UNIVERSITÀ DI SIENA 1240

Dipartimento di Scienze Fisiche, della Terra e dell'Ambiente

Dottorato in Scienze e Tecnologie Ambientali, Geologiche e Polari

38° Ciclo

Coordinatrice: Prof.ssa Letizia Marsili

Small-scale Ambient Noise Tomography and rapid Vs estimation for near-surface seismic characterization in complex environments: the case of Nirano Mud Volcano system

Settore scientifico disciplinare: GEO/10

Candidato
Nicolò Carfagna
DSFTA – Università di Siena

Supervisore
Prof. Dario Albarello
DSFTA – Università di Siena

Anno accademico di conseguimento del titolo di Dottore di ricerca
2025/2026

Università degli Studi di Siena
Dottorato in Dottorato in Scienze e Tecnologie Ambientali, Geologiche e Polari
38° Ciclo

Data dell'esame finale
16/06/2026

Commissione giudicatrice
Matteo Picozzi
Andrea Zunino
Marcello Viti

Supplenti
_____
_____

Table of Contents

ABSTRACT
Chapter I. Introduction 1
Chapter II. Ambient seismic noise 11
2.1 Origin of ambient noise 12
2.2 Ambient noise as a stochastic and stationary process 14
2.3 Surface waves 18
2.3.1 Rayleigh and Love waves 22
2.3.2 Phase and group velocity 24
Chapter III. Dispersion curves from seismic noise 27
3.1 Introduction to surface waves methods 28
3.2 Correlation of Ambient seismic noise 33
3.2.1 Spatial Auto-Correlation methods 34
3.2.2 Seismic Interferometry 39
3.3 SPAC and SI in a two-station application 46
Chapter IV. Inverse problems in surface-wave analysis 53
4.1 Introduction to inverse problem theory 54
4.1.1 Properties and characteristics of inverse problems 57
4.1.2 Misfit function and optimization 59
4.1.3 Regularization and prior information 64
4.2 A linear inverse problem: seismic tomography 68
4.2.1 Forward model and ray theory 70
4.2.2 Model solution for seismic tomography 74
4.3 A nonlinear inverse problem: from Rayleigh dispersion curve to Vs model 76
4.3.1 Forward model 78
4.3.2 Dispersion curve inversion 79
Chapter V. Codes Implementation 83
5.1 Dispersion curves extraction 84
5.1.1 Phase curves 85
5.1.2 Group curves 89
5.1.3 An automatic picking algorithm 92

5.1.4 Validation test for dispersion curves extraction 96
5.2 Seismic Tomography 101
5.2.1 Synthetic test 106
5.3 Direct dispersion curve inversion for shear wave velocity profile 112
Chapter VI. Case study 119
6.1 Context 120
6.2 Data acquisition 127
6.3 Data processing 128
6.3.1 Estimation of Rayleigh wave dispersion curves 129
6.3.2 Phase velocity maps through SWT 134
6.3.3 Local inversion dispersion curve inversion for 3D Vs model 144
6.4 Models Interpretation 155
Chapter VII. Conclusions 161
References 167
Appendix A 179

ABSTRACT

This thesis presents the development, implementation, and application of a novel methodology for near-surface seismic characterization based exclusively on the analysis of ambient seismic noise. The work focuses on the extraction of dispersive information about surface waves (Rayleigh waves) using minimal acquisition configurations, with particular attention to two-station approaches. The main objective is to evaluate the reliability of such method in complex and heterogeneous contexts, where the use of extended arrays or active seismic sources can be impractical.

After a discussion of the theoretical foundations of surface wave-based tomographic strategies and the main assumptions and limitations of ambient noise-based methods, the thesis describes the main codes and algorithms developed during the work. These include a partially automatic program for extracting phase dispersion curves from pairs of stations, which includes an innovative picking algorithm based on the joint use of group and phase velocity curves, capable of overcoming the ambiguities typical of two-station methods. A linear 2D travel-time tomography inversion approach, implemented in a single step while simultaneously considering all analysed frequencies, is also illustrated, along with a novel dispersion curve inversion algorithm designed to efficiently and quickly handle a large number of curves, based on a frequency-depth link to produce the subsoil profile.

The different strategies of the methodology were validated through synthetic tests and subsequently applied to a real case study represented by the Salse di Nirano area (Northern Apennines, Italy), a site characterized by the presence of active sedimentary volcanism. The area, already the subject of previous geological and geophysical investigations, is characterized by the presence of lateral heterogeneities even over small distances, thus it constitutes a significant testbed in order to evaluate the reliability of simplified acquisition configurations and station-pair approaches. The results show that significant phase dispersion curves can be obtained even from short-duration records and closely spaced station pairs.

The final 2D and 3D shear wave (Vs) velocity models show significant lateral and vertical variations even at small scales, consistent with a complex structural arrangement dominated by the presence of preferential pathways for fluid and gas migration within the investigated site.

Overall, despite the limitations of basic two-station methods, the approximation of the adopted tomographic inversion schemes, and the extreme simplicity of the implemented algorithm for dispersion curve inversion, the proposed approach proved robust and effective, returning physically plausible and interpretable results. The developed procedure represents a flexible and rapid tool for passive seismic investigations in constrained operational contexts and is potentially applicable to numerous scenarios, both in urban and natural contexts.

Chapter I.
Introduction

Surface waves are extensively used in seismology since their propagation along the Earth's surface is directly sensitive to the elastic properties of the underlying subsoil (Aki & Richards, 1980). This ability makes them particularly suitable in several seismic application at multiple spatial scales, as the Earth's surface is continuously crossed by surface waves generated by both natural and anthropogenic sources.

This wide range of applications arises essentially from their geometrical dispersion behaviour (expressed by the dispersion curve that link frequency and surface wave velocity), that makes different frequencies (or periods) sensitive to different portions of the subsoil, with higher frequencies sampling the uppermost subsoil and lower frequencies penetrating deeper. This peculiar property is the basis for basically all surface-wave methods (SWMs), that have undergone extremely rapid evolution thanks to the non-invasive and highly informative nature for subsurface characterization. Early quantitative surface wave applications were developed exploiting long-period Rayleigh and Love waves (tens to hundreds of seconds) generated by teleseismic earthquakes and recorded at multiple stations to infer large-scale Earth velocity models (e.g., Brune et al. 1961; Toksöz & Anderson, 1966; Nakanishi & Anderson, 1982; Dziewonski, 1984). At the same decades, several studies (Gutenberg, 1958; Aki, 1965) focused on the analysis of the ambient vibrations (or ambient noise), showing that surface waves contained within ambient vibrations could be used to retrieve information on subsurface properties without the need for localized sources.

Their use has been also extended at higher frequency bands (usually >1Hz) for near-surface applications, such as engineering investigations at the scale of buildings, urban hazard mapping, and seismic site characterization applied at many different geological contexts (Garofalo et al., 2016; Foti et al., 2018 and references therein). Common workflow of surface wave-based methods consists essentially of three steps: field data acquisition (through controlled sources or ambient noise records), extraction of experimental dispersion information through commonly adopted methods (this topic will be discussed in Chapter III), and some kind of inversion to retrieve spatial distribution of main elastic parameters (Vp, Vs, ρ) or attenuation-related parameters (e.g. quality factors). At regional scale, surface waves are particularly used to map crustal heterogeneities over very long distances, while in near-surface applications are commonly used to retrieve 1D vertical variations of seismic velocities.

Focusing on the latter, propagation of surface waves in 1D vertically stratified medium is theoretically well understood: each frequency component of surface waves not only samples different depth ranges, with longer wavelengths penetrating deeper into the subsurface, but can also propagate with different velocities, making surface-wave propagation basically a multimodal phenomenon (Thomson, 1950; Haskell, 1953). As a result, for a given frequency multiple velocities may exist, each associated with a different mode of propagation. These modes can coexist and contribute simultaneously to the recorded wavefield. The relative energy carried by the different propagation modes is not constant but depends on the subsurface stratigraphy and the depth and nature of the source. Although the fundamental mode is often the most energetic over a wide frequency range, higher modes may become significant or even dominant in presence of complex layering or velocity inversions through the vertical profile, producing a weighted experimental dispersion curve produced by the combination of the fundamental and higher modes, an 'effective curve' (Tokimatsu et al., 1992). Ignoring these

effects can therefore lead to incorrect interpretations of the experimental dispersion curves. Within an elastic medium, the propagation of surface waves is predominantly controlled by shear-wave velocity (Vs) variations; this marked sensitivity to Vs is the main reason for using surface wave methods in the reconstruction of subsoil Vs models and provides the basis for extending the analysis from purely vertical stratification to the study of lateral heterogeneity of elastic properties.

The literature of recent decades has progressively shifted attention from purely 1D schemes to pseudo-2D or 3D imaging strategies (Socco et al., 2010), often achieved by combining 1D inversions along multiple profiles or via explicit imaging methods.

Indeed, although both active and passive methods used to infer Vs profiles rely on well-established techniques (e.g., Aki, 1957; Nazarian & Stokoe, 1984; Park et al., 1999; Ohori et al., 2002), near-surface applications are inherently limited by the assumption of a 1D subsoil geometry, which neglects the presence of possible lateral variations in the subsoil.

When lateral heterogeneities become significant, however, the propagation of surface waves is affected not only by the vertical stratification, but also by horizontal variations in elastic properties, with visible effects on the phase, apparent velocity and modal content of the curves representing the surface propagation. Consequently, the interpretation of dispersion curves as the purely expression of a local 1D configuration may be incomplete or misleading, and methods based on this assumption can easily fail. To overcome these resolution limitations in the horizontal directions, surface wave tomography (SWT) has received increasing attention over the past decades, especially the development of Ambient Noise Tomography, hereafter ANT (Shapiro et al., 2005; Yao et al., 2006; Bensen et al., 2007; Lin et al., 2008). ANT takes advantage of the continuous ambient vibrations propagating at the top of the Earth's crust, which are largely dominated by surface waves and doesn't require for localized sources such as earthquakes or active shot.

Two-station dispersion technique (Ikeda et al., 2021; Magrini et al., 2020) is surely the core of most ANT methods to retrieve phase or group velocity maps, and it is essentially based on the correlation of seismic noise recorded at two different points. In common strategies, surface wave propagation is approximated as a process within a simple 2D domain (the Earth's surface), allowing lateral variations in propagation velocity to be explicitly considered. Seismic tomography through ANT has been shown to be particularly effective in recovering even significant lateral variations, providing high-resolution velocity models when the stations spatial coverage is adequate.

These imaging approaches were formulated based on the conversion of phase (or group) velocity measurements to travel-time observations: in these implementations SWT is often a 'two-step' procedure. First, dispersion information is extracted from pairs of receivers, providing frequency-dependent phase velocity estimates. These measurements are then inverted frequency by frequency within a 2D tomographic domain to obtain lateral velocity maps, which can subsequently be combined as local dispersion curves to reconstruct quasi-3D models of the parameter of interest through simple vertical inversion.

More recently, alternative approaches have been proposed that aim to directly invert surface wave observations for 2D or 3D shear wave velocity models, bypassing the intermediate tomographic phase (e.g., Boschi &

Ekström, 2002). Mohammadi et al. (2020) proposed STP method based on the simultaneous inversion of all sampled frequencies in order to account also for the correlation over the analysed frequency range.

Although most of mentioned studies have demonstrated the effectiveness of SWT in practice providing high resolution velocity maps, extending the general 1D interpretation of surface wave propagation to laterally heterogeneous media requires a broader conceptualization, since propagation in a non-horizontally stratified medium alter significantly dispersive and modal behaviour of the surface waves, changing the nature of the 1D original problem. In simple stratified media, surface wave propagation can be described in terms of well-defined propagation modes, each characterised by a specific dispersion relation, allowing for a largely analytical treatment of the problem. When lateral heterogeneities are introduced, this modal description is no longer strictly valid, as the propagation modes become spatially variable and can exchange energy along the propagation path. Consequently, surface wave propagation cannot be described by a simple set of modes, and the observed dispersion becomes dependent on the propagation path. In this context, the study of surface wave propagation in laterally heterogeneous media relies on approximate theoretical descriptions or numerical modelling approaches (Keilis-Borok & Levshin, 1989). When lateral heterogeneities are considered, the elastic properties of the medium, expressed in terms of Lamé parameters and density, are no longer function of depth alone. Accordingly, they vary depending on all the spatial coordinates x, y, and z:

$$\lambda = \lambda(x, y, z), \quad \mu = \mu(x, y, z), \quad \rho = \rho(x, y, z).$$

This apparent minor generalization has deep implications in the mathematical treatment of surface wave propagation, as it precludes the separation of the problem into independent vertical and horizontal components. Even small local variations of the medium properties (density, Lamé moduli or topography) results into changes in the propagation behaviour of surface wave: phase perturbations of the signal, changes in wave amplitudes (therefore also connected to their potential depth penetration) and also the 'mode coupling' phenomena between different modes, which therefore might lead the fundamental mode to be confuse with the energies of higher modes, becoming indistinguishable and increasing differences between fundamental curve and effective dispersion curve.

These changes in the behaviour of surface waves as response of lateral heterogeneities along the propagation path constitute the physical foundation that makes such waves sensitive to lateral structure. Wavefield perturbation can be in this way interpreted as the response of the system to a local change in elastic properties and can therefore be exploited for tomographic purposes.

In a small perturbation regime of the medium properties, propagation can be treated as a linear system where the wavefield variation is proportional to the local perturbations in the elastic properties, such as seismic velocity. In this context, weak lateral heterogeneities affect the wave propagation producing only small phase perturbations (forward or backward shift) proportional to the properties variation with no significant changes in the propagation direction (refraction and reflection can be avoided). These perturbations can be seen as the cumulative effects of the local variations along the propagation path and can be expressed in the form of

sensitivity integrals that quantify the contribution of different portions of the medium to the observed variation. This relationship between phase perturbations and elastic properties allows phase measurements to be treated as integrals (or sums in a discrete domain) of the phase velocities encountered by the wave along the propagation trajectory, formally justifying the use of ray tomography (Woodhouse, 1974; Woodhouse & Dziewonski, 1984).

This physically meaningful approximation therefore consists in assuming that lateral variations in elastic properties are weak, in the sense that they occur on spatial scales much larger than the dominant wavelengths of propagating surface waves (high-frequency approximation). Under this assumption, the wavefield can be described in terms of locally defined modal surface waves, which evolve smoothly along the propagation path. Although this approximation does not allow a completely exact solution, it provides a conceptual framework for interpreting observations of surface waves in laterally heterogeneous media. Applied over large spatial extents where lateral variations in elastic properties therefore occur over dimensions to tens up to hundreds of kilometres, the wavelengths of surface waves are much smaller than the scale of expected medium heterogeneities, allowing the geometric ray theory to be applied with good approximation. When lateral variations of the medium are gradual with respect to the dominant wavelength, surface waves propagate with no evident scattering processes, and the modes remain substantially decoupled and well distinguishable (Marquering et al., 1996). In this regime, the reconstruction of lateral variations is reliable and correctly represented by classical tomographic-based inversion.

Propagation in this context must be seen as an evolutionary process, in which phase and energy progressively accumulate as the wave propagates through regions characterized by different elastic parameters. Consequently, dispersion and locally observed curves represent path-dependent effective quantities that reflect the integrated response of the medium based on contributions along the propagation path rather than the properties of a single local structure.

On the contrary, in the presence of more marked or rapid heterogeneities on wavelength-comparable scales, more complex effects, such as mode coupling and multiple scattering (both forward and backward scattering) emerge instead, which may require lower frequencies, the inclusion of higher modes in the dispersion analysis, or advanced propagation models based on numerical simulations (Maupin, 1988; Friederich et al., 1993). Basically, under these conditions, account only for the fundamental mode to reconstruct the lateral structure becomes challenging and may also produce errors in the estimation of surface wave velocities. Moreover, in such contexts, experimental dispersion curve is barely influenced only by the fundamental mode, and it must be interpreted as effective curve conditioned also by the higher modes.

Many standard surface wave-based approaches, while providing stable and reliable results, produce models that clearly represent spatial averages of subsurface properties. This is particularly true when dispersion curves are extracted over extended spatial windows or when lateral homogeneity of the model is implicitly assumed. In laterally complex environments, such averaging can obscure localized variations and mask structurally significant features.

Overall, the effectiveness of SWT relies on phase variations along the propagation path, which are directly sensitive to the local elastic properties of the medium. Propagation and scattering theory therefore can be assumed valid and the observations, in terms of phase velocity or phase travel-times, can be interpreted, with increasing levels of complexity depending on the scale of heterogeneity and the propagation regime considered.

In recent years, attention has gradually shifted to the near-surface application of dispersion-based tomographic methods. When moving down to local spatial extensions (few hundred metres or less) some of the assumptions made so far (Woodhouse, 1974) at regional and global scales may no longer be satisfied. At such small spatial scales, the identification of buried bodies and lateral variations in subsurface properties is of great interest for several different fields, but at the same time is more challenging in presence of strong heterogeneity and scale-dependent propagation effects. In these contexts, dominant wavelengths of surface waves can be comparable to the scale of subsurface heterogeneities, as a result, wavefront might no longer be effectively described as simple rays. On the contrary, scattering effects become dominant, with energy deviations from the wave path and therefore with non-unique trajectories. In such contexts lateral and vertical scattering can produce significant variations in phase and amplitude and the mode coupling can become predominant, making classical travel-time-based interpretation unsuitable.

As for large-scale investigations, reliability of approximated surface wave propagation at local scale depends on the graduality of subsoil property variations. When heterogeneities evolve uniformly over distances greater than the dominant wavelength, propagation assumptions made before remaining valid, even at higher frequencies (>1 Hz). Conversely, sharp lateral variations, as often occur in near-surface contexts, becomes significant. Near-field effects, such as multiple reflections, local variations in surface topography or layer stiffness can significantly modify phase and waveform within a few, covering coherent parts of the signal that are essential for imaging purpose.

In summary, what distinguishes the small scale from the large scale in surface wave tomography is not simply the size of the studied area, but a set of physical constraints that makes hypotheses of ray theory to fail.

The potential of local-scale tomographic-like approach using surface waves is well illustrated also by Barone et al. (2021), who studied a laterally heterogeneous urban waste deposit using a two-step approach combining local dispersion curve retrieved from active seismic surveys, pseudo-2D phase velocity sections, and subsequent vertical inversion to obtain site-specific Vs profiles. Although they do not produce a true 3D model, such strategies allow for consistent delineation of complex near-surface structures (within 10m depth) in small-scale contexts. A similar strategy, based on dispersion analysis through active sources, is adopted by Papadopoulou et al. (2021), demonstrating that high-resolution imaging at local scale can be achieved applying multimodal tomography, therefore considering also energy relative to higher modes.

In such small-scale context, even the use of ambient noise represents an advantage, as it simplifies the data acquisition phase and extends the investigable area without the need for controlled sources. Several studies based on the analysis of ambient seismic noise (Picozzi et al., 2009; Pilz et al., 2012; Szanyi et al., 2016) have

shown that, over small investigation spatial extent, valuable results can be achieved in terms of local phase velocity maps in presence of weak and gradual variations in the stratigraphic configuration of the subsurface. However, the application of ANT methods in near-surface contexts requires more strict constraints than regional or global applications. Unlike large-scale crustal investigations, where the spatial distribution of low-frequency noise sources is relatively stable and dominated by distant sources, near-surface applications are strongly influenced by local anthropogenic sources, which generate spatially and temporally variable high-frequency contributions. This leads to greater heterogeneity in the recorded noise field, with some stations being strongly influenced by local sources and others being poorly influenced. Moreover, unlike large-scale applications, near-surface surveys typically lack permanent stations, long ambient noise records, or reliable reference models useful to constrain dispersion measurements and velocity models. In ANT applications, dispersion curves provided by two-station approaches suffer from inherent ambiguity in the choice of the right phase dispersion curves. On a regional or global scale, these difficulties are often overcome by referencing the dispersion curves selection to a target curve (e.g., Verbeke et al., 2012; Magrini et al., 2022) or by working on group velocities rather than phase velocities (Lin et al., 2007). However, the use of group velocities implies a substantial loss of resolution, as they are less sensitive to structural details than phase velocities. In the near-surface, such strategies are usually difficult to apply, as there is often a lack of prior information useful for constraining the dispersion curves, and the assumption of a reference model is difficult precisely because in contexts with expected lateral variations, standard approaches fail when assuming a 1D model.

A further step forward must also be taken. Indeed, regardless of the employed acquisition geometry (linear or 2D), the type of source used (active or passive), and the specific aim of the survey, most studies based on surface waves converge towards the reconstruction of subsoil models generally expressed in terms of body wave velocities. Vp or Vs models (1D, 2D or 3D) are indeed more directly accountable from a stratigraphic and structural point of view than phase velocity information alone. This step necessarily involves into strongly non-linear inversion procedure, which is a key step in the entire workflow.

Several inversion algorithms have been devoted to handle with the strong non-linearity relative to surface wave dispersion curve inversion. Standard approaches include Simulated Annealing (SA, Kirkpatrick et al., 1983), Genetic Algorithms (GA, e.g., Sambridge & Drijkoningen, 1992) and Neighborhood Algorithm (NA, Sambridge, 1999). These methods, given the strong non-linearity between data and models, rely basically on stochastic sampling of the parameter space to explore multiple suitable solutions. While these methods are effective in addressing non-uniqueness of the solution, the extensive exploration of the parameter space inevitably require a significant computational cost. When the goal is to reconstruct a 3D Vs model (or Vp) inversion must be performed on tens or even hundreds of local dispersion curves, and in the absence of optimal priori constraints, the computational effort for such a large number of inversions can become prohibitive, limiting the applicability of the approach in real-world applications. As an alternative to direct reconstruction of Vs profiles, several authors have proposed approaches based on the use of averaged representative parameters, such as the Vs_z (time-averaged shear-wave velocity up to the depth z). These parameters are clearly less informative but also less sensitive to inversion instability and non-uniqueness, while they still provide

meaningful information on lateral variations in the subsurface (e.g., Socco et al., 2017; Socco & Comina, 2017). Although these average parameters might represent an acceptable compromise between stability and interpretability, particularly in heterogeneous lateral contexts near the surface, they do not capture the full complexity of the medium.

This thesis develops within the concepts discussed in this Introduction. It is dedicated to the implementation and application of a small-scale ANT based on the dispersion analysis of Rayleigh waves. The approach adopted is based on two-stations method, through which Rayleigh phase dispersion curves are extracted. To reduce phase ambiguities in the selection of the correct curve, an automatic picking procedure based on the use of experimental group velocity curve as a reference is adopted, allowing the choice of the true phase curve to be robustly constrained.

The phase velocities thus obtained are interpreted as effective observed travel-time. These measurements are then inverted using a linear tomographic procedure, providing two-dimensional phase velocity maps over the entire frequency range of interest. The inversion is performed by explicitly considering a frequency correlation term, avoiding independent treatment of individual frequencies. The local dispersion curves obtained from the tomographic maps are finally inverted to obtain local shear wave velocity (Vs) profiles, allowing the reconstruction of 3D models of the subsoil.

Dispersion curves inversion provided Vs profiles through a dedicated strategy based on the association between frequency and sensitivity depth of surface waves. This new and simple approach led to stable and effective solutions in terms of Vs models, ensuring good consistency between experimental and theoretical dispersion curves, with low computational costs. The automatic picking procedure for phase dispersion curves and the rapid inversion algorithm for Vs profiles are two novel methodological elements introduced in the study, providing reliable results even from relatively short recording times, in the order of about one hour.

The proposed methodology is applied to an active mud volcano area (Salse di Nirano reserve), where vertical and lateral heterogeneity in terms of seismic velocity are expected, related to the fluid dynamics and outgassing processes clearly evident from their surface manifestations. Field acquisitions include two 2D seismic arrays, consisting of two parallel geophone lines, and one L-shaped array, with branches of approximately 100 m each. For the first two arrays, 3D Vs velocity models were reconstructed, highlighting the heterogeneous distribution of surface velocities and the presence of restricted anomalies potentially associated with fluid activity. The third array provided two shear wave velocity sections covering a significant portion of the most active area of the site, highlighting further anomalies attributable to the possible presence of surface degassing conduits beneath one of the main surface vents.

This thesis is structured in six Chapters:

Chapter II provides a general overview of the spectral and statistical characteristics of ambient seismic noise. Ambient noise is described as a stochastic and stationary process, with particular attention focused on the origin and distribution of noise sources and to the surface wave content of the noise wavefield.

Chapter III review the fundamental methods developed to analyse the dispersive behaviour of surface waves, focusing on their application for subsurface characterization. Special attention is given to spatial autocorrelation and seismic interferometry techniques, as they form the core of the procedure adopted in this thesis. The chapter includes a concise theoretical formulation for both approaches, building on the noise field characteristics introduced in Chapter 1. **Chapter IV** is devoted to the fundamental principles of inverse problem theory, considering main characteristics and the methodologies commonly adopted for their solution, with particular attention to the stability and reliability of the procedures. Two types of inverse problems are analysed: first, tomographic inversion, representative of a classic linear inverse problem, and on the other hand, dispersion curve inversion for shear wave (Vs) velocity profiles, which instead constitutes a strongly nonlinear problem. The conceptual and operational differences between the two approaches are discussed, highlighting their implications for resolution, stability, and interpretation of the results.

Chapter V describes the main implemented algorithms for this work, starting from the phase dispersion curve extraction procedures, with particular attention to automatic picking, up to the definition of the parameters used in the linear tomographic inversion for phase velocity maps. The chapter also presents the inversion strategy adopted for recovery Vs profiles, optimized to reduce the computational cost. The three described procedures constitute the core of the entire workflow and prepare for the interpretation of the results illustrated in the next chapter.

Finally, **Chapter VI** presents an example of the application of the entire workflow described in Chapter 4 to a real case study, within a geological context characterized by expected marked near-surface lateral heterogeneities, made particularly evident by the surface volcanic manifestations within the investigated area. The obtained results relative to all three phases of the work are illustrated, accompanied by an evaluation of their reliability, stability and geological coherence. Final outcomes, in terms of 2D and 3D Vs models, are then discussed and interpreted within the geological frame of the study area, and compared with results from previous investigations, giving emphasis to consistency of the proposed approach with existing observations.

Chapter II.
Ambient seismic noise

2.1 Origin of ambient noise

Even in absence of earthquakes, the Earth's crust is continuously crossed by ubiquitous vibrations propagating in all directions along the free surface, typically with amplitudes ranging between 10^{-4} and 10^{-2} mm (Okada, 2003). These persistent vibrations, commonly referred to as ambient seismic noise (or background seismic noise, or microseisms) constitute an omnipresent and almost random ambient wavefield not strictly related to discrete seismic events occurring within the Earth's crust, generated by multiple sources of diverse nature. Ambient noise is nowadays exploited in several seismology fields, from the near-surface characterization and evaluation of local site seismic response, to monitoring volcanic and slope instability contexts, to large-scale applications for crustal high-resolution imaging. Yet, despite its proven value, we still call it "noise", a throwback to when these vibrations were commonly neglected from the recorded wavefield, since seismologist were interested in signals generated by localized sources, such as earthquakes.

It is nowadays widely accepted that this ambient noise wavefield originates from several kinds of sources, broadly classified as natural or anthropogenic. Natural sources primarily include atmospheric phenomena, wind-induced vibrations, and interactions between ocean waves and coastal lines typically located at greater distances from seismic recording sites and resulting in a relatively stable amplitude pattern.

Conversely, anthropogenic sources are essentially related to human activities such as industrial machinery and vehicular traffic that generally emit signals that are recorded only by closer stations and exhibit distinct diurnal and weekly variations directly correlated with local human activities (Bonnefoy-Claudet et al., 2006).

These signals, whether from natural processes or human activities, propagate just like conventional seismic waves and suffer the same distance-dependent attenuation. However, unlike localized sources, such as earthquakes or controlled shots, their exact location is unknown, indeed, their exact location is irrelevant when analysing ambient noise, since it represents a complex wavefield produced by the superposition of multiple propagating wave coming from multiple directions (Aki, 1965; Asten, 1983) and generated by sources all over the world for which the exact location is not under study.

Although these ongoing ambient vibrations were observed and studied as early as the late 19th century, when for example Bertelli in 1872 suggested that the long-period oscillations of his pendulum were extremely correlated with the air pressure variations in the atmosphere, detailed understanding remained elusive until the 20th century, when advances in sensitive instruments, denser station networks, and quantitative data-processing techniques transformed studies from descriptive observations into quantitative analyses.

Numerous studies on seismic noise have been conducted since that time. Gutenberg's pioneering reviews (1911, 1958) marked the first comprehensive efforts to categorize hundreds of microseisms observations from around the globe, despite many reports being confined to local-language publications. His works established the concept, nowadays completely accepted, that different period bands could be related to different source mechanisms and that the regional geology affecting the wave propagation could change characteristics of the wavefield. In his work, Gutenberg (1958) extended his analysis to a truly global scale, synthesizing measurements from seismic stations worldwide to show how ambient noise properties, both amplitude and

dominant period, vary with geographic and tectonic context. This work not only introduced the discretization between natural and anthropogenic noise sources but also highlighted the controlling role of large-scale geological structures on the spatial behaviour of these ambient vibrations.

Fig. 2.1 (Bonnefoy-Claudet et al., 2006) provides a comprehensive summary of part of the pioneering classification made by Gutenberg (1958) and then almost totally confirmed by Asten & Henstridge (1984), illustrating the different type of noise sources and their related frequency content.

	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

Figure 2.1 – Summary of main ambient noise natural and anthropic sources according to their dominant frequency content. Figure is adapted by Bonnefoy-Claudet et al. (2006), based on earlier works by Gutenberg (1958), Asten (1978) and Asten & Henstridge (1984).

The high-frequency content (short-period) of seismic noise, roughly above 1Hz, is essentially produced by human activities posed at the Earth's surface. Indeed, this part of the spectral content of ambient noise is dominated by diurnal and weekly fluctuations that produce a bimodal shape (Nakata et al., 2019) of the spectrum in which amplitudes are higher during the day respect to the night and during the working week respect to the weekend. This behaviour suggests that the frequency content is mostly due to vehicular traffic, industrial machinery, and urban activity, which feed high-frequency noise into the ground motion (Okada, 2003). Yamanaka et al. (1993), through their experiment at the USC campus in Los Angeles, extracted spectral amplitudes at short and long periods, and observed a pronounced midday peak in the short-period band (traffic and human activity) alongside a long-period signal presumed to be effect of offshore ocean waves at ~100 km far. Conversely, at low frequencies the dominant natural sources produce a much more stable spectral content over time, with fluctuations that are far less periodic and more diffuse than the sharp diurnal peaks associated with anthropogenic high-frequency noise. The lowest frequency band (long-period) below approximately 0.5 Hz is dominated by energy from large-scale meteorological and oceanic phenomena, correlated with atmospheric pressure systems. In the intermediate band around 1 Hz (mid-period), locally generated wind and other weather-related phenomena contribute significantly to the noise wavefield.

This practical separation between natural and cultural sources is clearly recognizable observing seismic stations located deep within continental lands and stations located closer to the coastlines. First group tend to record high-frequency band spectrum dominated by cultural noise (traffic, industry, and other human activities) while the latter progressively exhibit stronger low-frequency energy associated with natural sources, such as microseisms due to the ocean-coast interaction (Peterson, 1993).

During colder months, ambient seismic noise below 1 Hz consistently exhibits elevated amplitudes, a reflection of more frequent and intense meteorological events and ocean effects. Seasonal analyses reveal that microseisms power often peaks in winter season, closely linked to variations in storm activity and large-scale pressure systems. Moreover, high-resolution barometric records Sorrells et al. (1971) demonstrate a robust relationship between atmospheric pressure and noise intensity: microtremor power rises rapidly following a pressure drop and subsides as pressure recovers, confirming after almost hundred years the first qualitative hypothesis made by Bertelli (1872).

More recently, Poli et al. (2020) analysed continuous seismic records from multiple Italian seismic stations before, during, and after the COVID-19 lockdown, documenting a marked decrease in high-frequency ambient noise (1–10 Hz) corresponding to the period of severely reduced road traffic and industrial activity. Noise levels reached their lowest during the strictest lockdown weeks, then gradually recovered as restrictions were lifted. This natural experiment confirms the strong dependence of the 1–10 Hz band on anthropogenic sources produced by daily work of people.

Although the 1Hz-threshold is not absolute, it serves as a practical delineation: natural processes prevail at frequencies below 1Hz, while human activities govern the above frequency content of ambient noise spectrum. However, although it is frequently used as a practical cutoff between natural and anthropogenic noise, local site conditions, such as sediment thickness, basin geometry, and proximity to the coast, can shift this boundary by several hertz. For example, analysing continuous ambient noise measurements within the Grenoble basin, Bonnefoy-Claudet et al. (2004) showed that daily changes in the spectral amplitudes of the seismic noise produced by changes in human activities could occur also at frequencies lower than 1Hz as response of deep soft materials in the stratigraphy, or more in general based on the local geological setting.

Technological advancements in instrumentation and data processing have significantly enhanced the capacity to quantitatively study ambient seismic noise. The global deployment of permanent seismic stations has further accelerated this progress by providing continuous, long-term records from diverse geological and cultural settings, enabling researchers to discretize source contributions and monitor tiny temporal variations. Such combined improvements have led to the development and wide application of seismic exploration techniques entirely based on the acquisition and processing of ambient noise. Despite these advances, the fundamental associations between noise frequency bands and their underlying source types, first outlined by Gutenberg, remain largely still accepted, with only minor refinements introduced over subsequent decades.

2.2 Ambient noise as a stochastic and stationary process

All above-mentioned observations have demonstrated that ambient noise can be characterized both spatially and temporally, and that its analysis depends on how its properties vary over time and space.

Several pioneering field studies helped establish the global consistency and local variability of seismic noise. Frantti et al. (1962) and Frantti (1963) investigated 48 geological different sites, recording two-minute “quiet” windows, and showed that noise levels tend to decrease between about 0.6 and 1 Hz then remain essentially flat up to several tens of hertz, revealing a remarkably uniform spectral shape across very different geological contexts. In Fig. 2.2 (reproduced from Peterson, 1993) a clear example of spectra relative to noise measurements made at different sites around the world is shown: beyond the different spectral amplitudes which are obviously correlated with the noise intensity and the local site conditions, all spectra show the same spectral behaviour.

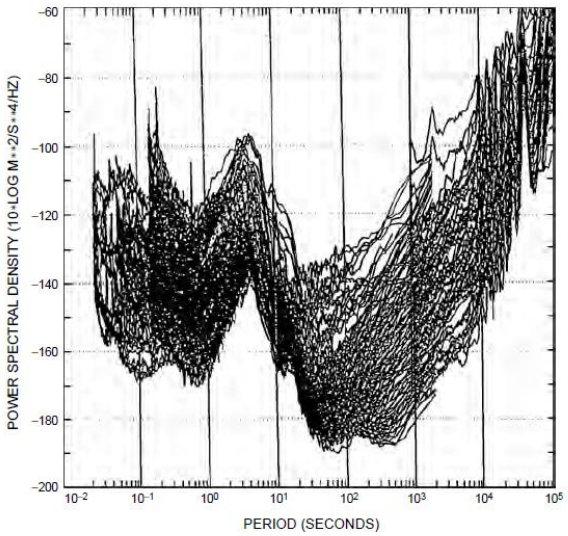

Figure 2.2 – Power spectral density of ambient seismic noise measured at different sites worldwide. Despite, differences in spectral amplitudes, a clear consistency in the ambient noise spectral behaviour is visible over a wide frequency range (Peterson, 1993).

This attitude of the ambient noise spectra indicated that, regardless of the recording site and the stratigraphic and geological conditions, it exhibits the same spectral shape with only small deviations in their amplitudes (so the energy) of the recorded noise.

This spectral invariance suggests that ambient seismic noise, when recorded at least over intervals of tens of minutes, acts as a stochastic and stationary process (Aki, 1957) produced by continuous, overlapping contributions of countless natural and human sources. As random process, the exact waveform cannot be predicted from event to event, and this requires us to characterize the noise field by its statistical (mean, variance, autocorrelation) and spectral properties. This leads us to model the noise as a collection of random realizations characterized by a time-invariant form of the power spectral density and sharing the same probability laws for which the statistical distribution of amplitudes within each frequency band approximates a Gaussian distribution (Nogoshi & Igarashi, 1971). In this context, the probability function of the amplitude values of the noise takes more or less the same mean value if measured for different periods of time and this means that, although ambient seismic noise varies both spatially and temporally, and it is therefore potentially

informative of the local subsoil characteristics, when it's recorded over a sufficiently long period (at least tens of minutes), effects of individual sources are theoretically cancelled out (Okada, 2003).

This implies that investigating such phenomena calls for an approach markedly different from conventional once, the emphasis moves away from individual seismic arrivals toward the study of the signal's average characteristics, especially in the frequency domain. Indeed, stationarity and stochastic behaviour play a key role in the application of most ambient noise-based techniques, which rely in the assumption that the contribute of each noise source can be neglected and the noise time series can be analysed such a perfect random process. Understanding the essential properties of a stochastic process is of main importance to provide the essential aspects of the ambient noise application as discussed in this thesis. Spectral representation is especially useful because it lets us break the noise down into individual frequency components, revealing how energy is distributed across the spectrum and making it easier to compare and analyse its properties over time and between different sites.

A stochastic process represents a mathematical framework for describing time-dependent quantities in probabilistic terms; each specific time series generated is referred to as a realization of that process. Unlike a deterministic process, identical in each shot, a stochastic process admits infinitely many non-repeating realizations that share only statistical characteristics. This distinction is essential in ambient noise analysis, since no two different time realizations of the process are identical, yet they generally reproduce stable spectral properties. In this sense, recorded ambient seismic noise *x(t)* can be modelled as a zero-mean, stationary, ergodic stochastic process (Eq. 2.1). This framework not only provides a brief mathematical description of the properties of the noise but also explains why passive recordings, even without impulsive sources, can be used to provide the same information as active sources, in the assumption of long enough time records able to capture the full statistic properties of the process.

Following Priestley (1981) and Okada (2003), each realization of the process can be expressed in the form of a stochastic spectral integral:

$$x(t) = \int_{-\infty}^{\infty} e^{i\omega t}\, dZ(\omega). \qquad [2.1]$$

In this formulation, *dZ(ω)* is a complex-valued random measure denoting an infinitesimal complex random increment of the noise process in the frequency band *[ω, ω + dω]*. Its ensemble mean is zero and its variance is *S(ω) dω*, where *S(ω)* is the power spectral density. By integrating $e^{i\omega t}\, dZ(\omega)$ over all frequencies, we can reconstruct a time-series *x(t)* whose statistical properties (zero-mean, stationarity, spectrum) exactly match those of the ambient noise wavefield and for infinitesimal increments over frequency satisfy:

$$E[dZ(\omega)] = 0, \qquad [2.2]$$

$$E\left[dZ(\omega)\,\overline{dZ(\omega')}\right] = S(\omega)d\omega, \qquad for\ \omega = \omega' \qquad [2.3]$$

$$E\left[dZ(\omega)\,\overline{dZ(\omega')}\right] = 0. \qquad for\ \omega \neq \omega' \qquad [2.4]$$

In Eq. 2.2, 2.3 and 2.4 $E[\cdot]$ denotes the ensemble mean across all possible realizations of the process and the overbar denotes the complex conjugate. The term $E[dZ(\omega)\overline{dZ(\omega)}]$ represents the ensemble expectation of the product between the infinitesimal noise increments at frequencies ω and ω'. Physically, it quantifies the average cross-spectral density: it vanishes for $\omega \neq \omega'$, indicating uncorrelated frequency bands, and it is equals to $S(\omega)\, d\omega$ when $\omega = \omega'$, thereby recovering the power spectral density in that band. Because each $dZ(\omega)$ has zero mean, the overall process $x(t)$ is also zero-mean (Eq. 2.5):

$$E[x(t)] = 0. \tag{2.5}$$

The above orthogonality condition ensures that components at different frequencies are uncorrelated, which in turn implies stationarity, so that all statistical moments of $x(t)$ are invariant under time shifts. In particular, the autocovariance $R(\tau)$ as below expressed in Eq. 2.6 depends solely on the time-lag τ and not on the absolute time t:

$$R(\tau) \approx E[x(t)x(t+\tau)], \tag{2.6}$$

In practice, this means that two different time windows recorded at different times provide almost unchanged local conditions. According to the Wiener-Khinchin theorem (Bendat & Piersol, 2010), the power spectral density $S(\omega)$ is given by the Fourier transform of the autocovariance function (Eq. 2.7):

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-i\omega r}\, d\tau \tag{2.7}$$

Likewise power spectral density remains constant, so that the familiar 'U-shaped' spectrum provided by Peterson (1993) reappears no matter when the data are recorded.

By discretizing the integral formulation (Eq. 2.1) into a finite sum of frequency bands, we draw a small random contribution for each band whose average energy matches the target power spectral density $S(\omega)$. Summing these oscillatory components in the time domain produces a synthetic time series (Fig. 2.3a) which, despite being a numerical approximation, preserves proximity to the zero-mean, stationarity, and spectral characteristics of ambient seismic noise. Power spectral densities (Fig. 2.3b) estimated from different, non-overlapping time windows exhibit a consistent spectral shape, supporting the theoretical assumption of stationarity.

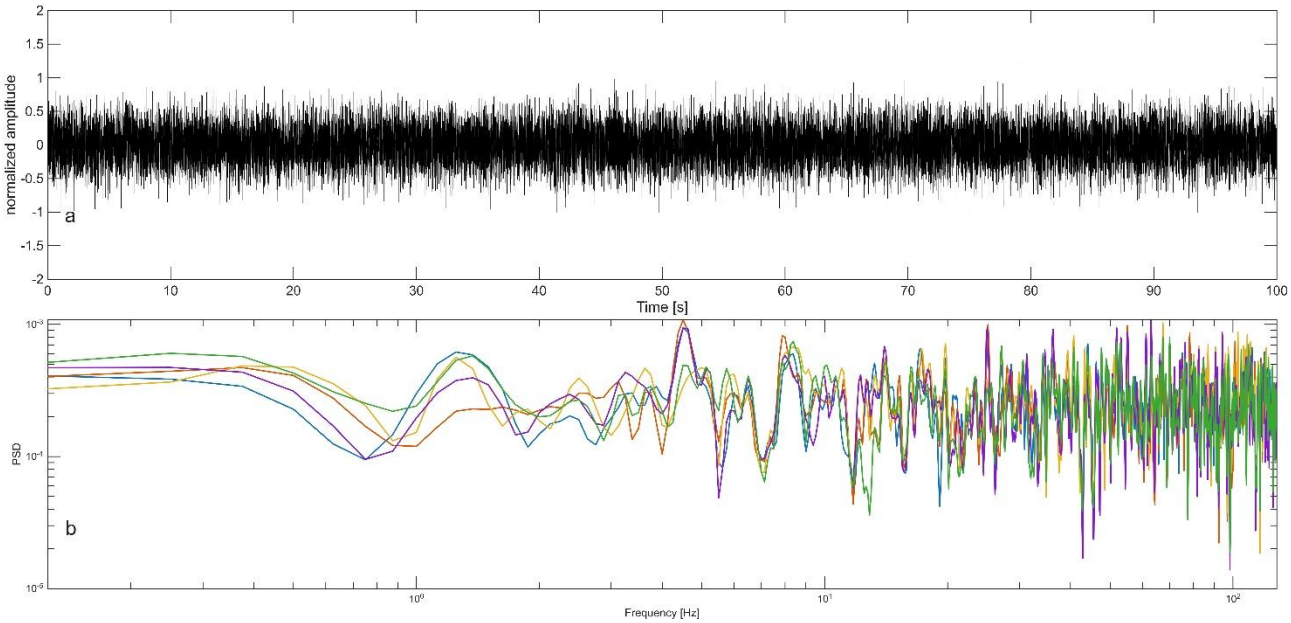

Figure 2.3 – Synthetic ambient seismic noise time series (a) and corresponding power spectral densities (b) computed for different time windows following Eq. 2.1. The relative stability of the PSDs across the different windows confirms the assumption of stationarity of the synthetic signal.

All the concepts introduced in this section provide the essential foundation to understand most ambient noise-based methods, which will be extensively discussed and described in the Chapter III. Based on these foundations, we now focus on the main components of ambient noise: surface waves. In the following section, we examine their properties, origin, propagation characteristics, and modal behaviour, because nearly all ambient-noise processing techniques rely on the fact that surface waves dominate the recorded ambient noise.

2.3 Surface waves

In last decades of 20th century up to now, many publications have been devoted to the study of ambient seismic noise: a restrict number of them focused to the study of the nature of this noise but most of them focused on the applicability of the recorded seismic noise in several ambient seismic methods (Bonnefoy-Claudet et al., 2006). Most of these ambient noise techniques work in the assumption that a substantial portion of microtremor energy propagate along the Earth's surface in all directions. These seismic waves confined to propagate along the Earth's surface have been recognized since the 19th century and we now call them surface waves.

When body waves encounter the Earth's free surface, they reflect and interfere with it, producing sustained motion that propagate confined near the surface with velocity depending on the body waves velocities, which in turn depend on the properties of the medium below the surface. In general, these superficial waves arise through the coupling of compressional and shear waves at the free-surface and by the trapping of shear energy

in near-surface layers, resulting in wavefields that are both long-lived and able of transport substantial energy over large distances.

Although ambient noise is composed of both body and surface waves, the latter generally dominate the recorded signals. This predominance is due not only to the fact that most microtremors sources act at the Earth's surface or on the seabed, but also to the lower attenuation of surface waves compared to body waves. Indeed, surface waves, as body waves, show a loss of energy during propagation, that is a progressive decay of amplitude as function of the distance. Unlike body waves, however, surface waves are characterized by a slower geometrical spreading, therefore, these latter attenuate less rapidly with distance than body waves (Lay & Wallace, 1995). This is mostly related to the tri-dimensional propagation domain of body waves, whereas surface waves propagate along the Earth's surface in a 2D domain.

As a result, surface waves contribute significantly to the ambient noise wavefield, both from local and distant sources, whereas body-waves energy is rapidly attenuated, especially at frequencies above few Hertz.

This implies that at distances of many times the considered wavelength, the contribution of the body waves at the ground motion decreases, while surface waves dominate the wavefield, especially at frequencies above few Hertz. Even in the case of an earthquake, although surface waves are the slower waves to be recorded, they show the higher amplitude, and these differences of amplitude and velocity make them also more recognizable with respect to propagating body waves (direct, reflected or refracted).

Contrary to what the name suggests, the displacement of particles produced by propagating surface waves is not limited to just the Earth's surface. Surface waves also impact particles below the surface, with the displacement amplitude decreasing exponentially as depth increases in a uniform half-space (Aki & Richards, 1980). Mathematically, in a homogeneous half-space the particle-motion amplitude decay approximately as:

$$A(z) \propto e^{-2\pi z f}. \qquad [2.8]$$

In Eq. 2.8 *A(z)* is the motion due to surface waves as function of the depth z and f is the frequency of the considered wave. Consequently, considering the frequency as the inverse of the wavelength, short wavelengths remain confined in the shallow portion of the subsoil, decaying rapidly with depth, while long wavelengths penetrate much deeper, with motion extending deeper. In a homogeneous medium, whose properties are invariant with depth, each harmonic component propagates with the same velocity regardless of its wavelength. Consequently, the investigated depth has no influence on wave velocity.

Now consider the subsoil model in Fig. 2.4 (from Foti et al., 2014). This is not homogeneous anymore, at least if one accounts only for vertical variations of the elastic properties. In this stratified model, each layer has its own P and S wave velocity (Vp and Vs) and its own density. Considering Eq. 2.8, and that surface wave velocity strongly depend on the Vp and Vs velocity of body waves, it becomes clear that the higher frequency will remain confined in Layer 1 of the model, while lower frequency will penetrate much deeper, with appreciable motion extending through Layers 2 and 3. This also means that the propagation of the wave travelling with higher frequency will depend only on the properties of the first layer (more generally from the

properties of the shallow subsoil), while the low-frequency waves will be influenced also by the deeper layer of the model.

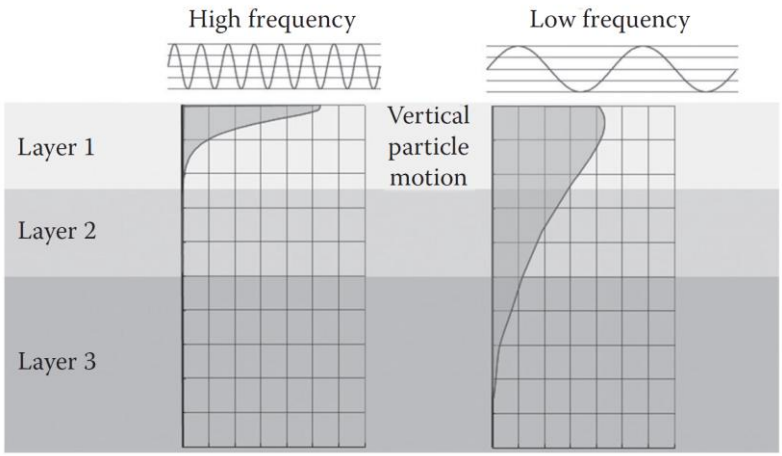

Figure 2.4 – Schematic representation of frequency-dependent depth penetration of surface waves in a vertical multilayered elastic medium. High-frequency waves are extremely sensitive to the first subsoil, while low-frequency harmonics are influenced also by the elastic properties of underlying layers of the model. Reproduced from Foti et al. (2014).

This depth-dependent sensitivity of surface waves underpins the fundamental characteristic of surface waves, that is the dispersion property. Dispersion of surface waves means that each frequency (or each wavelength) "samples" a different effective depth range, its velocity is influenced by the body wave velocities over that depth. Measuring frequency-velocity relation of surface wave allows to construct the so-called dispersion curve, which shape, and values necessarily depends on the subsoil properties in terms of variation of Vs, Vp and density.

Practically, short wavelengths (high frequencies) carry information about the uppermost layers, making them ideal for shallow site characterization, and long wavelengths (low frequencies) inform on deeper structures, enabling investigation of the sediment–bedrock transition and deeper crustal features.

Thus, Fig. 2.4 not only illustrates the physical origin of surface waves dispersion but also highlights why surface-wave analysis can noninvasively investigate a wide depth range simply by analyzing the frequency-dependent velocities. In the case of a simple homogeneous model, all harmonics will propagate with the same velocity, and the resulting dispersion curve will be flat with no changes across the broad frequency range. Dispersion character of surface waves is nowadays used as practical tool in almost every surface wave method (that will be discussed in the next Chapter). Most of them are devoted to retrieve the shear-wave velocity profile of the first subsoil in the assumption of a 1D velocity model. In contrast, recently the extraction of inter-station surface waves velocity from dispersion analysis has become the primary observed data aimed to the imaging of the upper and lower part of the Earth's crust.

