted from different days and different hours. After detrend and tapering, for each window the cross-correlations and its time derivatives were computed considering one-bit normalized data (e.g. Campillo and Paul 2003, Bensen *et al.*, 2007). The Rayleigh wave group velocity values were estimated at 12 different frequencies (i.e. 14, 12, 11, 10, 9, 8, 7.5, 7, 6.5, 6, 5.5 and 5 Hz) following this procedure:

1. the cross-correlation functions were filtered by applying a Gaussian filter with a narrow bandwidth (e.g. Dziewonski *et al.* 1969);
2. for each pair of stations, all estimated cross-correlations at the considered frequencies were stacked together and the resulting Green's functions normalized to a maximum amplitude of ±1;
3. the propagation delay between the stations was estimated by picking the maximum on the Green's functions envelope, computed through a spline interpolation. Consequently, considering the inter-station distance, the Rayleigh wave group velocity value was known.

In this work, the stability of the estimated Green's functions was also investigated: for this purpose, the root mean square (*rms*) difference between Green's functions obtained from increasing the number *n* of stacked cross-correlations was computed for frequencies 14 and 6 Hz. As it is possible to note in Fig. 3.13, regardless of the frequency, after stacking 50 cross-correlation functions, the *rms* no longer significantly changes: this means that, using windows 30 s wide, about 20–30 min of signal is sufficient for obtaining an estimate of the Green's functions between the two sensors, suggesting the possibility of an application of this method in urban areas, where long instrumental installations would not be possible (Picozzi *et al.*, 2009).

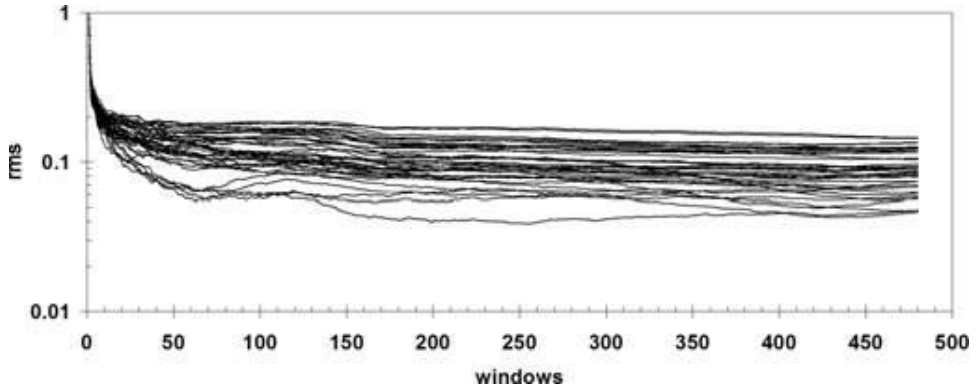

Fig. 3.13 - Stability (*rms*—root mean square) of the estimated Green's function with respect to the number of stacked windows for 14 and 6Hz (from Picozzi *et al.*, 2009).

Fig. 3.14 shows the computed Green's functions against distance for the frequency of 14 Hz: unlike the other considered frequencies, the Green's functions shows a clear asymmetry (Fig. 3.14-a) related with the non-uniform distribution of ambient vibration sources (with the largest energy contribution coming from the southern and western directions) as showed by the 2-D wavenumber plot obtained by *f-k* analysis carried out at this frequency (Fig. 3.14-b). Plotting the

estimated time delays as a function of the distance and azimuth between each pair of stations, the authors showed that the traveltime variations are dependent mainly on the interstation distances and that no significant changes with azimuth are present (Fig. 3.14-c). This feature witnesses that the traveltime estimates are stable even if the majority of sources are clustered in a narrow azimuthal range, provided that enough energy is coming from the remaining azimuthal sectors (Picozzi *et al.*, 2009).

This result is in agreement with Louie (2001) and Roux *et al*. (2005): these authors demonstrated that the hypothesis of a purely diffuse wavefield is not strictly necessary to estimate the Green's function. In fact, this procedure only requires the presence of at least some wavelet traveling along the direction nearly parallel to the one connecting the relevant sensors.

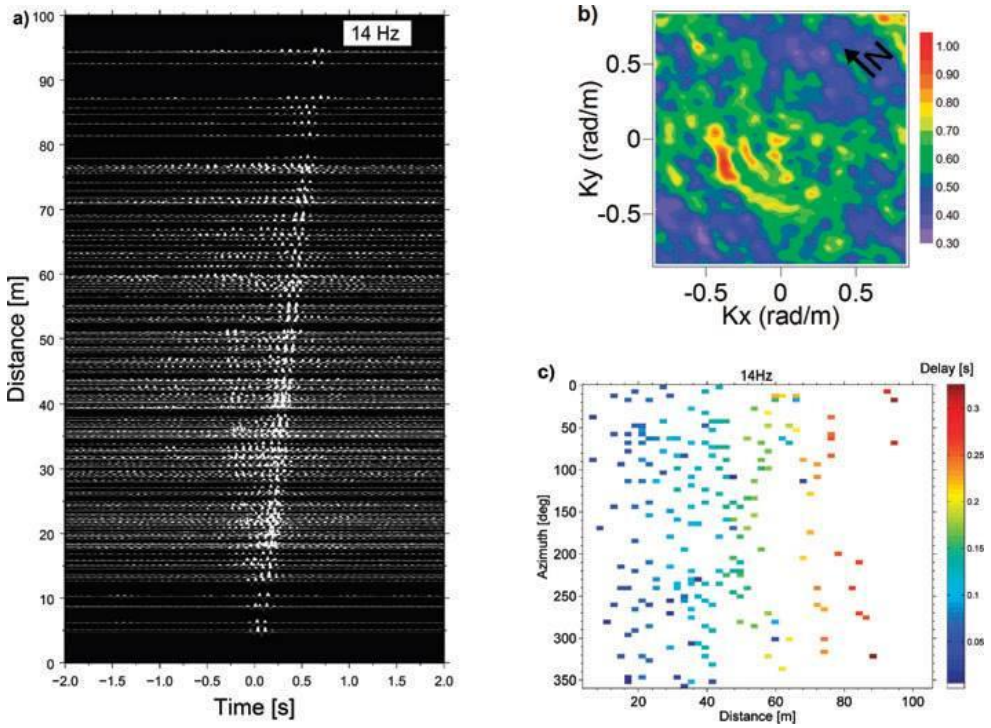

Fig. 3.14 - (a) Normalized cross-correlations computed for 14 Hz. (b) 2-D-wavenumber plot showing the distribution of the noise sources as determined by *f –k* analysis results for 14 Hz. (c) Time delays as a function of distance between each pair of stations and orientation of the pair with respect to north. The delay time is indicated by the color of the symbol (from Picozzi *et al.*, 2009).

Finally, the interstation traveltimes estimated from the Green's functions were inverted by a tomographic approach to derive a 2-D Rayleigh wave velocity model for each considered frequency. Moreover, the authors carried out an inversion procedure in order to estimate the S-wave velocity model (Fig. 3.15) over the section showed in Fig. 3.12-a: the resulting structure is very similar to that one obtained by geoelectric and GP radar survey (Fig. 3.12-b-c).

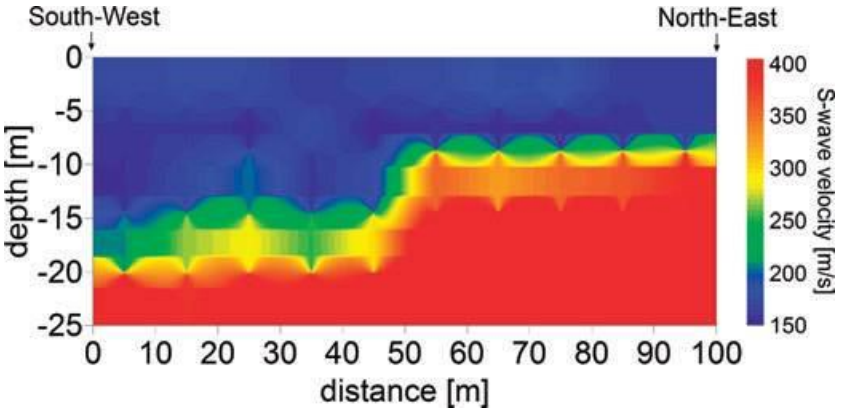

Fig. 3.15 - *S*-wave velocity section along the same geoelectric and GP radar profiles (from Picozzi *et al.*, 2009).

Following Picozzi *et al.* (2009), Renalier *et al.*, (2010) applied the ambient vibration cross-correlation in time domain technique to infer the Rayleigh wave velocity tomography and the 3-D V_S model of a landslide located in Avignonet (France). In this case, the authors analyze the vertical component of the ambient vibration recordings obtained by an array with irregular geometry and with minimum and maximum inter-sensors distance of tens of meters and 1000 meters respectively. Differently by Picozzi *et al.* (2009), after computing the cross-correlations analyses, they estimated the Rayleigh wave group velocity dispersion curves for each pair of stations by using the time-frequency analysis based on *S* transform (Stockwell *et al.*, 1996). Moreover, the authors noted that the cross-correlations were not stable at frequencies higher than 5 Hz: to overcome this problem, they performed a cross-correlation in time domain analysis from active sources to infer the dynamic characteristics of the shallowest part of the subsoil. Finally, the retrieved Rayleigh wave phase velocity maps and 3-D V_S model of the landslide are in good agreement with previous investigations.

The above described studies witness that the analyses based on the cross-correlation in time domain method can be useful tools to point out the Rayleigh wave phase velocity (and consequently the S-wave velocity) differences within of a local scale seismic array. Moreover, since Picozzi *et al*. (2009) showed the reliability of the Green's functions estimated with just 30 min of recordings and in presence of a non uniform distribution of noise sources as well as using a limited number of seismic stations (about 20), the use of this technique could be effective in urban framework and therefore for Seismic Microzonation studies. On the other hand, Renalier *et al.* (2010), witnessed that the Green's functions can be characterized by instable behavior for frequencies higher than 5 Hz also for long period acquisitions (about 15 days), hampering the retrieving of the shallow subsoil information. This problem could be related with the poor sampling of wavelengths shorter than hundreds meters: in fact, in the used seismic array, only one pair is characterized by inter-station distance of tens of meters.

3.3 The effect of S-wave velocity inversion on the dispersion curve

One of the purposes of this dissertation is the study of the effect of a sharp shear wave velocity inversion in the subsoil on the passive seismic measurements and, moreover, to verify the reliability and capability of these techniques in estimating the V_S profile in this unusual geological setting.

As was done for the HVSR curve (see paragraph 2.6), the aim of this paragraph is a review of the main works present in literature dealing the influence of the S-wave velocity inversion on the Rayleigh wave dispersion curve.

As mentioned in paragraph 1.3, Tokimatsu *et al*. (1992a) and Tokimatsu (1997) performed some tests with synthetic data and considering three different input models (Fig. 3.16-a). For each input model he showed the theoretical Rayleigh wave phase velocity curves for the fundamental up to the fifth higher mode and the phase velocity curves estimated between two sensors from synthetic data (Fig. 3.16-b-c-d). Together with these curves, he showed the effective (or apparent) phase velocity values $c_e(\omega)$ for the vertical component of the ambient vibrations defined as (Tokimatsu *et al.*, 1992a):

$$c_e(D,\omega) = \frac{\omega D}{\cos^{-1}\left[\frac{\sum_{n=1}^{M} {A_n}^2(\omega) c_n(\omega) \cos\left(\frac{\omega}{c_n(\omega)} D\right)}{\sum_{n=1}^{M} {A_n}^2(\omega) c_n(\omega)}\right]}, \tag{3.28}$$

where $c_n(\omega)$ and $A_n(\omega)$ are the phase velocities of the n^{th} mode Rayleigh wave and the corresponding medium response. $A_n(\omega)$ is related to the power spectrum density function of the n^{th} mode, *M* is the maximum order of mode for each frequency, and *D* is the distance between the sensors. Eq. 3.16 is a generalization of Eq. 3.28. In fact, the above equation expresses the effective phase velocity considering a signal propagates along the direction connecting the sensors, while the Eq. 3.16 takes into account an azimuthally homogeneous wavefield around the receivers.

a)

Layer number (1)	Thickness H (m) (2)	Density ρ (Mg/m³) (3)	V_p (m/s) (4)	V_s (m/s) Case 1 (5)	Case 2 (6)	Case 3 (7)
1	2	1.8	300	80	180	80
2	4	1.8	1,000	120	120	180
3	8	1.8	1,400	180	180	120
4	—	1.8	1,400	360	360	360

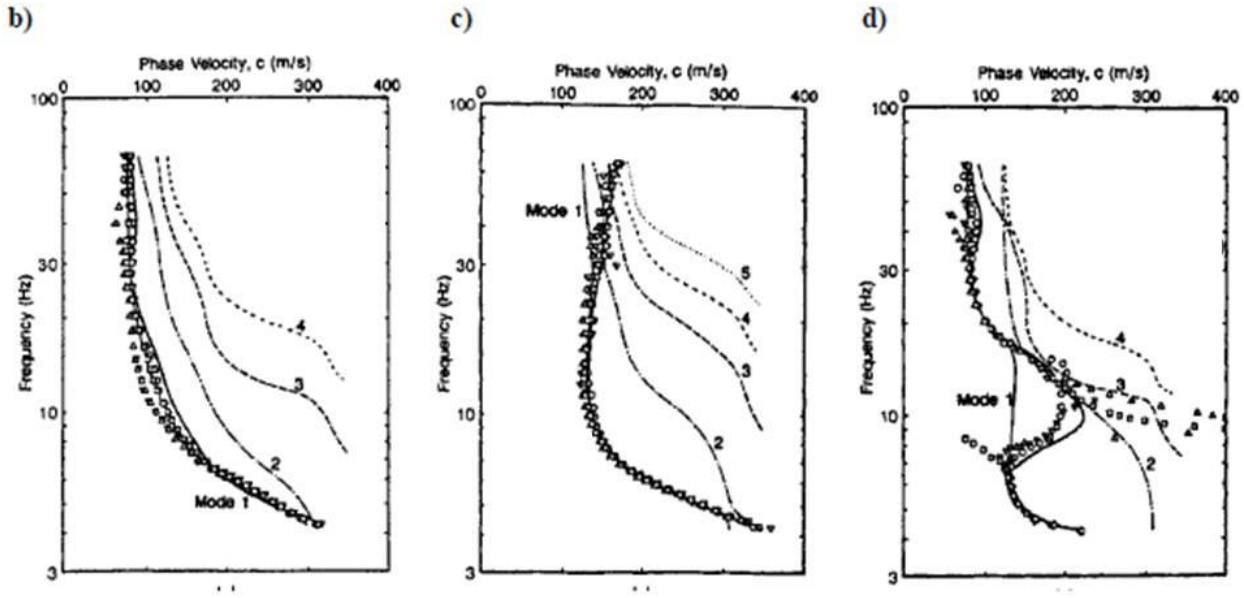

Fig. 3.16 - (a) Soil input models. Variation of the phase velocity with frequency for three different models: case 1 (b), case 2 (c), case 3 (d). All three figures show: dispersion curves from synthetic data using different sensor locations (dots, triangles and squares), theoretical phase velocities for the fundamental up to the fifth higher mode (dashed lines), and phase velocity for the apparent dispersion curve (thick line) obtained by Eq. (3.28) (modified from Tokimatsu *et al.*, 1992a).

Fig. 3.16-b-c-d shows the results of these studies. In particular, if the S-wave velocity of the subsoil varies regularly with depth, the experimental dispersion curves tend to be coincident with the fundamental mode of Rayleigh waves, showing a normally dispersive character (Fig. 3.16-b). Instead, if the S-wave velocity profile presents low velocity layers embedded between high velocity ones, a higher mode or even multiple modes dominate certain frequency ranges; in this case the effective Rayleigh wave phase velocity dispersion curve is characterized by an inversely dispersive trend in these frequency ranges (Fig. 3.16-c-d).

An interesting work was proposed by Zhang and Lu (2003). The authors analyzed the Rayleigh wave dispersion curve shape plotting the phase velocity values V_R as a function of the wavelength λ, assuming that λ value is equal to the depth value z. In particular, they analyzed the V_R-λ functions related by configurations where low velocity layers are contained in a stratified half-

space: these curves, defined as non-continual "zigzag" curves (Fig. 3.17), can be exploited as a possible tool in order to identify the positions of the low velocity structures.

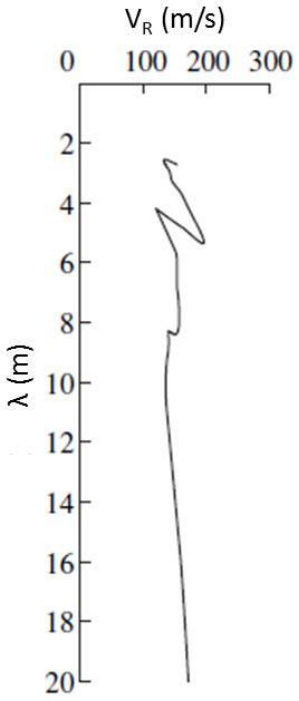

Fig. 3.17 - An example of theoretical zigzag dispersion curve computed by a model with two low shear velocity layer located at 5 m and 10 m respectively (modified from Zhang and Lu, 2003).

The authors, using numerical simulations based on the propagator matrix technique (Haskell, 1953), demonstrated that the observed zigzag V_R-λ dispersion curves are the result of the leap of the modes: in fact, as showed previously, higher modes than the fundamental one play an important role in presence of S-wave velocity profile characterized by low velocity layers included between high velocity ones. This does not occur in the cases where the shear wave velocities increase or decrease gradually from the top to the bottom layers: in these conditions only the fundamental mode of the signal has the dominant intensity and the V_R-λ curves are characterized by smooth and continual shape.

Finally, the authors conclude stating that the depth of the shallowest low velocity layer coincides with the λ value of the inflection point of the first zigzag shape. Conversely, this can not occur for the subsequent deeper low velocity layers (Fig. 3.17), where the wavelength value of the corresponding zigzag shape should be multiply by a proper weight coefficient. The value of this coefficient, less than 1, is variable and it is usually given by work experience and the soil properties of the local work area. The coefficient is generally about 0.5 in engineering seismic exploration, because the compressional and shear wave velocities are relatively small, while it is higher in the formation with higher velocities of compressional and shear waves.

CHAPTER 4
Shear wave velocity profile estimate

4.1 Introduction

As mentioned in the Introduction, the V_S profile estimate is of paramount importance for the Seismic Microzonation studies. In particular, it allows to assess the dynamic characteristics and to derive a quantitative analysis of seismic amplification of the investigated site. This amplification effect is often based on the average shear wave velocity estimate in the first 30 m of subsoil; this parameter, object of several scientific disputes (e.g., Gallipoli and Mucciarelli, 2009), is commonly defined as V_{S30} (Eurocode 8, EN-1998, 2004).

It is well clear from the previous chapters that HVSR and Rayleigh wave velocity dispersion curves are strictly related to the V_S profile of the site, which can be extracted from these measurements by means an inversion procedure. The inversion techniques can provide detailed quantitative information about the dynamic characteristics of the subsoil; on the other hand they present significant drawbacks:

- the obtained results are largely non-univocal due the strong non-linearity of the inverse problem;
- they are numerically troublesome and highly time consuming procedures.

To supply a preliminary and fast interpretation of the experimental curves, some simplified procedures can be applied: these methods are able to provide a dynamic characterization of the subsoil in terms of average shear wave velocity and an estimate of the resonant interface depth. The main advantages of these techniques are the quickness and the numerical simplicity that permit their use in the field just after the measurement has been taken (Albarello *et al.*, 2011) and, moreover, in the framework of the first and the second level of the Seismic Microzonation studies, where semi-qualitative information are required.

The aim of this chapter is to describe the main procedures that lead to the identification of the subsoil V_S configuration compatible with the observations obtained at the studied site. In the first part, the simplified approaches based on the preliminary interpretation of the Rayleigh wave dispersion curves are treated. In the second part, the inversion procedures are dealt, describing in particular the joint inversion of the HVSR and the dispersion curves with a specific focus on the most commonly used inversion algorithms.

4.2 Simplified approach for average V_S profile estimate

A simplified procedure aimed to infer the average V_S profile estimate and the resonant interface depth exploiting the Rayleigh wave dispersion curve was proposed by Albarello *et al.* (2011). This approach, defined by the authors as “fast and dirty” procedure, is characterized by an expeditious nature and larger uncertainty than to the inversion techniques.

Two assumptions underlie this approach:

- Rayleigh waves in the fundamental mode dominate ambient vibration wavefield;
- presence of a nearly homogeneous soft layer characterized by V_S values smoothly increasing with depth.

In particular, the V_S profile in the resonant layer is tentatively assumed in the form of a power law (Albarello *et al.*, 2011):

$$V_S(h) = V_0(h+1)^x \tag{4.1}$$

where $V_S(h)$ is the S-wave velocity as a function of the depth (1+*h*) ·1 m, V_0 is the S-wave velocity at 1 m under the ground surface and *x* is a suitable exponent. V_0 and *x* are empirical parameters that depend by subsoil nature.

In this approximation, it is possible to establish a relationship between the fundamental resonance frequency f_0 and the depth *H* of the resonant interface by an *f-H* conversion rule in the form (Ibs Von Seht and Wohlemberg, 1999):

$$H \cong \left[\frac{V_0\,(1-x)}{4f_0} + 1\right]^{\frac{1}{1-x}} - 1. \tag{4.2}$$

It is clear that this approach requires an estimate of the parameters V_0 and *x*. Experimental results obtained in different sedimentary settings (Ibs von Seth and Wohlenberg, 1999; Delgado *et al*., 2000a, 2000b; Parolai *et al*., 2002; Scherbaum *et al*., 2003; Hinzen *et al*., 2004; D'Amico *et al*., 2006a, 2006b) suggest that (Albarello *et al.*, 2011):

1. granular soils (sands) present values of *x* near to 0.25;
2. lower values of *x* should correspond to fine soils (clays and mud), since in these cases the dominant effect is the electrostatic interaction among the particles, which results in a minor effect of the lithostatic load on the form of interaction among them;
3. greater values of *x* can be expected in rework soils (landslides) or in fan zones with big non cemented clasts;
4. fluids under pressure can reduce the effect of the lithostatic load, decreasing the exponent *x* also in granular grounds.

Moreover, a negative correlation between the values of V_0 and *x* was observed: in particular, the stronger the lithostatic load effect (i.e., as *x* is higher), the lower is the expected value of V_S at the surface. An approximate parameterization of this correlation can lead an indication of possible realistic couples of these parameters (Albarello *et al.*, 2011):

- V_0 210 m/s and *x* 0.20 for compact soils;
- V_0 170 m/s and *x* 0.25 for sands;
- V_0 110 m/s and *x* 0.40 for rework or very recent soils.

Information about V_0 and *x* can be provided by a direct interpretation of available effective Rayleigh wave dispersion curves, that, in the frame of the approximation considered above, coincides with the fundamental Rayleigh mode. In this case, it is assumed that the average S-wave velocity up to a depth *h* ($\bar{V}_S(h)$) roughly corresponds to the Rayleigh waves phase velocity $V_R(f)$ corresponding to a fraction *c* of the relevant wavelength. That is:

$$\bar{V}_S(h) \cong V_R(f) \quad \text{with} \quad h \approx c \cdot \frac{V_R(f)}{f}. \tag{4.3}$$

On the basis of empirical evidence, several authors (Konno and Kataoka, 2000; Martin and Diehl, 2004) suggested that *c* value varies from 0.3 to 0.5; on the other hand, numerical simulations (Albarello and Gargani, 2010) show that this parameter can be included in the range 0.8-0.9.
The assumptions on Eq. 4.3 imply that the effective dispersion curve deduced from array measurements could be use directly to constrain V_0 and *x* values. In particular, V_R values relative to highest frequencies can be considered as a rough estimate of V_0, while *x* can be estimated from the literature or by best fitting the $\bar{V}_S(h)$ curve with the V_S profile defined in Eq. 4.1.
In order to obtain a first-glance information on the S-wave velocity profile, Albarello *et al*. (2011) applied the Eq. 4.3 to the dispersion curves obtained from the array measurements in the Abruzzo region (Italy) in this form (Fig. 4.1):

$$\bar{V}_S(h) \equiv 1.1 \cdot V_R(f) \quad \text{with} \quad h \equiv \frac{V_R(f)}{2f}. \tag{4.4}$$

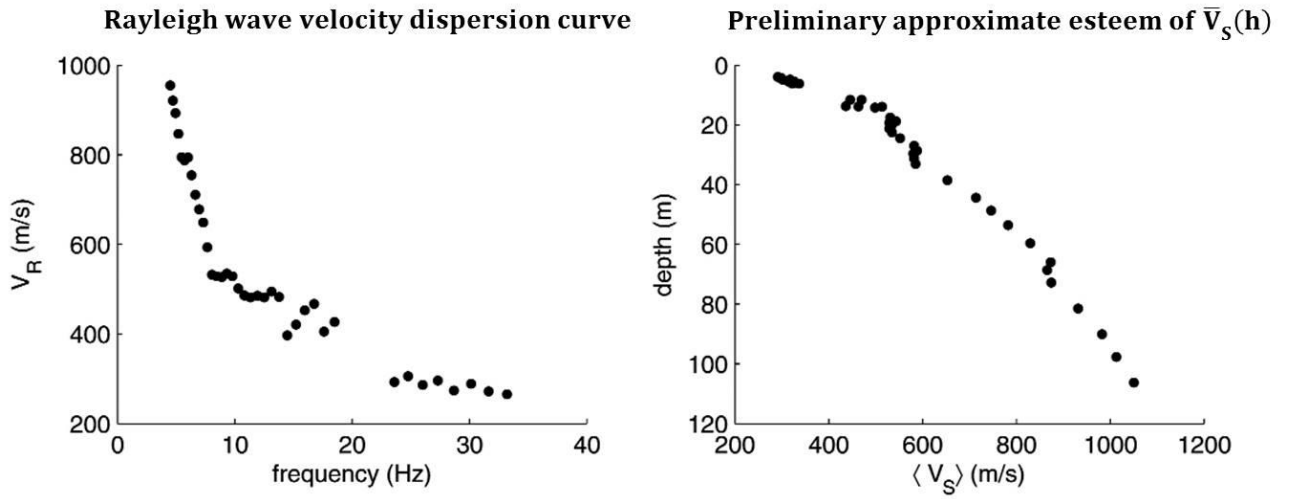

Fig. 4.1 - Example of the "fast and dirty" procedure on the dispersion curve obtained by Albarello *et al*. (2011) at one of the investigated sites in Abruzzo region (Italy). Left: experimental V_R curve. Right: approximate pattern of average S-wave velocity profile obtained by Eq. 4.4 (modified from Albarello *et al.*, 2011).

As showed in D'Amico *et al*. (2008), the V_S profile defined in Eq. 4.1 can be expressed in form of a piecewise power law model. In order to estimate the parameters V_0 and *x*, the authors

constrained the Eq. 4.1 using borehole data and the V_S profile obtained by jointly inverting the dispersion and the HVSR curves (Fig. 4.2). The average V_S profile adopted by the authors for the investigated area reveal the presence of two different trends (one relative to the shallow subsoil and one relative to the deeper layers; solid black line in Fig. 4.2) with two different pairs of V_0 and *x* (i.e. V_1 and x_1 for the shallower part and V_2 and x_2 for the deeper part). In this case, the Eq. 4.2 to be used to compute the deposits thickness *H* (i.e., the depth of the resonant interface) from f_0 estimates assumes the form:

$$H \cong \left[\frac{V_1\,(1-x_1)}{4f_0} + 1\right]^{\frac{1}{1-x_1}} - 1 \quad \text{if} \quad f_0 > \frac{V_1(1-x_1)}{4\left[(1+h^*)^{(1-x_1)}-1\right]}; \tag{4.5}$$

$$H \cong \left[\frac{V_2\,(1-x_2)}{4f_0} + C\right]^{\frac{1}{1-x_2}} - 1 \quad \text{if} \quad f_0 \leq \frac{V_1(1-x_1)}{4\left[(1+h^*)^{(1-x_1)}-1\right]}, \tag{4.6}$$

Where h^* indicates the depth of the transition between the two different subsoil portion. The parameter *C* is defined as:

$$C = (1+h^*)^{(1-x_2)} - \frac{V_2(1-x_2)}{V_1(1-x_1)}\left[(1+h^*)^{(1-x_1)} - 1\right]. \tag{4.7}$$

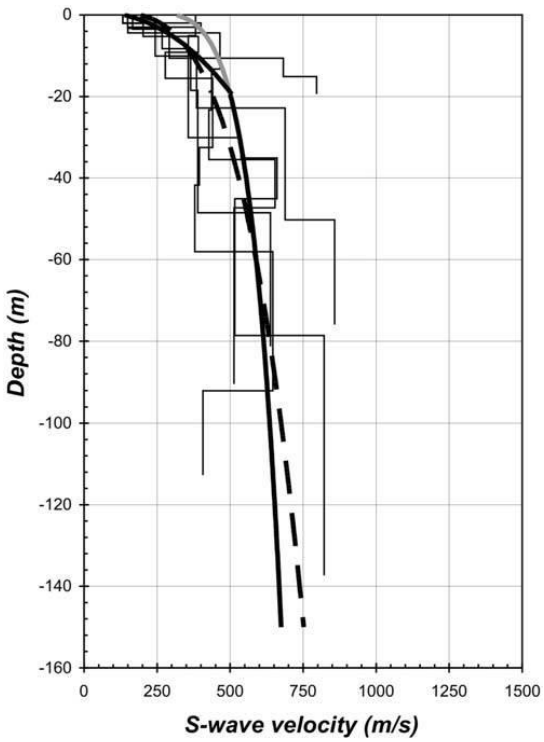

Fig. 4.2 - S-wave velocity profiles derived from joint inversion of dispersion and HVSR curves (thin lines) along with the average V_S models (Eq. 4.1) deduced from the same depicted profile (dashed line) and from borehole data (gray line). The solid black line represents the piecewise power law model (from D'Amico *et al.*, 2008).

4.3 Inversion procedures

Inverting experimental data is a normal and fundamental procedure used in geophysics. Considering the effects of a certain physical phenomenon, the aim of this procedure is to retrieve its underlying causes, i.e., the variation with the depth of the subsoil dynamic properties. Such causes can be studied on the basis of the appropriate choice of different types of effects (e.g., the Rayleigh-wave velocity dispersion curve or the HVSR function) defined as output response.

Once known the physic expression of the phenomenon, i.e. the model, the inversion procedure substantially is a problem of parameters identification. This problem is often illposed, especially in non linear cases like the inversion of the Rayleigh wave velocity dispersion curve.

The ill-conditioning of the Rayleigh wave inversion problem is mainly related with the non uniqueness of the solution: in fact, for a certain experimental dispersion curve, more than one subsoil model can correspond.

In general, the non uniqueness of the solution of the inversion procedures is related to a lack of information of the acquired data or to the intrinsic nature of the problem. In order to cope these issues, two strategies are commonly followed:

1. maximizing the available information;
2. adding limitations to the possible solutions.

As concerns the Rayleigh wave inversion, the first point can be pursued reducing the number of parameters to be inverted, adding possible available a priori information (e.g., P-wave velocity) or jointly inverting a different output response function (i.e., adding another set of experimental data, although dependent on the same subsoil parameters).

The second point can be pursued fixing the solution limits of some parameters (e.g., constraining between a minimum and a maximum the possible V_S and/or V_P values), blocking the ratio between some of them (e.g., setting the Poisson's coefficient value) or adopting restrictions on the overall trend of the solution (the V_S profile) making sure that this one does not present sharp irregularities.

The successful outcome of an inversion procedure can be obtained by solving an optimization problem, i.e., the research of the minimum points of a function defined in a multi-dimensional space that can be subject to constraints (Parker, 1994).

In particular, the simplest approach for the solution of an inverse problem consists in minimizing the squared Euclidean norm of the difference between the observed (d_i) and theoretical (f_i) data:

$$M = \sum_{i=1}^{N} (d_i - f_i)^2. \tag{4.8}$$

The function M is defined as "objective", "cost", "loss", or "misfit" function.

In order to define this function, it is evident that it is needed to solve the problem of the forward modeling. In particular, once that all the physic laws and the parameters of the process under study are known and well modeled, the deduction of the theoretical values of the interested

output response function is a computational task only. Hence, the forward modeling is a fundamental part of the inversion procedure (Fig. 4.3).

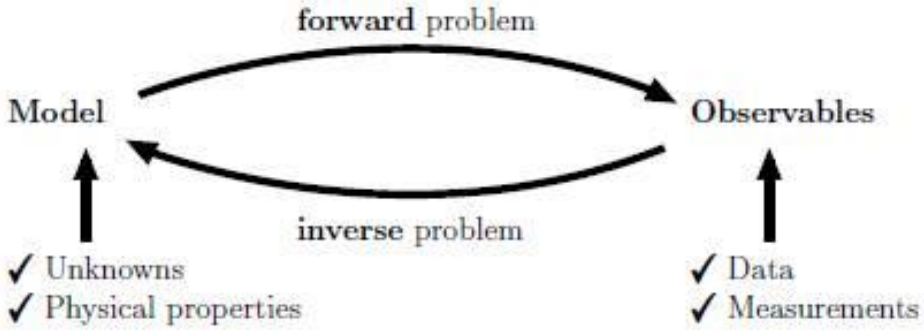

Fig. 4.3 - Definition of an inversion problem (from Wathelet, 2005).

The strategies aimed to perform the research of the misfit function minimum are divided in two main classes: Local Search Method and Global Search Method. This difference is due to the fact that, in general, in non linear problems the misfit function is characterized by multiple minima: among these, there will be an absolute minimum, called the global minimum, that is related to the best model. Nevertheless, there will be also several other relative minima in *M* function, called local minima, related to models more or less similar to the best one.

Starting from an assumption of initial model, the local search methods are iterative schemes which implement a sequence of perturbations to the parameters that allow to converge to a minimum point. In non linear cases these techniques advance by successive linearizations and are defined as Linearized Methods (LIN) (Menke, 1989; Parker, 1984; Tarantola, 2005). These methods require a smooth misfit function such as to be differentiable with respect to the various parameters of the model. However, even if this condition is accomplished, the convergence to the minimum occurs only if the initial model is close enough to the solution. Moreover, even when these procedures are able to converge to a minimum point, it is not possible to realize if it is the absolute minimum.

This limitation is coped by the Global Search Methods, where the whole solutions space is explored. These procedures are more robust and stable than the local ones, but more troublesome from the computational point of view. One of the most widespread global procedures is the Monte Carlo method (MC), developed in the 50s (Metropolis and Ulam, 1949; Press, 1968; Tarantola, 2005).

Sambridge and Mosegaard (2002) proposed to classify different Local and Global Search Methods on the base of the trade-off between “exploration” and “exploitation” criteria (Fig. 4.4). “Exploration” means that one tries to minimize the objective function by looking (randomly) in different regions of the parameter space, without considering what has been already learned from previous samples (Picozzi, 2005). “Exploitation” criterion has the opposite meaning: in fact, it means that one decides where to sample next by using the previous samples or the current best fit

sample only (Picozzi, 2005). Sambridge and Mosegaard (2002) made a comparison and classified various inversion approaches on the basis of the efforts made for satisfying the two criteria. In general, the more explorative methods (e.g., Uniform Search, Genetic Algorithms, Simulated Annealing, Neighbourhood Algorithm, Importance Sampling) are less efficient to converge on a solution, but is less likely that they fall into local minima. On the other hand, the exploitative algorithms (e.g., Steepest Descent, Newton-Raphson, Amoeba Search) are generally more efficient at converging, but are inclined to be entrapped in local minima.

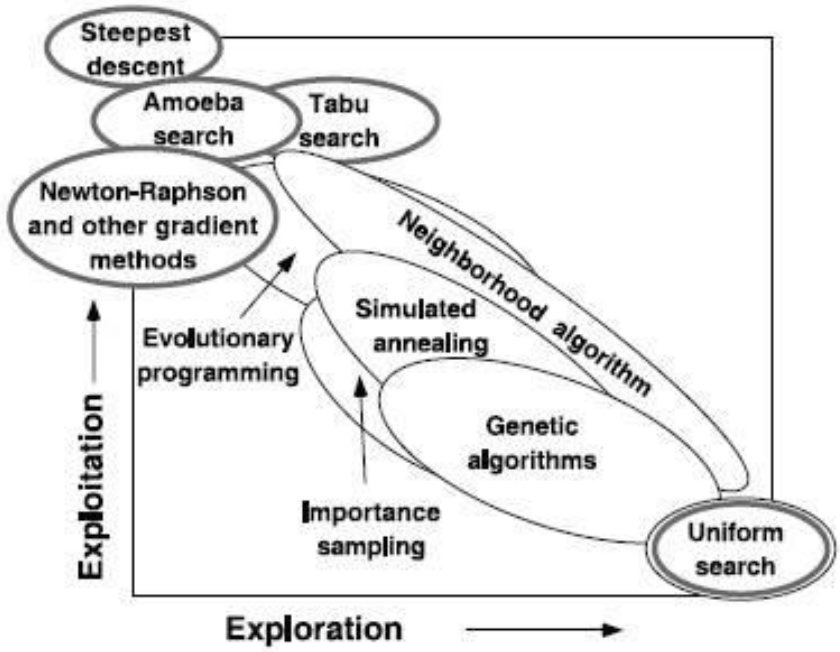

Fig. 4.4 - A schematic representation of various search/optimization algorithms in terms of the degrees to which they explore the parameter space and exploit information. Shaded borders indicate a deterministic (non-Monte Carlo) method. Uniform search includes the deterministic grid search (from Sambridge and Mosegaard, 2002).

In conclusion, only by the knowledge of the nature and complexity of the data-model relationship of the problem (i.e., the degree of non-linearity), the number of unknowns, and the computational resources available, it will be possible to choose the most suitable inversion technique. Nevertheless, for highly non-linear problems with multiple minima/maxima in the objective function, like for example the inversion of surface wave dispersion curves where various models might explain the data set with equal misfit, a combination of exploration and exploitation should be more appropriate (Picozzi, 2005).

The inversion of the Rayleigh wave velocity dispersion curve (obtained both by active and passive measurements) is dealt in several studies, where different procedures are used:

- LIN methods: e.g. Dorman and Ewing, 1962; Horike, 1985; Tokimatsu *et al.*, 1992b; Tokimatsu, 1997; Lai, 1998; Xia *et al.*, 1999; Herrmann, 2002; Ohori *et al.*, 2002; Arai and Tokimatsu, 2005.

- MC methods:
  - Simulated Annealing: e.g. Martinez *et al.*, 1999, 2000.
  - Genetic Algorithm: e.g. Yamanaka and Ishida, 1996; Parolai *et al.*, 2005; Picozzi *et al.*, 2005b.
  - Neighbourhood Algorithm: e.g. Wathelet *et al.* 2004, Wathelet, 2005; Wathelet *et al.*, 2005; Di Giulio *et al.*, 2012; Hobiger *et al.*, 2013.
- Hybrid strategy (MC and LIN methods). These strategies can be useful to face highly non linear problems. In particular, the use of global method allows an inversion analysis that does not depend upon an explicit starting model; the best-fitting model of MC is then used as the starting model for the LIN procedure that is able to drive the inversion to the global optimal minimum of the cost function. Two kinds of strategies has been used:
  - Simulated Annealing and LIN: Liu *et al.*, 2005.
  - Genetic Algorithm and LIN: Picozzi and Albarello, 2007.

4.3.1 Joint inversion

The problem of the non uniqueness of the solution that affects the Rayleigh wave dispersion curve inversion can be relieved by means a joint approach, i.e. taking in account also the HVSR curve.
Scherbaum *et al.* (2003) showed that the H/V ratio and the surface wave dispersion curves are characterized by different sensitivity to the S-wave velocity and thickness of the sedimentary layers. In particular, the dispersion curve represents the main constraint in the definition of the S-wave velocity of the soft sediments, while the fundamental frequency estimated from the H/V peak constrains mainly the total thickness of the sediment cover (Foti *et al.*, 2011). Hence, if an inversion procedure is applied to these curves separately, there is an un-resolvable trade-off between the model parameters, which hampers the analysis results. To overcome this problem, Parolai *et al.* (2005) and Arai and Tokimatsu (2005) proposed a joint inversion of Rayleigh wave phase velocity and HVSR curves (Fig. 4.5). They showed that with this approach, the trade-off between the model parameters can be reduced and a reliable evaluation of the local S-wave velocity structure can be obtained (Foti *et al.*, 2011).
On the other hand, a difficulty that arises from this procedure is the higher complexity of the joint misfit function, due to the even more non-linear contribution provided from the HVSR curve. This drawback can be face with good outcomes using an inversion procedure based on a Global Search method.
In the following sections, two different Global Search algorithms generally used to perform the joint inversion are described:

- the Genetic Algorithms (GA), adopted in the framework of this thesis and in recent PhD dissertation (Picozzi, 2005; Baliva, 2006; Pileggi, 2013) as well as in several research studies (e.g., Albarello *et al.*, 2011; Pileggi *et al.*, 2011; Albarello *et al.*, 2016).

- Neighborhood Algorithms (NA), which are becoming a widespread inversion method in Seismic Microzonation studies due to their implementation in the free software GEOPSY (www.geopsy.org).

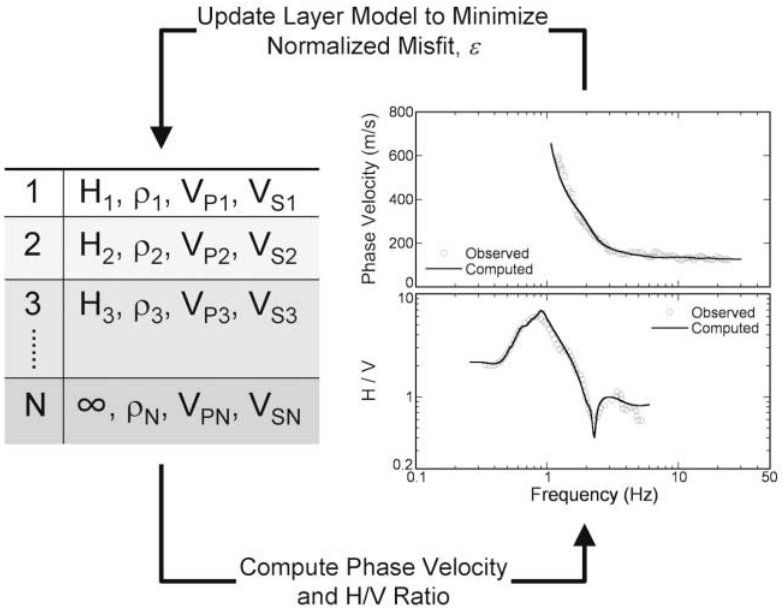

Fig. 4.5 - Schematic diagram showing the joint inversion using both dispersion curve and H/V ratio. H, $\boldsymbol{\rho}$, V_P and V_S represent respectively the thickness, density, P-wave and S-wave velocity of the model layers (from Arai and Tokimatsu, 2005).

4.3.2 Genetic Algorithms (GA)

Genetic Algorithms (GA) were originally devised by Holland (1975) as a model of adaptation in artificial systems. GA use random numbers to control the characteristics of the search process: this feature is typical of MC techniques. Several strategies were proposed to perform the inversion process with GA and, in general, an ensemble of specific operators is created for each particular application.

The basic idea of GA is to mimic the natural optimization process of biological evolution. Indeed, processes based on evolution are particularly suitable for solving computational problems, and especially those that require to find a best solution among a huge number of possible alternatives. Moreover, GA require only the computation of the objective function to be minimized: for this reason, they are inherently stable and robust procedures.

In seismology field, several papers dealing with GA have been produced in the early 1990s (e.g., Goldberg, 1989; Stoffa and Sen, 1991; Sambridge and Drijkoningen, 1992). Starting from this period, these inversion algorithms have been applied widely.

The basic structure of GA is summarized in Fig. 4.6. They use stochastic processes to produce an initial population of models from a defined search area (parameter space) and then, those able to

satisfy some particular criteria (objective function) are selected (selection operation) and used as parents for the generation of offspring (a new population of models). This generation stage happens generally by means of crossover and mutation operations (Picozzi, 2005).
Finally, the process described is iterated until a suitable model or group of models evolve and the optimization criteria are satisfied.

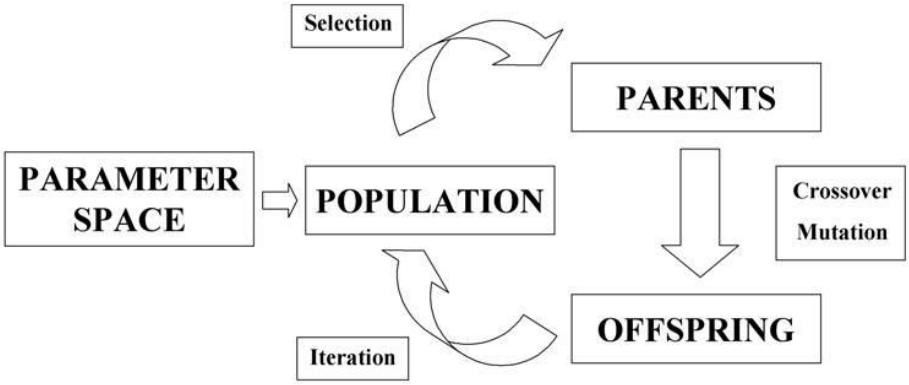

Fig. 4.6 - Schematization of the basic structure of GA (Picozzi, 2005).

The first stage of building any evolutionary algorithm is to define a suitable genetic representation of a candidate solution to the problem. Several approaches have been proposed for this step of the algorithm. One of the most common one is to adopt the binary representation, where each parameter in the search area is digitized with a *n*-bit binary string of 2^n. In particular, parameters at the lower and upper limits of the search area become (00…0) and (11…1), respectively and any other parameter to be examined is transferred to a binary string between these two limit strings (Picozzi, 2005). Subsequently, similarly again to the genetic operation, the binary strings relevant to all the parameters forming the model are concatenated in order to generate a "chromosome" to be used in the genetic operation.
After this first step, a set of models (*Q*) has been randomly extracted from the parameter space in order to assess their ability to explain the observed data. However, it is not merely sufficient to define which the better models are, but it is necessary to quantify how much one model is better than another. Thus, a fitness function (f_j) must be evaluated for each model *j*. To this purpose, the differences between observed (d_0) and theoretical (d_t) data are evaluated by the misfit function (ϕ) (Yamanaka and Ishida, 1996):

$$\phi_j = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{(d_0)_i-(d_0)_i}{\sigma(d_0)_i}\right]^2, \tag{4.9}$$

where *N* is the number of data and $\sigma(d_0)_i$ is the standard deviation associated to the observed data $(d_0)_i$. Thus, the fitness function is defined as (Yamanaka and Ishida, 1996):

$$f_j = \frac{1}{\phi_j}. \tag{4.10}$$

After this estimation, the parent selection operation is carried out in order to define which model should produce the offspring of the new population. In particular, models with higher values of f_j (or equivalently lower ϕ_j) have higher probability of proceeding to the next generation producing a set of *Q* new models (Gallagher and Sambridge, 1994). This stage is performed by means of the roulette rule (Goldberg, 1989): roughly speaking, it can be seen as the spinning of a one-armed roulette wheel where the sizes of holes are proportional to the probability of each model of being reproduced (Eiben and Smith, 2003). Using this approach, the population size does not change and multiple copies of models with higher fit will survive at the expense of those with lower fit, which could be extinguished completely. In conclusion, this operator introduces the element of survival of the fittest into the algorithm (Sambridge and Mosegaard, 2002).

These reproduced models were modulated by the crossover and mutation operations.

The crossover process allows to merge the chromosome structure relevant to the parents, by the exchange and recombination of each parent structure (Gallagher and Sambridge, 1994). In other words, the offspring obtained are a combination of the parents' characteristics and might develop some different features. Usually, this operator starts from two parents and creates two offspring, although in the case of specific applications any number of parents can be used (Eiben and Smith, 2003).

In practice, a threshold called "crossover probability", p_c, is determined. Then, all the *Q* parent strings are randomly paired and a random number ranging from 0.0 to 1.0 is generated. At this stage, if the random number is smaller than p_c, one of the forms of crossover takes place and the parents interchange a part of each bit string creating two new offspring (Fig. 4.7). Otherwise, if the random number is greater than p_c, the two individuals in the original pair remain in the next generation (Fig. 4.7) (Picozzi, 2005).