Surface-wave propagation in layered media exhibits also a multimodal behaviour, with multiple "standing-wave" patterns, known as modes, coexisting at a single frequency. Each mode has its own vertical motion

profile and thus samples different depth intervals, which in turn gives it a different velocity despite sharing the same oscillation frequency. The simplest, fundamental mode concentrates motion near the surface, while higher modes extend sensitivity deeper into the medium. Because each of these patterns interacts differently with the layer stack, they travel at different velocity even though they all share the same frequency. Fig. 2.5 shows an example of a Vp and Vs velocity profile, together with the corresponding dispersion curve, in which the higher modes are also highlighted. Characteristic dispersion behaviour is clearly recognizable: in such a normal dispersion model on the left panel of Fig. 2.5, lower-frequency waves propagate with velocities close to the Vs velocity of the deeper layers, while higher-frequency waves are influenced only by the uppermost layers and propagate with lower velocity.

In practice, the fundamental mode usually dominates the low frequency band and carries most of the energy, but at higher frequencies the higher modes can become excited and contribute noticeably to the recorded wavefield (Tokimatsu et al., 1992; Lai & Rix, 1998). Recognizing modal behaviour is an essential step in surface wave analysis because it means a single frequency measurement can, in principle, shows multiple depth intervals, each mode providing an independent look into the subsurface structure.

Because each mode travels with its own velocity, one practical method for isolating individual modes is to observe recordings at large epicentral distances, that is where higher modes arrive at distinctly different times, allowing them to be more easily separated.

In normally dispersive layered media (with velocities increasing with depth) with small impedance contrasts, the relative propagation of Rayleigh- and Love-wave is largely dominated by the fundamental mode. As contrasts increase or in an inverse dispersive medium (velocities decrease with depth), however, significant energy shifts into the higher modes, and the soil medium response includes energy contributions from higher modes (Tokimatsu, 1995; Lai & Rix, 1998).

In medium with low-velocity interlayers or other irregular S-wave contrasts, higher modes often carry most energy in specific frequency bands, producing inversely dispersive trends in the observed dispersion curve. As a result, the measured surface wave velocities must be viewed as apparent values results of the superposition of multiple modes (Ohori et al., 2002). Indeed, considering only the fundamental mode while neglecting higher modes often yields apparent dispersion curves that do not reflect the true subsurface structure, since the interplay among modal curves produces an effective experimental curve (Tokimatsu et al., 1992; Parolai et al., 2006).

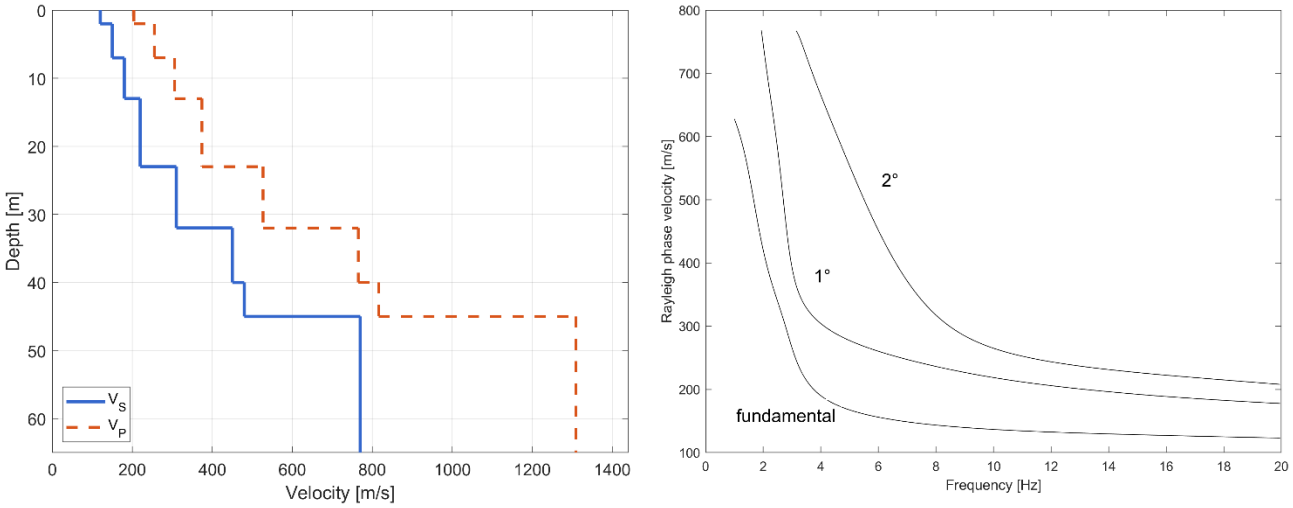

Figure 2.5 – Left: Stratified velocity model expressed in terms of V_S and V_P. Right: Theoretical dispersion curves (of Rayleigh waves) associated with the model, with evidence for the fundamental mode and two higher modes.

2.3.1 Rayleigh and Love waves

When body waves reach a boundary between two different media, depending on whether the arriving energy is in the form of P-waves, vertically polarized shear waves (SV), or horizontally polarized shear waves (SH), and on the nature of the boundary they encounter (for instance, an air-rock surface, a water-rock interface, or a soft sediment layer atop bedrock), these interactions give rise to different types of waves. In each case, a portion of the incoming energy is converted into a guided mode that remains trapped near the interface, decaying as it moves away from it. The result is a rich variety of surface and interface waves; each one is produced by specific combinations of wave polarization and boundary conditions. For example, coupling between P and SV motions at a free surface produces Rayleigh waves, whereas the trapping of SH energy within a normally dispersive medium gives rise to Love waves. Both Rayleigh and Love waves show their maximum amplitudes at the surface and exponentially decay with increasing depth. Other guided modes waves can occur under more specific conditions: Scholte waves generate at a at solid–fluid boundary (such as the ocean floor) with amplitudes decreasing exponentially from the interface into both fluid and elastic solid, while Stoneley waves are boundary waves that are created at the interface between two elastic solid media and decay exponentially away from the interface (Aki & Richards, 1980).

Rayleigh and Love waves were identified in the late 19th and early 20th centuries and since then they remained the most widely used seismic waves in seismology, prized for their easy applicability. In 1887, Lord Rayleigh published his seminal work demonstrating that elastic media could support a wave confined to their free surface, a discovery that fundamentally expanded our understanding of seismic propagation. Later, in 1911, A.E.H. Love identified a second surface-guided wave mode in layered materials, and the “Love wave” has thus become the complementary counterpart to Rayleigh’s original surface wave. Together, these two waves,

born from distinct body waves and boundary interactions, have served as the cornerstones of surface-wave seismology ever since.

Rayleigh waves, generating through the coupling of compressional (P) and vertically polarized shear waves (SV) at the free surface, give life to a typically elliptical particle motion in the vertical plane, with retrograde rotation respect to the propagation direction. These waves travel with velocity (V_R) slightly less than the shear-wave velocity (V_s) of the medium. In an isotropic, homogeneous half-space, the motion of Rayleigh waves satisfies the elastic 2D wave equation with the boundary conditions expressed by the stress-free condition at the surface and the radiation condition (which impose vanishing displacement) as the depth increase.

Love waves, by contrast, result from horizontally polarized shear waves (SH) trapped in a shallow low-velocity layer over a higher-velocity half-space and propagate with velocity V_L similar to V_R. Unlike Rayleigh waves, Love waves involve purely transverse motion in the horizontal plane, with particle displacement perpendicular to both the direction of propagation and the vertical axis. They do not exist in a simple homogeneous media and require specific conditions for which $V_{s1} < V_{s2}$, where V_{s1} is the shear-wave velocity of the surface layer and V_{s2} that of the underlying half-space. As Rayleigh wave, also Love waves share the same boundary conditions at the top and bottom of the model (Thomson, 1950; Haskell, 1953).

Given these considerations, it turns out to be clear that, neglecting the effect of body waves and considering only surface waves, vertical records of the ground motion are entirely due to the Rayleigh wave propagation, while horizontal ground motion is a complex mix of both types of surface waves.

In a perfectly homogeneous, isotropic half-space, where shear-wave velocity is the same at all depths, Rayleigh waves are nondispersive: every frequency component of the wave travels at the same velocity, that is a fixed fraction of the shear-wave velocity. By contrast, Love waves require velocity contrasts to exist at all. They originate only in a “waveguide” created when a softer, lower-velocity layer overlies a stiffer, higher-velocity half-space. In such a configuration, horizontally polarized SH energy becomes trapped and builds up through constructive interference. If the shear-wave velocity were uniform with depth, no such trapping could occur, and no Love-type modes would be supported (Kennett & Kerry, 1979).

Early studies conducted on arrays showed that the wave field of ambient seismic noise cannot be attributed to a single type of wave but rather results from the superposition of surface and body waves, with contributions that vary depending on frequency, subsoil structure and source characteristics (Toksöz & Lacoss, 1968; Horike, 1985; Arai & Tokimatsu, 1998; Yamamoto, 2000; Cornou, 2002). At low frequencies (~ 1Hz) microseisms are often dominated by Rayleigh and Love waves fundamental modes, especially Rayleigh waves, at least as far as the most superficial structures are concerned. As the frequency increases the wavefield often becomes a mixture of body waves and surface waves, both fundamental and higher modes, with the exact balance sensitive to source distance and impedance contrasts. Overall, the Love-to-Rayleigh and body-to-surface energy ratios are site-dependent, controlled by layering, impedance contrasts and the spatial distribution of sources, underscoring the need to identify wave type and mode before inverting noise data for shear-wave velocity. Bonnefoy-Claudet et al. (2006) present a thorough synthesis of dozens of fields and modelling studies,

highlighting both convergent and conflicting findings, and quantify how the different phases and modal branches (body waves, Love, and Rayleigh waves) contribute to the ambient-noise wavefield.

2.3.2 Phase and group velocity

We have thus far reviewed the key attributes that justify treating ambient seismic noise as a stationary stochastic process, introduced the fundamental concepts of dispersion and modal behaviour in surface waves, both direct consequences of the two-dimensional wave equations that govern their propagation and highlighted the persistent challenge of isolating the relative contribution of each wave phase when analysing ambient-noise records. Although these phenomena directly affect how surface-wave velocity varies with frequency and mode, we have so far referred to a single, generic "surface-wave velocity". Actually, surface waves can be identified by two different types of velocity, each related to different physical aspects of the surface waves.

Consider a general monochromatic source generating a single oscillation with its own phase velocity *c*. While phase velocity alone would be sufficient to describe the propagation of a perfectly monochromatic wave, any real source generates a superposition of harmonics spanning a range of frequencies, each travelling at its own phase velocity. This frequency-dependent propagation causes the different harmonic components to gradually travel at slightly different velocities, producing interference patterns: constructive interference where the components reinforce each other, and destructive interference where they cancel. The result is a modulated signal, a wave packet, whose overall envelope travels at a velocity different from that of the individual harmonics. This is the group velocity U, which governs the propagation of the packet's energy.

While phase and group velocities are conceptually different, as they will be defined in this section, they share the same dispersion and modal behaviour described before. In other words, both velocities vary with frequency and mode in much the same way, even though they differ in their interpretation: phase velocity relates to the motion of a single harmonic, whereas group velocity is related to the propagating envelope of a broadband wave packet. Because these two velocities generally differ, they give rise to two separate dispersion curves, each carrying valuable information about subsurface structure. In the following section, we will define phase and group velocities formally and explore their mathematical relationship. Following Lay & Wallace (1995) we can start introducing two different cosine waves with slightly different frequencies ω_1 and ω_2 and wavenumber k_1 and k_2 propagating in space *x* and time *t*. If we record the displacement due to the two waves, it will be the sum of the two contributions as in Eq. 2.9:

$$u(x,t) = cos(k_1 x - \omega_1 t) + cos(k_2 x - \omega_2 t), \qquad [2.9]$$

where the two wavenumbers k_1 and k_2 are function of two different phase velocities c_1 and c_2, such that:

$$k_1 = \omega_1/c_1 \quad ; \quad k_2 = \omega_2/c_2$$

It is easy to see that applying trigonometric identities of the kind $\cos(A)\cos(B) = 2\cos((A+B)/2)\ \cos((A-B)/2)$ to Eq. 2.9 leads to the definition of new terms ω, k, $\delta\omega$ and δk:

$$\omega = \frac{\omega_1 + \omega_2}{2};\ k\ = \frac{k_1 + k_2}{2};\ \delta\omega = \frac{\omega_1 - \omega_2}{2};\ \delta k = \frac{k_1 - k_2}{2}.$$

Applying this simple manipulation, we can define a central angular frequency ω, representing the mean between ω_1 and ω_2, and a mean wavenumber k between k_1 and k_2, such that $\omega = k/c$, where c is the mean phase velocity between c_1 and c_2. The two delta terms represent the distance of the mean values from angular frequencies ω_1 and ω_2 and wavenumbers k_1 and k_2. The total displacement $u(x,t)$ can be easily re-written as:

$$u(x,t) = 2\cos(\omega t - kx)\ \cos(\delta\omega t - \delta k x) \qquad [2.10]$$

Looking in detail at the two terms in Eq. 2.10, these are two cosine terms with the second one representing an oscillation with lower velocity then the first one. This simplified representation of $u(x,t)$ describes a wave with an angular frequency ω and wavenumber k, modulated by a cosine term representing an oscillation with lower velocity then the first one (black wave in Fig. 2.6). This formulation represents a carrier wave (Brillouin, 1960) moving with a phase velocity $c = \omega/k$, which is subject to amplitude modulation that propagates with the group velocity $U = \delta\omega/\delta k$. The phase velocity refers to the individual wavelet velocities for a fixed frequency. In contrast, group velocity corresponds to the velocity at which energy is transmitted across a frequency band, which can be interpreted as the propagation of the energy envelope that contains wavelets with phase velocity *c(ω)* and with close surrounding frequencies.

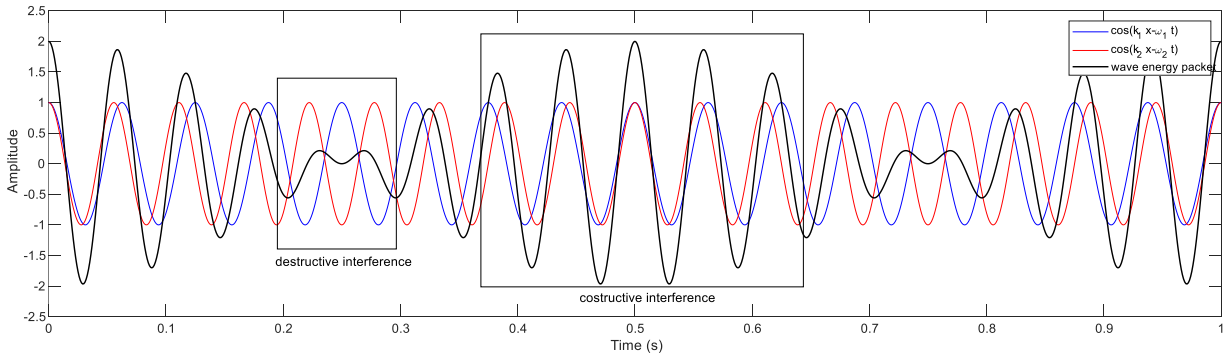

Figure 2.6 – Example of two diverse waves (red and blue curves) propagating with two phase velocities. Interference between the waves generate an amplitude modulation that propagate with group velocity U slightly different from the phase velocities of the two original waves.

In Fig 2.6, the two rectangles represent respectively areas where the two initial cosine waves interfere. In the first one between 0.2 and 0.3 seconds, the two waves interfere destructively, and the resulting amplitudes of

the modulated waves decrease drastically with respect to the amplitudes of the two original waves. In the second rectangle the two original waves are almost in phase, and they interfere increasing the resulting amplitudes of the modulated wave. Initially, the source pulse can be viewed as a compact packet in which each frequency component is launched with its own amplitude and phase. As the wave propagates, frequency-dependent velocities cause the packet to stretch into a longer wave-train in time. The energy remains concentrated in the constructive-interference envelope of these components, and that envelope travels at the group velocity, not at any of the individual phase velocities.

If we express Eq. 2.10 as a limit for $\delta\omega \rightarrow 0$ and $\delta k \rightarrow 0$, we see that group velocities can be computed as a simple derivation from phase velocity values:

$$U = \frac{d\omega}{dk} = \frac{d(kc)}{dk} = c + k\frac{dc}{dk} \qquad [2.11]$$

Group velocity U is strongly influenced by the phase velocity c and above all by the variation of phase velocity with respect to the wavenumber. If the derivative $dc/dk = 0$, that is characteristics of a homogeneous and non-dispersive medium, group and phase velocity are identical, and they do not depend on the frequency. Switching from group velocity to phase velocity is more complex; indeed, deriving phase velocity from group velocity is a delicate process due to the intricate, non-linear relationship between the two. However, in a normal dispersive medium, such as the Earth crust, velocities rise with depth and thus phase velocities decrease with frequencies. In this context group velocities are usually lower than phase velocities.

In surface wave methods, phase velocity measurements are generally preferred because they resolve individual modes with higher spectral detail and thus yield refined, depth-sensitive constraints. For broader-scale objectives, such as regional tomography, group velocities are often used because they are easier to estimate, less sensitive to mode misidentification and provide a convenient summary of energy transported across large apertures. The next Chapter therefore will present a comparative overview of the main techniques used to extract phase- and group-velocity dispersion curves, highlighting their respective advantages, practical complexities and inherent limitations, focusing on two different approaches, one to achieve phase dispersion curve (spatial auto correlation methods) and one to extract group dispersion curve (Seismic Interferometry).

Chapter III.
Dispersion curves from seismic noise

3.1 Introduction to surface waves methods

Surface wave methods rely on the analysis of the dispersive behaviour of surface waves propagating parallel to a free-surface (e.g. the Earth's surface) and have been used, historically and in current practice, primarily to infer the subsurface shear-wave (Vs) profile (Foti et al., 2011). This is because Rayleigh and Love waves are strongly influenced by the shear wave velocities and by how they change with depth, with only secondary sensitivity to P-wave and density-profile.

Surface waves travelling along the surface of a homogeneous half-space produce motion in terms of amplitudes that decay approximately exponentially with depth, so that most energy is confined within a depth on the order of one wavelength of the considered wave (Achenbach, 1999; Foti et al., 2014). In stratified media, as just mentioned in the previous Chapter II, dispersion behaviour of surface waves allows for different wavelengths sampling different depths. As a result, the phase velocity at a given frequency reflects the properties of the specific depth interval explored. In this context, such as in layered media, amplitude decay with depth cannot be predicted reliably in the absence of a priori constraints on the velocity structure (Foti et al., 2018).

The relationship between phase velocity (or group velocity) and frequency (or wavelength) defines what is known as the dispersion curve. This peculiar behaviour makes surface waves able to characterize the subsoil at several different scales, from meters to kilometres, only depending on the frequency range analysed and the sampling geometry of the seismic receivers.

Furthermore, surface wave propagation in layered media is inherently a multimodal process: multiple vibration modes can occur at each frequency, each characterized by its own distinct velocity of propagation. Nevertheless, in the case of a normally dispersive profile, where shear-wave velocity increases with depth and no significant impedance contrasts are present, energy tends to be concentrated in the fundamental mode. While this mode is most commonly analysed, higher modes may also be detectable and provide valuable information, such as in case of strong impendence contrasts at depth. That is why, in modern surface-wave surveys, attention is paid not only to the fundamental mode but also to higher modes when constructing the effective (composite) dispersion curve. For most frequency bands this curve coincides with the fundamental one, yet, depending on the stratigraphic layering and the shear-wave velocity profile, it can display jumps or deviation from the fundamental mode whenever a greater share of energy is transported into the higher modes (Tokimatsu et al., 1992). Surface wave dispersion analysis is most performed using vertical component of ground motion (Foti et al., 2018), which account only for motion due to the propagation of Rayleigh waves, although in some cases horizontal components are also considered. This preference arises from the relative simplicity of the vertical motion which can generally be attributed only to Rayleigh waves, while horizontal motion is a more complex mix of Rayleigh and Love waves. Given the superposition of both Rayleigh and Love waves in the horizontal recorded motion, the analysis become more complex and potentially ambiguous (Malagnini et al., 1993).

A wide range of surface wave-based analysis can be performed using different strategies (e.g. data acquisition and processing), and if executed correctly and used within their valid assumptions, they should lead equivalent levels of reliability. Surface wave methods can be broadly classified into active- and passive techniques.

In active methods, seismic wavefield is directly generated at the surface for the purpose of the experiment, typically through mechanical impacts or controlled sources such as hammer shots. Among active-source techniques, the most widely adopted methods are SASW (e.g., Stokoe et al. 1994) and MASW (Park et al., 1999) methods. Both rely on controlled seismic sources and linear array geometries: SASW (Spectral Analysis of Surface Waves) exploit phase differences (phase of the complex cross-spectrum) of the signal generated by the source and recorded at two sensors to extract phase velocity information. In contrast, MASW method (Multichannel Analysis of Surface Waves) takes advantage of being applied across several receiver located along a straight line. As other multi-station analysis procedure, MASW technique rely on transforming, typically by a double Fourier transform, the recorded signals from the original time–space domain into alternative spectral representations, such as frequency-wavenumber or frequency-slowness domain, where surface wave properties can be more clearly isolated and interpreted and the expected phase velocities can be selected by picking the spectral maxima (Foti et al., 2014). These methods are particularly effective for shallow investigations where high-frequency energy is needed, and the survey geometry can be precisely controlled. These spectral peaks correspond to coherent surface wave components detected at the receivers and can be used to extract phase velocity as a function of frequency. Such domain transformations are especially useful in array-based methods, as they enhance the signal-to-noise ratio and allow the separation of overlapping wave modes or interfering wavefronts, which would otherwise be difficult to resolve in the time domain. However, a clear limitation of active methods remains their restricted penetration depth, which depends on the energy of the seismic source. To address the limited low-frequency content of active surveys, passive surface-wave techniques are increasingly adopted as a complement, extending sensitivity to lower frequencies, so to deeper layers. Unlike active sources, the approaches rely on the ambient noise wavefield continuously present at the surface.

While this ambient noise was once regarded as a source of contamination in seismic records, it is now recognized as a valuable signal, especially due to the predominant presence of surface waves, which are highly sensitive to the shallow structure of the Earth's crust. The appeal of passive methods lies in their non-invasive nature and their operational flexibility. They require no artificial sources, making them particularly suitable for densely urbanized or geologically complex areas where active-source surveys are logistically difficult, expensive, or restricted. Passive techniques can be deployed over extended periods and with minimal environmental impact, enabling long-term monitoring and exploration with reduced cost and infrastructure.
As already mentioned in Chapter II, these methods operate under the assumption that the wavefield behaves as a stationary and stochastic process in both time and space, an assumption that, while often approximately satisfied for temporal stationarity, may be more critical regarding spatial isotropy, particularly in near-surface

contexts dominated by directional anthropogenic sources. This condition enables ensemble averaging over time and space, which helps suppress incoherent energy and enhance coherent wave components to be reliably extracted even in the absence of a known or repeatable seismic source. However, since the signal of interest is included within ambient noise with unpredictable characteristics, passive techniques generally require significantly longer recording times (tens of minutes for near-surface applications up to several months for broad-scale investigations) compared to active methods (often few seconds). Whereas active-source surveys can improve the signal-to-noise ratio by repeating the same impulsive source multiple times and stacking the results, passive approaches must rely on time averaging over long durations, or over several realization of the process, to achieve similar levels of stability and resolution. Within passive seismic methods, two main approaches are commonly employed to extract seismic site parameters and subsurface structure information: standalone measurements and array-based techniques.

Standalone or single-station methods often require only a three-component sensor, and they are often used to compute the Horizontal-to-Vertical Spectral Ratio (Nogoshi & Igarashi, 1971; Nakamura, 1989), so the ratio between spectral amplitude of the motion recorded in the horizontal direction (both north-south and east-west) and spectral amplitude of the vertical motion. These techniques are not made directly for the extraction of dispersion curves but instead HVSR curve provide information on site resonance frequencies and the underground structure.

The HVSR technique is based on the assumption that the ratio between horizontal and vertical motion carries information about the frequency-dependent elastic properties of the subsurface, and in particular about its fundamental resonance frequencies (Bard, 1998). In this framework, several authors (Arai & Tokimatsu, 1998; Konno & Ohmachi, 1998) have assumed that ambient noise is composed predominantly of Rayleigh surface waves, linking the H/V peak directly to the ellipticity of Rayleigh waves at the site resonance frequency.

In this context, a sharp peak in the HVSR curve should represent an accentuate velocity contrast between different layers with the peak amplitude depending on the difference in terms of velocity. This behaviour is widely accepted and more than one peak in the curve can be observed in presence of more impedance contrast, even in this case with peaks at lower frequencies marking contrasts at higher depths. However, over the years, several studies have challenged this single interpretation of the H/V ratio. In particular, some procedures have modelled the ambient noise as a contribution of body waves, interpreting the HVSR curve and the resonance in a layered medium as the result of incident P- and S-wave (e.g., Herak, 2008). Other more recent approaches have instead assumed that the H/V ratio reflects the response of a complex full-wavefield (García-Jerez et al., 2016), generated by the combination of both body and surface waves, whose relative contribution depends both on the structure of the subsoil and on the characteristics of the ambient seismic noise.

Beyond the different physical interpretations of the H/V ratio, standalone measurements are now more frequently used in combination with seismic array-based methods (Arai & Tokimatsu, 2005; Parolai et al., 2005), with the aim of further constraining the subsurface structure in terms of 1D profiles of Vs and Vp. These

array-based approaches, which are the focus of this Chapter, are based on conceptually similar principles but adopt different methodological strategies to estimate the elastic properties of the subsurface.

Array-based approaches, often grouped under the name Microtremor Array Measurements (MAM), analyse the spatial coherence and propagation characteristics of surface waves within ambient noise recorded simultaneously at multiple stations. This category of techniques includes methods such as **SPAC** (Aki, 1957, 1965), its extended version **ESAC** (Ohori et al., 2002), frequency–wavenumber (**f–k**) analysis (Capon, 1969; Lacoss et al., 1969), and **tau–p** transform (e.g., McMechan & Yedlin, 1981). These techniques require spatially unaliased arrays and with defined assumptions about the noise wavefield and the subsoil model, results obtained from a specific array geometries are commonly used to estimate shear-wave velocity profiles (e.g., Foti et al., 2011 and Wathelet et al., 2005) starting from the extracted dispersion curve, often Rayleigh dispersion curve.

All these passive array-based methods rely on the statistical analysis of a stochastic process, which approximates the noise wavefield to a sum of horizontal plane wave propagating with different frequencies and in different directions. These waves are also assumed to propagate with the same phase velocity for a fixed frequency and that waves with different direction and frequency to be statistically independent (Boxberger et al., 2011) as introduced in Chapter II. Over the past decades, a great number of experiments, most based on the original Aki's spectral formulation, have been conducted to study the effects of source distribution of ambient noise and stations geometry on the resulting dispersion analysis in terms of phase velocity accuracy (e.g. Asten, 2006 and Asten & Hayashi, 2018). MAM are nowadays widely used given their easy applicability and because their reliability has been confirmed by tens of years of studies. Unlike active-source methods, they carry out non-invasive approaches particularly suited for near-surface characterization. These methods are commonly used in seismic microzonation and hazard assessment, especially in urban settings where active-source surveys may be impractical (Foti et al., 2011). Unfortunately, they require the use of a great number of stations to extract a site-specific dispersion curve, and the choice of the array geometry plays a critical role and can significantly affects the reliability of the results.

It is important to note that signals from active and passive sources generally exhibit different spectral characteristics: active sources deliver higher energy at short periods (high frequencies) and are therefore suited to high-resolution imaging of shallow structures, whereas ambient noise is more informative of long-period (low-frequency) energy, making it ideal for probing deeper subsurface layers, in the way that the two approaches can be often seen as complementary in the depth range they can explore. Equally important is that both active and passive surface-wave methods rest on the assumption of lateral homogeneity within the investigation area, i.e. a 1D medium whose seismic properties vary only with depth beneath the receivers; any departure from this 1-D condition can distort the observed dispersion curves and propagate significant errors into the resulting velocity models.

The frequency band in which active-source and passive data overlap usually spans only a few hertz, yet within this narrow window the phase-velocity picks obtained from passive records can extend the depth of investigation beyond what active techniques alone can reach. When both data sets are available, combining

them yields the most stable dispersion information; however, where a single approach must be chosen, SPAC often outperforms active MASW technique in near-surface applications (Asten & Hayashi, 2018). Using omnidirectional ambient noise over long recording periods, SPAC can deliver high-quality dispersion curves without artificial sources or logistical activities that require a lot of fieldwork. Thus, while MASW and SASW remain dependable for high-resolution imaging of the very shallow subsurface, SPAC and other passive array methods are increasingly favoured for their operational efficiency, greater penetration depth, and adaptability to a wide range of field conditions.

Another branch of methods is based on discovery made in the acoustic field at the beginning of this century according to which the correlation of a diffuse acoustic field gives rise to the Green's function of the medium through which waves propagate (e.g., Lobkis & Weaver, 2001). In subsequent years, these principles were also applied in the seismology with the advent of Seismic Interferometry: with seismic waves replacing acoustic waves and the subsoil replacing the acoustic medium (e.g., Campillo & Paul, 2003; Shapiro & Campillo, 2004; Snieder, 2004).

Developed more recently and with different objectives, such as crustal imaging on larger scales or in non-homogeneous media, Seismic Interferometry is not inherently designed for small-scale applications near the surface. Indeed, while it's also based on the recording of ambient noise, it focuses on the reconstruction of empirical Green's functions through the cross-correlation of long-duration time records between pairs of stations. One distinctive feature of this approach is that it allows for the use of spatially aliased sensor configurations, making it more flexible in deployment compared to original SPAC or f–k methods, which require dense, unaliased arrays. Spatial aliasing occurs when the inter-station distance exceeds half the shortest wavelength of interest, a condition that would introduce ambiguities in the wavenumber domain and compromise the spectrum resolution of array-based methods such as SPAC or f–k (Foti et al., 2018). Seismic interferometry, operating through cross-correlation in the time domain rather than through wavenumber-domain beamforming, is inherently less sensitive to this constraint, allowing for more flexible and irregular sensor deployments. Furthermore, unlike most near-surface methods, which typically assume a 1D subsurface model beneath the stations, seismic interferometry is often applied in contexts where lateral heterogeneities are expected and even desired for imaging purposes. As such, it does not strictly rely on the assumption of lateral homogeneity, and in some applications, the violation of the 1D assumption is not considered a limitation but rather an opportunity to extract structural complexity.

It is worth mentioning that, alongside the methods discussed above, a more recent and different approach has emerged in the literature: ambient noise Full Waveform Inversion (FWI). Rather than relying on the extraction of Green's functions or on the assumption of equipartition of the noise wavefield, ambient noise FWI treats inter-station correlation functions directly as self-consistent observables, modelling the full seismic wavefield for arbitrary noise source distributions (Tromp et al., 2010; Sager et al., 2018). This approach offers the advantage of extracting structural information without requiring stationary, isotropic, or equipartitioned noise conditions, making it potentially more robust in complex environments. However, its application in near-

surface contexts remains computationally prohibitive, as each iteration requires full numerical simulation of wave propagation across the entire domain, a cost that scales poorly with the high frequencies and fine spatial resolution typical of shallow investigations.

In the following sections, the focus will move on two methodological families aimed to extract surface wave dispersion information from ambient seismic noise: **Spatial Autocorrelation** (SPAC) and **Seismic Interferometry** (SI). Although these approaches differ in their formulation and processing domain, indeed SPAC operates in the frequency domain and SI in the time domain, they find their roots in the same fundamental principles: the correlation of ambient seismic noise to isolate coherent features of the wavefield produced by the propagation of surface waves across the stations. Under appropriate conditions, the cross-correlation of continuous noise data between seismic stations can reveal meaningful information about wave propagation, effectively approximating the Green's function (both in frequency and time domain) between the sensors in the absence of any active source. These methods thus provide practical implementations of ambient noise correlation theory and are widely used to extract dispersion curves for several applications. In the next sections, each method is presented through a dedicated theoretical description, aimed at clarifying their physical foundations, highlighting their assumptions and limitations, and illustrating how, despite differing mathematical paths, they often lead to consistent conclusions.

3.2 Correlation of Ambient seismic noise

Correlation of ambient seismic noise has now become a powerful tool for extracting information about the Earth's subsurface without relying on active seismic sources. Several methods, most notably Spatial Autocorrelation-based methods and Seismic Interferometry (SI), take advantage of the continuous wavefield generated by unknown sources to estimate surface wave dispersion. Despite their different formulations, these approaches share a common theoretical background based on the statistical properties of the ambient noise field and the physics of wave propagation in a dispersive medium.

Indeed, although these two methods originate from different conceptual backgrounds, SPAC finds its roots in the frequency-domain array processing of Aki many decades before the conceptualization of Seismic Interferometry, they are now understood to be closely related and, under certain assumptions, mathematically equivalent within the classical Green's function retrieval framework (Tsai & Moschetti, 2010). Specifically, under diffuse-field conditions (Lobkis & Weaver, 2001) the time cross-correlation of ambient-noise recorded at two receivers converges to the time-domain Green's function that assumed one of the two station acting as a source of the signals, whereas the normalised cross-spectrum obtained at the same inter-station offset yields the complex coherence (the cross-spectrum is simply the Fourier transform of the cross-correlation function),

which can be interpreted as the Green's function expressed in the frequency domain and thus carries the dispersion information exploited by the SPAC method.

3.2.1 Spatial Auto-Correlation methods

Among the earliest and most elegant array-based techniques developed to extract dispersion curves from ambient seismic noise is the spatial autocorrelation (SPAC) method. This key insight, introduced by Aki (1957), enabled the direct estimation of Rayleigh wave phase velocities, typically dominant component of ambient seismic noise, by observing how spatial coherence of a signal behaves with respect of receiver spacing. Aki's original formulation laid the foundation for a new class of passive seismic techniques that rely entirely on ambient noise rather than controlled sources. Over the following decades, the SPAC method was refined and adapted for real-world applications. The theoretical basis of the SPAC method, as introduced earlier, relies on modelling the ambient wavefield as a superposition of plane waves propagating horizontally in different directions. At a given frequency, all waves are assumed to share the same phase velocity, although they may have different amplitudes and directions. In addition, it is assumed that waves traveling along different azimuths and at different frequencies are statistically independent. Aki demonstrated that under these strong assumptions, the spatial autocorrelation function from sensors arranged in a circular array is mathematically related to the zero-order Bessel function of the first kind.

To understand the theoretical basics behind the SPAC method, it's easy to begin by modelling a simple seismic wavefield generated by a single monochromatic plane wave and recorded by a seismic station. The displacement at x (which is the location of the seismic sensor) at the time t can be described in its complex form as:

$$u(x,t) = A\, e^{i(\omega t - kx)}, \qquad [3.1]$$

where x is the spatial coordinate of the sensor, ω is the angular frequency of the plane wave, k is the plane wave wavenumber expressed by $\omega/c(\omega)$. A represents the amplitude of the propagating wave with phase velocity c. Considering the x position as the origin of our reference system and given a second receiver at location $x + r$, or simply r, the displacement induced by the same plane wave at $x + r$ can be defined as:

$$u(x+r,t) = A\, e^{i[\omega t - kr\cos(\theta-\varphi)]} \qquad [3.2]$$

In Eq. 3.2 θ represents the azimuthal angle of the propagating wave (that is the direction from which each wave plane arrives), while φ is the azimuth of the vector connecting the two sensors with respect to the north. The phase delay term $[\omega/c\ r\cos(\theta - \varphi)]$ represents the real distance travelled by the plane wave from one sensor to

the other along its propagation direction. It highlights the directional dependence of the observed arrival time differences. The apparent velocity depends not only on the true phase velocity *c*, but also on the angle between the wavefront direction θ and the stations orientation φ.

The complex coherency, as introduced by Aki (1957; 1965), between the stations in *x* and *x* + *r* crossed by a single plane wave can be defined as in Eq. 3.3:

$$\rho(r, \varphi, \omega) = u(0, \omega) \cdot u^*(r\ cos(\varphi), \omega)\ . \qquad [3.3]$$

Where *u*(0) and *u*(*r*cos(φ)) are frequency-domain representation of the ground motion recorded at the two stations at the same time *t*, *r* is inter-station distance and φ denotes the azimuth of the direction between the two receivers (angle with the north) and where * denotes the complex conjugate. Substituting Eq. 3.2 in Eq.3.3, let us to introduce the inter-station coherency:

$$\rho(r, \varphi, \omega) = e^{i\omega r cos(\theta - \varphi)/c}\ , \qquad [3.4]$$

where ρ is the coherency frequency spectrum which contains information on the phase of the recorded signal, that depends on the interstation distance and the direction of wave propagation. That is, Eq. 3.4 represents the spatial correlation between two sensors recording the same plane wave propagating with phase velocity *c* (as function of the angular frequency ω) and with azimuth θ. Assuming a real-world scenario, ambient seismic noise consists of multiple plane waves propagating from all possible directions. In this situation Eq. 3.4 takes the following integral form over all the possible propagating azimuth θ for the seismic noise field:

$$\overline{\rho}(r, \omega) = \int_0^{2\pi} e^{i\omega r cos(\theta - \varphi)/c}\ d\theta\ . \qquad [3.5]$$

We can therefore define $\overline{\rho}(r, \omega)$ in Eq. 3.5 as the spatial autocorrelation function obtained as the spatially averaged coherency over all possible direction of propagation of the plane waves. The above expression suggests also that the integration computed over the azimuth θ led to the same results if computed over the angle between the receivers φ:

$$\overline{\rho}(r, \omega) = \int_0^{2\pi} e^{i\omega r cos(\theta - \varphi)/c}\ d\theta = \int_0^{2\pi} e^{i\omega r cos(\theta - \varphi)/c}\ d\varphi = \overline{\rho}(r, \omega)\ . \qquad [3.6]$$

Given Eq. 3.6, it becomes clear that equivalent results can be obtained either by using an array of sensors that densely sample the entire azimuthal space through multiple sensor pairs recording a single plane wave, or by using only two sensors that record multiple plane waves generated by a homogeneous source distribution all around the stations. Mathematically, in the first case, integration is performed over the azimuthal angle between

the sensors (φ), while in the second case, it is carried out over the propagation azimuths of the incoming waves (θ). Despite their apparent differences, in an ideal setting both configurations should yield the same spatial autocorrelation measurement. This theoretical equivalence highlights the flexibility of the SPAC method and its potential to produce reliable estimates even with simplified sensor setups. However, effectiveness of the SPAC method is strongly influenced by the real azimuthal distribution of ambient noise sources.

Ensuring that the ambient wavefield is sufficiently isotropic is essential for a robust SPAC analysis: when microtremor energy is concentrated along a single azimuth, frequency–wavenumber (f-k) or beam-forming techniques generally outperform SPAC because they resolve both velocity and direction of the signals (Claprood et al., 2011). Conversely, in the more common situation where noise arrives from many unresolved directions, the scalar nature of SPAC, that provides phase velocity estimates that are independent of azimuth, becomes an advantage. As first noted by Aki and later demonstrated in field studies by Asten (1983; 2001), a rich, omnidirectional wavefield enhances SPAC performance, making it particularly effective in urban environments where conventional directional methods often struggle. Studies by Cho et al. (2004) and Asten (2003; 2006) highlighted discrepancies between observed and theoretical coherence spectra when the azimuthal coverage of ambient noise sources is limited (~30°). These findings emphasize the critical role of assessing the directional distribution of the microtremor wavefield as part of the SPAC analysis workflow (Aki, 1957). Ensuring sufficient azimuthal sampling is essential to minimize bias and improve the reliability of SPAC-derived dispersion measurements.

The basic array configuration for the SPAC method consists of multiple receivers positioned uniformly along the perimeter of a circle, with a reference station located at the centre. This layout enables the computation of spatial auto-correlation functions between the central sensor and each surrounding station at a fixed radius. By averaging the spectral correlation results across all station pairs, an approximation of the azimuthal average is obtained in order to effectively stabilizing the estimate of the phase velocity. Naturally, increasing the number of stations along the perimeter improves the sampling of azimuths and leads to a better approximation of the theoretical azimuthal average.

This means that experimentally coherency is computed as the right-hand side of Eq. 3.6, integrating over the station's azimuth. This configuration ensures coverage of the entire azimuthal range, in the way that the more stations surrounding the central receiver, the better the experimental spatial correlation function will approximate the theoretical integral formulation. Complex coherency is computed for each station with the central one and then the total spatial correlation is achieved averaging all the different coherencies. Another array-based approach, the ESAC (Extended Spatial AutoCorrelation) method holds the frequency constant, computing coherency function for different inter-station distance r. This latter approach yields to more accurate results, due to the intrinsic frequency dependence of the phase velocity $c(\omega)$ (Ling & Okada, 1993).

Looking in more detail at the Eq. 3.6 representing the spatial autocorrelation function in the case of homogeneous ambient noise wavefield (considering both integrating factors), it takes the same shape of the integral form of a zeroth order Bessel function of the first kind (Aki, 1957). Consequently:

$$\overline{\rho}(r,\omega) = \int_0^{2\pi} e^{i\omega r\cos(\theta-\varphi)/c}\, d\theta \approx \int_0^{2\pi} e^{iz\cos(\theta)}\, d\theta = J_0[z] \qquad ; \qquad z = \left[\frac{\omega r}{c(\omega)}\right]. \qquad [3.7]$$

This representation, commonly used in plane wave theory, is fully equivalent to the real-valued cosine representation, which corresponds to the real part of the complex phase expression (Asten, 2001; 2006).
This also shows that considering the real part of the complex function in Eq. 3.7, under the assumption of isotropic and stationary ambient noise wavefield, it corresponds directly to the Bessel function $J_0(z)$. All this results in a proportionality between these two functions. Indeed, the spatial correlation function typically accounts only for the real part of the above complex expression, which corresponds to the cosine component. As a result, the formulation simplifies to the Eq. 3.8:

$$\overline{\rho}(r,\omega) \propto J_0\left[\frac{\omega r}{c(\omega)}\right] \approx \int_0^{2\pi} \cos\left(\omega r\cos(\theta-\varphi)/c\right)\, d\theta \qquad [3.8]$$

This formulation can be easily used to retrieve Rayleigh wave phase velocity recorded in vertical component of motion. The same principle behind Eq. 3.8 can be also used to retrieve Rayleigh and Love waves dispersion curve in the horizontal component (Boxberger et al., 2011). In Love's case, although the principle of extracting phase velocities is practically the same, slight changes must be made considering that the horizontal component of the motion in the horizontal direction is a mix of both Rayleigh and Love phases.
Overall, in the theoretical framework of spatial autocorrelation methods, the shape of the autocorrelation function is analytically related to Bessel functions of the first kind. In particular, for surface waves propagating in a homogeneous and isotropic medium, the spatial correlation of seismic signals recorded at two sites depends on the inter-station distance and the angular distribution of wave propagation. SPAC method (which uses a fixed r in a circular geometry with a reference station at the centre) involves in computing frequency cross-correlation function for a large azimuthal distribution of receivers and take the average coherency between all pairs. The solution in terms of phase velocity $c(\omega)$ is commonly found through a grid search procedure, comparing the estimated real spatial average coherency function with the theoretical Bessel function as in Eq. 3.8. The phase velocity $c(\omega)$ is adjusted as the argument of the theoretical Bessel function to minimize the residual error between these two functions. An illustrative example of comparison between experimental coherency function and zero-th order Bessel function is shown in Fig. 3.1. The red dots represent the experimental SPAC coefficients estimated for a single pair of stations, i.e. at a fixed inter-station distance, while the black curve corresponds to the Bessel function that best approximates the observed data, obtained by assuming a phase velocity of c=180m/s. Since the simulated medium is homogeneous and characterised by a constant velocity, the same phase velocity can describe the behaviour of the coherence function over the entire frequency range, in accordance with the assumption of a non-dispersive medium.

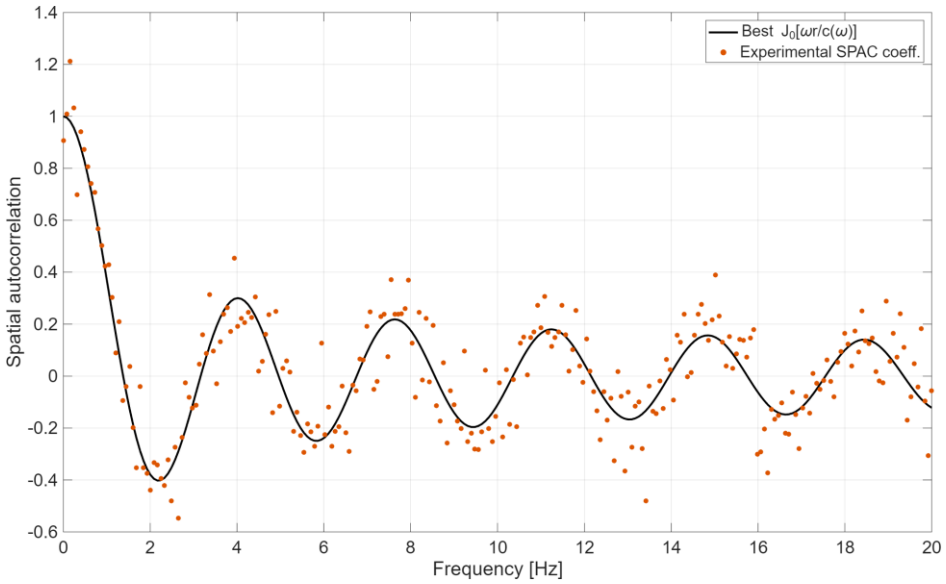

Figure 3.1 - Example of comparison between experimental SPAC coefficients (red dots) and theoretical Bessel function (black curve), calculated considering a homogeneous and non-dispersive medium with constant phase velocity c=180 m/s. Bessel function corresponding to c=180m/s provides the best fit to the experimental data, minimizing the root-mean-square residual error.

The reliability of SPAC-based methods anyway relies on the assumption that the ambient vibration wavefield is both temporally and spatially stationary, allowing it to be modelled as a stochastic process (Aki, 1957). This condition is typically met when recordings are sufficiently long, though generally tens of minutes, and the array aperture samples several possible azimuths of the noise. Ensuring stationarity is critical, as violations of this assumption can lead to time-dependent power spectral characteristics, thereby invalidating the interpretation of microtremors purely as a function of frequency. When ambient vibrations are recorded over extended time periods using a well-distributed sensor array, the wavefield can be treated as a composition of coherent surface waves propagating in multiple directions across a broad frequency range. This foundational assumption must be carefully evaluated in any analysis involving microtremor data Claprood et al. (2011).
In the case of a heterogeneous noise wavefield recorded by a couple of sensors following Cox (1973), the spatial autocorrelation function can be also defined accounting for the azimuthal distribution of the recorded sources:

$$\overline{\rho}(r,\omega) = \int_0^{2\pi} \xi(\omega,\theta)\, e^{i\omega r cos(\theta-\varphi)/c}\, d\theta\,. \qquad [3.9]$$

In this formulation, the integration is made respect to the direction of propagation of the plane waves and an additional term is considered. $\xi(\omega, \theta)$ is a normalized azimuthal density function, that in the case of a diffuse wavefield remains constant for all possible azimuth θ and then can be re-written as a function only dependent on the angular frequency. In the case of a heterogeneous noise wavefield the term $\xi(\omega, \theta)$ is not constant anymore with respect to the azimuth and then it cannot be neglected in the computation of averaged coherency

function. (Asten, 2003; 2004) through numerical approaches simulate spatial autocorrelation, specifically investigating the effects of using a finite number of sensors and an incomplete azimuthal coverage of propagating wave energy. The findings of Asten's numerical simulations indicate that, when the azimuthal distribution of seismic energy exceeds approximately 30°, even simple array geometries, such as a triangular layout, are generally sufficient to provide adequate azimuthal averaging. Under these conditions, SPAC-derived phase velocities remain reliable at least up to the first minimum of the SPAC spectrum across all wave propagation directions, and in some cases even beyond, depending on the azimuth. However, the study also highlights that a limited azimuthal coverage combined with a small number of array stations can introduce significant perturbations in the SPAC function, which may lead to biased phase velocity estimates. The bias becomes particularly pronounced in scenarios where the ambient wavefield is strongly directional, for example, due to nearby machinery or focused sources such as volcanic vents. In such cases, deviations from the theoretical $J_0(\omega r/c)$ curve can be substantial, especially at higher wavenumbers.

3.2.2 Seismic Interferometry

Seismic Interferometry (SI) has greatly increased its popularity in lasts decades following the discovery of the emergence of empirical Green's function from correlations of a diffuse acoustic field (Lobkis & Weaver, 2001) and in helioseismology (e.g., Duvall et al., 1993). Advances in the acoustic field have led many Authors (Shapiro & Campillo, 2004; Snieder, 2004; Campillo, 2006) to use the same principles so as to obtain information on the diffuse and random wave field generated at the surface of the earth, which is obviously composed of waves from every potential direction and with random amplitudes.

SI is therefore based on the assumption that the cross-correlation of ambient seismic noise recorded at two different stations at the same time approximates the Green function between the two stations, giving therefore information about the response of the medium at a source located at one of the receivers recorded at the other receiver. In this context, one receiver acts as a virtual source, while the other serves as the receiving station. The resulting ensemble of cross-correlations is then used to retrieve the inter-station velocity. Differently from frequency-based methods just described and based on spatial correlation, which allow phase velocities to be estimated, this type of approach, based on the desired context and detail, allows both phase and group velocities to be extracted. Mathematically, given two signals $u(x,\omega)$ and $u(y,\omega)$ recorded at two different sites x and y, the normalized cross-correlation function C_{xy}, depending on time and frequency, can be expressed from the integral in Eq. 3.10:

$$C_{xy}(\tau,\omega_0) = \frac{1}{2T}\int_{-T}^{T} u(x,t,\omega_0)\cdot u(y,t+\tau,\omega_0)\,dt, \qquad [3.10]$$

where T is the length of the cross-correlated time window, τ is the time lag between the signals and ω_0 is the considered angular frequency. In real case of a sampled recorded signals the cross-correlation function is not more a continuous function, and the integral is replaced by a normalized summation of the product of samples of both the recorded signals within a shifted time window. Moreover, real recorded signals contain all frequencies combined together and therefore it is the result of an integration also over the frequencies.

Common procedures of seismic interferometry (Bensen et al., 2007) consider long time records (months or years of registration) performed at large scales; indeed, the inter-station distances are often of the order of hundreds of kilometres. For the analysis of ambient noise, the time series are split in thousands of subsequent time windows (generally from tens of minutes up to a few days long) and for simultaneously windows recorded at two different sites, the time cross-correlation is established. It is estimated for each pair of synchronous windows, and then these are averaged in order to achieve the empirical Green's function with low values of the signal to noise ratio (S/N). As with the SPAC method, the main reason why the original signals are subdivided into contiguous (and possibly overlapping) windows of shorter length and then stacked (also refer to it as *ensemble average*) is obviously also to approximate as much as possible the wavefield to be diffuse. This is not true at any time but can be approximated with ambient noise recorded over a very long time.

Early seismic interferometric studies were initially based on the correlation of earthquake coda waves (e.g., Campillo & Paul, 2003), and only later extended to the study of seismic ambient noise (Shapiro & Campillo, 2004). The wavefield associated with earthquake coda is indeed subject to several scattering processes due to small-scale heterogeneity, reverberations and mode conversions of the seismic phases within the subsoil which allow it to be approximated to a statistically random noise. The use of coda waves requires the occurrence of a seismic event which is obviously unpredictable and presupposes good quality recordings at both the stations, even over long distances.

Nowadays, these conditions are more easily found recording ambient noise for long time. The concept of approximate empirical Green's functions cross correlating diffuse random wavefield can be mathematical treated starting from the normalized cross-spectral density between two receivers (Roux et al., 2005):

$$C_{xy}(\omega) = \frac{\sin(\omega \Delta x/c)}{\omega \Delta x/c}. \qquad [3.11]$$

Although different in the formulation, Eq. 3.11 describes the same coherency introduced with the spatial correlation method in Section 3.2.1, it takes the same form of a zero-th order Bessel function of the first kind as in Fig. 3.1. It's easy at this point to provide the expression for the time domain cross-correlation function taking the inverse Fourier transform of the Eq.3.11, which can be simply represented as

$$C_{xy}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_{xy}(\omega)\, e^{i\omega t}\, d\omega\,. \qquad [3.12]$$

This continuous integral can be decomposed into the sum of two integral terms centred at $t = \pm \mathrm{r}/c$, where r is the inter-station distance and c is the wave velocity. These terms represent the time-symmetric arrival of energy from reciprocal paths between the two stations. Introducing Eq. 3.12 into Eq. 3.11 of cross-correlation expression leads to:

$$C_{xy}(t) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega\left(t+\frac{\mathrm{r}}{\mathrm{c}}\right)}}{i\omega r/c} \, d\omega - \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega\left(t-\frac{\mathrm{r}}{\mathrm{c}}\right)}}{i\omega r/c} \, d\omega \,. \qquad [3.13]$$

Starting from the integral expression of the normalized cross-correlation function between two stations separated by a distance r, one can take the time derivative to recover its simplified form. By differentiating each term in the integral with respect to time, the factor $i\omega$ in the denominator cancels with the time derivative of the exponential term. As a result, the integrals reduce to two delta functions (Dirac delta function), each centred at $t = \pm \mathrm{r}/c$ (Roux et al., 2005).

$$\frac{d}{dt} C_{xy}(t) = \frac{1}{4\pi \mathrm{r}/c} \left[\delta\left(t + \frac{\mathrm{r}}{c}\right) - \delta\left(t - \frac{\mathrm{r}}{c}\right) \right]. \qquad [3.14]$$

Physically, these correspond to the symmetric arrival times of the Green's function propagating between the two stations. The two delta functions correspond to respectively the backward and forward Green's function traveling in opposite directions, as expected from the principle of reciprocity in isotropic media.

Physically, this reflects the fact that an impulsive source $\delta(t)$ located at one station would generate an undispersed pulse that arrives at the other station after a travel time r/c. Therefore, the delta functions that appear in the time derivative of the cross-correlation function can be directly interpreted as the Green's functions propagating between the two stations in opposite directions. This formal result underpins the core principle of seismic interferometry: that the correlation of random noise under suitable conditions reveals the impulse response of the medium.

In such an ideal medium, a signal propagating at constant velocity exhibits a temporal correlation composed of two Dirac delta functions at $\pm r/c$, corresponding to the Green's functions in both directions. In the frequency domain, this leads to a normalized cross-spectrum that oscillates as $sin(\omega r/c)/(\omega r/c)$, which corresponds to the real part of the complex coherency discussed earlier.

To give a further illustrative example of the link between the Green's function and cross-correlation, we can give up on the frequency-domain correlation function and instead begin with the ideal case of a single plane-wave of angular frequency ω, propagating in a homogeneous medium and recorded by a single station pair (Tsai, 2009; 2011). The ground motion produced by such wave recorded at x at time t can be described by a monochromatic signal of the form of Eq. 3.15:

$$u(x,t) = A(x,\omega) \cdot \cos(\omega t - kx), \qquad [3.15]$$

where k is the wavenumber defined as $k = \omega/c(\omega)$, c is the phase velocity of the propagating wave and $A(x, \omega)$ defines the amplitude term that is inversely proportional to the square root of the considered inter-station distance. The same plane wave recorded at y is represented by:

$$u(y,t) = A(y,\omega) \cdot \cos[\omega(t + t_d) - kx] \,. \qquad [3.16]$$

Eq. 3.16 represents the same formulation for ground motion recorded in x with only additional term represented by the time delay t_d between the arrival time at the stations. The time delay takes negative values when the plane wave propagates from y to x, while it takes positive values when the wave propagates from x to y. The time delay can be also expressed as function of the phase velocity c, the inter-station distance r and the angle of propagation of the wave θ. In that case the Eq. 3.16 becomes:

$$u(y,t) = A(y,\omega) \cdot \cos\left[\omega(t + rcos(\theta)/c) - kx\right] \qquad [3.17]$$

Expression in Eq. 3.17, assuming a propagating plane wave, the time delay between the two stations depends on the azimuth of the wave (θ). The term describing the time delay assumes different values, the maximum value when the wave propagates parallel the stations direction and the minimum value when the wave propagates orthogonal to the station's direction (time delay goes to zero and the receivers record the signal at the same time). At this point we might put Eq. 3.15 and 3.16 describing ground motion in x and y within the definition of the cross-correlation to obtain:

$$C_{xy}(\tau) = \frac{A(x)A(y)}{2} \int_{-T}^{T} \cos(\omega\tau - kx) \cos[\omega(\tau + t + t_d) - kx] \; d\tau \,. \qquad [3.18]$$

In Eq. 3.18 all the information about the amplitude of the signal, and then the information about attenuation of the signal between the sensors is out of the integral and it's totally contained in the term containing $A(x)$ and $A(y)$. By applying standard trigonometric identities, the product of cosines can be rewritten as a sum of cosines, leading to two terms: one that oscillates at frequency ω and another at frequency 2ω (see Boschi et al., 2013). The second term contains rapid oscillations and becomes negligible when integrated over enough long-time windows. Under this approximation, and after simplifying the result, the expression converges to Eq. 3.19 (Eq. 12 of Boschi et al., 2013), which highlights the dominant coherent contribution to the cross-correlation.

$$C_{xy}(t) \approx \frac{A(x)A(y)}{2} \cdot \cos[\omega(t + t_d)] \,. \qquad [3.19]$$

In the case of a single, localized source propagating in a 2D space without attenuation or scattering, the cross-correlation $C_{xy}(t)$ reduces to a cosine function centred at $t = -t_d$ if $t_d > 0$, or $t = t_d$ if $t_d < 0$. This behaviour reflects

the time delay associated with the wave traveling from one receiver to the other that depends on the wave propagation azimuth. The resulting signal can thus be interpreted as the response at one station due to a theoretical sinusoidal source located at the other station. Because the resulting cross-correlation is a purely cosine function, it will naturally exhibit multiple periodic peaks. However, in practical applications, particularly when considering a finite-time cross-correlation, the dominant peak will appear at $t = -t_d$, while the surrounding oscillations will exhibit decreasing amplitude symmetrically around this central lag. This decay reflects the finite duration and bandwidth of the recorded signal and highlights the temporal localization of coherent energy between the two receivers. In Fig. 3.2, the amplitude term which contains the amplitude information about the signal recorded at the two points is neglected and considered equal to 1, the time delay t_d between the two sensor is 0.5 and the simulated frequency is 1Hz. It's clear as the peak coincides with $t = -t_d$ which represents a wave propagating from the station in x to the station in y.

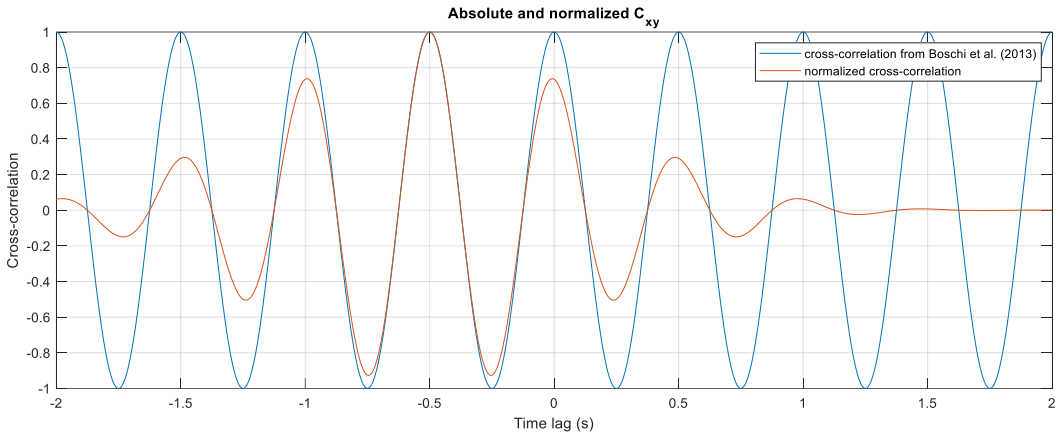

Figure 3.2 - Example of cross-correlation function for a single wave propagating in a 2D homogeneous medium with no attenuation or dispersion. The blue curve shows the cross-correlation function computed following Boschi et al. (2013), while the red curve represents the normalized cross-correlation. The dominant peak occurs at $t=-t_d$, corresponding to the propagation travel-time of the wave from station x to station y. The symmetrical decay of the surrounding oscillations reflects the finite duration and bandwidth of the signal. In this example, the amplitude term is set to one, the time delay between stations is td=0.5 s, and the simulated frequency is 1 Hz.

Even working in time domain, a clear example of two stations recording a single plane wave is crucial to understand what mathematically happens in the hypothesis of diffuse wavefield generated by a homogeneous distribution of sources around the stations. In this case the ambient noise can be assumed to be a cumulative effect of several sources. So, Eq. 3.15 and 3.16 of the ground motion recorded at sites x and y become:

$$u(x,t) = \sum_i A_i(x,\omega) \cos(\omega t - k_i x)\,, \qquad [3.20]$$

$$u(y,t) = \sum_i A_i(y,\omega) \cos[\omega(t + t_d^i) - k_i x)]\,. \qquad [3.21]$$

Both recorded ground motions can be expressed as a sum over the entire number of sources *i*. t^i_d represents the time delay between the two stations related to the i-th source and, as before in the case of single plane wave, it depends on the propagation azimuth of the wave. At this point we can again substitute the general equation of cross-correlation with these two last equations, obtaining Eq. 3.22 that describes the cross-correlation function of two different seismic stations recording seismic noise generated by multiple sources.

$$C_{xy}(t) = \frac{1}{2T} \sum_{i,j} \left\{ A_i(x) A_j(y) \int_{-T}^{T} \cos(\omega\tau - k_i x) \cos\left[\omega\left(\tau + t + t_d^j\right) - k_j x\right] d\tau \right\}. \qquad [3.22]$$

In Eq. 3.22, both *i* and *j* indices refer to discrete monochromatic sources emitting at the same frequency. The recorded wavefield is modelled as a superposition of these sources, each contributing to the signals at stations in *x* and *y*. Iteratively, each contribution is composed of the product of the cosines terms where for each contribution the first cosine term refers to the source *i* and second refers to source *j*. The terms $A_i(x)$ and $A_j(y)$ represent the amplitudes of the signals from sources *i* and *j* as recorded at the two stations. The total cross-correlation is computed as a sum over all sources.

When evaluating the full expression for the noise cross-correlation under the assumption of a discrete set of monochromatic sources, it becomes evident that only the terms involving the same source ($i = j$) contribute constructively to the result. In contrast, the terms corresponding to different sources ($i \neq j$) include rapidly oscillating cosine functions with different phase delays and arrival times, which vary from source to source. As a result, their integration over a sufficiently long-time window leads to cancellation due to the lack of coherent phase alignment. This implies that the cross-correlation function is primarily shaped by coherent arrivals from the same source recorded at both stations, whereas incoherent contributions from different source pairs effectively average out to zero. The resulting cross-correlation function corresponding to a set of uniformly distributed sources is composed of a sum of terms of the form of Eq. 3.23:

$$C_{xy}(t) \approx \sum_{i} \frac{A_i(x) A_i(y)}{2} \cos\left[\omega\left(t + t_d^i\right)\right]. \qquad [3.23]$$

The above cross-correlation represents the same cross-correlation form of a monochromatic wave from a single source (Eq. 3.19), except for the sum of terms that represents the contribution of the whole set of distributed sources.

In this discrete representation of ambient noise cross-correlation, each source generates a signal characterized by a specific propagation direction and contributes individually to the total correlation. The amplitude terms denote the energy of the signal generated by the *i*-th source as recorded at stations *x* and *y*, respectively, while t^i_d represents the interstation travel-time delay associated with the wave's propagation from that source.

In the case of continuous distribution of sources, it is clear that the above expressed sum over the sources becomes an integration over all the possible azimuth of the propagating waves from 0 to 2π. In this continuous

context, the discrete amplitude term is also replaced by a continuous function that describes the distribution of the energy arriving with time delay t_d and angular frequency ω. This new function can be defined as function of t_d and ω and it reflects the number of sources that returns a specific time delay, as well as the average amplitude of their contributions to the recorded wavefield:

$$C_{xy}(t) \approx \int_{-\frac{r}{c}}^{\frac{r}{c}} \xi(t_d, \omega) \cos[\omega(t + t_d)]\, dt_d \,. \qquad [3.24]$$

The integration is made within the limits -r/c and r/c that represent sources producing the same time delay but with different sign, or then to waves with same direction but opposite arrival angle. As done before, the same continuous equation can be expressed instead of respect to the time delay, as function of the azimuthal angle of arriving of the waves.

$$C_{xy}(t) = \int_0^{2\pi} \xi(\theta, \omega) \cos\left[\omega\left(t - \Delta x \frac{cos\theta}{c}\right)\right] d\theta \qquad [3.25]$$

In this formulation the source density function (that can be seen as the azimuthal density function introduced in Eq. 3.9) describing the distribution of the energy is expressed as function of the azimuth θ and t_d becomes $r\cos(\theta)/c$. Assuming an isotropic ambient noise field, the distribution of sources is considered uniform with respect to the azimuth. This implies a constant source density function $\xi(\theta,\omega)$ defined over the full azimuthal range $[0, 2\pi]$. While this uniformity holds in azimuth, the corresponding distribution with respect to interstation time delay $\xi(t_d,\omega)$ is not constant anymore, due to the nonlinear relationship $t_d(\theta) = r\cos(\theta)/c$. As a result, the delay-dependent source density function must be derived through a change of variables that accounts for the compression and expansion of azimuthal intervals.

The final expression of the cross-correlation function can be naturally decomposed into two parts: a causal component corresponding to positive interstation delays ($t_d > 0$) and an anti-causal component for negative delays ($t_d < 0$). In the ideal case of an isotropic noise field, where source contributions are uniformly distributed in azimuth and temporally uncorrelated, the two parts are expected to be symmetric with respect to zero, producing a cross-correlation function that is even in time.

So, following Boschi et al. (2013), and Tsai (2009) before them, the cross-correlation function obtained from two seismic stations recording a set of homogeneous distributed cosine sources generating plane waves can be written, at a fixed frequency ω, as in Eq. 3.26:

$$C_{xy}(t, \omega) = \int_{-r/c}^{0} \xi(t_d, \omega) \cos[\omega(t + t_d)]\, dt_d + \int_0^{r/c} \xi(t_d, \omega) \cos[\omega(t + t_d)]\, dt_d \,. \qquad [3.26]$$

Each term represents the contribution of plane waves arriving from opposite azimuth. When the source distribution is symmetric with respect to $t_d = 0$, so that $\breve{\xi}(t_d, \omega) = \breve{\xi}(-t_d, \omega)$, this corresponds to the assumption of an isotropic azimuthal distribution of ambient noise sources, and then the two integrals are mirror images, and the resulting cross-correlation function becomes symmetric with respect to time:

$$C_{xy}(t) = C_{xy}(-t)\,. \qquad [3.27]$$

Eq. 3.26 has the same structure as Eq. 3.13 and 3.14 provided by Roux et al. (2005), although it is obtained from a different source model. In both cases, the cross-correlation function can be broken down into two symmetrical contributions associated with opposite propagation directions, which give rise to arrivals centred at times t=±r/c. This formal equivalence highlights how, regardless of the specific description of the sources, by averaging a sufficiently large set of cross-correlation functions, the resulting Green's function reflects the average properties of the ambient noise field and can therefore be interpreted as a representative estimate of the medium response between the two stations. The average cross-correlation contains, in principle, all the information necessary to characterize the dispersive behaviour of surface waves dominating the recorded signal. As mentioned above, it can provide both phase velocity and group velocity information.

The phase velocities information is contained in the phase of the cross-correlation function, which describes the phase differences between the signals recorded at the two stations at the different frequencies investigated, given the inter-station distance. In this context, the transition to the frequency domain using the Fourier transform of cross-correlation allows to analyse the phase difference spectrum of the function and estimate the phase velocities starting from the phase spectrum as a function of frequency.

However, in practical applications, once empirical Green's function has been estimated, it is often preferred to operate in the time domain in order to retrieve group velocities. The total cross-correlation function (Eq. 3.26 integrated also over the frequencies) indeed represents the sum of the individual cross-correlations associated with the diverse harmonic component of the recorded surface waves, and each of them should show an amplitude peak in the cross-correlation depending on their propagation velocity. Since the resulting function incorporates contributions from a wide frequency range, a multi-filter analysis allows to isolate the arrival times relative to specific spectral component. The relationship between inter-station distance and estimated arrival times thus provides the propagation velocities as a function of frequency, the group velocity dispersion curve.

3.3 SPAC and SI in a two-station application

In practice, SPAC and SI follow rather different workflows even though they rest on the same theoretical foundation on the ambient noise as a stochastic and stationary process. SPAC method exploits these properties

in the frequency domain: sensors are arranged on a geometrically isotropic array, typically one or more concentric rings, and the complex coherency is averaged over all azimuths in order to obtain the spatial autocorrelation describing the site in a 1D approximation (Aki, 1965). SI, on the other hand, is applied in its original form as a two-station technique. The temporal stacking replaces the azimuthal integration required by SPAC approach: as noise sources migrate in space and time, their cumulative contribution should approximate a homogeneous distribution, so the stacked cross-correlation yields the full Green's function (Shapiro & Campillo, 2004). Tsai & Moschetti (2010) show that, far from being alternative approaches, these strategies are complementary. SPAC provides a compact workflow that directly returns phase-velocity estimates when a dense, isotropic array can be deployed, and the noise field is reasonably omnidirectional. SI, by contrast, is more tolerant of arbitrary station geometries and gains its robustness from long recording times and temporal averaging, at the cost of additional preprocessing and stacking. Despite these differences, both approaches rest on the same physical principle and, given the ideal assumptions of isotropy and stationarity of the noise sources, converge on identical dispersion curves if just two stations are implemented. Both approaches provide phase velocity estimations derived from the assumption of an azimuthal average of the propagation properties of the medium. Furthermore, the SPAC method applied to only two stations does not require strong approximations regarding the sampled wavelengths and works also in contexts where the heterogeneities have a size comparable to that of the wavelength investigated (Ekström et al., 2009).

From a mathematical point of view, the close relationship between the SPAC approach and seismic interferometry emerges clearly by comparing the expressions describing the spatial autocorrelation and the cross-correlation function for a fixed frequency (Tsai & Moschetti, 2010).

In particular, Eq. 3.9 that describes SPAC method can be interpreted as a special case of Eq. 3.25 relative to the interferometric formulation, obtained by evaluating the cross-correlation function at zero-time lag (t=0), in that case the temporal term goes to zero and the cross correlation and real part of the spatial correlation functions for a fixed frequency share the same form. Under the ideal assumptions of stationarity, isotropy and lateral homogeneity, both expressions yield identical phase-velocity information: the time dependence of the cross-correlation disappears, and the resulting expression reduces to a spatial function, which depends exclusively on the inter-station distance and phase velocity.

Cross-correlation at t=0 takes the form expressed in Eq. 3.28:

$$C_{xy}(0) = \int_0^{2\pi} \xi(\theta, \omega) \cos\left[\omega\left(\frac{\omega r cos\theta}{c}\right)\right] d\theta. \qquad [3.28]$$

Now considering Eq. 3.9, assuming φ=0 and a distribution of sources consistent with an isotropic and stationary noise field, the expression of the cross-correlation evaluated at t=0 is formally identical to the real part of the SPAC function.

$$\Re[\,\overline{\rho}(r,\omega)\,] = \int_0^{2\pi} \xi(\theta,\omega)\ \cos\left[\omega\left(\frac{\omega r cos\theta}{c}\right)\right] d\theta\ . \qquad [3.29]$$

Under these conditions, both formulations depend exclusively on the inter-station distance and phase velocity, demonstrating that under ideal assumptions the two approaches applied at two stations share the same information content.

Recently, increasing attention has been devoted on extracting surface wave dispersion information using two-station approaches applied to ambient noise records. This approach aims to infer subsurface information from just two sensors, with the possibility of expanding the data using multiple receivers in a tomographic framework.

Seismic interferometry was first conceived as a *two-station* technique that retrieves the inter-station Green's function by cross-correlating long records of ambient noise essentially for large-scale tomographic purpose (e.g., Campillo & Paul, 2003 and Shapiro & Campillo, 2004). Since then, numerous studies have also focused on adapting Aki's spatial autocorrelation to the two-station configuration (Ekström et al., 2009; Ekström, 2014; Ikeda et al., 2021), developing and applying SPAC-based methods for the same purpose at many different spatial scales (Gouédard et al., 2008; Picozzi et al., 2009; Pilz et al., 2012; Boschi et al., 2013).

As discussed earlier, even SPAC theory can converge on the interferometric solution if its spatial average is replaced by a sufficiently long temporal average, allowing the method to operate on a single receiver pair.

In a two-station method, the distance between the stations is clearly fixed and this significantly limits the resolution of both spectral and time analysis. The ideal azimuthally homogeneous distribution of sources is theoretically obtained by collecting continuous data for long time, which allows receivers to capture seismic waves propagating from multiple directions. However, the inter-station distance r plays a crucial role in the reliability of the resulting dispersion curve, particularly for long periods.

Focusing on SI, the analysis of ambient seismic noise between station pairs can lead to both group and phase velocity estimation depending on the processing approach. Many studies focus on group velocities through the analysis of the time-frequency evolution of the signal envelope, using techniques such as multiple filter analysis (Dziewonski et al., 1969). In these applications, once the ensemble cross-correlation is achieved, and the obtained Green's function is filtered for subsequent narrow-band, for each filtered function the peak in the envelope is expected to represent the first arrival time of the propagating wave with wavelength $\lambda = 1/f_c$, where f_c is the central frequency of the filter. In practical applications, selecting the peak of the envelope of the filtered cross-correlation function is easier than isolating individual waves at a fixed frequency, as is done in phase velocity analysis. However, this approach often results in less accurate velocity estimates compared to phase velocity curves. In Fig. 3.3 clear example of filtering process outcomes is shown. When observing the subsequent peaks in the envelope of the filtered cross-correlations, it appears clear that the frequencies are propagating with different group velocities, however, within the envelope that incorporates more frequencies, individual waves propagate with their own phase velocity (Brillouin, 1960). By converting the travel-time, that

correspond to the peak in the cross-correlation, into velocity through the inter-station distance, it is possible to reconstruct a frequency-velocity relationship, that is the group dispersion curve.

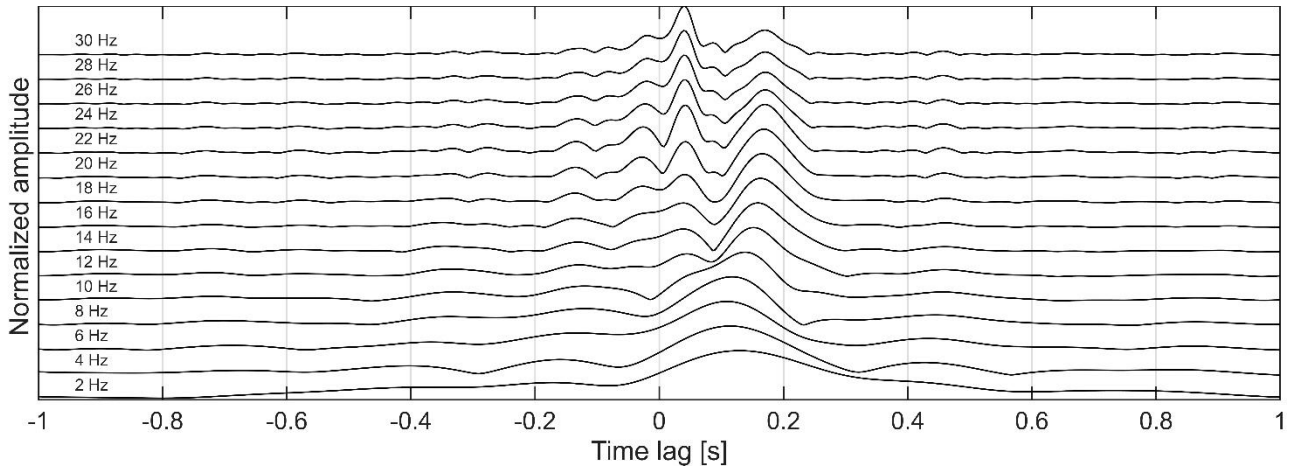

Figure 3.3 - Example of multi-filter analysis applied to the ensemble cross-correlation function. The figure shows a series of narrowband filtered cross-correlations at increasing centre frequencies, plotted with vertical offset. The different time lag of the envelope peaks with frequency highlights the dispersive behaviour of surface waves, with different frequency components propagating at different group velocities. The peak of each filtered envelope is associated with the travel time of the corresponding frequency band and can be used to estimate the group velocity.

The advantage of working in the time domain to retrieve group velocity estimation allows for the complete avoidance of the complex spectral phases of the signals. On the contrary, retrieving phase velocities within a two-station approach is more challenging and, even if it turns out to be more informative than group velocities, extraction of surface waves phase velocities is always subjected to enormous complications due to the intrinsic phase ambiguity in the choice of the true phase velocity (Magrini et al., 2020; 2022).
The concept of phase ambiguity problem can be clearly illustrated within either time- or frequency-approach: in the former case, once Empirical time Green's function is obtained, all information about the phase differences for the sampled frequencies are included in the phase spectrum of the average cross-correlation.
Applying the Fourier transform to the stacked cross-correlation aim to retrieve the cross-spectrum of the two recorded signals and therefore the phase difference spectrum is expressed by:

$$\Delta\phi(\omega) = \tan^{-1}\left[\frac{\Im\left(C_{xy}(\omega)\right)}{\Re\left(C_{xy}(\omega)\right)}\right] \qquad [3.30]$$

In Eq. 3.30, $\Re(\,)$ and $\Im(\,)$ represent respectively the real and imaginary part of the cross-spectrum. As a result, given the nature of the arctangent function, the raw phase spectrum is discontinuous and defined only in the range $-\pi$ to $+\pi$, and therefore it should be unwrapped to obtain the real phase difference value for each frequency. Operationally, however, phase unwrapping is prone to fail when the signal-to-noise ratio is low and, even in presence of well-quality data, the ambiguity must be handled explicitly by recognising that each

frequency-dependent phase shift is correct only up to an integer multiple of 2π. Accordingly, the phase-velocity can be expressed as:

$$c(\omega) = \frac{\omega \mathrm{r}}{\Delta\phi(\omega) + 2\pi n}. \qquad [3.31]$$

In the above equation r is the inter-station distance and n defines the possible 2π phase wraps (Verbeke et al., 2012). By evaluating Eq. 3.31 over a suitable range of n values, e.g. from −4 to +4, one obtains a set of potential dispersion curves, some of which are more plausible than others. Fig. 3.4 shows a representative example: all candidate curves are plotted, highlighting the complexity introduced by the phase ambiguity and the need for solid constraints in the selection of the true curve.

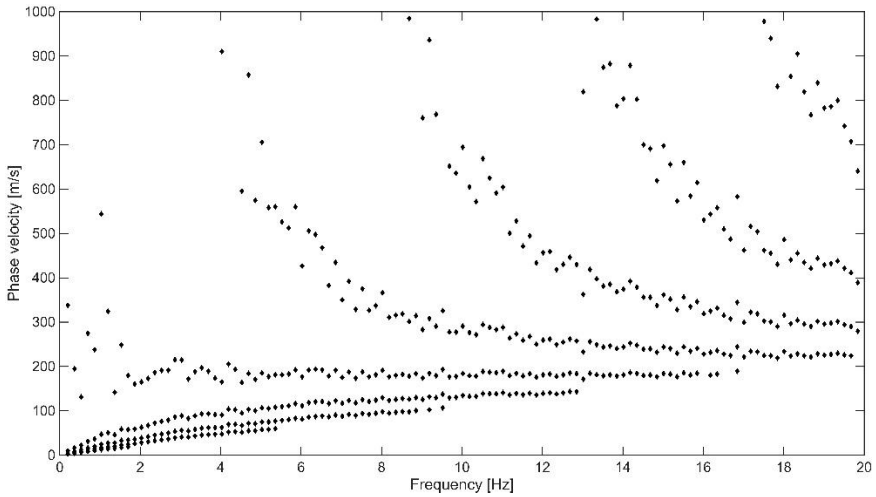

Figure 3.4 – Set phase velocity dispersion curves obtained for different values of the integer parameter n in Eq. 3.31, showing the strong phase ambiguity which the two-station approach involves. The resulting curves represent alternative solutions rather than higher modes, and only one corresponds to the true phase dispersion curve.

The same phase ambiguity becomes evident when the two-station dispersion analysis is performed based on the Aki's spectral formulation. After the time-stacking procedure, the experimental real coherency function (Fig. 3.1) for a single station pair is fitted to the theoretical curve $J_0(\omega r/c)$.

Because of the oscillating nature of both experimental and theoretical functions, for a fixed frequency (or for a specific frequency range) the fit can be satisfied by diverse phase velocity values: every time the oscillatory real part $\Re(\rho)$ happens to coincide with a lobe of the Bessel curve, an additional solution appears (Ekström et al., 2009). In practice, this means that the fit at one frequency does not yield a unique velocity but a set of possible phase velocities, each separated by a full 2π phase cycle, exactly the ambiguity detected before with the arctangent function.

One pragmatic way to simplify the choice is to focus not on the function as a whole, but instead in its zero-crossing points, since noise rarely shifts the position of a zero point of the coherence function.

This practical criterion usually reduces the ambiguity of the solution to a reduced set of plausible dispersion curves; Kästle et al. (2016) report that the zero-crossing rule can, in many data sets, halve the number of admissible phase-velocity curves. However, even zero-crossing method is far from eliminating the underlying ambiguity in selecting the right dispersion curve. Moreover, the use of only the zeros of the function results obviously in a lower frequency resolution of the dispersion curve, furthermore the number of zeros extremely depends on the velocity involved and the distance between the stations. On smaller scales or with relatively low phase velocities, the number of zeros decreases dramatically, and the zero-crossing method also tends to fail. Furthermore, even in ideal contexts, zero crossing is also susceptible to jumps between different curves along frequencies. These cases demonstrate that, whether the analysis is carried out in the time domain (via cross-correlation) or in the frequency domain (via Bessel-function fitting), the 2π-periodicity of phase information inevitably introduces non-uniqueness in the resulting phase-velocity estimates.

For large-scale phase dispersion analysis, this intrinsic ambiguity is typically resolved by using a reference dispersion curve (e.g., Dziewonski & Anderson, 1981) as a target and selecting the phase velocity curve that is closest to the target curve (e.g., Verbeke et al., 2012 and Boschi et al., 2013). However, at smaller scales, this approach is rarely feasible due to the absence of a reference curve, and alternative methods must be employed to identify the correct dispersion curve from among the multiple possible solutions.

Furthermore, working at regional or global scales often allows for the processing of a large amount of data from long time-series records of permanent seismic stations. In contrast, at smaller scales (on the order of tens to hundreds of meters), permanent seismic stations are generally unavailable, and thus only short time-series records can be analysed to extract dispersion information about the subsurface.

However, moving from regional to local and engineering scales, several works show that reliable dispersion information can be obtained also from short ambient-noise records (on the order of tens of minutes): in time domain, this has been demonstrated by studies that provided group-velocity curves from stacked empirical Green's functions (e.g., Picozzi et al., 2009 and Szanyi et al., 2016), while in frequency domain, phase-velocity retrieval from interstation coherences has been stabilised with **zero-crossing** picking to mitigate phase ambiguity (e.g., Pilz et al., 2012). Taken together, these results indicate that, even without the long acquisitions typical of regional surveys, careful stacking and quality control can yield dispersion measurements, and corresponding tomographic images, of sufficient quality for local-scale investigations.

In this Chapter we have shown that frequency-domain SPAC and time-domain Seismic Interferometry arise from the same assumptions about ambient noise and converge more or less on the same dispersion information when applied as two-station procedures. Despite their different workflows (azimuthal averaging for original SPAC; temporal stacking for SI), they are best viewed as complementary: each compensates for limitations of the other in terms of array geometry, data length, and directional bias: original SPAC method work based on the assumption the subsoil structure is purely 1D and that no lateral variations exist in terms of velocity. In contrast SI, even if is based on the extraction of simple 1D dispersion curves from couple of stations, has been widely used to retrieve lateral heterogeneities both at small- and large-scale applications. Despite these

differences, we've also seen that in the application of two-station method, under the assumption of a diffuse, stationary and stochastic noise wavefield, both techniques should lead to the same results in terms of inter-station phase velocity. We have also clarified why phase selection is intrinsically ambiguous in both domains and outlined practical ways to constrain it. Nevertheless, this ambiguity remains a major effort and is only partially mitigated in practice; the most reliable constrain typically comes from guiding the phase curve selection with an external reference dispersion curve. Taken together, these considerations establish a coherent processing pathway that is applicable at both regional scale and the local/engineering scales of primary interest here, even when recording durations are modest.

Building principles introduced in this Chapter, Chapter IV details both components of the workflow: the automatic retrieval of inter-station phase dispersion curves, (and group curves) and the tomographic inversion yielding surface-wave velocity maps at different frequencies.

Chapter IV.
Inverse problems in surface-wave analysis

In the following section, the fundamental concepts associated with inverse problems will be outlined, from the basic distinction between linear and nonlinear problem formulations, as well as the classification of problems according to their degree of determination (under-, over-, or well-determined systems) and the need for regularization and prior information in the search for a stable solution to the problem.
Emphasis will then be placed on the specific application of inverse problem theory first to seismic tomography, a classic linear inverse problem, and in the optimization of a 1D shear-wave velocity profile starting from observed phase dispersion curve, which is instead a common non-linear problem.
The discussion is framed within the general theoretical background provided by common references in the field of inverse problem (Menke, 2012; Aster et al., 2018; Fichtner, 2021), with attention to the probabilistic and model-based perspectives introduced by Tarantola & Valette (1982) and Tarantola (2005). These works constitute a foundational framework for understanding how observations can be used to retrieve model parameters, and how uncertainty, data noise, and model regularization affect the stability and resolution of inverse solutions in geophysical applications.

4.1 Introduction to inverse problem theory

A generic inverse problem involves three fundamental components: a set of observed data, a set of model parameters, and a physical relationship linking the two objects. When solving an inverse problem, the aim is, given the effects produced by a determined phenomenon (observed data) within a physic model (physical relationship), to retrieve the causes that produced that phenomenon (model parameters).
The model parameters, which are the unknown data in the inverse problem, represent physical properties of the system that we seek to infer. The physical relationship is typically expressed in the form of set of governing equations or numerical models that allow the prediction of the observed data given a specific set of model parameters.
The estimation of model parameters from observed data is equivalent to resolve an inverse problem, while the computation of synthetic data from known model parameters means to resolve a forward problem (Fig. 4.1). Solving an inverse problem inherently requires the ability to solve the corresponding forward problem.

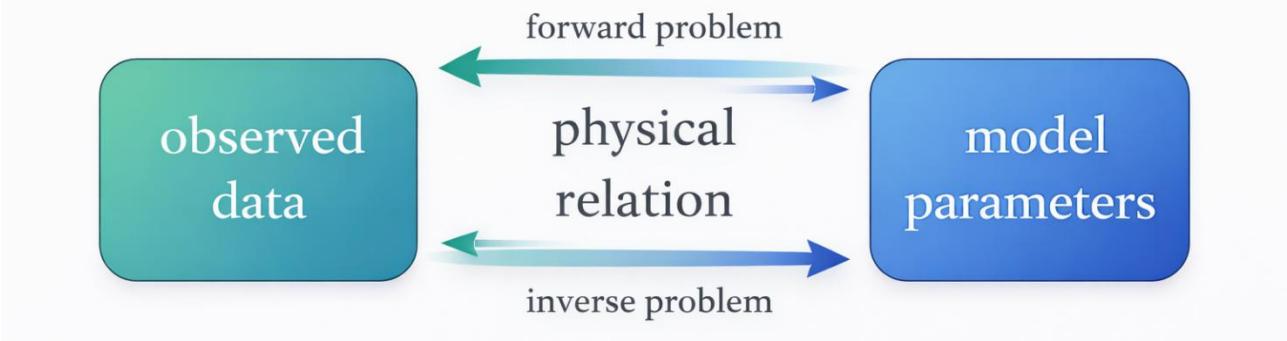

Figure 4.1 – Basic concept about forward and inverse problem. Moving from observed data to model parameters is equal to solve an inverse problem, while producing synthetic data through known model parameters involves in a forward problem.

For simplicity, the theoretical basis of inverse problems will be described making use of compact mathematical notation. Formally, if we denote *m* as the model vector containing the model parameters and *d* the vector containing the observed data, the basic inverse problem can be expressed as:

$$d = G(m), \tag{4.1}$$

where d is a column vector defined as $d = [d_1, d_2, d_3, \dots, d_N]^T$, m is the model column vector $m = [m_1, m_2, m_3, \dots, m_M]^T$ and G represents the forward operator, which may be both linear and nonlinear depending on the physical context. N and M are real integer numbers representing the total number of observations and the total number of model parameters, and the superscript 'T' denotes the vector transpose.
A more general class of inverse problems, however, involves implicit formulations of the type $G(m, d) = 0$, for which no explicit expression of the forward operator exists. This occurs, for instance, when the relationship between model and data is defined through a system of differential equations, where model parameters and observed data appear simultaneously as unknowns. In such cases, the forward problem cannot be separated from the inverse problem in a straightforward manner.
Given this general formulation, an inverse problem can be broadly classified as linear or nonlinear, depending on the nature of the forward operator *G*. In the simplest case of a linear inverse problem *G* is a linear operator, and the physical relationship involves in a simple matrix multiplication between elements of *G* and elements of the vector *m,* and the problem can be simply re-written as:

$$d = Gm \tag{4.2}$$

Linear inverse problems expressed in this form represent the most fundamental and well-characterized class of inverse problems. A wide range of significant problems in the physical sciences, including most seismic tomography problems, can be directly formulated in this form and even those governed by more complex relationships are frequently subjected to linearization, allowing similar solution strategies to be applied.
If we expand the compact formulation of Eq. 4.2, it takes the form of a simple system of linear equations of the type:

$$\begin{aligned} d_1 &= G_{11}m_1 + G_{12}m_2 + \cdots + G_{1M}m_m, \\ d_2 &= G_{21}m_1 + G_{22}m_2 + \cdots + G_{2M}m_m, \\ &\dots \\ d_N &= G_{N1}m_1 + G_{N2}m_2 + \cdots + G_{NM}m_m, \end{aligned} \tag{4.3}$$

that can be also represented in the form:

$$d_i = \sum_j G_{ij} m_j, \qquad i = 1,2,\dots,N; j = 1,2,\dots,M. \tag{4.4}$$

Looking at Eq. 4.4, it's also clear that if the considered inverse problem consists of N observations and M model parameters, the forward matrix G will be composed of N rows and M columns.
Considering instead a non-linear problem, the forward operator G is non-linear itself and it does not represent a simple mathematical operation such as multiplication, but it is usually characterized by physical relationship governed by complex non-linear equations. So, G is no more a simple multiplicator matrix and the system in Eq. 4.3 can be re-written as below:

$$\begin{aligned} d_1 &= G_1(m_1, m_2, \dots, m_M), \\ d_2 &= G_2(m_1, m_2, \dots, m_M), \\ &\dots \\ d_N &= G_N(m_1, m_2, \dots, m_M). \end{aligned} \tag{4.5}$$

A common strategy for tackling nonlinearity is linearization: the nonlinear forward operator is locally approximated by its first-order Taylor expansion around an initial guess model m_0. This allows the use of linear techniques nearby m_0 and, within an iterative process, updating the model step by step until convergence is achieved. Understanding whether a problem is linear or nonlinear, and to what degree, is fundamental in choosing the appropriate inversion approach and in assessing the expected convergence behaviour and computational cost. Several inverse problems are actually nonlinear and of course, the consequence of nonlinearity of the problem is that you cannot collect the Gm products into a single kernel matrix anymore: you must specify each $G_i()$ explicitly, and the solution strategy will typically require iterating a forward solver plus some form of local linearization (e.g. Newton, Gauss–Newton, quasi-Newton, etc.) to update the model m.

In general, when dealing with an inverse problem, given a guess starting model and given the physical relationship expressed by the forward operator *G*, predicted data *d* can be computed. This represents the first step of any inversion procedure, which practically involves solving the forward problem for an a priori assumed model. At this point one has two different kinds of data: the observed d_{obs} and the predicted *d* given the guess model m.
The aim of most inversion procedures is to minimize a misfit function, also called objective function or cost function, measuring the discrepancy between observed and predicted data. It should be noted, however, that a good fit to the observed data does not necessarily imply that the recovered model reflects the true properties

of the real system, as inverse problems are often affected by non-uniqueness. Alternative frameworks also exist, most notably the probabilistic or Bayesian approach, in which the goal is not to find a single optimal model but to explore the range of models that are consistent with the observed data, naturally incorporating uncertainty quantification.

In general, a perfect match between observed data and predicted data cannot be achieved. The differences observed arise both from measurement errors in the experimental data and from structural limitations of the model, which inevitably represents a simplification of the real studied phenomenon. Contrary to what might be intuitively expected, within the overall inversion procedure, solving the forward problem is often the most computationally intensive and mathematically complex step, far more so than the inversion steps that follow. Solving a forward problem often requires addressing complex physical formulations, such as systems of partial differential equations, integral equations, or algebraic systems, depending on the nature of the investigated phenomenon. Analytical solutions are rare in realistic settings, and numerical methods such as finite difference, finite element, or spectral approaches are generally employed. Despite being conceptually direct, the forward problem is not immune to difficulties. It may involve modelling approximations, discretization errors or incomplete knowledge of the physical system. Furthermore, the computed solution can be sensitive to these uncertainties, especially in nonlinear systems, in the way that errors and approximations related to the forward problem must necessarily affect the quality of the solution to the inverse problem and the ease with which the solution is found.

4.1.1 Properties and characteristics of inverse problems

A key concept in inverse theory is the distinction between the data space and the model space. The data space refers to the set of all possible measurements that can be made, typically represented by the vector $d \in R^N$,where N is the number of observations. The model space, on the other hand, includes all possible configurations of the unknown parameters, expressed as $m \in R^M$, with M the number of model parameters.

Practically, the dimensions of these spaces are often different and the rate between the number of unknowns and the number of observations (M and N) is a critical aspect in inversion procedure, just because based on the number of observations and model parameters, several kinds of inverse problems can be defined, and they often require different specific approaches to be solved.

When the number of observed data exceeds the number of unknowns ($N > M$), the problem is said to be overdetermined, meaning that in the ideal case of noise-free data one might be able to find a single set of model parameters that exactly satisfies all measurements. In practice, however, due to the inevitable error affecting the data, no exact solution exists and an optimization procedure is required to minimize the overall misfit and identify the solution that best approximates the observed data.

Instead, when the number of model parameters exceeds the number of available observations ($M>N$), the problem becomes underdetermined: there are infinitely many sets of parameters within the model space that

can reproduce experimental data with comparable quality, meaning that the available data are not sufficient to constrain a unique solution. In both linear and non-linear problems, this lack of reliable observations with respect to the model parameters results in the presence of large regions of the model space characterized by similar misfit values, and for which additional constraints must be introduced to identify a meaningful solution among the plausible ones.

Inverse problems can be also classified well-posed and ill-posed problems. A mathematical problem is said to be well-posed if it admits a solution, the solution is unique, and it responds in a stable way to small changes in the input data. These conditions, formulated by Hadamard (1923), are rarely all satisfied in inverse problems and if any of these criteria is violated, the problem is said to be ill-posed. In practice, most of inverse problems are ill-posed: either no solution exists, multiple solutions fit the data equally well, or small errors in the observations can produce large variations in the solution. This instability is particularly problematic when data are noisy, incomplete, or only indirectly related to the model parameters.

Understanding a well- and ill-posed inverse problem is only the first step. Even when a problem is theoretically well-posed, it can still be numerically unstable due to the structure of the equations involved in the discretization of the problem. This leads to the concept of conditioning, which is equally important when addressing real-world problems.

A problem is defined well-conditioned if small changes in the input data produce proportionally small changes in the solution. Conversely, an inverse problem is ill-conditioned when even tiny perturbations in the data, such as measurement noise, lead to large fluctuations in the solution. In inverse problem, the terms ‘ill-posed’ and ‘ill-conditioned’ are often used to describe the same kind of instability: the former refers to the continuous problem, while the latter is related to its discrete numerical approximation.

Indeed, ill-conditioning is often a numerical problem, arising when the matrix G, in the discrete linear case, has a large condition number, that is, when its singular values span many orders of magnitude or approach zero, making the solution extremely sensitive to errors in the data and therefore unstable. In the nonlinear case, the notion of ill-conditioning is more subtle and cannot be characterized by a single scalar measure, as it depends on the local geometry of the forward operator in the model space.

A fundamental issue concerning the stability of the solution to an inverse problem is the inherent trade-off between resolution and stability itself.

High-resolution solutions tend to amplify noise and instability, whereas strongly regularized models are more stable but may lose fine details of interest. Regularization should therefore be understood in a broad sense: while a common choice is to promote smoothness of the solution, penalizing rapid variations in the model parameters, alternative strategies enforce different properties, such as sparsity, which favours solutions where only a limited number of parameters are significantly different from zero. Finding the right balance between data fit and regularization constraints is a central task in inversion design and interpretation.

This trade-off is easily understood by anticipating one of the problems that will be discussed in detail below: when inverting a dispersion curve to retrieve a 1D Vs profile, the model is commonly discretized in a finite number of layers that compose the vertical profile. Increasing the number of layers surely enhance the

resolution of the final model, but it can also produce instabilities in the inversion, leading to solutions that deviate significantly from the true profile. To better understand which parts of the model are constrained by the data, one often resorts to sensitivity analysis. In nonlinear or linearized frameworks, sensitivity kernels describe how variations in the model parameters affect the predicted data. These kernels allow to evaluate which parameters of the model have greater influence on the final solution of the inversion.

Typically, all these possible numerical troubles (underdetermination, ill-conditioning, and ill-posedness) tend to drive the model away from the best solution, leading many times to unrealistic oscillations of the estimated model parameters, just making the solution completely unstable with respect to the observed data.

These challenges are commonly addressed by regularization strategies (such as Tikhonov regularization) which purpose is to suppress potential extremely variations of the parameters and thus to stabilize the inversion, providing slightly smoothed and stable solutions to the detriment of the final associated misfit.

Other regularization strategies include the incorporation of prior information (Jackson, 1979) within the inversion scheme, which are used to better constrain the solution in terms of model parameters within reliable boundaries, preventing the solution from excessively deviate from these a priori information.

The role of regularization and priori information is particularly clear in the application of seismic tomographic inversion. In this context, model regularization procedures allow to obtain a smoother velocity model, in which the parameters are estimated considering also spatial correlation between adjacent parameters. This approach is based on the hypothesis that physically close regions of the model cannot assume excessively different values, thus reducing unphysical oscillations in the final solution. Instead, a typical example of a priori information is represented by the introduction of a reference model, which constraints the solution to be associated with an initial reference model considered plausible on the basis of independent knowledge. In this case, the final solution is more anchored to the starting model, limiting unjustified deviations from the data.

Additional constraints might be imposed on the basis of realistic physical considerations. For example, in the case of elastic parameters such as shear wave velocity (Vs) or density, positivity conditions can be imposed, ensuring that the estimated parameters remain physically based.

All these strategies, which will be discussed detailed later, have the task of guide the solution toward a particular class of models that reflect a priori expectations or logical considerations, of course increasing the expected misfit.

4.1.2 Misfit function and optimization

So far, we have referred to a misfit function associated with the solution of an inverse problem. The misfit (or cost) value represents a quantitative index of the agreement between the observed data, generated by the real model, and the synthetic data produced by the estimated model, therefore its choice must obviously be weighted based on the type of inverse problem and the nature of the model.

Even a perfectly posed and well-determined system can yield misleading results if the input data contain noise.

Indeed, all inverse problems must contend with the imperfect nature of the observed data: they are affected by random fluctuations, instrumental limitations, and sometimes systematic bias. As a result, the goal of inversion is not to exactly reproduce the data, but to find a model that is consistent with them within an acceptable level of error.

To formalize the concept of misfit, starting from the general form of an inverse problem (Eq. 4.1), we can assume the presence and influence of measurement errors and uncertainties in the observed data:

$$d_{obs} = G(m) + e \qquad [4.6]$$

In Eq. 4.6 the discrepancy between observed and predicted data is included in the errors vector *e* which has clearly the same dimension of the data vector *d* and each element e_i of the error vector represents the expected error associated with the measure d_i.

A particularly common choice in inversion applications is to use the L_2-norm of the residual vector *e*, which provides an overall measure of the error in terms of Euclidean distance within model space. In this sense, the misfit is expressed as:

$$\left|\left|e\right|\right|_2 = \sqrt{\sum_{i=1}^{N} e_i^2}. \qquad [4.7]$$

The concept of norm of a vector allows also to introduce the conventional least square misfit coefficient used in several inversion applications. The square of the L2-norm naturally leads to the formulation of the least square problem and introduces fundamental basis for solving a fundamental class of linear and non-linear inverse problem:

$$\left|\left|e\right|\right|_2^2 = \sum_{i=1}^{N} e_i^2. \qquad [4.8]$$

This can also be conventionally expressed in matrix notation as:

$$\chi(m) = e^T e = (d_{obs} - Gm)^T (d_{obs} - Gm), \qquad [4.9]$$

leading directly to the familiar least-squares misfit formulation. Thus, minimizing Eq. 4.9 is precisely what we mean by solving an inverse problem in the least-squares sense. The choice of a least-squares misfit is typically justified under the assumption that data errors are independent and normally distributed. In such a statistical context, the maximum likelihood estimator corresponds to the solution that minimizes the sum of squared differences between observed and predicted data. When observational errors are assumed to be independent

and normally distributed with zero mean, the least-squares estimator corresponds to the maximum likelihood solution.