Mutation is instead a random process that allows the introduction of new characteristics unrelated to the parents into the offspring. It is a particular operator that considers each gene separately and has been introduced in GA with the aim to maintain a degree of diversity in the population. Also in this case a threshold value of mutation probability, p_m, is defined. Then, a series of random numbers ranging from 0.0 to 1.0 is generated and assigned to each gene: if these numbers are smaller than p_m, a bit in an individual string will flip from 0 to 1 and vice versa (Fig. 4.7) (Picozzi, 2005). In general the selected value of p_m is relatively low.

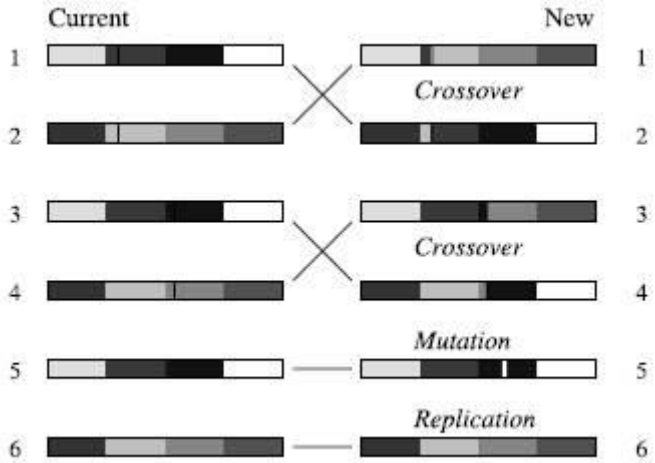

Fig. 4.7 - The effect of crossover and mutation operators in a genetic algorithm. In cases 1–4, the crossover operator cuts, swaps, and rejoins pairs of strings. In case 5, mutation changes a single bit. In case 6 a whole string is replicated. The shading represents the values of four different variables encoded into the bit string (from Sambridge and Mosegaard, 2002).

Yamanaka and Ishida (1996) observed that using the simple GA an individual with the lowest misfit could not be selected for the next generation. Moreover, with this inversion scheme the crossover might also destroy the chromosome of the best individual. Therefore, to overcome these problems these authors proposed a modified version of GA taking advantage of two further operators: the elite selection and the dynamic mutation.

The elite selection operator has been introduced in GA with the goal to ensure that the best model of a generation will be present also in the next one. For that reason, it simply replaces the worst model in the current generation with the best individual of the previous generation (Goldberg, 1989).

Despite the advantages, the use of the elite selection might cause a premature convergence towards the model introduced by this operator that has a local minimum of the misfit. To avoid the risk of a bias in the inversion procedure, a new kind of mutation operator was introduced. It is called dynamic mutation, and is a generation-variant mutation probability that may help the inversion avoiding premature convergence (Sen and Stoffa, 1992). In practice, when most of the models in the current population have a similar fit and chromosome patterns, the parameter p_m is increased. After the variety of the chromosomes in the population is assured, p_m assumes again the initial value (Picozzi, 2005). The average coefficient of variation (γ) is evaluated for all parameters in all individuals of each generation as

$$\gamma = \frac{1}{M}\sum_{i=1}^{M}\left(\frac{\sigma_i}{\bar{x}_i}\right), \tag{4.11}$$

where M is the number of parameters, $\bar{x}_i$ is the average of the i^{th} parameter, and σ_i is the standard deviation (Yamanaka and Ishida, 1996). According to γ, the mutation probability p_m is defined as

$$p_m = \begin{cases} p_i, \ \gamma > 0.1 \\ 0.1, \ 0.02 < \gamma < 0.1 \\ 0.2, \ \gamma < 0.02 \end{cases}, \tag{4.12}$$

where p_i is the initial mutation probability.
After these adaptations, Yamanaka and Ishida (1996) showed that the modified GA found a better solution than the simple GA, as indicated by the lower misfit value at the end of the inversion process (Fig. 4.8).

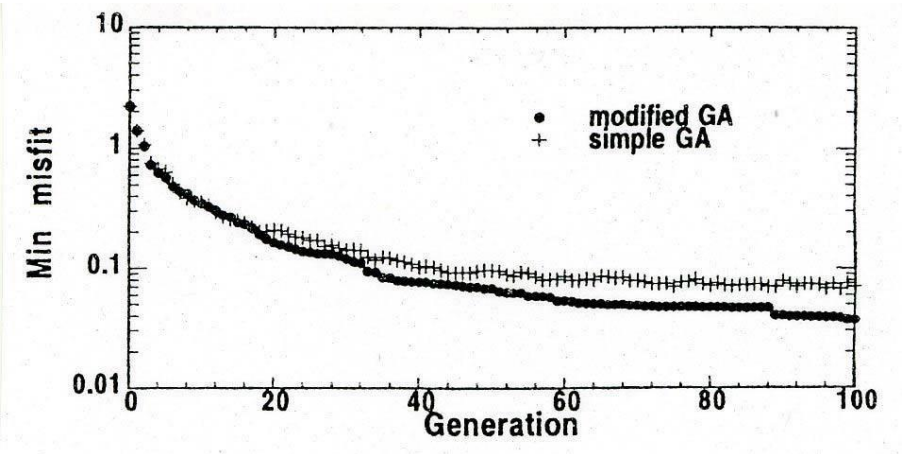

Fig. 4.8 - Minimum misfit value as a function of generation for the inversion of synthetic Lowe-wave group velocities (from Yamanaka and Ishida, 1996).

4.3.3 Neighborhood Algorithms (NA)

The Neighborhood Algorithm (NA), proposed by Sambridge (1999), is a stochastic direct search method for non-linear inversion. Similar to the GA, this technique tries to guide the random generation of samples by the results obtained so far on previous samples. In order to model the misfit function across the parameter space, NA makes use of the Voronoi cells, i.e. a unique decomposition of the space into n cells around n points p_i. The cell around point p_i is defined by the ensemble of points that are closer to p_i than to whatever other point p_j where $j \neq i$.
The misfit function is supposed to be known for n_{s0} samples randomly distributed with a uniform probability inside the parameter space at the beginning of the inversion. A Voronoi cell centered

around one of these samples is the nearest neighbor region defined under a suitable distance norm (usually Euclidean). The union of all cells with a low misfit is the area of interest where new samples with small misfits are expected: the size of this ensemble is defined by the tuning parameter n_r, i.e. the number of best cells to consider (Wathelet, 2008). The second tuning parameter, n_s, denotes the number of the new samples that are generated and added to the original population: in particular, n_s/n_r samples per cell are added. The geometry of the initial Voronoi cells is modified to include these new n_s samples. The process above described is repeated it_{max} (last tuning parameter) times until an acceptable sampling of the solution is obtained (Wathelet, 2008).

Fig. 4.9-a is an example of a two-dimensional parameter showing the models (black dots) and the limits of the Voronoi cells. Random n_{s0} models (nine in this case) are generated and the grey cell has the lowest misfit. In this example, seven new models are generated in one cell (n_r = 1, and n_s = 7). Fig. 4.9-b depicts the Voronoi geometry after the first iteration. It is possible to note that the size of the original cell decreases as the sampling rate increases. If the cell with the grey outline has the lowest misfit, the density of sampling will not decrease systematically after each iteration. This is an interesting property of the Voronoi geometry that allows the center of sampling to jump from place to place, whilst always sampling the most promising n_r regions simultaneously (Wathelet, 2005).

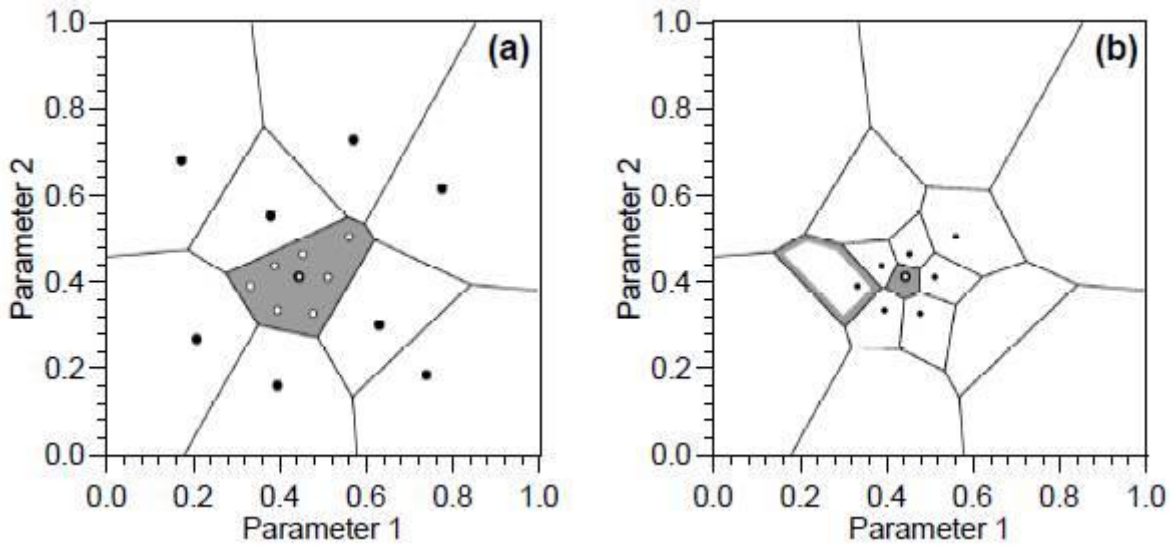

Fig. 4.9 - Example of Voronoi cells for a two-dimensional parameter space (from Sambridge, 1999).

Sambridge (1999) proposed a simple but very efficient way to generate new random samples inside a Voronoi cell based on a Gibbs sampler (a Markov chain algorithm), i.e. a random walk. A walk is a sequence of perturbations to a model along all axes. The modified model is statistically independent of the original model. Asymptotically, the samples produced by this walk will be uniformly distributed inside the cell regardless of its shape. To confine the random walk inside a

particular cell it is mandatory to calculate the multi-dimensional limits of the cell. Calculating the complete Voronoi geometry for high dimensional spaces becomes practically impossible when the number of models increases (Wathelet, 2005). Sambridge (1999) proposed an original algorithm to compute only the limits along lines which are parallel to axis: these lines support the successive segments of the random walks.

As concerns the control parameters, the NA is more exploratory if the n_s new samples are distributed on many cells and it optimizes more if they are restricted to the very few best cells. Typical values for the tuning parameters are 100 for n_{s0}, n_s, n_r. To force a better optimization, n_r may be set to 5, 10 or 50. Tests show that generally better misfits are obtained with less iterations (low it_{max}) if n_r is low, but the inversion is more trapped in local minima (Wathelet, 2005).

The more exploratory mode (e.g. n_r=100 and n_s=100) usually provides better final misfits if the inversion is conducted with a great number of iterations, e.g. from 50 to 200. Since with these settings the procedure makes a total of 5000 to 20000 generated models, highly optimized forward computations are required in order to reduce the total computation time of the inversion process.

In its original form, the Neighborhood Algorithms handles a parameter space with orthogonal boundaries, where all parameters have a uniform probability within prior fixed limits. Wathelet (2005) and Wathelet (2008) proposed an improvement of the classical NA form considering the parameter space bounded by a hyper-box (classical limits) and by irregular limits, due to physical conditions, numerical limitations or prior information. By contrast with the hyper-box, the irregular limits may have no explicit definition, in fact they are settled by a list of conditions that links several parameters together (e.g., the value of the parameter p_1 must be lower than the parameter p_2).

In this case, the misfit computation is possible only inside the intersection of these two ensembles. Some problems can be arise: for instance, when the active region (i.e., the union of all best n_r cells) is close to one of the complex boundaries, Voronoi cells where new samples are generated can be cut by one of them and only a small percentage of their multi-dimensional volume may be included inside the valid region (e.g. cell *l* in Fig. 4.10). Thus, there might be only very little chance to generate one good sample even after a lot of trials (Wathelet, 2008). A way of solving this problem is to count all accepted and rejected models per cell: if the proportion of rejected models exceed a threshold (e.g. 90%), the cell is thrown away from the active region and replaced by the cell with the best misfit currently outside the active region (Wathelet, 2008).

Moreover, when there are a lot of conditions to satisfy, the original random generator is not very efficient and a lot of invalid models must be rejected before accepting just one. To overcome this drawback, Wathelet (2008) proposed a modified version of the Gibbs sampler in order to always return a valid model. In particular, assuming that at least one model has successfully passed all conditions (model *A* in Fig. 4.10), the parameter *i* can take any value from x_j to x_l (according to the original NA) in order to stay within cell *k* (Fig. 4.10). To fulfill the complex conditions x_l is replaced by x_b (Fig. 4.10), which is computed exactly by the intersection of all admissible ranges given by all

conditions available for parameter *i* keeping all other parameters constant. Hence, model *A* can be perturbed along axis *i* and the obtained model *B* is also satisfying all conditions (Fig. 4.10). The process is repeated for all axes as in the original algorithm (Wathelet, 2008).

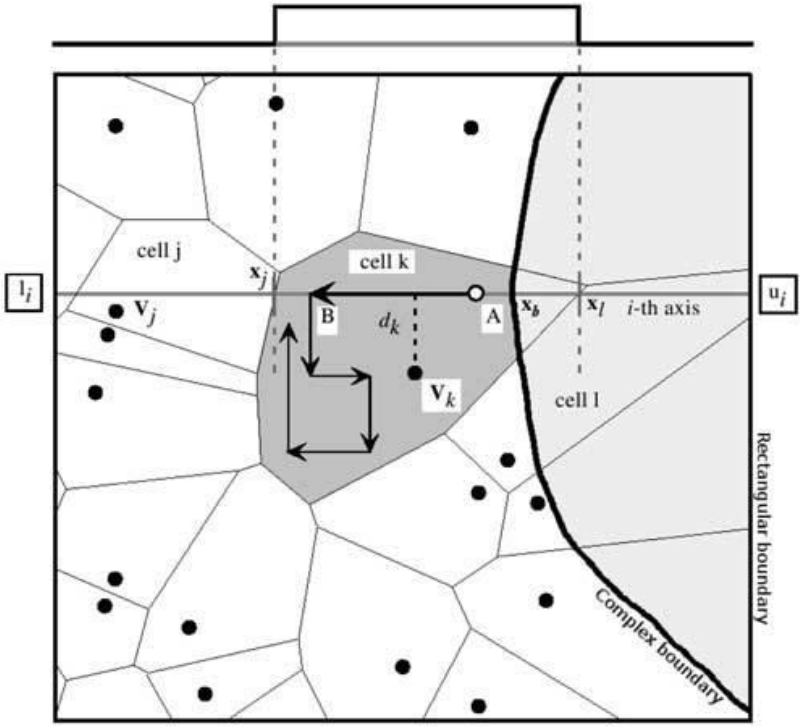

Fig. 4.10 - A uniform random walk restricted to a Voronoi cell and by complex boundaries. Starting from a sample inside the cell *A*, a Markov-Chain random walk is achieved by introducing random perturbations along all axes successively. Each random perturbation (for instance along axis *i*) is bounded by the rectangular boundary (l_i and u_i), by the limits of the Voronoi cells (x_j and x_l) and by the intersection of axis *I* passing by *A* with the complex boundary (x_b). The light grayed area is the region outside the parameter space still inside the rectangular boundary (from Wathelet, 2008).

CHAPTER 5
Influence of shear wave velocity inversion on passive seismic measurements: the case of Malta

Based on the paper published by Geophysical Journal International (Farrugia et al., 2016)

5.1 Introduction

One of the aims of this dissertation is to verify and evaluate the reliability of the passive seismic techniques in the contexts characterized by complex geological settings, i.e. situations not related with simple 1-D scheme where the S-wave values increase with depth. In particular, this chapter is focalized on the effects due to a sharp S-wave velocity inversion in the subsoil on the HVSR and Rayleigh wave phase velocity dispersion curves. Moreover, the reliability of the joint inversion procedures in the estimating of V_S profile in this particular context was explored.

This study is characterized by an experimental nature and the investigated sites are located in the northwestern part of the Malta's Archipelago. In this area, a layer of clays and marls, known as Blue Clay (BC), may exceed 70 m in thickness and underlies the topmost, youngest layer of the sedimentary sequence, a hard reef limestone varying in thickness from a thin covering of a few meters to more than 80 m.

The presence of a shallow thick low-velocity layer is not so uncommon, especially in sedimentary environments encompassing clay deposits. However, studies of the effects of such stratigraphies on ambient vibration measurements and multimode inversion procedures as well as on V_{S30} estimate are limited. A literature review concerning this argument was treated previously in Chapter 2 and 3.

Previous studies on the islands have yielded shear-wave velocity profiles at a limited number of sites using the active Multichannel Analysis of Surface Waves (MASW) technique (Panzera *et al.,* 2013). However, the depth resolution of these studies was limited to around 30 m, and the low-velocity clay layer was not adequately sampled in most of the cases. Other models of seismic-wave velocity profiles were made by inverting single-station H/V data in one area along the NE coast of the islands characterized by the same shear-wave velocity inversion (Pace *et al.,* 2011; Panzera *et al*., 2012; Vella *et al.*, 2013).

Hence, this work is also aimed to improve the information obtained by the previous investigations by extending the number of measurement sites as well as by using more robust acquisition and inversion methods. In particular, the multi-station measurements are analyzed with the ESAC technique and the obtained Rayleigh wave phase velocity dispersion curve has been used jointly with the Horizontal-to-Vertical Spectral Ratio to infer one-dimensional V_S profiles at seven sites with the stratigraphy explained above, i.e. with the presence of a thick, buried low-velocity layer, on the two principal islands. The investigated sites were chosen in order to have a reliable

comparison with the achieved data and the real geological configuration, which is clearly visible in outcropping and easily deduced by using the geological map.
This study was accomplished with the collaboration of Department of Geosciences of Malta University and it was published by Geophysical Journal International (Farrugia *et al.*, 2016).

5.2 Geological settings

Located in the Central Mediterranean, the Maltese archipelago comprises three main islands (Malta, Gozo and Comino) having a total area of around 320 Km^2 (Fig. 5.1-a). These islands were formed as marine sediments during the Oligocene and Miocene epochs. The geology, that reflects abrupt changes in the depositional environment during this period, consists of four main strata of lime-rich sedimentary rocks; the lime content being present mainly due to the fossil shells of animals and plants which are found in abundance in the strata (Pedley *et al*., 2002 and references therein). Starting from the oldest, the sequence consists of (Vella *et al.*, 2013; Fig. 5.1-a,b):

- the Lower Coralline Limestone (LCL) formation (Chattian). This is the oldest unit of rock visible above sea level. It is a hard and compact pale grey limestone, rich in coralline algae, and can be over 140 m thick. The LCL forms the sheer cliffs bordering the southwestern coast of the Maltese islands;
- the Globigerina Limestone (GL) formation (Aquitanian-Langhian). This layer is made up of soft yellowish fine grained limestone. Its thickness varies between 20 and 200 m. The fine grained GL is lightly cemented making this layer soft and easily cut, sculpted and eroded. It is subdivided into three sublayers: the Lower, the Middle and the Upper Globigerina Limestone. These three layers are separated by two thin hardground conglomerate layers, which are much harder and more resistant to erosion, and easily identified when exposed;
- the Blue Clay (BC) formation (Serravallian to early Tortonian). This layer can be considered as a continuation of the Globigerina Limestone being composed of very fine grained sediment and containing deep water planktonic organisms. The main difference is the higher clay mineral content which gives this layer the banded bluish grey color. The mix of clay between the lime rich sedimentary grains prevent the binding of the particles hence it is the softest layer of all and is eroded very quickly with rain water. Very often it provokes landslides and in various parts of the islands it forms 45° talus slopes over the underlying strata. Its depth varies between 70 and 20 m, thinning notably from west to east of the archipelago (Pedley *et al.*, 1976). Above the Blue Clay Formation is a thin layer of bioclastic limestones named the Greensand Formation (Mgg). This layer varies from 1 to 11 m in depth and is present only in a limited number of areas of the islands;
- the Upper Coralline Limestone (UCL) formation (late Tortonian-early Messinian). This is the uppermost and youngest layer and it was formed in very similar conditions to the Lower Coralline Limestone, in shallow waters. This hard pale grey layer can reach thicknesses of

more than 80 m in certain parts of the islands. As with the LCL, the layer presents a variety of facies ranging from reef limestones to cross-bedded sands and fine-grained muds. The UCL consistency is also highly variable, ranging from fractured and friable to highly compact.

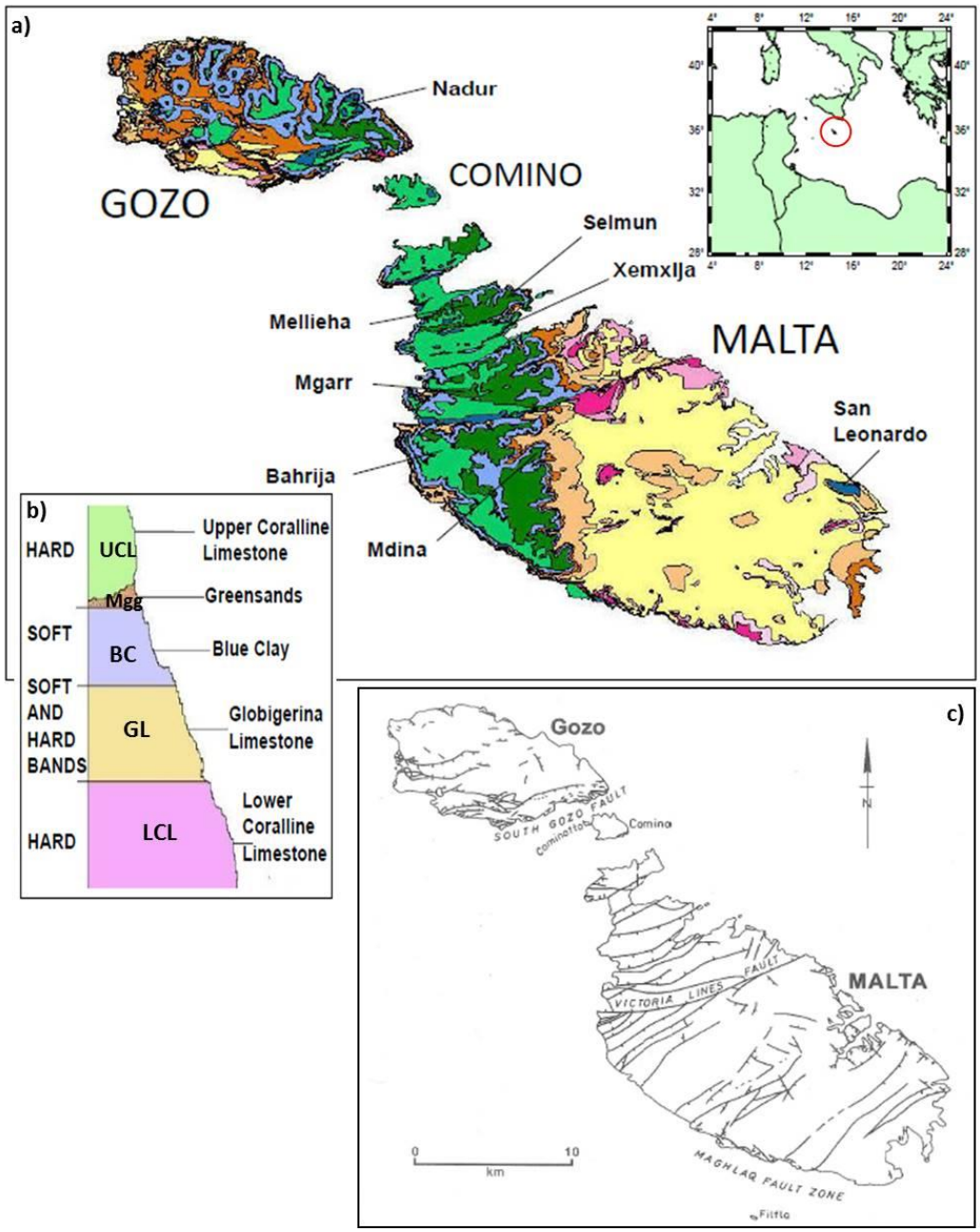

Fig. 5.1 - a) The geology map of the Maltese islands together with the location of the chosen sites for investigation. The different shades of violet, yellow and green correspond to the different members of the LCL, GL and UCL respectively. Inset shows the position of the Maltese islands (red circle) in the Mediterranean Sea (modified from Farrugia *et al.*, 2016). b) A sketch of the full sedimentary sequence making up the islands above sea level. c) Structural map of the Maltese Islands showing the pattern of the major faults (from Pedley *et al.*, 2002).

No marine sediments younger than the UCL are found on the islands, indicating that the time of uplift of the islands above sea level together with the northern flank of the Sicily Channel Rift (Pedley *et al.*, 2002) occurred around 6 Ma ago. This uplift coincided with the closure and partial desiccation of the Mediterranean during the Messinian. Some patchy deposits of the Pleistocene epoch containing remains of terrestrial fauna were laid down on land by fluvial action and are mostly found near shorelines and in caves (Vella *et al.*, 2013).

Tectonic processes, accompanied by differential weathering and erosion have largely controlled the relief and landscape of the Maltese islands, which exhibit an intricate pattern of topography. Tilting toward the northeast is mainly due to the uplift of the islands on the shoulder of the graben system (Illies, 1981). The faulting pattern is due to the different regional stress phases in the Central Mediterranean. An older NE–SW trending set of extensional block faults divides the central part of the archipelago into a system of horsts and grabens, giving it its characteristic ridge–valley topography. The master faults of this system are the Victoria Lines Fault that bisects Malta from SW to NE coast, and the South Gozo Fault (Fig. 5.1-c). Along the sub-vertical sides of these valleys, the sedimentary sequence is exposed and subjected to differential weathering. In particular, some ridges and hills composed of UCL resting on an open-sided layer of BC are rendered quite unstable as a result of the plasticity of the clay (Vella *et al.*, 2013). A more recent NW–SE trending fault pattern runs parallel to the Sicily Channel Rift graben-bounding faults; the main fault of this system (Maghlaq Fault) running along the southern coastline of the islands (Fig. 5.1-c) and exhibiting a vertical throw of some 240 m to the seaward side.

As can be seen from the geological map of the islands (Fig. 5.1-a), in the eastern half of Malta the two youngest layers (BC, UCL) are missing and the area is characterized mostly by outcropping GL, giving rise to a flat, rolling landscape in this part of the islands. On the other hand, the western half of Malta and some areas in Gozo retain the full sedimentary sequence (Pedley *et al*., 2002), with UCL hillcaps and plateaus overtopping BC gentle slopes being a dominant feature in the landscape of north-western Malta and north-eastern Gozo (Gigli *et al*., 2012;). Throughout much of this area, the geomorphology is mostly dictated by the fact that the UCL forms a rigid rock slab resting on a much weaker layer of clayey material (Fig. 5.2-A). This geological situation creates stresses in the upper slab resulting in fracturing brittle collapse and lateral spreading phenomena (Gigli *et al.*, 2012). When the exposure of the geological cross-section is along the coast, the additional weathering effect of marine processes accelerates the destabilization by erosion of the clay layer. This results in the formation of large cliff-parallel surface fractures (Fig. 5.2-B) that produce partially isolated blocks having volumes of the order of thousands of cubic meters (Fig. 5.2-C). Horizontal and vertical mass movement of such blocks forming part of the cliff face, and boulder detachment and collapse, result in a fractured and boulder-strewn coastline (Fig. 5.2-B,C), that is typical of much of the northwest coast of Malta (Galea *et al.*, 2014).

Panzera *et al.* (2012) and Galea *et al.* (2014) investigated the site dynamic characteristics of this kind of areas by means ambient vibration single station measurements analyzed with HVSR and frequency-dependent polarization (e.g., Burjanek *et al.*, 2010, 2012) technique. In particular, with

this latter analysis, the authors showed that the particle motion associated with the higher frequencies is characterized by strongly directional polarization generally normal to the cliff edge or to the large scale fractures. They conclude stating that this feature is probably related with the predominant directions of detached blocks vibrational effects.

Fig. 5.2 - Geological features of the cliffs near Selmun area (northwest coast of Malta). Typical landscape of this part of the island, where almost the whole geological succession is clearly visible (A). Large cliff-parallel surface fractures (B). Detachment and collapse of UCL blocks along the coast (C).

The Maltese archipelago is affected by low-to-moderate seismic hazard. Since 1530, at least four earthquakes of intensity VII or VII-VIII on the European Macroseismic Scale (EMS-98) were experienced, with the major contributor to the seismic hazard being the northern segment of the Malta Escarpment (Galea, 2007). Felt and damaging earthquakes can also be attributed to active fault zones in the Sicily Channel and the Hellenic Arc (e.g., Agius *et al.,* 2015).

5.3 Data acquisition and analysis

5.3.1 The investigated sites

In order to verify the reliability of the passive seismic techniques and to retrieve the V_S profile in presence of a thick buried low velocity layer, seven sites have been selected (six in Malta and one in Gozo, shown in Fig. 5.1-a). All of these location are characterized by the full sedimentary sequence, i.e. the Blue Clay is embedded between the Upper Coralline Limestone above and the Globigerina Limestone from below (see Fig. 5.2-A). By choosing sites of similar stratigraphy in different parts of the islands, we could also investigate any spatial geophysical variations within a particular stratum. The sites for array measurements were chosen to have no major topographical slopes or irregularities.

One of the chosen sites, Mdina, is the former capital city of Malta and it is a heavily urbanized fortified town built on a hill with a thin outcropping UCL layer (with a maximum thickness of around 6 m) protecting the erodible BC. A very thin layer (less than 1 m) of Greensand is also present (Gigli *et al.,* 2012). The city has previously suffered serious damage from major earthquake events, in particular, the Sicily Channel earthquake of 11th January 1693 (estimated moment magnitude of 7.4). This event produced an intensity of VII-VIII degree of the MCS scale, and various buildings suffered serious damage (Galea, 2007; Gigli *et al*., 2012). Other sites were in a more rural environment, but always close to inhabited areas. At each site, a number of single-station ambient noise measurements were conducted jointly with the array measurements.

No direct measurements like borehole seismic tests (e.g., down-hole and cross-hole) that can witness the effective presence of a V_S inversion between the two youngest geological formations exist in the selected sites and in this part of the Maltese Archipelago. As well as by the clearly different geological characteristics of the two lithologies, this feature is inferred by previous active (Panzera *et al.*, 2013) and passive (Farrugia *et al.*, 2015) seismic measurements carried out directly on BC: the achieved V_S values (about 300-400 m/s) are significantly lower than the ones retrieved by inverting the MASW curves (700-900 m/s) obtained directly on UCL (Panzera *et al.*, 2013).

In addition, measurements were also conducted on a small area in the SE of Malta, called San Leonardo (Fig. 5.1-a), which is the only known site on the islands where the UCL directly overlies the GL, i.e. the BC is not present in the geological sequence (Pedley, 2011; Zammit-Maempel, 1977). This area offers the best opportunity to validate the results obtained in the presence of the buried BC.

5.3.2 Passive seismic survey

The passive seismic survey was carried out by performing, in each of the above showed sites, a number of single station recordings located close to the deployed seismic array (Fig. 5.3).

Fig. 5.3 - Passive seismic survey at Bahrija (A) and Xemxija (B). The red circles in the sketch indicate the position of the investigated sites respect to the urban areas. The red dots show the position of each geophone of the seismic arrays while the Gxx at the corners of the two array configurations indicate the number of the geophone found in that corner. The green points show the location of the single-station acquisitions (modified from Farrugia *et al.*, 2016).

The single-station measurements were carried out using the three-directional digital tromograph Micromed Tromino™ (www.tromino.eu). Tromino is composed of three orthogonal velocity sensors and 24-bit digitizer, whose sampling frequency extends up to 1024 Hz. The ambient vibrations were acquired for 20 minutes with a sampling frequency of 128 Hz and the H/V curves were computed following the procedure described in Paragraph 2.3. In particular, the time series were divided into 60 non-overlapping windows, each 20 s long, as suggested by the SESAME guidelines (SESAME, 2004). Moreover, any window containing spurious signals was removed and the quality of the resulting HVSR curve was evaluated following the criteria described in Paragraph 2.4.

In order to determine whether the sampled area beneath the array consists of approximately homogeneous strata, various single-station acquisitions were carried out close to the array itself (Fig. 5.3). If all H/V curves peaked at the same frequency and had similar shapes, this was taken as an indication that the sediment cover was uniform over the extent of the array, satisfying one of the key assumptions of the array methods. An example of this is given in Fig. 5.4.

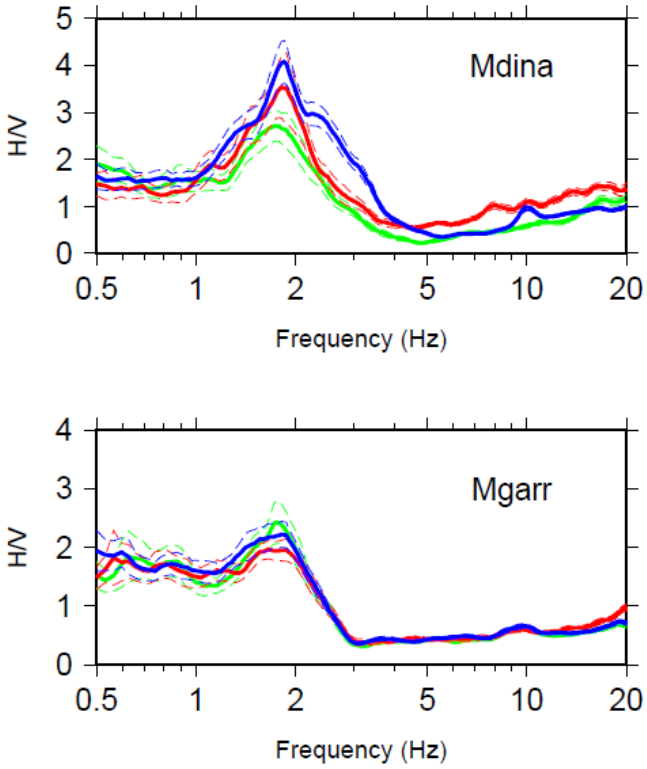

Fig. 5.4 - All the H/V curves obtained at Mdina and Mgarr site. The dashed lines represent their confidence intervals (from Farrugia *et al.*, 2016).

As concerns the seismic arrays, these measurements were conducted using Micromed SoilSpy Rosina™ seismic digital acquisition system equipped with 4.5 Hz vertical geophones. A total of 42 geophones were used, and placed in an L- or C-shaped configuration with a regular interstation distance of 5 m and maximum dimension of 120-150 m. At the Mdina and San Leonardo sites, only 17 geophones were employed because of space limitations; in these cases the maximum inter-station distance was about 70 m. Fig. 5.3 shows the deployment at Bahrija and Xemxija site. Since only vertical sensors were used, the signals detected are interpreted as plane Rayleigh waves in their fundamental and higher propagation modes. The recordings, each 20 minutes long and sampled at 256 Hz, were analyzed using the ESAC technique (see Section 3.2.1.2) in the form proposed by Parolai *et al.* (2006). As explained in Section 3.2.1.1 and 3.2.2.3, this approach allows the determination of the effective (or apparent) dispersion curve instead of the modal ones. This characteristic could represent an important limitation of this procedure with respect to other

approaches (e.g., *f-k* techniques) but, on the other hand, this makes the ESAC method more robust with respect to the alternative procedures, since it does not require troublesome picking of existing propagation modes (Albarello *et al.*, 2011).
Moreover, as mentioned in Section 3.2.2.3, the use of this method allows to monitor wavelengths even higher than 2-3 times the largest inter-geophonic distance. Because maximum depth of investigation is about 1/2-1/3 of the largest wavelength analyzed, the adopted array configurations allowed us to reach depths of up to 120-150 m for the largest arrays and 70 m approximately for the shortest ones.

5.4 Results and discussion

5.4.1 HVSR curves

Fig. 5.5-A shows the representative H/V curves obtained at the seven sites characterized by a velocity inversion together with the one obtained at San Leonardo. The former sites exhibit a peak between 1 and 2 Hz, the amplitude of which varies between 2 and 5. Other studies carried out on the islands (Pace *et al*., 2011; Panzera *et al.*, 2012; Panzera *et al.*, 2013; Vella *et al.,* 2013; Galea *et al.,* 2014), have also observed this ubiquitous peak wherever the soft BC is found underlying the compact UCL. This resonance is presumably associated with the boundary separating the BC and the GL (Panzera *et al*., 2012) and decreases in frequency with increasing depth of this boundary (Pace *et al*., 2011).
Moving to high frequency values, the peak is immediately followed by a drop below 1 in the H/V spectrum over a wide frequency range (Fig. 5.4 and 5.5). As explained in the described studies (Di Giacomo *et al*., 2005; Castellaro and Mulargia, 2009) in Paragraph 2.6, this feature can be attributed to the presence of a velocity inversion in the stratigraphy and is also evident and consistent in all previous studies of areas of similar lithostratigraphy on the islands.
As pointed out by Castellaro and Mulargia (2009), a drop of H/V below 1 does not only occur in the presence of low shear velocity layer in the subsoil. However, comparing the results obtained from extensive H/V surveys carried out on directly on the other formations of the Archipelago (Panzera *et al.*, 2013; Vella *et al.,* 2013), this characteristic appears more evident in the measurements achieved on UCL. This can be seen in Fig. 5.5-B, where a comparison between the H/V curves of Fig. 5.5-A and some representative H/V curves obtained on sites where the BC formation is outcropping is showed.
Finally, observing Fig. 5.5-A, it is possible to note that the curve obtained at San Leonardo site (where UCL outcrops but the BC layer is missing) contrasts with those where the BC is present as it is flat with no valid peaks above two units. Although the curve drops below 1, this is not observed for a wide range of frequencies. In view of this, it is possible to conclude that the characteristics of the previous curves may indeed be attributed to the presence of the BC layer.

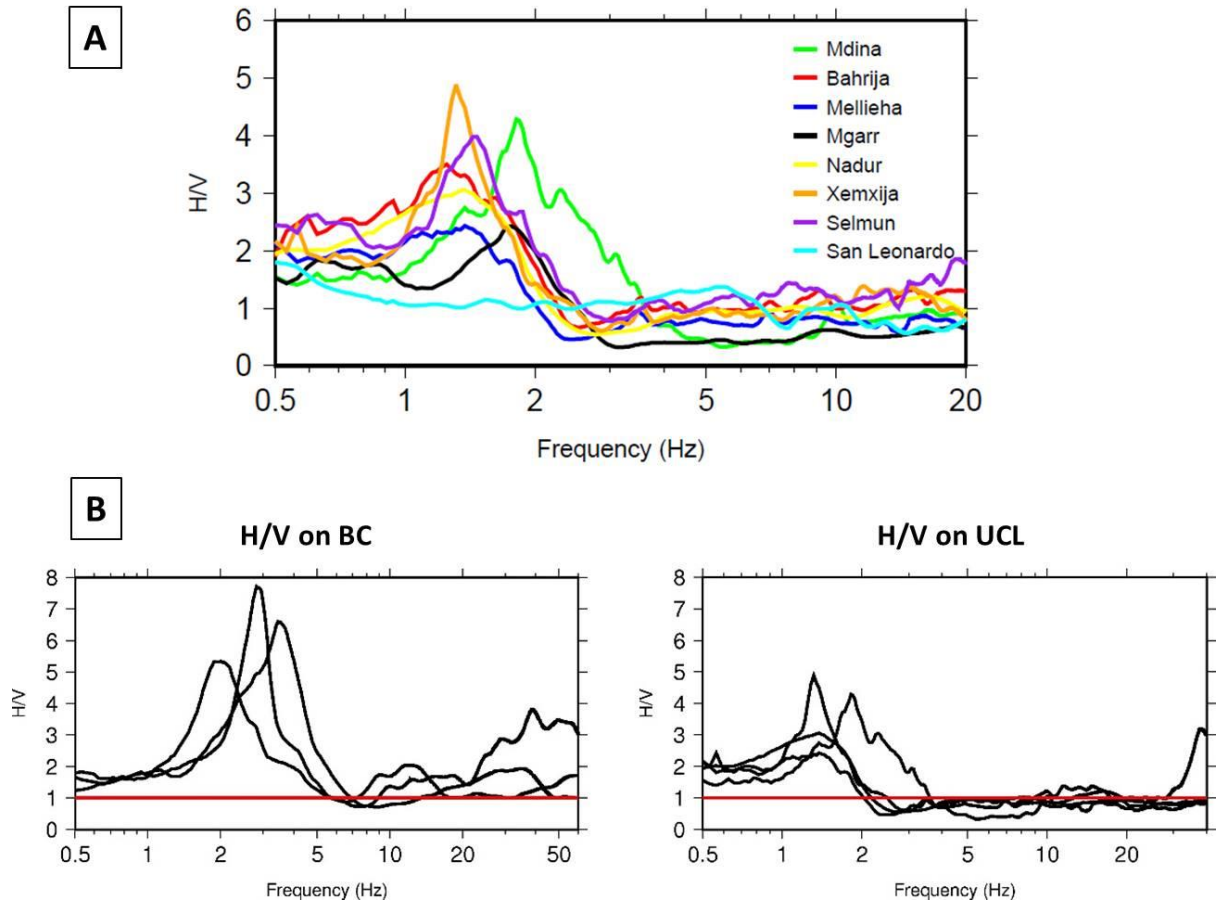

Fig. 5.5 - Panel A shows the representative H/V curves obtained at the seven sites characterized by a low velocity layer and one obtained at San Leonardo site (from Farrugia *et al.*, 2016). Panel B shows the comparison between some representative H/V curves obtained directly on BC from previous studies (Panzera *et al.*, 2013; Vella *et al.*, 2013) and the ones achieved on UCL depicted in Fig. 5.5-A.

5.4.2 Dispersion curves

Fig. 5.6 and Fig. 5.7 show the resulting effective dispersion curves obtained by ESAC analysis at each of the seven sites and San Leonardo. At lower frequencies, in the ranges that generally vary between 3-8 Hz and 3-5 Hz, the curves exhibit normal dispersion characteristics because the Rayleigh wave phase velocity decreases with increasing frequency. At higher frequencies, this trend changes to an overall inversely dispersive one, i.e. the velocity values increases with increasing frequency. As showed in Paragraph 3.3, various theoretical studies (e.g., Tokimatsu *et al.*, 1992a; Tokimatsu, 1997; Arai and Tokimatsu, 2005) attribute this latter dispersion curve shape to the presence of a stiff layer overlying a softer one: in view of this, this feature could further demonstrate that a shear wave velocity inversion occurs between the UCL and BC.

In Mellieha (Fig. 5.6), only an inversely dispersive curve was obtained suggesting that the combined UCL and BC layers are too thick for the GL to be adequately sampled with the given array configuration. Using an array with a larger aperture and geophones with a lower eigen-frequency, might improve the low-frequency part of the dispersion curve.

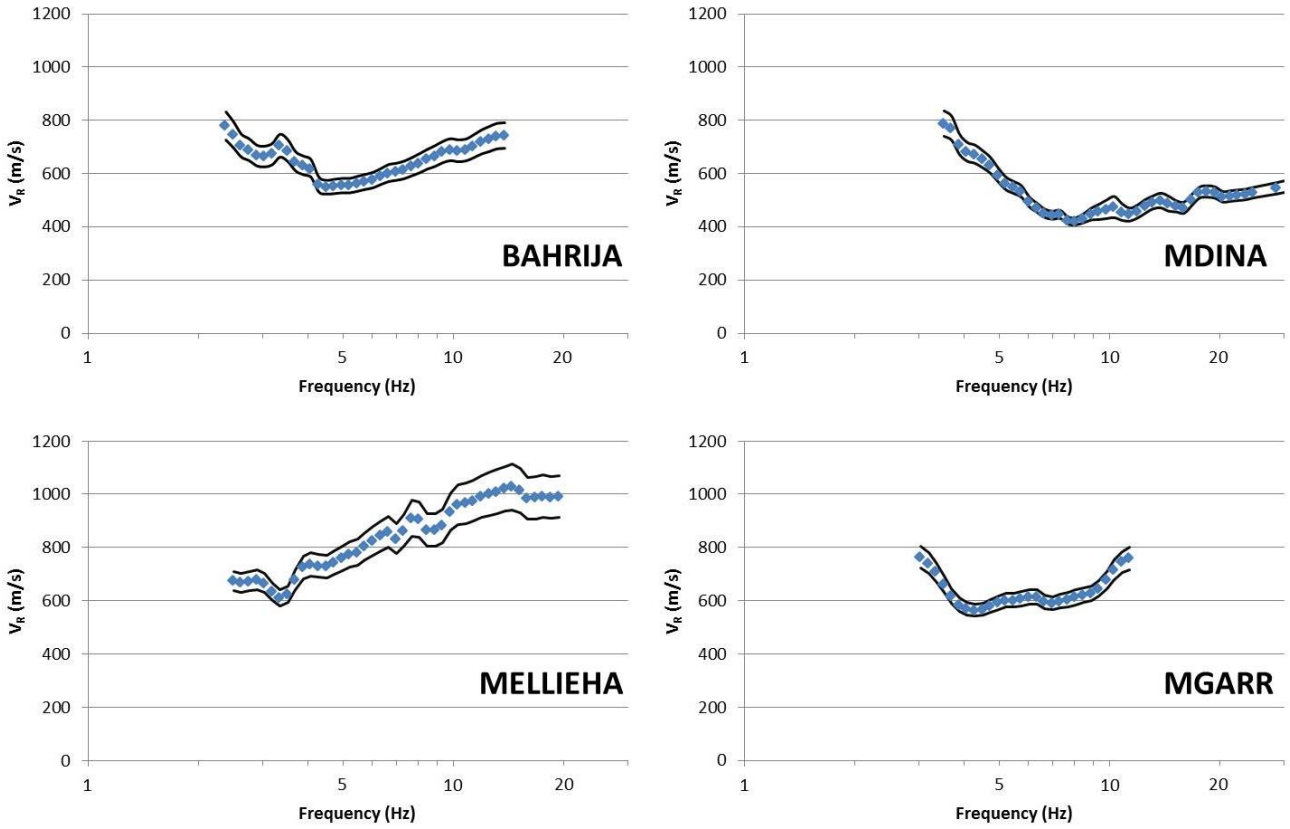

Fig. 5.6 - The effective Rayleigh wave phase velocity (V_R) dispersion curves (blue dots) and their 95% confidence interval (solid black lines) obtained at the Bahrija, Mdina, Mellieha and Mgarr sites by means the ESAC method. The frequency values are represented in logarithmic scale (modified from Farrugia *et al.*, 2016).

In contrast with these sites, the curve obtained at San Leonardo (Fig. 5.7) shows a general decrease in phase velocity with frequency. At the highest attainable frequency, the obtained Rayleigh wave velocity is around 600 m/s. At lower frequencies, the velocity continues to increase until it reaches 1000 m/s and never goes below 600 m/s, indicating that the topmost layer has the lowest velocity in the stratigraphy, and confirming the lack of the BC layer (and therefore of the V_S inversion) in this area.

The saw-tooth shape that characterizes the latter dispersion curve at frequency higher than 10 Hz is not linked with the geological features of the subsoil. In fact, this particular shape is an artefact that can mask the underlying smooth pattern and it commonly affects the dispersion curve when the array configuration is not able to detect the relatively small phase differences expected in rock

site (where large wavelengths are expected). This error could be reduced increasing the sampling frequency (Foti *et al*., 2011).