Considering Eq. 4.9 within linear inverse problem, or a locally linearized nonlinear problem, the misfit function takes a quadratic form with respect to the model parameters and can be explained as follows:

$$\chi(m) = m^T G^T G m - 2 d_{obs}^T G m + d_{obs}^T d_{obs} \qquad [4.10]$$

The resulting expression clearly exhibits a quadratic form with respect to the model vector m, as indicated by the term $m^T G^T G m$. This quadratic structure is a general feature of all linear inverse problems solved using the least-squares approach: the objective function to be minimized always takes the general form:

$$\chi(m) = m^T A m + b^T m + c \qquad [4.11]$$

In Eq. 4.11, $A = G^T G$ is a symmetric matrix and the remaining terms, $b = -2 d_{obs}^T G m$ and $c = d_{obs}^T d_{obs}$, are linear and constant with respect to m. If the matrix A is also positive definite, the misfit function $\chi(m)$ is strictly convex, meaning that a unique solution corresponding to the absolute minimum of $\chi(m)$ can be found. Fig. 4.2 (from Sambridge & Mosegaard, 2002) represents the typical least-square misfit space produced by a two-parameters inverse problem. It's also clear that working with more than two parameters leads to a more complex form that is not easily reproducible in a graph but maintaining the quadratic form of Eq. 4.11.

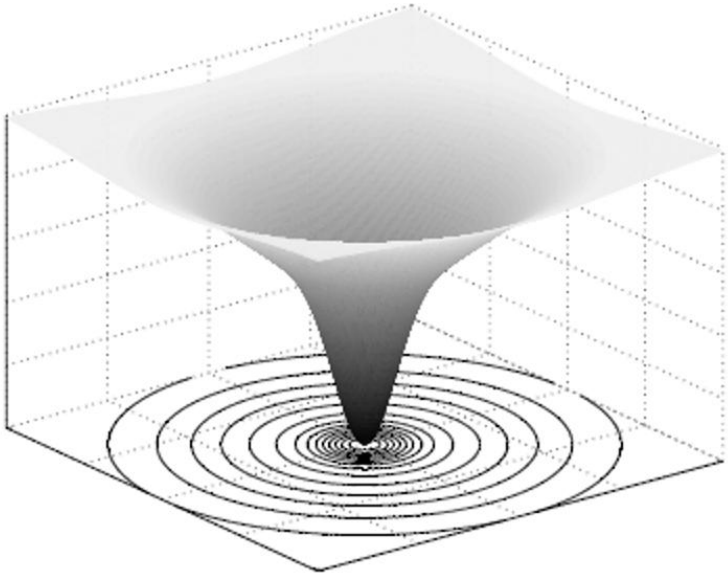

Figure 4.2 - Typical misfit space associated with a two-parameter linear inverse problem in the least squares formulation. The quadratic nature of the misfit as function of two variables produces a convex surface characterized by a unique global minimum, corresponding to the best solution to the problem (Sambridge & Mosegaard, 2002).

This property is particularly advantageous in inverse theory, as it ensures that the least-squares solution is not only stable but also uniquely defined. As a result, finding the optimal model reduces to solving a linear system derived by setting the gradient of Eq. 4.10 to zero. In linear problems this implies that a unique solution can be achieved in one-step inversion without the need for iterative and optimization procedures.

Assuming that the matrix G^TG is positive definite, finding the minimum of the misfit function simply involves identifying the set of model parameters m that sets the gradient of the function to zero, which occurs precisely at its unique global minimum. The gradient of χ(m) with respect to m is:

$$\nabla_m \chi(m) = 2G^T Gm - 2G^T d_{obs} \quad [4.11]$$

Setting this gradient equal to zero provides the condition for minimizing the least square misfit and leads directly to the Eq. 4.12:

$$m_{est} = (G^T G)^{-1} G^T d_{obs} \quad [4.12]$$

Eq. 4.12 represents the optimal linear solution in the least-squares sense, under the assumption that G^TG is invertible, and the problem is not underdetermined or ill-conditioned. The matrix G^TG is commonly referred to as the pseudoinverse matrix, or more precisely, the Moore–Penrose inverse.

It follows that the analytical form of the least-squares solution in Eq. 4.12 formally requires the inversion of a square matrix and there is no guarantee that the matrix is actually invertible in cases of rank deficiency or severe ill-conditioning of *G*. The assumption that G^TG is invertible, we will see is almost always not satisfied, depending on the dimension of the problem, the matrix *G* is often a sparse matrix, with large number of elements equal to zero. This makes the matrix *G* often singular, leading to numerical instabilities in the inversion procedure.

The optimization process can be approached, depending on the nature of the problem, in a single step or through an iterative procedure. For nonlinear inverse problems, the context is quite different. In the linear case, the quadratic form of the least-squares misfit function, which implicitly assumes Gaussian-distributed errors associated to the observed data, guarantees convexity and the existence of a unique global minimum. In the nonlinear case, this quadratic structure is lost regardless of the error distribution, and the misfit function can be highly complex, containing multiple local minima (local maxima in Fig. 4.3 from Sambridge & Mosegaard, 2002), making the optimization process significantly more challenging.

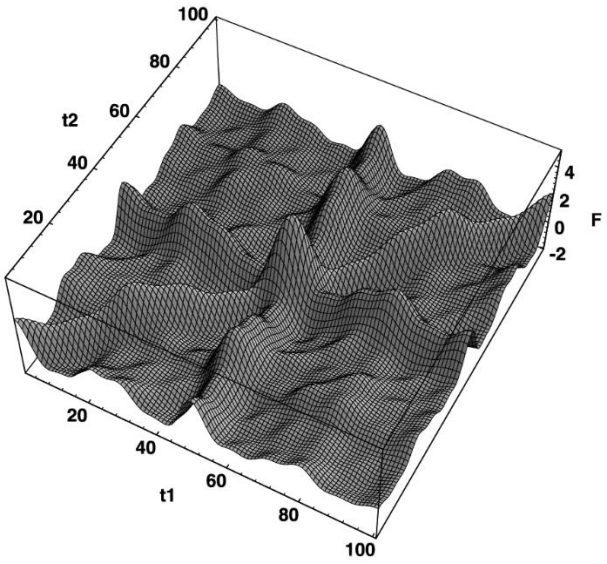

Figure 4.3 – Example of a misfit space produced by a two-parameter nonlinear inverse problem, characterized by a complex surface, in which numerous local maxima may be present (the maxima correspond to the model that best explains the observed data). Under such conditions, the optimization process is highly dependent on the initial model and requires the adoption of iterative procedures (figure reproduced from Sambridge & Mosegaard, 2002).

Consequently, for nonlinear problems numerical approaches must be adopted for which the optimization is achieved iteratively, through subsequent updates of the model toward the global minimum of the misfit function. A wide range of optimization techniques has been developed for this purpose, each based on a different philosophy of how to explore the solution space. Local search methods, also known as linearized inversion methods, typically start from an initial model and iteratively update it by minimizing the misfit function through a sequence of linear approximations (Tarantola, 2005). These methods often rely on the numerical computation of the gradient of the misfit function, which requires the misfit function itself to be differentiable. Iterative solvers such as conjugate gradient (e.g., Hestenes & Stiefel, 1952) or LSQR (Paige & Saunders, 1982) offer a more scalable alternative for large-scale problems, where the explicit formation and inversion of the normal equations matrix, such as $(G^TG)^{-1}$ in regularized problems, becomes computationally infeasible due to memory constraints or problem size. In such cases, these methods rely solely on progressive matrix-vector products to iteratively converge to the solution, without ever explicitly forming or inverting the full system matrix.

Moreover, the matrix G may be ill-conditioned, even in linear settings, especially when the data do not provide uniform sensitivity to all model parameters.

While they can be computationally efficient and converge rapidly under favourable conditions, they generally require the starting model to be reasonably close to the true solution. Otherwise, the inversion risks to be trapped in local minima (see Fig. 4.3), leading to non-optimal results.

On the other hand, a class of methods has been designed to explore the solution space more comprehensively through random sampling of the model parameters, these are broadly referred to as Monte Carlo methods. Rather than seeking a single optimal model, these methods aim to characterize the distribution of models that

are consistent with the observed data, remaining less sensitive to the choice of initial model and more robust in the presence of multiple local minima..

Early Monte Carlo algorithms were based on purely iteratively random exploration which implies a higher computational cost. More sophisticated variants, such as Simulated Annealing (Rothman, 1985; 1986) or Genetic Algorithms (Sambridge & Drijkoningen, 1992), incorporate mechanisms that guide the search process based on information accumulated during iterations, mimicking physical annealing and biological evolution respectively. Rather than sampling a probability distribution, these methods perform stochastic optimization, aiming to locate the global minimum of the misfit function while avoiding entrapment in local minima. These strategies therefore find a better balance between thorough exploration of the solution space and computational feasibility.

The wide variety of optimization methods developed and their ability to find a solution even in complex model spaces has meant that, even when the inverse problem is well-posed, linear, and theoretically admits a unique least-squares solution, it is often not solved analytically in a single computational step, and iterative methods are frequently preferred due to both computational efficiency and numerical stability. For large-scale problems, solving the normal equations $G^TGm=G^Td_{obs}$ may become infeasible due to memory constraints or computational cost, making iterative solvers the only viable option.

the explicit inversion of the matrix G^TG may be infeasible due to memory constraints or computational cost.

Finally, iterative approaches offer greater flexibility in incorporating regularization terms, such as damping, smoothness constraints, often more efficiently than closed-form solutions for large-scale problems.

4.1.3 Regularization and prior information

As seen in the previous section, even when an inverse problem is theoretically solvable in one step through the least-squares solution, practical limitations and uncertainties often require more robust strategies. In this sense, it is often advisable to use iterative procedures also for linear problems and therefore find the solution to the problem through an optimization process with one of the methods briefly listed above. Regardless of the chosen inversion method and the type of problem (linear or non-linear), regularization strategies introduced earlier are often necessary to constrain the solution, especially when implemented with methods based on quasi-random sampling. The concepts of regularization and a priori information introduced in this section are general in nature and can be applied to both linear and nonlinear inverse problems. However, for clarity of exposition, the mathematical formulation here presented starts from formulation in Eq. 4.12, representative of linear or locally linearized problems.

In the case we instead want to treat (mathematically) the problem as linear and therefore based on the quadratic form of the misfit function, various operations can be performed to modify the expected solution to make the inversion of the G^TG matrix feasible. One such strategy is regularization, which plays a central role in stabilizing solutions to ill-posed or ill-conditioned problems. In many real-world contexts, the matrix G is

nearly singular or poorly conditioned, and small perturbations in the data can cause large variations in the model. Furthermore, in underdetermined problems, multiple models may fit the data equally well. Without additional criteria, there is no unique way to choose among them. Regularization addresses both issues by introducing additional constraints or prior expectations on the model parameters, helping to select a more stable, plausible, and interpretable solution.

Without any constraints, the best-fitting solution of the inverse problem might oscillate extensively to match every fluctuation in the data, leading to a solution that would be mathematically correct but physically meaningless. Regularization introduces a *penalty* that discourages such behaviour, to the benefit of smoother or simpler models.

The most common and well-known regularisation strategy is undoubtedly the Tikhonov regularization. The key idea is to modify the standard least-squares objective function by adding a penalty term that discourages features not desired in the model, such as excessively large values or rapid variations. The regularized objective function becomes:

$$\chi(m) = \left|\left|d_{obs} - Gm\right|\right|^2 + \lambda^2 \left|\left|Lm\right|\right|^2, \qquad [4.13]$$

or equivalently in other form:

$$\chi(m) = (d_{obs} - Gm)^T (d_{obs} - Gm) + \lambda^2 (Lm)^T (Lm)\,, \qquad [4.14]$$

where d_{obs}, G and m are still the vector of observed data, the kernel matrix and the vector with the model parameters. λ is the regularization parameter controlling how strongly the regularization influences the solution, which means that higher values of λ encourage smother models suppressing strong changes in the parameters, while lower values of λ gives more importance to fitting the data, at the expense of possible brutal fluctuations of the model parameters. The matrix L represents the regularization operator and defines what aspect of the model has to be penalized.

The first term in Eq. 4.14 is the classical data misfit formula, measuring how well the model reproduces the observations, while the second term is the regularization term and plays a fundamental role in stabilizing the solution.

If for example $L = I$ (identity matrix), then the regularization term becomes $||m||^2$, that is the sum of the squared model parameters. Minimizing this term means penalize models with large magnitude parameters and allows to prevent unphysically high values. On the other hand, L usually can be expressed as derivative operator (e.g. finite difference between neighbouring parameters), opposing to roughness or variability of the model parameters, allowing for smoother models.

The purpose of this regularization term is to reduce the sensitivity of the solution to noise and to introduce stability in ill-posed or underdetermined problems. Basically, we are no longer minimizing just the misfit to the data, but also the "complexity" of the model, according to some predefined criterion. By minimizing this

combined cost function, we obtain a solution that not only fits the data reasonably well but also exhibits controlled behaviour in terms of magnitude or smoothness. The solution becomes:

$$m_{est} = (G^T G + \lambda L^T L)^{-1} G^T d_{obs}. \quad [4.15]$$

This is often referred to as the *Damped Least Squares solution*, where the damping term λ acts as a stabilizing term against possible numerical instabilities of the standard least-squares approach. Several inverse problems in seismology are solved by resorting to the use of damping terms that allow sudden oscillations of solutions to be avoided (e.g. many localization problems). Choosing an appropriate value for λ is essential. If it is too small, the solution remains unstable and sensitive to noise. If it is too large, the solution may become overly smooth or may fail to fit the data adequately.

Stabilization of the problem, on the other hand, can be often guided by priori information, such as assumptions or knowledge about the model that are not derived directly from the data. This information helps constrain the inversion by incorporating expectations about the structure or behaviour of the model. For example, we may know that certain physical parameters must be positive, or that they lie within a specific range.

Incorporating this kind of prior knowledge can significantly reduce the set of admissible solutions and, in some cases, restore uniqueness. Priors can be used, for example, to guide the solution toward a reference model, m_0, that represents prior expectations of the solution. In this case, the regularization term is designed to penalize strong deviations from m_0, rather than from zero. The last equation of χ(m) becomes:

$$\chi(m) = (d_{obs} - Gm)^T (d_{obs} - Gm) + \lambda^2 [L(m - m_0)]^T [L(m - m_0)] , \quad [4.16]$$

where L can still be the identity (for penalizing amplitude deviations) or a derivative operator (for penalizing differences in structure or smoothness). Eq. 4.16 represents a damped least squares solution with encapsulated a priori information. This formulation encourages the estimated model to not only fit the data but also stay close to the reference model in terms of overall structure or value and it is especially useful when a prior model is available from previous studies or from physical knowledge about it.

Just as a priori knowledge can help constrain the model parameters, it is equally important to consider the quality of the data. In real applications, measurements are affected by varying levels of noise and uncertainty. By considering this information, the inversion process can weight each observation, accordingly, ensuring that more reliable data have greater influence on the final solution. To take measurement uncertainty into account, the contribution of each data point to the overall misfit must be adjusted based on its expected reliability. This is typically done by introducing a data weighting matrix, W, which assigns lower weights to less reliable measurements and higher weights to more accurate ones. The weights are usually derived from the data covariance matrix, C_d, which quantifies the statistical variance and potential correlations between observations. To simplify the expression of the misfit function, the matrix W is defined such that:

$$W^T W = C_d^{-1}, \qquad [4.17]$$

where C_d is the data covariance matrix, which quantifies the expected variance and correlations of the data errors. This transformation allows the data misfit to be rewritten in a normalized form, where each residual is effectively scaled by the inverse of its expected variance. Taking Eq. 4.17 and putting in it the information about the reliability of the observed data leads to a weighted least-squares formulation based on regularization:

$$\chi(m) = (d_{obs} - Gm)^T C_d^{-1}(d_{obs} - Gm) + \lambda^2[L(m - m_0)]^T[L(m - m_0)] \qquad [4.18]$$

In this form, data points with greater uncertainty (i.e., larger variances) have less influence on the solution. This ensures that the inversion is not unduly affected by noisy or unreliable measurements and enhances the overall robustness of the model estimate.

All this additional information discussed so far, ranging from the incorporation of data uncertainty to the inclusion of regularization terms and prior knowledge, leads to a generalized form of the misfit function. In Eq. 4.18, each term plays a specific role in improving the robustness and stability of the solution. The data covariance matrix (or, equivalently, the weighting matrix W) modifies the original least-squares term by adjusting the contribution of each residual according to its expected uncertainty. This directly affects the way the data misfit is computed and ensures that more reliable observations carry more influence in the solution. On the other hand, the regularization and prior information terms form a separate additive component that helps stabilize the solution in cases of non-uniqueness or ill-conditioning. These terms represent the expectations about the model, such as smoothness, proximity to a reference model, or physical plausibility. Importantly, both the weighted misfit and the regularization term preserve a quadratic form, which means that the mathematical structure of the optimization problem remains the same. As a result, the same procedure used earlier to derive the least-squares solution can be applied here as well, ultimately leading to the following expression for the estimated model parameters (Tarantola, 2005):

$$m_{est} = \left(G^T C_d^{-1} G + \lambda^2 L^T L\right)^{-1}(G^T C_d^{-1} d_{obs} + \lambda^2 L^T L m_0) \qquad [4.19]$$

Eq. 4.19 is the generalized form of the regularized and weighted least-squares solution, centred around a reference model m_0. It combines all sources of information, data fit, uncertainty, and prior assumptions, into a unique framework that improves both the stability and interpretability of the final model.

In conclusion, regularization, data weighting, and a priori information are fundamental tools in solving inverse problems. Regularization stabilizes solutions and prevents overfitting, while data weighting reflects the varying reliability of measurements. A priori constraints allow us to incorporate physical knowledge and

expectations about the model. Together, these methods provide the necessary structure to obtain solutions that are not only mathematically valid but also physically meaningful and geologically interpretable.

While the regularized least-squares formulation presented above allows for a direct one-step solution in the linear case, real-world problems are often non-linear or only approximately linear. Nevertheless, the same mathematical structure remains useful when embedded in an iterative procedure toward the best solution. At each iteration, the forward problem is linearized around the current model m_k , and a linearized version of the misfit function is minimized to obtain an update δm with respect to the current model. This leads to a sequence of regularized least-squares problems, solved progressively adjusting the model parameters until convergence is reached. In this way, the tools developed for the linear case, including data weighting, regularization, and prior information, can still be applied effectively in non-linear settings by updating the linear system at each step.

4.2 A linear inverse problem: seismic tomography

The word tomography is derived from the Greek roots tomos (τόμος), meaning "slice" or "section," and –graphy (–γραφία), meaning "writing" or "representation." In practice, tomographic techniques reconstruct the internal structure of an object by combining information from multiple sectional views, effectively *imaging* its interior one thin slice at a time.

Today, tomographic methods play a central role across numerous scientific and engineering disciplines, offering powerful tools to "see" inside otherwise opaque materials. Perhaps the most familiar example is medical imaging CT (computed tomography) and MRI (magnetic resonance imaging) scans, where successive "slices" through the human body are reconstructed from measurements of X-ray absorption or radiofrequency response. Regardless of the physical modality, whether photons in medicine, electrons in materials science, or seismic waves in geophysics, the goal of tomography remains the same: to reconstruct the internal structure and physical properties of a target medium by combining observations made along many different paths through it.

Seismic tomography is undoubtedly one of the most powerful tools in seismological field and it is a powerful tool for imaging the internal structure of the Earth, based on the principle that propagation of seismic waves is affected by the elastic properties of the materials they travel through.

By measuring specific characteristics of seismic waves, such as propagation velocity or wave amplitude, and analysing how they change during propagation within the Earth, it is possible to infer 1D, 2D or 3D distribution of physical parameters such as seismic velocity, density and attenuation. In this way, seismic tomography turns the planet itself into a natural laboratory using its internal vibrations to illuminate hidden features from crustal faults and magma chambers to mantle plumes and the core-mantle boundary.

The simple concepts outlined above place tomographic methods squarely within the class of inverse problems. Inverse problems are at the heart of much of seismological research: they all rest on the same basic principle, namely inferring the physical properties of a model, whether Earth's crust, mantle, or core, by observing external measurements (e.g., travel times, amplitudes, and waveforms) that have been modified by those properties. By treating seismic imaging as an inverse problem, we transform raw wavefield data into quantitative maps of velocity, density, and attenuation, thereby turning indirect observations into direct insights about the planet's hidden interior.

After theoretical framework for solving inverse problems have been described, including their linear and nonlinear formulations, regularization strategies, and the role of data uncertainty and prior information, we now shift our focus to a well-defined field application: seismic tomography.
Seismic tomography includes several classes of inverse problems in which the spatial distribution of seismic properties, typically seismic wave velocities (or density of the medium), is inferred by analyzing how seismic waves propagate through the Earth (or how their amplitudes attenuate during the propagation). Over the years, a variety of tomographic approaches have been developed, each characterized by different types of waves, data acquisition strategies, and resolution targets. In this section, we briefly review the main types of seismic tomography, with emphasis on those based on surface waves and ambient seismic noise.

Most traditional strategies of seismic tomography involve in recording arrival times of body waves, especially P waves, generated by natural earthquakes (e.g., Aki & Lee, 1976). The waves generated by the sources propagate through the Earth's interior and the time the signal takes to travel from the sources to each receiver is obviously influenced by the distribution of medium properties along the path within the medium. In this context, each travel-time measurement can be modelled as an integral of the slowness (reciprocal of the wave velocity) along the ray path:

$$tt = \int s \, dl, \qquad [4.20]$$

where *tt* is the travel-time from the sources to the receiver, *dl* denotes the infinitesimal ray path element along the wave propagation and *s* is the slowness value along that element. These approaches, although still widely used today, inherently require a well-defined point source, either natural or artificial. As a result, the quality and spatial distribution of the available sources (e.g., earthquakes or controlled explosions) become limiting factors in terms of both the density of ray coverage and the resolution of the final model. In regions with low seismicity, or where active surveys are not feasible, the number of usable ray paths may be insufficient to ensure adequate coverage, potentially compromising the stability and interpretability of the tomographic solution. Given these limitations, alternative approaches have been developed that do not rely on discrete sources. Among them, surface-wave tomography has gained widespread adoption due to its ability to exploit the continuous wavefield (either from regional surface-wave propagation or from background seismic noise)

to achieve dense path coverage and high-resolution imaging, particularly in the shallow crust, only depending on the number and location of the seismic stations.

Surface-wave tomography exploits the dispersive nature of Rayleigh and Love waves and their frequency-dependent sensitivity to shear-wave velocity variations in the shallow subsurface (see Section 2.3), making them particularly suitable for near-surface characterization. An increasingly widespread implementation of surface-wave tomography is known as Ambient Noise Tomography (e.g., Ekström, 2014; Kästle et al., 2016; Ritzwoller et al., 2011), hereafter ANT. Rather than relying on active sources or earthquakes, ANT exploits the continuous background seismic noise recorded at dense seismic arrays. The key idea is to correlate long-term ambient noise recordings between pairs of stations (as discussed in Chapter III) to retrieve information about the noise wavefield and the subsurface properties, as if one station were acting as a virtual source and the other as a receiver.
The resulting correlated signals predominantly contain Rayleigh and Love waves, whose dispersive characteristics can be analysed using the same techniques employed in traditional surface-wave tomography.
ANT enables high-resolution imaging of shallow structures over broad regions, particularly in areas with low seismicity or where deploying active sources is not feasible.
In practice, even ANT, despite being based on surface-wave propagation and dispersion within the ambient noise, is often implemented as a standard travel-time tomography. The phase or group velocities extracted at a given frequency (or period) between station pairs are commonly converted into equivalent travel-times through the measured velocity. These frequency-dependent travel times are then used as observed data within a 2D domain, producing lateral velocity maps at different frequencies. In this way, the inversion process reflects the more classic travel-time tomography performed in a simplified 2D model.
Initially developed for deep Earth studies (e.g. Dziewonski, 1984), seismic tomography has since evolved to incorporate a wide range of scales and methodologies, from global crust and upper mantle imaging to near-surface characterization. While large-scale surface-wave tomography often relies on data from regional or global arrays (see the works mentioned above), there is a growing interest in applying tomographic principles at smaller scales, using data from temporary deployments or dense local networks (Hayashi & Suzuki, 2004; Boiero & Socco, 2010; Bergamo et al., 2012; Socco et al., 2014). In such contexts, traditional array-based methods may be limited by geometry or logistics, motivating the development of alternative approaches based on inter-station active measurements and ambient noise-based approaches.

4.2.1 Forward model and ray theory

Considering a general case of travel-time tomography, which as before state can be used both for body and surface waves tomography, and looking at the aforementioned general form of inverse problem in least square sense (Eq. 4.19), the vector d with the data is represented by the arrival time of the seismic waves recorded at

the seismic stations and the vector *m*, representing the model of the system, denotes the medium which is crossed by the seismic waves. The model parameters express a well-defined characteristic of the medium, that is in most seismic tomographic applications, the seismic velocity distribution of the medium, and indirectly the density itself. Finally, the matrix *G* even in this case represents the physical relation between the observed arrival times and the velocity of the medium, the higher the density of the medium, the greater the velocity of propagating waves, therefore the shorter the travel-times between source and stations.

These changes in the subsoil properties are recognized indirectly by following the diverse travel-time paths of seismic waves from a source to seismic stations and, depending on the assumed propagation model (forward problem), evaluating the relative arrival times between one station and another. Today, various increasingly realistic numerical methods exist to approximate the wave propagation, although none can fully capture the whole physical complexity of seismic wave behaviour in heterogeneous media.

Finite-difference methods, for instance, enable accurate simulation of wave propagation by numerically solving the elastodynamic equations, particularly effective for modelling wavefields including scattering, diffraction, and interference phenomena. These methods discretize both space and time and are especially valuable in media with lateral heterogeneities, although their application to irregular geometries or complex topography may require alternative approaches such as finite-element or spectral-element methods (e.g., Komatitsch & Tromp, 1999). More recently, alternative approaches have gained popularity for specific applications, such as the Fast-Marching Method (Sethian, 1996), which solve the eikonal equation efficiently to estimate first-arrival travel times in smoothly varying media (e.g., Vidale, 1988; 1990). These methods are computationally efficient and provide accurate travel-time estimates, making them suitable for large-scale tomographic inversions.

All such methods, however, shift seismic tomography into the field of non-linear problems, due to the complexity of wave propagation and the dependence of travel times on the velocity structure. One of the simplest ways to approximate propagation from a source to a station is by assuming that the wavefront travels along a straight-line path, both for 2D or 3D models, like an arrow pointed from the source to the receiver (in the case of ANT from one receiver to the other one).

In this simple formulation, the physical model underlying tomography relies on the assumption that geometrical ray theory is valid. This theory, which corresponds to the high-frequency approximation of the wave equation (Aki & Richards, 1980), allows wave propagation between two points to be represented as occurring along infinitesimally narrow ray paths.

In this approximation, the wave does not change trajectory, and the total travel time depends solely on the number of blocks (that constitute the 2D model parameters of the inverse problem) it crosses, and the velocity assigned to each. The more the ray travels through high-velocity zones, the shorter the travel time; conversely, crossing slower blocks results in longer travel times.

Given the discrete form of the inverse problem in the matrix definition of the model parameters (a discrete set of parameters), it is self-evident that the set of parameters that describe the model must be discretized from the continuous formulation of Eq. 4.20. In practical tomographic studies the medium is subdivided into several

2D, or 3D blocks (depending on the considered space dimensions) and each of these block (or cell) represent a single model parameter with its own characteristic velocity (or slowness) and given the considered wave propagation through the medium the goal is to find the most reliable velocity value for the blocks. So, Eq. 4.20 can be discretized and written in matrix form as:

$$tt_i = \sum L_{ij} s_j, \qquad [4.21]$$

where tt_i (d_i in the general form) represents the i-th element of the vector containing the travel-time data at the receivers, s_j (m_j in the general form of inverse problem) is the j-th element of the vector containing the slowness of the cells (2D or 3D) into which the domain has been subdivided and L represents the forward matrix (G) that relates the observed travel-time to the slowness of the blocks along the ray path and thus governs the physics of the wave propagation within the considered domain.

In general, the elements of L represent the sensitivity of each data point to perturbations in each model parameter. In more formal terms, the forward operator can be thought of as the Jacobian matrix of the forward problem, where each entry L_{ij} approximates the partial derivative:

$$L_{ij} = \frac{\partial tt_i}{\partial s_i}. \qquad [4.22]$$

This means that L_{ij} quantifies how a small change in the j-th model parameter s_j would affect the i-th observation tt_i. In many physical problems, such as seismic tomography, this sensitivity is determined by the physics of wave propagation and is inherently spatially variable. The construction of G therefore depends on the physical model chosen (e.g., straight ray-path), the parametrization of the model space (e.g., block-wise discretization) and the type of data considered (e.g., first arrivals, phase shifts, amplitudes).

When seismic wave propagation is approximated as rectilinear and first rivals are considered; the forward problem simplifies considerably. Assuming that the model parameter s_j represents the slowness in the j-th cell, the travel time for each path can be expressed as the sum:

$$tt_i = \sum_j l_{ij} \cdot s_j \qquad [4.23]$$

In Eq. 4.23 l_{ij} is the length of the considered ray segment passing through the cell j along the i-th propagation path. Even in this case, the forward operator L (that contains the elements l_{ij}) is fully linear and contains the geometric path lengths, which in all respects represents the derivative of the travel time tt_i with respect to the model parameter s_j in a finite difference sense. So, the forward matrix L can still be considered as representative of the Jacobian of the problem relating data and model parameters:

$$L_{ij} = \frac{\partial tt_i}{\partial s_j} = l_{ij} \qquad [4.24]$$

To illustrate the idea behind the construction of the forward operator L in ray-based seismic tomography, we can consider a simplified 2D model in the *x, y* plane discretized into square cells of same size *nx* and *ny* (with *nx* = *ny*). Each cell is associated with a constant slowness value s_j which represents the model parameters and are grouped into the vector $s = [s_1, s_2, \ldots, s_N]^T$, where N is the total number of cells.

In the Fig. 4.4, a single source (red star) generates seismic waves recorded at multiple receivers (blue triangles). For each source-receiver pair, a straight ray path is assumed. The corresponding travel time tt_i can be discretized as the sum of the slowness values in the cells crossed by the ray, weighted by the distance travelled within each cell. This formulation gives the same linear relation expressed in Eq. 4.23. In matrix notation, the system takes the same form of a linear inverse problem.

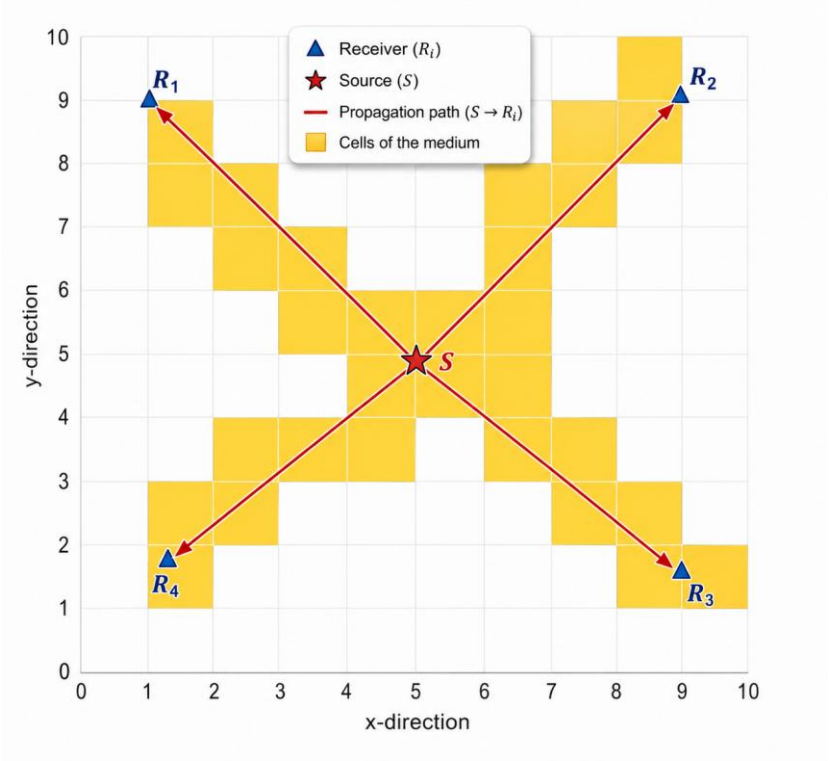

Figure 4.4 – Conceptual diagram of 2D straight-ray seismic tomography: a source (S) and four stations (R_1–R_4) define the distribution of the propagation paths; the grid cells crossed by the rays are highlighted, representing the model parameters that contribute to the estimation of the solution.

Considering this practical example, the forward matrix *G* will be composed of four rows, one for each observed travel-time, and number of columns due to the number of model parameters, one hundred in this case. If we consider the i-th observed travel-time, following Eq. 4.23, it will be represented by the linear combination of all the products:

$$tt_i = L_{i1}s_1 + L_{i2}s_2 + \cdots + L_{iM}s_M. \quad [4.25]$$

From this example, it becomes clear that only parameters associated with cells crossed by the source-receiver paths are nonzero within the matrix L *(*highlighted cells in Fig. 4.4*)*, and the corresponding matrix element L_{ij} stores the length of the ray segment that lies within that cell. Conversely, for cells not intersected by the rays, the associated matrix entries are set to zero. This structure significantly simplifies the summation involved in computing each travel time: rather than requiring a full matrix–vector product over all model parameters, the sum reduces to a subset of terms involving only the slowness values in the cells along the ray path.

This localized sensitivity pattern explains why the matrix L is typically sparse in ray-based tomography: each row of L (corresponding to one travel-time observation) contains only a few non-zero entries. While this sparsity is computationally advantageous, it also reflects the limited information content of each observation, reinforcing the need for adequate ray coverage and supplementary constraints such as regularization.

It is important to note that the quality and resolution of the inversion strongly depend on the structure of L matrix. Well-conditioned and well-populated forward operators allow for stable and meaningful inversions, while poor coverage, parameter redundancy, or data noise can render the system ill-posed or ill-conditioned. This has direct consequences on the stability of the inversion, the amplification of noise, and the ability to achieve reliable features from the final model. In practice, the forward matrix is often constructed explicitly through forward modelling, using numerical solvers or semi-analytical expressions, depending on the chosen level of approximation. Its size and shape (i.e., number of rows and columns) are determined by the number of data points and the number of model parameters, respectively. Understanding its structure is essential before attempting any inversion strategy.

4.2.2 Model solution for seismic tomography

When the matrix G is sparse and contains several zero entries, as is often the case in travel-time tomography, its associated normal matrix G^TG will also be sparse and may be ill-conditioned or even non-invertible. Attempting to solve the inverse problem directly using the standard least-squares formulation (Eq. 4.12) can lead to numerical instabilities, unreliable solutions, or a complete inversion failure.

As previously discussed, the strategies introduced to stabilize the inversion (regularization, prior information, and data weighting) now become essential. In the following, these elements will be explicitly reintroduced and reinterpreted within the specific framework of 2D seismic tomography, providing a clear physical meaning for each term in the generalized inversion formulation.

All these concepts regarding model stability, sparse data coverage, and the need for additional constraints converge in the standard scheme to solve linear or weakly non-linear problems (Eq. 4.19).

In this expression, each term reflects one of the key ingredients introduced to overcome the intrinsic limitations of inverse problem, a 2D seismic tomography in this example. The matrix G (L in this subsection) remains the core of the formulation, consisting of the geometric relationship between sources, receivers, and the model space. Its structure reflects how the wave paths sample the model, and its sparsity is a direct consequence of the localized sensitivity of each observation.

To mitigate the effect of noisy or unevenly distributed data, a weighting matrix C_d^{-1} is introduced. This matrix scales the contribution of each travel-time measurement in proportion to its reliability, effectively down-weighting uncertain observations. The term C_d^{-1} in the formulation is equivalent to the inverse of the data covariance matrix and ensures that the inversion accounts for errors among the input measurements data. In this context, uncertainties about the data reflect uncertainties in the travel-time estimation, and in classical travel-time seismic tomography they are directly correlated to the uncertainties in the velocity estimation during the data acquisition step. In the inversion process, more influence will be given to that measurements that show narrower confidence intervals, at the expense of those observations with great uncertainty.

However, even in presence of weighted data, the problem often remains ill-conditioned or underdetermined due to insufficient ray coverage. This is where regularization becomes essential. The term $\lambda^2 L^T L$ introduces a stabilizing constraint on the model, typically penalizing large gradients or deviations from smoothness. The parameter λ, known as the damping or regularization parameter, balances the importance of fitting the data versus enforcing stability and simplicity in the model. In this context, the regularization aims to stabilize the slowness of the model parameters assuming that two close model parameters should not show strong differences in terms of slowness (or velocity). This regularization aims to smooth the slowness values along the model domain.

Finally, the presence of a reference model m_0 allows the inversion to be guided toward a plausible model, especially when the data are too sparse to constrain the model independently. This prior model might represent a homogeneous velocity model, or a more articulated structure known from independent information. Its influence prevents the final solution in terms of seismic velocity from deviating too far from this reference unless the data clearly support this behaviour. Considering the tomographic inverse problem, several constrains in the form of prior information could be considered: velocities can take only positive numbers; from a mathematical point of view each parameter can take any value if it involves in a better data-fitting, even if negative velocities have no sense to exist. Additionally, if the medium under study consists essentially of rock, the density can be expected to lie within a reasonable geological range, such as between 1,000 and 10,000 kg/m³. Incorporating such constraints into the inversion process can significantly reduce the set of reliable solutions, and it may even reintroduce a unique solution in original underdetermined or non-unique problems.

Together, these elements create a robust inversion framework, whose solution is not only mathematically stable, but also physically interpretable under real-world data conditions.

So, approximating the wave propagation as a straight-line between source and receiver (or receiver to receiver in case of ANT), this trajectory necessarily corresponds to the minimum travel-time path between the two points. In this simplified approximation the tomographic problem remains linear. Therefore, it can be resolved through the standard linear inversion scheme described above, finding the least square solution that best approximates the experimental data.

4.3 A nonlinear inverse problem: from Rayleigh dispersion curve to Vs model

Surface wave dispersion curve represents one of the most powerful and, at the same time, “indirect” measure of near-surface properties. In Chapter III, we provided a brief review of the commonly adopted surface wave methods that allow to estimate surface wave dispersion at a site. We highlighted the difference between active and passive methods and focused on the latter class of approaches, which are based on the analysis of ambient seismic noise as a source of information on the dispersive properties of surface waves. In most cases, all these techniques used to retrieve dispersion curves are almost always followed by the application of an inversion process of the curve itself. In fact, each dispersion curve is an “indirect” expression of the main elastic properties of the subsoil, which, once a reliable curve has been achieved, are obtained precisely through inversion. Unlike direct measurement of the main properties of the medium, such as layer velocity or interface depth, the dispersion curve is the result of a complex propagation process that affects the entire volume sampled by the wavelength and returns, in compact form, the effect of the whole vertical elastic structure of the subsoil. In this context, as first introduced theoretically and then illustrated in the tomographic case, solving the inverse problem means estimating the characteristics of the profile in terms of Vp, Vs of the layers composing the model as well as their thickness and density, starting from the observed dispersion curve. Conversely, solving the direct problem means calculating a theoretical dispersion curve based on a specific subsoil model expressed strictly in terms of a 1D vertical configuration (Fig. 4.5 from Foti et al., 2011).

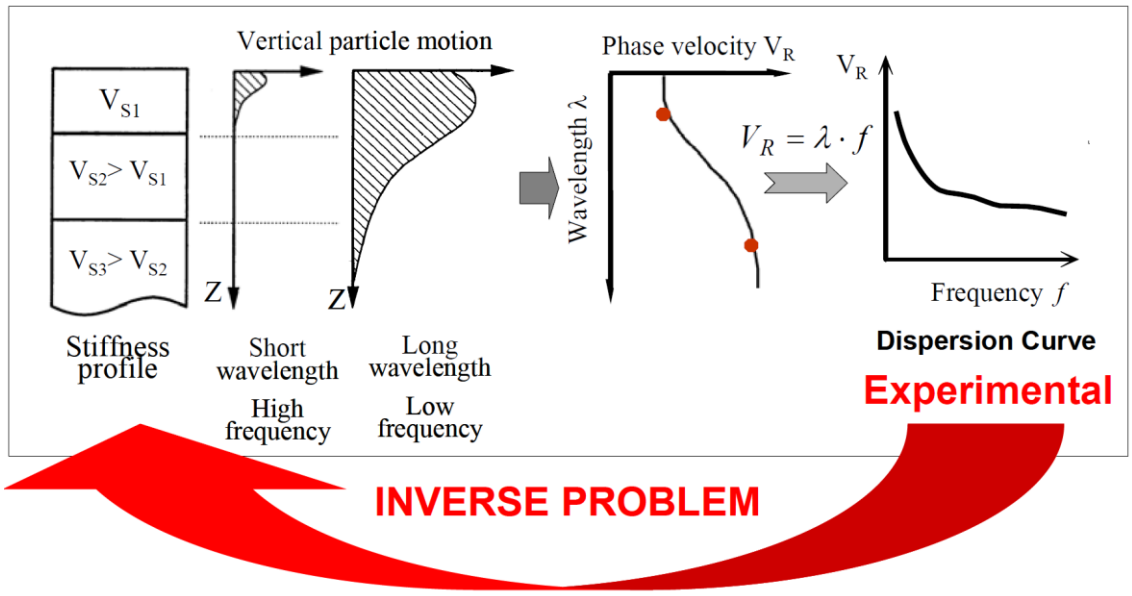

Figure 4.5 - Scheme of the inverse problem associated with the inversion of surface wave dispersion curves. The 1D variation of elastic properties with depth controls the energy distribution of surface waves with different wavelengths (and so frequencies), defining the observed phase velocity. The experimental dispersion curve therefore represents an indirect measure of the vertical structure of the subsurface which inversion should converge toward the velocity model. Reproduced from Foti et al. (2011).

The standard formula for the inverse problem introduced at the beginning of this chapter (Eq. 4.1) is still formally valid even when considering the dispersion curve problem. In this case, d is populated by the phase velocities observed from the experimental curve (either Raleigh or Love), m denotes the parameters that describe the layered model, which are the sets of Vp, Vs and density of the layers that build up the vertical profile. Unlike the linear (or linearized case) problem, here the forward operator can no longer be expressed as a simple matrix operator G, instead, it must be expressed as a non-linear operator $\boldsymbol{\mathcal{F}}$ acting on the model parameters:

$$d^{obs}(\omega) = \boldsymbol{\mathcal{F}}(m) \qquad [4.26]$$

As already mentioned, this type of problem is highly non-linear and, unlike the tomographic problem formulated in section 4.2, it is a hard task to express the inverse problem in a closed form (as in Eq. 4.19). This intrinsic complexity is not only due to the number of parameters involved in the formulation of the problem, which, based on the specific problem, is generally lower than the number of parameters to be solved in a tomographic inversion, but especially to the way these parameters (*Vp*, *Vs*, density and thickness) enter into the forward formulation, which is a complex relation in which every observed phase velocity $c(\omega)$ depends practically on almost all parameters that compose the model.

Although the misfit can be expressed in terms of least squares error also in this case (Eq. 4.27), this formulation does not produce any kind of advantage: the resulting parameter space is not a convex surface but contains numerous local minima, in no way guaranteeing stable convergence toward a global solution.

$$\chi(m) = \left|\left|d^{obs} - \boldsymbol{\mathcal{F}}(m)\right|\right|^{\mathbf{2}} \quad [4.27]$$

Basically, in forward formulation, the observed phase velocities at each frequency are the results of a complex contribution of the elastic parameters of individual layers and the distribution of wave energy at depth, hard layers may increase the phase velocity and therefore to pull up the dispersion curve, on the contrary slow and weak layers reduce the phase velocity, but the relationship between cause and effect remains inherently non-linear. Among the controlling parameters, the shear wave velocity Vs plays a fundamental role in the dispersion and modal propagation of surface waves, while the effect of Vp and density is usually secondary. Not surprisingly, most surface wave-based methods aim at reconstructing a Vs(z) model.

4.3.1 Forward model

To understand why this transition from dispersion curve to subsoil velocity model is so delicate, it is useful to look carefully at the nature of the forward problem, i.e. the calculation of the dispersion curve from a known set of model parameters. From a theoretical point of view, Rayleigh waves (and Love waves) in a vertically stratified medium are solutions of the elastodynamic equations for vertical and horizontal displacements in the considered propagation planes (Aki & Richards, 1980), P-SV for Rayleigh wave and SH for Love wave.
Importing them as "guided waves" on the free surface leads to a linear differential eigenvalue problem (Eq.4.26):

$$\frac{dl(z)}{dz} = A(z)l(z), \quad [4.26]$$

where the state vector *l(z)* contains displacement and stress as a function of depth and the matrix *A(z)* depends on the elastic parameters of the layers. Given a stratified subsoil model, there exist only some specific admissible phase velocities for each frequency that simultaneously satisfy the equations of motion and the considered boundary conditions (free-surface condition at the top and radiation condition at the bottom of the model).
In the context of continuous variation of the elastic parameters with depth, the solution of the problem of Eq. 4.26 might be found by integration through numerical methods (e.g. Takeuchi & Saito, 1972) considering the boundary conditions that must be imposed at the top and the bottom of the model. In common approaches, however, the continuous vertical profile is discretized in a stack of homogeneous layers characterized by constant values of the model parameters (Vp, Vs and density). Most of these methods are based on the transfer

matrix method (Thomson, 1950; Haskell, 1953), later formalized by Gilbert & Backus (1966) through the introduction of the propagator matrix method, which consists essentially in finding the roots of the characteristic dispersion equation for Rayleigh and Love waves.

The dispersion equation is constructed through sequentially matrix multiplication that propagate the fundamental state vectors *l(z)* across the layers, maintaining physical consistency with body and surface wave propagation and satisfying the boundary conditions of the differential problem, which must be also accompanied by the imposition of continuity of displacements and stress at the interfaces between adjacent layers. The size of the matrix depends on the type of surface wave considered: in the case of Love waves the formulation leads to a 2×2 system (since Love waves is influenced only by Vs), while for Rayleigh waves the problem takes a more general 4×4 form. The matrix propagation through the layers allows to link the boundary conditions imposed at the bottom of the model (typically the radiation condition in a homogeneous half-space) with the free-surface conditions at the top of the stratified model (zero-stress at the surface).

The roots of the dispersion equations represent the wavenumbers associated with the different surface waves mode of propagation: for a given frequency, only a few sets of wavenumbers (or phase velocities) represent a feasible solution only if, once propagated through the entire model, it simultaneously satisfies all the boundary conditions. Consequently, the forward problem, from a mathematical point of view, comes down to a procedure of root finding of the characteristic dispersion equations of Love and Rayleigh waves. For each frequency, diverse roots can provide multiple distinct solutions, each associated with a different propagation mode, including the fundamental mode and higher modes. The dispersion curve is nothing more than the set of phase velocities, as function of the frequency, that satisfy all above conditions.

On this theoretical basis, several refinements of Thomson–Haskell's original method have been proposed in last decades, primarily with the aim of improving numerical stability in the context of Rayleigh wave forward computation, for which classical formalism of the original Thomson-Haskell method can lead to significant numerical instabilities in solving the eigenvalue problem.

Among them, the best known are the delta matrix method proposed by Pestel & Leckie (1963), the Schwab–Knopoff method (Knopoff, 1964), the Dunkin matrix method (Dunkin, 1965) and the reflection–transmission RT method introduced by Kennett (1974), then implemented in several computer programs (e.g., Herrmann, 2013 and Wathelet et al., 2020). All these techniques have been developed over the years with the aim of reducing numerical instabilities, especially associated with the roots-finding procedure for Rayleigh wave, and improving computational efficiency in solving the forward problem. A systematic review of the main approaches proposed and used was provided by Buchen & Ben-Hador (1996).

4.3.2 Dispersion curve inversion

After describing the forward problem for computing dispersion curves in a layered model, it is possible to understand why dispersion curve inversion is a highly nonlinear problem.

The function that associates the related dispersion curve with a stratified model is strongly nonlinear: a small perturbation of Vs at a given depth can produce complex, disproportionate changes across multiple frequency bands of the dispersion curve, making simple linearized approximations unreliable and motivating the use of dedicated nonlinear inversion strategies.

In the forward problem, indeed, the relationship between model parameters and observed data cannot be expressed in linear form: for each frequency the phase velocity is obtained by solving a dispersion equation through a root search procedure, and the forward operator can only be expressed in terms of a non-linear operator linking the predicted dispersion curve with the model. Consequently, even if the general formulation of the inverse problem remains unchanged, the misfit function turns out to be a non-linear function of the model parameters. In this context, minimizing the differences between the observed and the computed curve requires necessarily an iterative optimization procedure, which must handle the presence of multiple local minima in the misfit space and strong trade-offs between the model parameters. A classic example is the trade-off between Vs and layer thickness (an increase in velocity can be compensated by a lower thickness) which is often addressed by fixing the number and thickness of the layers within the model. Indeed, increasing the number of layers or treating layer thickness as a free parameter dramatically increases the overall complexity of the problem, especially in terms of uniqueness of the solution, increasing the chances of inversion failure.

The choice of parameterization and a priori constraints therefore becomes critical in the final interpretation of the solutions, since there might be infinitely many model configurations able to reproduce the experimental curve and this unfortunately is an insurmountable obstacle, which causes inversion methods based on the gradients computation (such as those mentioned above) to fail, as the solution space is a complex higher-dimensional surface (as many as the parameters), with valleys and discontinuities that can make classical inversion procedure to remain trapped in local minima without sampling more alternative solutions.

Comparative studies (e.g., Foti et al., 2014; Lai & Rix, 1998) show that a significant portion of the variability in inverted profiles is related not exclusively to the inversion algorithm, but also to the choice of a good parameterization of the problem: number of layers, thickness distribution, physical constraints related to realistic velocity range. Too rigid parameterization can prevent real variations in the structure from being captured, while weakly parameterizations might generate unphysical oscillations and amplify non-uniqueness. A balanced choice of parameterization is essential to obtain interpretable models, and its quality often matters as much (or more) than the choice of optimization algorithm.

All this means that the goal of a dispersion curve inversion cannot be just to find a single “true” model, but rather to identify a family of plausible models, compatible with the a priori data and knowledge integrated from independent information or geological constraints into the inversion scheme, in order to reduce the space of feasible solutions and enhance the stability of the overall inversion.

These challenges in identifying a unique and stable solution have led to the adoption of previously mentioned global search methods designed to explore the entire parameter space (see Sen & Stoffa, 1995). The basic idea is to avoid the use of local optimization methods, which can provide information on model resolution or the existence of trade-offs between parameters (Foti et al., 2011), but which are inevitably constrained to the

vicinity of the initial model, at the expanse of global search methods. The inversion problem is then formulated as a stochastic sampling problem, in which the model space is random sampled extensively. This of course reflects in a computationally more expensive process that nevertheless allows the set of feasible solutions to be explored globally. Such approaches are generally more robust to the nonlinearity of the problem, since they do not require for matrix inversion or gradients calculation.

In this context, the class of Monte Carlo methods represent the most widely used methods to infer subsoil velocity configuration starting from observed dispersion curves. In the simplest Monte Carlo implementation, random models are generated within fixed constraints according to subjective criteria, the theoretical dispersion curve is then calculated for each model and the misfit between the observed and predicted curve is evaluated in order to preserve the best models as potential solutions of the problem. This strategy, while crude, highlights very well the global structure of the expected misfit, but at the same time reflects the need for subjective constraints that guide the inversion to sample effective portions of space.

The main limitation of pure random sampling lies in its poor computational efficiency, since no information about the obtained solutions transmitted from one iteration to the next one.

To overcome this problem, guided exploration strategies have been introduced, in which the information contained in the misfit progressively computed during the iterative process is used as a dynamic constraint to focus the search towards more promising regions of the model space.

Examples of such approaches used in dispersion curve inversion include Simulated Annealing (e.g., Beaty et al., 2002), Genetic Algorithms (e.g., Picozzi & Albarello, 2007), and the Neighbourhood Algorithm (Sambridge, 1999a, 1999b). Despite adopting different strategies to explore the model space, these methods share the common goal of locating the global minimum of the misfit function while maintaining a balance between thorough exploration and computational efficiency.

Simulated Annealing (SA) is an optimization technique based on thermal cooling of materials. At the beginning (when the temperature is high) the algorithm broadly explores the parameter space, and even poor solutions are accepted, as the iterations go by (temperature decreases), sampling becomes more selective, focusing only on some solutions in the space that improve the misfit. This sampling rate depends on a temperature parameter that, as the iterations proceeds, modifies the sampling probability distribution that allows within certain limits to consider even worse solutions in terms of misfit. This also reduce the chances of being trapped in local minima. Genetic algorithms (GA), on the other hand, is based on the iterative generation of populations of models. The algorithm starts not from a single initial model but generating a set of candidate models evaluated in terms of misfit. In this context the best-fitting models are selected as the population of models that best approximate experimental data and in the next iteration, a new population of model is generated by combining the characteristics of the best models selected before and introducing small random perturbations. A yet conceptually related approach is provided by NA, which key idea is to divide the model space into discrete cells associated with the models already evaluated, and to use the misfit value to decide where to densify the sampling. In this way, the space is adaptively sampled: promising areas are explored in greater detail, while regions with high misfit are progressively neglected.

Of course, none of these global methods are infallible, they all need knowledge and a priori information to be effective. In the absence of reasonable constraints for the parameter, the model space can become extremely vast making any sampling almost impractical.
Increasingly attention, especially in near-surface applications, has been made to the combined use of dispersion curve and independent measures into inversion procedure (e.g., Parolai et al., 2005 and Picozzi et al., 2005). Such an example is the integrated use of HVSR curve and dispersion curve to better constrain the iterative process toward the solution. Since dispersion is extremely sensitive to Vs variations and HVSR constrains the depth of the main impedance contrasts, their joint inversion allows for a more precise identification of the geometry of the 1D configuration. This approach is particularly useful in cases where the available frequency band does not allow both the surface part and the depth of the bedrock to be independently constrained. While joint inversion of dispersion curves and other independent measures can significantly improve the stability and resolution of the solution, it also introduces additional practical challenges and, in any case, the problem of computational cost remains. Each forward modelling requires solving the eigenvalue problem associated with the considered model, with dispersion curves calculation over a wide range of frequencies. A global search algorithm, even if efficient, typically requires thousands of evaluations to produce a meaningful set of models. If this process is to be repeated for tens or hundreds of local curves, as occurs when attempting to reconstruct a quasi-2D or -3D image of the subsurface from surface wave observations, the time required can quickly become prohibitive.

In Chapter V, the theoretical concepts discussed in Chapters II, III and IV will be formalized into implemented algorithms. The aim is not to directly repropose, or replicate methodologies already consolidated, but rather to coherently integrate the approaches that have been focused on throughout the theoretical discussion. In particular, the two main methods for extracting dispersion curves from pairs of stations (seismic interferometry and autocorrelation methods), the use of a linear tomographic approach for phase velocity inversion, and the implementation of a simplified procedure for dispersion curve inversion will be integrated. The latter will not be intended as a global search algorithm, but as an essential tool, operating in the neighbourhood of an initial model and aiming primarily to verify the consistency between the observed dispersion curves and the proposed velocity models, rather than to exhaustively explore the solution space. The Code Implementation chapter will then describe in detail the operational choices adopted and the ways in which the various theoretical components were put into practice. Subsequently, these implementations will be applied to a real case study, in order to discuss the results obtained and evaluate their overall effectiveness in the application context considered.

Chapter V.
Codes Implementation

Chapter V is devoted to the implementation of the theories introduced in the previous chapters. In particular, three code implementations are discussed, in response to the main phases of the workflow adopted in this work.
Section 5.1 describes principles and implementation of the joint extraction procedure of dispersion curves between pairs of stations (both phase and group curve). After the basic signal preprocessing, cross-correlation is computed in time- and frequency-domain, then phase and group velocity information are retrieved through automatic picking algorithm which is made on purpose of handle with inherent phase ambiguity of two-station dispersion analysis and exploits group curve to better constraint the choice of the right phase curve. Quality in the two-station approach is verified applying the methodology to different real context to define both quality and stability of the extracted dispersion curve. Picking algorithm is also tested for diverse two-station analysis, comparing dispersion curve automatic selected with the same curve picked manually.
Section 5.2 is primarily devoted to the tomographic imaging procedure, in which phase velocities are reformulated in terms of travel-times between pairs of stations and then inverted to obtain frequency-dependent Rayleigh wave velocity maps. The forward operator and the adopted model parameterization (data weighting and regularization strategy) are described, then the inversion in tested with two synthetic model in order to obtain reliable information about model resolution.
Finally, Section 5.3 focuses on the algorithm used for local dispersion curve inversion, that provide effective shear wave velocity profiles in a computationally efficient way, while maintaining a good consistency between experimental and theoretical data. Also, the dispersion curve inversion algorithm is extensively tested, since in real application it must handle with several local curves of various nature produced by the previous dispersion and tomographic analysis.
For all the implementations presented, validation tests are necessary because, as emerges from the previous chapters and will be further highlighted in this chapter, each procedures requires assumptions, approximations and intrinsic limits due to uncertainty, which must be carefully considered when interpreting the application results.

5.1 Dispersion curves extraction

Dispersion curve extraction (both phase and group velocities) is performed based on common two-station methods. Seismic signals are first pre-processed in order to remove the average from the records and applying an amplitude normalization with respect to the root-mean-square (RMS) amplitude of each individual signals. In this context, we're interested only in phase and group velocities, and amplitude information can be neglected, however, maintaining the phase of the original signals unchanged. All traces are then divided into a sequence of overlapping time windows, whose length and overlap are defined primarily based on the inter-station distances, indeed the distance between the receivers plays a key role within the dispersion curve extraction procedure. Distant stations typically require the use of larger time windows, to enhance the

frequency resolution. Conversely, for sufficiently close stations, shorter time windows can be used without compromising the quality of the estimates. The choice of the length of the time windows therefore represents a compromise: increasing the window length reduces the number of windows available for subsequent dispersion analysis, while shorter windows allow for greater data redundancy, which for ambient noise analysis is usually essential to improve the stability of measurements.

Once the traces have been divided, an amplitude threshold (defined as multiple of the signals average amplitude) is applied, and each window may be accepted, if the mean amplitude within the window is below the threshold, or rejected if the mean amplitude is above. In the subsequent step of windows correlation, only those windows whose average amplitude is below the threshold in both traces will be considered. In Fig. 5.1, an example of accepted 20s time-windows accepted according to an amplitude criterion is shown. Red portions of the signals correspond to windows not exceeding two times the signals standard deviations.

Once all useful windows have been identified and selected, each pair of synchronous windows is cross-correlated in the time and frequency domains. In the following sections, the two approaches are described separately.

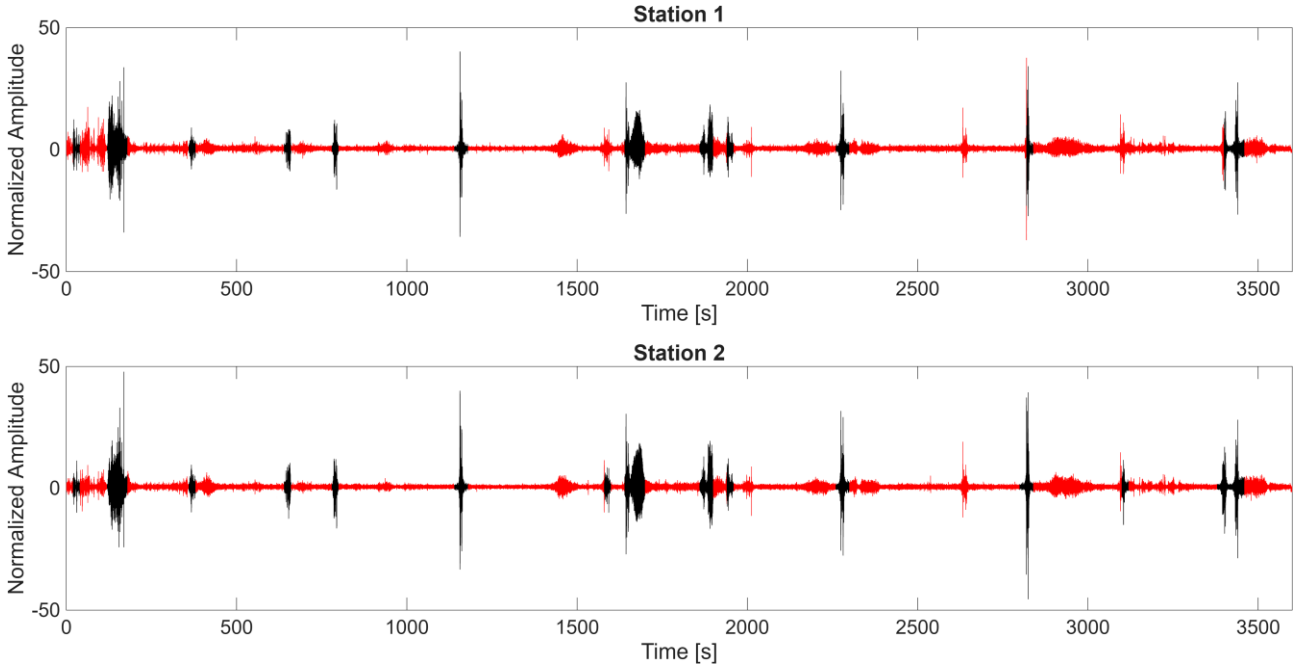

Figure 5.1 – Example of windows selection based on amplitude threshold for two seismic traces. Red portions of the signals have been accepted, while black windows have been rejected, showing a RMS amplitude greater than the threshold.

5.1.1 Phase curves

Frequency cross-correlation is performed between each couple of synchronous windows in order to retrieve the normalized cross-spectrum $\rho(\omega)$ (e.g., Malagnini et al., 1993; Parolai et al., 2006; Boxberger et al., 2011).

Considering the subscripts *i* and *j* denoting two different stations, the normalized cross-spectrum between two windows will be described as:

$$\rho(\omega) = \frac{\left(S_i(\omega) \cdot S_j(\omega)^*\right)}{\sqrt{\left|\left(S_i(\omega)\right)\right| \cdot \left|\left(S_j(\omega)\right)\right|}}. \qquad [5.1]$$

In Eq. 5.1, $\rho(\omega)$ is a complex function in which the whole amplitude information is removed by the normalization and only the phase of the signals is considered. $S_i(\omega)$ and $S_j(\omega)$ are respectively the auto-spectra of the signals and the superscript '*' denotes the complex conjugate. Then $\rho(\omega)$ is averaged over the N selected time windows and its real part is considered:

$$\bar{\rho}(\omega) = \frac{1}{N}\Re[\sum_{n=1}^{N} \rho_n(\omega)], \qquad [5.2]$$

In Eq. 5.2 the entire $\bar{\rho}(\omega)$ function represents the azimuthally averaged spatial autocorrelation (also known as coherency) computed over N windows and the individual values of the function are the SPAC coefficients obtained for each frequency.

For each time window, the magnitude squared coherence (MSC, which is not to be confused with experimental coherency) defined between 0 and 1 is also computed as a quantitative index of the phase coherence between the traces within the windows. Accordingly, the values of $\rho(\omega)$ are selected based on the corresponding MSC value at that frequency: when the coherence value at the angular frequency ω_0 is above a defined threshold, the relative value of $\rho(\omega_0)$ is used in the averaging process. On the contrary, if the coherence value at ω_0 is below the threshold, the $\rho(\omega_0)$ value in the considered window is rejected and it will not be used in the averaging process. This strategy of course reduces the number of observed $\rho(\omega)$ for each frequency based on the value of the coherence threshold, however, it allows to enhance the quality of the final autocorrelation function, since the contribution of incoherent frequencies within the windows can be progressively reduced. The effect of the coherence threshold on the averaged spatial correlation is clearly visible in Fig. 5.2: in the frequency band 15-20 Hz the procedure let two oscillations of the experimental function to emerge, whereas before they were masked by incoherent noise. It is worth noting that the entire procedure acts only on the amplitude of the experimental function, while the zero-crossing points remain practically constant, meaning that the phases of the signal are not changed.

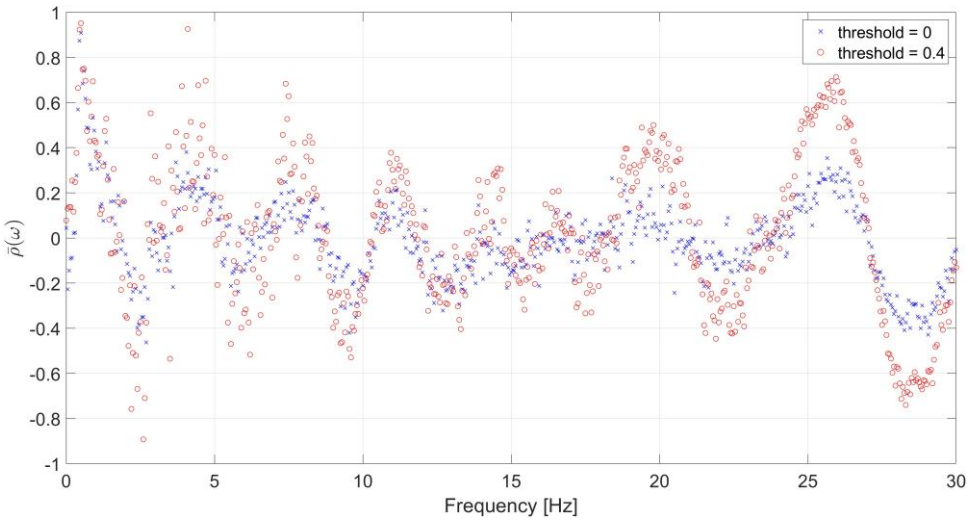

Figure 5.2 - Comparison between the azimuthally averaged spatial autocorrelation obtained without applying any coherence threshold (blue markers) and the one computed by imposing a threshold of 0.4 (red markers). The application of the threshold allows to improve the quality of the estimated function at higher frequencies, highlighting oscillations of the experimental curve that are partially masked by noise in the absence of threshold.

Once experimental autocorrelation has been achieved, a linear combination of cubic splines is used to fit the observed coherency (in a least square sense), yielding a smoothed function (see Boschi et al., 2013) that is employed for subsequent phase velocity estimation.

Phase velocity estimation is performed through a set of overlapping frequency windows across the observed coherency. For each window, a set of trial phase velocities, defined in a range c_{min}-c_{max}, is used as arguments of the theoretical Bessel function $J_0(r\omega/c)$ and the fit between experimental and theoretical functions is then associated with the central frequency (ω_0) of each window. The fit coefficient $\Phi(\omega_0, c_0)$ is formulated as a linear combination of two different coefficients, involving in a linear correlation coefficient and a root mean square residual (*rms*) parameter. The first one serves as a direct evaluation of the likelihood between experimental and theoretical functions, taking into account whether the functions are in or out of phase. The *rms* value acts as a weighting factor, increasing the correlation coefficient when the two functions are similar in terms of amplitudes, that is, when the oscillations of the experimental function have pretty the same scale as those of the theoretical one. Eq. 5.3 illustrates the fit coefficient formulation:

$$\Phi(\omega_0, c_0) = k_1 * corr[\phi(\Delta\omega), J_0(\Delta\omega, c_0)] + k_2 * (1 - rms[\phi(\Delta\omega), J_0(\Delta\omega, c_0)]) \qquad [5.3]$$

In Eq. 5.3, the fit $\Phi(\omega_0, c_0)$ depends on the considered frequency and the trial phase velocity. $\Delta\omega$ and ω_0 represent, respectively, the frequency window and the central frequency of that window and c_0 is the trial phase velocity for which the fit is evaluated. The fit is formulated normalizing both correlation coefficient and rms in the range [0,1], and linearly combining them through weighting factor k_1 and k_2, such as k_1+k_2=0.

In Fig. 5.3, a schematic example of the fitting procedure is shown. In panel (a) three different frequency windows are selected for central frequencies 3Hz, 12Hz and 22Hz, while, in the bottom panels (b), (c) and (d),

the best fit between experimental (black curve) and theoretical (red curve) functions is plotted within the windows. The best fitting values will correspond to three different phase velocities 3Hz, 12Hz and 22Hz.

Thus, by applying a moving frequency window and exploring a set of trial phase velocities, a frequency-velocity spectrum is constructed (Fig. 5.4), where the colormap directly represents the fit between coherency and Bessel function.

From Fig. 5.4, the concept of phase ambiguity previously introduced in Chapter III is slightly evident, especially at higher frequencies. Multiple curves do not represent fundamental or higher modes, among them only one curve is the true one and it should be considered as the effective curve produced by the combination of the propagation modes according to the real subsoil configuration. In normally dispersive contexts with no important velocity contrast ad depth this phase curve should be representative of the fundamental mode, however, the effects of higher modes may appear as slight deviations along the curves over the considered frequency range.

Experimentally, the inter-station distance of the considered receivers also strongly influences the phase ambiguity mainly acting on the shape of the coherency. The effect is clearly evident thinking to the inter-station distance as an argument of the theoretical Bessel function. As the distance between stations increases, this results in a larger number of oscillations of the theoretical function and the more the oscillations increase in the same frequency range and the more the number of potential dispersion curves increases.

In this sense, low velocities and small inter-station distances tend to mitigate the phase ambiguity, allowing for an easier identification of the correct curve at the expense of a lower resolution over the frequency range of interest. Conversely, when the distance between stations increases a shorter frequency window is typically more practical in the search for phase velocities, so as to highlight the details of the curve.

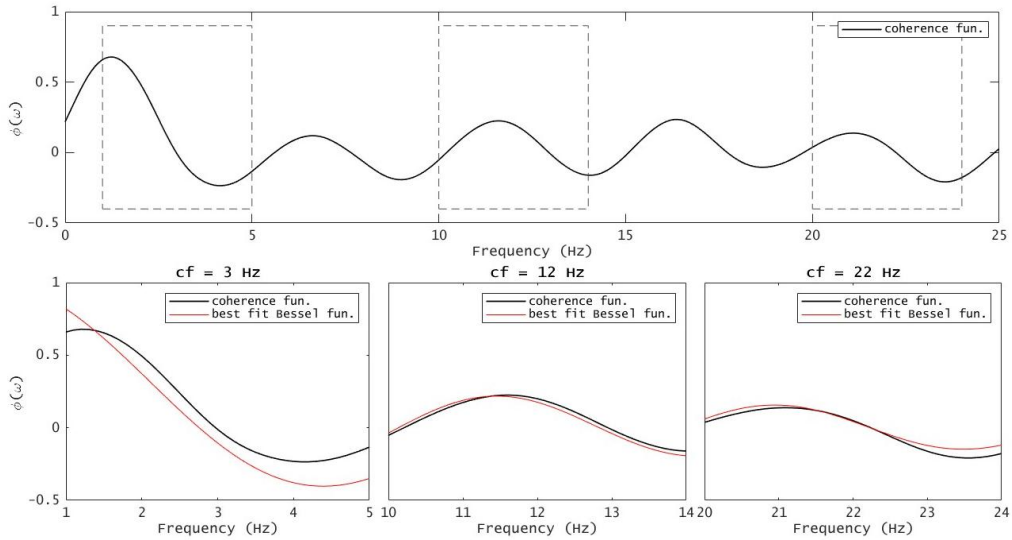

Figure 5.3 - Example frequency moving window applied to experimental coherency function for phase velocity estimation. The top panel shows the coherency, with three 4Hz-frequency windows highlighted by dashed rectangles. The lower panels illustrate, for three central reference frequencies (ω_0 = 3 Hz, 12 Hz and 22 Hz), the comparison between

the observed coherency (black line) and the theoretical Bessel function providing the best fit (red line) within each frequency window.

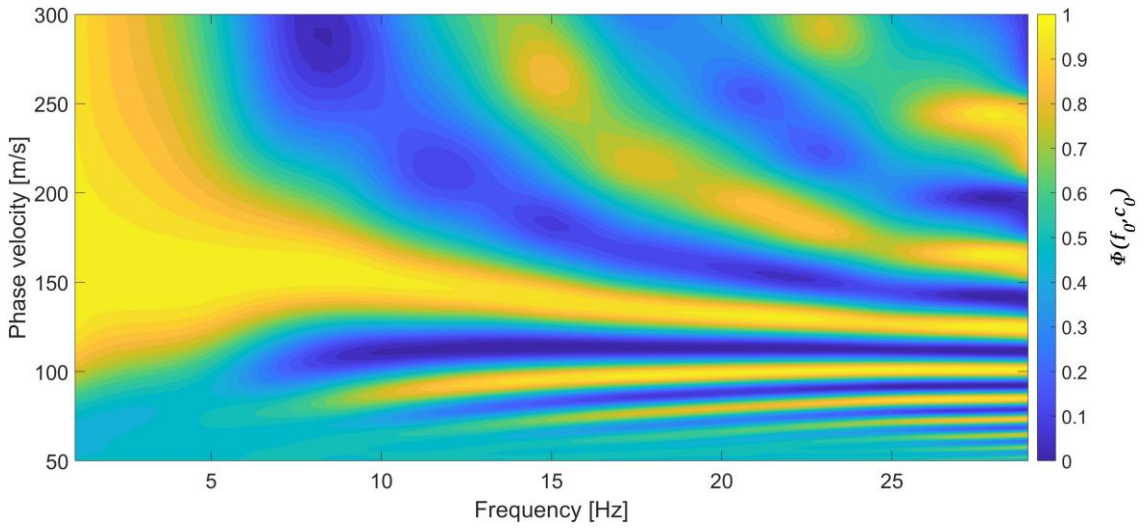

Figure 5.4 – Frequency-velocity spectrum computed through fit evaluation of experimental coherency using a 2Hz-frequency moving window. The colormap represents the value of the fit parameter $\Phi(\omega_0, c_0)$ between experimental function and theoretical Bessel function, calculated for a set of trial phase velocities in the range 50-300m/s. The higher fit values identify potential Rayleigh waves dispersion curves.

Once, the frequency-velocity plane (Fig. 5.4) is obtained, accurate automatic picking of the dispersion curve must be performed. For extensive application, such as the creation of phase velocity maps, tens or even hundreds of phase dispersion curves must be analysed, and this reflects the need of an efficient and reliable tool able to identify the right dispersion curve among the possibilities for each two-station analysis.
Furthermore, as already mentioned, increasing the interstation distance also dramatically increases the intrinsic ambiguity of the problem and multiple possible curves appear in the frequency-velocity plane in the same investigated frequency and velocity range. This greatly complicates the choice of the true dispersion curve compared to the example in Fig. 5.4. The algorithm for automatic dispersion curve selection will be discussed in Section 5.3, after presenting the procedure for extracting group velocities.

5.1.2 Group curves

While the extraction of phase dispersion curves is performed in the frequency domain, following the typical spectral approaches of the SPAC method applied to pairs of stations, the extraction of group curves, implemented simultaneously with the frequency domain analysis, focuses on the time domain cross-correlation between pairs of receivers. Within the time-domain approach, the goal is to directly process the recorded ambient noise so that the averaged time cross-corelation function approaches to a Rayleigh-wave Green's. As

in the frequency domain analysis, the time domain procedure to retrieve group dispersion information is based on the temporal stacking of ambient noise windows, i.e. on averaging independent noise realizations at both receivers. This procedure, unlike frequency-domain analysis, is entirely performed in the time domain, allowing the reconstruction of an ensemble cross-correlation that, as a first approximation, behaves as the impulsive response of the medium, providing insight about the propagation of detected surface waves from two distinct sites.

As already introduced in Chapters II and III, both phase and group velocities can be estimated from the cross-correlation between two stations. The Fourier transform of the cross-correlation function provides the cross-spectrum, which contains information about the phase delay between stations for each sampled frequency. However, even in this context the phase is ambiguously defined (it is reconstructed via an arctangent-type relation), making a phase unwrapping necessary to obtain a continuous representation. This step introduces an inevitable uncertainty related to the selection of the correct phase curve, just as in the case of the SPAC-type spectral approach. Furthermore, its automatic implementation is extremely challenging, making the application of automatic dispersion curve selection particularly uncertain.

This makes the ensemble cross-correlation for the extraction of group curves particularly useful, because it does not expressly require a change of domain since all the information about the dispersion of group velocities is contained in the function expressed in the time domain.

Whereas for phase velocity estimation, the cross-correlation between two stations is performed in the frequency domain as normalized cross-spectral product; to retrieve group velocities, the cross-correlation is calculated in the time domain for sequential synchronous time windows, based on the definition of cross-correlation function of signal u(t) recorded at x and y (Eq. 3.10).

Given the discrete nature of the recorded signals, the cross-correlation is expressed as a discrete function, where integration is replaced by summation. In Eq. 5.4 the function $C_{xy}[k]$ represents the normalized cross-correlation for a single time window.

$$C_{xy}[k] = \frac{1}{N} \sum_{n=1}^{N-1-|k|} u_x[n] \cdot u_y[n+k] \qquad [5.4]$$

Once the signals have been windowed and the cross-correlation is computed for each pair of synchronous windows as in Eq. 5.4, the resulting empirical Green's function (EGF) is chieved by averaging the cross-correlations over the windows (Eq. 5.5):

$$EGF[k] = \frac{1}{N} \sum_{i=1}^{N} C_{xy}^{i}[k] \qquad [5.5]$$

EGF is the empirical Green's function computed over N time windows, while the index k in both Eq. 5.4 and 5.5 is the vector index representing the lag-time τ according to the length of the cross-correlation window.
The empirical Green's function thus obtained contains information about the propagation of surface waves, but it is a complex combination of all spectral components propagating with their own velocity and to highlight these components, a multiple-filter technique is applied to it. This process allows narrow frequency components to be progressively isolated.
Time peaks of the filtered signals represent arrival times of coherent surface wave harmonics and provide the basis for group dispersion curve extraction.
Unlike the SPAC approach, where different plausible dispersion curves represent alternative solutions, approaching in the time-domain should allow us to identify contributions associated with diverse modes. Indeed, the multi-filtering process can highlight for a single frequency multiple arrivals attributable to different propagation modes. In practice, however, this is not frequent: for higher modes to be visible, they must propagate well separated in time and be sufficiently excited by ambient noise.
The time-based analysis consists of applying multiple Gaussian filters to the empirical Green's function, over a set of increasing central frequencies ω_0. Each filter is defined as a normal distribution in the frequency domain:

$$H(\omega) = e^{-(\omega-\omega_0)^2/2\sigma^2}. \qquad [5.6]$$

In Eq. 5.6, ω_0 is the central frequency of the filter, ω is the frequency vector, and σ is the standard deviation of the Gaussian window that controls the filter bandwidth. In order to highlight the energy maxima of the propagating waves, the envelope of each filtered signal is computed, the time-lag is converted in group velocity through the inter-station distance and the frequency-velocity plane, which provides a direct representation of the dispersive nature of surface waves. In Fig. 5.5, a scheme of the workflow for group curve extraction is shown, starting from the ensemble cross-correlation function in the top panel, the middle panel illustrates the set of filtered function and their envelope describing the energy maxima, while in bottom panel the same set of function is converted from frequency-time to frequency-velocity domain.
Once the frequency-velocity planes are obtained, automatic picking can be applied to retrieve both phase and group dispersion curves. The picking algorithm is the argument of next Section.

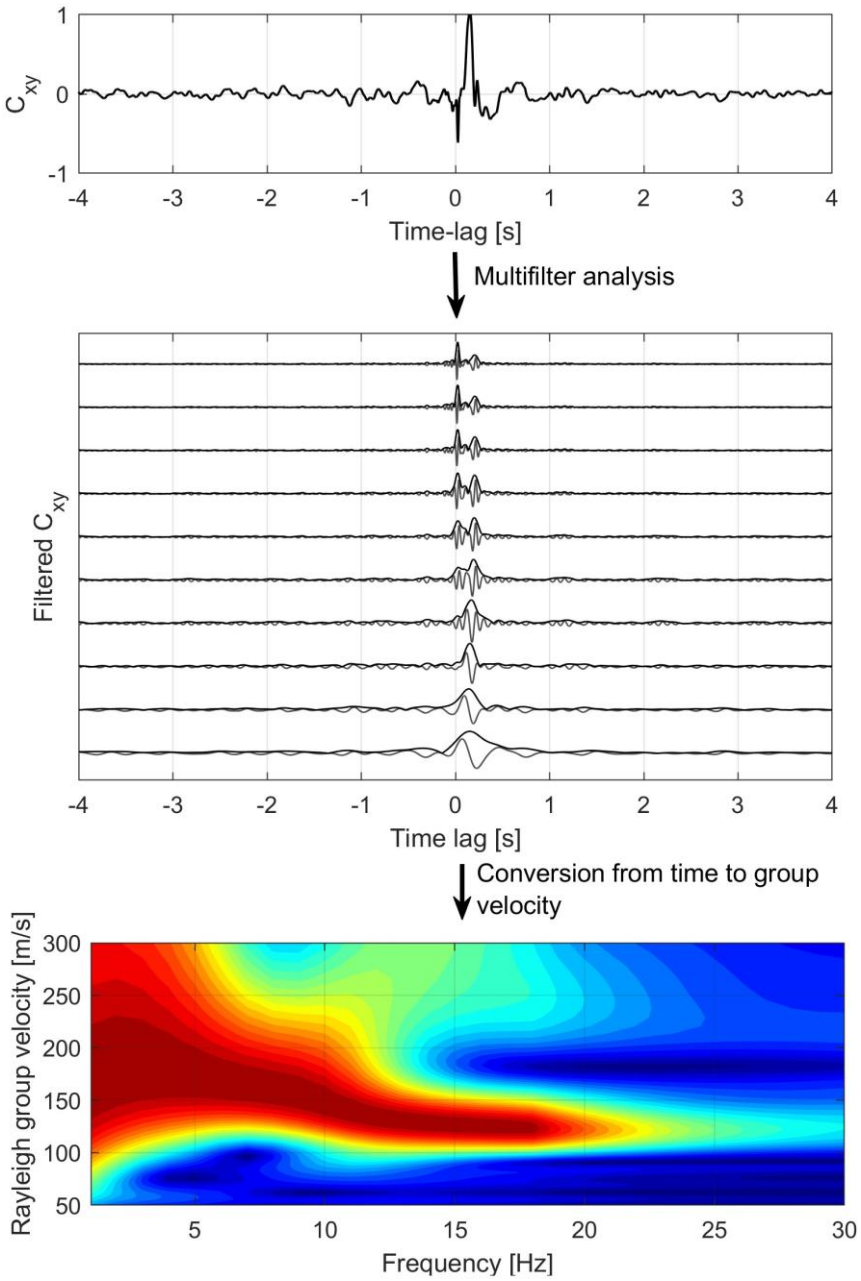

Figure 5.5 - Time-domain processing flow scheme for group velocity estimation. The top panel shows the normalized cross-correlation function between two stations. The central panel shows the filtered correlation functions for different central frequencies, obtained by multifilter analysis. The bottom panel shows frequency-velocity plane obtained converting time lag values into group velocities.

5.1.3 An automatic picking algorithm

The dispersion image obtained through Eq. 5.3 can be more challenging than the one in Fig. 5.4, where few potential phase dispersion curves can be identified. An automatic picking algorithm must consider the overall complexity of the frequency-velocity spectrum when many multiple dispersion curves exist, especially at

higher frequencies where the inherent ambiguity is much more pronounced and when the inter-station distance becomes larger.

Furthermore, it is not guaranteed that the different potential dispersion curves are always well separated over the whole investigated frequency range, as the example in Fig. 5.4. More frequently, a non-uniform distribution of noise in specific frequency bands can compromise the continuity of the curves, inducing local deviations of the experimental curve toward higher or lower phase velocities. In these cases, the picking may start correctly on a given curve but, at one or more frequencies, deviate to the wrong branch.

An automatic picking algorithm must therefore consider this behaviour by evaluating the continuity of the curve across adjacent frequencies and estimating the probability of remaining on the same dispersion branch, or to move to a different branch.

Typically, in large-scale applications, the phase ambiguity is addressed using a target curve to which the picking procedure can be linked, in order to select the curve that lies closest to the reference one: the picking generally starts at low frequencies, where experimental data are affected by minor ambiguity and, given the starting frequency, the selected phase velocity is the closest value to the reference curve. Once a branch is selected, if the space solution has good quality, the picking process may continue up to the higher investigated frequency.

It is therefore clear that a picking algorithm must necessarily deal with two different kinds of problem during the curve selection process: one of intrinsic nature, associated with the choice of the correct curve among the many possible ones, and another one of experimental nature, which is the uncertainty due to the presence of noise or non-ideal measurement conditions. These two factors can easily cause the algorithm to ‘jump’ from one curve to another one within the frequency band of interest, leading to completely incorrect results.

In the present work, the picking of the correct dispersion curve is not carried out solely according to the frequency-phase velocity spectrum, but also uses the information provided by the frequency-velocity plane of group velocities. In the latter case, the automatic selection of the group curve is generally simpler, since there are no significant ambiguities: at the same frequency, the algorithm must in fact choose only between three possible contributions, corresponding to the causal, acausal part or to the spectrum generated by folding the cross-correlation function. In ideal conditions of isotropic and uniformly distributed noise, the choice between these components would be substantially indifferent. However, in real-cases application, the noise field is often directional, and one of the two components (causal or acausal) tends to have a higher energy content, making the group curve more clearly identifiable. The picking of the group velocity is then performed frequency by frequency by selecting, at each frequency, the velocity value corresponding to the maximum energy in the frequency-velocity plane. In this way, it is possible to reconstruct an experimental group curve directly from the observed energy distribution frequency by frequency.

Once the group dispersion curve is obtained, the phase velocities can be selected. In the adopted implementation, the algorithm is designed to identify all possible dispersion curves in the frequency-phase velocity domain. For this reason, the picking process starts from the higher frequencies, where the ambiguity

is greater but where, at the same time, the maxima associated with all the possible dispersion branches are all generally present.

The first step involves in applying a mask to the frequency-velocity domain, defining a threshold below which the values of the fit parameter are set to zero. This operation allows for a distinct separation of the different dispersion curves. The threshold value is chosen empirically and represents the parameter that allows to distinguish the different curves in the domain. Once the mask is applied, the frequency-velocity plane is no longer interpreted in terms of individual peaks, but as a set of probability ranges for each frequency, which represent areas of plausible values of phase velocities.

Starting from higher frequencies, the algorithm identifies all these probability bands and, for each of them, starts a procedure to reconstruct a dispersion curve towards the lowest frequency. The transition from one frequency to the next is based on heuristic criteria. In particular, the selection does not take place simply by comparing the peak values, but by evaluating the degree of overlap between the probability bands at adjacent frequencies. For example, in the transition from 25 Hz to 24 Hz, the extent of the overlap between the two probability bands is verified: if the overlap is sufficient, the transition is accepted and the 24 Hz band is included in the curve under reconstruction. When overlap is small or negligible, the intermediate frequency is "freezed" and the continuity criterion is applied by directly comparing the starting frequency with the next available one (e.g. 25 Hz and 23 Hz). If the criterion is also met in this case, the 23 Hz probability band is selected, and the process continues towards next frequency. If more than two consecutive frequencies are frozen, the picking process for that specific curve is stopped and the algorithm moves on to the next curve, re-starting from the highest frequency. Conversely, if the process can continue uninterrupted up to the minimum sampled frequency, the curve can be considered as completely reconstructed in terms of range values of phase velocity for each frequency. Dispersion curves will be selected for each frequency as the peaks values within each probability range. By repeating this procedure for all probability bands identified at the first higher frequency, it is thus possible to obtain a set of automatically selected phase dispersion curves starting from the same frequency–velocity domain.

Fig. 5.6 shows the final result of the frequency-based dispersion analysis conducted on a pair of stations approximately 53 m apart. The automatic picking procedure allows to reconstruct the overall set of potential dispersion curves in the frequency range 2–20 Hz.

At lower frequencies, where the phase ambiguity between the different curves is less pronounced, the automatic algorithm is able to effectively separate the curves. Conversely, in higher frequency bands the contribution of noise can compromise the continuity of the curves: selection is sometimes interrupted already in the early stages, due to the lack of sufficient overlap between the probability bands associated with consecutive frequencies. It is also possible to observe how, among the higher velocity curves, a sudden jump around 17 Hz appears probably due to an increase in the noise level in that specific frequency band, rather than to a real effect related to the propagation of higher modes.

The presence of concentrated noise around 17 Hz leads to a loss of coherence making continuous selection of the dispersion curve more challenging and producing jumps in the automatic picking. These discontinuities do not reflect real variation in the properties of the medium and would be probably discarded during manual selection.

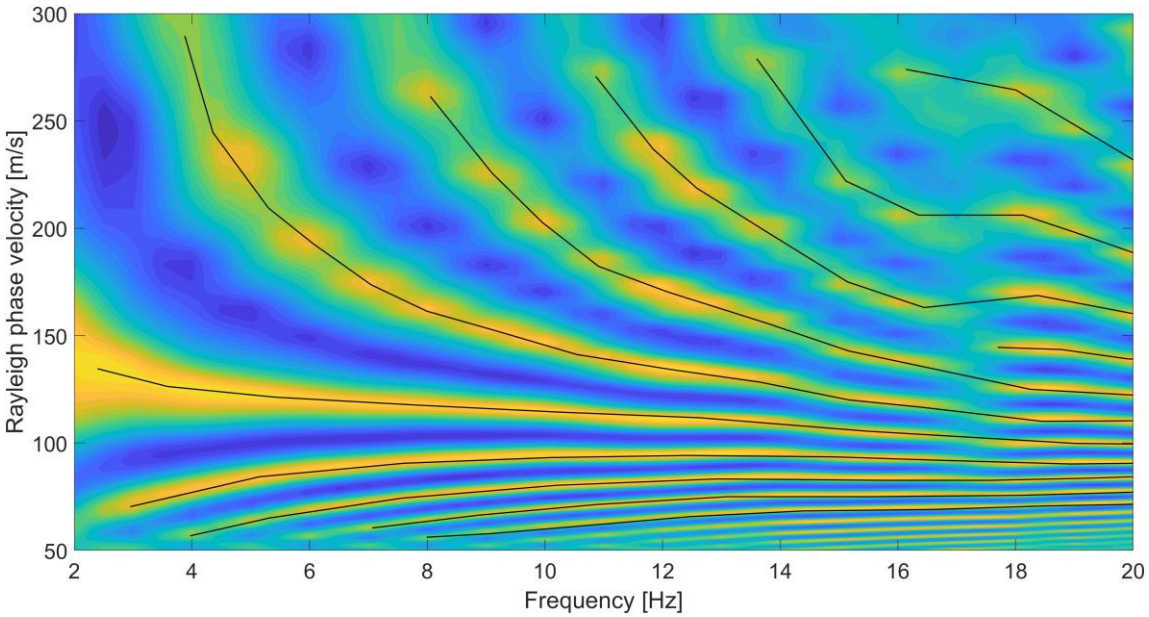

Figure 5.6 - Frequency-phase velocity spectrum obtained from the frequency-based two-station approach, with the automatically selected dispersion curves superimposed. The colormap represents the value of the fit parameter, while the black lines indicate the curves reconstructed by automatic picking.

Once all possible phase velocity curves have been selected, the corresponding theoretical group velocity curve is calculated for each of them by numerical differentiation, as reported in Eq. 2.11, thus obtaining a set of theoretical group velocity curves associated with the different phase solutions identified.
This set of theoretical group curves is then compared with the experimental group velocity curve derived from the time-domain analysis. The comparison is made in terms of the mean squared difference, and the curve that minimizes this residual is selected as the best solution, consequently providing the phase velocity curve considered most representative of the observed data. In the final selection, however, additional quality criteria are introduced: the extension of the phase curve with respect to the investigated frequency range and its continuity are also considered. Curves that are too short or limited to narrow portions of the frequency spectrum are discarded (in Fig. 5.6 for example a curve is composed of just two points), as are those that exhibit obvious discontinuities or unrealistic jumps. In this way, the automatic procedure favours solutions not only consistent with experimental observations, but also physically plausible over a large frequency range. Once the best solution among the phase dispersion curves has been selected, it is linearly interpolated in order to obtain a discretized curve with a step size of 0.5 Hz. Interpolation allows us to derive phase velocity values at the same frequencies for all pairs of stations analysed, making the data homogeneous and directly comparable. In this

way, the curves thus obtained can be used as input for the subsequent tomographic inversion step, aimed at constructing phase velocity maps.

5.1.4 Validation test for dispersion curves extraction

The dispersion curve extraction procedure, including the stability of automatic picking algorithm, was subjected to a series of tests in order to evaluating its behaviour under different experimental conditions. Two main tests were conducted in this framework: (1) curves extracted from individual pairs of stations were compared with theoretical curves derived from sites characterized by known velocity profile (or considered sufficiently reliable based on independent surveys, such as downhole surveys) and compared with experimental dispersion curves obtained from active-source multichannel analysis (MASW). This comparison made it possible to directly assess the reliability of the individual measurement with respect to a reference model considered reliable. (2) at an approximately 1D and laterally homogeneous site, the automatic picking algorithm was used to select right dispersion curves from numerous pairs of geophones. This test was aimed to verify the stability of the automatic curve estimation within the variability of distance and orientation of the pairs, so as to quantify the statistical reliability of the method.

DOWNHOLE TEST

In this first sub-section, three different field surveys using pairs of sensors were carried out at the same site at different times. As a first step, the repeatability of the measurements was evaluated comparing the results obtained with theoretical one computed by forward modelling of the velocity profile produced by independent downhole investigations. The Vs and Vp profile obtained by downhole investigations was considered as reliable reference model of the local stratigraphic configuration.

Fig. 5.7 shows the subsoil Vs and Vp model derived from the downhole tests (a) and, alongside (b), the corresponding Rayleigh dispersion curves, fundamental and first two higher modes. The latter represents the expected behaviour under conditions of lateral homogeneity, that is considered plausible for the area considered for the test. Seismic acquisitions were carried out on single stations using a Tromino™ 24-bit three-directional digital tomograph (Moho srl) and the sampling rate of 256 samples per second (256 sps) was used for all recordings. The three different two-station acquisitions were also processed considering different possible configurations, varying both inter-station distance and relative orientation. The experimental curves thus obtained were subsequently compared with the theoretical reference curve. Such comparison allowed to verify the consistency between the experimental results and the known model, as well as assessing the influence of the distance and orientation of the station pairs on the quality of the dispersion curve estimate. Fig. 5.7 shows three frequency-velocity spectra obtained from stations separated respectively by 20, 30 and 40m. The black lines represent Rayleigh wave modes derived starting from the velocity profile in Fig. 5.7(s).

Based on the comparison performed, it emerges that dispersion curves extracted from the different pairs of receivers, even at variable distances and orientations, are overall like each other. Moreover, the dispersion curves obtained from measurements carried out at different times show a acceptable stability, indicating that the individual analysis can be considered reliable.

One noticeable aspect relates the influence of the distance between receivers on the frequency–velocity spectrum. Given the same propagation conditions, by progressively reducing the inter-station distance, the spectrum preserves the ability to well identify the correct curve, with no abrupt and significant loss in resolution. This aspect is encouraging, since the effect reflecting the intrinsic difficulty in separating the phases should become predominant when the receivers are too close: waves propagating along different paths tend to overlap, reducing the clarity of the information contained in the spectrum.

Small deviations from the fundamental mode are clearly visible at frequencies around 6-10Hz for analysis conducted at 20m and 40m, where the fundamental mode curve takes velocity values slightly lower than that produced by the two-station approach. This can be interpreted as the potential effect of higher modes that tend to lift up the experimental velocities. Dispersion analysis carried out with inter-station of 30m seems to be the one which better captures the variability of the fundamental mode, suggesting that inter-station distances of that order might be the best compromise between uncertainty and resolution.

Overall, the area of uncertainty associated with the experimental curves can still incorporate the trend of the fundamental mode, confirming the validity of the procedure.

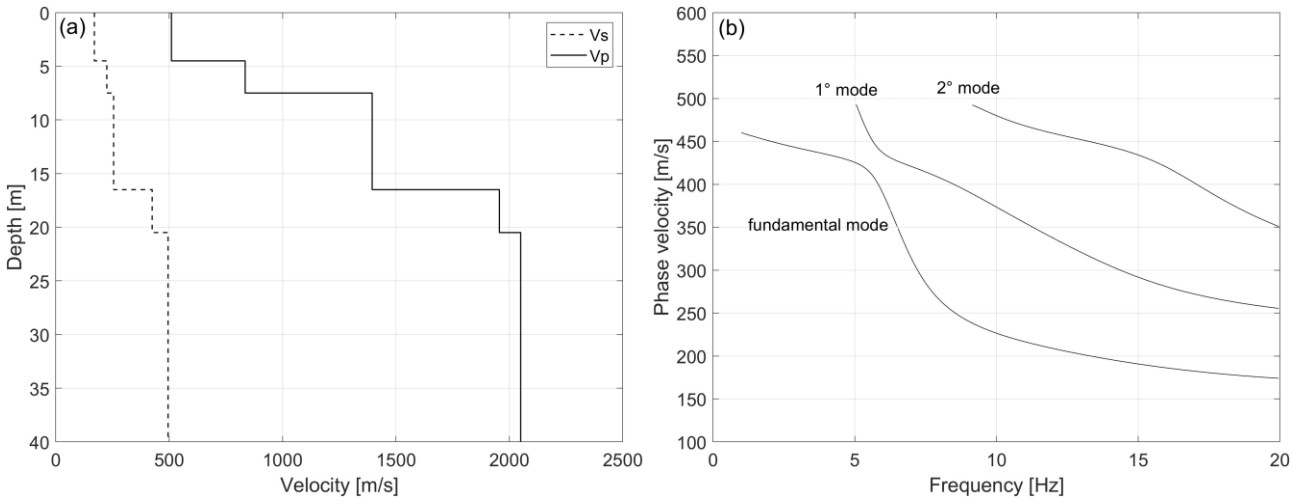

Figure 5.7 - (a) Stratigraphic velocity model used to compute theoretical Rayleigh wave dispersion curves provided by Down-Hole surveys. (b) Fundamental and first two higher modes of Rayleigh waves calculated for the stratified model in (a).

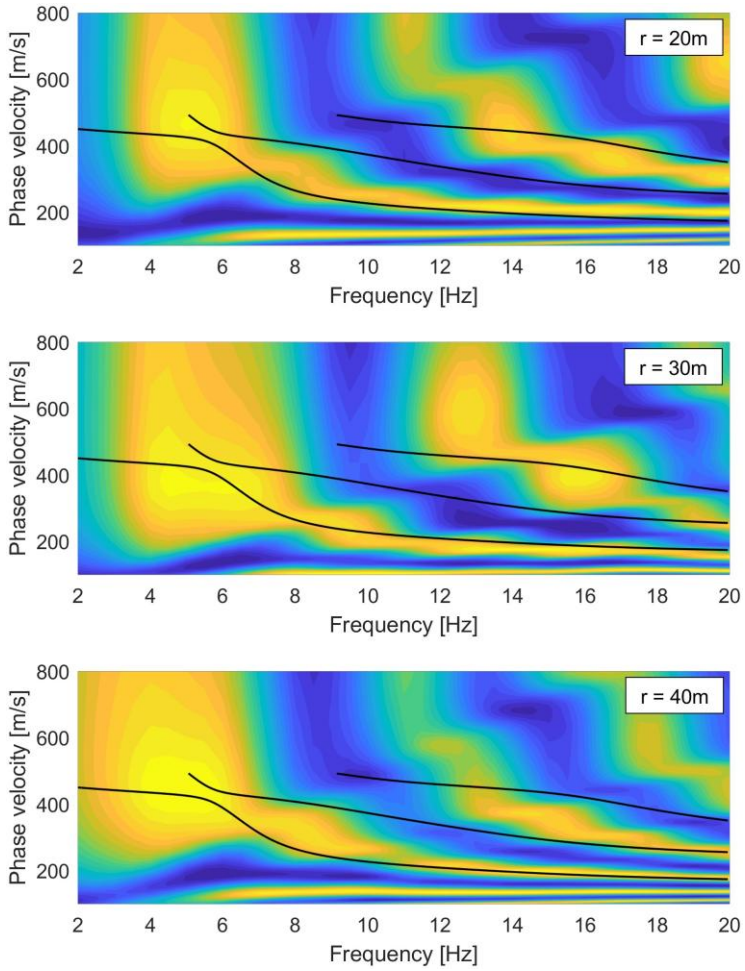

Figure 5.8 - Frequency–phase velocities spectra of Rayleigh waves obtained for different inter-station distances: r = 20 m, 30 m and 40 m (from top to bottom). The colormap represent the value of the fit parameter in the frequency–velocity domain, while the black lines indicate the theoretical dispersion curves calculated for the stratigraphic model considered, fundamental and first two higher modes.