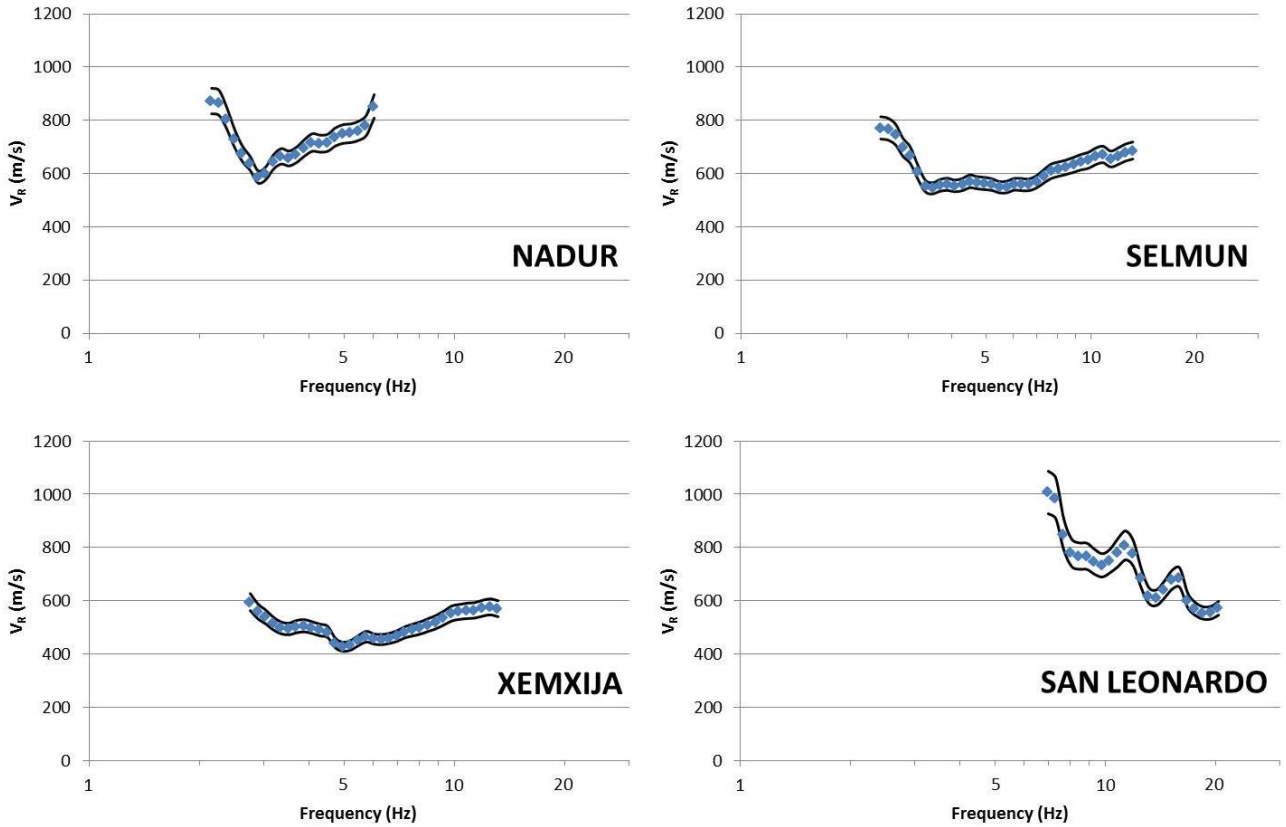

Fig. 5.7 - The effective Rayleigh wave phase velocity (V_R) dispersion curves (blue dots) and their 95% confidence interval (solid black lines) obtained at the Nadur, Selmun, Xemxija and San Lenardo sites by means the ESAC method. The frequency values are represented in logarithmic scale (modified from Farrugia *et al.*, 2016).

As described in Paragraph 3.3, the inversely dispersive trend that can characterize the effective dispersion curves in some frequency range is indicative of the presence of higher modes of surface waves. In order to detect the presence of the higher modes, an attempt was done analyzing the multi-station acquisitions with the Beam-Forming *f-k* method (see Section 3.2.2.2). Fig. 5.8 shows an example of this analysis for the Selmun site: in this figure, the dispersion curves obtained by *f-k* (solid black line) and ESAC (blue line) technique are depicted on the *f-k* spectrum together with the wavenumber limits obtained from the theoretical array response (see Section 3.2.2.2). Even though a higher mode is clearly visible, the *f-k* curve still has a similar shape to the ESAC one. This might indicate that not all the higher modes were resolved: one reason for this could be due to the limited number of geophones, as shown by Foti (2000) (Fig. 5.9).
Moreover, it is worth noting that:

- the two curves are in good agreement at high frequencies until around 4 Hz, i.e. the frequency value that corresponds to $k_{min}/2$;

- below this limit the *f-k* method provide larger estimates of phase velocity than ESAC. This bias in the phase velocity estimate could be related with the "degeneration phenomenon of *f-k* spectra" described in Section 3.2.2.3. For this, the ESAC method usually provides a more reliable dispersion curve estimate in the lower frequency regions.

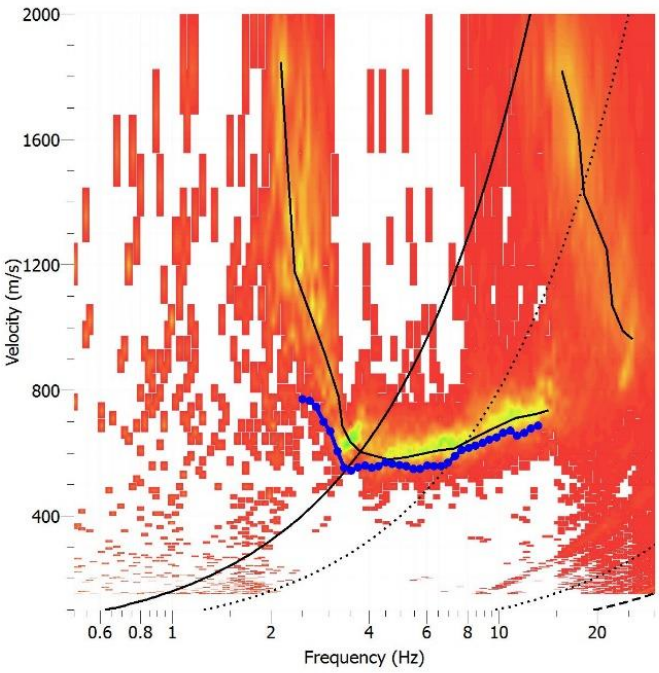

Fig. 5.8 - The f-k (solid black line) and ESAC (blue line) Rayleigh phase velocity dispersion curve obtained at Selmun site together with the wavenumber limits K_{min} (dashed line) and K_{max} (solid black line with similar shape). These curves are depicted on the *f-k* spectrum, where the yellow and green areas represent the higher power spectral density values.

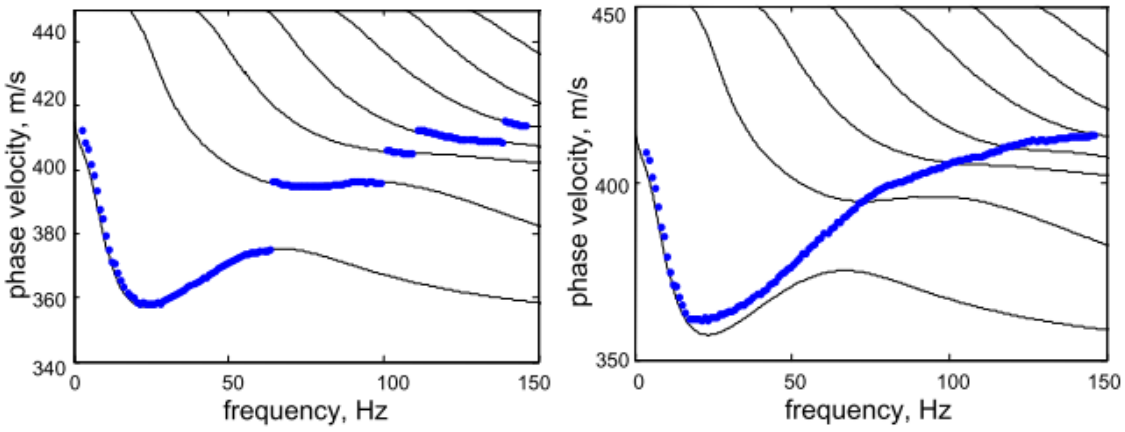

Fig. 5.9 - Dispersion curves obtained using the *f-k* method (blue) with 256 receivers (left) and 24 receivers (right) together with the theoretical dispersion curves (black) (from Foti, 2000).

5.4.3 Data inversion and V_S profiles

5.4.3.1 Joint inversion settings

The one dimensional V_S profiles were obtained by inverting the site-representative H/V curve and the effective Rayilegh wave velocity dispersion curves using the joint inversion approach described in Section 4.3.1. This inversion procedure was based on the modified version of the Genetic Algorithm (GA) proposed by Yamanaka and Ishida (1996): as described in Section 4.3.2, this method requires a series of tuning parameters to work properly. Therefore, as a preliminary operation, a search space is defined for mechanical parameters of each stratum: thickness, P- and S-wave velocities, density, damping factors. This operation allows restricting the parameter space to a delimited volume, from which all the tested models are selected.

Then, an initial population models is generated by means of a random criterion: the initial population consists of 100 individuals. The following step is an allotment of a score to each model, according to its fit, i.e., inversely to its misfit function value (see Section 4.3.2): this is a measure of the distance between the experimental curves (HVSR and V_R as function of the frequency, in the present case) and the ones produced by the theoretical model via the forward simulation code (Pileggi, 2013).

Afterwards, a new generation with the same size is created by applying the genetic operators described in Section 4.3.2, i.e. crossover, mutation and elite selection. In particular, the crossover probability (p_c) was settled to 0.9, the mutation probability (p_m) was initially settled to 0.1 (this value is dynamically varied during the inversion) and the number of elite individuals was settled to 5.

The iterations were repeated up to the 150th generation, except in the case of Nadur where 300 generations were needed to observe no further significant reduction of the misfit function respect to the previous generations. Therefore, this inversion scheme involves the analysis of 15000 S-wave velocity models (30000 for the Nadur site).

Since this inversion technique ultimately relies on a probabilistic approach using random numbers, at each site the whole inversion procedure was repeated 10 times and the best fitting profile for each was saved. The final result was chosen as the one characterized by the minimum misfit value from all 10 inversions: this means that the best model was selected among all the 150000 (300000 for the Nadur site) models tested. The other best results are useful to estimate the inversion result variability and robustness.

As concerns the mechanical parameters of the search space strata, sensitivity analyses (Tokimatsu, 1997; Xia *et al.,* 1999) have shown that the Rayleigh-wave dispersion and H/V curves are mostly influenced by V_S and thickness H, rather than by the density ρ, P-wave velocity V_P, and damping factors. As a consequence, in order to reduce the variability of the inversion results, it has been decided to link the value of ρ to the V_P one (Boore, 2015). In particular, density value is defined by an uniform distribution in the range [ρ_{min} , ρ_{max}], where

$$\rho_{\text{min}} = \begin{cases} 1500, & \text{if } V_P < 1{,}5 \\ 1740 \cdot V_P^{0,25} - 200, & \text{if } 1{,}5 \leq V_P < 6 \\ 1661{,}2 \cdot V_P - 472{,}1 \cdot V_P^2 + 67{,}1 \cdot V_P^3 - 4{,}3 \cdot V_P^4 + 0{,}106 \cdot V_P^5 - 200, & \text{if } V_P \geq 6 \end{cases} \quad (5.1)$$

with V_P in km/s, ρ_{min} in kg/m^3 and the further limit $\rho_{\text{min}} \leq 4600$ kg/m^3. The maximum limit ρ_{max} is equal to 2100 kg/m^3 in the first case and to ρ_{min} + 400 kg/m^3 in the other ones. In the cases examined in this paper, just the first two classes are of concern, since V_P was allowed to vary between 200 m/s and 3000 m/s only.

The range of values of the most significant parameters in the inversion, namely layer thickness *H* and shear-wave velocity V_S, for each site were guided by previous knowledge of the site geology, from the available geological maps or from the previous studies carried out on the Archipelago above cited. In each inversion, the number of layers, of variable thickness, was kept fixed; a minimum of four layers (representative of the investigated geological context) were inserted above the half space. The shear-wave velocity in each layer was allowed to vary over a wide range of values, with no *a priori* assumption about the presence or nature of a low-velocity layer. This enabled us to assess the ability of the experimental data and the GA to correctly identify and characterize the shear-wave velocity inversion. Tab. 5.1 shows one example of the limits used in the Mellieha case.

Layer number	*H* limits (m)	V_S limits (m/s)
1	5 - 15	400 - 1500
2	10 - 60	400 - 1500
3	10 - 60	400 - 1500
4	20 - 80	700 - 1800
5	50 - 100	800 - 2000
6	200 - 300	1000 - 2500
7	100 – 1000	1000 - 2500
8	0	1000 - 3000

Tab. 5.1 - The inversion limits used for the Mellieha data concerning thickness (*H*) and V_S values. In this case, the first two layers correspond to the Upper Coralline Limestone, the third, fourth and fifth represent the Blue Clay, Globigerina Limestone and Lower Coralline Limestone respectively. Layers 6 and 7 are added so as to avoid biases in the forward computation, while layer 8 represents the half space (from Farrugia *et al.*, 2016).

An important aspect of the inversion is the forward modeling implemented in the procedure. Theoretical HVSR and effective dispersion curves have been modeled on the basis of the DSS model (see Section 2.2.3) proposed by Lunedei and Albarello (2009). In particular, the authors

speculate the subsoil as a flat stratified viscoelastic medium where surface waves (Rayleigh and Love) propagate only. From this model, both theoretical HVSR and effective dispersion curves can be computed from the above-mentioned set of parameters representative of the subsoil (thickness, V_S, V_P, density and damping factors).

In the implemented forward modeling, higher propagation modes were considered. In fact, as mentioned previously, in sites where the V_S value varies irregularly with depth, the higher modes can play an important role in certain frequency ranges. The inclusion of their effect in the inversion procedure stabilizes the process and increases the resolution of the retrieved V_S profiles. Moreover, higher modes are more sensitive to the deeper structure than the fundamental one and thus they can increase the resolvable depth of the profile (Xia *et al*., 2003). In particular, according with Tokimatsu (1997), higher modes up to the 10^{th} one are taken into account.

Furthermore, when the presence of higher mode is significant, their sudden disappearance (modal truncation) results in unrealistic jumps in the computed dispersion and H/V curves. As suggested by Picozzi and Albarello (2007), an additional couple of layers (of the order of hundreds or one thousand of meters) were added at the bottom of the model (layer 6 and 7 in Tab. 5.1) over the half-space to prevent some artefacts. Of course, the parameters of these layers cannot be resolved by the experimental curves and simply have the role of preventing modal truncation effects (Foti *et al.*, 2011).

5.4.3.2 Obtained V_S profiles

The results of the joint inversion for all sites are summarized in Fig. 5.10 and 5.11, where the final profiles are displayed. In particular, the red profile indicates the best fitting (i.e. lowest misfit) profile, the green ones represents the profiles whose misfit are not greater than 50% of the best model one, and the yellow profiles depict the remaining inversion outcomes. The inversion concerns the San Leonardo site was not performed because the normally dispersive trend of the V_R curve (and therefore the lack of the clay formation) witnesses the absence of the shear wave velocity inversion, which is the main object of this study.

First of all, a good match between the theoretical and experimental dispersion curves and H/V peak can be observed in all cases, except in Nadur (Fig. 5.11), where the theoretical resonance frequency is slightly lower than the experimental one.

The UCL and BC thickness obtained in Mdina (Fig. 5.10) is in agreement with the borehole and the cross-section showed by Gigli *et al*., (2012). Moreover, the resulting thicknesses of these formations in the other sites are in consonance with the borehole data available in the country (Government of Malta, 1958) and are strictly comparable with the geological configuration clearly visible in outcropping and easily deductible by the geological map of the Maltese Islands (Oil Exploration Directorate, Office of the Prime Minister, Malta, 1993).

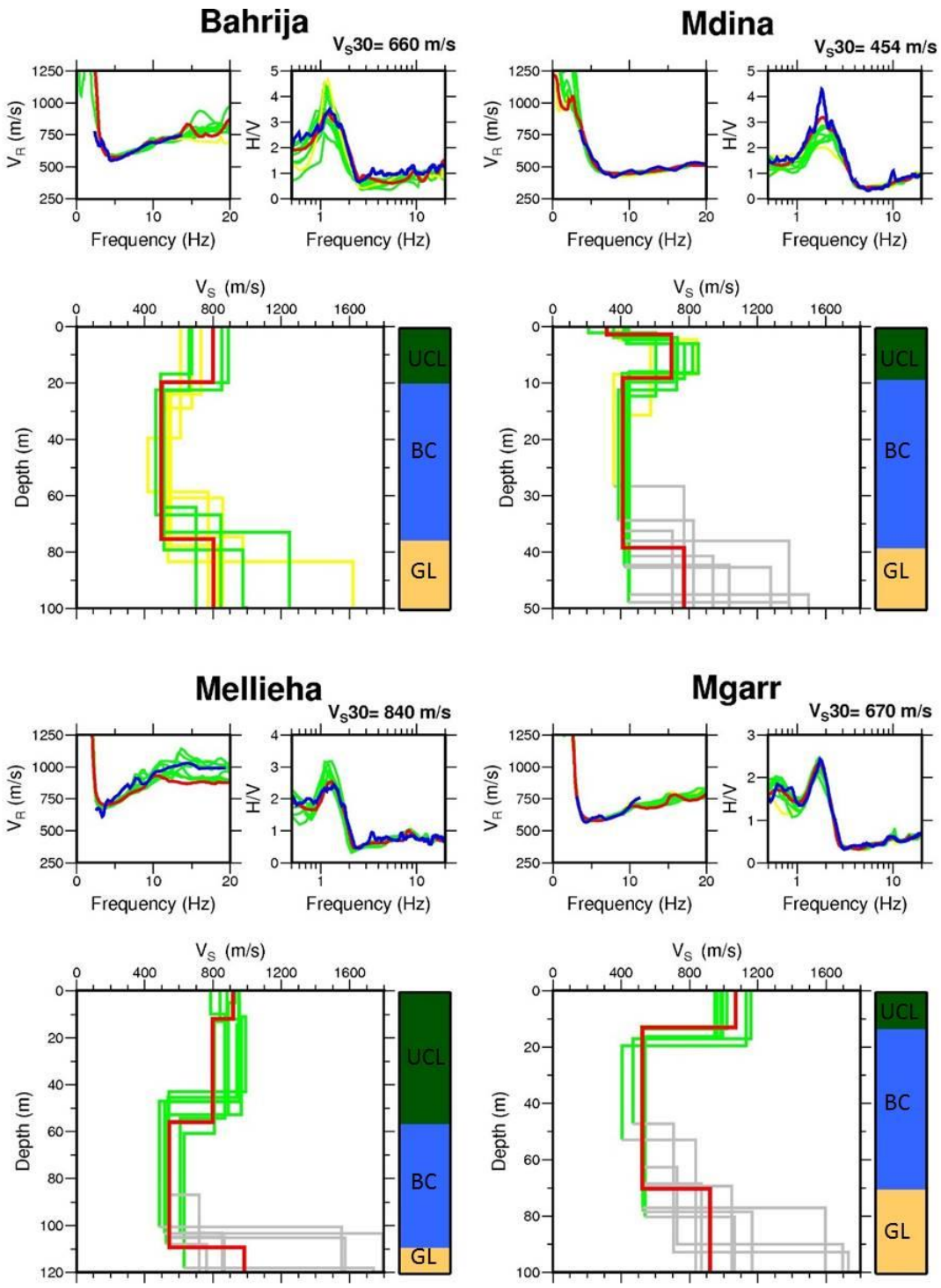

Fig. 5.10 - The joint inversion results and stratigraphic interpretation for Bahrija, Mdina, Mellieha and Mgarr sites. The red profile indicates the best fitting profile, the green ones represents the profiles whose misfit are not greater than 50% of the best model one, and the yellow profiles depict the remaining inversion outcomes. The layers interpreted as GL are displayed in grey since the values are less reliably constrained by the data (see the text). Shown in the upper panel for each site are (from left to right) the dispersion and H/V curves. The blue curve represents the experimental one, the red curve shows the best fitting theoretical curve while the rest (green and yellow) correspond to the other 9 profiles. The calculated V_{S30} for each site is displayed in the top right corner (modified from Farrugia *et al.*, 2016).

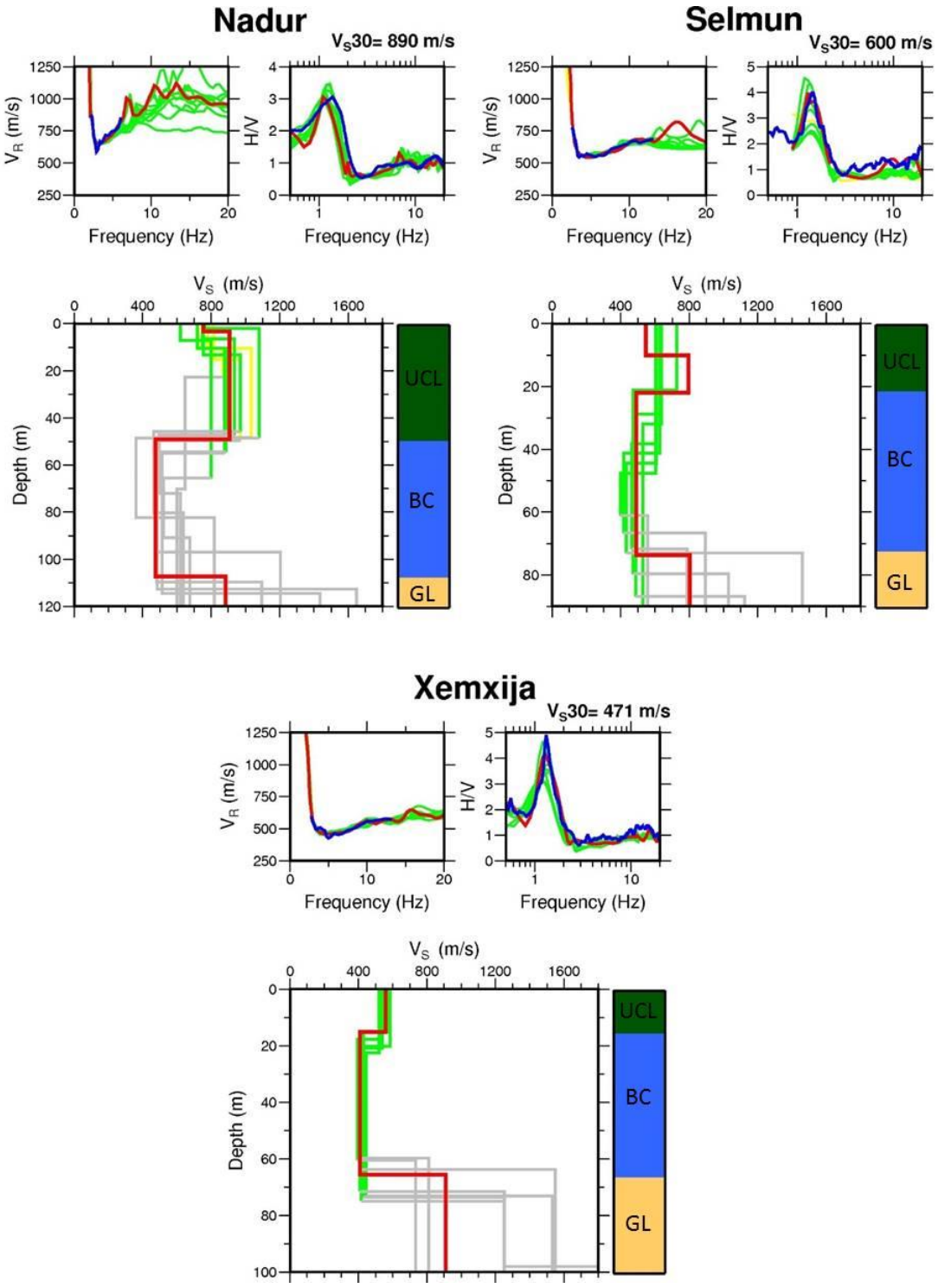

Fig. 5.11 - The joint inversion results and stratigraphic interpretation for Nadur, Selmun and Xemxila sites. Shown in the upper panel for each site are (from left to right) the dispersion and H/V curves. For colors and symbols see caption of Fig. 5.10. The calculated V_{S30} for each site is displayed in the top right corner (modified from Farrugia *et al.*, 2016).

A notable feature of the ten final profiles in each inversion is that they are all in agreement on both the position of the low-velocity layer and its velocity. Considering the fact that broad exploration ranges were initially associated with the layers in the parametrizations and that no *a priori* constraint of a low-velocity layer was made, such an agreement in all the models at the respective sites shows the robustness of the inversion and again indicates the sensitivity of the curves to the presence and properties of the low-velocity layer. In addition, this justifies the use of global search methods, such as the GA, which are able to retrieve reasonable profiles without the need of an initial profile close to the solution.

On the other hand, this consistency diminishes in the prediction of the velocity of the UCL and more so of the Globigerina Limestone formation layer, where overall values between 700-1800 m/s were obtained. The latter inconsistency can be mainly attributed to the available array conditions, especially maximum length and resonance frequency of geophones which limit the observable depth leaving less constrained the inversion solutions regarding the V_S velocity of the GL formation. This difficulty in sampling this calcareous layer (that can be considered the seismic bedrock) can be also related with the presence of the soft BC layer, which can act as a high-pass filter partly confining the propagation of the surface wave along the depth.

The V_{S30} values of the best models were also calculated for each site and are displayed in Fig. 5.10 and Fig. 5.11. The resulting values classify Mellieha and Nadur as a class A (i.e. rock site), and the rest as class B sites (i.e. deposits of very dense sand, gravel, or very stiff clay) according to the EC8 classification (Eurocode 8, EN-1998, 2004). It is possible to note that this classification strongly depends to the presence of the low velocity layer within 30 m, which can sharply decrease the average velocity value also in the presence of high V_S for UCL (such as in Mgarr site). In some cases (Mdina, Selmun and Xemxija) the average V_S values for UCL are sufficiently low to insert these sites in class B. Obviously, the presence of resonance peaks in the H/V curves indicates possible amplification phenomena also in the class A sites, due to the buried BC, which need to be accounted for site response studies.

The final profiles related to the best models reveal a variation in the UCL and BC shear-wave velocities at the different sites. Tab. 5.2 lists the overall best model values and the ranges corresponding to the best-fit models from the 10 inversions. Whilst the V_S value in clay is around 400 m/s in Mdina, where the thickness of the UCL is less than 10 m, this value increases at the other sites and reaches a maximum value of 550 m/s in Mellieha, where the thickness of the overlying hard layer is around 56 m. This phenomenon can presumably be related to the overburden of the hard UCL layer on the BC, increasing the compactness of the particles, and thus the V_S of the layer. The BC shear wave velocity values retrieved in Mdina site are comparable with the ones (300-400 m/s) obtained by Panzera *et al*. (2013) and Farrugia *et al*. (2015) near this city, where the clays are outcropping.

The variation in UCL shear wave velocity observed at the different sites is more remarkable. In fact, the V_S values in this formation (which generally are about 700-800 m/s) are characterized by a minimum of 560 m/s in Xemxija and a maximum of 1070 m/s in Mgarr site. This feature is

probably due to the different structural and geological characteristics of the UCL, which are mostly related with its fracturing condition and different geological facies.

SITE	UCL S-wave velocity (m/s)	BC S-wave velocity (m/s)
Bahrija	800 (610 – 890)	490 (410 – 540)
Mdina	700 (580 – 860)	410 (360 – 460)
Mellieha	800 (800 – 980)	550 (470 – 630)
Mgarr	1070 (950 – 1160)	520 (380 – 530)
Nadur	900 (760 – 1080)	480 (370 – 660)
Selmun	700 (600 – 720)	490 (400 – 530)
Xemxija	560 (500 – 600)	400 (390 – 460)

Tab. 5.2 - The UCL and BC shear-wave velocity values obtained for the best fitting models and the ranges (in brackets) given by all the best-fit models of the 10 inversions (from Farrugia *et al.*, 2016).

Finally, the theoretical Rayleigh-wave dispersion curves up to the second higher mode were computed for the best models so as to compare with the experimental effective dispersion curve. Fig. 5.12 shows the plots for the Bahrija and Mdina best models, including the computed theoretical effective dispersion curve, which fits very well with the observed data. These examples confirm that the effective Rayleigh wave velocity dispersion curve is indeed the superposition of different modes: in particular, it is possible to note that the higher modes play an important role in the frequency range when this curve shows an inversely dispersive character.

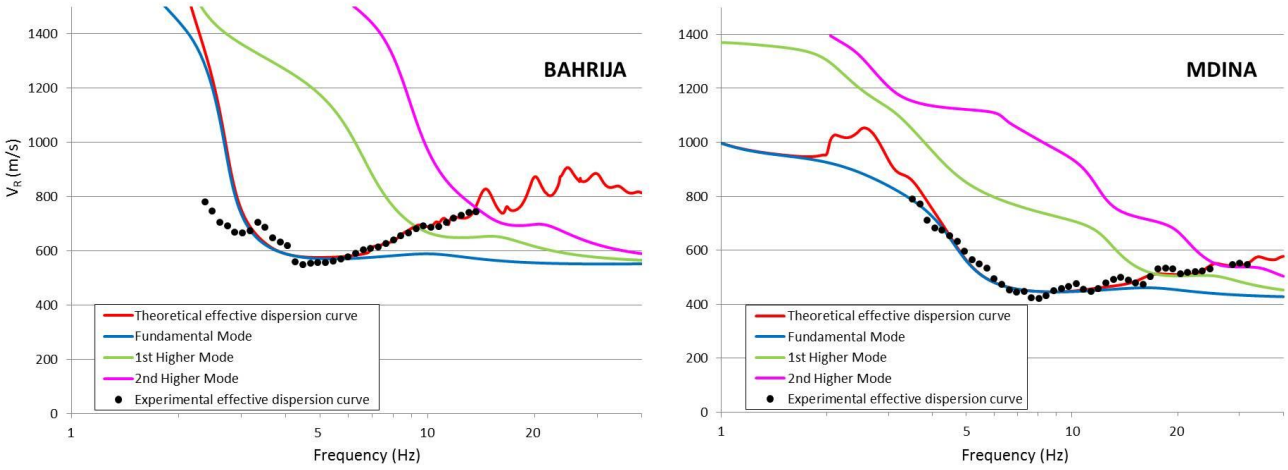

Fig. 5.12 - Comparison of the experimental effective Rayleigh-wave phase velocities (black dots) with the theoretical effective phase velocities and the first three Rayleigh-wave modes for the Bahrija and Mdina sites (modified from Farrugia *et al.*, 2016).

CHAPTER 6
Extensive use of ambient vibration prospecting and identification of large scale heterogeneities: the case of Emilia Romagna, Italy

Based on the paper published by Bulletin of Earthquake Engineering (Paolucci et al., 2015)

6.1 Introduction

One of the aims of this dissertation is to check the informative capability of the extensive use of the ambient vibration prospecting. In particular, this chapter is devoted to achieve the dynamic characteristics of the subsoil at the scale of about thousands of square kilometers identifying the possible lateral large scale heterogeneities (of the order of tens-hundreds square kilometers) by means the information inferred from the passive seismic measurements.

The study area, which extends for about 2500 Km^2, is located in the central Po Plain (Northern Italy) and was hit by a damaging seismic sequence in 2012 (Fig. 6.1). In particular, from May 2012 until the end of June 2012, this area was affected by a seismic sequence characterized by two mainshocks (Scognamiglio *et al.*, 2012): the first occurred on May 20 (M_L = 5.9) and the second (located about 12 km West of the former) on May 29 (M_L = 5.8). The sequence produced serious injuries to buildings and infrastructures with severe losses to the economy of the region; as a whole, 27 causalities occurred.

To accomplish the objectives described above, 204 single station three component acquisitions and 34 seismic arrays were taken into account. These ambient vibration measurements were performed by teams from different institutions and universities in the whole damaged area in order to assess the possible site effects related to the amplification of the ground motion. Many of these acquisitions were located within the areas characterized by effects corresponding to or larger than VI MCS (Fig. 6.1). Some acquisitions were carried out during the seismic sequence, after the first main shock by teams of University of Siena (UNISI), University of Basilicata (UNIBAS), University of Malta (UM) and Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (OGS); some preliminary results from OGS and UNIBAS have been discussed in Priolo *et al*. (2012). Many measurements were also performed in 2013 on behalf of "Servizio geologico, sismico e dei suoli" of the Emilia-Romagna regional administration (RER) in the frame of the Seismic Microzonation project of the area (http://ambiente.regione.emilia-romagna.it/geologia/temi/sismica/speciale-terremoto/sisma-2012-ordinanza-70-13-11-2012-cartografia).

In order to identify the large scale heterogeneities, the HVSR curves obtained by the single station acquisitions are taken into account: their cheapness and quickness allowed to satisfactorily cover the whole investigated area (with higher focus to the urban zones) detecting sub-areas where analogous seismic response is expected. Moreover, the differences between H/V curves obtained

in different acquisition periods are treated: in particular, a correlation between H/V peak amplitudes and the sea wave activity in the Central Mediterranean was attempted.

In view of the large number of multiple-station data collected, a simplified approach such as that described in Paragraph 4.2 was adopted in order to estimate the S-wave velocity of the shallow subsoil materials and the resonant interface depth. In fact, this approach (based on the direct interpretation of the Rayleigh wave phase velocity dispersion curves provided by the seismic arrays) results less numerically troublesome and time consuming than the inversion procedures and therefore more attractive in order to manage a large amount of data. Furthermore, its use is particularly suitable in contexts where no great geological and geomorphological changes for the shallow subsoil are expected.

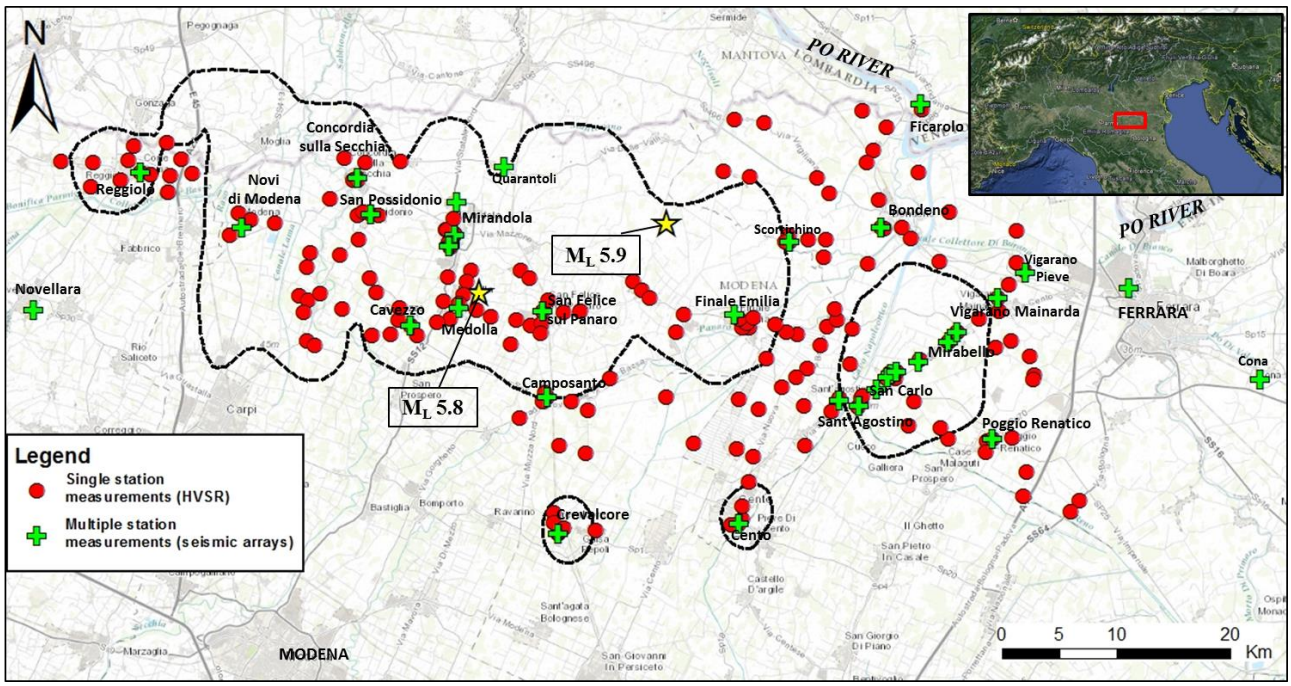

Fig. 6.1 - Localizations of the passive seismic measurements: green crosses related to the seismic arrays, red dots to the single stations measurements. Dashed black lines indicate 6 MCS isoseismal (modified from Galli *et al.*, 2012), the yellow stars are the epicenters of two mainshocks (May 20 and 29) (from Paolucci *et al.*, 2015).

6.2 Geological settings

The Po Plain includes the foreland of the S-verging central-southern Alps and of the N-NE-verging northern Apennine fold-and-thrust belts (Carminati and Doglioni, 2012); its physiographic boundaries coincide to the contact between the Quaternary alluvial sediments of the plain and the pre-Quaternary rocks outcropped along the mountain fronts.

The epicentral area of the 2012 seismic sequence is located in the central-northern sector of Emilian Po Plain, southward Po River, between Ferrara, Modena and Reggio Emilia municipalities.

From the morphological point of view, this area is extremely regular and smooth, with altitudes ranging from W to E from 25-30 to 8-10 m above sea level.

The local seismicity is mainly due the compressional tectonic activity of the faults belonging to Ferrara Arc (Pieri and Groppi, 1981; Toscani *et al.*, 2009). This structure is a fold-and-thrust system which has given rise to two main buried ridges, which together describe on map a N-verging arc extending between Reggio Emilia and the Adriatic coast, reaching northwards the zones near and over the Po River (Bigi *et al.*, 1990). In particular it is possible to distinguish an internal ridge, with its maximum uplift between Novi di Modena and Medolla, known as Mirandola's structure (or Mirandola's ridge), and a more external one, that reaches the most high point between Bondeno and Ferrara (Martelli and Romani, 2013).

As concerns the superficial geology, the outcropping materials are represented by alluvial Holocene sediments (Fig. 6.2), particularly alternances of sand, silt and clay of different environments, like river channels and levees, meander and floodplain deposits (Martelli *et al.*, 2014).

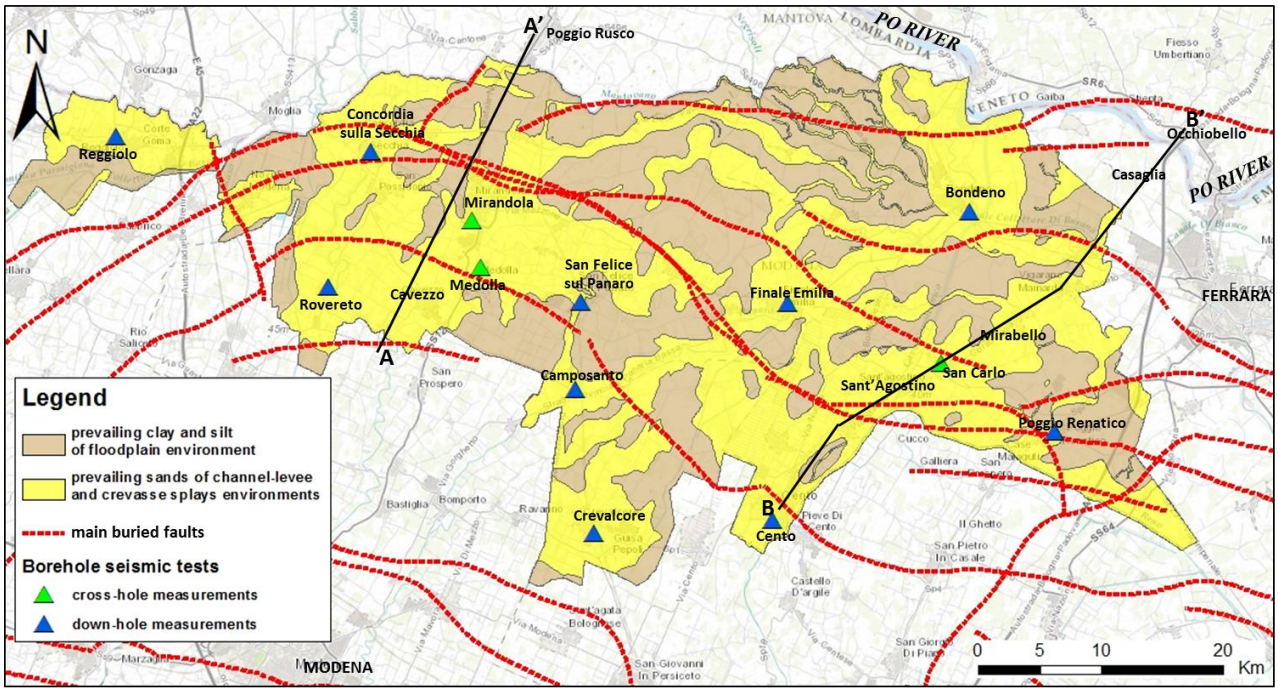

Fig. 6.2 - Geological map of the study area; the age of outcropping materials is Holocene everywhere. The red dashed lines indicate the traces of the fronts of the main buried faults, the black continuous lines represent the geological cross-section traces; triangular placemarks indicate the borehole seismic tests localizations (from Paolucci *et al.*, 2015).

The whole alluvial succession, belonging to Pleistocene-Holocene age, is attributable to two main depositional cycles: Emiliano-Romagnolo Lower Synthem (AEI), included between 700.000-650.000 and 450.000-400.000 years (Middle Pleistocene), and Emiliano-Romagnolo Upper

Synthem (AES), included between 450.000/400.000 years and the present (RER & ENI-Agip, 1998; Fig. 6.3). This continental succession lies on a Lower-Middle Pleistocene deposits (mainly sands) of marine and transitional origin (RER & ENI-Agip, 1998; Martelli *et. al*, 2014; Fig. 6.3); the whole Quaternary succession lies directly on marine Oligocene-Upper Pliocene deposits characterized by more evident tectonic deformations (Fig. 6.3).

The depth of the bottom of the alluvial succession changes due to Ferrara Arc configuration: in synclinal zones the continental deposits thickness can exceed 500 m, instead in the anticlinal areas it reaches about 100 m of depth (RER & ENI-Agip, 1998). In the structural high areas, the Quaternary marine succession is often heavily eroded and the deposits of the Emiliano-Romagnolo Lower Synthem (AEI) have reduced thickness as well. Moreover, in these zones, there are wide stratigraphic gaps (Martelli and Romani, 2013): for example, between Novi di Modena and Medolla (Mirandola's ridge; Fig. 6.3-A), continental deposits lie directly on Lower Pliocene marine sands and clays, instead between Bondeno and Ferrara (northeast zone) the Low Quaternary transitional sediments lie sometimes on Oligo-Miocene marls (Casaglia's ridge, Fig. 6.3-B).

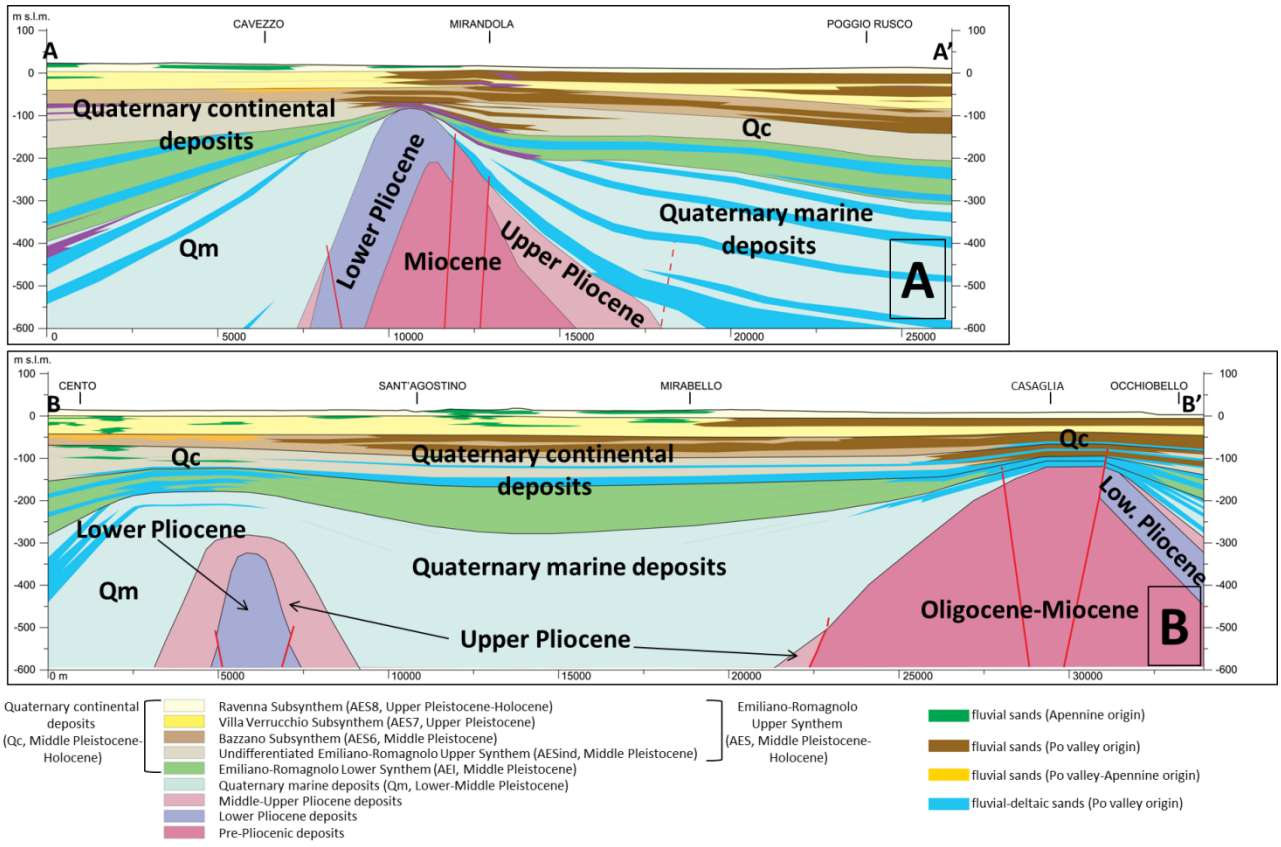

Fig. 6.3 - Geological cross-sections through the study area scale 1:50000 with vertical exaggeration 12,5x; the relative traces are indicated in Fig. 6.2. In the left part of the legend, the main sedimentary cycles colors are represented; in the right part, the main sand bodies detected by borehole stratigraphic profiles are reported (from Paolucci *et al.*, 2015).

6.3 Ambient vibration prospecting

6.3.1 Single station measurements

Most of single station acquisitions were performed by a three-directional digital tromograph Tromino Micromed with a sampling frequency of 128 Hz and an acquisition time of 15-20 minutes. At each site, the HVSR curve was obtained dividing the time series into 60 non-overlapping windows with 20 s length and following the procedure described in Paragraph 2.3. Just in the case of seismic stations managed by OGS (located mainly in the Ferrara province), ambient vibrations were recorded in continuous mode for about two months (Priolo *et al.*, 2012).
As in the study described in Chapter 5, any window containing spurious signals was removed and the quality of the resulting HVSR curve was evaluated following the criteria described in Paragraph 2.4. In particular, the quality classification proposed by Albarello *et al*. (2011) was adopted in order to identifying unreliable experimental results.
Examining the overall trend of the HVSR curves by means a visual inspection, 3 main patterns were identified which are representative of 3 geographic macrozones (Fig. 6.4):

- Zone 1, including the south and west parts of the investigated area, is characterized by the most widespread pattern. The envelope of HVSR curves relative to this zone shows a main broad peak in the frequency range 0.8-0.9 Hz (hereafter the “main frequency peak”) and a secondary peak around 0.25-0.3 Hz (hereafter the “low frequency peak”). Similar low frequency HVSR maxima are identified in the frequency range 0.1-0.2 Hz in a preliminary study realized in the same area with different acquisition instruments (Bordoni *et al.*, 2012) and in the frequency range 0.2-0.5 Hz in a study achieved near the bank areas of the Po River (Martelli *et al.*, 2014).
- Zone 2, localized in the northeast part of the area, shows an HVSR pattern similar to the previous one (the main broad peak in the frequency band 0.8-0.9 Hz and the low frequency peak at about 0.25 Hz) with the addition of a slight “hump” between 1.5 and 2 Hz;
- Zone 3, including the municipality of Mirandola, is characterized by HVSR curves with a single sharp peak in the range 0.8-1.4 Hz.

Within Zones 1 and 2, the frequency of the main HVSR maximum shows irregular variations in the relevant frequency range (0.8-0.9 Hz) without any clear spatial pattern; these random spatial variations are observed also at the single settlement scale. This could also be the effect of the broadness of the relevant peaks that may hamper the precise identification of the relevant resonance frequency. As concerns Zone 3, sharpness of the main peak allows a better identification of the resonance frequency. Thus, the observed frequency variations appear more significant. Furthermore, in this zone, the frequency shift shows a more defined pattern: in the range 0.8-0.9 Hz in Mirandola maintown, to 1.4 Hz moving westward (in the area between San Possidonio and Rovereto), to 1.2 Hz moving southward (San Giacomo in Roncole) and then to 0.8 Hz again moving towards Medolla.

The same frequency values were found in Mirandola town by Gallipoli *et al*. (2014); in this work a comparison between the ambient vibrations HVSR curve and those ones obtained from earthquake recordings was performed.

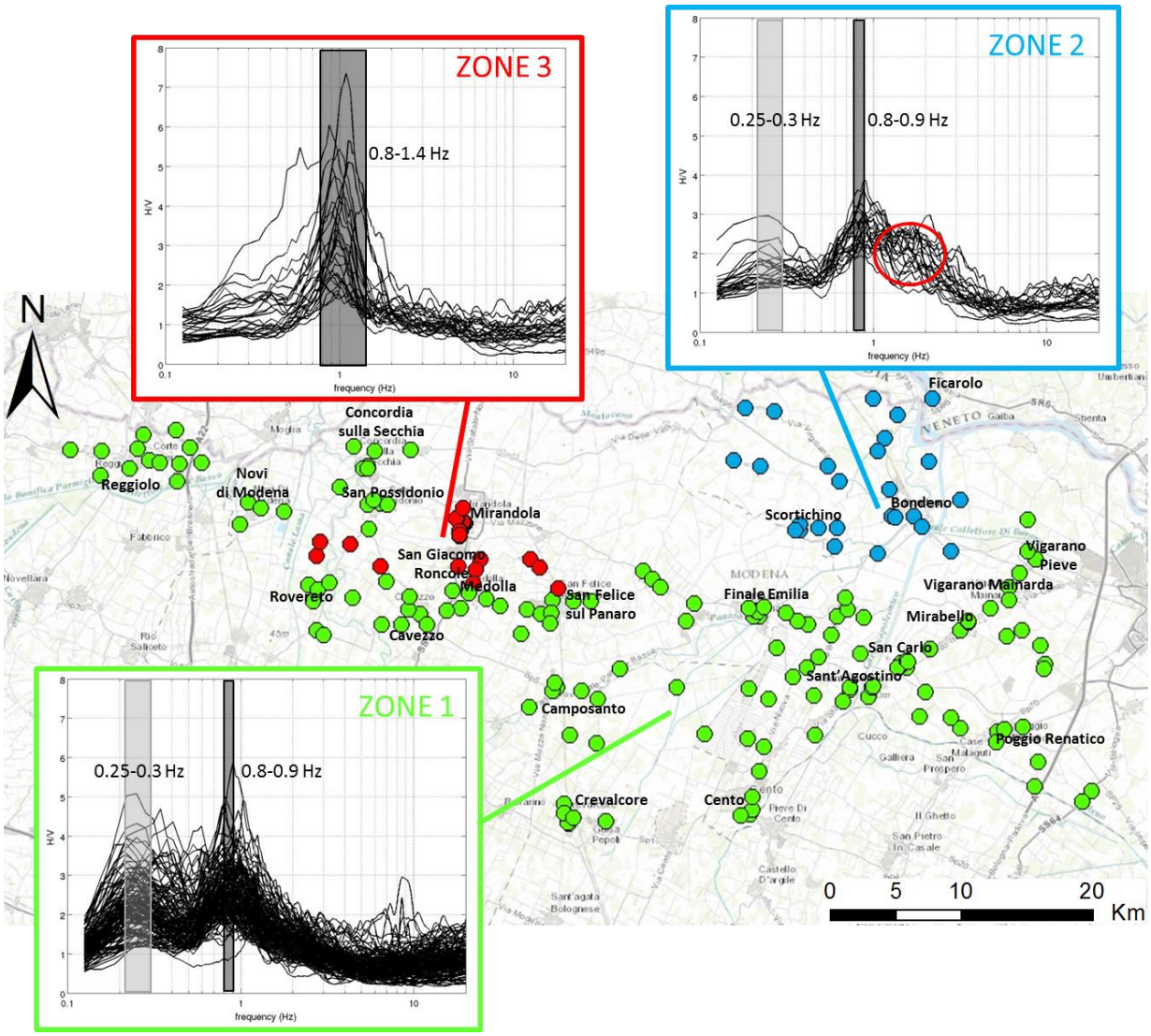

Fig. 6.4 - Localizations of the 3 macrozones representative of the 3 main patterns identified examining all HVSR curves. The dots indicate the HVSR measures positions: the green ones belong to the Zone 1, the blue ones to the Zone 2 and the red ones to the Zone 3. Black bars show main peak frequency range, light gray bars indicate low frequency peak range; red circle is referred to the slight "hump" between 1.5-2 Hz identified by Zone 2 HVSR pattern (from Paolucci *et al.*, 2015).

It is worth noting that many of the obtained HVSR curve are characterized by a clear drop below 1 over a wide frequency range, often included between 2 and 20 Hz (Fig. 6.4). This pattern is particularly evident in the Zones 1 and 2 and it do not appear to be related to a sharp shear wave velocity inversion on the subsoil. Conversely to the situation studied in Chapter 5, in this area no

geological elements exist that can confirm this hypothesis. Moreover, the direct V_S estimates provided by the borehole seismic tests present in the whole investigated area are not characterized by significant S-wave inversions in the shallowest part of the subsoil (*vide infra*); in fact, the maximum V_S decrease and the involved thickness are of the order of 50-100 m/s and few meters respectively.

6.3.2 Seismic arrays

In order to constrain the local V_S profile, 34 seismic arrays were deployed at sites uniformly distributed in the whole survey area (Fig. 6.1). Ambient vibrations were recorded for 20 minutes at 128 Hz sampling rate by using vertical geophones (4.5 Hz) and a digital acquisition system (BrainSpy 16 channel acquisition system by Micromed for UNISI and RER and Geode 24-channel modular acquisition system by Geometrics for UNIBAS). Sensors were irregularly spaced (in the range 0.5-40 m) along two crossing perpendicular branches (UNISI) or along L and C shape branches (UNIBAS), with maximum dimensions between 100 and 250 m.

As in the study described in the previous Chapter, the determination of the Rayleigh wave phase velocity (V_R) dispersion curves was obtained by using the ESAC technique (see Section 3.2.1.2) in the form proposed by Parolai *et al.* (2006).

As can be seen from Fig. 6.5, all the dispersion curves are included in the frequency range between 1.5-2.5 Hz and 25-45 Hz and show very similar trends, with velocity values that monotonically decrease when frequency increases (i.e. normally dispersive trends). In particular, minimum V_R values are very low (100-150 m/s) while maximum values reach 300-400 m/s (only Medolla and San Felice sul Panaro's curves show slightly higher velocity values). Differently from HVSR curves, it was not possible to identify different patterns of the relevant dispersion curves relative to the geographic zones described above.

These features suggest that, at least as concerns the first 50-100 m of subsoil (the maximum investigation depth of seismic arrays), the entire area can be considered nearly homogeneous from the seismostratigraphic point of view.

In view of this, it is possible to state that the differences pointed out by HVSR curves are mainly related with the dynamic and geological characteristics of the detected resonant interfaces.

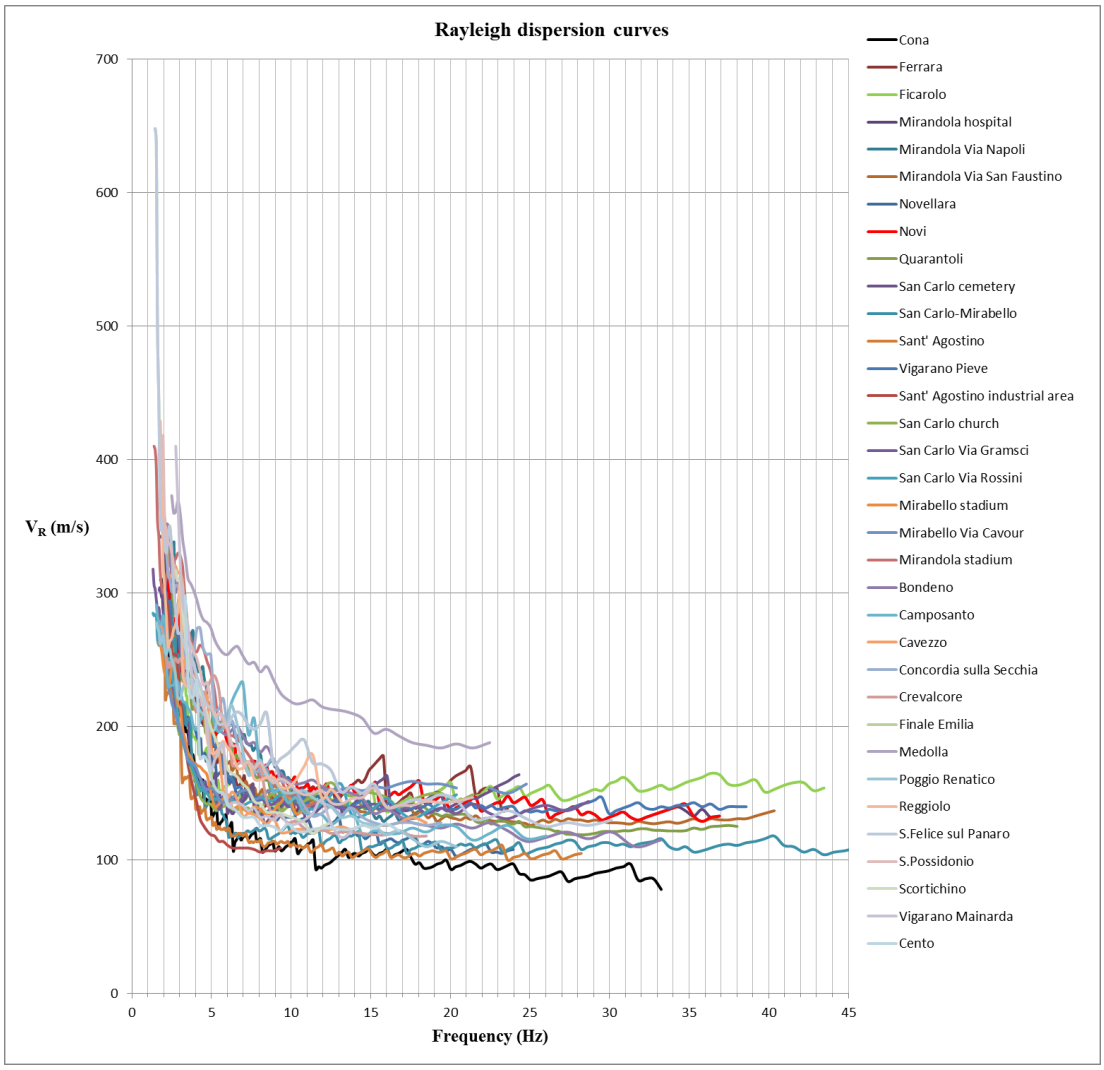

Fig. 6.5 - Rayleigh wave dispersion curve trends obtained by ESAC analysis of the 34 seismic arrays carried out in the study area (from Paolucci *et al.*, 2015).

The seismostratigraphic homogeneity of the shallowest subsoil materials is also confirmed by the shear wave velocity values obtained from 13 borehole seismic tests (10 down-hole and 3 cross-hole, Fig. 6.2) subsequently carried out by RER with the aim of obtaining more information about the dynamic characteristics of the sediments present in the area damaged by the earthquake. As one can see (Fig. 6.6), in the first 50 m of depth, V_S profiles show maximum velocity differences of 150 m/s, with values fully comparable with the V_R values retrieved from array measurements. Of particular importance appears borehole information obtained at the Mirandola and Medolla sites, where Lower Pliocene deposits reach the bottom of the alluvial deposits at a depth of about 100 m. Thus, direct V_S values have been retrieved for both Quaternary and Pliocenic formations. This information is of paramount importance for the characterization of seismic properties of the Po Plain deposits.

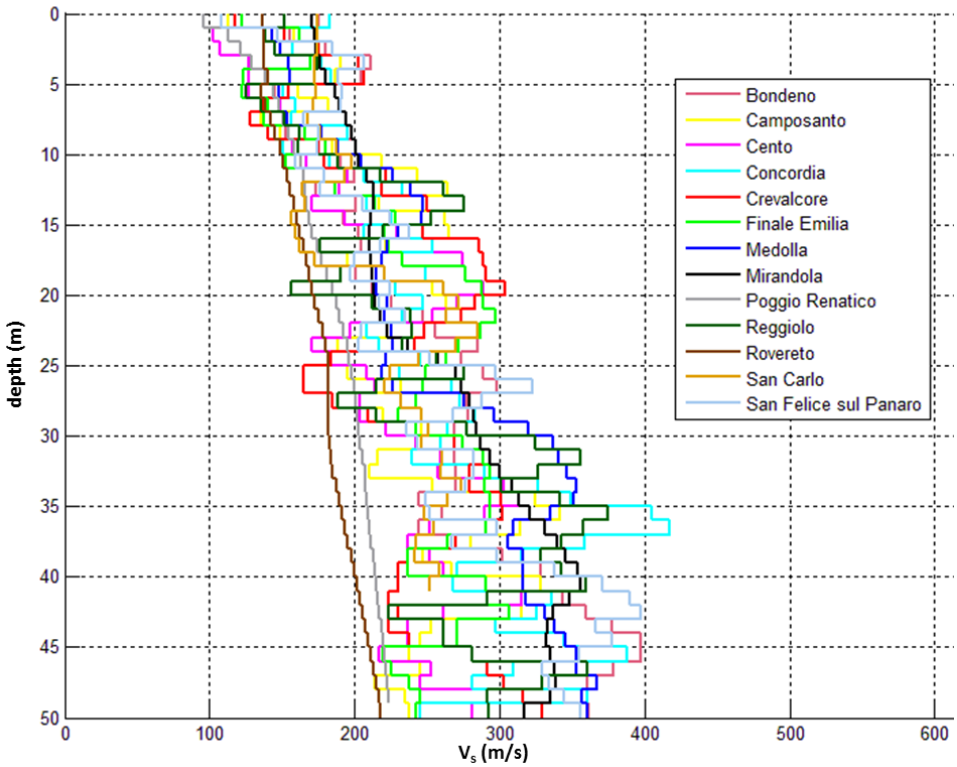

Fig. 6.6 - Shear wave velocity profiles obtained from 13 borehole seismic tests carried out by RER (from Paolucci *et al.*, 2015).

6.4 The effect of local meteo-climatic conditions

Beyond very general patterns leading to the zonation provided above, peaks of the HVSR curves are characterized by significant and spatially erratic variations of amplitude values, both in the low and the high frequency range (Fig. 6.4). This variability occurs also at the scale of a single settlement. These variations mainly affect the amplitude of the HVSR maxima, while relevant resonance frequencies of both peaks appear quite stable within the frequency ranges detected.

In principle, resonance frequencies depend on the ratio between the average V_S velocity in the resonant layer and the thickness of this layer (Eq. 3). The amplitude, instead, depends on a number of factors (see, e.g., Lunedei and Albarello, 2010; Albarello and Lunedei, 2011), including both seismostratigraphic features (the impedance contrast at the resonant interface) and the structure of the ambient vibration wavefield (e.g., the presence of a source-free domain around the recording station, see Section 2.2.3).

The relative homogeneity of V_S values deduced from ambient vibration measurements and borehole data suggests that variations observed in the amplitudes of HVSR peaks do not depend on significant variations of the V_S profile in the uppermost part of the subsoil. Furthermore, available geological data allow excluding significant lateral variations of the stratigraphical

configuration at wavelengths comparable with the ones characterizing the observed variations of HVSR amplitudes.

To explore the possibility that possible effects independent from subsoil condition may be responsible for the observed changes in the HVSR shape, amplitude of HVSR values corresponding to the resonance frequencies were correlated with the sea-wave activity in the seas surrounding the area (Tyrrhenian and Adriatic Seas). In fact, as described in Paragraph 1.2, it is known that the nature of the ambient vibration wavefield in frequency ranges lower than 1 Hz is strictly related to the oceanic-sea wave activity, in particular to the waves striking along the coasts and the occurring cyclones.

Data provided by ISPRA—Servizio Mareografico ‘‘Rete Ondametrica Nazionale’’ (Bencivenga *et al*. 2012; http://www.telemisura.it/) were considered on purpose: in particular, sea-wave average height obtained each half hour at the measure stations of Ancona, Cagliari, Civitavecchia, La Spezia, Palermo, Ponza and Venice were taken in account (Fig. 6.7).

Fig. 6.7 - Localizations of the “Rete Ondametrica Nazionale” stations (orange wave icons) (from Paolucci *et al.*, 2015).

The results of this analysis enlighten the presence of a statistically significant correlation between sea-wave activity and amplitude of low frequency peaks; specifically the maximum correlation coefficient values (about 0.5-0.6) are verified with Tyrrhenian Sea data (La Spezia and Cagliari) and with the oceanographic stations (Ancona and Venice) nearest to the study area (Tab. 6.1; Fig. 6.8). No significant correlation was found between the amplitude of the main frequency HVSR maximum with global or local meteorological conditions (Tab. 6.1; Fig. 6.8). This result, however, does not exclude the possible impact of local sources of ambient vibrations related to the anthropic activity; the possible interaction between this kind of sources and the HVSR amplitude is

showed by Cara *et al*. (2010). Thus, in order to avoid possible interpretative biases, HVSR amplitudes have not been considered in the following analysis and the interpretation of HVSR curves focused on the resonance frequencies only.

Station name	Correlation coefficient A1	Correlation coefficient A0
Ancona	0.55	0.1
Cagliari	0.58	0.17
Civitavecchia	0.25	0.24
La Spezia	0.60	0.16
Palermo	0.56	0.4
Ponza	0.44	0.09
Venezia	0.49	0.08

Tab. 6.1 - Pearson correlation coefficient values obtained among the sea-wave height measured in each of the seven stations belonging to "Rete Ondametrica Nazionale" and amplitude of the low frequency range peak (A1) and that of main frequency peak (A0) (modified from Paolucci *et al.*, 2015).

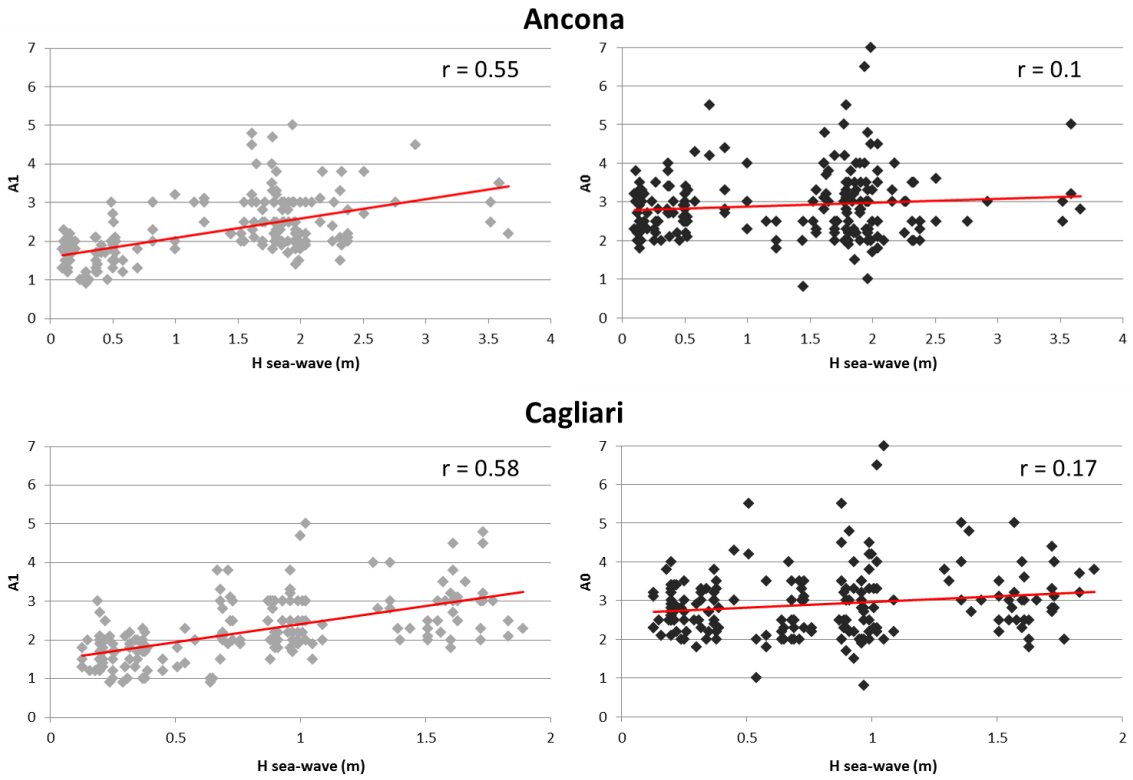

Fig. 6.8 - Amplitude values of the low frequency range peak (A1, on the left) and of the main frequency peak (A0, on the right) as a function of the sea-wave height (H sea-wave) for the Ancona (Adriatic Sea) and Cagliari (Tyrrhenian Sea) stations of ''Rete Ondametrica Nazionale''. Red lines represent the linear regression curves, r is the Pearson correlation coefficient (Tab. 6.1) (from Paolucci *et al.*, 2015).

6.5 Estimating depth and geological interpretation of resonant interfaces

In order to associate a possible depth to the seismic interfaces responsible for resonance frequencies detected by the HVSR measurements in the three zones described above, the simplified approach described in Paragraph 4.2 has been adopted. In particular, the velocity profile $V_S(h)$ is assumed in the form of the power law in Eq. 4.1. This assumption is not in contrast with available information about the V_S profiles in the study area: in fact, as we can see in Fig. 6.6, the V_S profiles obtained by borehole seismic tests show a gradual velocity increasing without significant impedance contrasts at least in the depth range considered. From Eq. 4.1, one can see that the average velocity $\bar{V}_S(h)$ up to the depth *h* has the form

$$\bar{V}_S(h) \approx V_0(1-x)h^x. \tag{6.1}$$

where V_0 and *x* are empirical parameters (related with the subsoil nature) to be computed case-by-case (see Paragraph 4.2).

In this approximation, the depth *H* of the resonant interface responsible for the resonant frequency f_0 at any site can be estimated by the *f–H* conversion rule of the Eq. 4.2.

Dispersion curves deduced from array measurements and borehole data have been jointly considered to evaluate actual feasibility of the simplified model in Eq. 6.1 to represent average V_S values and to estimate the values of V_0 and *x* best fitting observations.

6.5.1 Calibration of f–H conversion rule

Estimation of V_0 and *x* values can be performed in two ways. The first is by considering Eq. 4.1 and choosing values best fitting V_S profiles values at a number of sites. The other way, is considering Eq. 6.1 and best fitting the $\bar{V}_S(h)$ curve of the average V_S values up to any value of the depth *h*. In principle, these approaches are equivalent. However, the second curve is generally smoother than the first one, and this makes fitting easier. Furthermore, the $\bar{V}_S(h)$ values can be directly retrieved from the dispersion curves: in fact, the adopted approach (see Paragraph 4.2) assumes that the Rayleigh wave phase velocity roughly corresponds to the average S-wave velocity up to a depth *h*, where *h* corresponds to a fraction *c* of the relevant wavelength (Eq. 4.3).

The estimate of the parameter *c* was achieved by comparing values obtained with Eq. 4.3 and direct experimental evaluations of $\bar{V}_S(h)$ provided by borehole data. Thus, dispersion curves and V_S profiles from borehole data have been considered to constrain V_0 and *x* values to be used in the study area to retrieve *H* values from f_0 measurements via Eq. 4.2.

To this purpose, borehole data have been considered first. In principle, there are two ways to compute $\bar{V}_S(h)$ from the $V_S(h)$ profile:

$$\bar{V}_S(h) = \frac{h}{\sum_{i=1}^{n} \frac{h_i}{V_i}} \tag{6.2}$$

or

$$\bar{V}_S(h) = \frac{\sum_{i=1}^{n} V_i h_i}{h} \tag{6.3}$$

where h_i and V_i are respectively thickness and V_S interval velocity of the i^{th} seismo-layer and $h = \sum_{i=1}^{n} h_i$. These two equations result in different estimates of $\bar{V}_S(h)$. In order to identify the best performing equation, the cross-hole seismic tests of Mirandola and Medolla (Fig. 6.9) have been considered. These boreholes are the only ones in the study area where a sharp impedance contrast has been intercepted at the depths of 95 m (Medolla) and about 115 m (Mirandola), corresponding to the bottom of the alluvial Quaternary sediments at the top of Pliocene marine deposits (Fig. 6.9).

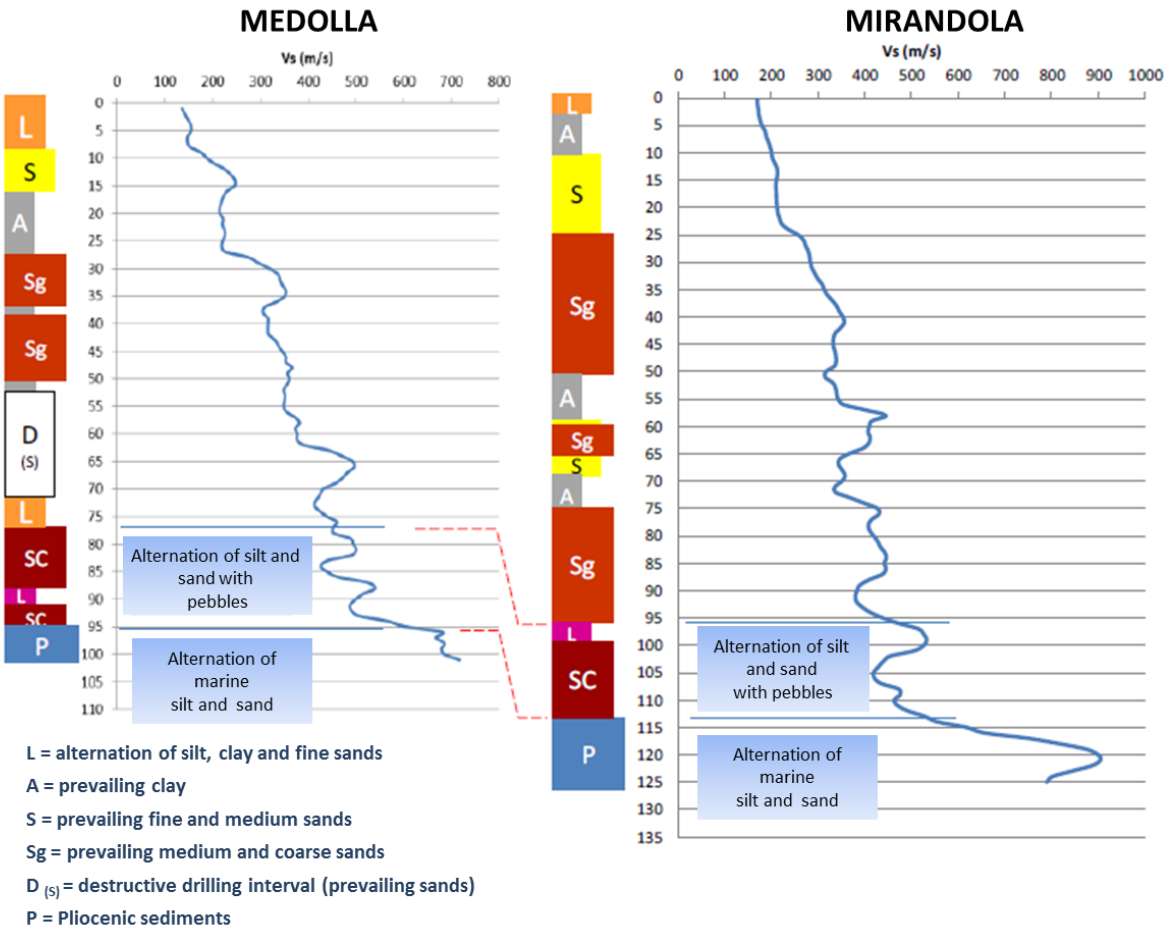

Fig. 6.9 - V_S profiles obtained by cross-hole tests in Medolla and Mirandola and respective borehole stratigraphy (from Paolucci *et al.*, 2015).

At these two sites, the $\overline{V_S}(h)$ curves have been determined by considering Eqs. 6.2 and 6.3. Best fitting curves in the form of Eq. 6.1 were obtained having respectively parameters V_0 = 80, x = 0.37 (Fig. 6.10-A) and V_0 = 80, x = 0.42 (Fig. 6.10-B). At these two sites f_0 were also determined from ambient vibration measurements and resulted in the range 0.8-0.9 Hz. When these values are used in Eq. 4.2 with parameters V_0 = 80 and x = 0.37, H values of 70 m at Medolla and 70–85 m (accounting for the uncertainty in the experimental f_0) at Mirandola were obtained. When the parameters V_0 = 80 and x = 0.42 are used, the estimates of H result 90 m for Medolla and 90–110 m for Mirandola, best fitting observations. This suggests using Eq. 6.3 to compute $\overline{V_S}(h)$ from borehole data available in the study area (Fig. 6.11).

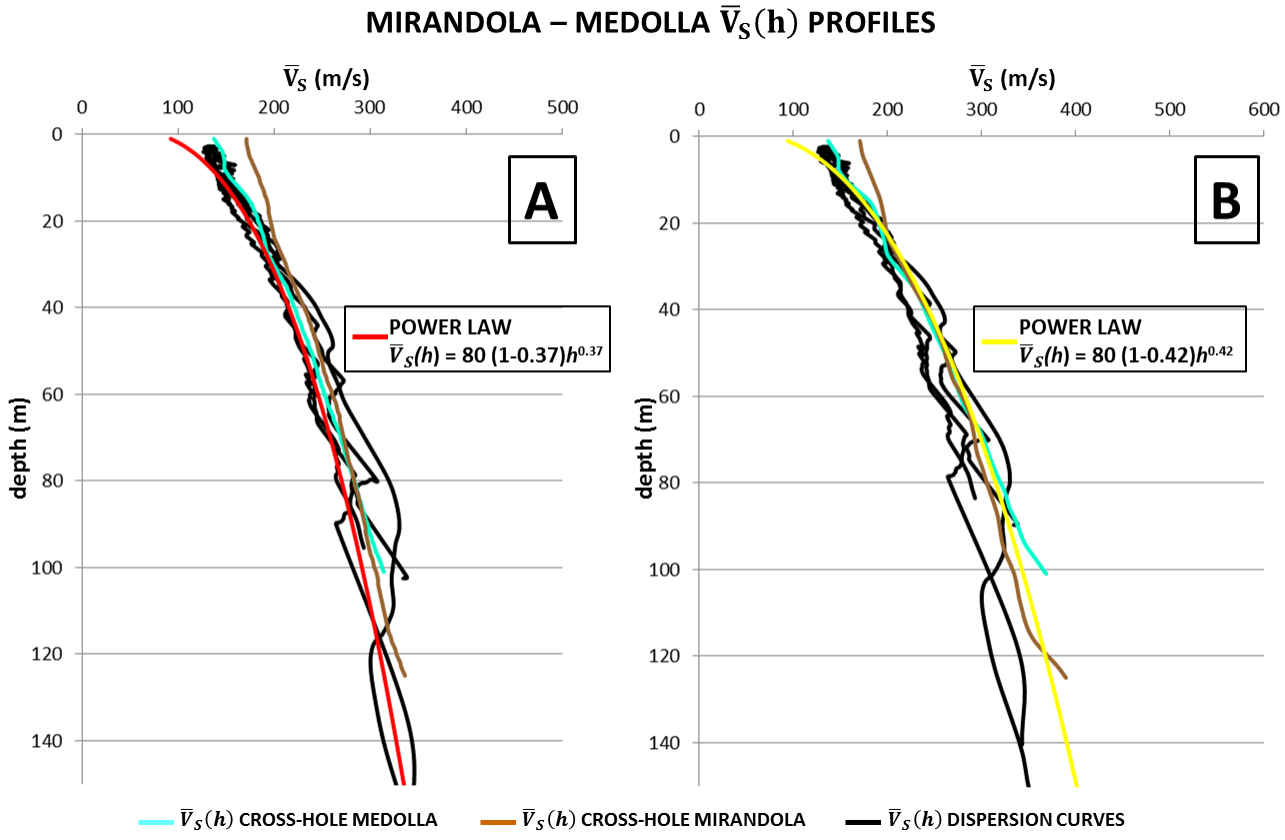

Fig. 6.10 - Average V_S profiles achieved in Mirandola-Medolla area. Light blue curve is the $\overline{V_S}(h)$ profile of Medolla cross-hole, the brown one is the $\overline{V_S}(h)$ profile of Mirandola cross-hole; these profiles are obtained using Eq. 6.2 (A) and using Eq. 6.3 (B). Black curves represent $\overline{V_S}(h)$ profiles obtained by the dispersion curves of the seismic arrays performed near the two boreholes; these profiles are obtained using c = 0.8 (A) and using c = 0.7 (B). Red (A) and yellow (B) curves are the power law pattern best fitting the observed $\overline{V_S}(h)$ profiles (from Paolucci *et al.*, 2015).

The $\overline{V_S}(h)$ profiles deduced at the Medolla and Mirandola sites were also considered to constrain the value of c to be used in Eq. 4.3 to deduce $\overline{V_S}(h)$ values where experimental dispersion curves were available. To this purpose, dispersion curves available at the two sites of Mirandola and

Medolla were considered. When Eq. 6.2 is used, a *c* value of 0.8 best fits observations (Fig. 6.10-A). A value equal to 0.7 is the best performing one where Eq. 6.3 is used (Fig. 6.10-B). On the basis of results obtained above, the value of 0.7 was considered the best performing and thus adopted to retrieve all the $\overline{V_S}(h)$ curves in correspondence of array measurements.

The envelope of the resulting $\overline{V_S}(h)$ curves is reported in Fig. 6.11 and reveals rather similar trends up to depths of the order of 150 m, with differences between average velocity values about 100 m/s.

On the basis of the above results, two power-law patterns were determined to bound observed $\overline{V_S}(h)$ profiles and capture observed variability. The lower bound is marked by a power law with parameters V_0 = 55 and x = 0.43; upper bound is marked by a power law with parameters V_0 = 120 and x = 0.3 (Fig. 6.11). For each value of the resonance frequency, these two curves allow to compute upper and lower bounds of possible depth of the resonant interface.

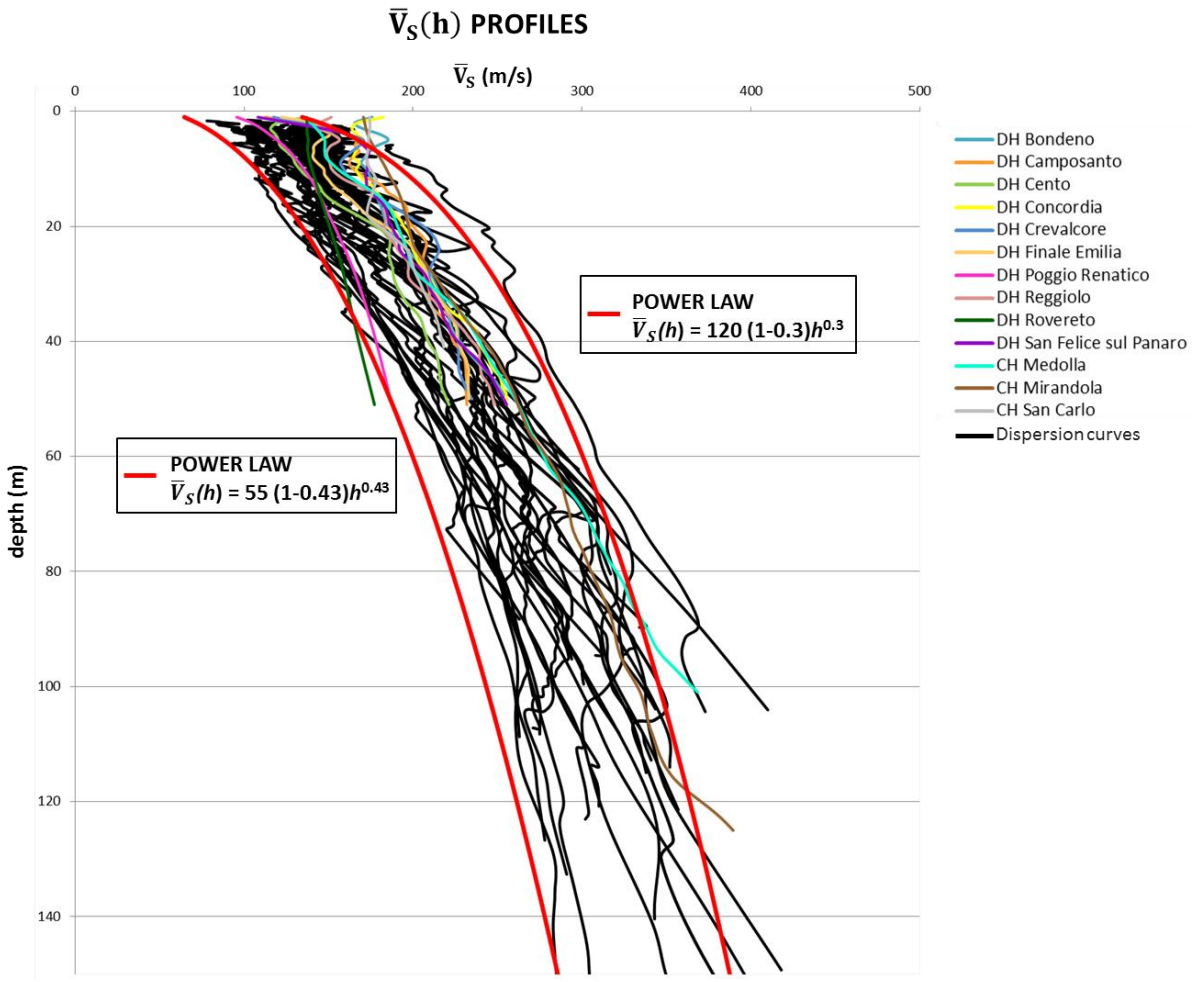

Fig. 6.11 - Average V_S profiles of the whole study area. Colored curves are related to $\overline{V_S}(h)$ profiles obtained by borehole seismic tests and black lines are $\overline{V_S}(h)$ profiles obtained by dispersion curves; red lines are minimum (V_0 = 55, x = 0.43) and maximum (V_0 = 120, x = 0.3) power law taken in account to assess the model variability (from Paolucci *et al.*, 2015).

6.5.2 Estimating depths of the resonant interfaces

By using the *f-H* conversion rule determined above, it results that the resonant surface responsible for the main peak in the Zones 1 and 2 is located at depth in the range 60 and 110 m. Broadness of this HVSR peak does not allow distinguishing significant differences in the whole study area and the above variability only depends on uncertainty in the parameters of Eq. 4.2. In other terms, this resonant interface can be considered as nearly flat and located at a depth in the range 60–110 m.
Within the Zone 3, instead, as mentioned previously, it has been possible to note clear frequency shifts due to sharpness of the main HVSR peak. This allowed the estimate of lateral depth variations of resonant interface. In particular, starting from Mirandola area, the depth of the interface gradually switches from 115 m (Fig. 6.9) to about 50 m moving westward (up to the area between San Possidonio and Rovereto) and to 50–70 m moving southward (up to San Giacomo in Roncole). At the boundary of Zone 1, this interface returns to larger depths (about 100 m), as shown in Medolla borehole (Fig. 6.9).
It can be inferred that the resonant interface relative to the low frequency peak (0.25-0.3 Hz) is deeper than the one relative to the main peak. To evaluate this depth, the *f-H* conversion rule cannot be used since it should imply the extension downward of a smooth power-law pattern from the surface down to this resonant interface. This, however, is not compatible with the presence of the impedance contrast associated to the main HVSR peak. Furthermore, data relative to deeper layers are not available and thus, any piecewise power law model (see Paragraph 4.2) cannot be adopted. Thus, very rough depth estimates for this layer can be provided. In particular, by assuming as a minimum V_S estimate the one relative to the deeper layer detected by the borehole test in Mirandola (about 500 m/s) and by applying the simple rule $H = V_S/4f$ (where f is the resonance frequency for the low frequency peak) one obtains depth values about 500 m.

6.5.3 Numerical testing

In order to evaluate feasibility of the above depth estimates in interpreting HVSR pattern in the three Zones, numerical simulations have been carried out by considering the last version of the full-wavefield DSS model (see Section 2.2.3) proposed by Lunedei and Albarello (2015). Three ‘‘representative’’ HVSR curves were obtained for each zone, by considering the average shape of the HVSR curves included in the relevant Zone (Fig. 6.12).
In particular, these seismostratigraphic models were used:

- as concerns the Zone 1 (Fig. 6.12-A), it was possible to simulate an HVSR curve very similar to the experimental curves average trend attributing to materials overlying the main resonant interface average V_S values of 270 m/s and a thickness of 90 m. Regarding to those ones located between the two interfaces, average V_S values of 600 m/s and a thickness of 500 m were considered. Beneath the stack of strata a halfspace with V_S values of 2000 m/s was inserted;

- as concerns the Zone 2 (Fig. 6.12-B), a similar model to the previous one was used. In particular, for the shallow materials average V_S values of 220 m/s and a thickness of 80 m were used; for to those ones located between the two interfaces, average V_S values of 500 m/s and a thickness of 500 m were considered. In this case a half-space with V_S values of 1500 m/s was inserted. To reproduce the slight ''hump'' between 1.5 and 2 Hz, a weak impedance contrast at about 20 m depth was added (Fig. 6.12-B);
- as concerns the Zone 3 (Fig. 6.12-C), the average HVSR curve was reproduced attributing to materials overlying the main resonant interface average V_S values of 280 m/s and a thickness of 60 m; regarding underlying materials average V_S values of 600 m/s and a thickness of 90 m were considered. Insertion of a further impedance contrast at 150 m depth, represented by a half-space with V_S values of 1000 m/s, allowed simulating HVSR peak sharpness.

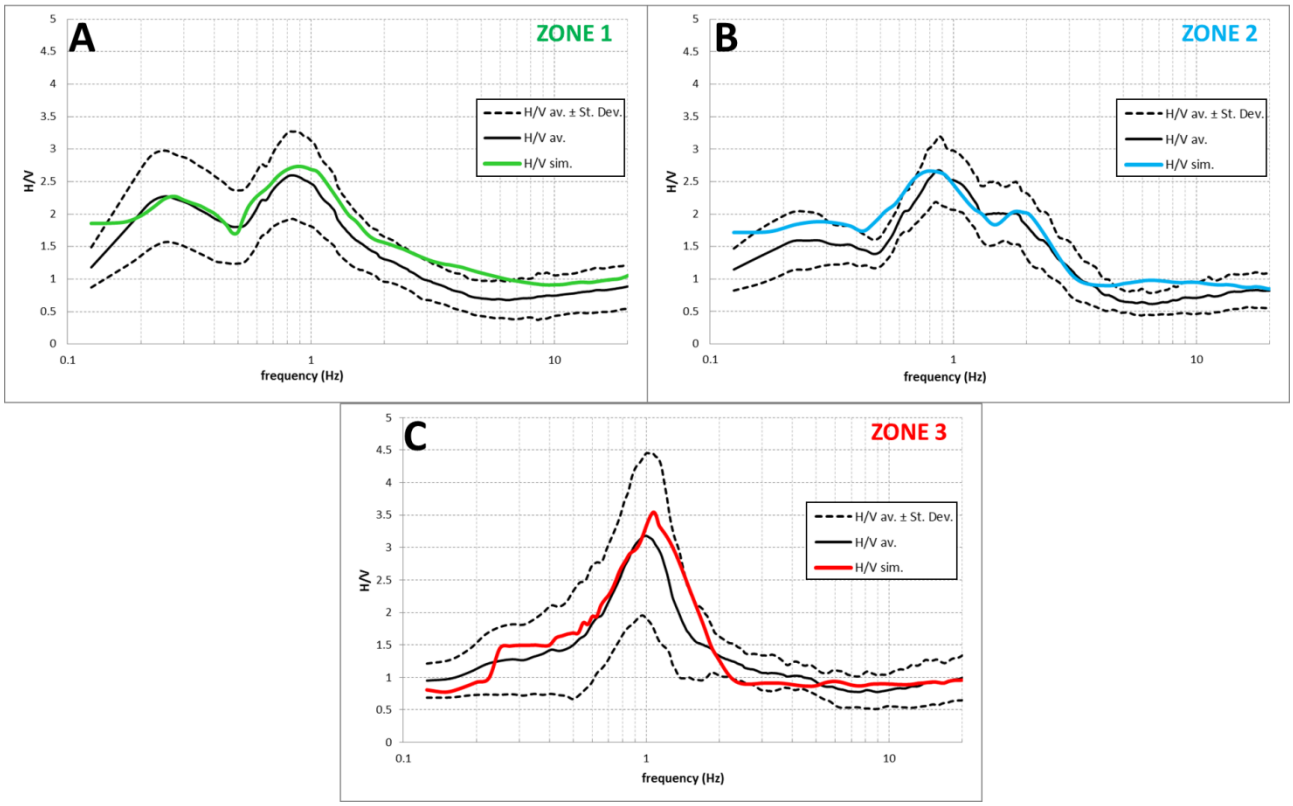

Fig. 6.12 - Outcomes of the numerical simulations achieved to simulate the HVSR curves average trend (continue black curves) of the Zone 1 (A), Zone 2 (B) and Zone 3 (C). Green curve (A) represents the simulated HVSR curve for the Zone 1, the blue one (B) for the Zone 2 and the red one (C) for the Zone 3. Dashed black curves delimit an interval about the mean with half-width equal to the standard deviation (from Paolucci *et al.*, 2015).

6.5.4 Geological interpretation

From the geological point of view, the nature of the resonant interface related to the main frequency peak is very similar to that one detected in the bank areas of the Po River (Martelli *et al*. 2014). Taking in account the Zone 1 and 2, this interface is likely related with a sub-horizontal unconformity within the Quaternary alluvial sequence. In particular, observing the geological sections of the investigated area (Fig. 6.3), the main impedance contrast could correspond to the contact, located in the depth range 60–110 m, between two Middle Pleistocene sedimentary cycles: the Bazzano Subsynthem (AES6) and the Undifferentiated Emiliano-Romagnolo Upper Synthem (AESind), both belonging to higher rank sedimentary cycle represented by Emiliano-Romagnolo Upper Synthem (AES).

Within these two Zones, broadness of the main frequency peak could suggest the presence of a relatively broad transition in the seismic impedance values at depth: in view of this, the resonant interface could be linked to a gradational geological contact and not to a real surface.

As concerns the low frequency peak, it can be attributed to a relatively deep interface (500–600 m) that is probably linked to an impedance contrast within the Quaternary marine deposits or to the presence of pre-Quaternary deposits (Fig. 6.3). It is unlikely that this peak is related to the contact between the Quaternary continental deposits and the underlying marine ones. In fact, this contact is located at 200–300 m depth (Fig. 6.3) and in the south part of the section A–A' (Fig. 6.3-A) it shows a gradually deepening from 150 m to 400 m: this feature is not detected by the HVSR measurements.

Regarding the geological interpretation of the slight ''hump'' identified by Zone 2 HVSR curves, most likely this weak resonant interface is linked to an interdigitation zone created by the alternation between sandy sediments of channel environments (related to the Po river proximity, Fig. 6.2; Regione Emilia Romagna, 1999) and clay-silty materials of floodplain environment.

Within the Zone 3, the subsoil geological setting changes in a more remarkable way: in fact, due to the presence of a buried ridge related with the Ferrara Arc structure (see Paragraph 6.2), the pre-Quaternary deposits rise up to arrive close and even to coincide to the main resonant interface detected in the Zone 1 and 2 (Fig. 6.3-A, 6.9). As suggested by the numerical simulations (Fig. 6.12-C), this feature could explain the presence of the single HVSR maximum detected in this area and, moreover, its sharpness and its higher amplitude values than the main HVSR peaks identified in the rest of the study area. Comparable geological setting is present in Casaglia zone (Fig. 6.3-B), just outside of the NE part of the study area; some single station measurements performed in this zone show HVSR curves very similar to those ones obtained within Zone 3, with a single sharp peak at 0.9 Hz (Picozzi and Albarello, 2007).

CHAPTER 7
Automatic identification of large scale heterogeneities by using PCA analysis of HVSR data

7.1 Introduction

The study carried out in Chapter 6 witnesses the good capability of the HVSR technique in identifying large scale areas (of the order of tens-hundreds square kilometers) where analogous seismic response is expected in case of future potentially damaging earthquakes. Obviously, this characteristic is strictly connected with the different geological configurations present in the whole study area, that in the case of Emilia Romagna are mainly related with the geological characteristics of the buried resonant interfaces rather than with the outcropping materials.