MASW TEST

Two-station dispersion results were also compared with reliable outcomes from MASW techniques conducted at two different sites to assess the program's accuracy. Also for current analysis, phase dispersion curves were manually picked, which is highly recommended when handling with few curves.

Approximately 30 minutes of recordings were used for dispersion analysis. Within the MASW lines, which spans a distance of about 100 meters in both examples, multiple station pairs were tried to assess potential differences and test the stability of the results across varying inter-station distances. For all manually picked phase dispersion curves, synthetic group dispersion curves were computed to verify their reliability and ensure that the correct phase dispersion curve was selected.

In the first example, three analyses were conducted for different station pairs with varying inter-station distances: approximately 20m, 40m, and 80m, all along the MASW branch. The resulting Rayleigh phase dispersion curves from the three analyses were plotted and compared with the MASW-derived curve in Fig.

5.9. The dispersion curve derived from the intermediate station pair (black line in the upper panel), considering both the central curve and the upper and lower velocity limits, provides the best fit with the MASW results. As expected, velocity estimation exhibits the highest uncertainty at lower frequencies, where the central curve velocity values deviate from those of the MASW curves, particularly for the most distant and closest station pairs (red and green curves). Differences in velocity values at higher frequencies between the estimated (green curve) and MASW dispersion curves may be attributed to noise or real heterogeneities in the shallow subsoil between the receivers, such as the measurement at 20m. The results suggest that station pairs about 40m apart yield the highest reliability at both lower and higher frequencies, as they appear to better capture phase velocities provided by MASW.

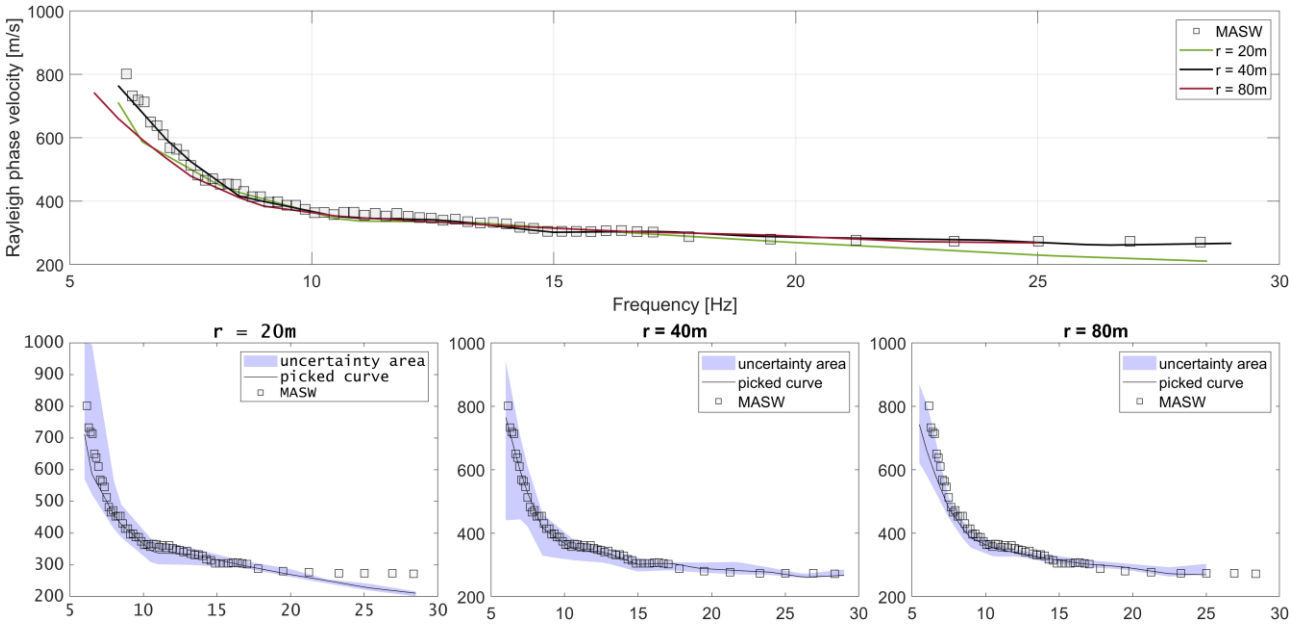

Figure 5.9 - Comparison of the phase dispersion curves obtained with the two-station approach and the results of the MASW analysis. The top panel shows the selected phase curves for different inter-station distances (r = 20 m, 40 m and 80 m) together with the MASW results (squares). In the lower panels each selected phase curve (solid black line) is compared with the MASW data (squares), while the shaded area represents the estimated uncertainty range.

Second example follows the same line of the first one. Even in this case three different station pairs have been tested (10, 30 and 50m) in order to check the stability of the dispersion results. In Fig. 5.10 the obtained dispersion curves are compared with MASW results.

Even in this case, considering uncertainty ranges for picked phase velocity, two-station technique has proven to get reliable results in terms of phase velocity. Unexpectedly, even with very close stations, it is still possible to identify the correct dispersion curve, despite a significant increase in uncertainty. Typically, such small inter-station distances would result in very low resolution, leading to a large uncertainty in the results. This is especially true considering that the lack of spatial sampling should theoretically cause difficulties in capturing the full range of frequency-velocity relations. However, despite this, phase analysis performed in the frequency domain still manages to reduce the inherent ambiguity and successfully identifies the correct dispersion curve.

This outcome suggests that, even with limited resolution, the phase analysis remains effective in distinguishing the true phase velocity, especially when compared to synthetic group velocity curves. As shown in the bottom left panel of Fig. 5.10, this phenomenon highlights the robustness of the analysis in mitigating phase ambiguity, even under challenging sampling conditions.

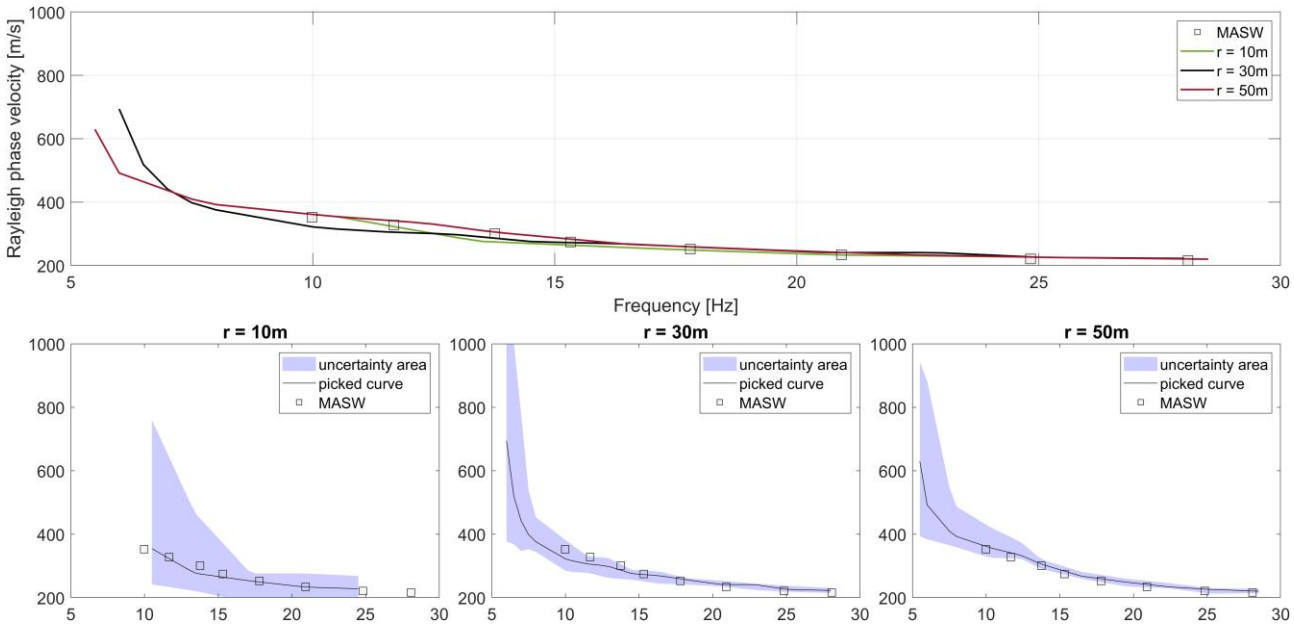

Figure 5.10 - Phase dispersion curves obtained by two-station analysis with inter-station distance 10 m, 30 m and 50 m. The top panel shows the selected phase curves against the dispersion curve provided by MASW approach. The lower panels show a more detailed comparison for each pair, highlighting the selected phase curve and its uncertainty range.

HOMOGENEOUS 1D MODEL

In order to further evaluate the reliability and robustness of the proposed approach, an additional test was conducted by applying the two-station analysis above a site characterized by a substantially homogeneous 1D model. For this purpose, an already available seismic acquisition, carried out independently of the present study, was used, which allowed the algorithm to be tested on a large set of station pairs included in the same array. For each selected station pair, the entire dispersion curve extraction, including the automatic picking procedure, was applied according to the criteria described in the previous sections. This test was designed with a dual objective: on the one hand, to verify the stability of the dispersion measurements for many different distance and orientation of the station pairs; on the other, to evaluate the ability of the automatic picking algorithm to correctly identify the dispersion curve even in the absence of manual supervision. The results obtained, shown in Fig. 5.11, highlight how the ability of automatic picking procedure to consistently reproduce the real dispersion curve over numerous pairs of stations of the same seismic array, confirming both the robustness of the two-station approach and the reliability of the automatic picking method when applied in a practically 1D context.

Approximately 20 pairs of stations characterized by different inter-station distances, were analysed. As shown in Fig. 5.11, the results indicate that the investigated site can be considered, as first approximation, almost 1D.

Beyond some variations over the investigated frequency band, which are generally inevitable for indirect measurements, the extracted dispersion curves present an overall coherent trend, all reproducing the same general dispersion behaviour. In the assumption of a 1D configuration, such a test also provides valuable indication of the expected experimental uncertainty about the phase velocity measurement. Considering the set of extracted curves in Fig. 5.11, an average curve representative of the 1D context can be estimated and the general variability of the dispersion curves around the average curve therefore provides an empirical estimate of the expected uncertainty associated with the phase velocity measurements.

By considering the average curve and estimating the maximum variability around it for each frequency, it is possible to further quantify the stability of the obtained measurements. The set of dispersion curves shows a general variability lower than 10% of the phase velocity value at the considered frequency. At lower frequencies (around 4 Hz), the relative variability can reach values around 12%, while at higher frequencies it tends to decrease, generally remaining in the range 4–7% of the phase velocity.

Overall, this test also confirmed two key aspects: on one hand, the good reliability of the automatic dispersion curve extraction and picking procedure; on the other, the ability of the proposed approach, also based on comparison with group velocity curves, to correctly identify significant phase dispersion curves at least in a 1D context, even in the absence of manual supervision.

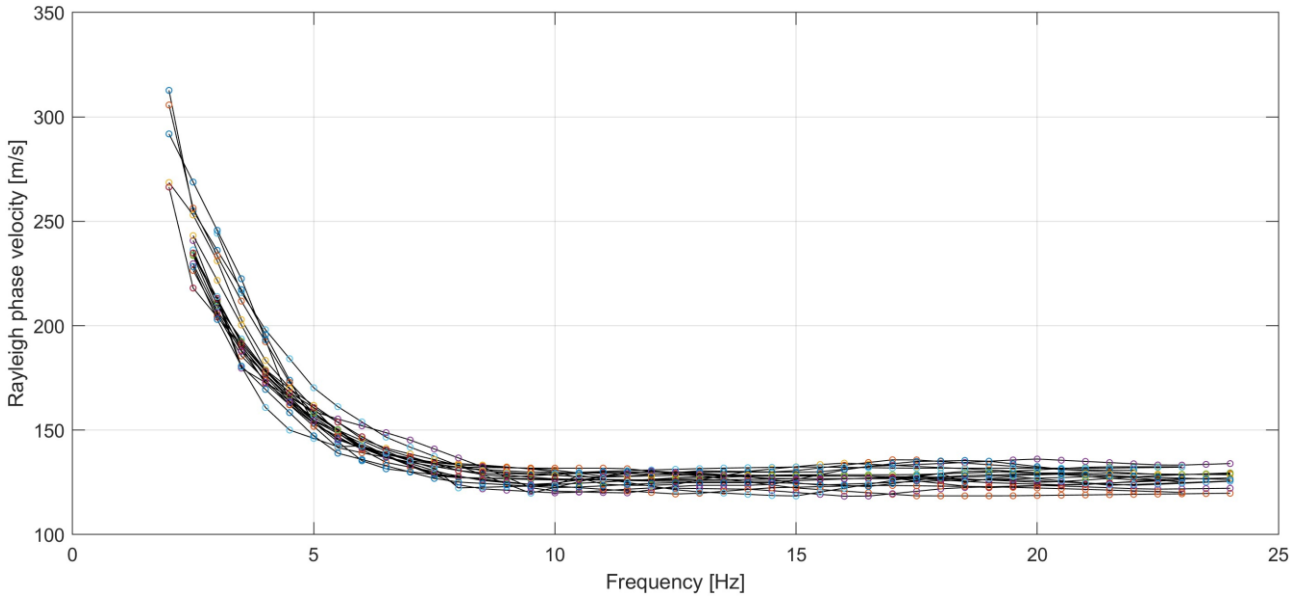

Figure 5.11 - Set of phase dispersion curves obtained by two-station analysis for 20 station pairs belonging to the same seismic array. The curves, automatically extracted for different inter-station distances and orientations, show an overall coherence across the investigated frequency range, suggesting an approximately 1D context.

5.2 Seismic Tomography

Implementation of seismic tomography is based on a relatively simple approach, following a common 2D linear travel-time tomography scheme, fully described in Chapter IV. Although the choice of describing the tomographic problem as a purely linear problem is very simplified compared to more refined approaches, its reliability in determinate context, with due approximations, has been widely proven.

The adopted forward implementation is based on ray tracing theory, assuming a purely linear ray path between the source and receiver (or, in this case, between receiver and receivers). This assumption reduces the overall complexity of the problem while maintaining a strong explanatory value. The 2D model is described in terms of a regular grid composed of square cells distributed over the entire spatial domain, each cell represents fundamental unit of the model (a model parameter), and the length of the ray path through it allows to populate the corresponding element of the forward matrix G.

The phase velocities extracted from the dispersion curves are converted into travel-times by dividing the inter-station distance by the phase velocity at each frequency. Given *n* stations, this yields up to *n(n-1)/2* travel-time observations per frequency, corresponding to all possible station pairs. The conversion from phase velocity to phase travel-time allows to treat a single picked phase velocity as a monochromatic wave with its own frequency propagating between two stations in the time *tt*. In this way, we obtained the first fundamental block of the tomographic linear inversion, which consist of the observed travel-time for each pair of stations.

Eq. 4.18 and Eq. 4.19 in Chapter IV represent a general form of a least squares problem, in which numerical stability is improved by the use of regularization strategy and prior information, and the quality of the final solution is constrained by the inclusion of weighting factors controlling the reliability of the observed data.

In this Section, the main quantities involved in the inversion implementation are described in detail: the starting point is the forward matrix G, which translates the geometry of the problem into the numerical terms required for analysis.

Subsequently, attention is focused on the data covariance matrix C_d, which allows different weights to be assigned to observed travel-time according on their uncertainty, and on the model covariance, which controls the degree of spatial correlation between model parameters and introduces a smoothing constraint on the final solution..

Finally, the role of the reference model, as a prior information, is discussed. It provides a target model which the inversion can converge to, linking the estimated model to the reference one and thus suppressing abrupt variations within it.

The forward matrix G, as a matrix linear operator, is constructed so that each element G_{ij} of the matrix represents the length of the path associated with observed travel-time i within a specific cell j of the domain (i.e. a specific model parameter j). Considering the total set of station-to-station paths, some cells are crossed more frequently, resulting in densely over-determined parameter, while others are almost never sampled by the propagation and produce zero-valued elements of the matrix $\boldsymbol{G}$ that do not influence the inversion process. In this sense, the matrix $\boldsymbol{G}$ is often a sparse matrix, containing many null elements. This sparsity, according to the geometry of the problem and the distribution of the stations within the 2D domain, directly reflect the spatial coverage provided by the stations network, as well as the resolution adopted to discretise the model in

individual parameters. A too fine discretisation, in the absence of a sufficient number of ray-paths, could produce a poorly constrained system and numerical instabilities, while a coarser discretization strongly reduces the resolution, but ensures higher numerical stability in the inversion.

The forward matrix $\boldsymbol{G}$, even within a linear inverse problem, can be subjected to strong numerical instabilities given its sparsely nature, and this aspect makes the joint use of model and data covariance matrices generally necessary to better constrain the solution during the one-step inversion.
The data covariance matrix $\boldsymbol{C_d}$ introduces a weighting factor into the inversion problem based on the effective reliability of the observed data. Here, observed data are indirectly derived from phase velocity observations, extracted frequency by frequency from the dispersion curves selection and subsequently converted into travel-times. Upper and lower uncertainty bounds are assigned to each curve, then they're propagated to travel-time uncertainties and used to populate $\boldsymbol{C_d}$.
As introduced in Section 4.1.3, implementation of data covariance matrix assumes a normal distribution of the errors around the average travel-time. Within this framework, all observations might be defined by just two numbers, mean and variance (or standard deviation). In real cases, this normal distribution of uncertainties cannot be demonstrated and is, in fact, unlikely. Within a given uncertainty range of phase velocities, the picked velocity value does not necessarily coincide with the centre of that range, as the uncertainty bounds are generally asymmetric with respect to the selected point. The travel-time uncertainties are obtained by propagating those in the picked phase velocities through the relation $t = d/c(\omega)$, assuming independent errors. Where d is the inter-station distance and c is the selected phase velocity, which is function of the frequency, the travel-time uncertainty can be expressed as:

$$\sigma_{tt} = \left|\frac{\partial tt}{\partial c}\right| \sigma_c \approx \frac{d}{c^2}\sigma_c. \qquad [5.7]$$

For each frequency, the uncertainties associated with phase velocity are estimated from the upper and lower bounds (c_u and c_l) and propagated to the travel-time using a linear approximation. The variances thus obtained are finally used to construct the covariance matrix of the C_d data, assumed to be diagonal, so that no correlation exists between different observations. Here, to mitigate the asymmetry of the uncertainties associated to the phase velocities, σ_c is computed as the average between c-c_l and c_u-c, since there're no guarantees about the symmetry in the phase velocity error bars. Moreover, even in the case the phase velocity c coincides exactly with the mean value between c_l and c_u, the nonlinear conversion still introduces a symmetry breaking of the uncertainty intervals once expressed in terms of travel-times.
Focusing on the model covariance, instead of common Tikhonov regularisation $\lambda^2\boldsymbol{L}^T\boldsymbol{L}$, the inversion code adopts a prior Gaussian information about the model, $m \sim N(m_0, C_m)$, which consists in the introduction of the term $(m-m_0)^T\boldsymbol{C_m}^{-1}(m-m_0)$ into the cost function (Eq. 4.18). The model solution becomes the model

minimizing the sum of the weighted quadratic residuals and the regularisation term defined by the model covariance $\boldsymbol{C_m}$:

$$m_{est} = \left(G^T C_d^{-1} G + C_m^{-1}\right)^{-1} (G^T C_d^{-1} d_{obs} + C_m^{-1} m_0) \quad [5.8]$$

The model covariance matrix, which represents a prior on the model, is described as an isotropic Gaussian spatial covariance matrix in the form of Eq. 5.9 (Fichtner, 2021):

$$C_m(i,j) = \sigma_m^2 \cdot e^{-\frac{|x_i - x_j|^2}{2l_c^2}} \quad [5.9]$$

where $|x_i - x_j|$ is the Euclidean distance between the centres of cells i and j in the 2D space model, σ_m is the model standard deviation controlling the amplitude of the expected variability around the reference model m_0, and l_c is a smoothing parameter (also known as correlation length) that governs how quickly the correlation decreases with distance across the model (the larger l_c, the smoother and more consistent the solution over longer distances). This choice makes it explicit and easy to interpret *how much* we want the model parameters to be correlated in space each other.

Compared to a classic Tikhonov regularization (which uses a damping parameter associated to a differential operator L), here we adopt a Gaussian prior (Tarantola & Valette, 1982) based on the correlation length and expected model standard deviation: in Eq. 5.8, we therefore replace the term $\lambda^2 L^T L$ with C_m^{-1}. The cost function is still defined by the quadratic form, the difference lies in how one regularises: instead of using derivatives (which impose local smoothing), we impose a Gaussian covariance on the model. Thus, neighbouring cells in the solution will be more similar (correlated), while distant ones will be roughly uncorrelated; the correlation length defines how smooth the solution should be. Operationally, σ_m is chosen according to the expected dispersion around the model parameters, while l_c is calibrated according to the resolution of the problem (typically some multiple of the cell size or based on the station coverage). Note that C_m thus defined is dense: unlike C_d, which is a diagonal matrix with no covariance between the observed data, C_m has no-zero element also off the diagonal. This results in more direct control of spatial correlation, at the expense of higher computational cost (which become negligible since we apply a one-step inversion).

Although σ_m and l_c have a precise interpretation in the Bayesian framework, representing the prior standard deviation of the model parameters and their spatial correlation length respectively (Tarantola, 2005), in practice they can be chosen empirically to achieve a stable and geologically plausible solution, balancing data fit and model smoothness

In many cases, we would like to impose physical constraints (e.g., velocities always positive or within a realistic range according to the geological context), but these constraints are not Gaussian: the cost function is no longer quadratic and there is no closed solution; constrained/iterative methods would be needed, or we could change the variable (e.g., work on log-velocities) so that positivity is guaranteed *by nature*. Since we

usually do not have reliable information on spatial correlation, we adopt a weak prior: we stabilize the solution without creating artifacts not supported by the data. In other words, we insert a very uninformative prior, favouring the spatial continuity of the parameters of the model.

In practice, the goal is to obtain models that are compatible with the data and stabilizing the inversion using prior realistic information about the model: the correlation length l_c (how similar two neighbouring cells tend to be) and the amplitude σ_m (permitted variability around the prior model). By reducing l_c, the model can vary more rapidly from cell to cell: the effect of the data on the model is improved, but the risk of overfitting increases; by increasing l_c, a smoother and more robust solution is obtained, but detail is lost (over-smoothing could hide small structures within the domain) and the general fit will increase as well. Similarly, a small σ_m produces a strongly damped (over-damped) problem, in which the solution remains close to the prior; a large σ_m disadvantage the constraint and, especially with a small l_c, can favour overfitting.

In practice, as first step before the inversion a trial set of values for l_c and σ_m must be explored (as illustrated in Fig. 5.12 from Fichtner, 2021), in order to find the best values. Useful operational criteria include, for example, the misfit-roughness curves (L-curve criterion), which tries to find the best compromise between data fitting and model smoothing. simple synthetic resolution tests, and a visual check of the plausibility of the structures. The key idea is to keep the prior as weak as possible, sufficient to stabilize the inversion without imposing features not justified by the data.

To conclude, it is useful to introduce the so-called posterior covariance matrix C_m^{post}, which quantifies the uncertainty associated with the estimated model parameters:

$$C_m^{post} = \left(G^T C_d^{-1} G + C_m^{-1}\right)^{-1}. \qquad [5.10]$$

The diagonals of C_m^{post} provide the variances (i.e. standard deviations) of the individual parameters in the final solution, while the off-diagonal elements describe the correlations between the parameters.

In our linear-Gaussian model, the posterior covariance is always less than or at least equal to the prior covariance C_m, because the data simply add useful information about the model. For this reason, it makes sense to compare, cell by cell (parameter by parameter), the variances on the diagonal of C_m with those of C_m^{post}: where the ratio $C_m^{post}(i,i)/C_m(i,i)$ is small, the uncertainty has been significantly reduced; where the ratio is close to 1, the data provide little information and the result remains dominated by the prior. Operationally, C_m^{post} allows to map uncertainty for the parameters and to possibly generate a population of quasi-random models $m \sim N\ (m_{est},\ C_m^{post})$ according to mean (m_{est}) and variance ($diag(C_m^{post})$) of the parameters distribution, in order to assess the stability of the final solution and propagate uncertainty on estimated parameters.

5.2.1 Synthetic test

In order to evaluate the robustness of the inversion of travel-times, two different tests were conducted on synthetic data. Two different synthetic models were evaluated, for each of them synthetic travel-times were computed between pairs of stations, assuming a straight-path propagation. Gaussian noise was added to the synthetic travel-times, in order to simulate possible uncertainty associated with the real data.

The first test model consists of a common checkerboard model, widely used as a benchmark in tomographic problems. It allows to verify the ability of the algorithm to identify periodic velocity variations and to highlight limitations in terms of spatial resolution. The second trial model, instead, is based on a more realistic scenario, designed to simulate a subsurface configuration closer to a plausible geological context characterized by lateral heterogeneities. This latter approach allows testing the effectiveness of the inversion procedure under conditions similar to those expected in real applications. In both cases, number and location of the receivers, as well as the size of the cells within the domain, defined according to the expected scale for the real-case application, were kept constant. For each model, the key inversion parameters, such as model standard deviation and correlation length, were chosen via *trial-and-error* approach, evaluating the most effective combination as function of the cell size and the coverage of the model considering the available travel-time paths. It should be also noted that, for each of the test inversions, the same reference model (m_0) was used as prior for the inversion. In Fig. 5.11 both synthetic models used for the tests are presented. In the checkerboard model, Fig. 5.11a, the model assumes only two velocity values, 100 and 200m/s, while the model shown in Fig 5.11b takes velocity that can vary gradually within the same velocity range (100-200m/s). The considered 2D domain takes the same dimension along the x- and y-directions (100m × 100m), and it is discretized into cells of 5m. The size of the cells could be of course reduced, but this would increase potential instabilities during the inversion, since the rays would sample a minor number of parameters and the solution space would risk not being explored adequately. As for the stations, 16 virtual receivers were generated, distributed along two parallel branches at the right- and left-boundary of the 2D domain.

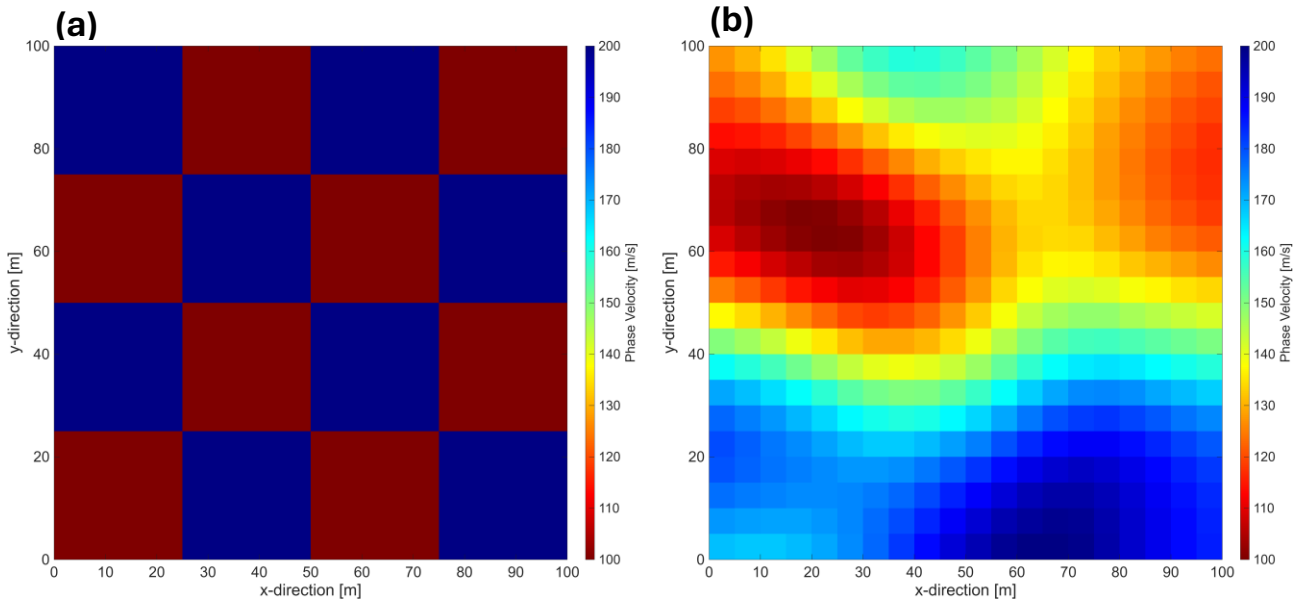

Figure 5.11 - True velocity models adopted for the two synthetic tests. The left panel (a) shows the model with an alternating distribution of high- and low-velocity cells, while the right panel (b) reports the smooth velocity model used to evaluate the inversion handling with more plausible lateral variations.

Considering the checkerboard model (Fig. 5.11a), which is characterized by clear velocity contrasts between 100 and 200 m/s, the estimated model after the inversion (Fig. 5.12a) appears able to recover the general pattern of the anomalies and therefore to reproduce the alternating high- and low-velocity structures of the model. It is clearly visible that the general alternation trend is practically returned in the same way as the synthetic true model. However, the combined effect of data noise (σ_d= 2e^{-2}s) and regularization term (σ_m = 0.0045s/m and l_c = 7.5m) produces significant smoothing boundaries between the checkboard elements: the sharp transitions of the true model become more gradual and less defined. The inversion tends to underestimate extreme amplitudes, producing intermediate values that attenuate the sharp contrasts of the true model. This behaviour is direct effect of the regularization strategy, which yields a more realistic model to be obtained compared to the true one, but at the same time suppressing velocity impedance contrasts.

Overall, therefore, the checkerboard experiment confirms that the inversion is able to reproduce the general pattern of anomalies and the main differences in velocity, it also highlights the typical limits of a regularized parameterization: loss of resolution corresponding to under-sampled parameters and smoothing of contrasts, effects that become more pronounced also depending on the amount of noise injected into the data.

Inversion applied to the model characterized by gradual velocity variations produces overall more satisfactory estimated solution (Fig. 5.12b) than the checkerboard test. The best solution is found implementing σ_m = 0.0018s/m and l_c = 7.5m as regularization terms, maintaining the same expected data uncertainty (σ_d= 2e^{-2}s). The two main velocity features, the low-velocity anomaly at the top and the high-velocity anomaly one at bottom of the 2D model, are reproduced clearly, with amplitudes and spatial distributions close to those of the true model. The gradual nature of the variations makes the problem more suitable to be treated with a regularized approach: the constraints imposed by the prior does not translate into a significant loss of

information but rather contributes to enhance the coherence of the final model. A further aspect to underline is that, even in the areas less sampled by the rays, the solution remains plausible: the cells not directly sampled by the data do not take anomalous values but adapt to surrounding trends (correlation between model parameters) or to the prior reference. This is a positive effect of the joint use of correlation length and model standard deviation, which the unsampled parameters to be constrained by the prior model and the surrounding cells. In summary, the realistic model test highlights how the approach adopted manages to reproduce not only the main anomalies but also the intermediate variations, with a good compromise between adaptation to the data and spatial coherence. Compared to the checkerboard model, the greater similarity between inversion and true model demonstrates that the methodology is particularly effective in contexts where the heterogeneities of the medium are distributed gradually, where common assumption about surface waves propagation can be considered.

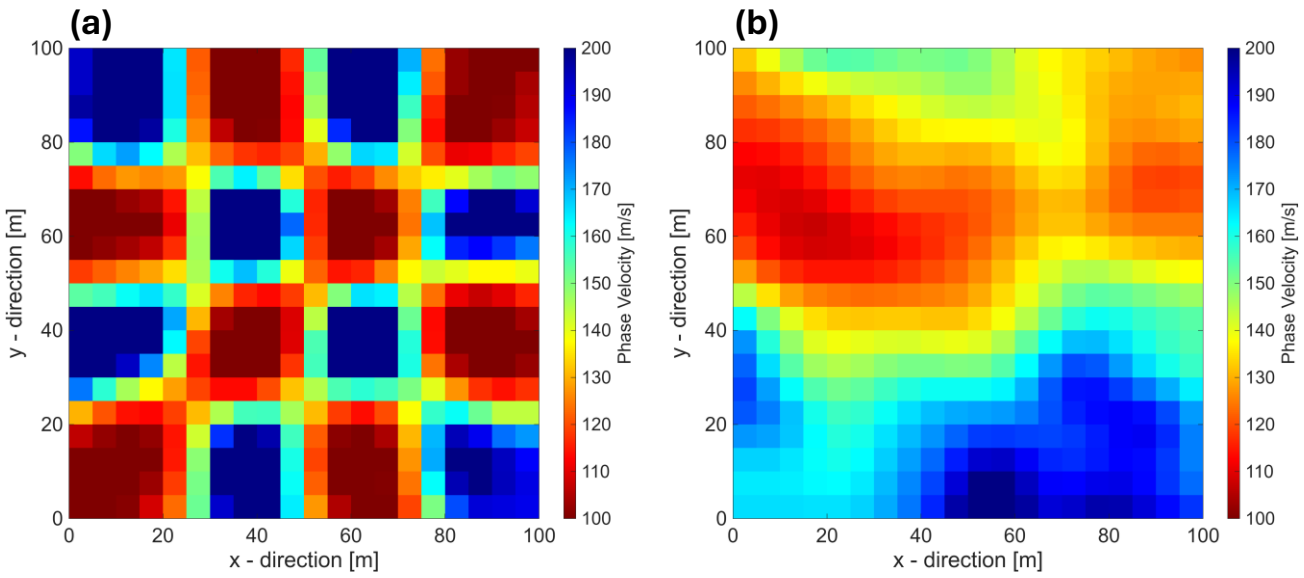

Figure 5.12 - Estimated velocity models for the two synthetic tests based on the true models shown in Fig. 5.11. Panel (a) shows the result obtained for the test with alternating high- and low- velocity blocks, while panel (b) shows the estimated model starting from the smooth model with gradual lateral variations.

The results show that, although the inversion is able to correctly reproduce the main features of the true models, the quality of the reconstruction is highly dependent on the regularization parameters, which values must be selected according to a qualitative procedure (such as L-curve criterion).

To discuss the effects of the regularization terms, dedicated simulations were performed for one of the synthetic models in which, in turn, one parameter between σ_m and l_c is kept constant while the other is changed. Fig. 5.13 shows the effect of changing the model prior variance at a fixed correlation length: too low values produce an over-damping, with excessively smooth models and increasing misfit, while too high values lead to under-damping and produced noisy models but reducing the overall misfit.

On the other hand, Fig. 5.14 shows the effect of gradually change the correlation length at fixed σ_m: for values that are too small there is under-smoothed models characterized by unrealistic and unstable variations, while too large values lead to over-smoothing, with loss of the main model anomalies. Here again, misfit grows above unity. Here again, misfit increases.

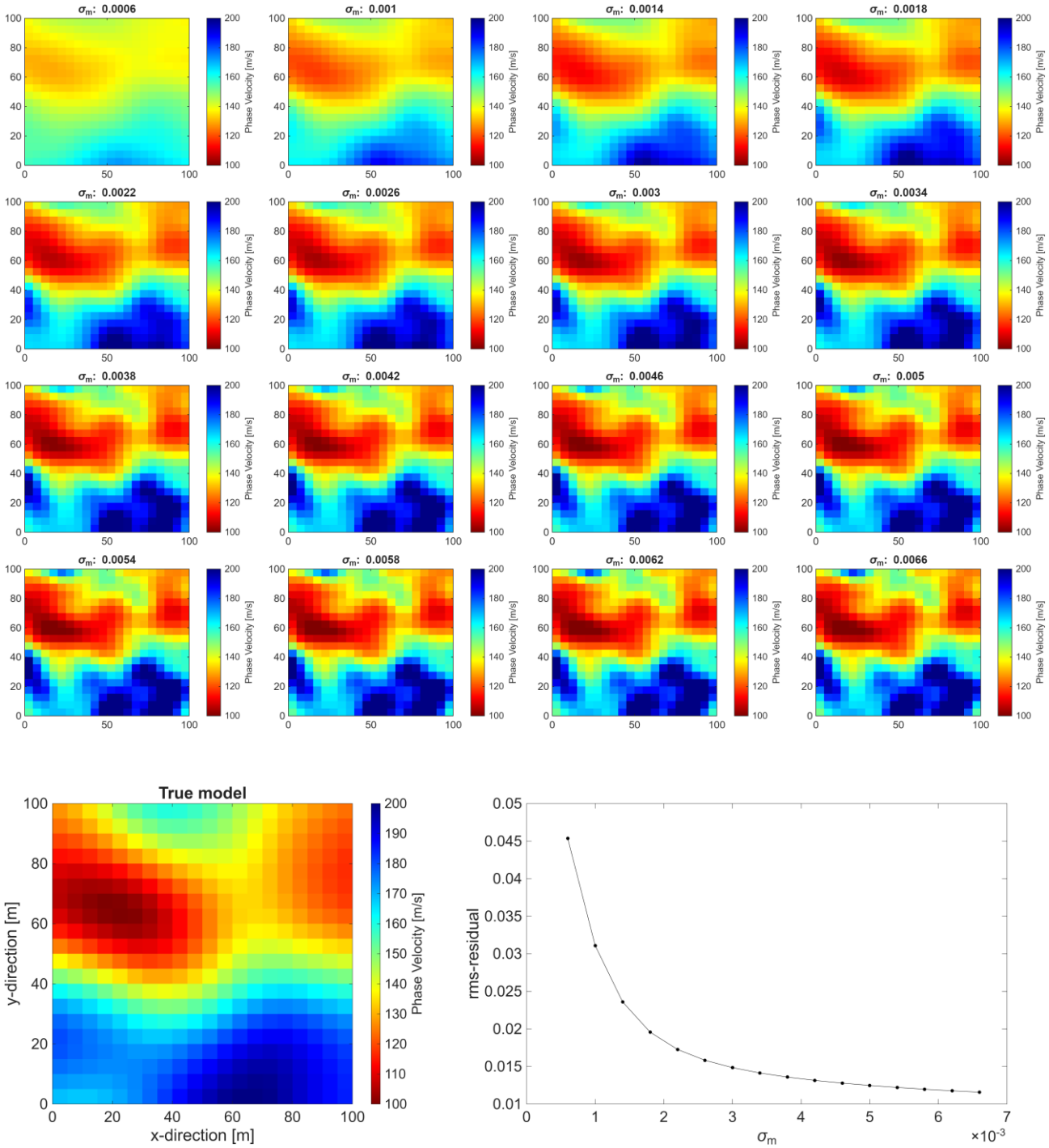

Figure 5.13 - Effect of changing prior variance on the estimated velocity model and the final inversion misfit. The upper panels show the models estimated for increasing values of the prior variance σ_m, highlighting the progressive increase in model complexity that reproduces more detailed structures. The true model is shown in the bottom left panel, while the rms-residual as function of σ_m is shown in the bottom right panel.

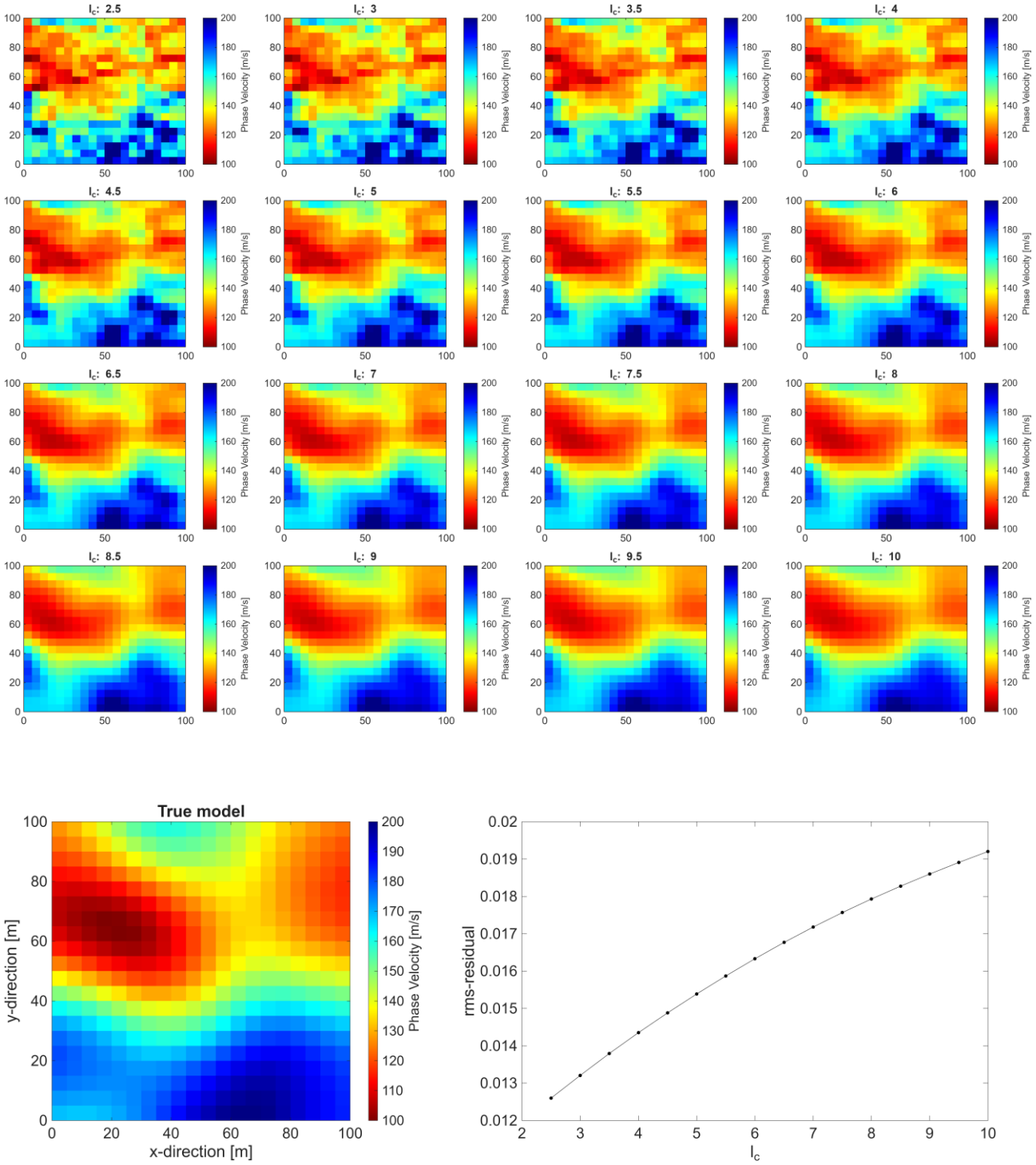

Figure 5.14 - Influence of different values of correlation length lc on the inversion results, keeping the a priori variance σm constant. The estimated models obtained for increasing values of lc (top panels) show a higher spatial coherence, accompanied by also higher values of the rms-residual (bottom right panel). The real model is shown in the bottom left panel.

In both experiments it emerged that the optimal values for the regularization terms do not coincide with the absolute minimum of the cost function, indicating that best solutions correspond to misfit values slightly higher than the minimum.

Allowing the model to freely adapt to the data, results in lower misfits but also to less reliable solutions: over-sampled parameters can be correctly estimated at the expense of the parameters located in less constrained areas that assume values inconsistent with the context. On the contrary, a stronger constraint to the prior or towards higher correlation values limits any kind of fluctuations, allows the model to maintain overall coherence even in under-sampled areas, but suppressing the effect of observed data.

It should be noted that, although the formulation adopted here borrows the notation of Tarantola (2005), the parameters σ_m and l_c are not derived from independent physical measurements as prescribed in the strict Bayesian framework. Instead, following common practice in applied seismic tomography, they are treated as empirical regularization parameters whose choice is guided by the characteristics of the problem rather than by direct physical constraints.

In the context of this work, σ_m and l_c are treated as numerical parameters that control the stability and reliability of the solution rather than as quantities derived from physical measurements Their choice is therefore strictly dependent on the type of involved problem and not on the absolute values of the model parameters, such as the velocity in tomographic problems. In particular, the overall size of the domain, the discretization into cells and the geometry of the seismic paths play a more important role.

The correlation length must be interpreted in relation to the discretization adopted: values equal to many times the size of the cell imply that velocity in adjacent cells is considered strongly correlated, imposing a certain degree of spatial regularity within the domain. σ_m, on the other hand, balances the relative weight of the data relative to the prior, stabilizing the inversion when the number of parameters grows (decreasing the cells size) or when the paths coverage is less uniform. Significant differences in the noise level or in the distribution of seismic paths may possibly require further adjustment of the parameters.

Finally, in this implementation an adjoint correlation parameter is included in the inversion as additional term in the model covariance construction. Following Eq. 5.9, we focused on the spatial correlation introduced by the correlation length parameter lc, which links model parameters each other assuming a natural smoothed model as solution of the inverse problem, damping sharp heterogeneities in terms of velocity of the model. But here, we are applying a surface waves tomography, which implies the use of phase velocities discretized frequency by frequency. In common approaches, as introduced in the Introduction of this thesis, inversion of phase velocities is performed for separately frequencies, and the result of the inversion is a set of phase velocity maps for each sampled frequency. In this study work we propose an algorithm developed to simultaneously consider all frequencies within the inversion scheme (as done before by Mohammadi et al., 2020): in this implementation a single model parameter (cell in the 2D domain) is divided into as many parameters as the number of investigated frequencies. Each set of phase velocities, specific of a frequency, has still its own data vector d_i, forward matrix G_i and model vector m_i, but the overall solution is found merging all vectors and

matrices in order to achieve a single inverse problem. So, given a N_f number of sampled frequencies, we can formulate the problem as in Eq. 4.2, but were d, G and m are expressed as:

$$d = \left[d_1, d_2, d_3, \dots, d_{Nf}\right]^T,$$
$$m = \left[m_1, m_2, m_3, \dots, m_{Nf}\right]^T,$$
$$G = \begin{bmatrix} G_1 & 0 & 0 & 0 & 0 \\ 0 & G_2 & 0 & 0 & 0 \\ 0 & 0 & G_3 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & G_{Nf} \end{bmatrix}. \quad [5.11]$$

The model and data covariance matrix are subjected to the same chaining procedure over the frequencies, but the model covariance matrix will no more be represented by Eq. 5.9. Conversely, a frequency correlation term is included in the formulation of C_m and model covariance becomes:

$$C_m(i,j) = \sigma_m^2 \cdot e^{-\frac{|x_i - x_j|^2}{2l_c^2}} \cdot e^{-\frac{|f_k - f_l|}{2l_f^2}} \quad [5.12]$$

In Eq. 5.12, $|f_k - f_l|$ represents the frequency distance between two parameters, that can be 0 up to the entire investigated frequency range, while l_f is the frequency correlation length which acts as the normal correlation length l_c. In this formulation, all phase velocities (or phase travel-times) are simultaneously inverted, and the minimization must also consider a frequency correlation between the parameters. All phase maps are obtained in only one step, and the effect of the frequency correlation parameter is to suppress strong fluctuations of the phase velocity values across the frequency range, assuming the natural continuous nature of dispersion curves.

5.3 Direct dispersion curve inversion for shear wave velocity profile

The traditional approach to surface-wave dispersion curve inversion involves in two main steps. First, dispersion curve is estimated experimentally from array-based methods such as SPAC or frequency–wavenumber analysis in the assumption of a simple 1D velocity profile without any lateral variations in terms of velocity and thickness of the layers. In the second step, the dispersion curve is inverted in order to infer a stratified model representing subsoil 1D structure (see for example Foti et al., 2011). As already introduced in Chapter 4, the problem of dispersion curves inversion is strongly nonlinear. For this reason, commonly adopted approaches are based on global search methods, which do not identify a single solution, but rather they explore a set of plausible models that can reproduce the experimental dispersion curve. Although such methodologies are generally robust, they are also extremely computationally expensive. In this context, where a large number

of local dispersion curve must be inverted, on the order of hundreds of curves, the systematic application of global search techniques becomes effectively impractical.

In this sense, the results of the tomographic inversion scheme are a group of cells composing the 2D domain, each of one characterized by a set of frequency-phase velocity pairs. Practically, each cell has its own local dispersion curve provided by the linear tomographic inversion. This procedure also implies that the number of final local dispersion curves is a function of the number of model parameters used in the tomographic inversion and handling with a 2D domain means that the number of the cells grows inversely to their size. A small and simple square domain of 50m size, two different cell dimensions just like 5m and 2.5m lead to 100 and 400 model parameters, which reflect in the same number of local dispersion curves to be inverted.

An alternative approach is here proposed to retrieve an effective Vs model solution for each obtained local dispersion curve. It starts from the local dispersion curves, and it is based on an iterative procedure, applied independently to each point of the tomographic grid. The goal is not to solve a global inverse problem in a statistical sense, but to obtain, for each model parameter, a shear-wave vertical model Vs(z) consistent with the Rayleigh phase velocity curve derived from tomography, while maintaining an acceptable computational cost.

The starting point of the inversion is the local dispersion curve. This curve is considered as the observed data to be reproduced using a one-dimensional stratified purely elastic model. The local dispersion curve is exploited not only as an observation to be reproduced, but also as a preliminary indication of an initial model. With this aim, some simple assumptions on the link between dispersion curve and stratigraphic structure are introduced. In particular, a linear relationship between frequency (or wavelength) and investigated depth is assumed, which allows the experimental curve to be converted into a set of depths associated with phase velocity values.

The relationship between frequency and depth is assumed to be constant over the entire frequency band and this represents a strong constraint on the method, since it implies a constant penetration coefficient not independent of the subsurface structure. Previous studies (Socco et al., 2017; Socco & Comina, 2017) have shown that the penetration depth of surface waves is not at all constant but depends significantly on the stratigraphy: velocity distribution (especially Vs) and the presence of strong impedance contrasts, produce a wavelength/depth relationship that is far from being linear. In this context, the penetration coefficient cannot be interpreted as a unique physical quantity, but rather as an effective parameter that allows to provide an initial parameterization of the model consistent with the scale of the observed wavelengths.

For these reasons, its value is not chosen arbitrarily, but it is therefore evaluated based on the ability to reproduce the observed dispersion curve, while remaining aware of the approximate nature of this assumption. Starting from this depth conversion, the stratigraphy can be expressed as a series of overlapping layers, each characterized by a thickness defined as the difference between successive depth values. In parallel, the observed phase velocity is converted into an initial estimate of the average shear wave velocity across the profile. Vs values are obtained accordingly on the assumption that, in near-surface contexts, the velocity of the fundamental Rayleigh mode is systematically lower than the effective shear velocity of the medium. The

velocity profile thus obtained does not yet represent an interval velocity in the strict sense, but a cumulative average velocity VsH(z), which is subsequently converted into a stratified Vs(z) profile by imposing consistency between thicknesses, depths and vertical travel times:

$$Vs_i(z) = \frac{h_i}{\left(\frac{z_i}{VsH_i(z)}\right) - \sum_{z_0}^{z_i}\left(\frac{h_i}{Vs_i(z)}\right)} \quad [5.13]$$

In Eq. 5.13, h_i represents the thickness of the considered layer i and z_i is the depth up to the same layer i. At this point $Vs_i(z)$ is the initial model in terms of shear wave velocity profile.
In Fig. 5.15 an example of derived $Vs_i(z)$ from experimental dispersion curves is provided. Dispersion curve is first converted in depth-phase velocity pairs and then in depth-VsH(z) profile. Through Eq. 5.13 VsH(z) is transformed in a stratified Vs(z) model.