The cheapness and the quickness of the single-station ambient vibration measurements allowed to collect hundreds of HVSR curves after the Emilia Romagna seismic sequence and after other damaging earthquakes (e.g. L'Aquila earthquake in 2009: see Albarello *et al.*, 2011; Gallipoli *et al.*, 2011), as well as in the framework of the Seismic Microzonation studies that are involving a large number of municipalities all over the Italian country. To manage this huge amount of data, the availability of automatic procedures aimed to group HVSR curves may be of great help in order to point out the possible large scale geological heterogeneities. To perform this kind of procedure, some studies based on cluster analysis exist (e.g., Rodriguez and Midorikawa, 2002; Bragato *et al.*, 2007; Ullah *et al.*, 2013; Capizzi *et al.*, 2015) and they will be shortly outlined in the next Paragraph.

The aim of this Chapter is to investigate the automatic grouping of the HVSR curve by using the Principal Component Analysis (PCA). With respect to cluster analysis, a main advantage of PCA is that it allows the grouping of sites characterized by similar HVSR curves by also identifying the relevant "characteristic" HVSR patterns. In order to explore the reliability, the effectiveness and the drawbacks of the PCA application for the purposes described above, this technique was applied in different geological contexts where several single station measurements were collected.

After the brief description of the previous works based on cluster analysis, a general description of the PCA technique is done. Then, the application of this method on the HVSR curves is treated, describing in particular the adopted criteria to optimize the partition of the dataset. Finally, the results of the PCA application in different dataset are described. In particular, the first attempt is done in the Chapter 6 study area (Emilia Romagna region, Italy), with the aim to compare the PCA outcomes with the HVSR Zones detected by visual inspection (see Section 6.3.1). After that, the PCA reliability is checked in two study areas in the Central Italy characterized by more complex geological settings, i.e. Collesalvetti and Montecatini Terme municipalities.

7.2 HVSR grouping using cluster analysis

This paragraph is devoted to a short review of the works where an automatic zonation of the HVSR curves based on cluster analysis is performed. This kind of automatic grouping is here briefly described in order to carry out a comparison with PCA technique described in the next paragraphs, with the aim to point out the drawbacks and the benefits of both analyses.

The cluster analysis is a multivariate statistic technique that allows to group the statistical units identifying some subsets of objects, called clusters, in order to maximize the similarity within them. The level of similarity (internal cohesion) is evaluated according to some criteria taking into account the values that a number of variables chosen take in each statistical unit. Generally, in the analysis for grouping is not necessary to have in mind any interpretative model (Fabbris, 1983). The partition is successful if the objects within the clusters are closer to each other than other in different clusters (Barbarito, 1999).

Many clustering algorithms exist (e.g., Gan *et al.*, 2007; Everit *et al.*, 2011). Essentially, they can be categorized into two main types: hierarchical clustering and non-hierarchical clustering methods.

The first ones make subsequent mergings of the data: for this, they are explorative methods and it is not necessary to define a priori the number of clusters. The main characteristic of the hierarchical methods is that an object can not be assigned to another cluster during the procedure: that is, once an object is joined to a cluster, it is not longer removed. The whole procedure relies on the definition of the assignment criterion of the unity to the clusters, which is chosen so as to be suitable with the studied subject and the data nature. One of the results of hierarchical clustering methods is the dendrogram which shows the progressive grouping of the data: by this representation it is possible to gain an idea of a suitable number of classes into which the data can be grouped, but on the other hand it is not always easy to choose the suitable level of similarity where the dendrogram has to be truncated.

Unlike the procedures described previously, the non-hierarchical methods need to define a priori the number of clusters: this operation allows to reduce the computational efforts of the procedure. In this way, the algorithm generates a single partition of the statistical units and it does not need to explore the best aggregation between the clusters at each step. Unlike the hierarchical methods, the assignment of an object to a cluster is not definitive. Usually, after choosing the initial partition, the algorithm attempts to improve it as a function of the considered internal variance minimization criterion.

Rodriguez and Midorikawa (2002) used the Ward hierarchical clustering algorithm to group both the HVSR and the SSR (Standard Spectral Ratio; Borcherdt, 1970) curves obtained at 145 sites located in Yokohama city (Japan). This procedure was aimed at increasing the objectivity of the outcomes interpretation and to verify the agreement between the two kind of spectral ratios. The mutual similarities amongst the sites are provided by the calculation of Euclidean distances between the two spectral ratios. Therefore, the shorter the distances, the stronger is the resemblance among sites and vice versa. The Euclidean similarity coefficient establishes

similarities as the distance (d_{ij}) between two sites *i* and *j* within a space E_K, with dimensions defined by the number of *K* attributes (A_K) employed in the analysis, as expressed by this relationship (Rodriguez and Midorikawa, 2002):

$$d_{ij} = \sqrt{\sum_{K=1}^{34}\left(A_{Ki} - A_{Kj}\right)^2}. \tag{7.1}$$

The 34 parameters *K* in this equation are represented by the predominant periods and maximum spectral ratio values attained from HVSR and SSR methods, as well as 15 spectral ratios values representing the contribution at different frequencies from 0.5 to 10 Hz. The computation and arrangement of all possible Euclidean distances among sites according to Eq. 7.1 takes the form of a similarity or dissimilarity matrix. The representation of the results of this hierarchical clustering process in graphical form leads to the attainment of the dendrogram illustrated in Fig. 7.1-A. This graph illustrates at which similarity levels the 145 sites appear to gradually group into increasingly similar sub-clusters, until reaching the bottom level, where all sites stand alone and similarity equals 1.

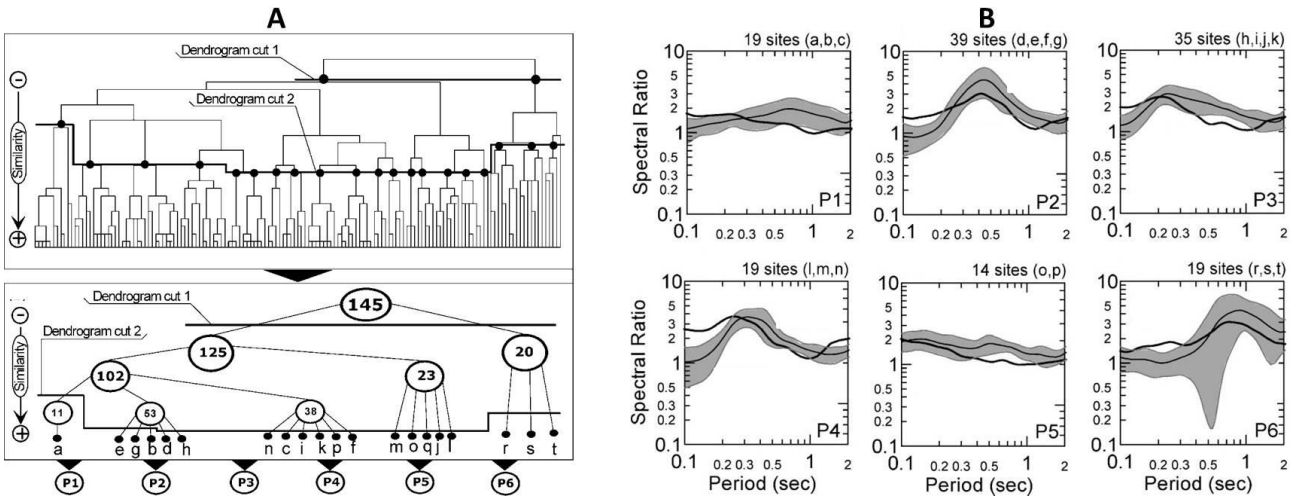

Fig. 7.1 - A: grouping attained from the application of the cluster analysis. The upper panel shows the dendrogram with the relative similarity amongst sites; the lower panel shows the grouping scheme with the division into 20 groups of sites sharing very similar characteristics. B: generalization of the cluster analysis results into six patterns. The thick line represents the average SSR curve, the thin line the average H/V curve and the shaded area the H/V spectra ±1 standard deviation (modified from Rodriguez and Midorikawa, 2002).

At dendrogram cut "2" the clusters composed of 11 and 20 sites (Patterns 1 and 6) appear robust in terms of both coincidence and shape of the spectral ratios. However, for the rest of the clusters this similarity of spectra does not appear meaningful until the authors group the remaining 53, 38, and 23 sites into smaller sub-clusters, as shown at dendrogram cut 2 (Fig. 7.1-A). Finally, the

authors merged the dendrogram analysis results with their own academic judgement to regroup the resulting 20 clusters into the generalized patterns illustrated in Fig. 7.1-B and obtained averaging the relevant spectral ratios. The locations of the measure points belonging to these final patterns seem to show a good agreement with the local geology, pointing out its heterogeneities.

Bragato *et al*. (2007) proposed a more sophisticated approach where the automatic zonation problem is posed as one of optimal partitioning. As the above mentioned work, the authors try to subdivide the available set of measurement points into disjointed subsets so that the similarity among the HVSRs in each subset is maximized, trying to keep the number of subsets (clusters) as low as possible. This approach has the further requirement that each subset must correspond to a connected region of the study area. This geometric constraint is formalized by concepts derived from graph theory: the measurement points are connected in a Delaunay triangulation and the only acceptable subsets are those that correspond to a connected subgraph of the triangulation. To reduce the computational complexity of the problem, the authors restrict the search to a particular class of subgraphs of the triangulation, called its spanning trees (e.g., Kruskal, 1956). In particular, they assume that each point of the triangulation is represented by a vector $y_i = (y_{i,1}, \ldots, y_{i,M})$ whose components are the HVSRs computed at the M frequencies $f_1, \ldots, f_M$. By removing edges, the authors partition the Delaunay triangulation into N_C connected subgraphs or clusters, each one including N_k measurement points ($k = 1, \ldots, N_C$). The edge of the triangulation with vertices x_m and x_n is weighted with the Euclidean distance of the corresponding HVSR vectors (Bragato *et al.*, 2007):

$$\|y_m - y_n\| = \sqrt{\sum_{j=1}^{M} \left(y_{m,j} - y_{n,j}\right)^2}. \tag{7.2}$$

After that, in order to give a measure of the internal dissimilarity of the k^{th} connected cluster in a partition, they used the non-hierarchical *K*-means clustering approach (MacQueen, 1967) and considered the within-group sum of squares $WGSS_k$, defined as (Bragato *et al.*, 2007):

$$\mathrm{WGSS}k = \sum_{i=1}^{N_k} \|y_i - y_{Ck}\|^2, \tag{7.3}$$

where y_{Ck} is the centroid of the vectors y_i in the cluster. The optimal partition for fixed N_C is that minimizing the sum of $WGSS_k$ and the Bayesian Information Criterion (BIC; Schwarz, 1978). In this approach, these two functions are minimized over the space of the connected partitions of a Delaunay triangulation using a procedure based on the Genetic Algorithm (see Section 4.3.2), where increasing values of N_C are explored for different runs.

The authors applied this approach in a dataset of 102 HVSR measurements collected in Vittorio Veneto area (Northeastern Italy). Each run of the whole algorithm for a fixed N_C value took about 1 hour. The resulting zonation (Fig. 7.2-A) is characterized by a good agreement with the local

geological changes; the corresponding HVSR patterns (obtained averaging the curve belonging to each zone) are showed in Fig. 7.2-B. As noted by the authors, the automatic procedure is not able to create a separated area for the curves of the pattern 3b (even using a higher number of clusters), whose positions are pointed out in the map by diamonds with black circles; indeed, these curves were inserted by the algorithm within the pattern 3 (Fig. 7.2-B).

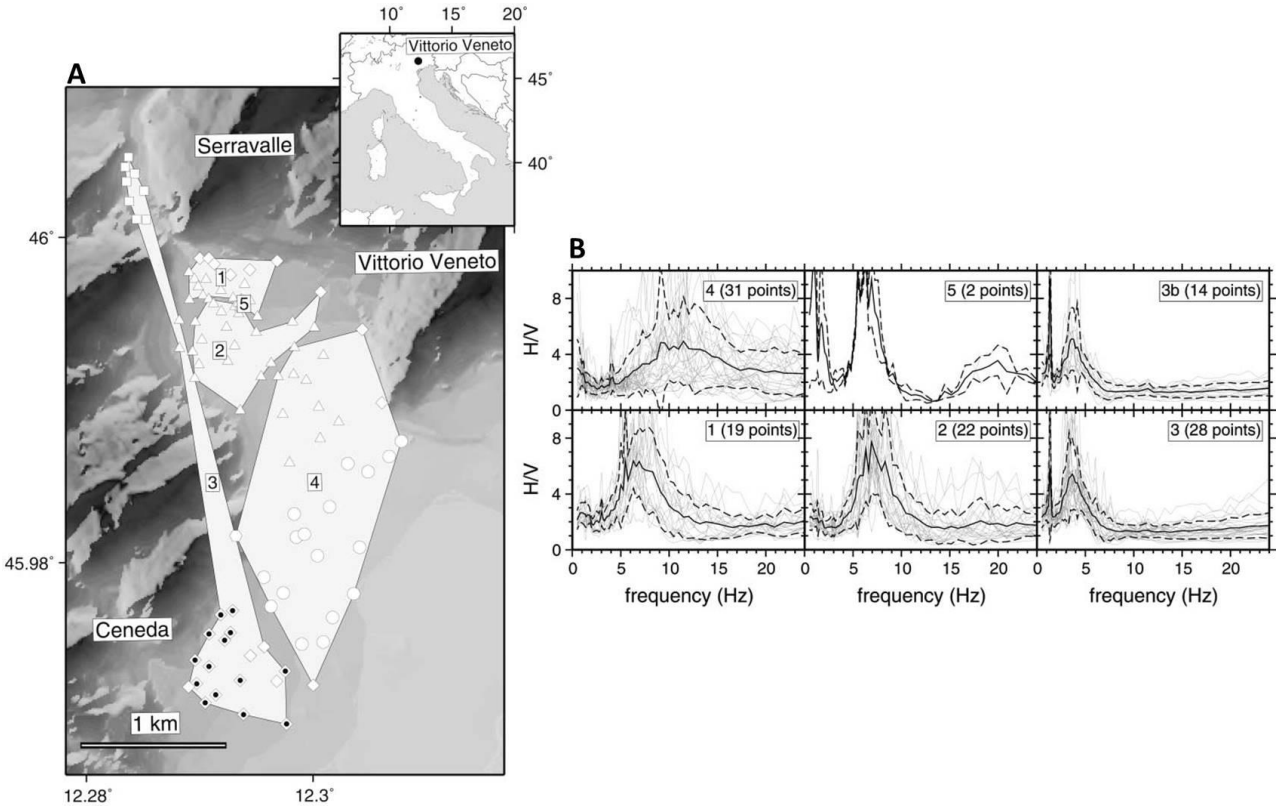

Fig. 7.2 - Panel A shows the optimal zonation estimated for the Vittorio Veneto area (five zones, with zone 5 including only two points). The symbols of the measurement points represent different geologic conditions. The diamonds with black circles indicate the points of class 3 that have a strong peak at about 1.3 Hz (class 3b in panel B). Panel B shows HVSRs for each zone (thin gray lines). The thick and the dashed lines show the average HVSR ± one standard deviation in each zone (modified from Bragato *et al.*, 2007).

Ullah *et al.* (2013) applied the non-hierarchical *K*-means clustering algorithm in order to group the both HVSR and SSR curves obtained by 15 temporary stations installed in Bishkek (Kyrgyzstan). Using this approach no a priori geographical location information was taken into account. As described previously, this algorithm requires that the number of clusters (*k*) is defined a priori; therefore *K* centroids are then defined randomly to represent each cluster at each frequency. The dissimilarity between an object (in this case the value of the spectral ratios at each frequency) and the centroid of each cluster is calculated by means the squared Euclidean distance. After the assignment of all the objects to the cluster whose centroid is nearest to the object, a new series of *k* centroids for each frequency are re-calculated based on the already formed clusters.

Consequently, the objects are again assigned to the new centroids based on the shortest Euclidean distance. The process continues iteratively until there is no further update of the centroids. In particular, the aim of this cluster algorithm is to minimize the sum of the squared error over the whole dataset given as (Ullah *et al.*, 2013):

$$\mathrm{E} = \sum_{i=1}^{k} \sum_{l=1}^{n} \sum_{j=1}^{m} \left\| x_{lj,i} - c_i \right\|^2, \tag{7.4}$$

where *m* is the number of objects (in this case, the number of frequencies examined) in a data set, *n* is the number of observations (or spectra) in each cluster, *k* the number of clusters, c_i is centroid of the cluster *i*, and $x_{lj,i}$ is the value of *j* object in cluster *i* of observation number *l*. The authors repeated the analysis 10 times, using different randomly chosen starting centroids and the resulting solution that provided the minimum of the sum of the squared error was chosen as the best one. The resulting zonation correlates well with the main geological features in the area, pointing out variability from south to north, consistent with both the changes in the thickness of the sedimentary cover and in the Quaternary material outcropping at the surface. Moreover, the zonation relative to the HVSR curves shows a very similar spatial distribution to that one relative to the SSR curves.

Another interesting approach is that one proposed by Capizzi *et al.* (2015), where the authors applied a two step cluster analysis on the HVSR data.

The first step is devoted to identify the correct time windows to be used for the calculation of the mean HVSR for each measure, in order to avoid the difficulties that can arise by a simple visual inspection. In particular, the authors used a hierarchical clustering algorithm on the HVSR curves obtained in each time window for each measure using as proximity measure the Standard Correlation (SC_{xy}) defined as (Capizzi *et al.*, 2015):

$$SC_{xy} = \sum_{i=1}^{n} x_i \cdot y_i / \sqrt{\sum_{i=1}^{n} {x_i}^2 \cdot \sum_{i=1}^{n} {y_i}^2}, \tag{7.5}$$

where x_i and y_i indicate the values of the spectral ratios relative to the i^{th} frequency and the generic pair of analysis windows. In this approach the cutting level of the resulting dendrogram was chosen on the basis of some criteria that involve the width of the gap identifiable between two successive levels of the hierarchy detectable in the dendrogram (Gan *et al.*, 2007; Everit *et al.*, 2011).

The second step, devoted to perform a zonation grouping similar HVSR curves, is carried out by means a non-hierarchical centroid-based algorithm. This procedure is similar to that one performed by Ullah *et al*. (2013), but in this case the distance *D* of each unit from the initial centroids is calculated as the weighted sum of the Euclidean normalized distances of different kind of variables. These are coordinates (*x*, *y* and *z*), frequency (*f*), amplitude (*A*) and lithology (*L*) (Capizzi *et al.*, 2015):

$$D = a\sqrt{(dx^2 + dy^2 + dz^2)} + b\sqrt{df^2} + c\sqrt{dA^2} + d\sqrt{dL^2}, \quad (7.6)$$

where *a*, *b*, *c* and d are the weights. The choice of weights and of the optimal number of *k* classes is optimized maximizing *R* parameter, taking into account the intra-cluster (DEV_{IN}) and inter-cluster (DEV_{OUT}) variances (Capizzi *et al.*, 2015):

$$R = DEV_{OUT}/DEV_T, \qquad DEV_T = DEV_{OUT} + DEV_{IN} = 1. \quad (7.7)$$

This approach seems strongly constrained by the location of the measurements and it is mainly based on the characteristics of the peak (frequency and amplitude) individuated as local maxima in the HVSR function. The use of such features could be not optimal if, for example, the HVSR curves are characterized by broad maxima; that is, the procedure should work for "standard" curves (with one or more well defined peaks) and not for the generality of cases.

7.3 The PCA approach

Principal Component Analysis (PCA) is a technique used in the framework of multivariate statistics and is adopted in many scientific fields, like astronomy, atmospheric physics and geophysics in general. The main purpose of this technique is the simplification of a source dataset containing a large number of variables creating a new dataset with fewer (optimistically many fewer) new variables. These new variables are linear combinations of the original ones, and these linear combinations are chosen to represent the maximum possible fraction of the variability contained in the original data. That is, given multiple observations of a (1 $\times K$) data vector $\boldsymbol{x}$, PCA finds (1 $\times$ *M*) vectors $\boldsymbol{u}$ whose elements are linear combinations of the elements of the vectors $\boldsymbol{x}$, which contain most of the information in the original collection of vectors $\boldsymbol{x}$ (Wilks, 2006). PCA is most effective when this data compression can be achieved with $M << K$. The elements of these new vectors $\boldsymbol{u}$ are called the Principal Components (PCs).

The linear transformation of the variables projects the original ones on a new Cartesian system where the variables are ordered in descending order of variance: therefore, the variable with higher variance is projected on the first axis, the second one on the second axis and so on. Complexity reduction occurs by merely analyzing the main new variables, that is that ones with higher variance. Unlike other linear transformations used for statistical purposes, in this technique the transformation vectors are determined by the source data itself.

Extensive treatments of PCA are given by Preisendorfer (1988) and Joliffe (2002); moreover, descriptions of this method concerning the geoscience field are present in Davis (2002) and Wilks (2006).

Beyond a simple data compression, PCA can be a very useful tool for exploring large multivariate data sets, in order to provide insights into both the spatial and temporal variations exhibited by the analyzed variables. Moreover, new interpretations of the original data can be suggested by the nature of the linear combinations that are most effective in compressing the data.
Usually, the principal components are computed as linear combinations of centered data $\boldsymbol{x}'$, that is $\boldsymbol{x}' = \boldsymbol{x} - \overline{\boldsymbol{x}}$, where $\overline{\boldsymbol{x}}$ is the mean of the variable observations.
The first PC, u_1, is the linear combination of $\boldsymbol{x}'$ characterized by the largest variance of the original variables. The successive principal components u_m (where *m* = 2, 3,…., *M*) subsequently represent the largest possible variances of $\boldsymbol{x}$ after the first component. The PCs are subject to the condition that they are mutually uncorrelated.
PCs (that is the elements u_m of $\boldsymbol{u}$ that maximize the joint variability of $\boldsymbol{x}'$) are uniquely defined by the eigenvectors of the covariance matrix of $\boldsymbol{x}$, $[S]$. In particular, the m^{th} principal component u_m is obtained as the projection of the data vector $\boldsymbol{x}'$ onto the m^{th} eigenvector $\boldsymbol{e}_m$ (Wilks, 2006):

$$u_m = \boldsymbol{x}'\boldsymbol{e}_m = \sum_{k=1}^{K} x'_k \, e_{km}\,, \quad m = 1, \ldots., M. \tag{7.8}$$

In the above equation, each m^{th} PC is computed from a particular set of observations of the *K* variables x_k; therefore, each of *M* PCs can be understood as a sort of weighted average of the x_k values. The weights are represented by the elements (the e_{km} values, usually named as "loadings") of the *M* eigenvectors $\boldsymbol{e}_m$; considering that each eigenvector has unitary module, the squared sum of its elements is equal to 1. The higher the e_{km} absolute value, the greater is the weight that the x'_k (and therefore x_k) have in determining the m^{th} principal component u_m. This means that the principal component u_m is much more characterized by the x'_k variables associated to the e_{km} coefficients with greatest absolute value. If the data sample consists of *n* observations (and therefore of *n* data vectors $\boldsymbol{x}$, or *n* rows in the data matrix $[X]$), there will be *n* values for each of the principal components (Wilks, 2006). The individual values of the PCs are usually defined as "scores".
From the geometrical point of view, the data matrix $[X]$ can be represented as *n* points in a *K*-dimensional space. Since that PCA is aimed to reduce the variable number being analyzed, this operation can be understood as the projection of the *n* point in an *M*-dimensional subspace. In particular, the orthogonal square matrix $[E]$ whose columns are the *K* eigenvectors defines this rigid rotation (Wilks, 2006):

$$\boldsymbol{u} = \boldsymbol{x}'[E], \tag{7.9}$$

which is the simultaneous matrix-notation representation of the Eq. 7.8. Thus, the PCs individuate a new coordinate system where its first axis (defined by the first eigenvector $\boldsymbol{e}_1$) represents the maximum variability of the original dataset and its second axis (constrained to be perpendicular to the first one and defined by the second eigenvector $\boldsymbol{e}_2$) represents the second largest variability

(Fig. 7.3). Subsequent axis (and therefore the subsequent eigenvectors $\boldsymbol{e}_m$, where m = 3, 4,…, M) in turn will be perpendicular to all the previous ones and they will continue to locate directions in which the original data jointly exhibit maximum variability. Obviously, each m^{th} eigenvector $\boldsymbol{e}_m$ is associated to the corresponding m^{th} eigenvalue λ_m of the covariance matrix $[S]$.

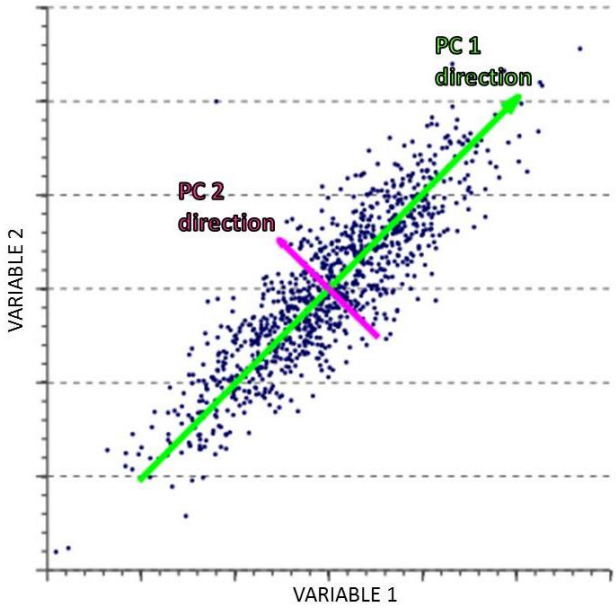

Fig. 7.3 - Example of principal components on bivariate dataset. The green axis identifies the first PC and it points in the direction of the largest variability of the data. The violet axis (perpendicular to the green one) represents the second PC, where the second largest variability is located.

The total variance of the original dataset is defined by the sum of the variable individual variances. Since the variances are located along the diagonal of $[S]$, this is equivalent to finding the trace $tr[S]$ of this square matrix. This parameter is equivalent to the sum of the K eigenvalues of $[S]$ and therefore to the trace of the diagonal matrix $[\Lambda]$ whose nonzero elements are the same eigenvalues of $[S]$ (Wilks, 2006):

$$tr[S] = \sum_{k=1}^{K} S_{k,k} = \sum_{k=1}^{K} \lambda_k = tr[\Lambda], \tag{7.10}$$

where $S_{k,k}$ and λ_k are respectively the K variances and the K eigenvalues of $[S]$.
Since the PCs are uncorrelated each other, the covariance matrix $[S_u]$ of the u_m's is obtained by the diagonalization of $[S]$ and it is simply equivalent to the diagonal matrix $[\Lambda]$ (Wilks, 2006):

$$[S_u] = [E]^T [S][E] = [\Lambda]. \tag{7.11}$$

In view of the two equations above, it is possible to state that the variance of the m^{th} principal component u_m is the m^{th} eigenvalues λ_m and that each PC represents a share of the total variation (R_m) in $\boldsymbol{x}$ that is proportional to its eigenvalue (Wilks, 2006):

$$R_m = \frac{\lambda_m}{\sum_{k=1}^{K} \lambda_k} \times 100\% = \frac{\lambda_m}{\sum_{k=1}^{K} S_{k,k}} \times 100\%. \tag{7.12}$$

Considering Eq. 7.9, if $[E]$ contains all *K* eigenvectors of $[S]$ as its columns, the resulting vector $\boldsymbol{u}$ will also have dimension (1 × *K*). This transformation can be reversed in order to reconstruct the original (1 × *K*) data vector $\boldsymbol{x}'$ from the principal component according to (Wilks, 2006):

$$\boldsymbol{x}' = \boldsymbol{u}[E]^T. \tag{7.13}$$

Consequently, Eq. 7.8 can be expressed as:

$$x'_k = \boldsymbol{u}\,\boldsymbol{e}_k{}^T = \sum_{m=1}^{K} u_m\, e_{km}\,, \quad k = 1, \ldots, K \tag{7.14}$$

Since in the above equation the full set of *M* = *K* PCs is used, the reconstruction is complete and exact. If *M* < *K* PCs are used, $[E]^T$ becomes a (*M* × *K*) matrix and $\boldsymbol{u}$ a (1 × *M*) vector. This means that the reconstruction of $\boldsymbol{x}'$ is approximate and the Eq. 7.14 can be rewritten as (Wilks, 2006):

$$x'_k \approx \boldsymbol{u}\,\boldsymbol{e}_k{}^T = \sum_{m=1}^{M} u_m\, e_{km}\,, \quad k = 1, \ldots, K. \tag{7.15}$$

It can often happen that the original available data are characterized by measure units not comparable among them or by very different subsamples amplitudes. In such conditions, it is necessary to standardize the original data normalizing the samples of the source variables with their standard deviation. In this case, the covariance matrix will coincide with the correlation matrix of the source dataset. It is worth mentioning that the eigenvectors (and therefore the PCs) obtained with the correlation matrix are different from those ones obtained considering the covariance matrix.

7.4 PCA on HVSR curves

7.4.1 Principal components computation

As mentioned previously, the aim of this Chapter is to investigate the automatic grouping of a HVSR curve dataset by means the PCA identifying its representative patterns.

In this case the source dataset is represented by a (*n* × *K*) matrix $[X]$ where the *K* columns refer to each measurement point (where the HVSR analysis was performed) and the rows indicate the *n* H/V ratio values computed for each considered frequency. Thus, in this application, the data matrix $[X]$ (and therefore $[X']$) is dimensioned as a (frequency × space) matrix, since data at *K* locations in space are "sampled" at *n* different frequencies.

Considering the form of Eq. 7.8 and 7.9, the original data can be arranged in one-dimensional (1 × *K*) vector $\boldsymbol{x}$, where each location (from 1 to *K*) refers to the appropriate element in the vector. Since that the original data consists of *K* spectra, the Eq. 7.8 and 7.9 can be expressed as a function of frequency *f*:

$$\boldsymbol{u(f)} = \boldsymbol{x'(f)}\ [E], \tag{7.16}$$

or in scalar form:

$$u_m(f) = \sum_{k=1}^{K} x'_k(f)\ e_{km}\,, \quad m = 1, \ldots., M. \tag{7.17}$$

In the above equations, the frequency index *f* runs from 1 to *n*. The reversed equations (Eq. 7.13, Eq. 7.14 and Eq.7.15) can be written using the same notation.

Considering the Eq. 7.17, it is possible to state that the e_{km} values (the so-called loadings) represent the contribution that the single H/V measurement $x_k(f)$ gives to the m^{th} principal component $u_m(f)$.

If the original data consist of a set of spectra, then the principal components are also expressed as spectra: in fact, also the resulting principal components matrix will be constituted by (*n* × *K*) elements (that is a (frequency × space) matrix). Plotting each element of the $u_m(f)$ principal component (the "scores") as a function of the respective frequency value, it is possible to visualize the pattern of an "apparent" HVSR curve: it follows that, once identified the representative *M* principal components according to their respective variance, it is possible to visualize the H/V patterns more representative in terms of the whole original dataset variance. These apparent curves (blue curves in Fig. 7.4) resemble the classical patterns of the HVSR curves (characterized by peak and troughs) but they present significant amplitude differences respect to the experimental ones. In fact, the high amplitude values (of the order of tens) reached by the "apparent" HVSRs are related to the variance of the PC taken into account, which is equal to the corresponding eigenvalue: it follows that the higher the overall amplitude values, the higher is the explained variance percentage of the considered PC. Moreover, the centering of the original data causes the evident oscillation of the apparent curves between positive and negative amplitude values.

From the computational point of view, the PCA has been carried out by using the "princomp" MATLAB routine, which uses the singular value decomposition (SVD) on the centered original data matrix: this algorithm is equivalent to the procedure explained before, but numerically more

stable (see, e.g., Joliffe, 2002; Wilks, 2006). This routine provides the values of the loadings and scores, as well as the vector containing the eigenvalues of the covariance matrix of $[X]$. The latter information is needed to accomplish the graph which shows the original dataset variance (and the cumulative one) percentage explained by each principal component. The whole algorithm is characterized by a remarkable quickness: in a dataset with 100-200 measurements, the computation takes only few seconds.

To perform the PCA, each HVSR curve was considered in the frequency range 0.1-10 Hz in order to do not account of spatial random geological heterogeneities related to the first 2-3 meters of the subsoil (often associated with the presence of anthropic material). Moreover, each H/V ratio was performed following the procedure described in Paragraph 2.3. An important characteristic connected with the H/V computation procedure is related with its frequency sampling: within the considered range, in the interval 1-10 Hz the frequency sampling is considerably greater than that one in the interval below 1 Hz. In order to make the sampling more uniform giving the same importance to the HVSR characteristics present in both interval, it was considered appropriate to perform a rescaling of the experimental curves based on logarithmic interpolation before accomplishing PCA.

Finally, since all the original variables (that is the *K* H/V spectral ratios) are expressed in the same measure unit with comparable amplitude values, it was not considered necessary to standardize them computing the correlation matrix of the source dataset.

7.4.2 Grouping criteria of the dataset

After that the main "apparent" HVSR patterns able to jointly represent almost the whole of the original data variance are identified, the next step was to look for similarities between these trends and those ones of the individual experimental HVSR curves. This step was needed to perform groupings between similar curves in order to point out the large scale geological heterogeneities of the investigated area.

Usually, in the contexts when the *K* elements of $\boldsymbol{x}$ are measurements at different locations in space, the loadings can be displayed graphically in a quite informative way. In fact, each eigenvector constituting the matrix $[E]$ contains exactly *K* loadings, which have a one-to-one correspondence with each of the *K* locations in the product from which the corresponding principal component is calculated (Eq. 7.17). Thus, each loading e_{km} can be plotted on a map at the same location as its corresponding $x_k(f)$ HVSR curve: such maps depict which locations contribute most strongly to the respective principal components. Substantially, the higher the e_{km} absolute value, the stronger is the measure contribute to the considered PC. If the e_{km} has negative sign, the associated experimental HVSR will resemble the opposite trend of the considered "apparent" H/V curve (red curves in Fig. 7.4), that for facilitation are plotted together in the subsequent analyses.

However, this kind of map visualization presents two practical drawbacks:

1. since the loading values associated to the respective measure point change depending on the principal component taken into account, a map for each considered $u_m(f)$ is needed. It follows that, the higher the number of principal components considered representative of the original dataset variance, the higher is the number of the maps to be displayed;
2. the choice of the number *M* of the principal components to retain may not be easy. The use of truncation criteria (e.g., Wilks, 2006) can not be enough to avoid subjective choices.

To overcome these problems, it is useful to see the formulation that relates the principal components and the original data in the form of Eq. 7.14 (considering the notation of Eq. 7.17). In this case, the e_{km} values represent the contribution that the m^{th} principal component $u_m(f)$ gives to reconstruct each single H/V measurement $x_k(f)$. Thus, taking into account this formulation, it is possible to individuate which of all the principal components gives the greater contribution to reconstruct each single HVSR curve and therefore which principal components is more similar to the considered H/V curve. In fact, also in this case, the *K* loadings constituting the eigenvectors of $[E]$ have a one-to-one correspondence with each of the *K* principal components u_m in Eq. 7.14.

To perform a correct identification of the dominant PC, it was considering appropriate to multiply the loading values with the maximum excursion of the scores of each PC (i.e., the amplitude maximum excursion of the "apparent" HVSR curve). In this way it is possible to take into account the large amplitude differences between each PC. Moreover, since the first PC is characterized by larger score maximum excursion than the second one and so on, this product is also related to the explained variance of each PC. The maximum between the absolute values of this product set allows to identify the dominant PC for each HVSR curve: this information can be easily showed in a single map. Obviously, if the e_{km} value that contributes to identify the dominant PC has negative sign, the "apparent" trend associated with the respective HVSR will be the opposite of the computed one. Thus, despite these opposite trends are defined by the same principal component, they will be treated separately in the subsequent analyses.

Once that the measure with the same dominant PC were grouped together, it was possible to note that this first procedure part was not able to discriminate the peak amplitude differences between the experimental HVSR curves but only the pattern similarity. It follows that curves characterized by high peaks and by negligible peaks (i.e., peaks with amplitude below 2) could be insert together in the same group. The discrimination of these features is paramount in order to identify the large scale geological heterogeneities of the study area: in fact, high H/V peaks are usually associated with soft sediment zones (where high seismic amplification phenomena are expected) and curves with negligible peaks (which can be considered as flat curves) are usually associated with rock sites or zones with consolidate sediments.

In this analysis, the amplitude differences are expressed by the loading values: the higher the e_{km} value, the higher is the peak amplitude of the associated experimental H/V curve. In order to identify the flat curves detecting an external group to those ones already identified, it was needed

to consider again the maximum value of the product between each e_{km} and the amplitude maximum excursion of the "apparent" HVSR curve defined by the PCs. In particular, the flat curves are characterized by a maximum value of this product lower than a certain threshold, that in all the analyzed cases it is lower than 0.6. The measure station with the maximum value of this product higher than 0.6 present HVSR curves with at least one peak with amplitude equal or higher than 2.

After the grouping of the flat curves, the last step of the procedure was to detect the peak amplitude variations within the groups identified by the dominant PC. In particular, the e_{km} values allowed to identify subgroups characterized by both similar peak amplitude values and similar pattern. The e_{km} value thresholds are defined on the basis of the numerical distribution of these values.

7.4.3 PCA application on the Emilia Romagna dataset

The first PCA application on a HVSR dataset is done in the Chapter 6 study area. The aim of this first attempt is to verify the capability of this procedure in identifying similar HVSR Zones to those ones detected by visual inspection and showed in Fig. 6.4 (see Section 6.3.1). The acquisition and analysis characteristics of the 204 HVSR curves taken into account are described in Section 6.3.1; the geological settings of the area are described in Paragraph 6.2.

After the logarithmic interpolation of the experimental curves, PCA was applied in the range 0.1-10 Hz using the procedure described previously. Thus, the first step was to determine the dominant principal component for each experimental curve. Fig. 7.4 shows the "apparent" HVSR curves defined by the dominant principal components identified by the algorithm, that is the first PC, the second one, the third one and the fourth one. The associated curves are named as PC 1, PC 2, PC 3 and PC 4 (blue curves in Fig. 7.4). PC -2 curve (red curve in Fig. 7.4) is defined by the opposite trend of PC 2 and it is associated to the experimental HVSR curves where the dominant principal component is the second one and the corresponding e_{km} value has negative sign. It follows that, for the other dominant principal component, all the measures are characterized by positive e_{km}.

The graph in Fig. 7.5 shows that the first four principal components are able to explain together almost the 95% of the original dataset variance. It is immediately evident that the first principal component explain almost the 80% of the overall variance and the apparent curve associated to this component (characterized by two peaks at 0.9 and 0.25 Hz) is very similar in terms of peak and trough position to the average pattern defined by the HVSR curves located within the Zone 1 (Fig. 6.12), which includes almost all the investigated area (Fig. 6.4). Despite PC 2 and PC -2 define two different "apparent" H/V curve, they explain the same percentage of the original data variance, i.e. almost 8%. It is worth nothing that PC 2 resembles the Zone 2 average HVSR pattern in the range 1-10 Hz: in fact, the "apparent" curve is characterized by a peak at almost 1 Hz and a

slight "hump" between 1.5 and 2 Hz, but it has not the maximum at 0.25-0.3 Hz. The latter maximum is clearly evident in PC -2 "apparent" curve, which shows no peak in the range 1-10 Hz. As concerns PC 3 and PC 4, their curve trends are quite similar to that one of PC 1, with the higher frequency peak shifted to 1.5 Hz and an additional peak located at 8-9 Hz in PC 4 curve.

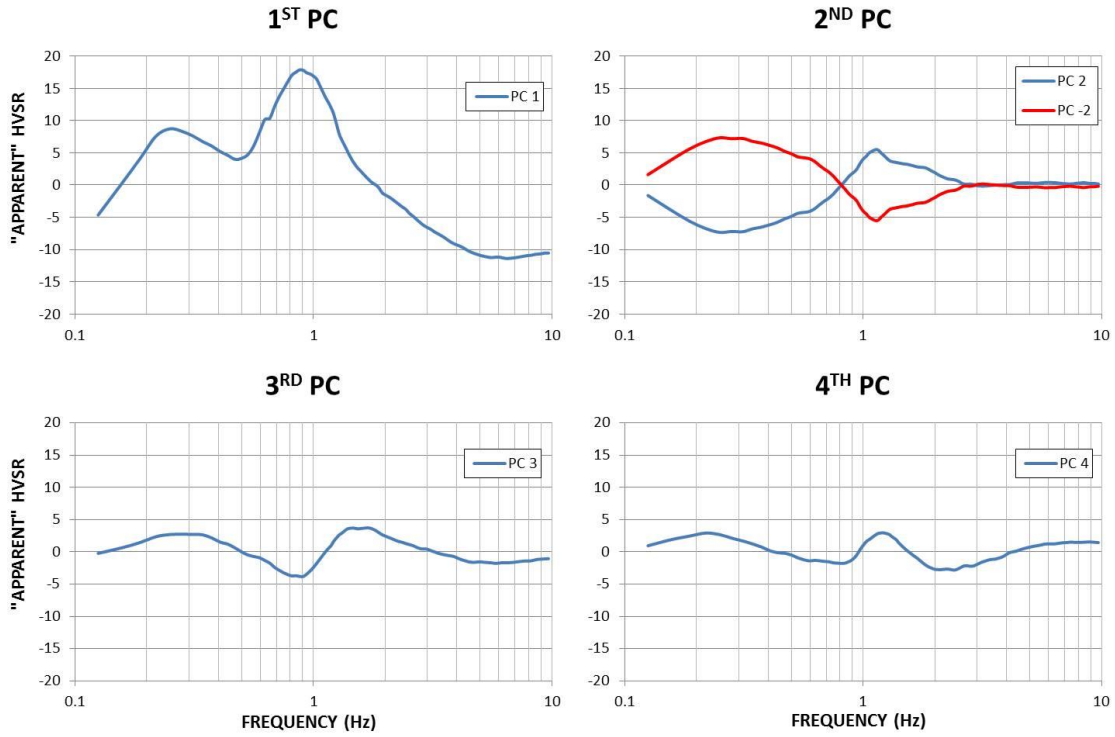

Fig. 7.4 - Blue curves represent the "apparent" HVSR curves defined by the found dominant principal components. Red curve is defined by the opposite trend of the corresponding blue curve and represents the "apparent" HVSR associated to the experimental HVSR curves where the e_{km} value that contributes to identify the dominant PC has negative sign.

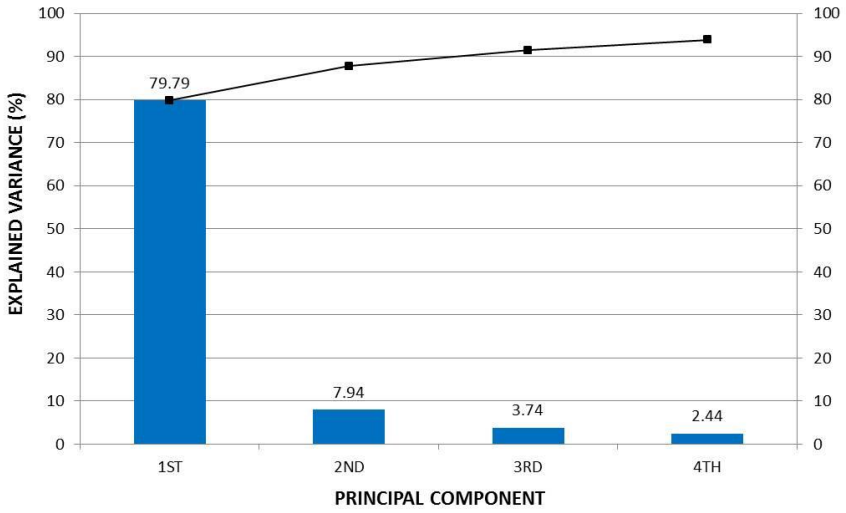

Fig. 7.5 - Original dataset variance explained by each dominant component. Black line represents their cumulative variance.

The following step of the procedure was to group together the experimental H/V curves characterized by the same dominant "apparent" HVSR pattern: this allowed to identify 5 main classes (Fig. 7.6-7.7), where the PC 1 represents the largest one (193 curves of 204). It was observed that the product value which detects the dominant principal components is always higher than 0.6: this is in keeping with the absence of flat curves in the whole study area.
In order to point out the amplitude differences, the main classes were then split up into subclasses (Fig. 7.6-7.7) determined by the values of the loadings. Considering the main class represented by PC 1 (the largest one), it was considered appropriate to detect four subclasses (A, B, C and D); the corresponding chosen e_{km} thresholds are showed in Fig. 7.8. In order to make a homogeneous sub-classification, the loading thresholds are retained for all the remaining main classes. Since PC 2, PC -2, PC 3 and PC 4 main classes contain far fewer measurements, one or more subclasses within them may be missing.
An important outcome of the procedure is the evident similarity between the experimental curves grouped in the subclasses (and therefore in the main classes) and the corresponding "apparent" HVSR defined by the dominant principal component. This result demonstrates the good capability of PCA in grouping similar HVSR patterns.
The map in Fig. 7.8 shows the spatial distribution of the main classes and their corresponding subclasses, indicated by the color and the size of the dots respectively. It is clearly evident that the PC 1 class covers almost the whole area including the Zone 2 and part of the Zone 3 identified in the previous study and showed in Fig. 6.4: it follows that almost all the measures present a high similarity with the pattern defined by PC 1 HVSR "apparent" curve. The PC 2 class is represented by four points located in the center of the Mirandola town (part of Zone 3) and the remaining classes are characterized by very few points randomly located in the investigated area.
In view of this, it is possible to state that the groupings achieved by PCA are not able to discriminate the slight differences pointed out by the previous zonation. This could be due to the fact that the Chapter 6 zonation, based on a visual inspection, took into account also a geographical criterion: in fact, the experimental curve belonging to PC 2 class were inserted previously in the Zone 3 rather than within Zone 2 despite being characterized by the slight "hump" between 1.5-2 Hz. On the other hand, this feature it was not discriminating to identify the Zone 2 in the PCA analysis: this could be related to the presence of the low frequency peak at 0.25 Hz, which may have been critical to insert the curves of this Zone within the PC 1 class.
It is worth noting that some information can be deduced by the distribution of PC1 subclasses: in fact, the subclass C and D, characterized by higher amplitude values, are mainly located in Mirandola-Medolla area (part of Zone 3) and in Concordia sulla Secchia-San Possidonio area (part of Zone 1). In the first case, the high amplitude is related with the sharp peak at about 1 Hz, while in the second case with the peaks at about 0.9 Hz and 0.25 Hz. On the basis of the available geological information, it is difficult to relate the Concordia sulla Secchia-San Possidonio area (not identified by the first visual zonation) with subsoil geological features different by those ones of the Zone 1: it follows that, the high amplitude values of these HVSR peaks are probably affected by

the ambient vibration wavefield conditions present during the acquisition (see Paragraph 6.4). Moreover, this phenomenon, that most probably affects both HVSR maxima in the rest of the area, causes spatially random amplitude variations that make difficult the identification by the sub-classification of further clear amplitude groups in relation to the surface and subsurface geological information.

In view of the obtained results, it is possible to state that the relative homogeneity of the HVSR experimental curves collected in the Emilia Romagna area make difficult a clear identification of large scale heterogeneities by PCA. For this reason, it was considered appropriate to test PCA in more complex geological setting areas, in order to have a HVSR original dataset consisting of curves more different each other.

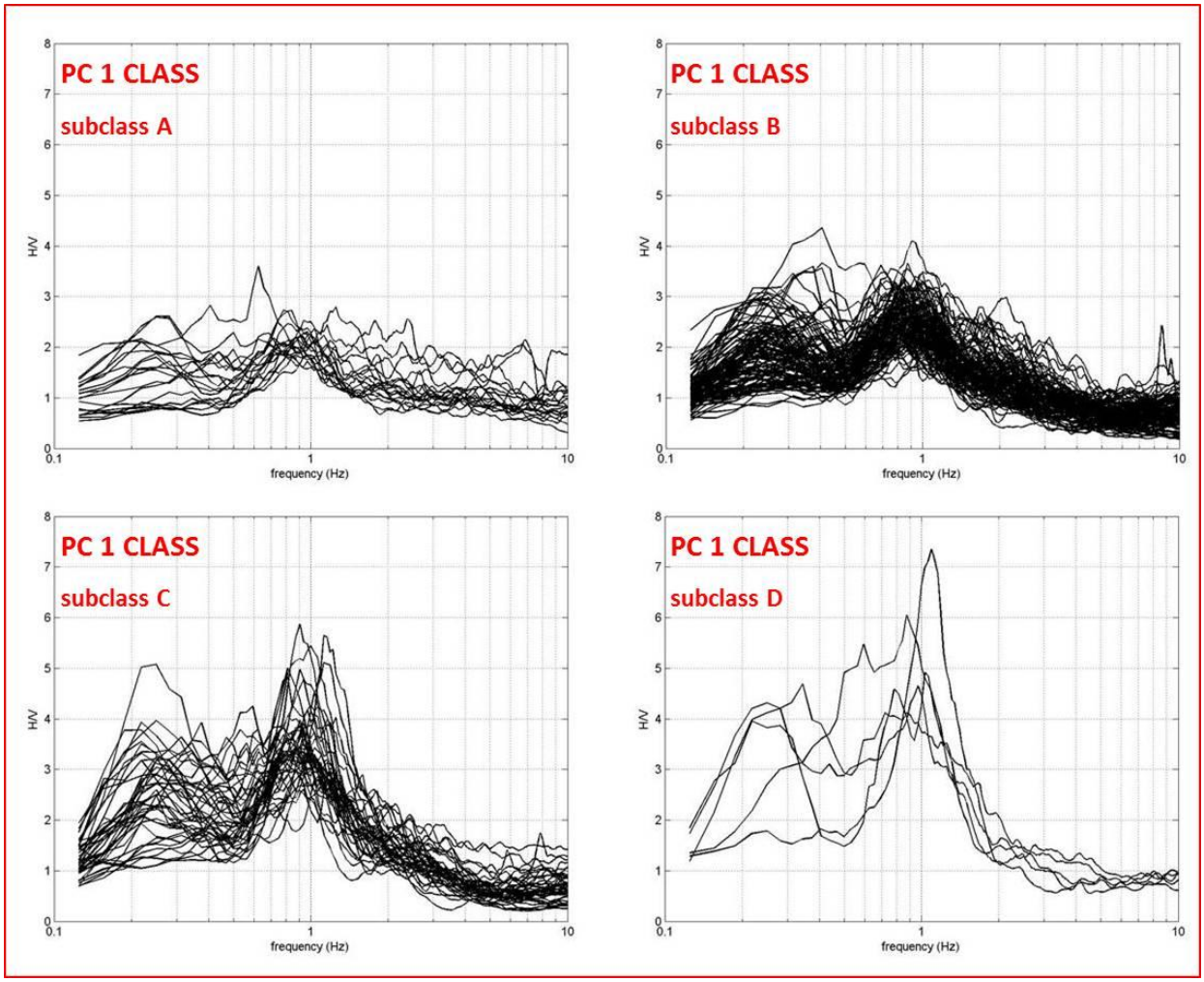

Fig. 7.6 - HVSR experimental curves belonging to PC 1 main class. The four panels show the curves belonging to each subclass (A, B, C, D) obtained by the values of the loadings. The used e_{km} thresholds are showed in 7.8. The border color is related with the dot color of PC 1 class in Fig. 7.8.

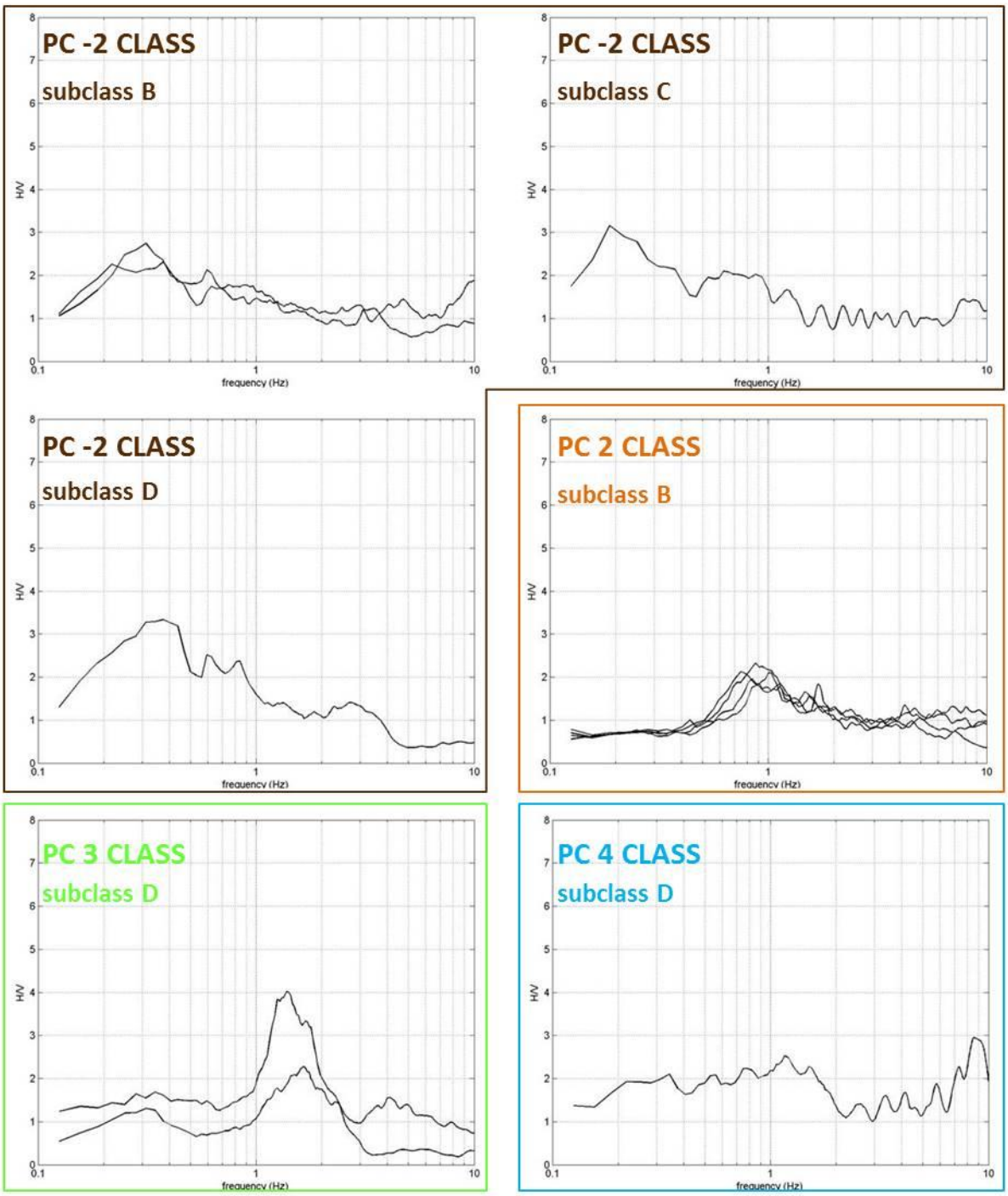

Fig. 7.7 - HVSR experimental curves belonging to PC -2 (brown border), PC 2 (orange border), PC 3 (green border) and PC 4 (light blue border) main classes. The panels associated of each main class show the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in 7.8. The border colors are related with the dot colors of the main classes in Fig. 7.8.

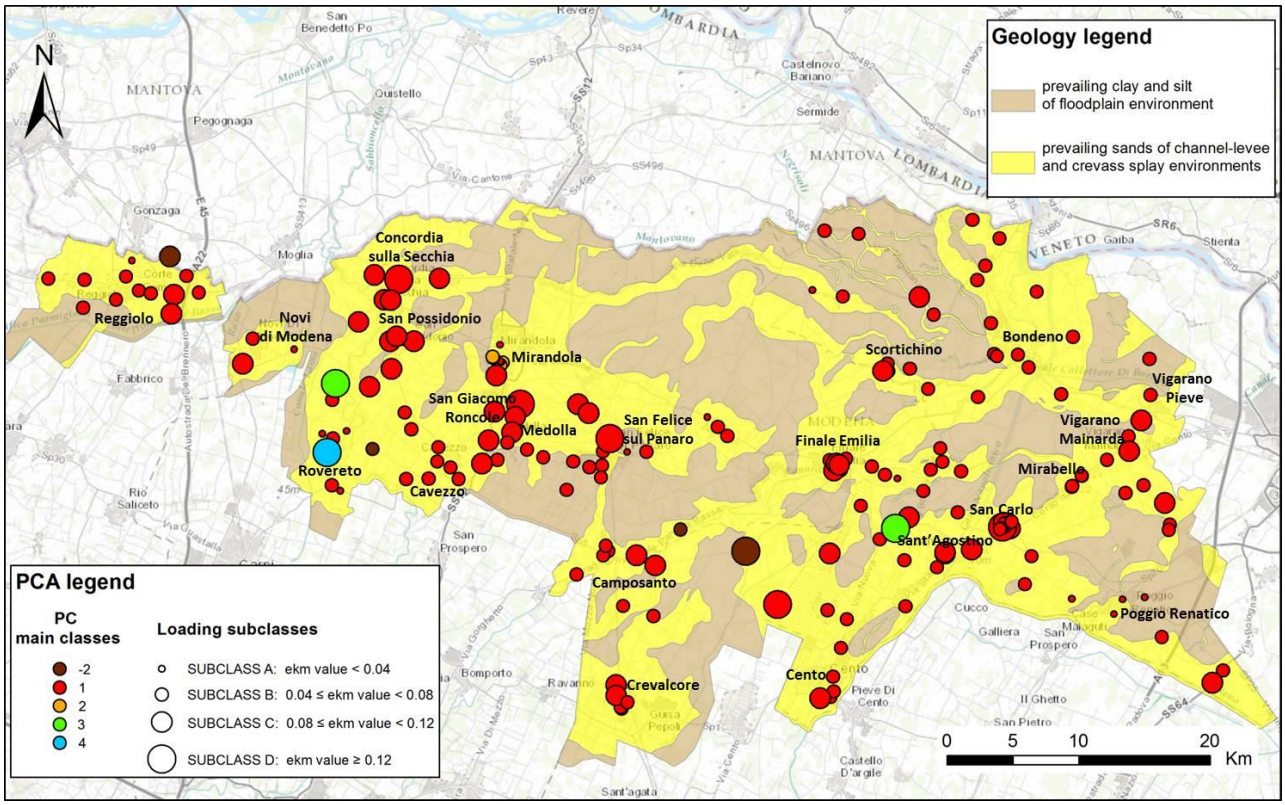

Fig. 7.8 - The map shows the spatial distribution of the main classes and their corresponding subclasses, indicated by the color and the size of the dots respectively. Moreover, the PCA legend on the bottom left shows the e_{km} thresholds used to perform the sub-classification.

7.4.4 PCA application on the Collesalvetti municipality dataset

7.4.4.1 Geological settings

Collesalvetti is a town located in Central Italy in Tuscany region about 15 Km from Livorno and Pisa cities (Fig. 7.9). The area of its municipality extends for about 110 Km^2 and includes three main morphologic sectors:

- the south-western sector, characterized by the higher reliefs of the whole area (about 400 m a.s.l) and slopes with high steepness;
- the hilly sector, which extends from south to north up to the Arno river plain and constitutes most of the municipality area. The elevation ranges between 100 and 50 m a.s.l.;
- the Arno River alluvial plain, which characterizes the northern part of the municipality area and it extends about 15 km from east to west. This sector is characterized by a flat morphology and elevations included between 5 m and 1 m a.s.l.

From the geological point of view, the whole area is included in the Northern Apennines domain. In the first sector, the formations outcropping belong to the Ligurian Complex (LC, Fig. 7.9), which includes rocks deposited within the Ligurian Ocean starting from Jurassic until Cretaceous age. After the deposition, these materials were intensely deformed and fractured by the subsequent compressive tectonic phases (Elter *et al.*, 1964; Elter and Marroni, 1991). This complex is characterized by magmatic and sedimentary rocks. Basically, the first ones are peridotites, gabbros and ophiolite breccias; the second ones are argillites, calcareous rocks and marbles of turbiditic origin.

The hilly sector is divided in four different geological zones. The first one, where the villages of Pietreto, Parrana San Martino, Parrana San Giusto and Le Case are aligned (Fig. 7.9), is located close the rocky ridge of the south-western sector and is characterized by the outcrop of Miocene formations (MI; Fig. 7.9). In particular, these sediments are fairly heterogeneous and consist of conglomerates, sand and marls with gypsum intercalations mainly deposited in a continental environment. The Miocene deposits overlay with angular unconformity relationship the LC with a maximum thickness of about 100 m. The second zone, where the village of Castell'Anselmo and Nugola Nuova are located (Fig. 7.9) is characterized by the outcrop of Pliocene sediments (PLI; Fig. 7.9). These materials consist of consolidated clay and sands deposited during the marine transgression that concerned the area during Pliocene age (Lazzarotto *et al.*, 1990a) and overlay with angular unconformity relationship the MI with a maximum thickness of about several hundred meters. The third geological zone of the hilly sector, where the Collesalvetti town, Vicarello and Nugola Vecchia villages are located, is characterized by Quaternary Pleistocene sediments (PLE; Fig. 7.9). These materials were deposited during the transgressive-regressive phases that affected the Pleistocene age (Lazzarotto *et al.*, 1990b) and consist of alluvial and transitional sands and gravels. As the underlying deposits, Pleistocene sediments lie on the PLI with angular unconformity relationship with a thickness decreasing from south (about 100 m) to north (about tens of meters). The fourth zone is represented by the valley bottoms that drain northwards on the Arno River plain; these valleys are filled by alluvial Holocene sands, silts and clays (HOL; Fig. 7.9) and colluvial deposits (CL; Fig. 7.9): the overall thickness of these materials ranges from tens to few meters.

Finally, The Arno River alluvial plain sector includes the village of Mortaiolo, Guasticce and Stagno. As the smaller valley bottom, this alluvial plain is filled by the Holocene deposits (HOL; Fig. 7.9) with the preponderance of silty and clayey materials. These sediments present an overall thickness of about 20-30 m.

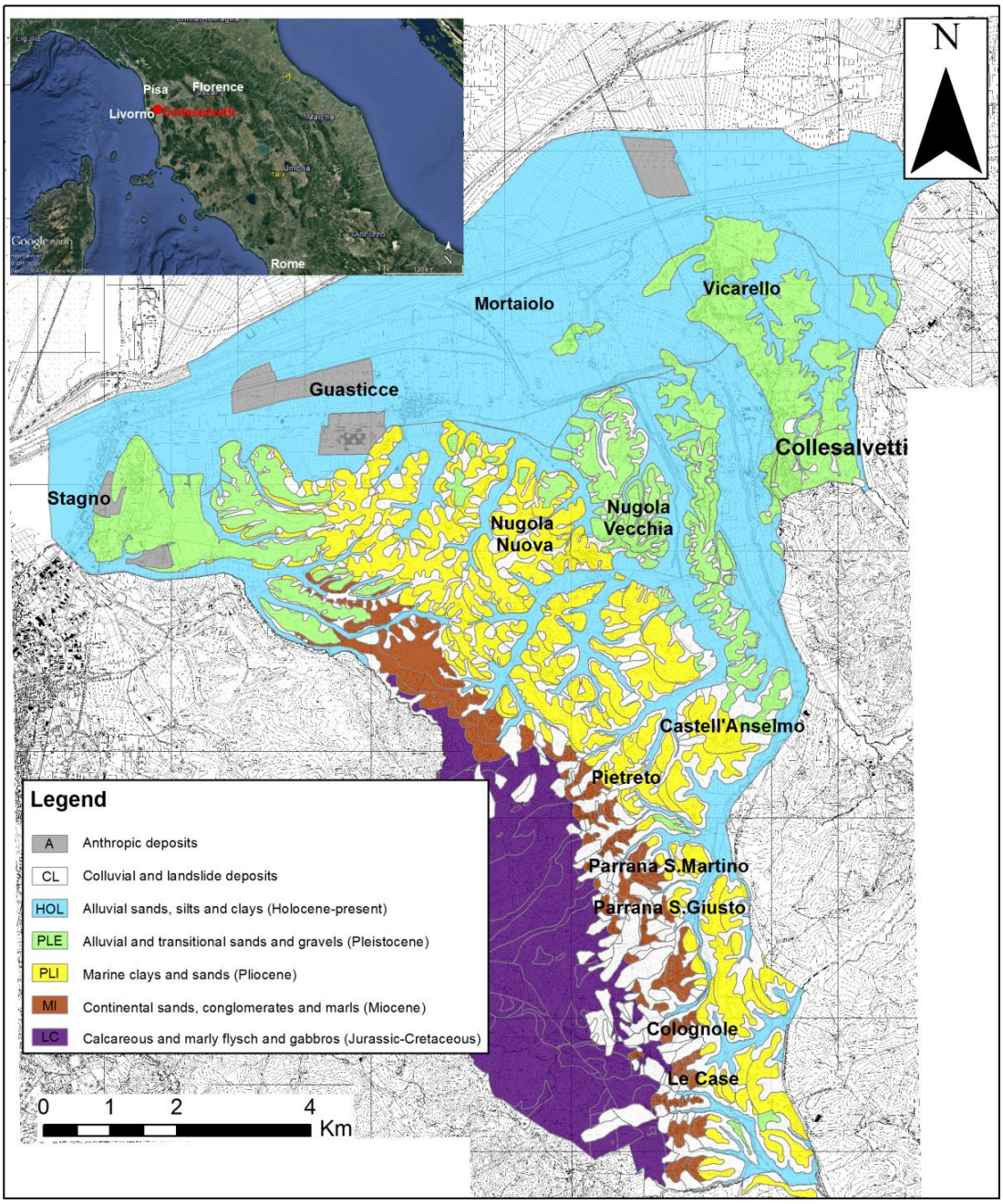

Fig. 7.9 - Geological map of the Collesalvetti municipality area.

7.4.4.2 Result description

The HVSR dataset collected in Collesalvetti Municipality consists of 100 measurements performed by a three-directional digital tromograph Tromino Micromed with a sampling frequency of 128 Hz and an acquisition time of 20 minutes. At each site, the HVSR curve was obtained dividing the time

series into 60 non-overlapping windows with 20 s length and following the procedure described in Paragraph 2.3.

PCA was applied in the range 0.1-10 Hz using the procedure previously described. Fig. 7.10 shows the "apparent" HVSR curves defined by the dominant principal components identified by the algorithm. In this case the dominant principal components are the first one, the second one, the third one, the fourth one, the sixth one and the eighth one. The related curves are represented in blue, while the red curves represent their opposite trends and are associated to the experimental HVSR curves where the e_{km} value defining the dominant principal component has negative sign. It is possible to note that the first principal component is only represented by the "apparent" opposite HVSR curve (PC -1 in Fig. 7.10).

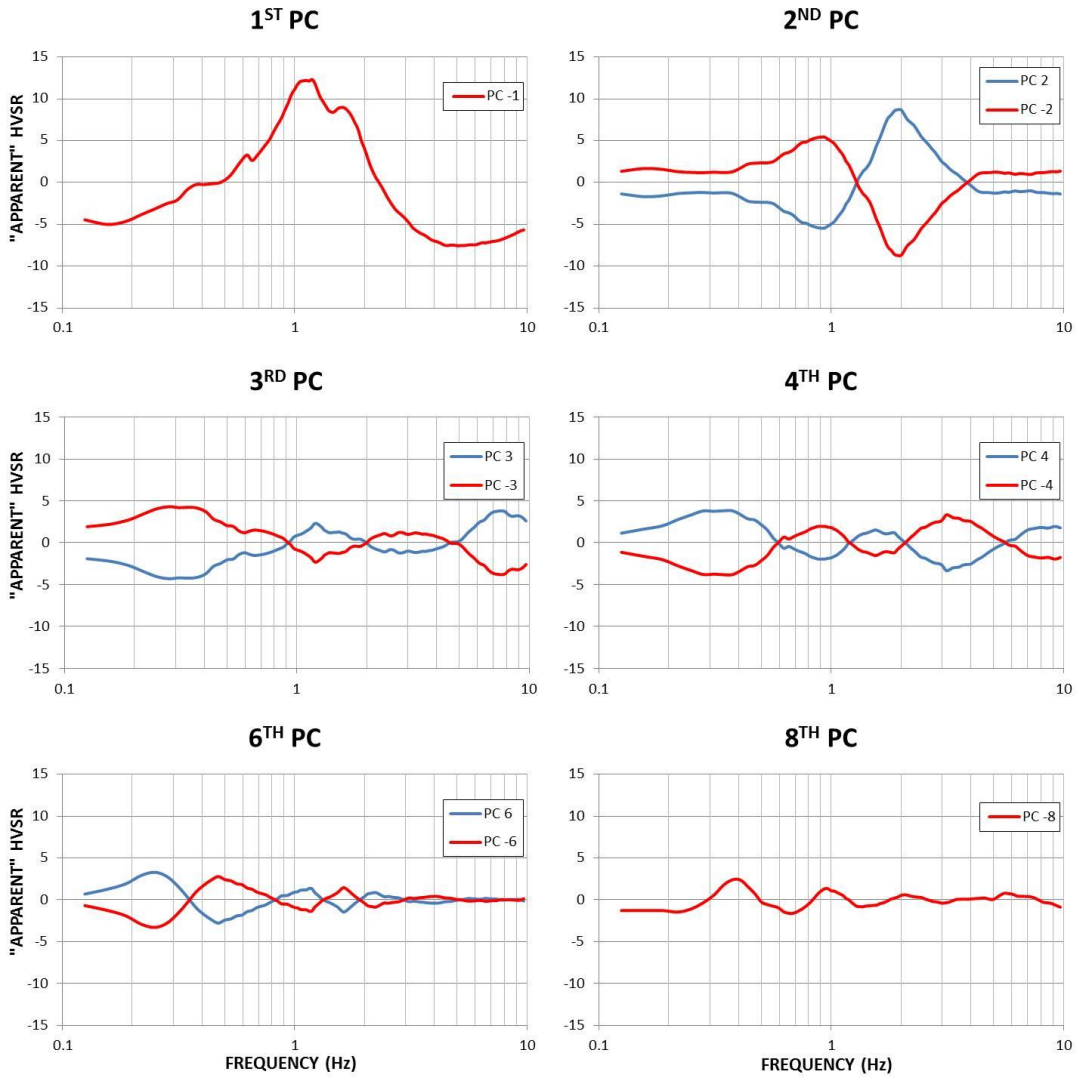

Fig. 7.10 - Blue curves represent the "apparent" HVSR curves defined by the found dominant principal components. Red curves represent the "apparent" HVSR associated to the experimental HVSR curves where the e_{km} value that contributes to identify the dominant PC has negative sign.

The graph in Fig. 7.11 shows that the first eight principal components are able to explain together almost the 95% of the original dataset variance. Observing this graph, it is possible to state that the original dataset variance is more equally distributed on the dominant principal components than the previous case. In particular, the first principal component explains almost the 60% of the overall variance and is represented an "apparent" HVSR curve characterized by a clear main peak at about 1 Hz. The second principal component, which explains almost the 20% of the overall variance, is represented by two "apparent" curves, PC 2 and PC -2 (Fig. 7.10), characterized by a clear peak at about 1 Hz and 2 Hz respectively. As concerns the remaining dominant component, they are represented by "apparent" HVSR with at least two clear peaks, some of which located in the range 0.25-0.5 Hz (Fig. 7.10).

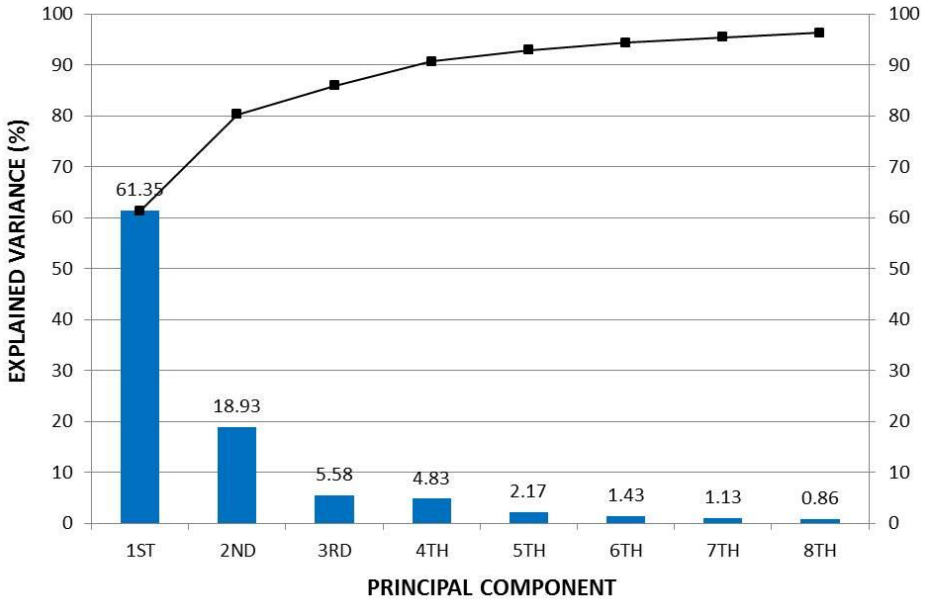

Fig. 7.11 - Original dataset variance explained by the first eight principal components. Black line represents their cumulative variance.

The grouping of the experimental H/V curves characterized by the same dominant "apparent" HVSR pattern allows to identify 10 main classes (Fig. 7.12-7.13-7.14-7.15): the PC -1 represents the largest one (57 curves of 100). As previously described, the experimental measure where the product value which detects the dominant principal component (i.e., the product between the associated e_{km} and the amplitude maximum excursion of the "apparent" HVSR curve defined by the PC) is lower than 0.6 allows to identify the flat curves. This further group is here formed by 5 experimental HVSR (Fig. 7.15).

As concerns the subsequent amplitude sub-classification, on the basis of the e_{km} values of the experimental curves belonging to the PC -1 main class (the largest one) it was considered appropriate to establish three subclasses (A, B and C; Fig. 7.12). The corresponding chosen e_{km} thresholds are showed in Fig. 7.16 and are retained for all the remaining main classes in order to

make a homogeneous sub-classification. As explained above, this may cause the absence of one or more subclasses within them (Fig. 7.13-7.14-7.15).

Also in this PCA application, the evident similarity between the grouped experimental curves and the corresponding "apparent" HVSR defined by the dominant principal component witnesses the good capability of the procedure in grouping similar HVSRs and in identifying their representative patterns. It is possible to note that a frequency difference of 1 Hz exists between the peaks of some HVSRs in PC -1 subclass B and C and the peak of the corresponding "apparent" curve: this means that the overall pattern similarity played an important role in grouping these curves within PC -1 class despite this frequency shift.

The map in Fig. 7.16 shows the spatial distribution of the main classes and their corresponding subclasses, indicated by the color and the size of the dots respectively. Observing the positioning of the PC -1 subclasses, it is possible to note that the subclass A measurements (characterized by low amplitude peaks) are mainly located on the Pleistocene sediments, while the subclass B and C measurements (characterized by high amplitude peaks) are located on Holocene alluvial sediments, where higher resonance effects are expected. A similar characteristic concerns the measurements of the PC 2 and PC -2 subclasses: their subclasses B and C are mainly situated on the northern alluvial plain, while their subclasses A are located on the Miocene sediments in the south-western part of the area. These materials are also identified by the position of the HVSRs belonging to PC 3 (characterized by two peaks; Fig. 7.13) and PC -4 classes. The presence of curves with amplitude and frequency differences on MI, denotes the heterogeneity characterizing these deposits. An evident similarity affects the experimental curve patterns of the PC -3, PC 4, PC 6, PC -6 and PC -8 classes, where the HVSRs are characterized by a clear peak at 0.3-0.4 Hz (Fig. 7.13-7.14-7.15). Almost all these measurements are located on the Pliocene sediments.

Finally, it is possible to note that the 5 identified flat curves (named as "no peak" in Fig. 7.16) are situated on the Jurassic-Cretaceous rocks and on Pliocene deposits in the south part of the municipality: in these materials, no significant resonance effects are expected.

In view of the above described results, it is possible to state that in this geological context (more various than the previous one) PCA on HVSR data is a useful tool to obtain a quick first identification of the large scale heterogeneities of the study area.

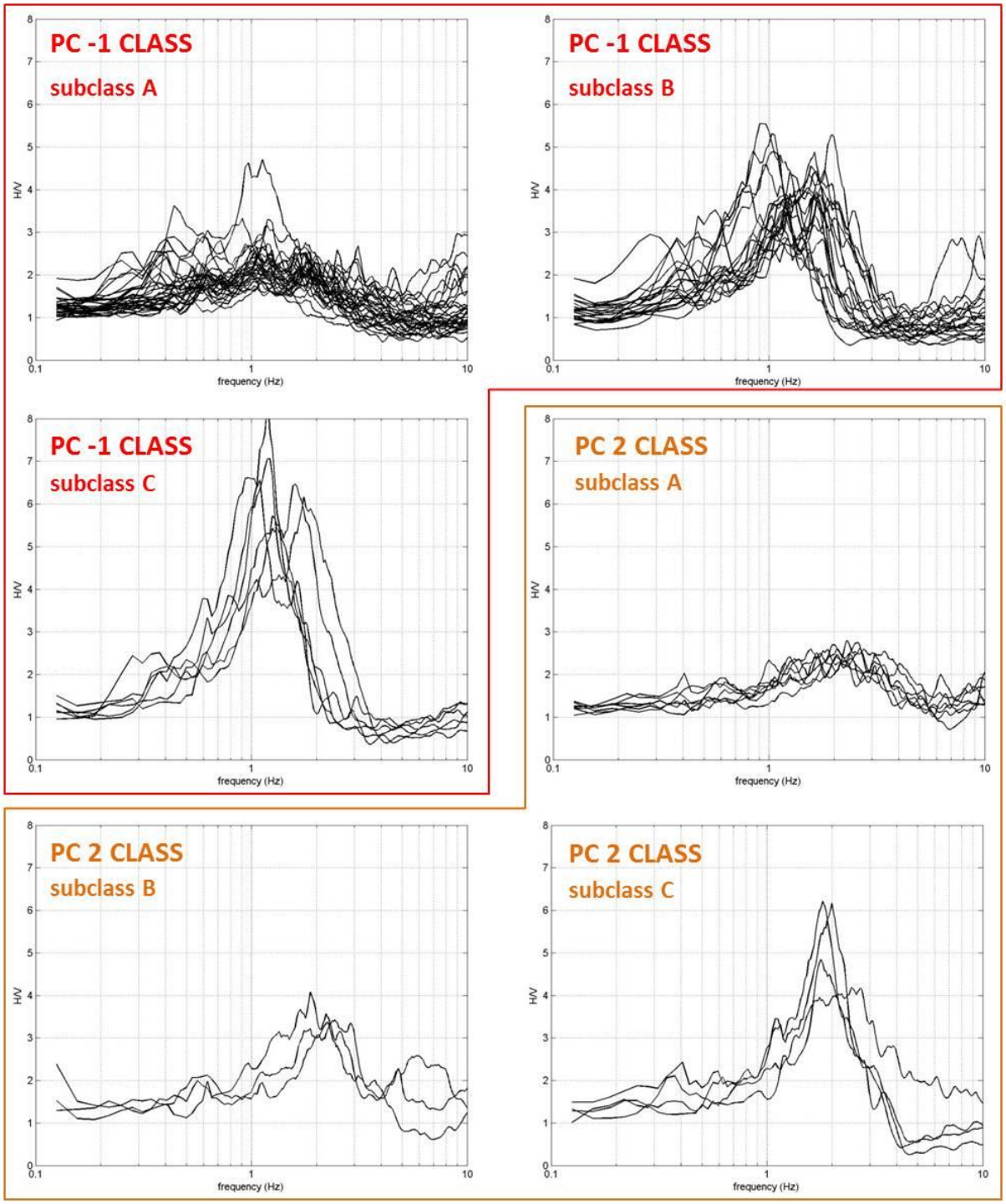

Fig. 7.12 - HVSR experimental curves belonging to PC -1 (red border) and PC 2 (orange border) main classes. The panels associated of each main class show the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in Fig. 7.16. The border colors are related with the dot colors of the main classes in Fig. 7.16.

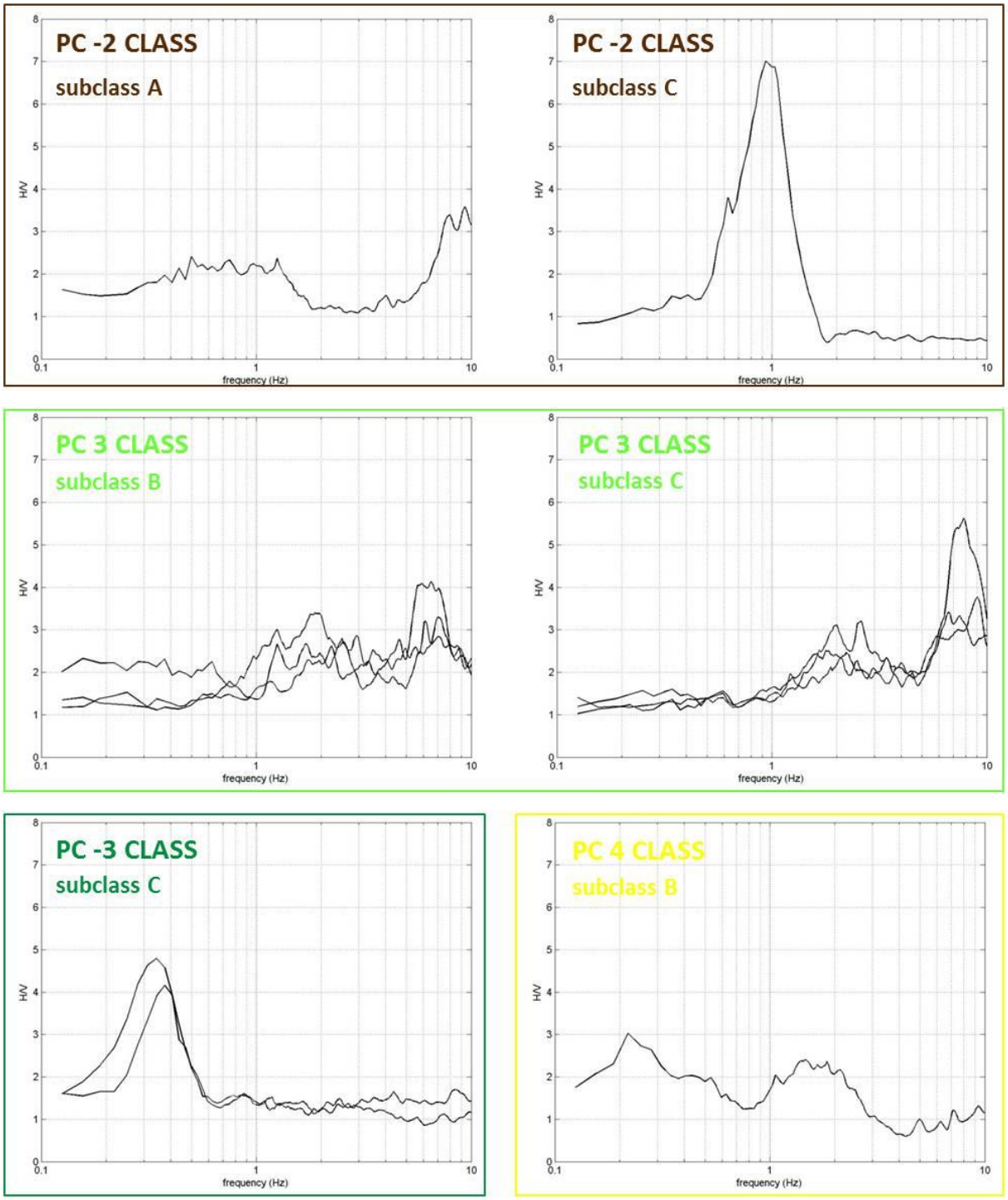

Fig. 7.13 - HVSR experimental curves belonging to PC -2 (brown border), PC 3 (light green border), PC -3 (dark green border) and PC 4 (yellow border) main classes. The panels associated of each main class show the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in Fig. 7.16. The border colors are related with the dot colors of the main classes in Fig. 7.16.

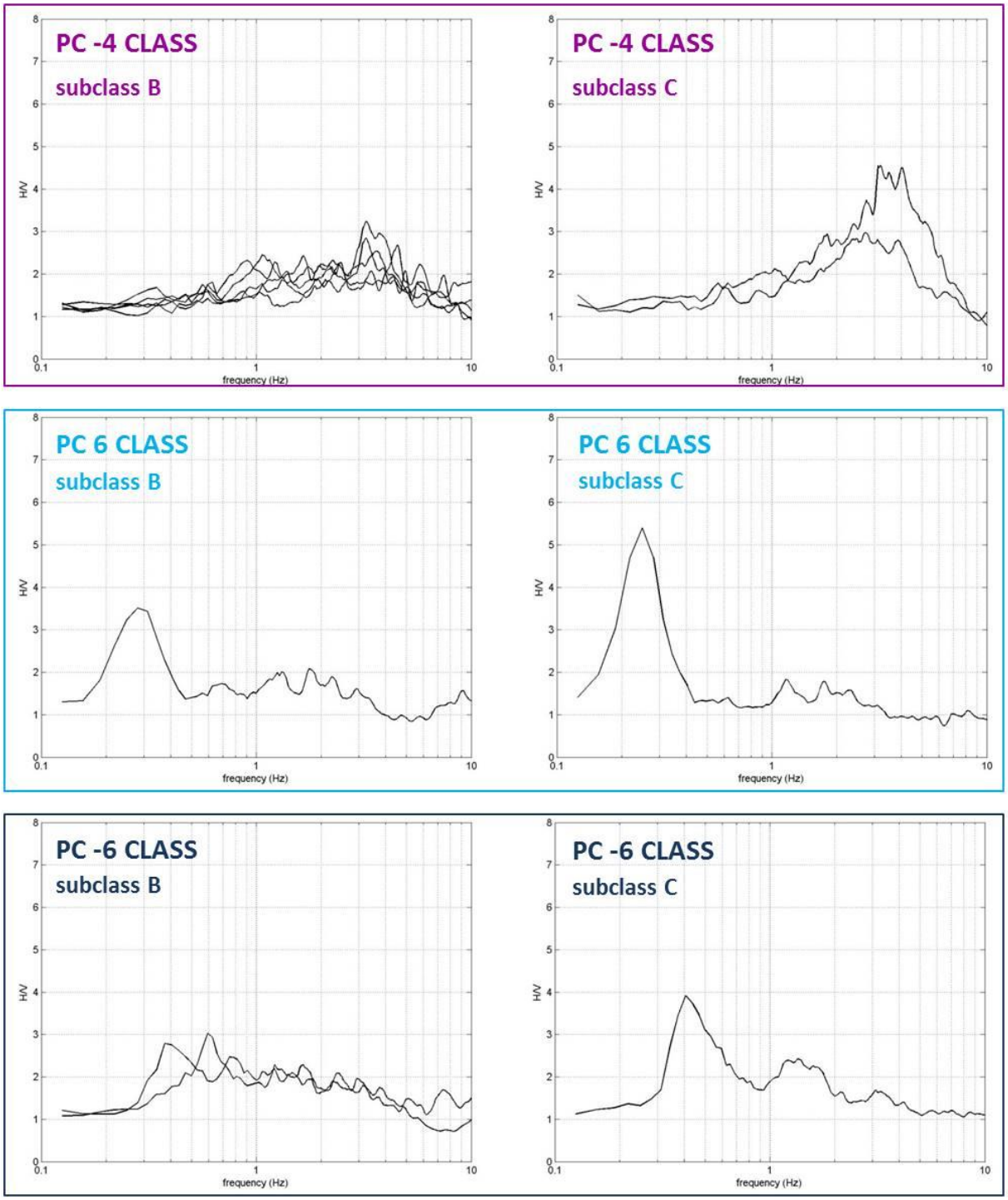

Fig. 7.14 - HVSR experimental curves belonging to PC -4 (violet border), PC 6 (light blue border) and PC -6 (dark blue border) main classes. The panels associated of each main class show the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in Fig. 7.16. The border colors are related with the dot colors of the main classes in Fig. 7.16.

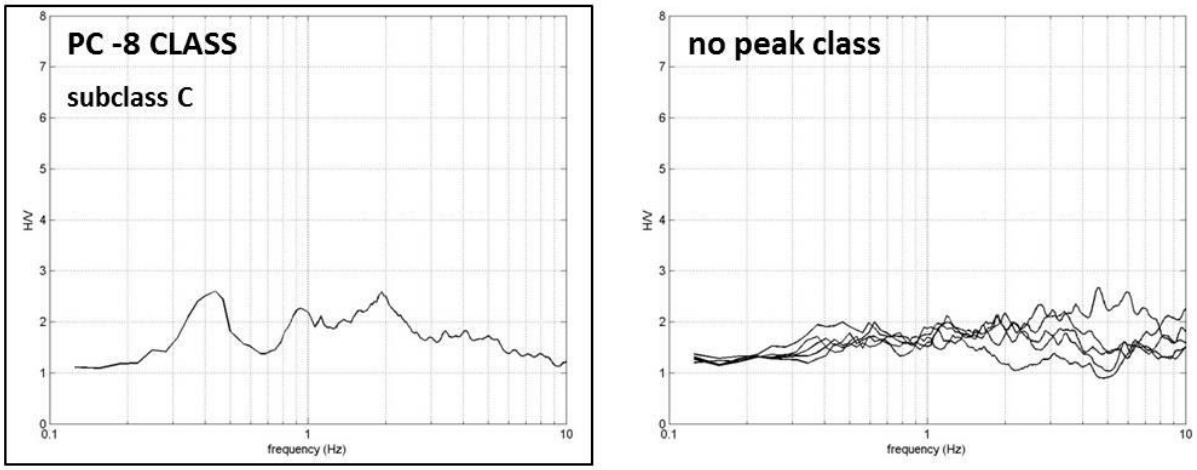

Fig. 7.15 - HVSR experimental curves belonging to PC -8 main class (black border) and the identified flat HVSR curves (no peak class). The panel associated to the PC -8 main class shows the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in Fig. 7.16. The border color of the PC -8 class is related with the dot colors of this class in Fig. 7.16.

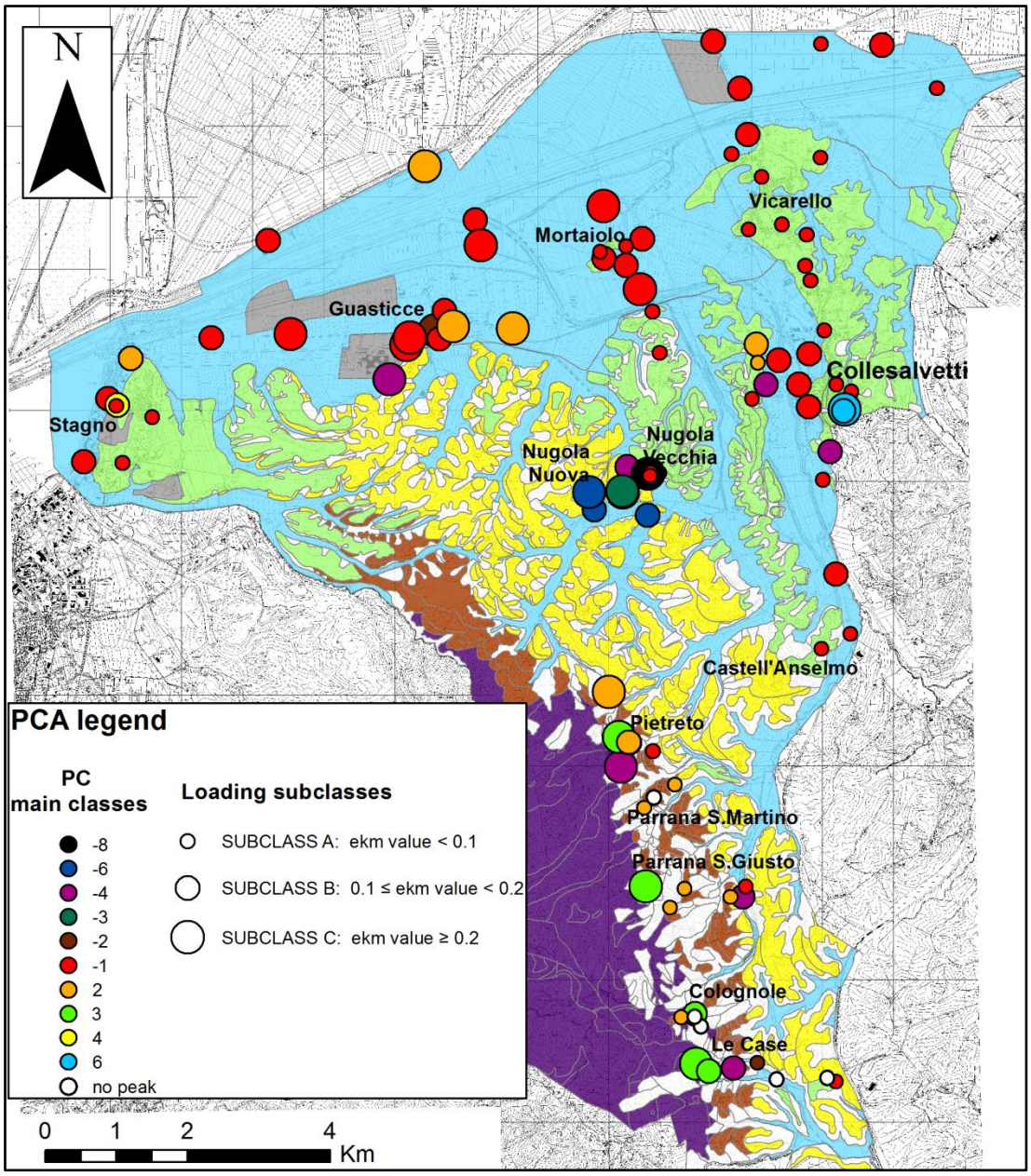

Fig. 7.16 - The map shows the spatial distribution of the main classes and their corresponding subclasses, indicated by the color and the size of the dots respectively. Moreover, the PCA legend on the bottom left shows the e_{km} thresholds used to perform the sub-classification. The geology legend is showed in Fig. 7.9.

7.4.5 PCA application on the Montecatini Terme municipality dataset

7.4.5.1 Geological settings

Montecatini Terme is a town located in Central Italy in Tuscany region about 30 Km from Pisa and 40 Km from Florence (Fig. 7.17). The area of its municipality extends for about 18 Km^2 and includes two main morphologic sectors:

- the hilly sector, located in the northern part of the area and characterized by reliefs of the order of about 200 m a.s.l.;
- the alluvial plain sector, located in the southern part of the municipality area and including most of the Montecatini Terme town. This sector is characterized by a flat morphology and elevations of about 20 m a.s.l.

From the geological point of view, the whole area is included in the Northern Apennines domain. Exhaustive treatments of the geological characteristics of this area are given by Trevisan (1954), Brandi *et al*. (1967) and Puccinelli *et al*. (2000).

In the hilly sector, the rock formations outcropping belong to two different tectonic units: the Tuscan Unit and the Ligurian Unit (Fig. 7.17). The latter Unit lies on the first one due the compressive tectonics during the Apennine formation. The Tuscan Unit consists of different kinds of geological formations, in particular jaspers (JAS, Middle Jurassic-Upper Jurassic), calcareous rocks (MAI, limestones and calcarenites; Upper Jurassic-Lower Cretaceous), alternations of argillites and calcareous material (SCA, Lower Cretaceous-Oligocene) and sandstones (MAC, Upper Oligocene). Ligurian Unit rocks, outcropping in the eastern part of this sector, mainly consist of chaotic alternations of argillites and calcareous materials of turbiditic origins (LIG, Cretaceous-Eocene). Moreover, the hilly sector is characterized by Quaternary alluvial terraced deposits (QU, Fig. 7.17; mainly located in the valley around the Nievole village) and numerous landslide deposits and debris slopes (DE, Fig. 7.17).

Most of the plain sector is characterized by Holocene and actual alluvial deposits with the preponderance of silty materials (HOL, Fig. 7.17). The more ancient Quaternary sediments (QU, Fig. 7.17) are located on the relief edges and consist of terraced alluvial deposits with the preponderance of silty and clayey materials. Both alluvial formations lie on a Quaternary lacustrine succession (not outcropping in the study area) formed of irregular alternations of clays, sands and gravels.