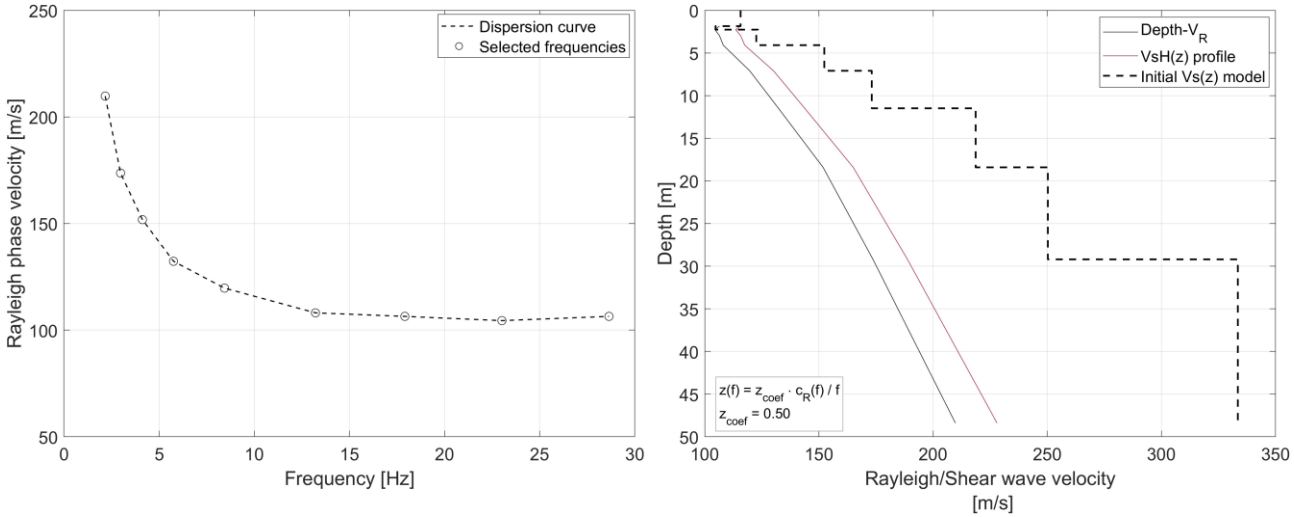

Figure 5.15 - On the left, a local Rayleigh wave dispersion curve obtained from a discrete set of selected frequencies (black circles). On the right, frequency–depth association derived assuming a linear relationship between wavelength and investigated depth. The V_R-z points are used to construct an initial profile of VsH(z) profile and then Vs(z) stratified velocity model.

Once the initial model is defined in terms of thickness layer and Vs profile, the remaining properties required for forward dispersion modelling are assigned. In particular, the Vp is derived from the Vs values via a constant Poisson's ratio, while the density is defined according to expected values for the investigated context.
This also represents a strong approximation, since in real contexts the Poisson's ratio can vary significantly as a function of the lithology, the degree of saturation and the state of consolidation of the materials.
Assuming a constant Poisson's ratio allows us to reduce the number of free parameters and make the inverse problem computationally advantageous, but it inevitably introduces an additional constraint into the model.
The forward problem is then solved for Rayleigh waves through Delta-Matrix method according to the current model (Vs, Vp and density), computing a theoretical dispersion curve based on the initial assumed model.

During forward modelling, both the fundamental and the effective curves can be considered. Accordingly, the inversion scheme will be based on the fit between observed curve and predicted curve, which may correspond either to the fundamental or the effective curve. The obtained theoretical curve is then interpolated on the same frequencies as the observed curve, so as to allow a direct frequency-by-frequency comparison.
The residuals between the observed and computed curve are computed as a function of frequency, in the form of normalized difference between the two curves:

$$res(f) = \frac{V(f)^{exp}_{Rayleigh} - V(f)^{teo}_{Rayleigh}}{V(f)^{teo}_{Rayleigh}} \quad [5.14]$$

The residuals vector will be composed of the individual misfit computed for each considered frequency across the dispersion curve. Starting from these residuals, a global misfit function is defined, as the root-mean-square value of the individual residuals over the entire frequency range. This synthetic indicator in Eq. 5.14 allows to progressively follow the residual variations at individual frequencies which describe the local fit and the contribution of individual layers and to quantify the global agreement between the model and the observed data. The Vs model is iteratively updated, using a correction rule proportional to the residuals obtained at each frequency. In practice, the shear wave velocity associated with each layer is increased or decreased as a function of the sign and amplitude of the residual at the corresponding frequency, applying a scaling factor that controls the magnitude of the step. As in other classical approaches, even in this implementation the layer thickness is kept fixed during the iterative procedure so as to avoid trade-off processes between velocity and thickness of the considered layer.
In order to improve the stability and convergence of iterative model updating a regularization has been introduced in the Vs correction based on the residual values. First, the calculated residuals in the frequency domain are reinterpreted in the depth domain (within the stack of stratified layers) through a smoothing procedure.
Instead of implicitly assuming that the residual at a given frequency acts only on a single “layer”, it is assumed that it represents the integrated effect of a depth range, consistent with the physics of surface wave propagation, whose sensitivity is vertically distributed and not localized.
This results in a vertical distribution of residuals more aligned with the not strictly one-to-one relationship between frequency-depth. In practice, after associating each frequency with an effective depth, the residuals are not used just as punctual measure of the misfit but redistributed along the vertical assuming that each residual represents the integrated effect of a depth range. In this way, the residuals are transformed from a frequency representation to a continuous depth distribution, which can be used for model updating. The final residual used in the update process is computed as a weighted combination of the residuals at nearby frequencies (in terms of effective depth). The iterative process continues until one of the convergence criteria is reached. In particular, inversion is stopped when the maximum number of iterations is reached or when the misfit values between successive iterations is lower than a defined threshold, indicating substantial

convergence to a model that reproduces observed data. A further criterion is reached when all percentage residuals associated with the frequencies involved in the inversion end up below a pre-established threshold, e.g. 1%. Such a condition indicates an overall good fit with experimental curves, for which all residuals are less than the 1% of the phase velocity values. Finally, in the presence of divergent or unstable behaviours of the global misfit, the inversion is stopped and the resulting profile is marked as unreliable.

At the end of each iterative process, the best solution is selected as the model corresponding to the minimum misfit value, that not necessarily coincides with the last iteration.

In addition to the final Vs(z) and Vp(z) profiles, diagnostic information such as the evolution of the misfit and the intermediate profiles generated are preserved. These elements allow for a critical assessment of the quality of the solution and allow for the identification of any problematic or ambiguous cases.

Within a 2D scheme, i.e. where each inversion represents a model profile defined over the domain, each profile is obtained via an independent inversion, but the spatial coherence provided by the tomographic inversion ensures continuity of the final model also in terms of Vs.

Although the proposed approach does not rely on an explicit formulation of sensitivity kernels as in common inversion strategies, the obtained Vs(z) represents an equivalent solution, that may reproduce the local dispersion curve within the accuracy of the data, rather than a unique and physically exact estimate of the elastic properties of the subsoil The observed convergence in several cases suggests that the procedure, while approximate, produces patterns consistent with the information content of the dispersion curves.

Despite the practical effectiveness of the proposed approach, it is necessary to underline some intrinsic limitations. First, the conversion from wavelength to depth is based on the assumption of a constant penetration coefficient, which represents a significant simplification of the real behaviour of surface waves, whose vertical sensitivity depends on the stratigraphy and is not constant over the frequencies. Second, updating the velocity profile from the curve residuals does not explicitly account for the distribution of sensitivity kernels, and therefore does not allow rigorous localization of Vs variations in depth. Furthermore, assuming a constant Vp/Vs ratio along the profile implies a uniform Poisson's ratio, which may be unrealistic in the presence of different lithologies or saturation variations. Finally, the method provides an effective velocity model that reproduces the observed dispersion but does not guarantee the uniqueness of the solution. These limitations must be taken into account when interpreting the results, which should be interpreted as an equivalent representation of the subsoil consistent with the information contained in the dispersion curves.

In Appendix A, Fig. A1, three tests for inversion code validation are performed using velocity profiles derived from downhole tests. Starting from these profiles, the corresponding dispersion curves were evaluated and subsequently inverted with the described algorithm. In each panel, the comparison between target dispersion curves (derived from the downhole profile) and modelled curves is shown, as well as between the velocity profile obtained from the inversion and the original one obtained by the downhole. The panels highlight how the inversion is overall faithful to the reference profile in the first surface levels, while divergences are observed at greater depths. This behaviour is partly attributable to the parameterization, which provides a progressive increase in layer thicknesses with depth, and a consequent reduction in vertical resolution compared to the

detail downhole profile. Despite this, the inverted model appears to be able to capture the general trend of the reference profile, well reproducing the vertical behaviour of the shear wave velocity (Vs) main variations.

Chapter VI.
Case study

Algorithms and procedures described in Chapter V are based on principles introduced in Chapter II, III and IV and constitute the reference methodological framework for real-data application presented here. Their detailed description provided a correct interpretation of the application results, allowing for a full understanding of the assumptions, approximations, and operational choices that characterize the entire workflow. In particular, the three main methodological procedures, the two-station dispersion curves extraction from seismic noise, the tomographic inversion of phase velocities and the local inversion of dispersion curves into shear wave velocity profiles, will be applied to a real-case study conducted at the Salse di Nirano reserve. In this context, the techniques are employed both in 2D array configurations, in order to obtain interpretable quasi-3D velocity models, and along linear profiles of geophones, to reconstruct 2D sections of the subsoil.
The entire analysis provides estimation of the shear wave velocity distribution beneath the site, a key parameter for subsurface characterization and for interpreting the active processes in the study area. Therefore, the following Chapter is devoted to a first description of the geological context, the presentation of the field data acquisition and the obtain results, which are finally interpreted in the framework of existing geological and geophysical knowledge, previous studies and observations.

6.1 Context

Mud volcanoes are surface manifestations of overpressured, fluid-rich sediments that rise toward the surface generally through faults and fracture systems. These processes are often linked to rapid diagenetic processes and the generation of fluids at depth (Kopf, 2002). As these fluids ascend through the sedimentary successions, they give rise to mud breccias or diapiric mélanges accompanied by saline water, hydrocarbons, and large quantities of gases (Accaino et al., 2007; Mazzini et al., 2009).
These systems can remain active for very long geological times and are largely influenced by the structural and stratigraphic conditions of the basin, which regulate both the local overpressures and the ways through which fluids rise towards the surface.
The activity of sedimentary volcanism systems, although potentially very long in time, is characterized by marked variability in time. Gas emissions associated with mud volcanoes can in fact occur either through continuous and low-intensity degassing processes or through short but energetically relevant eruptive events, during which large volumes of gas are released into the atmosphere or into the water column in very short times (Judd, 2005). This variability occurs as alternations of quiescent and paroxysmic phases of activity, as well as through transient events of outgassing or material emission. Such variations are commonly interpreted as the result of variations in the overpressure state of the fluids and the permeability of the medium, often associated with processes of fracturing, sealing or reactivation of the upwelling pathways (Mazzini et al., 2009; Mazzini & Etiope, 2017). Sedimentary volcanism systems also show considerable variability in terms of size and morphology, ranging from small surface structures, sometimes limited to a few meters cones or vents, to

large-scale structures that can extend for kilometres and therefore are fed by potentially enormous reservoirs. In some regions of the world, notably Azerbaijan, mud volcanoes represent large and complex systems, composed of multiple buildings and diffuse degassing areas spanning several kilometres (Kopf, 2002; Mazzini & Etiope, 2017).

In recent years, growing interest in mud volcanoes has emerged due to their implications in energy exploration, seismic hazard, and environmental impact (Martinelli & Panahi, 2003). In particular, it became clearer that the dynamics of mud volcanoes is mainly controlled by the migration and release of underground fluids and gases and the release of greenhouse gases (especially CO_2 and CH_4) from these structures raises questions about their contribution to the global atmospheric methane budget (Mazzini and Etiope, 2017). Nowadays, variations in fluid pressure and variations in outgassing flows towards the surface are certainly recognized as the main driver of eruptive processes, responsible for the generation of observable signals, such as surface deformations, gas emission variations and transient or permanent seismic signals (Mazzini et al., 2009).

This overpressure fluid migration is frequently associated with medium fracturing processes and the reactivation of pre-existing fractures and discontinuities within the sedimentary successions. Indeed, numerous studies (Milkov, 2000; Dimitrov, 2002; Kopf, 2002; Bonini, 2007) have highlighted how the location of these volcanic mud systems and their outgassing systems is closely controlled by the structural and tectonic context, in particular by the presence of major faults, folds and crustal weakness areas acting as preferential path for the fluids rising. These tectonic processes, at very different scales, not only influence the spatial distribution of the surface manifestations, but can also modulate their activity over time in terms of outgassing and represent a direct source of numerous seismic signals, which can manifest themselves either as impulsive events linked to brittle ruptures during fluid ascent, or as more continuous signals associated with phenomena of opening, closing or resonance of the fractures themselves (Mazzini & Etiope, 2017).

In some cases, changes in outgassing intensity or surface activity have been observed induced by the propagation of seismic waves generated by near seismic events, suggesting a plausible interaction between subsurface fluid dynamics and seismogenic processes that cause a sharp change in the stress state.

Most mud volcano systems are offshore, beneath the sea surface, where any kind of direct measurements are very challenging. Even in onshore settings, emissions often occur not only from the main cones, but diffusely across numerous minor vents and mud pools distributed over a broad area (Oppo, 2011). These emissions are driven by the complex dynamics of subsurface flow, which can laterally migrate the surface emission sites from the deep fluid reservoir by several kilometers (Thrasher et al., 1996).

Despite these difficulties, sedimentary volcanism systems have been intensively investigated through a wide variety of approaches and disciplines (see the edited volume by Martinelli and Panahi, 2003, and references therein), ranging from geological and structural studies, punctual geochemical investigations of the gas emissions, gravimetric and magnetometric investigations, as well as through active and passive geophysical measurement campaigns. Each approach provides complementary insights into degassing processes and subsurface structure, contributing to an integrated view of complex and highly heterogeneous systems.

Within this multidisciplinary framework, a growing number of studies have explored the use of seismic observations as proxies for monitoring gas release at mud volcanoes and similar degassing systems (Albarello et al., 2012; Antunes et al., 2022; La Rocca et al., 2023). These investigations have revealed analogies between cold (mud) and hot (magmatic) volcanic systems. In particular, 'drumbeat'-like seismic signals, typically related with episodic degassing events, have been observed in both types of settings (Iverson et al., 2006; Kendrick et al., 2014; Lupi et al., 2016; Giovani et al., 2017; Antunes et al., 2022; La Rocca et al., 2023).
These studies suggested that seismic monitoring may provide important information about the dynamic processes responsible for outgassing in mud volcano systems, making their application particularly attractive.

In Italy, approximately 60 continental mud volcanoes have been recognized, mainly concentrated along the outer edge of the Apennine chain and in Sicily region (Martinelli & Judd, 2004). Among these, the Natural Reserve of Salse di Nirano (Fig. 6.1a) site is one of the best-known mud volcano systems within the Apenninic thrust belt and can be considered as a small-scale natural laboratory, where the amount of pre-existing data and documented degassing activity provide the conditions for detailed investigations. The site is located approximately 40 km west of the city of Bologna.
Over the past few decades, the Salse di Nirano site have been the subject of several geophysical multidisciplinary studies aimed to understand the complex subsurface structure and the surface evidence of sedimentary volcanism, both due to the fluid migration processes.
The local geology consists of gently folded, Plio-Pleistocene marine clays (the "Argille Azzurre" Formation), occasionally interbedded with sandy lenses, which overlaps Epiligurian and Ligurian units, primarily composed of sandstones, claystones and conglomerates). At the base of the entire stratigraphy, around 2km beneath the surface, there is the Marnoso-arenacea Formation (Bonini, 2008).
In particular, the site and its evolutionary dynamics appear to have been strongly influenced by compressional structures and thrusts of the northern Apennines. Indeed, the site is located on the top of an anticline NW-SE oriented included into a larger monocline with a predominantly NE-dipping, in the direction of the Po Valley (Gasperi et al., 2005).
In this context, buried faults and zones of structural weakness are interpreted as preferential pathways for fluid migration towards the surface that interrupt the continuity of the Plio-Pleistocene clays, helping to explain both the concentration of outgassing evidence and their spatial distribution within the main oval depression (Bonini, 2009; Lupi et al., 2016). This complex setting suggests locally deformation even at very shallow depths consistent with an articulated fluid migration system.
Geomorphologically, the Nirano Sauces field shows an evident sub-elliptical depression (Fig. 6.1b), generally interpreted as the result of gravitational collapse processes linked to the persistent emission of mud and fluids from the underground. The flat bottom of the basin hosts a group of four main mud vents rising about 3 meters in height, together with several minor vents and pools, characterized by continuous bubbling activity, that are spread out across the site. These several smaller vents and mud pools are more diffusely distributed, reflecting the presence of a complex feeding system and multiple upwelling paths toward the surface, as also observed

in other Italian sedimentary volcanism systems (Martinelli & Judd, 2004; Kopf, 2002; Etiope et al., 2007). The ejected clay materials progressively cover the bottom of the depression, producing mudflow deposits of variable thickness, which contribute to mask surface tectonic structures and make the area to rapid changes its surface morphological setting.

This articulated surface morphology suggests that the entire system has experienced much more intense phases of activity in the past than today, with relevant implications for understanding its dynamic evolution and the functioning of the deep reservoir (Bonini, 2009; Capozzi & Picotti, 2010).

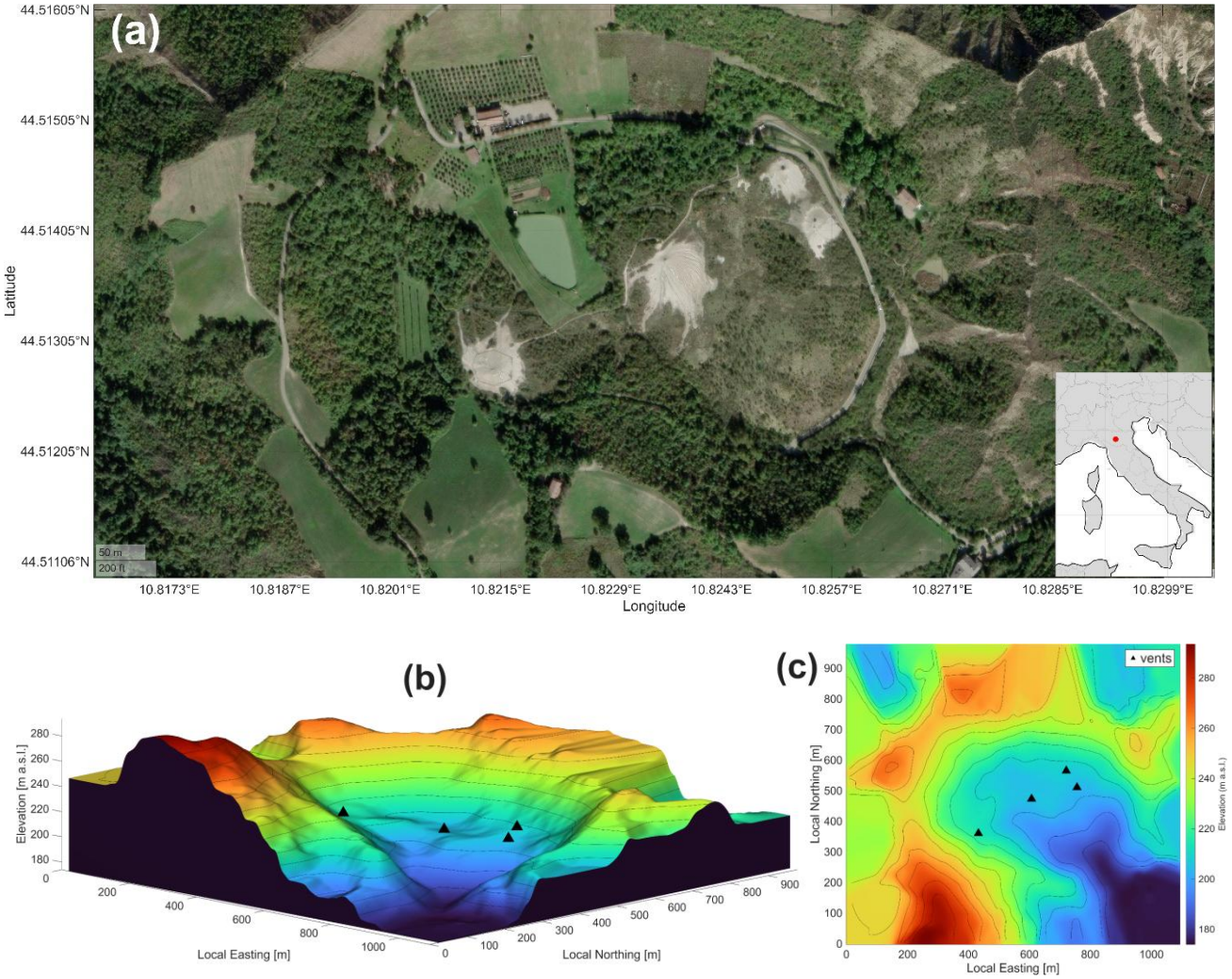

Figure 6.1 - Geographic representation of the Salse di Nirano site. (a) Satellite view of the study area, indicating the location of the Nirano site in the local geographical context; the insert shows the location of the site on a national scale. (b) Digital terrain model of the Salse di Nirano area represented in local coordinates (Easting and Northing). The topographic surface highlights the sub-elliptical depression morphology of the mud volcano field. The black triangles in both (b) and (c) panels indicate the position of the main vents. (c) Plan view of the same digital model, visualized using a colormap and elevation isolines, highlighting the lateral topographic variations and the spatial distribution of the vents with respect to the geometry of the depression. These representations allow to appreciate the morphological complexity of the study area, characterized by strong surface heterogeneities related to the processes of mud emission.

Over time, the Salse di Nirano system has been subjected of several studies carried out from different disciplinary fields. In particular, the site has been extensively investigated from a geological and structural point of view (Bonini, 2007; Bonini, 2008), in order to reconstruct its morphological evolution, as well as from a geochemical investigation (Capozzi & Picotti, 2010; Sciarra et al., 2019), with focus on the composition and variability of gas emissions. In particular, geochemical investigations have provided a detailed description of the spatial distribution and temporal variability of gas emissions. The works of Sciarra et al. (2017; 2019) highlighted a marked heterogeneity of CO_2 fluxes, in line with the presence of buried structural discontinuities. These outcomes suggested that outgassing at the Salse of Nirano reserve might be dominated by localized preferential fluid migration paths, rather than diffuse large-scale processes. These studies confirmed that gas emissions, even if relatively modest in intensity, are not concentrated at the main vents but rather distributed over a broader ENE–WSW-trending zone (Sciarra et al., 2019). A further contribution to the understanding of the surface system is provided by hydrogeological studies (e.g., Giambastiani et al., 2022), who showed how differences in mud levels between different mud volcanoes are mainly due to different gas–liquid ratios in the conduits of alimentation. The data also indicate the presence of surface aquifers at depths between 5 and 30 m, interpreted as temporary reservoirs for rising gas, whose overpressure state can control the system's activity.

More recently, attention has progressively shifted to the use of non-invasive geophysical techniques, in response to the need to characterize potential heterogeneities within the area, that are hard to investigate with only direct and punctual observations. In this context, integrated geophysical approaches, such as resistivity measurements (Accaino et al., 2007; Lupi et al., 2016; Oppo, 2011), gravimetric campaigns (e.g. Nespoli et al., 2023) and seismic monitoring studies (Giovani et al., 2017; Antunes et al., 2022; Brindisi et al., 2024), have progressively contributed to a plausible reconstruction of the subsurface geometry in the first tens of meters, helping to delineate a multi-level structured feed system at depth.

One of the first attempts to characterize the subsurface of the Nirano site through geophysical investigations was conducted by Accaino et al. (2007), which carried out an integrated approach based on seismic refraction and 2D geoelectric tomography. The study, focusing on one of the main cones in the southwestern portion of the field, highlighted the presence of sub-vertical structures in the first subsurface (first tens of meters deep), interpreted as fluid conduits, as well as low-velocity and low-electrical resistivity anomalies compatible with saturated materials and possible surface reservoirs. The study by Accaino et al. (2007) also highlighted how the structural and lithological complexity of the site makes it necessary to integrate different geophysical techniques, in order to reduce the interpretative ambiguity associated with each method considered individually.

Subsequent studies have also investigated the relationship between subsurface structure and outgassing manifestations. In particular, Lupi et al. (2016) combined geoelectric investigations with continuous seismic monitoring, highlighting the presence of surface conductive anomalies often associated with major vents and interpreted as fluid-saturated volumes or shallow reservoirs. Seismic recordings show background noise at low frequencies, sometimes interrupted by high-frequency pulse signals of the drumbeat type, observed for the first

time in a context of sedimentary volcanism. The same study also documented significant variations in the seismic signal occurring as response of a distant earthquake, suggesting a possible sensitivity of the system to external seismic stresses. The complexity of structure and elastic behaviour of the subsurface was further confirmed by Giovani et al. (2017), who showed that the local seismic response is far from being attributable to simple layered models, and it is indeed influenced by pronounced attenuation effect right below the mud field.

In recent years, the use of passive seismic (e.g., Brindisi et al., 2024; Carfagna et al., 2024) assumed an increasingly central role in the study of the site, trying both to investigate the stratigraphy beneath the subsurface and with the aim of studying the characteristics and location of the main seismic signals generated by near-surface volcanic activity. In particular, Antunes et al. (2022) identified and characterized different types of high-frequency pulse signals defined as drumbeat-like signals (see Iverson et al. 2006), located in the most active portion of the duct system. The sources of these signals were also located in correspondence with structural anomalies already known from previous studies (e.g., Lupi et al., 2016). From the study by Antunes et al. (2022) the recorded signals are essentially dominated by P-waves and show analogies with signals observed in common hot volcanic settings, suggesting an origin related to the migration of mud and gas within the conduits and a close relationship between seismic activity and surface morphological changes.

Brindisi et al. (2023) instead first distinguished both regular and irregular impulsive signals characterized by high frequency contents and dominated by shallow S-waves, interpreting them as the result of the mechanical interaction between solid ducts and a flow composed of mud and gas.

In the wake of these results, Carfagna et al. (2024) attempt to localize these seismic patterns at very shallow depths (generally <20 m) through an approximate and fast location procedure in the assumption of a homogeneous medium (Pujol & Smalley, 1990). The location procedure aimed to localized thousands of sources beneath the site (Fig. 6.2 readapted from Carfagna et al., 2024), especially below the mud vents, interpreting them as the expression of stick–slip mechanisms related to the interaction between gas, mud and the walls of the emission duct, and proposing a gas flow estimate consistent with previous geochemical measurements (Sciarra et al., 2017; 2019). Furthermore, they also showed that seismic activity, probably linked to outgassing processes, is nowadays mainly affecting the eastern portion of the site and highlighted two potential conduits directly beneath the volcanic cones further east (EW and NS sections in Fig. 6.2). Polarization analysis also revealed a complex particle motion caused by the detected signals, further suggesting that significant part of the recorded signals included the contribution of shear-wave, accordingly also to the estimated propagation velocities.

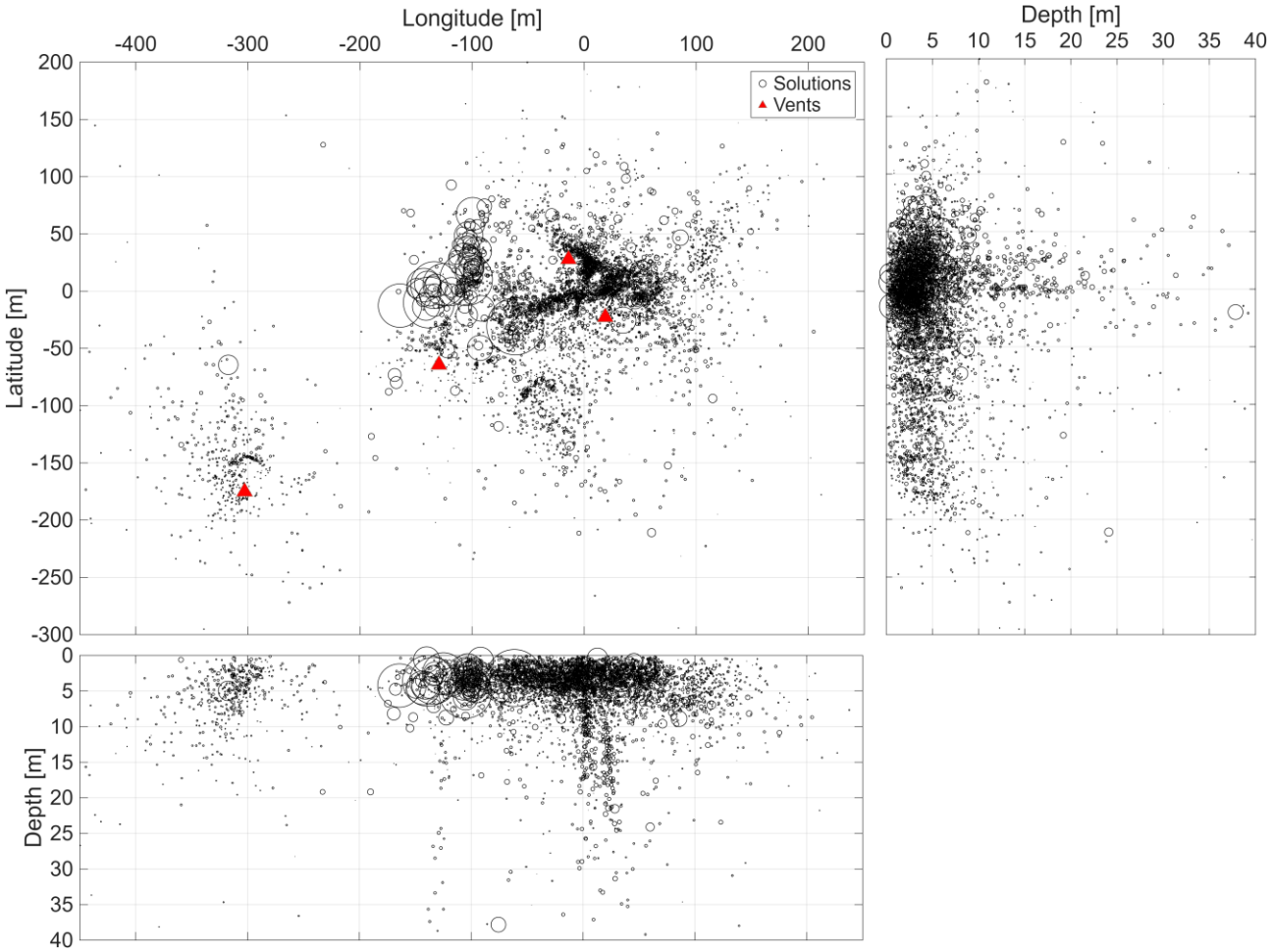

Figure 6.2 - Spatial distribution of seismic sources localized through passive-seismic acquisition readapted from Carfagna et al., (2024). Localized solutions (black dots) are shown in a x-y plane (longitude vs latitude) and in two vertical sections (longitude-depth and latitude-depth). The red triangles indicate the position of the main vents. The distribution of solutions highlights a prevailing concentration in the eastern area of the mud field and a laterally extended geometry, incompatible with a simple source model. Vertical sections show that the solutions are mainly located in the first tens of meters of depth, highlighting possible deeper sources extensions (conduits) beneath the vents.

Active seismic profiles were also conducted (Antunes et al., 2022; Brindisi et al., 2023), revealing low near-surface shear-wave velocities (Vs ≈ 150 m/s), which increase to about 300–400 m/s at depths of 40–50 m. A similar depth-dependent trend was inferred for P-wave velocities, with low Vp values in the uppermost layers progressively increasing within the first few tens of meters beneath the site.

In particular, Antunes et al. (2022), showed a not fully uniform distribution of Vs in the first subsurface, with the possible presence of velocity reversals superimposed on stiffer layers. Although such models are mainly used as interpretative support and are not discussed in terms of detailed elastic stratigraphy, they suggest a possible both vertical and horizontal heterogeneity of the surface subsurface.

Ambient vibration analyses showed no significant peaks in horizontal-to-vertical spectral ratios (HVSR), suggesting the absence of strong shallow 1D resonance effects, but also presenting peculiar low-frequency minima (or maxima in VHSR curves) possibly associated with mud reservoir.

Overall, the set of diverse investigations conducted at the Salse di Nirano area delineates a potential heterogeneous system, characterized by a complex network of conduits, shallow and deeper reservoirs, and preferential migration paths. This framework highlights the importance of integrated approaches and provides the scientific context within which this thesis work fits, with the aim of identify lateral heterogeneities and shallow structures associated with fluid and degassing process through passive seismic measurements.

6.2 Data acquisition

The seismic data acquisition was conducted in October 2025 at the Salse di Nirano site and consisted of three different seismic array configurations (see Fig. 6.3). For all investigations, vertical geophones with a natural frequency of 4.5 Hz were used, connected to a BrainSpy™ digital acquisition system, produced by Moho srl, which ensured high quality and stability of the records throughout the entire duration of the measurements.
The first two array configurations (N1 and N2 in Fig. 6.3) were specifically designed to allow for a subsequent 2D tomographic inversion of phase velocities. In both configurations, the seismic array consisted of two parallel branches, each composed of eight geophones, for a total of sixteen sensors. Within each branch, inter-distance of 5 m between the geophones were chosen, while the distance between the two branches was varied as a function of the investigation area.
The N1 array was placed in the central-southern portion of the mud volcano field, in an area characterized by denser vegetation. In this configuration, the distance between the two parallel branches of the array was 50 m, delimiting an investigated area of approximately 50 m × 35 m. The second array, N2, was instead performed in the eastern part of the site, in proximity of one of the main emission cones and therefore the most recently active zone. For N2 the distance between the two parallel branches was reduced to 40 m, corresponding to a survey area of approximately 40 m × 35 m.
The geometry of these N1 and N2 measurements was chosen to optimize the subsequent tomographic processing phase, allowing us to delimit a spatial domain with good trajectory coverage between the employed geophones. Although the dimensions of these domains are relatively small, they are adequate in the context of the Nirano site, where lateral variations in the elastic properties of the subsurface are expected even on very small spatial scale.
The third measurement, called N3 (see Fig. 6.3), was performed for different purpose than the other two arrays. In this case, the objective was to obtain extended pseudo-section of the investigated area, to be used as a qualitative reference and to support the interpretation of the results from N1 and N2 smaller-scale investigations. To this end, a ‘L-shaped’ seismic array consisting of two approximately perpendicular branches was deployed, oriented towards south and east, joined at the north-western corner. Both branches were approximately 100 m long with a constant inter-distance of 20 m between all geophones.

Only eleven geophones were used for N3 measurement. The number of employed geophones and the distance between them inevitably implies a lower spatial resolution than the N1 and N2 configurations; however, it allows to investigate a larger area and provide an overall observation of the subsurface structure. The results obtained from this configuration are therefore intended as complementary to those derived from smaller-scale arrays, which provide a greater level of detail.

During the data acquisition phase of the N1 configuration, geophone G12 presented operating problems that compromised the quality of the recorded data, as a result, this sensor was excluded from subsequent processing steps and was not used for the extraction of dispersion curves.

The overall recording duration was approximately 2 hours for the N1 and N2 configurations, and approximately 1 hour for the N3, ensuring in all cases a sufficient amount of data for subsequent analyses in the time and frequency domain. Overall, the acquisition configurations adopted allowed to investigate the subsurface of the Salse di Nirano by combining an extended view of the area with a higher-resolution analysis of specific portions of the field. Starting from this dataset, the next Sections illustrates the main data processing steps, describing the procedures adopted for seismic noise analysis and for extracting information useful for characterizing the elastic properties of the subsurface.

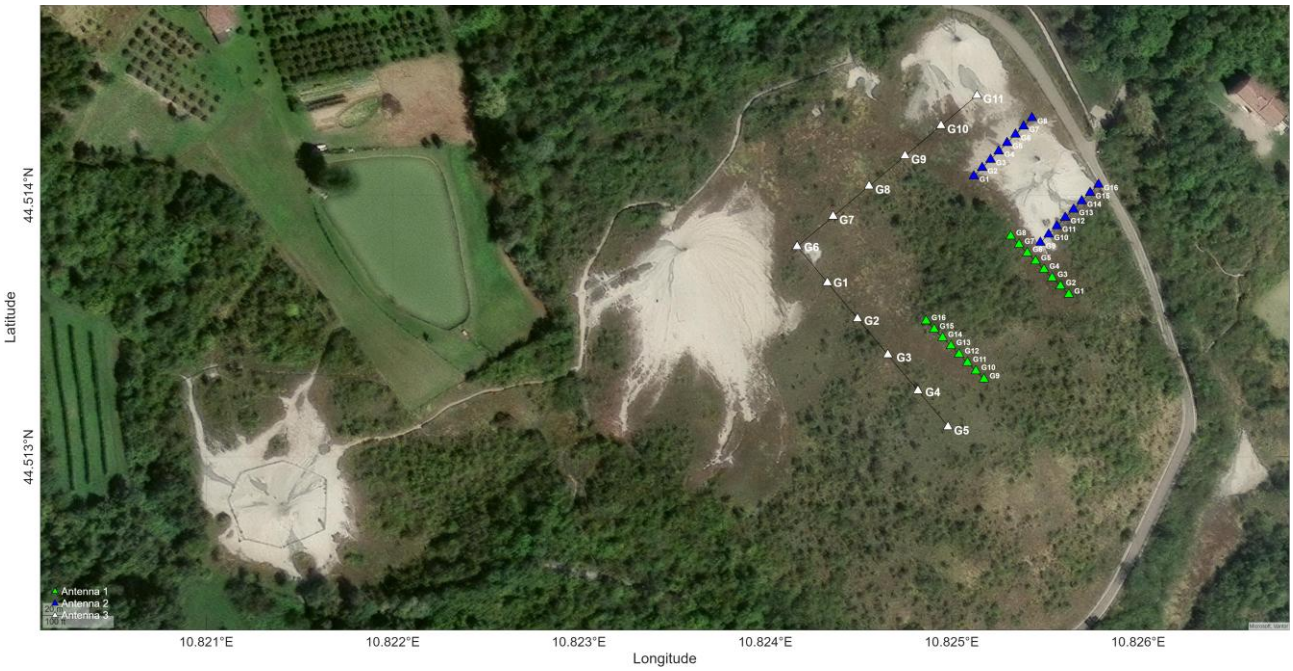

Figure 6.3 - Satellite view of the Nirano mud volcano field showing the layout of the three seismic arrays employed during the experiments. Arrays N1 and N2 consist of two parallel linear branches, while array N3 has an L-shaped geometry. Coloured markers indicate the position of individual geophones; 16 for array N1 and N2 and 11 for N3.

6.3 Data processing

Section 6.3 focuses on the general workflow through which data acquired at the Salse di Nirano reserve were processed. It is divided into three main sub-sections. First part is dedicated to the extraction of Rayleigh dispersion curves through ambient seismic noise records considering all pairs of geophones within each array (N1, N2 and N3). Dispersion curves extraction represents the first step of the entire process, and it is uniformly applied to all the seismic arrays, with minor adaptations depending on the acquisition geometries. In this step, particular attention was given to the main differences between the three arrays in terms of phase velocity distribution across the frequency range, exploring consistency of the picked curves and the role of group velocity information in the choice of the right curve through automatic picking.

The second part concerns the tomographic inversion of phase velocities. Considering arrays N1 and N2, specifically designed for this purpose, inversion was applied to obtain 2D Rayleigh phase velocity maps representative of the lateral variations of the subsurface elastic properties within the investigated domains. For array N3, characterized by an a 'L-shaped' geometry and a larger spatial extension, the same principle was adopted with the aim of obtaining pseudo-sections of Rayleigh phase velocity, useful to provide a more general view of the subsurface structure and to support the interpretation of the results at smaller scales.

The third part of this Chapter is devoted to the inversion of local dispersion curves obtained following tomographic inversion, that provided local shear wave velocity (Vs) profiles. For arrays N1 and N2, this procedure allows us to construct 3D models in terms of Vs, while considering N3, it allows to obtain 2D Vs sections consistent with the array geometry. In each subsection, all configurations are then considered together, evidencing similarities and differences in the acquired data and obtained results.

Finally, considering the dispersion curves extraction procedure, it was not possible to extract dispersion curves of adequate quality for all pairs of geophones. Consequently, in the subsequent tomographic inversion phases, only the curves considered adequately reliable were selected. Furthermore, during the inversion phase, different combinations of parameters were tested, particularly in terms of damping factors and spatial and frequency correlation, in order to evaluate their influence on the stability and quality of the final results.

6.3.1 Estimation of Rayleigh wave dispersion curves

For all three deployed seismic arrays (N1, N2 and N3), data processing began with the extraction of phase and group dispersion curves between all possible pairs of stations. The procedure to retrieve dispersion curves was initially conducted automatically (see Chapter V) in order to identify, for each pair, the most reliable dispersion curves within the frequency–velocity plane.

Although automatic picking allowed a large number of pairs to be quickly processed, the results were not accepted uncritically. For each array, a systematic verification of the extracted curves was performed, with the aim of evaluating their physical consistency and continuity in the frequency range of interest. This control allowed to identify clearly unrealistic curves and picking algorithm failure.

Manual intervention generally concerned the selection of alternative time windows and the adoption of different amplitude thresholds, with the aim of rejecting highly noisy portions of the records with no significantly coherent signals. In other cases, it was necessary to adapt the frequency window to the shape of the experimental spatial correlation function to improve frequency resolution across the curve. Manual intervention was also needed where the automatic picking selected incorrect portions of the cross-correlation function for group velocity estimation, compromising group curve comparison and providing incoherent phase curves according to the context. Overall, 68 dispersion phase dispersion curves were extracted for array N1, 92 for array N2, and 24 for array N3. In Fig. 6.4, all selected phase dispersion curves for arrays N1 and N2 are showed.

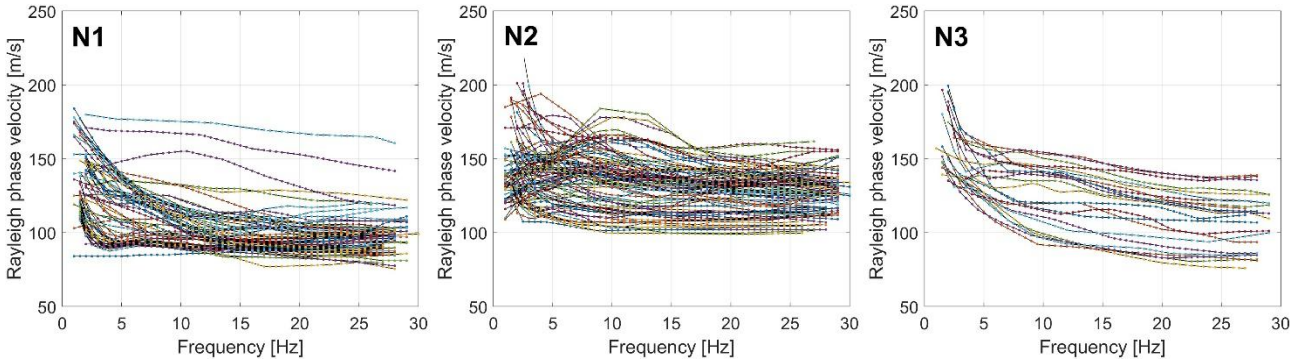

Figure 6.4 - Rayleigh wave dispersion curves extracted for the seismic arrays N1, N2 and N3. The curves show the phase velocity as a function of frequency for all pairs of stations considered, highlighting significant variability across a wide frequency range. Considering N1, N2 and N3, respectively 68, 92 and 24 curves composed the dataset.

The curve distributions show significant differences between the three arrays. Array N2 is characterized by moderately higher phase velocities, especially when compared to velocities captured by array N1.

In contrast, phase velocities from the array N3 cover a wider range, which includes both the lowest velocities observed within the array N1 and the highest values estimated in the study area. Although the curves associated with N3 suffer from lower resolution and therefore exhibit a more continuous and smoother trend, this result can be considered encouraging, as it indicates that the N3 array is able to capture the same overall variability of the subsoil properties. Also array N1 exhibits increased variability, indicative of lateral heterogeneity within a relatively small sampled area.

Overall, although the observed phase velocities are within a relatively small range, approximately always between 80 and 200 m/s, dispersion curves appear to highlight the presence of lateral heterogeneities, also at higher frequencies.

In both array N1 and array N2, recurring trends can be observed in the distribution of dispersion curves, which tend to cluster showing overall similar across the frequency range investigated. In particular, in the array N1 there is a large set of curves showing a slight minimum in the frequency range 3-4 Hz, and that are characterized by phase velocities below 100m/s, which are not observed at all in the closer array N2.

In the N2 array, instead, a large group of curves show a marked velocity decrease at low frequencies, suggesting the possible presence of low velocity structures beneath some portions of the investigated area.

In both arrays (N1 and N2), the presence of some dispersion curves deviating from the main group of extracted curves is also observed. This behaviour is particularly evident in array N1, where 2/3 curves are characterized by higher phase velocities than the rest of the curves. Although these curves appear as outliers compared to the general phase velocity distribution, they are associated with pairs of spatially close stations and show a good coherence each other, suggesting that they may truly reflect local conditions rather than processing artifacts.

Contrary to expectations, an important informative contribution was also provided by the curves extracted between pairs of stations with minimum inter-distance (5 m), i.e. along the same branch of the N1 and N2 arrays. Although the uncertainty associated with these curves was generally greater, they were still consistent with each other and with the local context reproduced by the pairs at greater distances. Considering the observed phase velocities, a surface wave propagating with phase velocity 150m/s takes approximately 0.03s to travel two geophones 5m apart. This value is around ten times the sampling time (0.0039s), indicating therefore that even two receivers at 5m of distance can be able to detect such wave.

Spatial consistency between dispersion curves is indeed a key element in the validation of the final dataset. A significant example is shown in Fig. 6.5, where the phase curves obtained from different pairs of geophones of array N2, involving in a reference geophone (G2) coupled with multiple geophones of the opposite branch, are reported. In particular, the pairs considered include combinations from 2–10, 2–11, up to 2–15, which correspond to increasing distances in that order (see Fig. 6.3). The set of curves highlights a systematic decrease of phase velocities from the closest receivers (2–10) to the progressively more distant ones (2–15), a coherent behaviour over a wide frequency range, with lower frequencies that span more velocities. The fact that this trend is consistently reproduced by all considered pairs suggests that the extracted dispersion curves are consistent with each other and therefore reflect a gradual change in the elastic properties of the subsurface.

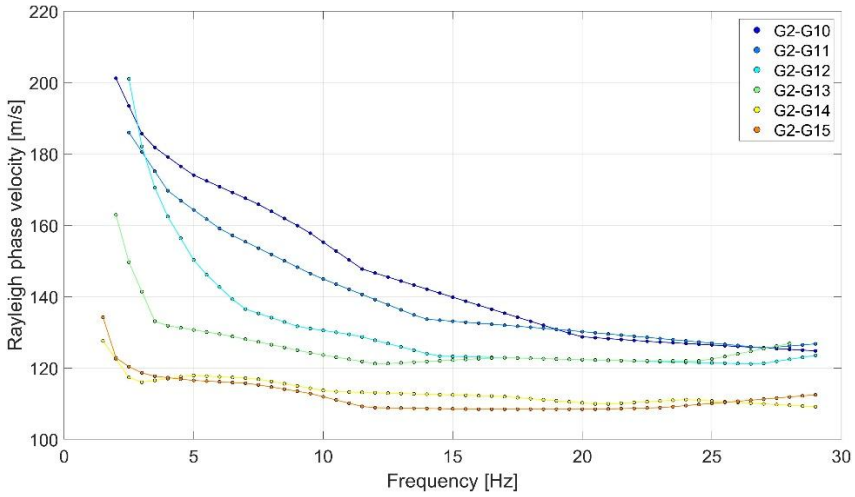

Figure 6.5 - Examples of Rayleigh wave dispersion curves extracted for different pairs of stations within array N2. Despite the differences between the individual curves, the overall trend is consistent and characterized by a gradual variation in phase velocity in a wide frequency band, even at higher frequencies.

Group velocity curves automatically extracted also played a fundamental role in assessing the reliability of the individual result in terms of phase dispersion curve. Group curves that deviate significantly from the general context might indicate critical issues in the picking process, making it necessary to evaluate the quality of group velocity estimates a posteriori. In this context, although also group velocity curves show appreciable dispersion in terms of velocity, covering an overall velocity range similar to the phase velocity curves within the same array, the lower detail associated with their propagation becomes evident focusing on curves extracted for arrays N1 and N2. Phase curves show a clear difference in terms of phase velocities (see Fig. 6.4); however, this distinction does not emerge as clearly from the analysis of the group curves alone. Group curves associated with the two arrays largely sample the same velocity range and don't allow for a preliminary discrimination of two different contexts. This behaviour is consistent with the lower sensitivity of group velocities to small-scale lateral variations compared to phase velocities. In this work, group curves are not the focus of the study, as they are primary used as a validation tool in the phase curve selection. Nevertheless, sufficiently reliable estimates of the group velocities suggested that the same approach adopted in subsequent tomographic inversion of phase curves could also be extended to group curves, that, although less sensitive than phase curves, are able to capture a certain lateral heterogeneity in the elastic properties of the subsurface.
In Fig. 6.6, for some analysis, phase velocity curves selected through automatic picking, accompanied by relative uncertainty bands, are shown, together with the relative estimated group curves.