An important characteristic of the plain sector is the presence of a significant thermal activity. In particular the thermal waters are confined within the JAS formation and the underlying calcareous materials (no outcropping in the north sector) situated beneath the Quaternary deposits. These waters rise up to the surface along fracture lines and give rise to the travertine deposit formation (TR, Fig. 7.17) in proximity of their emergences.

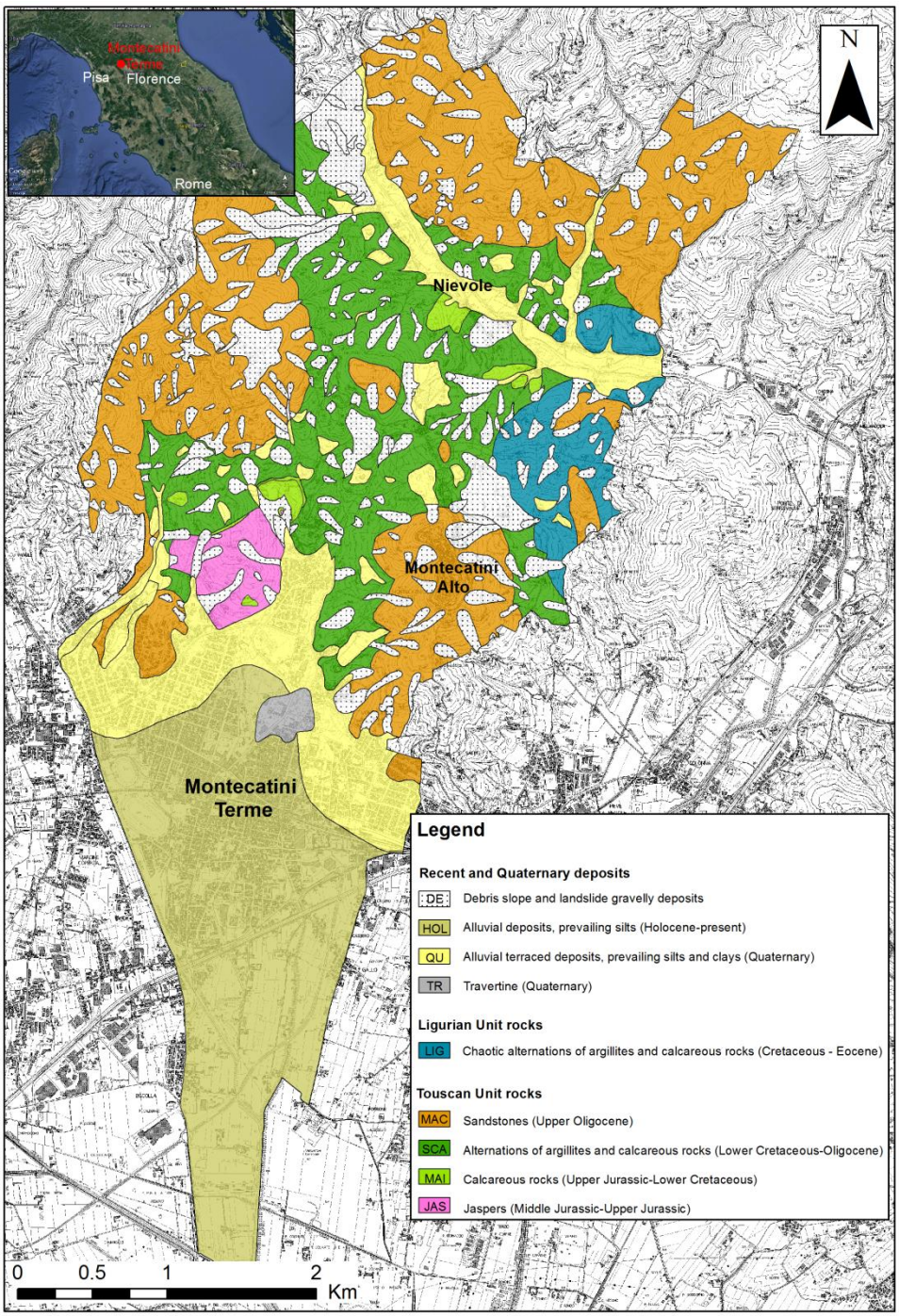

Fig. 7.17 - Geological map of the Montecatini Terme municipality area.

7.4.5.2 Result description

The HVSR dataset collected in Montecatini Terme Municipality consists of 85 measurements performed by three-directional digital tromographs Tromino Micromed and SR04HS Sara with a sampling frequency of 128 and 100 Hz respectively. The ambient vibrations were acquired for 20

minutes. At each site, the HVSR curve was obtained dividing the time series into 40 non-overlapping windows with 30 s length and following the procedure described in Paragraph 2.3.
PCA was applied in the range 0.1-10 Hz using the procedure previously described. In this case, the rescaling of the experimental curves with logarithmic interpolation allowed to overcome the differences due to the different sampling frequencies.
Fig. 7.18 shows the "apparent" HVSR curves defined by the dominant principal components identified by the algorithm. In this case the dominant principal components are the first one, the second one, the third one, the fourth one and the fifteenth one. The related curves are represented in blue, while the red curves represent their opposite trends and are associated to the experimental HVSR curves where the e_{km} value defining the dominant principal component has negative sign.

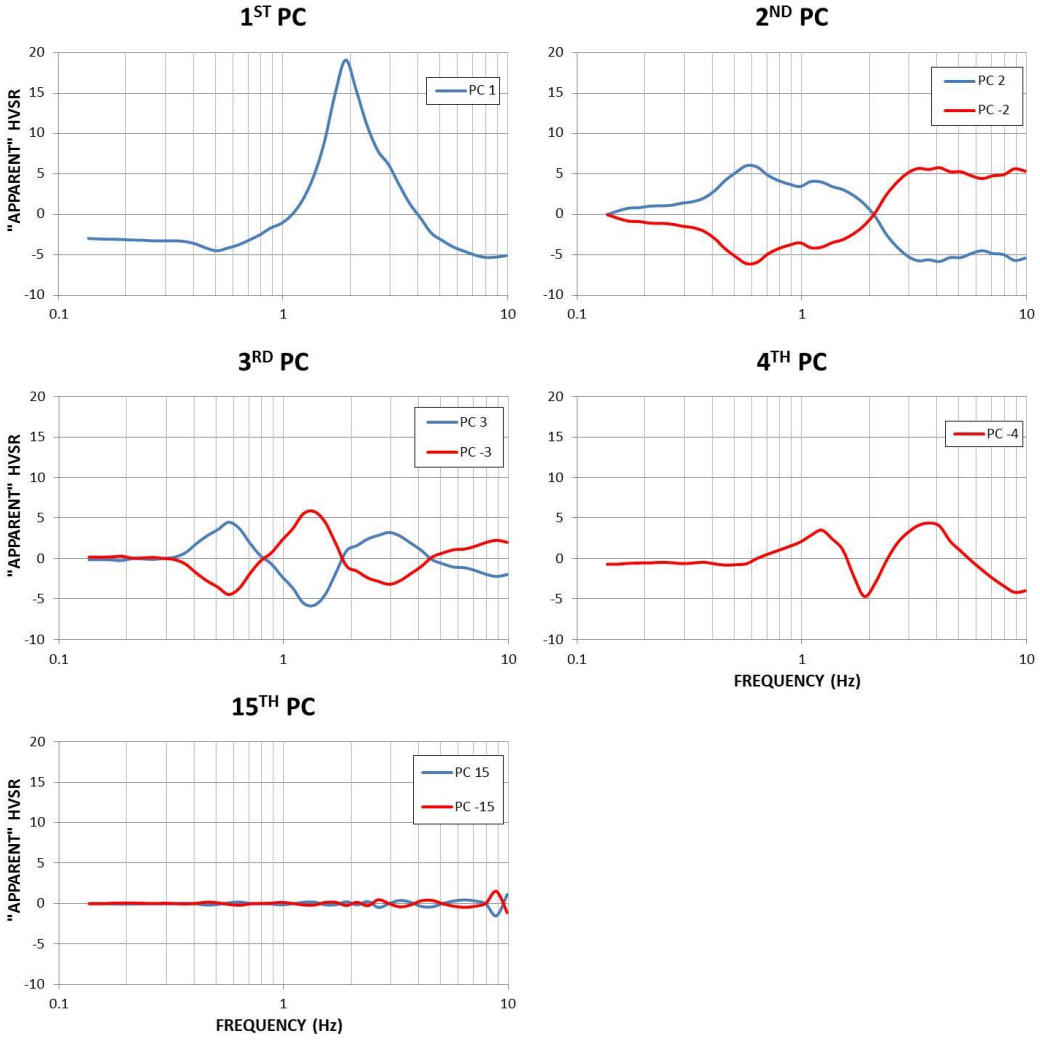

Fig. 7.18 - Blue curves represent the "apparent" HVSR curves defined by the found dominant principal components. Red curves represent the "apparent" HVSR associated to the experimental HVSR curves where the e_{km} value that contributes to identify the dominant PC has negative sign.

The graph in Fig. 7.19 shows that the first fifteen principal components are able to explain together almost the 100% of the original dataset variance. Similar to the previous case, it is possible to note that the original dataset variance is more equally distributed on the dominant principal components than the Emilia Romagna study case. In particular, the first principal component explains about the 50% of the overall variance and is represented an "apparent" HVSR curve characterized by a clear main peak at about 2 Hz. The second principal component, which explains about the 20% of the overall variance, is represented by two "apparent" curves, PC 2 and PC -2 (Fig. 7.18). The first one shows two peaks at 0.6 Hz and about 1 Hz merged together; the second one is characterized by a general rise in the frequency range 3-10 Hz. The third and the fourth principal components explain respectively about the 8% and 7% of the overall variance: it is possible to note that the fourth principal component is only represented by the "apparent" opposite HVSR curve (PC -4 in Fig. 7.18), characterized by two peaks (at about 1.5 Hz and 4 Hz) as well as the PC 3 apparent curve (with maxima at about 0.6 Hz and 3 Hz). Conversely, its opposite curve PC -3 shows a clear peak at about 1.5 Hz. As concerns the last found dominant PC, the fifteenth principal component explains about the 0.2% and presents two "apparent" H/V (PC 15 and PC-15 in Fig. 7.18) both characterized by high frequency maxima (about 9-10 Hz).

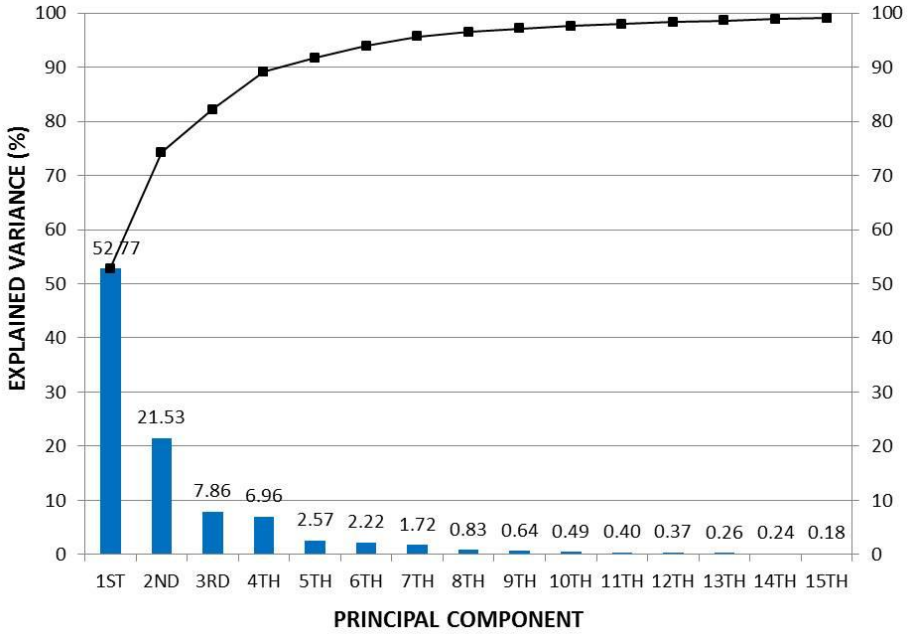

Fig. 7.19 - Original dataset variance explained by the first fifteen principal components. Black line represents their cumulative variance.

The grouping of the experimental H/V curves characterized by the same dominant "apparent" HVSR pattern allows to identify 8 main classes (Fig. 7.20-7.21-7.22-7.23): PC 1 represents the largest one (30 curves of 100). The flat curve group consists of 17 experimental HVSR (Fig. 7.23).

As concerns the subsequent amplitude sub-classification, on the basis of the e_{km} values of the experimental curves belonging to the PC 1 main class, it was considered appropriate to establish three subclasses (A, B and C; Fig. 7.20) again taking into account the e_{km} thresholds established for the Collesalvetti municipality study case. As usual these thresholds are retained for all the remaining main classes (Fig. 7.21-7.22-7.23).

As concerns the PC 1 main class, it is possible to note a good similarity between the grouped experimental curves and the corresponding "apparent" HVSR. As in the previous case, the frequency difference of about 1 Hz between the peaks of some HVSRs in subclass B and the peak of PC 1 "apparent" curve can be related to the significant role played by the overall pattern similarity in grouping these curves within PC 1 class despite this frequency shift. Regarding the remaining main classes, it is worth noting that some evident differences between the experimental and the "apparent" curves exist: in particular, within PC 2, PC 3 and PC -4 subclasses, the HVSRs never show both peaks. This fact could be related to the peak frequency shift that mainly characterizes the measures performed in the central and southern part of the plain sector: the apparent curves obtained by PCA for these dominant principal components could be understood as a sort of merging of the HVSR curves achieved in this zone.

The map in Fig. 7.24 shows the spatial distribution of the main classes and their corresponding subclasses, indicated by the color and the size of the dots respectively. It is clearly evident that the flat curves ("no peak" in Fig. 7.24) and the subclasses A of PC 1, PC 2 and PC -2(characterized by low amplitude peaks) are mainly located in the hilly sector in the northern part of the study area, where no significant resonance effects are expected. Moreover, in this sector also subclasses B and C of PC -2 and subclass C of PC 15 and PC -15 are situated: the high frequency peaks characterizing these HVSR curves are probably related with the debris slope and landslide shallow deposits present in this zone. The other important feature deducible by the map is the presence in the plain sector of PC 1, PC 2, PC 3, PC -3, PC -4 subclasses B and C. As expected, the whole sector is characterized by high resonance effects most probably related to the contact between the Holocene and Quaternary sediments and the underlying rocks belonging to the Tuscan Unit. It is possible to note that the PC 1 subclass B and C measures (which show HVSR maxima at about 2 Hz) are mainly concentrated in the northern-central part of the plain, while the other classes are distributed towards the lateral municipality borders and southwards. This distribution is related to the general deepening of the basin: in particular, starting from PC 1 class area, the frequency shifts westwards from 2 Hz to 1 Hz, where PC -3 subclasses B and C are located. Moreover, the peak frequency also gradually shifts southwards from 2 Hz to about 0.5 Hz in the zone of PC 3 and PC 2 classes. Conversely, a bedrock rise exists in the eastern part of the plain, where some measure of PC 1 subclass B and one measure of PC -4 subclass C are characterized by HVSR peaks at 3-4 Hz.

As mentioned above, PCA was not able to clearly identify all these frequency changes detecting all the groupings with the same peak frequency value. Despite this drawback that concerns only few measures, this case confirms the good capability of this procedure in obtaining a quick first

identification of the large scale heterogeneities of the study area and in identifying the main representative pattern of the HVSR dataset.

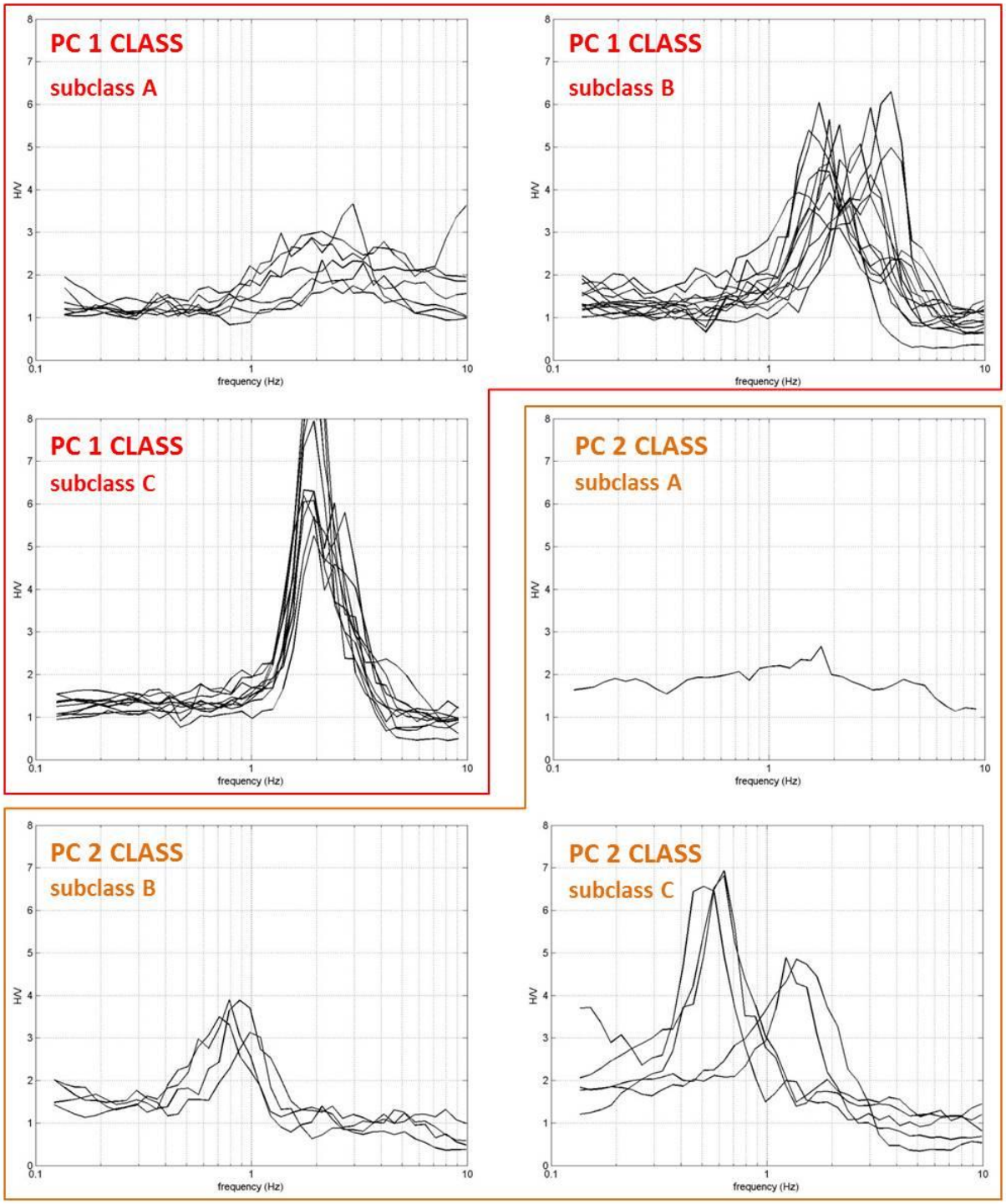

Fig. 7.20 - HVSR experimental curves belonging to PC 1 (red border) and PC 2 (orange border) main classes. The panels associated of each main class show the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in Fig. 7.24. The border colors are related with the dot colors of the main classes in Fig. 7.24.

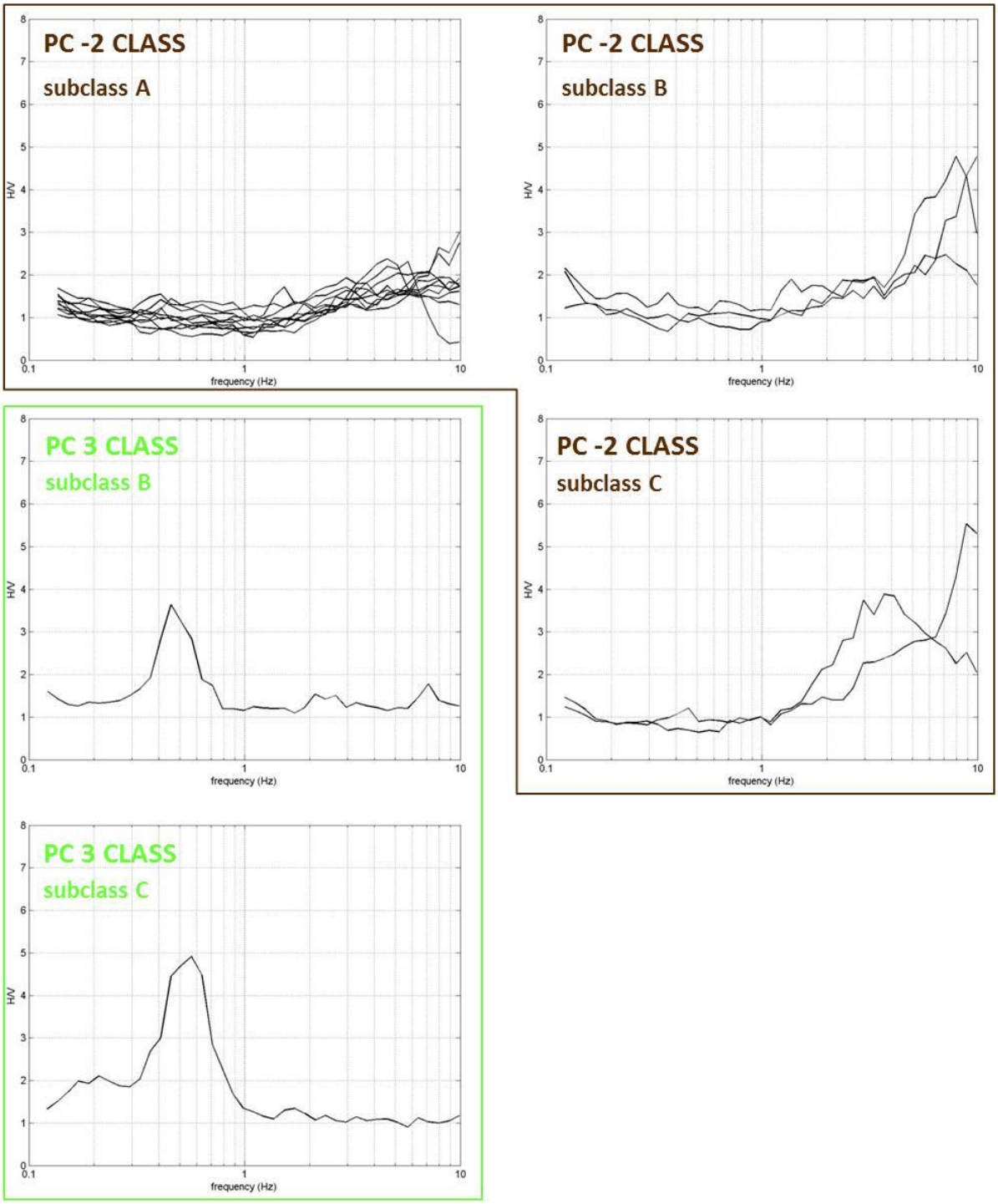

Fig. 7.21 - HVSR experimental curves belonging to PC -2 (brown border) and PC 3 (light green border) main classes. The panels associated of each main class show the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in Fig. 7.24. The border colors are related with the dot colors of the main classes in Fig. 7.24.

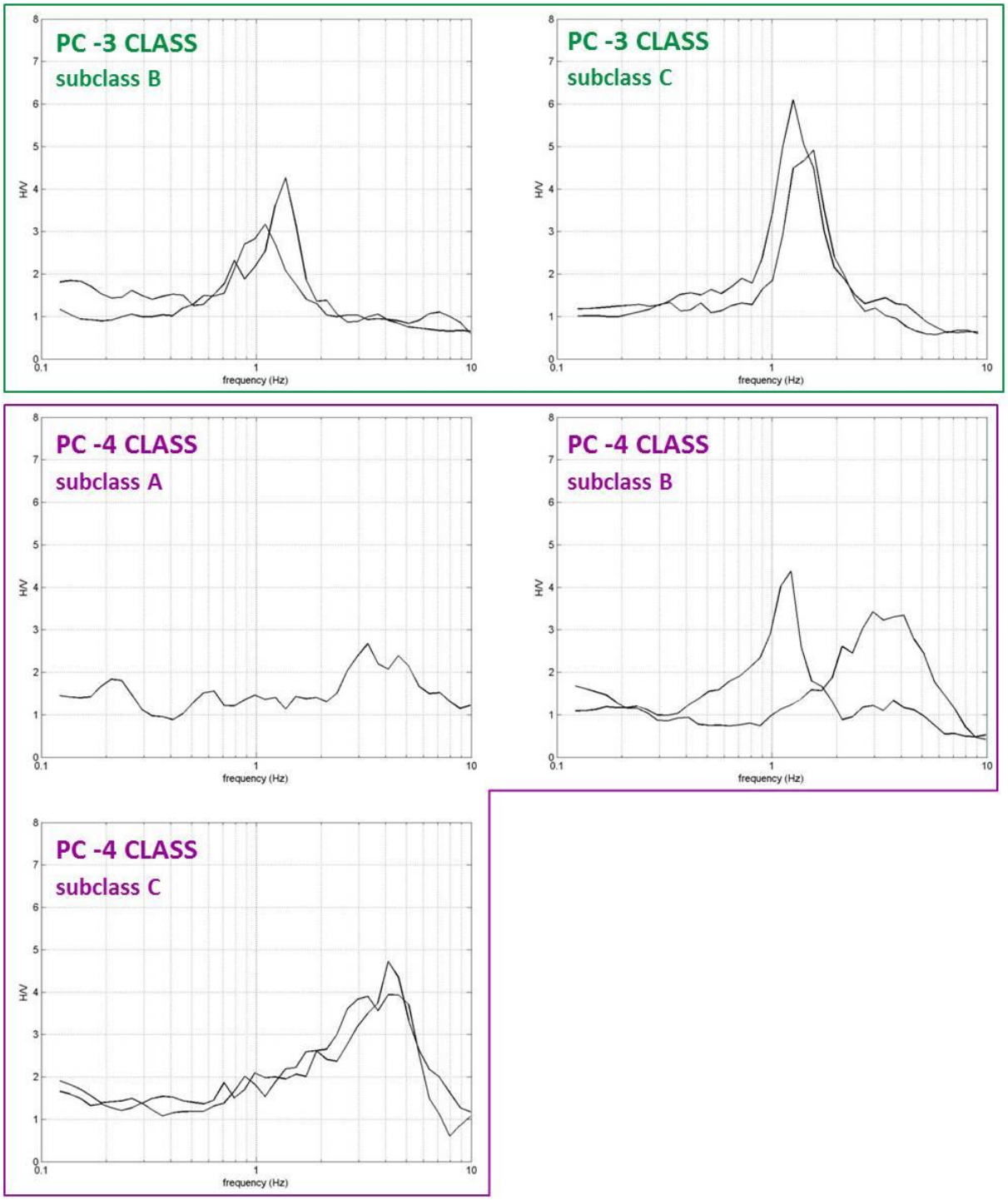

Fig. 7.22 - HVSR experimental curves belonging to PC -3 (dark green border) and PC -4 (violet border) main classes. The panels associated of each main class show the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in Fig. 7.24. The border colors are related with the dot colors of the main classes in Fig. 7.24.

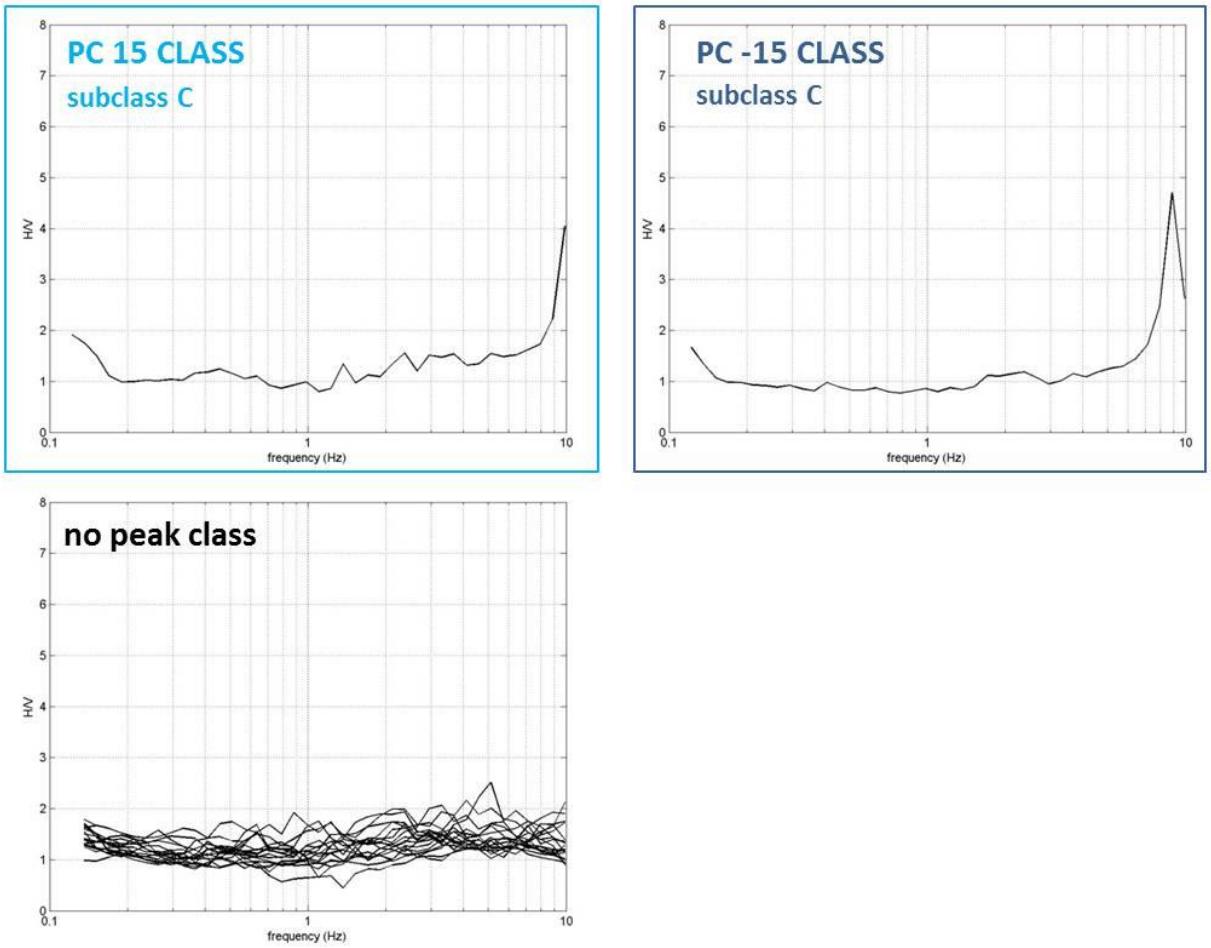

Fig. 7.23 - HVSR experimental curves belonging to PC 15 (light blue border), PC -15 (dark blue border) main classes and the identified flat HVSR curves (no peak class). The panels associated to the PC 15 and PC -15 main classes show the curves belonging to the indicated subclass obtained by the values of the loadings. The used e_{km} thresholds are showed in Fig. 7.24. The border colors of the PC 15 and PC -15 classes is related with the dot colors of these classes in Fig. 7.24.

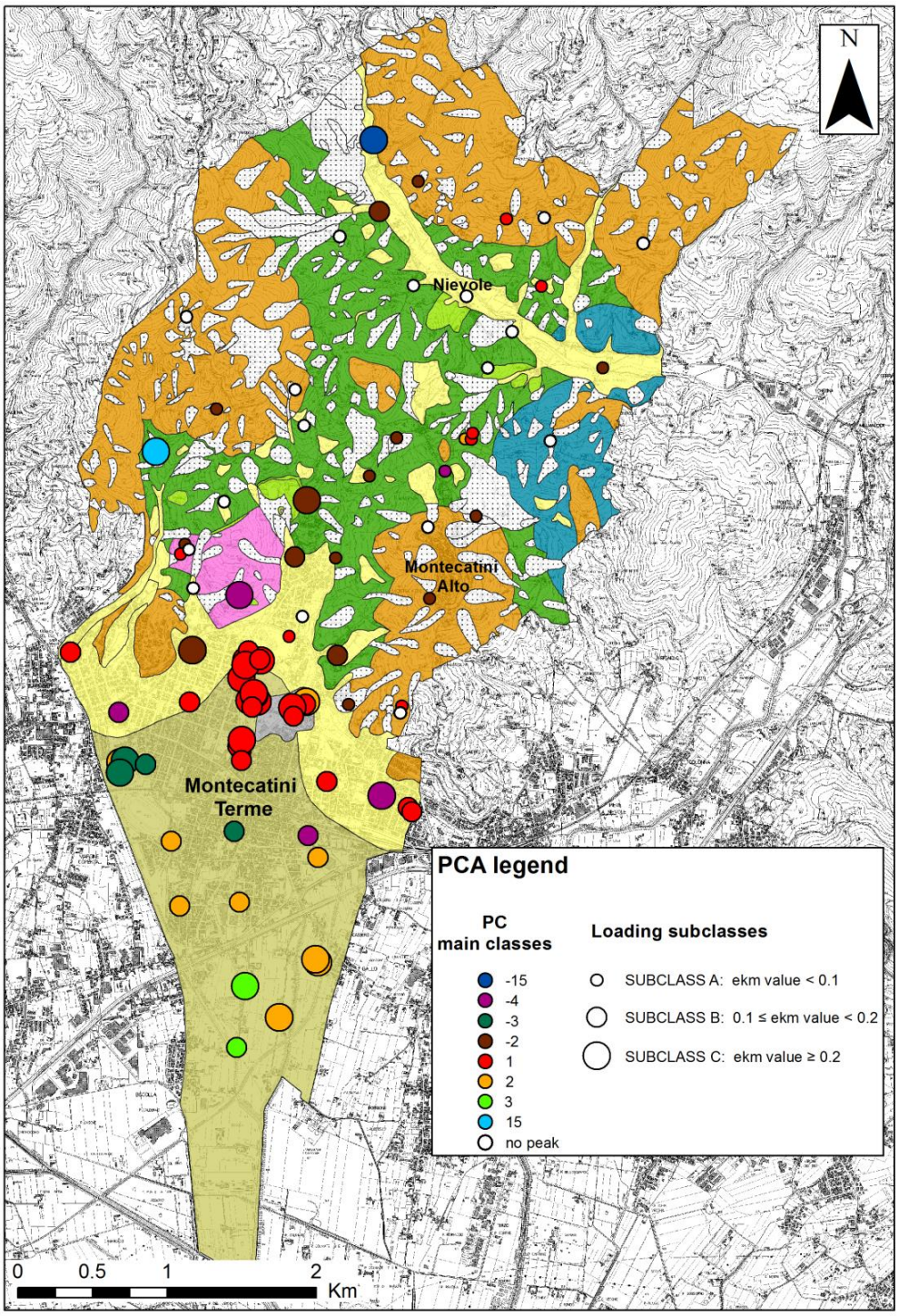

Fig. 7.24 - The map shows the spatial distribution of the main classes and their corresponding subclasses, indicated by the color and the size of the dots respectively. Moreover, the PCA legend on the bottom left shows the e_{km} thresholds used to perform the sub-classification. The geology legend is showed in Fig. 7.17.

CONCLUSIONS

The use of the passive seismic measurements for dynamic characterization of soils is significantly growing in the last few years. Moreover, thanks to the benefits provided by the exploitation of the ambient vibration wavefield, their application in the framework of Seismic Microzonation studies is of paramount importance in order to estimate the resonance frequency and the representative shear wave velocity profiles of the investigated area.

This PhD thesis is characterized by an applicative nature. Its main goals were:

- to verify and evaluate the reliability of the passive seismic techniques in presence of a sharp S-wave velocity inversion in the subsoil;
- to check the informative capability of the extensive use of the ambient vibration prospecting, also identifying the lateral geological heterogeneities at scale of some kilometers.

The objectives and the research lines dealt in this dissertation are devoted to improve the informative potential of the passive seismic measurements for purposes related to Seismic Microzonation studies. In particular, the first objective concerned the third investigation level of the Italian Seismic Microzonation guidelines (SM Working Group, 2015), where detailed and accurate procedures are required. The second objective concerned mainly the use of the data collected in the framework of the first investigation level, where information with extensive character as well as rougher and faster procedures are required.

Before the description of the applicative study cases, the first three Chapters of the thesis were devoted to review the state of the art concerning the origin and the nature of the ambient vibrations and the different analysis procedures adopted for the single station and seismic array measurements. Moreover, in the Chapter 4 a review of the methods aimed to estimate the V_S profile was given, both concerning the simplified approaches and the most common inversion procedures.

Regarding the first goal of the dissertation, the study carried out in the Maltese islands shows the optimal capability of the passive seismic techniques and the joint inversion procedure based on genetic algorithms in resolving both the presence and the characteristics of the thick buried low-velocity layer in the stratigraphy. The effective Rayleigh wave dispersion curves obtained using the ESAC method proved to be an excellent marker to identify this particular subsoil characteristic: in particular, these curves showed an inversely dispersive segment related to the S-wave velocity inversion and indicating the possible presence of higher mode surface waves. As concerns the HVSR curves, the drop below 1 over a wide frequency range which in this case is more evident where the whole stratigraphic succession is present, it is not always a reliable marker of a sharp S-wave velocity inversion in the subsurface. In fact, this feature is also clearly present in the HVSRs performed in the Emilia Romagna area, where the borehole seismic tests witnessed the absence of a significant S-wave velocity inversion in the first 50-100 m of subsoil. Finally, it is possible to state that this study case highlights the applicability and use of passive seismic methods in obtaining reliable V_S profiles not only in simple geological cases, but also in more challenging ones.

As concerns the second objective of this thesis, the extensive use of the passive seismic techniques in the large scale study carried out in Emilia Romagna (the investigated area is about 2500 km^2) allowed the application of a simplified procedure to assess the average shear velocity profile of the sedimentary cover and the depth of resonant interfaces present in the area. In this way, two main interfaces have been detected: at depths of about 50–100 and 500–600 m respectively. This model finds a plausible geological interpretation and has been also checked by considering advanced simulations. It is worth noting that both interfaces lie well below the depth range (0–30 m) considered in the engineering practice (down-hole, MASW, refraction, etc.). This suggests that simplified approach to site response (e.g., Eurocode8, EN-1998 2004) may reveal inadequate also in relatively simple seismostratigraphical configurations as the one considered in this work.

The extensive application of the single station measurements also allowed to emphasize the features of the HVSR curves achieved in different experimental conditions and in different acquisition periods. The most evident characteristic of the HVSR measurements is related to the variations of the amplitude values relative to HVSR peaks measured nearby: it's difficult to link these variations to the impedance contrast differences in the subsoil because they are spatially random, even at the scale of a single settlement. As described in Paragraph 6.4, the amplitude values of the low frequency peak seem to be significantly affected by the sea-waves activities. Regarding the main frequency peak, it is not clear if its amplitude values differences are related to other experimental conditions, like weather (for example the wind intensity) or anthropic activities. In view of these considerations, the amplitude differences in HVSR curves could make an ambiguous index of the impedance contrast in the subsoil: this feature can be critical in inversion procedures, where the bedrock shear wave velocity could be overestimated.

On the other hand, persistency of resonance frequency values observed in the whole area, indicate the robustness of HVSR peaks as experimental estimator of seismic resonance phenomena. The information showed by the identified HVSR Zones indicate the presence of significant seismic resonance phenomena in the range 0.8-1.4 Hz. These could have contributed to the collapses and to the serious damages of large dimension structure like industrial warehouse, the most damaged buildings in occasion of the mainshocks on May 20 and on May 29 (Galli *et al.*, 2012; Masi *et al.*, 2014). In almost the whole investigated area (except the Zone 3), this resonance peak is accompanied by another low frequency peak (0.25–0.3 Hz) of no interest for engineering purposes, but can result useful to reconstruct the deep seismic structure of the Po Plain.

The Zones identified by the visual inspection of the HVSR allowed to detect large scale areas (of the order of tens-hundreds square kilometers) where analogous seismic response is expected in case of future potentially damaging earthquakes. In order to perform an automatic identification of these areas, in Chapter 7 a procedure based on PCA was proposed. It was observed that this technique is able to group together similar HVSR curves also identifying the main patterns (the "apparent" HVSR curves) representing the overall variance of the original dataset. The analyzed cases demonstrate that this technique and the adopted grouping procedure discriminate especially well the large scale heterogeneities in contexts where significant geological changes occur: this is clearly evident in Collesalvetti and Montecatini Terme municipality study areas. As

seen in Emilia Romagna case, PCA are not able to identify the slight differences among the curves pointed out by the zonation performed by visual inspection. However, the procedure was able to clearly detect the main representative pattern (defined by the PC 1 “apparent” curve) which represents almost the 80% of the original dataset variance: this witnesses that the investigated area is characterized by a strong homogeneity among the HVSRs curves. Another evident drawback was found in Montecatini Terme study case: although concerning few measurements, the frequency shift which characterizes the curves located in the central and southern part of the plain caused the presence of unrealistic “apparent” HVSR curves (PC 2, PC 3 and PC -4). These dominant patterns, which explain respectively almost the 20%, 8% and 7% of the overall variance, probably represent a sort of merging of the experimental curves achieved in this zone.

The immediate representative pattern recognition is the main characteristic that distinguishes the PCA from the described techniques based on cluster analysis. In fact, in the latter methodologies the representative pattern recognition is obtained after the grouping identification averaging the relevant spectral ratios, while in PCA it represents the first step of the procedure. Moreover, PCA is characterized by a significant low computational complexity which makes this method definitely quicker than the cluster analysis procedures, in particular than that one proposed by Bragato *et al*. (2007): in fact, considering a dataset of about 100 HVSR curves, the latter procedure takes about 1 hour, while PCA takes only few seconds. Furthermore, PCA does not need a priori information (as the number of cluster in non-hierarchical clustering procedures) significant subjective choices (as the level of similarity where to cut the dendrogram in hierarchical clustering procedures) and it does not take into account any kind of geographical location information.

Finally, despite the grouping obtained by PCA can be refined by a subsequent visual inspection, it is possible to state that PCA adopted on HVSR data is a useful tool to obtain a quick first identification of the large scale heterogeneities of the study area and to identify the main representative patterns of the original HVSR dataset.

REFERENCES

AA.VV. (2011). *Contributi per l'aggiornamento degli "Indirizzi e criteri per la microzonazione sismica"*. Ingegneria sismica, 28 (2), Patron editore Bologna, Italy, available at: http://www.protezionecivile.gov.it/resources/cms/documents/aggiornamento_indirizzi_microzonazione_sismica.pdf.

Agius M.R., D'Amico S. and Galea P. (2015). *The Easter Sunday 2011 earthquake swarm off-shore Malta: Analysis on felt reports*. In D'Amico S. (Ed.), Earthquakes and their impact on society, Springer.

Akamatsu K. (1961). *On microseisms in frequency range from 1c/s to 200c/s.* Bull. Earthquake Res. Inst., Tokyo Univ., 39, 23-75.

Aki K. (1957). *Space and time spectra of stationary stochastic waves, with special reference to microtremors.* Bulletin of Earthquake Research Institute, 35, 415 – 456.

Aki K. (1965). *A note on the use of microseisms in determining the shallow structures of the earth's crust.* Geophysics, 30, 665-666.

Aki K. (1988). *Local site effects on ground motion.* Earthquake Engineering and Soil Dynamics II, 103 – 155.

Aki K. and Richards P.G. (2002). *Quantitative seismology.* Univ.Sci. Book, Sausalito, CA, USA, 700 pp.

Albarello D. and Baliva F. (2009). *In-situ estimates of material damping from environmental noise measurements*. In: Mucciarelli, M., Herak, M. and Cassidy, J. (Eds.) Increasing Seismic Safety by Combining Engineering Technologies and Seismological Data, NATO Science for Peace and Security Series C: Environmental, Springer, XVIII, 382.

Albarello D. and Gargani G. (2010). *Providing NEHRP soil classification from the direct interpretation of effective Rayleigh waves dispersion curves*. Bull.Seism.Soc.Am., 100 (6), 3284‑3294, doi: 10.1785/0120100052.

Albarello D., Cesi C., Eulilli V., Guerrini F., Lunedei E., Paolucci E., Pileggi D., Puzzilli L.M. (2011). *The contribution of the ambient vibration prospecting in seismic microzoning: an example from the area damaged by the 26th April 2009 L'Aquila (Italy) earthquake*. Boll. Geofis.Teor.Appl., 52 (3), 513-538, doi: 10.4430/bgta0013.

Albarello D. and Lunedei E. (2011). *Structure of an ambient vibration wavefield in the frequency range of engineering interest ([0.5, 20] Hz): insights from numerical modelling.* Near Surface Geophys 9:543–559. doi:10.3997/1873-0604.2011017.

Albarello D. and Lunedei E. (2013). *Combining horizontal ambient vibration components for H/V spectral ratio estimates.* Geophys J Int 194:936–951. doi:10.1093/gji/ggt130.

Albarello D., Socco L.V., Picozzi M., Foti S. (2015). *Seismic hazard and land management policies in Italy: the role of seismic investigation*. First Break, 33: 79-85.

Albarello D., Francescone M., Lunedei E., Paolucci E., Papasidero M.P., Peruzzi G., Pieruccini P. (2016). *Seismic characterization and reconstruction of reference ground motion at accelerometric sites of the Italian national accelerometric network (RAN).* Natural Hazards, http://link.springer.com/article/10.1007/s11069-016-2310-4.

Arai H. and Tokimatsu K. (1998). *Evaluation of local site effects based on microtremor H/V spectra.* Proc. of the Second International Symposium on the Effects of Surface Geology on Seismic Motion, Yokohama (Japan), 2, 673-680.

Arai H. and Tokimatsu K. (2004). *S-wave velocity profiling by inversion of microtremor H/V spectrum.* Bulletin of Seismological Society of America, 94, 1, 53-63.

Arai H. and Tokimatsu K. (2005). *S-wave velocity profiling by joint inversion of microtremor dispersion curve and horizontal-to-vertical (H/V) spectrum.* Bull. Seism. Soc. Am., 95, 5, 1766 - 1778.

Asten M. W. (1976). *The use of microseisms in geophysical exploration.* PhD Thesis, University of Macquarie, North Ryde (Australia).

Asten M. W. (1978). *Geological control of the three-component spectra of rayleigh-wave microseisms.* Bulletin of Seismological Society of America, 68, 1623-1636.

Asten M. W. and Henstridge J. D. (1984). *Array estimators and use of microseisms for reconnaissance of sedimentary basins.* Geophysics, 49, 1828-1837.

Asten M. W. (2006). *On bias and noise in passive seismic data from finite circular array data processed using SPAC method.* Geophysics, 71, V153-V162, doi:10.1190/1.2345054.

Baliva M. (2006). *Profilo di velocità delle onde S dallo studio dei microtremori* (in Italian). PhD. Thesys, University of Siena, 182 pp.

Barbarito L. (1999). *L'analisi di settore: metodologia e applicazioni*. Milano, Franco Angeli ed.

Bard P.Y. (1999). *Microtremor measurements: a tool for site effect estimation?* The Effects of Surface Geology on Seismic Motion, eds. K. Irikura, K. Kudo, H. Okada and T. Sasatani (Balkema, Rotterdam), 1251-1279.

Bencivenga M., Nardone G., Ruggiero F., Calore D. (2012). *The Italian Data Buoy Network (RON)*. In: Advances in Fluid Mechanics IX. Edited by WIT, UK, ISBN: 978-1-84564-600-4, 2012, pp. 321-332.

Bensen G.D., Barmin M.P., Levshin A.L., Lin F., Moschetti M.P., Shapiro N.M., Yang Y. and Ritzwoller M.H. (2007). *Processing ambient noise seismic data to obtain reliable broadband surface wave dispersion measurements*. Geophys. J. Int*.,* 169, 1239–1260, doi:10.1111/j.1365–246X.2007.03374.x.

Bigi G., Cosentino D., Parotto M., Sartori R., Scandone P. (1990). *Structural model of Italy, 1:500000*. CNR, Progetto Finalizzato Geodinamica.

Bindi D., Parolai S., Spallarossa D. and Cattaneo M. (2000). *Site effects by H/V ratio: comparison of two different procedures.* Journal of Earthquake Engineering, 4, 1, 97-113.

Birtill J. W. and Whiteway F. E. (1965). *The application of phased arrays to the analysis of seismic body waves*. Phil. Trans. R. Soc. of London, Math. and Phys. Sciences, A-258**,** No. 1091, 421-493.

Bonnefoy-Claudet S., Cornou C., Kristek J., Ohrnberger M., Wathelet M., Bard P.Y., Moczo P., Fäh D., Cotton F. (2004). *Simulation of seismic ambient noise: I. Results of H/V and array techniques on canonical models.* In: Proceedings of the 13th world conferences on earthquake engineering (WCEE), Vancouver.

Bonnefoy-Claudet S., Cornou C., Bard P.-Y., Cotton F., Moczo P., Kristek J. and Fäh D. (2006a). *H/V ratios: A tool for site effects evaluation. Results from 1-D noise simulations*. Geophysical Journal International 167, 827–837.

Bonnefoy-Claudet S., Cotton F. and Bard P. Y. (2006b). *The nature of noise wavefield and its applications for site effects studies: A literature review*. Earth Science Review 79, 205–227.

Bonnefoy-Claudet S., Kohler A., Cornou C., Wathelet M., Bard P.Y. (2008). *Effects of Love waves on microtremor H/V ratio.* Bull Seismol Soc Am 98(1):288–300. doi:10.1785/0120070063.

Boore, D. (2015). *Notes on relating density to velocity for use in site amplification calculations*. Available online: http://www.daveboore.com/daves_notes/daves_notes_on_relating_density_to_velocity_v3.0.pdf. Last accessed: 24 August 2015.

Borcherdt R.D. (1970). *Effects of local geology on ground motion near San Francisco bay*. Bulletin of the Seismic Society of America, 60: 29 –61.

Bordoni P., Azzara R.M., Cara F., Cogliano R., Cultrera G., Di Giulio G., Fodarella A., Milana G., Pucillo S., Riccio G., Rovelli A., Augliera P., Luzi L., Lovati S., Massa M., Pacor F., Puglia R., Ameri G. (2012). *Preliminary results from EMERSITO, a rapid response network for site-effect studies*. Annals of geophysics, 55, 4, 599-607, doi: 10.4401/ag-6153.

Bormann P. (1966). *Recording and interpretation of seismic events (principles, present state and tendencies of development)* (in German). Publ. Inst. Geodyn., Jena, Akademie Verlag, Berlin, Vol. 1, 158 pp.

Bormann P. and Wielandt E. (2013). *Seismic signals and noise*. In: Bormann, P. (Ed.), New Manual of Seismological Observatory Practice 2 (NMSOP-2), Potsdam: Deutsches GeoForschungsZentrum GFZ, pp. 1—62. DOI: 10.2312/GFZ.NMSOP-2_ch4.

Bragato P.L., Laurenzano G., Barnaba C. (2007). *Automatic zonation of urban areas based on the similarity of H/V spectral ratios*. Bull Seism Soc Am 97(5):1404–1412. Doi: 10.1785/0120060245.

Brandi G.P., Fritz P., Raggi G., Squarci P., Taffi L., Tongiorgi E., Trevisan L. (1967). *Idrogeologia delle Terme di Montecatini.* Collana scientifica delle Terme di Montecatini, 39, Montecatini Terme.

Burjanek J., Gassner-Stamm G., Poggi V., Moore J.R. and Fäh D. (2010). *Ambient vibration analysis of an unstable mountain slope*. Geophys. J. Int.*,* 180(2), 820–828.

Burjanek J., Moore J.R., Yugsi Molina F.X. and Fäh D. (2012). *Instrumental evidence of normal mode rock slope vibration*. Geophys. J. Int.*,* 188(2), 559–569.

Capizzi P., Martorana R., D'Alessandro A., Luzio D. (2015). *Contribution ofthe cluster analysis of HVSR data for near surface geological reconstruction*. Congress proceedings 34° Convegno Nazionale GNGTS, Atti - Tema 3: Geofisica Applicata, 50-56.

Campillo M. and Paul A. (2003). *Long-range correlations in the seismic coda*. Science, 299, 547–549.

Capon J. (1969). *High-resolution frequency-wavenumber spectrum analysis.* Proceedings of the IEEE, 57, 8, 1408 - 1418.

Cara F., Di Giulio G., Milana G., Bordoni P., Haines J. And Rovelli A. (2010). *On the stability and reproducibility of the Horizontal-to-Vertical Spectral Ratios on ambient noise: case study of Cavola, Northern Italy*. Bull. Seism. Soc. Am., 100, 3, 1263-1275, doi: 10.1785/0120090086.

Carminati E. and Doglioni C. (2012). *Alps vs. Apennines: The paradigm of a tectonically asymmetric Earth*. Earth Sci. Rev., 112, 67-96; doi:10.1016/ j.earscirev.2012.02.004.

Castellaro S. and Mulargia F. (2009). *The effect of velocity inversions on H/V*. Pure and Applied Geophysics, 166, 567–592.

Chavez-Garcıa F. J. and Luzon F. (2005). *On the correlation of seismic microtremors.* J. Geophys. Res., 110, B11313, doi:10.1029/2005JB003671.

Chavez-Garcıa F .J. and Rodrıguez M. (2007). *The correlation of microtremors: empirical limits and relations between results in frequency and time domains.* Geophys. J. Int., 171, 657–664.

Chouet B., De Luca G., Milana G., Dawson P., Martini M. and Scarpa R. (1998). *Shallow velocity of Stromboli volcano, Italy, derived from small-aperture array measurements of Strombolian tremor.* Bull. Seism. Soc. Am., 88, 3, 653-666.

Cooley J.W. and Tukey J.W. (1965). *An algorithm for the machine calculation of complex Fourier series.* Math. Comput., 19, 297–301.

Cornou C. (2002). *Traitement d'antenne et imagerie sismique dans l'agglomération grenobloise (Alpes françaises): implications pour les effets de site.* Phd Thesis, Joseph Fourier University, Grenoble (France), pp. 260.

Cox H. (1973). *Spatial correlation in arbitrary noise fields with applications to ambient sea noise.* J. Acoust. Soc. Am., 54, 1289–1301.

D'Amico V., Albarello D., Baliva F., Picozzi M. and Agili F. (2006a). *Site response characterization of the Florence urban area (Italy) using seismic noise measurements.* In: 1st European Conference on Earthquake Engineering and Seismology, Geneva, Switzerland, Abstract Book, pp. 333.

D'Amico V., Picozzi M., Baliva F., Albarello D., Menichetti M., Bozzano F., Martino S., Rivellino S. and Scarascia Mugnozza G. (2006b). *Test sites in Europe for the evaluation of ground motion amplification: site response of the Gubbio basin (central Italy) using geological data and seismic noise measurements*. In: 1st European Conference on Earthquake Engineering and Seismology, Geneva, Switzerland, Abstract Book, pp. 301.

D'Amico V., Picozzi M., Baliva F. and Albarello D. (2008). *Ambient Noise Measurements for Preliminary Site-Effects Characterization in the Urban Area of Florence*. Bull.Seism.Soc.Am., 98 (3), 1373‑1388, doi:10.1785/0120070231.

Davis J.C. (2002). *Statistics and Data Analysis in Geology*. Third edition; New York, Chichester, Brisbane, Toronto, Singapore: John Wiley & Sons, 656 pp., ISBN 0-471-17275-8.

Delgado J., Lopez Casado C., Estevez A., Giner J., Cuenca A. and Molina S. (2000a). *Mapping soft soils in the Segura river valley (SE Spain): a case study of microtremors as an exploration tool.* J. Appl. Geophys., 45, 19–32.

Delgado J., Lopez Casado C., Giner J., Estevez A., Cuenca A. and Molina S. (2000b). *Microtremors as a geophysical exploration tool: applications and limitations*. Pure Appl. Geophys., 157, 1445–1462.

Di Giacomo D., Gallipoli M.R., Mucciarelli M., Parolai S. and Richwalski S.M. (2005). *Analysis and modeling of HVSR in the presence of a velocity inversion: the case of Venosa, Italy*. Bulletin of the Seismological Society of America, 95, 2364–2372.

Di Giulio G., Savvaidis A., Ohrnberger M., Wathelet M., Cornou C., Knapmeyer-Endrun B., Renalier F., Theodoulidis N. and Bard P.Y. (2012). *Exploring the model space and ranking a best class of models in surface wave dispersion inversion: Application at European strong-motion sites.* Geophysics, 77, 3, B147-B166, 10.1190/GEO2011-0116.1.

Dong S., He R. and Schuster G. (2006). *Interferometric prediction and least squares subtraction of surface waves*. In Proceedings of 76th Annual International Meeting (expanded abstracts), Society of Exploration Geophysicists, 2783–2786.

Dorman J. and Ewing M. (1962). *Numerical Inversion of seismic surface waves dispersion data and crust-mantle structure in the New York-Pennsylvania area*. J. of Geophysical Research, 67, 5227-5241.

Douze E. J. (1964). *Rayleigh waves in short-period seismic noise.* Bulletin of the Seismological Society of America, 54, 4, 1197-1212.

Douze E. J. (1967). *Short-period seismic noise.* Bulletin of the Seismological Society of America, 57, 1, 55-81.

Dziewonski A., Bloch S. and Landisman M. (1969). *A technique for the analysis of transient seismic signals*. Bull. seism. Soc. Am., 59, 427–444.

Dziewonski A.M. and Hales A.L. (1972). *Numerical analysis of dispersed seismic waves.* In Method in Computational Physics, Volume 11, edited by B.A. Bolt, Academic, pp. 39–85.

Eiben A.E. and Smith J.E. (2003). *Introduction to Evolutionary Computing.* Sprinter, 3-540-40184-9.

Ekstrom G., Abers G.A. and Webb S.C. (2009). *Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation*. Geophys. Res. Lett., 36, L18301, doi:10.1029/2009GL039131.

Ekstrom G. (2014). *Love and Rayleigh phase-velocity maps, 5–40s, of the western and central USA from USArray data*. Earth and Planetary Science Letters, 402, 42–49.

Elter P., Gratziu C. and Labesse B. (1964). *Sul significato dell'esistenza di una unità tettonica alloctona costituita da formazioni terziarie nell'Appennino settentrionale*. Boll. Soc.Geol.It., 83(2):373-394.

Elter P. and Marroni M. (1991). *Le Unità Liguri dell'Appennino Settentrionale: sintesi dei dati e nuove interpretazioni*. Mem. Descr. Serv. Geol. Italiano, XLVI: 121-138.

EN 1998-1 Eurocode 8 (2004). *Design of structures for earthquake resistance – part 1: general rules, seismic actions and rules for buildings*. Authority: The European Union per regulations 305/2011, Directive 98/34/EC, Directive 2004/18/EC.

Everitt B.S., Landau S., Leese M., Stahl D. (2011). *Cluster Analysis*. Wiley Series in Probability and Statistics, 5th Edition, pp. 332, ISBN-10: 0470749911, ISBN-13: 978-0470749913.

Fabbris L. (1983). *Analisi esplorativa di dati multidimensionali*, Cleup ed.

Fäh D. (1997). *Microzonation of the city of Basel.* Journal of Seismology 1 (1), 87–102.

Fäh D., Kind F., Giardini D. (2001). *A theoretical investigation of average H/V ratios*. Geophys. J. Int. 145, 535-549.

Farrugia D., Paolucci E., D'Amico S., Galea P. (2015). *Site characterisation and response study in Rabat, Malta*. Congress proceedings "Georisks in the Mediterranean and their mitigation", Valletta (Malta), 2015.

Farrugia D., Paolucci E., D'Amico S., Galea P. (2016*). Inversion of surface-wave data for subsurface shear-wave velocity profiles characterised by a thick buried low-velocity layer.* Geophysical Journal International, 206(2): 1221-1231, doi: 10.1093/gji/ggw204.

Ferrari G., Albarello D., Martinelli G. (2000). *Tromometric measurements as a tool for crustal deformation monitoring.* Seismological Research Letters, 71 (5), 562-569, ISSN: 0895-0695.

Field E. and Jacob K. (1993). *The theoretical response of sedimentary layers to ambient seismic noise.* Geophys Res Lett 20(24):2925–2928.

Forbriger T. (2006). *Low-frequency limit for H/V studies due to tilt.* Sitzung der Arbeitsgruppe Seismologie des FKPE. Haidhof, Germany, Extended Abstract, 32.

Foti S. (2000). *Multistation methods for geotechnical characterization using surface waves*. PhD Thesis, Politecnico di Torino, Italy.

Foti S., Parolai S., Albarello D., Picozzi M. (2011). *Application of surface–wave methods for seismic site characterization*. Surv. Geophys., 32 (6), 777-825. DOI 10.1007/s10712-011-9134-2.

Frosch R. A. and Green P. E. (1966). *The concept of the large aperture seismic array*. Proc. R. Soc., London, Ser. A, 290, 368-384.

Galea P. (2007). *Seismic history of the Maltese islands and considerations on seismic risk*. Annals of geophysics, 50, 725‑740.

Galea P., D'Amico S. and Farrugia D. (2014). *Dynamic characteristics of an active coastal spreading area using ambient noise measurements - Anchor Bay, Malta*. Geophysical Journal International, 199, 1166-1175.

Gallagher K. and Sambridge M. (1994). *Genetic algorithms: A powerful tool for large-scale non-linear optimization problems.* Comput. Geosci., 20(7/8), 1229–1236.

Galli P., Castenetto S., Peronace E. (2012). *The MCS macroseismic survey of the Emilia 2012 earthquakes*. Annals of geophysics, 55, 4, 663-672, doi: 10.4401/ag-6163.

Gallipoli M.R. and Mucciarelli M. (2009). *Comparison of site classification from VS30, VS10, and HVSR in Italy*. Bulletin of the Seismological Society of America, 99, 340-351.

Gallipoli M.R., Albarello D., Mucciarelli M. and Bianca M. (2011). *Ambient noise measurements to support emergency seismic microzonation: the Abruzzo 2009 earthquake experience.* Boll.Geofis.Teor Appl., 52 (3), 539-559.

Gallipoli M.R., Chiauzzi L., Stabile T.A., Mucciarelli M., Masi A., Lizza C., Vignola L. (2014). *The role of site effects in the comparison between code provisions and the near field strong motion of the Emilia 2012 earthquakes*. Bull Earthq Eng 12:2211–2230. doi:10.1007/s10518-014-9628-7.

Gan G., Ma C., Wu J. (2007). *Data Clustering: Theory, Algorithms, and Applications*. Cambridge University Press, pp. 184, ISBN: 9780898716238.

Gerstoft P., Sabra K. G., Roux P., Kuperman W. A. and Fehler M. C. (2006). *Green's functions extraction and surface-wave tomography from microseisms in southern California.* Geophysics, 71, 23–32.

Gigli G., Frodella W., Mugnai F., Tapete D., Cigna F., Fanti R., Intrieri E. and Lombardi L. (2012). *Instability mechanisms affecting cultural heritage sites in the Maltese Archipelago*. Natural Hazards and Earth System Science, 12, 1883‐1903.

Goldberg D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, Reading, Mass.

Government of Malta (1958). *Reports on Geological Investigation of the Maltese islands*. R. Costain Ltd.

Gruppo di Lavoro sulla Microzonazione Sismica (2008). *Indirizzi e criteri per la microzonazione sismica.* Conferenza delle Regioni e delle Province autonome – Dipartimento della Protezione Civile, Roma, 3 vol. e DVD, available at: http://www.protezionecivile.gov.it/jcms/it/view_pub.wp?contentId=PUB1137.

Gruppo di Lavoro sulla Microzonazione dell'area Aquilana - MS-AQ (2010). *Microzonazione sismica per la ricostruzione dell'area aquilana.* Regione Abruzzo, Dipartimento della Protezione Civile, 796 (3 VV e DVD), available at: http://www.protezionecivile.gov.it/jcms/it/view_pub.wp?contentId=PUB25330.

Groos J. C. and Ritter J. R. R. (2009). *Time domain classification and quantification of seismic noise in an urban environment.* Geophys. J. Int., 179, 1213-1231; doi: 10.1111/j.1365-246X.2009.04343.x.

Gutenberg B. (1958). *Microseisms.* Advanced Geophysics, 5, 53-93.

Halliday D.F., Curtis A., Robertsson J.O.A. and Van Manen D.-J. (2007). *Interferometric surface-wave isolation and removal*. Geophysics, 72, A67–A73.

Halliday D.F., Curtis A. and Kragh E. (2008). *Seismic surface waves in a suburban environment - active and passive interferometric methods.* Leading Edge, 27, 210–218.

Haskell N.A. (1953). *The dispersion of surface waves on multilayered media*. Bull. Seismol. Soc. Am. 43, 17–34.

Herak M. (2008). *ModelHVSR – a Matlab® tool to model horizontal-to-vertical spectral ratio of ambient noise.* Comput Geosci 34(11):1514–1526. doi:10.1016/j.cageo.2007.07.009.

Herrmann R. B. (2002). *Computer Programs in Seismology. Surface waves, receiver functions and crustal structure.* Versione 3.25, Saint Louis University (Missouri, USA).

Hinzen K.G., Scherbaum F. and Weber B. (2004). *On the resolution of H/V measurements to determine sediment thickness, a case study across a normal fault in the Lower Rhine embayment, Germany*. J. Earthq. Eng., 8, 909–926.

Hobiger M, Cornou C., Wathelet M., Di Giulio G., Knapmeyer-Endrun B., Renalier F., Bard P.Y., Savvaidis A., Hailemikael S., Le Bihan N., Ohrnberger M. and Theodoulidis N. (2013). *Ground structure imaging by inversions of Rayleigh wave ellipticity: sensitivity analysis and application to European strong-motion sites.* Geophys. J. Int., 192, 207-229, doi: 10.1093/gji/ggs005.

Holland J.H. (1975). *Adaptation in Natural and Artificial Systems.* Univ. of Mich. Press, Ann Arbor, Mich.

Horike M. (1985). *Inversion of phase velocity of long period microtremors to the S-wave-velocity structure down to the basement in urbanized areas.* J. Phys. Earth, 33, 59-96.

Ibs Von Seht M. and Wohlemberg J. (1999). *Microtremor measurements used to map thickness of soft sediments.* Bull. Seismol. Soc. Am., 89, 250–259.

Illies J.H. (1981). *Graben formation - the Maltese islands - a case history*. Tectonophysics, 73, 151–168.

Jolliffe I.T. (2002). *Principal Component Analysis*, 2nd Ed. Springer, 487 pp.

Kanai K., Tanaka T. and Osada K. (1954). *Measurement of the micro-tremor I*. Bull. Earthquake Res. Inst., Tokyo Univ., 12, 192-210.

Kawase H., Sanchez-Sesma F.J., Matsushima S. (2011). *The optimal use of horizontal-to-vertical spectral ratios of earthquake motions for velocity inversions based on diffuse-field theory for plane waves.* Bull Seismol Soc Am 101(5):2001–2014. doi:10.1785/0120100263.

Kind F., Fäh D. and Giardini D. (2005). *Array measurements of S-wave velocities from ambient vibrations.* Geophysical Journal International, 160, 114-126.

Köhler A., Ohrnberger M., Scherbaum F., Wathelet M. and Cornou C. (2007). *Assessing the reliability of the modified three-component spatial autocorrelation technique.* Geophys J. Int., 168, 779-796; doi:10.1111/j.1365-246X.2006.03253.x.

Konno K. and Ohmachi T. (1998). *Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor.* Bulletin of Seismological Society of America, 88, 1, 228-241.

Konno K. and Kataoka S. (2000). *An estimating method for the average S-wave velocity of ground from the phase velocity of Rayleigh wave.* Proceedings of JSCE., 647, 415–423.

Kramer S.L. (1996). *Geotechnical Earthquake Engineering.* Prentice Hall, New Jersey, USA.

Kruskal J. B. (1956). *On the shortest spanning subtree of a graph and the travelling salesman problem*. Proc. Am. Math. Soc. 7, 48–50.

Kudo K. (1995). *Practical estimates of site response. State-of-art report.* Proceedings of the fifth Internationnal Conference on Seismic Zonation, Nice (France).

Kværna T. (1989). *On exploitation of small-aperture NORESS type arrays for enhanced P- wave detectability*. Bull. Seism. Soc. Am., 79, 888-900.

Lacoss K. T., Kelly E. J. and Toksöz M.N. (1969). *Estimation of seismic noise structure using arrays*. Geophysics, 34, 1, 21 - 38.

Lai C. G. (1998). *Simultaneous inversion of Rayleigh phase velocity and attenuation for near-surface site characterization.* PhD Thesys, Georgia Inst. Technology, Atlanta (Georgia, USA).

Lachet C. and Bard P.Y. (1994). *Numerical and theoretical investigations on the possibilities and limitations of Nakamura's technique*. J Phys Earth 42:377–397.

Larose E., Margerin L., Derode A., Van Taggelen B., Campillo M., Shapiro N., Paul A., Stehly L., Tanter M. (2006). *Correlation from random wave-field.* Geophys., 71(4), SI11–SI21.

Lazzarotto A., Mazzanti R., Nencini C. (1990a). *Carta Geologica dei comuni di Livorno e di Collesalvetti in scala 1:25:000.* Quaderni del Museo di Storia Naturale di Livorno, 11.

Lazzarotto A., Mazzanti R., Nencini C. (1990b). *Geologia e geomorfologia dei Comuni di Livorno e Collesalvetti*. Suppl. 2 ai Quaderni del Museo di Storia Naturale di Livorno, 11, 1-85.

Lermo J. and Chavez-Garcia F. J. (1993). *Site effect evaluation using spectral ratios with only one station.* Bull. of Seism. Society of America, 83, 5, 1574-1594.

Levshin A. L., T. B. Yanocskaya, A. V. Lander, B. G. Bukchin, M. P. Barmin, L. I. Ratnikova, and E. N. Its (1989). *Seismic Surface Waves in a Laterally Inhomogeneous Earth*. Edited by V. I. Keilis-Borok, Kluwer Acad., Norwell, Mass.

Li T. M. C., Ferguson J. F., Herrin E. and Durham H. B. (1984). *High-frequency seismic noise at Lajitas, Texas.* Bull. Seism. Soc. Am., 74, 5, 2015-2033.

Lin F., M. P. Moschetti and M. H. Ritzwoller (2008). *Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps.* Geophys. J. Int., 173, 281–298, doi:10.1111/j1365-1246X.2008.3720.x.

Ling S. and Okada H. (1993). *An extended use of the spatial autocorrelation method for the estimation of structure using microtremors* (in Japanese). Proc. 89th SEGJ Conference, Nagoya, Japan.

Liu Y., Luke B, Pullammanappallil S., Louie J. N. and Bay J. (2005). *Combining active- and passive-source measurements to profile shear wave velocities for seismic microzonation.* Congress proceedings of Geo-Frontiers 2005, Austin, Texas (USA).

Louie J. N. (2001). *Faster, better: shear-wave velocity to 100 meters depth from refraction microtremor arrays.* Bulletin of Seismological Society of America, 91, 2, 347-364.

Lunedei E. and Albarello D. (2009). *On the seismic noise wavefield in a weakly dissipative layered Earth.* Geophys J Int 177:1001–1014. doi:10.1111/j.1365-246X.2008.04062.x (Erratum: Geophys J Int 179:670. doi:10.1111/j.1365-246X.2009.04344.x).

Lunedei E. and Albarello D. (2010). *Theoretical HVSR curves from the full wave field modelling of ambient vibrations in a weakly dissipative layered Earth*. Geophysical Journal International 181, 1093–1108. DOI: 10.1111/j.1365-246X.2010.04560.x.

Lunedei E. and Albarello D. (2015). *Horizontal-to-vertical spectral ratios from a full-wavefield model of ambient vibrations generated by a distribution of spatially correlated surface sources.* Geophys J Int 201:1140–1153. doi:10.1093/gji/ggv046.

Lunedei E. and Malischewsky P. (2015). *A review and some new issues on the theory of the H/V technique for ambient vibrations.* A. Ansal (ed.), Perspectives on European Earthquake Engineering and Seismology, Geotechnical, Geological and Earthquake Engineering 39, DOI 10.1007/978-3-319-16964-4_15.

MacQueen J. B. (1967). *Some methods for classification and analysis of multivariate observations*. In Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, L. M. Le Cam and J. Neyman (Editors), Vol. 1, University of California Press, Berkeley, California, 281–297.

Malagnini L., Rovelli A., Hough S.E. and Seeber L. (1993), *Site amplification estimates in the Garigliano Valley, Central Italy, based on dense array measurements of ambient noise.* Bull. Seism. Soc. Am., 83, 1744-1755.

Martelli L. and Romani M. (2013). *Microzonazione Sismica e analisi della condizione limite per l'emergenza delle aree epicentrali dei terremoti della pianura emiliana di maggio-giugno 2012, relazione illustrativa*. Servizio geologico, sismico e dei suoli Regione Emilia Romagna,http://ambiente.regione.emilia-romagna.it/geologia/temi/sismica/speciale-terremoto/sisma-2012-ordinanza-70-13-11-2012-cartografia.

Martelli L., Severi P., Biavati G., Rosselli S., Camassi R., Ercolani E., Marcellini A., Tento A., Gerosa D., Albarello D., Guerrini F., Lunedei E., Pileggi D., Pergalani F., Compagnoni M., Fioravante V., Giretti D. (2014). *Analysis of the local seismic hazard for the stability tests of the main bank of the Po river (Northern Italy)*. Boll. Geofis. Teor. Appl., 55, 1, 119-134, doi:10.4430/bgta0094.

Martin A.J. and Diehl J.G. (2004). *Practical experience using a simplified procedure to measure average shear wave velocity to a depth of 30 meters (VS30)*. In: 13th World Conference on Earthquake Engineering, Vancouver, B.C., Canada, August 1–6, 2004, paper n. 952.

Martinez M. D., Lana X., Badal J., Canas J. A., Pujades L. (1999). *Simulated annealing in 3-D seismic modelling and elastic structure of the Mediterranean basin.* Phys. Chem. Earth, 24, 3, 253–260.

Martinez M. D., Lana X., Olarte J., Badal J. and Canas J. A. (2000). *Inversion of Rayleigh wave phase and group velocity by simulated annealing.* Physics of the Earth and planetary interiors, 122, 3-17.

Masi A., Santarsiero G., Gallipoli M.R., Mucciarelli M., Manfredi V., Dusi A., Stabile T.A. (2014). *Performance of the health facilities during the 2012 Emilia earthquake and analysis of the Mirandola Hospital case study*. Bull Earthq Eng 12:2419–2443. doi:10.1007/s10518-013-9518-4.

Menke W. (1989). *Geophysical data analysis: discrete inverse theory.* International Geophysical Series, Academic Press, Inc., vol. 45, pp. 289.

Métaxian J.-P. and P. Lesage (1997). *Permanent tremor of Masaya Volcano, Nicaragua: Wave field analysis and source location*. Journal of Geophysical Research, 102, 22529-22545.

Metropolis N. and Ulam S. M. (1949). *The Monte Carlo method.* J. Am. Stat. Ass., 44, 335-341.

Mucciarelli M. (1998). *Reliability and applicability of Nakamura's technique using microtremors: an experimental approach.* J. Earthquake Engin., 2, 4, 625-638.

Mucciarelli M. (2008). *Codes, models and reality: reductionism vs. holism in a review of microzonation studies in the Umbria-Marche region*. Ann.Geophys., 51 (2-3), 231-238.

Nakamura Y. (1989). *A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface.* QR Railway Technical Research Institute, 30, 25 - 33.

Nakamura Y. (1996). *Real-time information systems for hazards mitigation.* Proceedings of the XI World Conference on Earthquake Engineering, Acapulco (Mexico).

Nakamura Y. (2000). *Clear identification of fundamental idea of Nakamura's technique and its applications.* Proceedings of the XII World Conference on Earthquake Engineering, Auckland (New Zealand).

Nogoshi M. and Igarashi T. (1970). *On the amplitude characteristics of microtremor (part 1).* Journ. Seism. Soc. Japan, 23, 281-303.

Nogoshi M. and Igarashi T. (1971). *On the amplitude characteristics of microtremor (part 2).* Journ. Seism. Soc. Japan, 24, 26-40.

Ohori M., Nobata A. and Wakamatsu K. (2002). *A comparison of ESAC and FK methods of estimating phase velocity using arbitrarily shaped microtremor arrays.* Bulletin of Seismological Society of America, 92, 6, 2323-2332.

Oil Exploration Directorate, Office of the Prime Minister (1993). *Geological Map of the Maltese Islands: Sheet 1 — Malta; Sheet 2 — Gozo and Comino*. Valletta, Malta.

Okada H. (2003). *The Microtremor Survey Method.* Geophysical Monograph Series, SEG, pp. 129.

Pace S., Panzera F., D'Amico S., Galea P., and Lombardo G. (2011). *Modelling of ambient noise HVSR in a complex geological area- Case study of the Xemxija Bay Area, Malta*. Congress proceedings 30th Convegno Nazionale GNGTS, Trieste, Italy, 299-302.

Panzera F., D'Amico S., Lotteri A., Galea P. and Lombardo G. (2012). *Seismic site response of unstable steep slope using noise measurements: the case study of Xemxija Bay area, Malta*. Natural Hazards and Earth System Science, 12, 3421‐3431.

Panzera F., D'Amico S., Galea P., Lombardo G., Gallipoli M.R. and Pace S. (2013). *Geophysical measurements for site response investigation: preliminary results on the island of Malta*. Bollettino di Geofisica Teorica ed Applicata, 54, 111‐128.

Paolucci E., Albarello D., D'Amico S., Lunedei E., Martelli L., Mucciarelli M., Pileggi D. (2015). *A large scale ambient vibration survey in the area damaged by May-June 2012 seismic sequence in Emilia Romagna, Italy*. Bull. Earthquake Eng., 13(11): 3187-3206, DOI: 10.1007/s10518-015-9767-5.

Park C. B., Miller R. D. and Xia J. (1999). *Multichannel analysis of surface waves.* Geophysics, 64, 800-808.

Parker R. L. (1994). *Geophysical Inverse Theory.* Princeton University Press, New Jersey, pp. 386.

Parolai S., Bormann P. and Milkereit C. (2002). *New relationships between Vs, thickness of sediments and resonance frequency calculated from H/V ratio of seismic noise for the Cologne area.* Bull. Seismol. Soc. Am., 92, 2521–2527.

Parolai S., Richwalski S. M., Milkereit C. and Bormann P. (2004). *Assessment of the stability of H/V spectral ratios from ambient noise and comparison with earthquake data in the Cologne area (Germany).* Tectonophysics, 390, 57-73.

Parolai S., Picozzi M., Richwalski S.M. and Milkereit C. (2005). *Joint inversion of phase velocity dispersion and H/V ratio curves from seismic noise recordings using a genetic algorithm considering higher modes.* Geophys. Res. Lett., 32, L01303, doi: 10.1029/2004GL021115.

Parolai S., Richwalski S. M., Milkereit C. and Fäh D. (2006). *S-wave velocity profiles for earthquake engineering purposes for the Cologne area (Germany)*. Bull. Earth. Engineering, 4, 65-94.

Parolai S. (2012). *Investigation of site response in urban areas by using earthquake data and seismic noise*. In: Bormann, P. (Ed.), New Manual of Seismological Observatory Practice 2 (NMSOP-2), Potsdam: Deutsches GeoForschungsZentrum GFZ, pp. 1—38. DOI: 10.2312/GFZ.NMSOP-2_ch14.

Parolai S. (2014). *Shear wave quality factor Qs profiling using seismic noise data from microarrays*. J Seismol*,* 18, 695‑704, doi:10.1007/s10950-014-9440-5.

Pedley H.M., House M.R., Waugh B. (1976). *The geology of Malta and Gozo*. Proceedings of the Geologists' Association 87, 325–341.

Pedley M., Clarke M. and Galea P. (2002). *Limestone Isles in a Crystal Sea: The geology of the Maltese islands*. San Gwann: Publishers Enterprises Group Ltd.

Pedley M. (2011). *The Calabrian Stage, Pleistocene highstand in Malta: a new marker for unravelling the Late Neogene and Quaternary history of the islands*. Journal of the Geological Society, 168(4), 913-926.

Peruzzi G., Albarello D., Baglione M., D'Intinosante V., Fabbroni P., Pileggi D. (2016). *Assessing 1D litho-stratigraphical amplification factor for microzoning studies in Italy*. Bull. Earthquake Eng. 14:373–389. DOI 10.1007/s10518-015-9841-z.

Peterson J. (1993). *Observations and modeling of seismic background noise*. U.S. Geol. Survey Open-File Report, 93-322, 95 pp.

Picozzi M. (2005). *Joint inversion of phase velocity dispersion and H/V ratio curves from seismic noise recordings.* PhD Thesys, University of Siena (Italy), 170 pp.

Picozzi M., Parolai S. and Albarello D. (2005a). *Statistical analysis of noise Horizontal to Vertical Spectral Ratios (HVSR).* Bull. Seism. Soc. Am., 95, 5, 1779-1786.

Picozzi M., Parolai S. and Richwalski S. M. (2005b). *Joint inversion of H/V ratios and dispersion curves from seismic noise: estimating the S-wave velocity of bedrock.* Geoph. Res. Lett., 32, 11, doi: 10.1029/2005GL022878.

Picozzi M. and Albarello D. (2007). *Combining genetic and linearized algorithms for a two-step joint inversion of Rayleigh wave dispersion and H/V spectral ratio curves*. Geophys. J. Int., 169, 189–200.

Picozzi M., Parolai S., Bindi D., Strollo A. (2009). *Characterization of shallow geology by high-frequency seismic noise tomography*. Geophysical Journal International, 176, 1, 164-174.

Picozzi M., Parolai S. and Bindi D. (2010). *Deblurring of frequency-wavenumber images from small-scale seismic arrays.* Geophys.l J. Internat. doi: 10.1111/j.1365-246X.2009. 04471.x.

Pieri M. and Groppi G. (1981). *Subsurface geological structure of the Po Plain*. Pubbl.414, PF Geodinamica, C.N.R. pp. 23.

Pileggi D., Rossi D., Lunedei E. and Albarello D. (2011). *Seismic characterization of rigid sites in the ITACA database by ambient vibration monitoring and geological surveys.* Bull Earthquake Eng, 9, 1839–1854.

Pileggi D. (2013). *Intensive and extensive application of Ambient Vibration recordings: critical issues in different geological settings.* PhD Thesys, University of Siena (Italy), 164 pp.

Pilz M., Parolai S., Picozzi M. and Bindi D. (2012). *Three-dimensional shear wave velocity imaging by ambient seismic noise tomography*. Geophys. J. Int., 189, 501–512.

Pilz M., Parolai S., and Bindi D. (2013). *Three-dimensional passive imaging of complex seismic fault systems: evidence of surface traces of the Issyk-Ata fault (Kyrgyzstan).* Geophys. J. Int., 194, 1955–1965, doi: 10.1093/gji/ggt214.

Preisendorfer R.W. (1988). *Principal Component Analysis in Meteorology and Oceanography*. C.D. Mobley, ed. Elsevier, 425 pp.

Press F. (1968). *Earth models obtained by Monte Carlo inversion*. J. Geophys. Res., *73*, 5223–5234.

Press W.H., Teukolsky S.A., Vetterling W.T. and Flannery B.P. (1992). *Numerical Recipes in Fortran.* 2nd edn, Cambridge University Press, New York, 963pp.

Priolo E., Romanelli M., Barnaba C., Mucciarelli M., Laurenzano G., Dall'Olio L., Abu Zeid N., Caputo R., Santarato G., Vignola L., Lizza C., Di Bartolomeo P. (2012). *The Ferrara thrust earthquakes of May–June 2012: preliminary site response analysis at the sites of the OGS temporary network*. Ann. Geophys. 55(4):591–597. doi:10.4401/ag-6172.

Puccinelli A., Verani M. and Rossini V. (2000). *Nuovi dati sull'assetto idrogeologico dell'area termale di Montecatini Terme (Pistoia) e loro implicazioni nella pianificazione territoriale.* Quaderni di Geologia Applicata, 7 – 3(2000), pp. 33-48, 7 ff., Pitagora Ed. Bologna.

Regione Emilia Romagna (1999). *Carta geologica di pianura dell'Emilia Romagna 1:250.000*. Edited by Preti D., Regione Emilia Romagna. S.EL.CA., Firenze.

Renalier F., Jongmans D., Campillo M. and Bard Y.P. (2010). *Shear wave velocity imaging of the Avignonet landslide (France) using ambient noise cross correlation*. J. Geophys. Res., 115, F03 032, doi:10.1029/2009JF001538.

RER & ENI-Agip (1998). *Riserve idriche sotterranee della Regione Emilia-Romagna*. Edited by G. M. Di Dio. Regione Emilia-Romagna, ufficio geologico – ENI-Agip, Divisione Esplorazione & Produzione. S.EL.CA., Firenze, pp 120.

Ritzwoller M. H., N. M. Shapiro, M. P. Barmin and A. L. Levshin (2002). *Global surface wave diffraction tomography*. J. Geophys. Res., 107(B12), 2335, doi:10.1029/2002JB001777.

Rodriguez V.H., Midorikawa S. (2002). *Applicability of the H/V spectral ratio of microtremors in assessing site effects on seismic motion*. Earthquake Eng Struct Dyn 31(2): 261–279.

Rost S. and Thomas C. (2002). *Array seismology: methods and applications.* Reviews of Geophysics, 40, 3, 1-27.

Roux P., Sabra K.G., Kuperman W.A., Roux A. (2005). *Ambient noise cross correlation in free space: theoretical approach.* J. Acoust. Soc. Am., 117(1), 79–84.

Safak E. (2001). *Local site effects and dynamic soil behaviour*. Soil Dynamics and Earthquake Engineering, Elsevier Science Ltd., Vol. 21, pp. 453-458**.**

Sakaji K. (1998). *Temporal variation of the power spectra of microtremors observed at soil and rock sites.* Graduation Thesis, Hokkaido University.

Sambridge M. and Drijkoningen G.G. (1992). *Genetic algorithms in seismic waveform inversion.* Geophys. J. Int., 109, 323–342.

Sambridge M. (1999). *Geophysical inversion with a Neighbourhood algorithm, I, Searching a parameter space.* Geophys. J. Int., 138, 479–494.

Sambridge M. and Mosegaard K. (2002). *Monte Carlo methods in geophysical inverse problems.* Review of Geophysics, 40, 3, 1-27.

Sanchez-Sesma F.J. and Campillo M. (2006). *Retrieval of the Green's function from cross correlation: the canonical elastic problem.* Bull. Seismol. Soc. Am. 96(3):1182–1191. doi:10.1785/0120050181.

Sanchez-Sesma F.J., Rodrıguez M., Iturraran-Viveros U., Luzon F., Campillo M., Margerin L., Garcıa Jerez A., Suarez M., Santoyo M.A., Rodrıguez-Castellanos A. (2011). *A theory for microtremor H/V spectral ratio: application for a layered medium.* Geophys J Int 186:221–225.

Scherbaum F., Hinzen K.G. and Ohrnberger M. (2003). *Determination of shallow shear wave velocity profiles in the Cologne, Germany, area using ambient vibrations*. Geophys. J. Int., 152, 597-612.

Schuster G. T. (2001). *Theory of daylight/interferometric imaging: tutorial.* In: Proceedings of 63rd Meeting, European Association of Geoscientists and Engineers, Session: A32 (extended Abstracts).

Schwarz G. (1978). *Estimating the dimension of a model*. Ann. Stat. 6, 461–464.

Scognamiglio L., Margheriti L., Mele F.M., Tinti E., Bono A., De Gori P., Lauciani V., Lucente F. P., Mandiello A.G., Marcocci C., Mazza S., Pintore S., Quintiliani M. (2012). *The 2012 Pianura Padana Emiliana seismic sequence: locations, moment tensors and magnitudes*. Annals of geophysics, 55, 4, 549-559, doi: 10.4401/ag-6159.

Seed H.B. and Idriss I. M. (1969). *Influence of soil conditions on ground motion during earthquakes*. Journal of soil mechanics and foundation division, ASCE, vol. 95, N. SM1.

Seed R. B., Dickenson S. E, Riemer M. F., Bray J. D., Sitar N., Mitchell J. K., Idriss I. M., Kayen R. E., Kropp A., Harder L. F. and Power M. S. (1990). *Preliminary report on the principal geotechnical aspects of the October 17, 1989 Loma Prieta earthquake*. Report No. UCB/EERC-90/05, Earthquake Engineering Research Center, University of California, Berkeley, 137 pp.

Sen M.K. and Stoffa P.L. (1992). *Rapid sampling of model space using genetic algorithms.* Geophys. J. Int., 108, 281–292.

Shapiro N. M. and Campillo M. (2004). *Emergence of broadband Rayleigh waves from correlations of ambient seismic noise*. Geophys. Res. Lett., 31, L07614, doi10.1029/2004GL019491.

Shapiro N. M., Campillo M., Stehly L. and Ritzwoller M. (2005). *High resolution surface wave tomography from ambient seismic noise*. Science, 307, 1615–1618.

Site Effects Assessment using Ambient Excitations (SESAME) European project (2004a). *Nature of noise wavefield.* Final Report WP08.

Site Effects Assessment using Ambient Excitations (SESAME) European project (2004b). *Guidelines for the implementation of the H/V spectral ratio technique on ambient vibrations: measurements, processing and interpretation*. Deliverable D23.12.

SM Working Group (2015). *Guidelines for Seismic Microzonation*. Conference of Regions and Autonomous Provinces of Italy - Civil Protection Department, Rome, available at http://www.protezionecivile.gov.it/httpdocs/cms/attach_extra/GuidelinesForSeismicMicrozonation.pdf?.

Snieder R. (2004). *Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase*. Phys. Rev. E, 69. doi:10.1103/PhysRevE.69.046610.

Socco L.V., Foti S. and Boiero D. (2010). *Surface-wave analysis for building near-surface velocity models - Established approaches and new perspectives*. Geophysics, 75 (5), 75A83-75A102. doi:10.1190/1.3479491.

Stockwell R. G., L. Mansinha and R. P. Lowe (1996). *Localization of the complex spectrum: The S transform*. IEEE Transactions on Signal Processing, 44(4), 998‑1001.

Stoffa P.L. and Sen M.K. (1991). *Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane wave seismograms.* Geophys, 56, 1794–1810.

Stone W.C., Yokel F.Y., Celebi M., Hanks T and Leyendecker E.V. (1987). *Engineering aspects of the September 19, 1985 Mexico earthquake*. NBS Building Science Series 165, National Bureau of Standards, Washington, D.C., 207 pp.

Takahashi R. and Hirano K. (1941). *Seismic vibrations of soft ground* (in Japanese). Bulletin of the Earthquake Research Institute, 19, 534–543.

Tarantola A. (2005). *Inverse problem theory and methods for model parameter estimation.* SIAM, Philadelphia, pp. 342.

Tokimatsu K., Tamura S. and Kojima H. (1992a). *Effects of multiple modes on Rayleigh Wave dispersion characteristics*. Journal of Geotechnical Engineering, 118, 10, 1529-1543.

Tokimatsu K., Shinzawa K. and Kuwayama S. (1992b). *Use of short-period microtremors for Vs profiling.* Journal of Geotechnical Engineering, 118, 10, 1544-1558.

Tokimatsu K. (1997). *Geotechnical site characterization using surface waves.* Earthquake Geotechnical Engin., ed. Ishihara, Balkema, Rotterdam, 1333-1368.

Toksöz M. N. (1964). *Microseisms and an attempted application to exploration.* Geophysics, 29-2, 154-177.

Toksöz M. N. and Lacoss R. T. (1968). *Microseisms – Mode structure and sources.* Science, 159, 872-873.

Toscani G., Burrato P., Di Bucci D., Seno S., Valensise G. (2009). *Plio-Quaternary tectonic evolution of the northern Apennines thrust fronts (Bologna-Ferrara section, Italy): seismotectonic implications*. Boll. Soc. Geol. Ital. (Italian Journal of Geosciences), 128, 605-613, doi:10.3301.IJG2009.128.2.605.

Trevisan L. (1954). *La nuova sorgente Leopoldina di Montecatini Terme e le condizioni geologiche del sottosuolo.* Boll. Ingegneri Firenze II, n° 8-9, Florence.

Ullah S., Bindi D., Pittore M., Pilz M., Orunbaev S., Moldobekov B., Parolai S. (2013). *Improving the spatial resolution of ground motion variability using earthquake and seismic noise data: the example of Bishkek (Kyrgyzstan)*. Bull. Earthquake Eng., 11:385–399. DOI 10.1007/s10518-012-9401-8.

Vella A., Galea P. and D'Amico S. (2013). *Site frequency response characterisation of the Maltese islands based on ambient noise H/V ratio.* Engineering Geology, 163, 89-100.

Wang R. (1999). *A simple orthonormalization method for stable and efficient computation of Green's functions.* Bull. Seism. Soc. Am. 89, 733–741.

Wathelet M., Jongmans D., Ohrnberger M. (2004). *Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements.* Near Surface Geophys 2:211–221.

Wathelet M. (2005). *Array recordings of ambient vibrations: surface wave inversion.* PhD Thesis, University of Liege (Belgium).

Wathelet M., Jongmans D. and Ohrnberger M. (2005). *Direct Inversion of Spatial Autocorrelation Curves with the Neighborhood Algorithm.* Bulletin of the Seismological Society of America, 95, 5, 1787-1800.

Wathelet M. (2008). *An improved neighborhood algorithm: Parameter conditions and dynamic scaling.* Geophysical Research Letters, 35, L09301, doi:10.1029/2008GL033256, 2008.

Weaver R. and Lobkis O. (2001). *On the emergence of the Green's function in the correlations of a diffuse field.* J. acoust. Soc. Am., 109, 2410.

Weaver R.L. and Lobkis O. (2004). *Diffuse fields in open systems and the emergence of the Green's function.* J. acoust. Soc. Am., 116, 2731–2734.

Wilks D.S. (2006). *Statistical methods in the atmospheric sciences.* Second edition; Academic Press, 630 pp., ISBN 13: 978-0-12-751966-1.

Woods J.W. and Lintz P.R. (1973). *Plane waves at small arrays.* Geophysics, 38, 6, 1023-1041.

Xia J., Miller R.D. and Park C.B. (1999). *Estimation of near surface shear wave velocity by inversion of Rayleigh waves.* Geophysics, 64, 3, 691-700.

Xia J., Miller R.D., Park C.B., and Tian G. (2003). *Inversion of high frequency surface waves with fundamental and higher modes*. Journal of Applied Geophysics, 52, 45-57.

Yamamoto H. (2000). *Estimation of shallow S-wave velocity structures from phase velocities of Love and Rayleigh waves in microtremors.* Proceedings of the XII World Conference on Earthquake Engineering, Auckland (New Zealand).

Yamanaka H., Dravinski M. and Kagami H. (1993). *Continuous measurements of microtremors on sediments and basements in Los Angeles, California.* Bulletin Seismological Society of America, 83, 5, 1595-1609.

Yamanaka H., Takemura M., Ishida H. and Niwa M. (1994). *Characteristics of long-period microtremors and their applicability in exploration of deep sedimentary layers.* Bulletin of the Seismological Society of America, 84, 6, 1831-1841.

Yamanaka H. and Ishida H. (1996). *Application of Genetic Algorithms to an inversion of surface-wave dispersion data.* Bull. Seism. Soc. America, 86, 2, 436-444.

Yang Y., Ritzwoller M.H., Levshin A.L. and Shapiro N.M. (2007). *Ambient noise Rayleigh wave tomography across Europe*. Geophys. J. Int., 168, 259–274, doi:10.1111/j.1365–246X.2006.03203.x.

Yao H., Van Der Hilst R.D. and De Hoop M.V. (2006). *Surface-wave array tomography in SE Tibet from ambient seismic noise and two station analysis: I - phase velocity maps*. Geophys. J. Int., 166, 732– 744.

Yokoi T. and Margaryan S. (2008). *Consistency of the spatial autocorrelation method with seismic interferometry and its consequence*. Geophys. Prosp., 56, 435–451.

Young Ch.J., Chael E.P., Withers M.W., and Aster R.C. (1996). *A comparison of the high- frequency (>1Hz) surface and subsurface noise environment at three sites in the United States*. Bull. Seism. Soc. Am., 86, 5, 1516-1528.

Zammit-Maempel G. (1977). *An Outline of Maltese Geology*. Progress Press, Malta

Zhang B. and Lu L. (2003). *Rayleigh wave and detection of low-velocity layers in a stratified half-space*. Acoustical Physics, 49, 516–528.

Zhang S.H., Chan L.S. and Xia J. (2004). *The selection of field acquisition parameters for dispersion images from multichannel surface wave data*. Pure Appl. Geophys., 161, 185–201