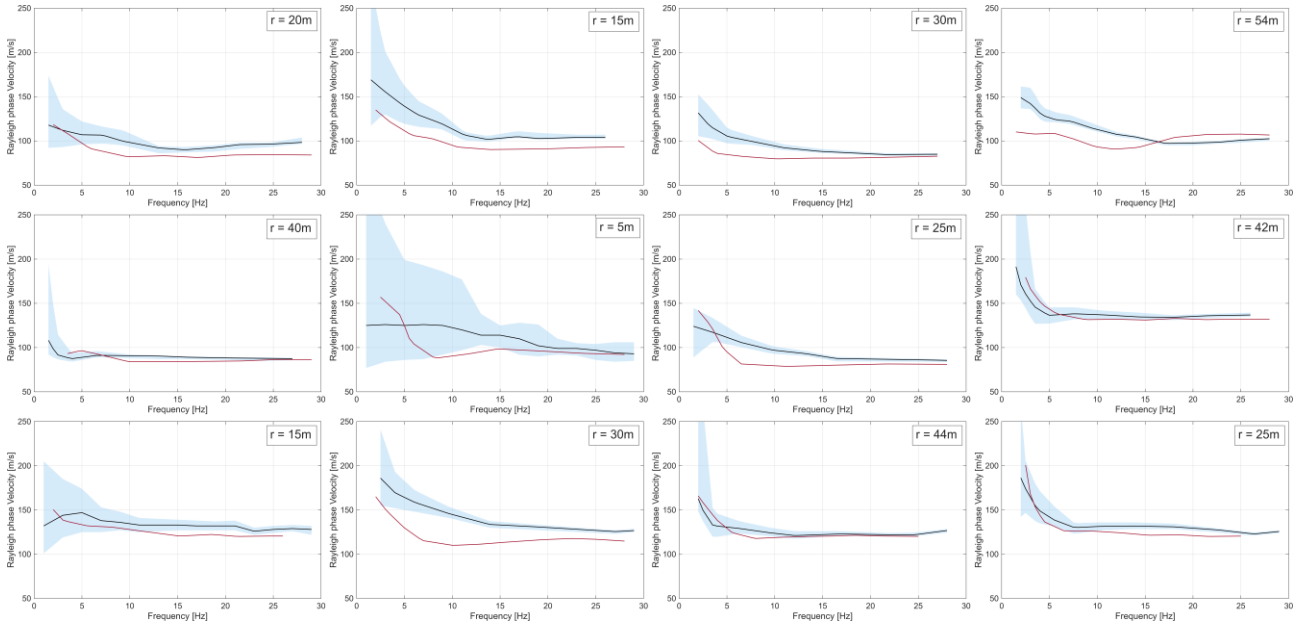

Figure 6.6 - Comparison of phase and group dispersion curves for a representative set of two-station analysis. Each panel shows the phase velocity curve selected by automatic picking (black line), its uncertainty band (blue area) and the estimated group velocity curve (red line) also provided by automatic picking. The interstation distance value is indicated at the top right in each panel.

Fig 6.6 shows the phase dispersion curves (black curves), represented within the relative uncertainty boundaries (blue area), and the group velocity curves (red curves). The set of curves highlights a clear relationship between inter-station distance and the amplitude of uncertainty associated with phase velocity estimates, particularly marked at low frequencies. This behaviour is largely expected, since for short distances between stations it becomes more challenge to accurately detect wave arrival, resulting in increased uncertainty in the estimates. Considering pairs of curves in Fig. 6.6, it may be also observed that experimental group curves sometimes lie entirely below the selected phase velocity curve, while in other cases group and phase curves intersect and partially overlap within restricted frequency bands.

In normally dispersive contexts, group curve is generally expected to lie below phase curve (see, for example, Eq. 2.11), while intersections or overlaps between phase and group curves can occur as response of more complex structures or in presence of low-velocity layers at depth, intercalated between denser materials.

However, both experimental curves must be interpreted taking into account the uncertainties that necessarily affect the estimation procedure and, which may produce relative shifts of both the curves within the confidence range.

Overall, the extraction of dispersion curves can be considered satisfactory. Taking into account all three arrays, little more 200 two-station analysis were performed through automatic procedure. Only for approximately 20% of cases further manual intervention was necessary in order to modify the selected curve or to improve the quality of the estimate. As already mentioned, even a manual intervention for a limited part of these low-quality data was not sufficient to obtain adequate phase and group dispersion curves. These curves were therefore rejected in order to avoid too nosily estimates that could significantly influence subsequent tomographic inversion.

In addition to the qualitative evaluation of the extracted dispersion curves, the stability of each two-station analysis was also explored considering the behaviour of the windows stacking process. Rather than focusing on individual curves, the reliability of the estimates was evaluated by analyzing the convergence of the two experimental functions as the number of employed windows increases. This check represents a good practice to control if the stacking procedure contributes to reduce possible bias related to the spatial heterogeneity of noise sources and the number of windows sufficient to ensure good convergence towards stable cross-correlation and coherence functions. This approach provided an additional criterion to evaluate the robustness of the obtained dispersion curves.

In Fig. 6.7 cross-correlation and coherency functions associated with three different dispersion analysis are shown. In each panel of Fig. 6.7, red function is obtained stacking 20 minutes of records and black functions are obtained through one hour of stacked records. The comparison highlights how the stability of experimental functions does not depend exclusively on the number of stacked windows, but mainly on their quality.

In particular, in presence of sufficiently coherent noise wavefield, even a few tens of minutes of signal are sufficient to obtain effective azimuthal coverage of the sources and to achieve good convergence of cross-correlation and coherency functions. Furthermore, it's clearly visible that sometimes, the amplitudes of the spatial autocorrelation function are more defined for some frequency ranges in the first 20 minutes of recording

than in longer intervals, such as an hour. This suggests that the inclusion of noisy time windows can degrade the overall quality of stacking process, highlighting the importance of window selection in pre-processing step.

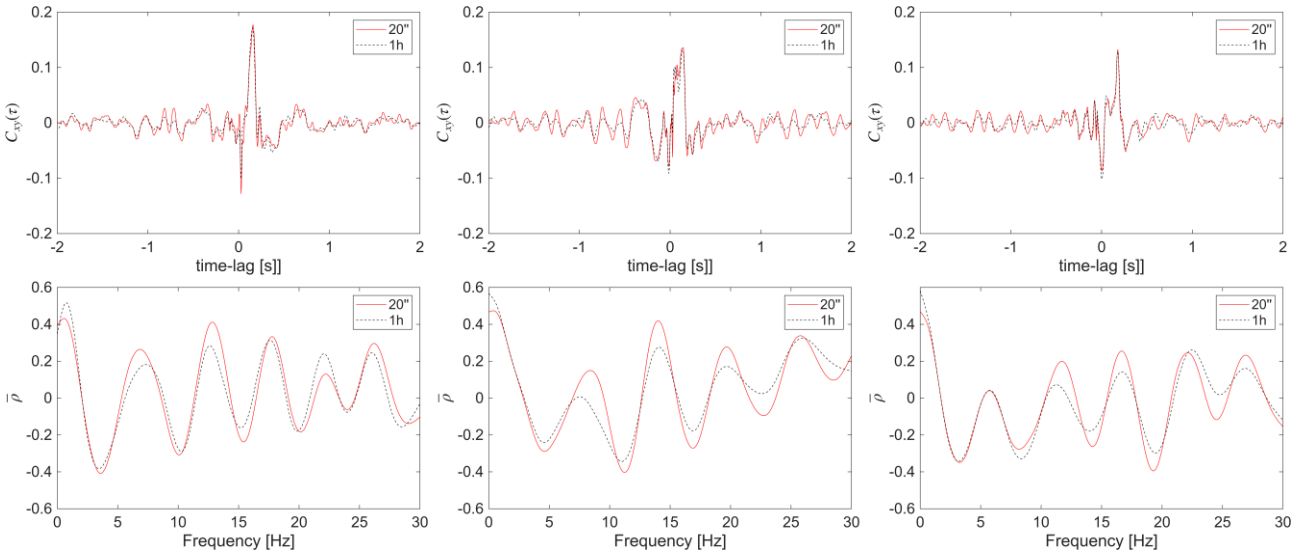

Figure 6.7 - Comparison of the time cross-correlation functions (top panels) and the corresponding frequency coherency functions (bottom panels), computed considering 20 minutes (red lines) and 1 hour (dashed black lines) of stacked time windows, for three pairs of stations. In the presence of coherent signals, even relatively short stacking intervals can provide stable results, comparable to those obtained with longer recordings.

Overall, the obtained dispersion curves highlight consistent spatial variability in the shallow subsurface. However, information coming from individual dispersion curves, even when viewed as a whole, is still limited to the geometry of the paths considered and can only give us preliminary indications of the possible expected variations. For this reason, to more comprehensively sample the distribution of phase velocities within the investigated areas, in the following paragraph the selected dispersion curves are used as input for a linear tomographic inversion in order to image the near-surface structure beneath the site.

6.3.2 Phase velocity maps through SWT

The previously extracted phase dispersion curves for the different pairs of stations were used as input for tomographic inversion of phase velocities. The analysis was conducted separately on the seismic arrays N1, N2 and N3, which were characterized by different geometries and spatial scales, and thus designed to investigate the subsurface with different levels of detail. The three arrays, whose geometry has already been shown in Fig. 6.3, are treated individually here and for each one a local reference system is defined, based on the array geometry. Later, the tomographic results in terms of phase velocity maps will be re-discussed together as a whole to give a meaning consistent within the Nirano context.

As anticipated in previous Sections, arrays N1 and N2 consisted of 2D geometry configurations composed of two parallel branches of geophones, with inter-distances respectively of 50m and 40m between the branches and fixed inter-station distance of 5m within the branches (Fig. 6.8a and 6.8b). Array N3, on the other hand, was made up of two almost orthogonal branches in a 'L-shaped' configuration, each approximately 100 m long, with geophones placed at a fixed distance of 20 m (Fig. 6.8c). For all the arrays, the same inversion procedure described in Section 5.2 was applied; however, given the geometry of N3 and the distribution of the stations, the goal of the inversion for N3 is to obtain phase velocity sections along the two branches of the array.

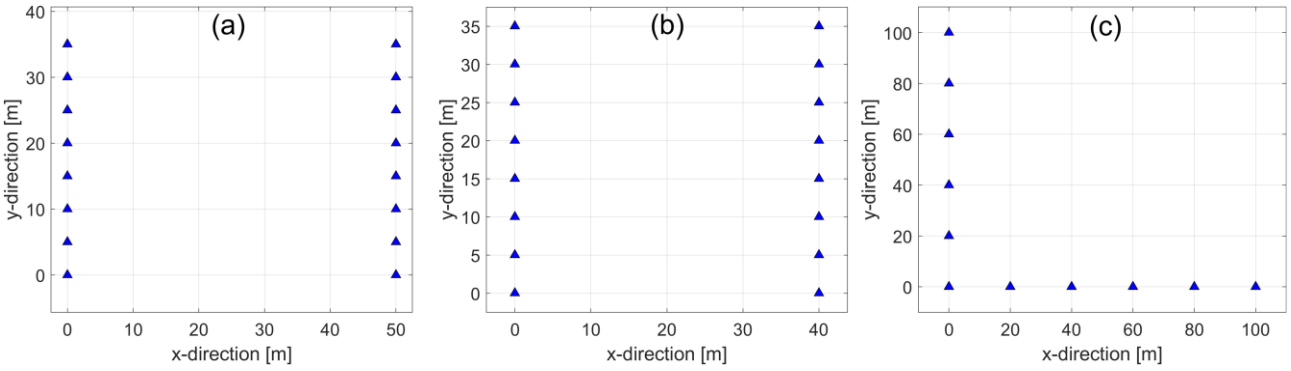

Figure 6.8 – Geophone configurations for arrays N1, N2, and N3. Panels (a), (b) and (c) show the spatial arrangement of the sensors for each array, highlighting the different geometries and spatial scales adopted in the acquisitions. For each array a relative coordinate system is used.

The inversion scheme described in Section 5.2 is implemented in order to jointly invert all sampled frequencies, instead of treating each frequency independently. As discussed above, this choice involves in the introduction of a correlation term along the frequency range of interest. In this context and given the regularization terms before introduced controlling the solution stability and data adherence, it is necessary to carefully validate the inversion parameters to identify the best set for tomographic inversion.

For each frequency, the prior model m_0 was defined using arrival times of the phase velocities and corresponding inter-station distance; a simple linear regression was then applied to estimate the best fitting velocity with respect to the observed data, subsequently assuming m0 as homogeneous for each parameter.

As shown above, inversion based only on the fit between observed and predicted data is sometimes discouraged, as it can lead to unstable solutions that may lose physical realism respect to the real-world context. For this reason, several tests were conducted by varying the main inversion parameters, such as the spatial and frequency correlation length and the model prior variance, in order to evaluate the most suitable combination for each array. At the end of this Chapter some practical examples of the effect of the regularization parameters on the results in terms of phase velocity solutions and misfit, will be presented in the same form of the test presented in previous Chapter (see Fig. 5.13 and 5.14).

Furthermore, the model discretization was also analysed, evaluating the best cell-size for an optimal compromise in terms of stability and smoothing level of the final model. The joint variation of the inversion

parameters and the cell-size allowed to verify a set of potential solutions characterized by different trade-off between data adherence and regularity of the final solutions. In all three arrays, however, it was observed that the obtained tomographic solutions did not differ substantially from each other, as the considered parameters varied. This result suggested that the reconstruction of phase velocity maps was mainly controlled by the information contained in the data, rather than by the regularization parameters.

For arrays N1 and N2, the inversion results are presented in the form of 2D Rayleigh phase velocity maps that provide a spatially approximated continuous representation of lateral phase velocity variability within the areas bounded by the two branches of each array. The inversion results for array N3 are instead reported as offset-frequency sections (or offset-wavelength section), highlighting Rayleigh phase velocity variations along the 100m long branches of the array.

Based on the considerations above discussed, Table 6.1 shows the inversion parameters adopted to obtain the tomographic images for the three arrays. It is important to highlight that the differences in optimal parametrization mainly reflect obvious differences in the scale of investigation and the different geometry of the arrays.

Array		σ_m	l_c	l_f	**cell size**
N1		0.008sm^{-1}	7.5m	2Hz	2.5m
N2		0.006sm$_{-1}$	7.5m	2Hz	2.5m
N3	**A**	0.01sm^{-1}	10m	2Hz	5m
	B	0.01sm^{-1}	10m	2Hz	5m

Table 6.1 - Best parameters adopted for tomographic inversion of arrays N1, N2 and N3. The value of the prior model variance (σ_m), the spatial correlation length (l_c), the frequency correlation length (l_f) and the cell size used in the model discretization are reported.

Based on the inversion parameters reported in Table 6.1, some representative phase velocity maps and cross-sections are presented below. In particular, phase velocity maps obtained for the arrays N1 and N2 are shown in Fig. 6.9 at three selected frequencies, while two offset-frequency sections obtained by array N3 are illustrated in Fig. 6.10. For each phase velocity map in Fig. 6.9, the associated rms-value is reported, computed considering only phase velocities relative to that specific frequency. In fact, the incorporation of the whole frequency range into a single one-step inversion has direct implications to the obtained global misfit, which necessarily considers all the analysed frequencies. In this context, a more detailed description of the inversion results can be investigated by isolating the contributions of individual frequencies considered. It was thus possible to evaluate the effectiveness of the frequency-by-frequency solutions.

The misfit values associated with the individual frequencies, expressed in terms of root-mean-square residual between observed and predicted data, systematically fell within a range between 0.0027s and 0.0085s for all the analysed arrays.

Generally, higher rms-values are observed at lower frequencies (see Fig. 6.9), an expected behaviour consistent with the fact that data associated with these frequencies sample deeper portions of the subsoil, thus being less constraining and showing greater uncertainty than data from the first few meters beneath the surface.

Fig. A2 in Appendix A shows the coverage of the model obtained from the trajectories used for the tomographic inversion of arrays N1 and N2. The good number of dispersion curves extracted in the previous step, combined with a generally continuous sampling along the investigated frequency range, made it possible to obtain an almost uniform number of observed phase velocities in the range between 3 and 28 Hz. Only at frequencies below 3 Hz and above 28 Hz a reduction in the number of data is observed.

This distribution of phase velocity data made it possible also to constrain, at least with one observation, almost all the model parameters within the domains investigated by the two arrays, ensuring overall adequate spatial coverage for the application of tomographic inversion.

Considering arrays N1 and N2, the obtained phase velocity maps show clear lateral variability within the explored domain, with velocity patterns that persist over a wide frequency range, while their spatial extent and contrast amplitude can vary with frequency, reflecting different depth sensitivities of surface waves.

The estimated values for both arrays fall within a range of approximately 70–80 m/s to 160–180 m/s. These latter values are consistent with those already reported in previous passive seismic studies conducted in the area, while the lower velocities (below 100 m/s) in the current context, could reveal low-velocity structures linked to the volcanic activity and the subsurface outgassing manifestations, especially referred to the array N2. Indeed, array N2, located precisely above one of the main eruptive vents in the north-eastern part of the volcanic field, provided phase velocity maps systematically showing the presence of a low-velocity structure approximately at the centre of the investigated area, near the eruptive vent. This low-velocity anomaly is durable over different considered frequencies and takes a spatial extension that change gradually, tending increase towards high frequencies, i.e. at lower depths. At the same time, the low-velocity structure appears to be surrounded by higher-velocity edges, suggesting a well-defined lateral contrast between the active system of the vent and the surrounding material. As frequency gradually increases, a further low-velocity anomaly progressively appears along the lower edge of the domain surrounded by the sensors, which represents a contact edge between arrays N1 and N2. In general, as one move to higher frequencies, the phase velocity maps become progressively more articulated, as expected. At higher frequencies, in fact, more marked velocity contrasts emerge, indicating that toward the surface the overall context becomes increasingly less regular.

The N1 array is instead located within an area characterised by thick vegetation and with no evident eruptive or surface degassing manifestations, thus presenting a more challenging scenario for interpretation. The estimated maps show phase velocities compatible to those obtained for the array N2, which is expected given the reduced spatial distance between the two configurations.

In the case of N1, the maps also show the presence of low-velocity areas that tend to become predominant at higher frequencies. In particular, the south-eastern part (in the local reference system of the array) is characterised by the highest velocities; this zone shows a partial spatial overlap with the high velocities

observed in the N2 array, indicating a possible lateral continuity of the more rigid structures between the two investigated areas.

Compared to results for array N2, phase velocity maps for array N1 show clear marked contrast between the southeastern portion of the area, characterized by higher velocities that decrease towards the surface, and the northwestern portion, associated with lower velocity values that become predominant in almost the whole area at high frequencies. As shown by the layout of the geophones in Fig. 6.3, the north-eastern corner of array N1 directly borders the south-eastern corner of array N2. In this spatial overlap zone, and considering the same frequency, the obtained tomographic maps are in partially agreement, indicating that the low- and high-velocity structures are consistently reconstructed regardless of the acquisition geometry, especially at lower frequencies.

However, it should be emphasized that the interpretability of the tomographic results differs between the two arrays. Considering the array N2, located precisely above the volcanic vent, this surface evidence makes the phase velocity maps more immediately interpretable at this preliminary step. The tomographic patterns obtained are indeed consistent with the surface geomorphological features. In contrast, a direct interpretation of the phase velocity maps for array N1 is more challenging. The absence of evident surface manifestations in the investigated area makes it less immediate to attribute a physical meaning to the lateral variations in velocity observed, at least at this stage of the workflow. For this reason, results from array N1 are mainly discussed in relation to those of array N2, which provides phase velocity maps more constrained by the morphological context. Overall, the results obtained for the two arrays suggested that, despite their modest spatial extent, are able to capture significant lateral heterogeneities. The adopted acquisition geometry and inversion implementation allow velocity variations to be resolved even over small spatial domains of the order of a few tens of meters, while maintaining consistently low misfit across the analysed frequency range.

In Appendix A, Fig. A3 and A4 report the whole set of Rayleigh-wave phase velocity maps obtained from the inversion of phase travel-times from arrays N1 and N2, in the frequency range considered. The maps are presented starting from 2 Hz and discretized with a 1 Hz step.

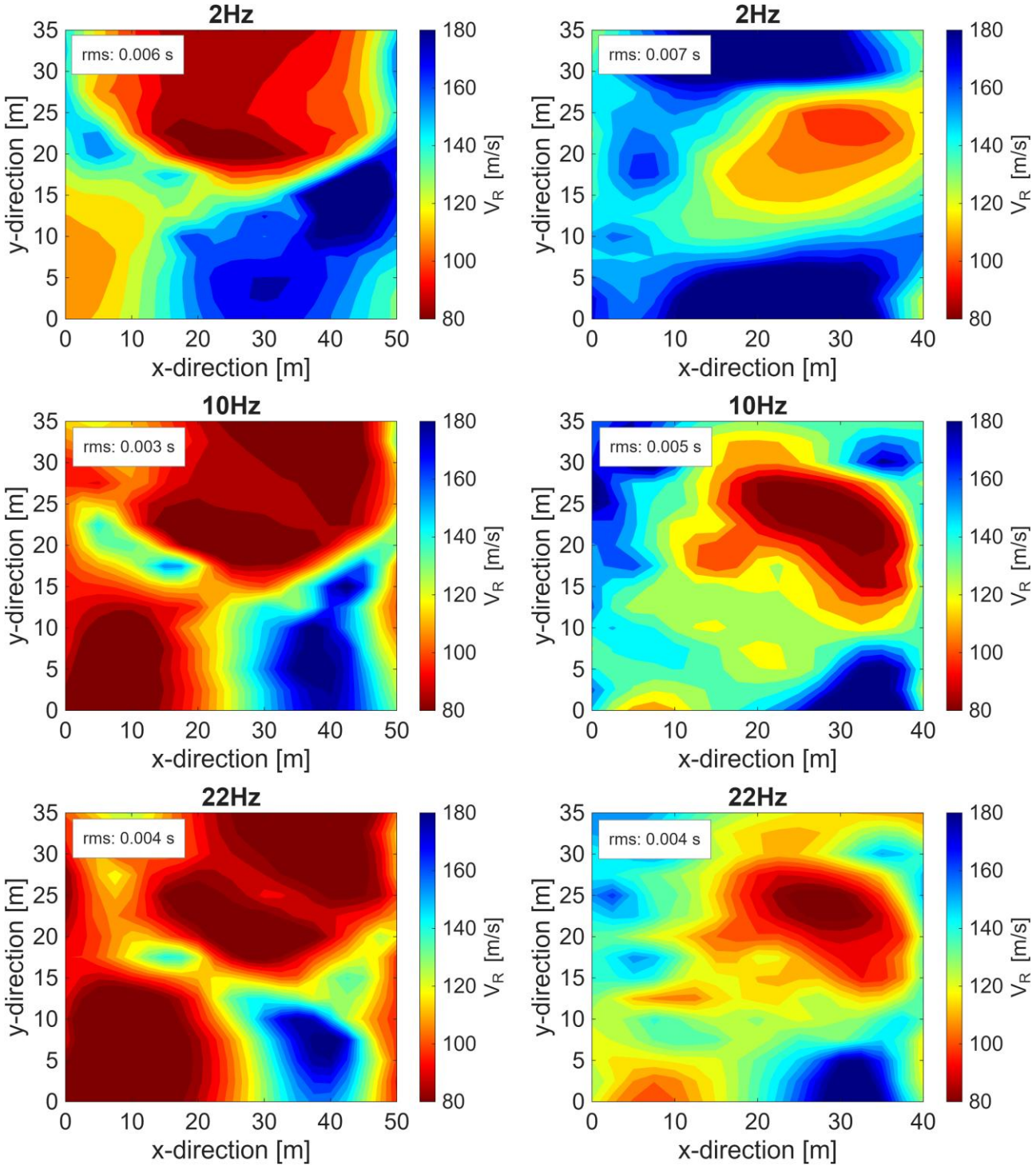

Figure 6.9 - Comparison of the phase velocity maps obtained for arrays N1 (left column) and N2 (right column) at three frequencies representative of the investigated range (2, 10 and 22Hz). The maps highlight the spatial evolution of phase velocities as function of the considered frequency and show patterns that are partly consistent between the two arrays.

Turning now to the array N3, the results of the tomographic inversion are presented in the form of offset-frequency pseudo-sections reconstructed along the two branches of the array (Fig. 6.10). In order to provide a more interpretable result at this step of the work, the offset-frequency sections are transformed into offset-wavelength (Fig. A5 in Appendix A) sections through a frequency-wavelength conversion and subsequent interpolation along the wavelength values.

Although, as already mentioned, there is no guarantee of a linear relationship between wavelength and penetration depth, the sections in Fig. A5 certainly allow to better analyse the tomographic results in terms of structures and lateral variations.

In both type of sections, branch A is the seismic profile oriented approximately from north to south, while branch B runs close to the west-east direction and the zero offset in both sections represents the junction point of the branches.

For array N3 as well, although the investigated distance and inter-stations distances are overall larger than those considered in arrays N1 and N2, this does not appear to have a significant impact on the spatial resolution of the final solutions. The residuals associated with inversion are in fact generally within one or two times the sampling time (0.0039 s), indicating a good level of adherence to the data.

It is worth noting, however, that the behaviour of the residuals differs slightly considering the two branches of the array. The branch oriented approximately in an east–west direction (branch B in Fig. 6.10), with origin (offset = 0) in the left side of the section, shows higher residuals at low frequencies, in an analogous way to what was observed for inversion of N1 and N2. In contrast, the quasi-north–south oriented branch (branch A in Fig. 6.10), with zero offset also in the left side of the section, exhibits excellent fit at low frequencies, while at higher frequencies, corresponding to shallower portions of the model, the inversion is less constrained, with residuals that reach values close to 0.01s in the frequency range above 24 Hz.

Focusing more on the phase velocity sections constructed as a function of offset and wavelength, it is evident how the larger extension of the seismic profiles translates into the observation of different portions of the subsoil. This aspect emerges clearly looking at the two branches of the N3 array, which appear to represent, at least in part, different contexts.

In particular, the branch B, approximately along the east–west direction, reveals a marked increase in phase velocities towards the east, which reach values of 180 m/s already at the smallest wavelengths, therefore close to the surface. In contrast, across the branch A comparable velocities are observed only at longer wavelengths. Looking in more detail at branch A, one can also note how, moving south (i.e. to the right in the section), surface velocities tend to decrease significantly. This behaviour is consistent with what was observed in the area investigated by the N1 array.

Overall, although the reconstructed sections for array N3 are characterized by lower resolution than the 2D maps obtained for arrays N1 and N2, they appear to confirm the main results emerging from previous analyses performed on 2D arrays. Furthermore, beyond the observed differences in the distribution of residuals with respect to the frequencies, the two branches of the array show good continuity. A significant example is represented by the areas close to the zero-offset point in both sections, which correspond to the contact zone between the branches. In there, a slightly reduction in velocities close to the surface is observed compared to laterally adjacent contexts along each profile.

This agreement between the two inversions is particularly encouraging, as the contact area between the two branches corresponds with one of the main vents of the study area. The lower phase velocities observed in this

zone beneath the surface can therefore be interpreted, in a manner consistent with what has been discussed previously, as related to shallow processes of rising fluid just below the vent.

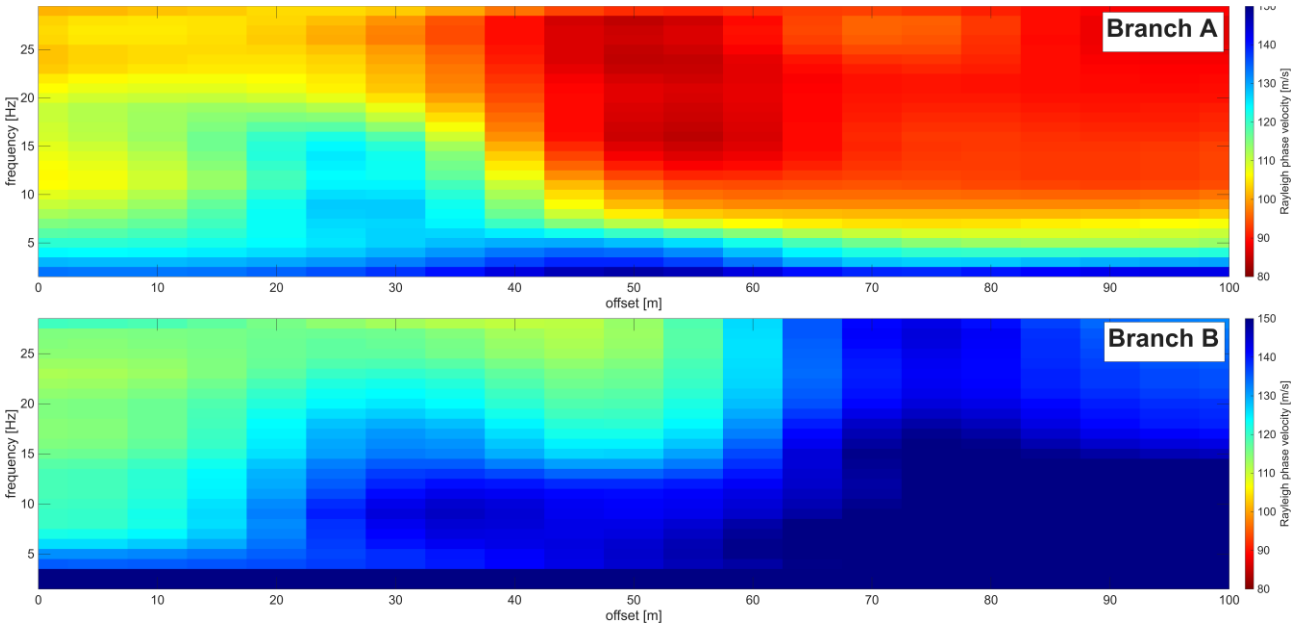

Figure 6.10 - Phase velocity sections as a function of offset and frequency for the two branches of array N3. The upper section refers to branch A, while the lower section refers to branch B. The maps show the variation of phase velocities along each branch, highlighting lateral differences and good continuity of patterns at the contact areas between the two branches.

To further support the interpretation of tomographic results, reliability of the retrieved solutions was evaluated based on major characteristics of the linear inversion scheme. In the context of the adopted tomographic inversion, the inverse problem as formulated assumes Gaussian-type noise on the data. The residuals distributions shown in Fig. 6.11 must therefore be interpreted within this conceptualization. In order to analyse the distribution of errors, Fig. 6.11 shows histograms of the final residuals for all station-station paths considered in the inversion and for the entire frequency range analysed, for each of the arrays studied.

Arrays N1 and N2 exhibit well-centred distributions around zero and are characterized by contained residuals, demonstrating that the data efficiently constrained the estimated phase velocity models. Considering array N3, the residuals are slightly higher overall, particularly looking at the branch A; however, they are still distributed substantially symmetrically around zero, with no evidence of systematic biases. This behaviour proved that, despite the larger interstation distances and different acquisition geometry, the inversion results remain statistically reliable.

The contribution of the observed data to the final solutions can also be evaluated looking at the behaviour of the model parameters within the inversion. In particular, comparing prior and posterior variances of the model parameters allows to quantify the contribution of the observed data in constraining the parameters within an effective range. Posterior variances are extremely influenced by the sampling guaranteed by the paths over the frequencies.

In Fig. 6.12 is shown, by way of example, the effect of the data in the individual variances of the model parameters considering three frequencies of both arrays N1 and N2. In each panel in Fig. 6.12, represents the spatial distribution of the ratio between posterior and prior variances at frequencies 3Hz, 15Hz and 22Hz; where the ratio approaches to 0 the data have well-constrained the parameter, while values slightly close to 1 indicate that data did not significantly affect the estimated parameters, and the latter were more influenced by spatial and frequency correlation or by the reference model used in the inversion.

As expected, the effect of the observed data is more pronounced along the two parallel branches of both arrays where the path coverage is denser, while more limited variations between prior and posterior variance are observed at the edges of the investigated area, especially at the top and bottom parts of the domains, in accordance with the lower sampling imposed by the array geometry in those areas. Array N2 shows the lowest values if compared with array N1, primarily due to the larger number of observed path and the reduced extent of the 2D domain.

Overall, for both arrays the ratio between posterior and prior variance always lies below 0.15 and this result suggested that, given the rms final values, the under-sampled model parameters within the 2D domain are pretty-well controlled by the regularization strategy adopted.

To complete the quality description of the tomographic inversion, the obtained posterior covariances were sampled in order to generate a set of alternative model solutions based on the probability distribution of the individual parameters.

Assuming a Gaussian approximation of the inverse problem, the tomographic solution obtained through inversion can be fully represented as a distribution with mean equal to the estimated model and variance expressed by the values along the diagonal of the posterior covariance matrix Cm^{post}, which therefore describes the final uncertainty on the parameters after the update provided by the data.

Starting from $m$est and Cm^{post}, several plausible model realisations, consistent with data and adopted regularizations, were generated. The solutions obtained by sampling the posterior covariance show that, while introducing local variations within the estimated uncertainty range, the main velocity patterns are reproducible between the different realisations and remain consistent with the final estimated solution. In particular, the position and extent of the main velocity structures were substantially stable, indicating that the obtained models are robust and do not depend critically on specific numerical choices or individual observation.

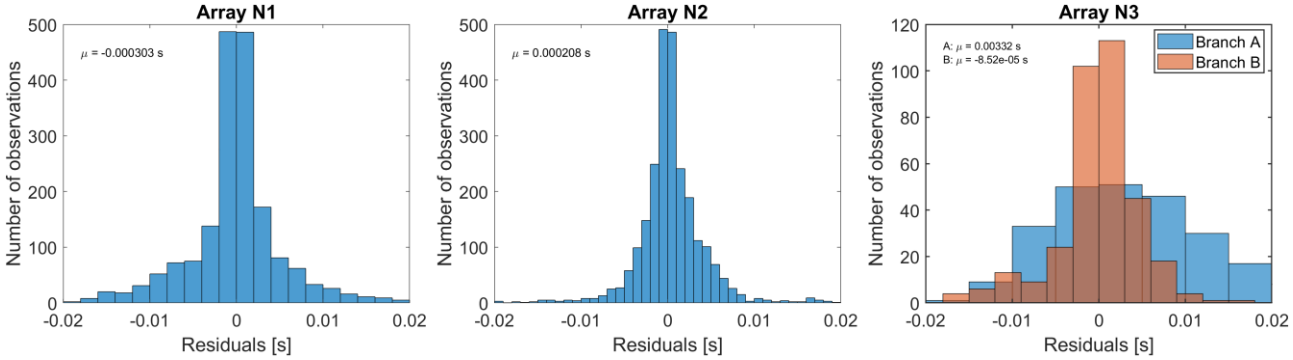

Figure 6.11 - Distributions of residuals between observed and modelled data. The distributions are well-centred around zero, indicating the absence of significant systematic biases, while dispersion provides a measure of the quality of the models fit to the data.

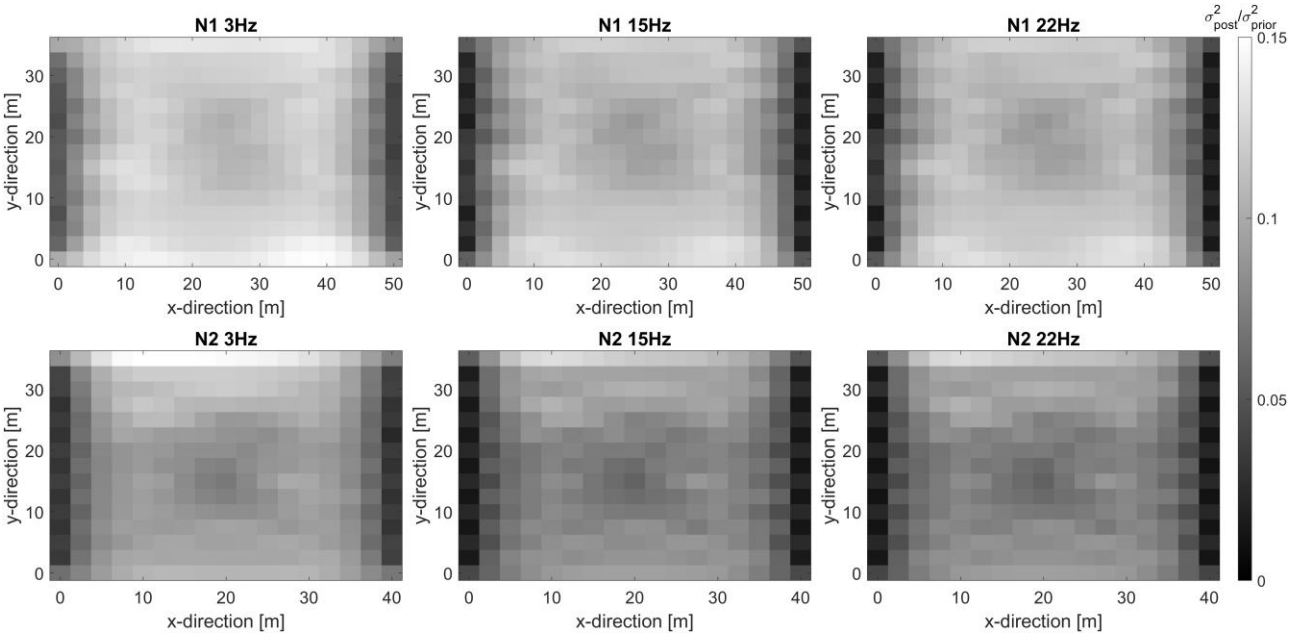

Figure 6.12 - Spatial distribution of the ratio between posterior and prior variance at different frequencies for arrays N1 and N2. Zones characterized by reduced ratio values correspond to regions where the information contained in the data produced a significant reduction in uncertainty compared to the assumed priori model. The colour scale is truncated to emphasize low relative variance obtained.

Although this approach provides a useful indication of the stability of inverted solutions within the model space defined by the inversion, a more quantitative and empirical assessment of the uncertainty associated with the tomographic models was obtained through a bootstrap analysis following Picozzi et al. (2009).

The bootstrap analysis consists of repeatedly performing the tomographic inversion on a large number of synthetic datasets (bootstrap samples) obtained by randomly sampling with replacement the set of observed travel-times. Each bootstrap sample has the same size as the original dataset, and the uncertainty estimate is provided by the standard deviation of the distribution of the reconstructed phase velocity models. For each array, 2000 bootstrap iterations were performed, and the resulting standard deviation map provides a cell-by-cell empirical estimate of the model uncertainty.

For array N1, the standard deviation (σ) remains below 40 m/s across the entire domain and all investigated frequencies. However, a zone of relatively higher uncertainty is observed, reaching approximately 20% of the estimated velocity, which persists consistently across the frequency range and follows a geometry related to the ray paths between specific station pairs. This persistent pattern suggests that the higher uncertainty in this area may be attributed to picking errors on specific inter-station dispersion curves rather than to an intrinsic lack of resolution of the tomographic model.

For array N2, the results show that the standard deviation σ remains within 10 m/s in the central areas of the domain, corresponding to approximately 5% of the estimated velocity, and does not exceed 20 m/s in the peripheral areas where ray coverage is less dense, corresponding to approximately 10% of the estimated velocity. An exception is represented by the 2 Hz map, for which σ values locally reach up to ~40 m/s along the upper and lower edges of the investigated domain,

where ray coverage is poorest. In general, uncertainty tends to decrease at higher frequencies, where inter-station coherence is greatest.
For array N3, bootstrap analysis was conducted separately for branches A and B. In both cases, standard deviation values almost always remain below 20 m/s across the entire offset-frequency domain, indicating considerable robustness of the obtained models. Small differences are observed between the two branches: Branch A shows slightly greater uncertainty at high frequencies, while Branch B shows slightly higher values at low frequencies, both of which are still contained within limits compatible with a reliable interpretation of the results.
Overall, the bootstrap analysis results confirm the robustness of the obtained tomographic models for all arrays. The main velocity anomalies are stable across replicas, indicating that the reconstructed velocity patterns do not critically depend on single observations or specific numerical choices. The standard deviation maps for arrays N1 and N2 and the offset-frequency sections of σ for branches A and B of array N3 are given in Appendix A from Fig. A6 to A9.

The offset–wavelength pseudo-sections provided by array N3, as well as the phase velocity maps obtained from arrays N1 and N2, can be used for an initial description of the expected lateral variability of the near-surface structures and allow main velocity patterns to be highlighted within the explored domains. However, a representation based purely on frequency (or wavelength) does not allow for direct interpretation in terms of depth or stratigraphic structure, making it more difficult to establish a direct relation between observed phase velocity anomalies and the real portions of the subsoil actually sampled by the surface waves.
For this reason, the obtained phase velocity maps and sections are exploited as input for the reconstruction of local shear wave (Vs) velocity profiles. This step allows the information expressed as a function of frequency or wavelength to be converted into more strictly depth-dependent models, providing a more consistent representation of the subsoil structures.
In particular, for arrays N1 and N2 the phase velocity maps are combined to obtain 3D Vs models, while for array N3 the phase velocity sections are used to reconstruct 2D Vs sections along the two branches of the array. This approach allows for a more complete and consistent description of the lateral and vertical heterogeneities in the elastic properties.

6.3.3 Local inversion dispersion curve inversion for 3D Vs model

The phase velocity maps and pseudo-sections obtained through tomographic inversion certainly provided a qualitative and partially quantitative description of the expected subsoil variability at the investigated frequencies. By converting the frequencies into wavelengths, it was also possible a more consistent visualization of these heterogeneities in greater detail along the branches of array N3. However, as already discussed in the previous paragraph, these results do not allow for a rigid interpretation in terms of stratigraphy or vertical distribution of elastic properties, since surface waves depth sensitivity depends on the considered wavelength which does not always show a strictly linear relationship with the investigated depth.

In order to transform the outcomes obtained through inversion into a uniform representation of the subsoil stratigraphy, local dispersion curves were inverted to estimate shear wave velocity (Vs) profiles. This step allows effective vertical models, representative of the local subsoil structure, to be obtained.
It is important here to remind again that the objective of this final step is not to reconstruct high-resolution Vs models in the strict sense, nor to exhaustively explore the solution space. On the contrary, inversion is used as a rapid and robust tool for obtaining effective Vs profiles consistent with the observed local dispersion curves and suitable for systematic application to a large number of curves.
In this regard, for arrays N1 and N2, the inversion of the local dispersion curves allowed the reconstruction of a 3D representation of the Vs near-surface structure composed respectively of 315 and 255 vertical profiles, while for array N3 it allowed to obtain Vs pseudo-sections along the two branches of 20 vertical profiles each.
Starting local dispersion curves are illustrated in Fig. 6.13 discretized based on the array.
Considering array N3, which covers a larger portion of the site through an 'L' geometry, the curves are less numerous and show a clear organisation into two distinct groups, corresponding to the two branches of the array.

While remaining confined to a fairly small phase velocity range (always belo 200m/s), the curves of the two branches are well discretized and characterised by consistent dispersion trends within each group. This separation suggested, as already observed in the offset–frequency and offset–wavelength pseudo-sections, that the two branches of the N3 array actually sampled portions of the subsoil with different elastic characteristics, linked to more or less sharp lateral variations.
For arrays N1 and N2, on the other hand, which are characterised by a 2D geometry and higher spatial resolution, the number of local curves is significantly higher. In this case, the number of curves makes it less straightforward to identify clear clusters, but it is also possible to highlight how both arrays explore a very similar velocity range, with the exception of some curves of the array N2 which towards low frequencies reach phase velocities up to 250m/s.
Considering N2, the group of curves characterised by lower velocities can be easily associated with the central portions of the studied domain, corresponding to the low-velocity anomaly persisting across the entire frequency range, with phase velocities consistently around 80–90 m/s. The N1 array, located in an area with no surface evidence of eruptive activity, shows a slightly more dispersed behaviour, without clearly distinct clusters, and characterised by slower velocities even at low frequencies (<5Hz), where some curves barely reach 100m/s.
A significant aspect that emerges from the comparison between the three configurations concerns the role of spatial resolution. The greater density of sensors in arrays N1 and N2 allows for a wider range of velocities to be resolved and greater detail to be preserved in the local dispersion curves. Conversely, the N3 array, characterised by a wider geophone spacing (20 m) and a substantially 1D geometry, provides essentially continuous curves, perhaps blind to small-scale structures, in which any anomalies on a scale smaller than the inter-station distance are inevitably smoothed and partially suppressed. In other words, localised anomalies

that would be resolvable with a spacing of a few metres may appear flattened when sampled over longer distances.

For these reasons, results obtained from the N3 array should be interpreted primarily as a tool to frame the general context of the site, useful to link local observations form arrays N1 and N2 with an overall structural picture and with the results of previous studies. The N1 and N2 arrays, on the other hand, represent higher level of detail and allow for a more direct assessment of the proposed methodology to resolve lateral velocity variations on spatial scales of a few tens of metres.

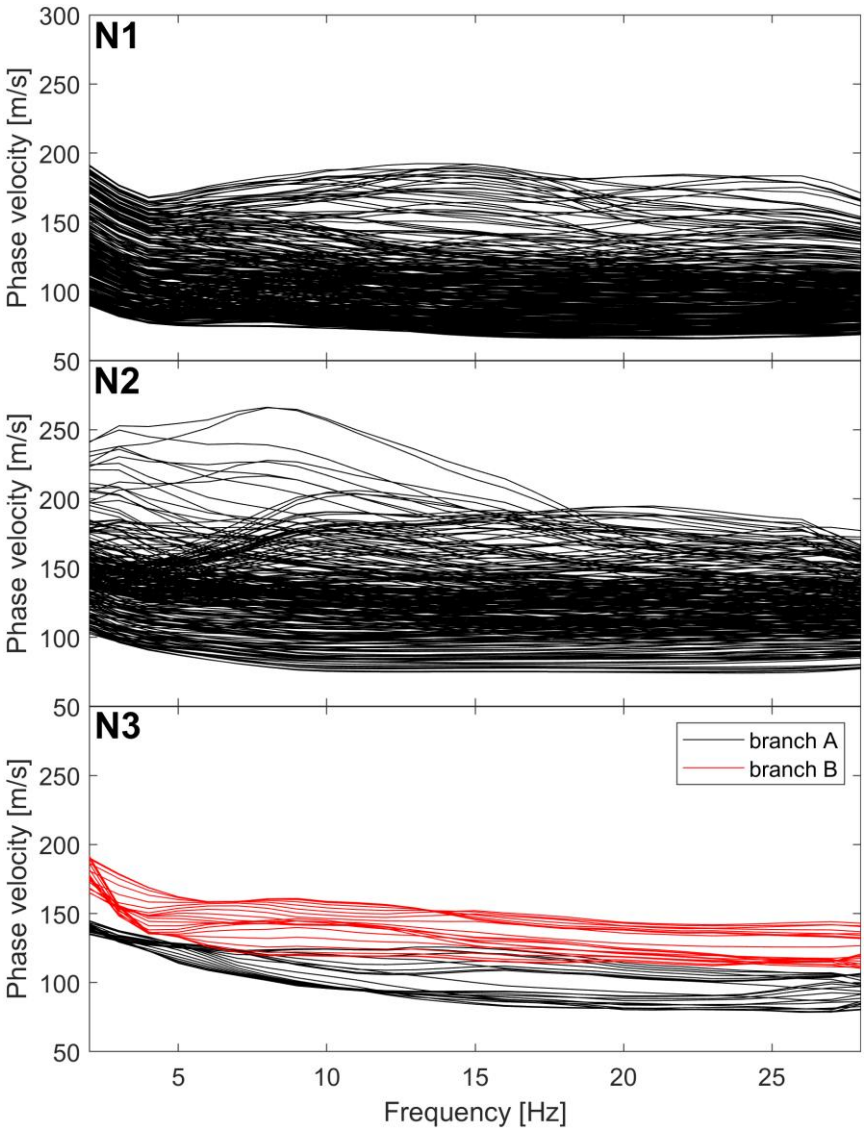

Figure 6.13 - Local Rayleigh phase dispersion curves at the end of the tomographic inversion for the three arrays N1, N2 and N3. For each array, the curves corresponding to the previous model parameters are reported. In the case of array N3, the curves for the two branches are shown separately (branch A in black and branch B in red). The distributions highlight the differences in the phase velocity fields reconstructed from the three investigations and reflect the subsurface heterogeneity of the studied area.

Based on the results presented in Fig. 6.13, every local dispersion curve was inverted in order to obtain reliable Vs profiles. Inversion was performed through procedure described in Section 5.4, providing effective Vs profiles from dispersion curves, assuming a link between frequency and sensitivity depth of Rayleigh waves.
For the inversion of the entire set of local dispersion curves (on the order of a few hundreds), a parameterization based on 10 layers was adopted, associated with ten discrete frequencies distributed in the range 2–28 Hz.
Within each inversion, the layer thicknesses, obtained by roughly converting the wavelengths into depth and then into layers, are kept fixed in order to avoid potential trade-off processes between updates in terms of velocity and thickness. This inversion scheme was applied consistently to all analysed arrays.
In most inversions, the observed local dispersion curves were well-reproduced, with a misfit value, expressed as normalized percentage difference between 0 and 1, always below the 1% for all frequencies, and the obtained Vs profiles were then considered valid and effective solutions to describe the local dispersion behaviour. However, in a limited number of cases, the fit between observed and modelled curves was found to be less optimal (between 2 and 5%). These discrepancies should not be interpreted as a procedural error on the tomographic inversion algorithm, but rather as a consequence of the indirect nature of local dispersion curves. Indeed, they represent the result of a procedures chain that includes specific assumptions, approximations and data regularizations. The combination of the uncertainties introduced in each step can lead to intrinsically more complex or less constrained local curves. This aspect is particularly evident for curves associated with parameters poorly sampled, and thus constrained by tomographic inversion, for which data sensitivity is particularly reduced. Consequently, it is not always possible to obtain an optimal reproduction of all local dispersion curves.
At the end of the first inversion, the curves characterized by the worst fit were subjected to a second inversion run. At this stage, some inversion parameters were varied, in particular, the number of layers was changed in order to change the number of free parameters; in parallel, different values of the penetration coefficient were tested to improve the model's ability to reproduce the observed trend of the curves.
The lack of prior information or constraints during the iterations, such as those used for tomographic inversion, obviously sharply contrasts with the non-uniqueness of the problem, whose solution for each local curve must necessarily be seen as one possibility among many possible solutions.
In the inversion process, both Poisson's ratio and the penetration coefficient used to link the wavelength of Rayleigh waves to the sensitivity depth were assumed to be constant for all frequencies and depths and took respectively the values 0.3 and 0.8.
Furthermore, the inversion was carried out considering only the shear wave velocities (Vs) as free parameters, while the Vp were recalculated at each iteration based on the assumed value of Poisson's ratio. As a result, the Vp/Vs ratio is also constant along each velocity profile.
Despite these assumptions, the approach adopted allows Vs models consistent with the observed dispersion curves to be obtained and provides an effective estimate of the main vertical and lateral variations in elastic properties, which is adequate for the exploratory objectives of this thesis.

Fig. 6.14 shows some of the results in terms of local curves to be reproduced and theoretical curves modelled. In normally dispersive media, the inversion procedure proved to converge quite easily towards a solution congruent with the observed data. Since the initial model derives directly from the observed local curve itself, it is easy to intuit that the solutions were necessarily found in the vicinity of that initial subsurface model without exploring other combinations of models further away from the initial model. Despite this, the observed data tends to be well reproduced, providing at least a plausible solution.

Inversion scheme was first performed for array N3 which, even characterized by lower lateral resolution, provides a general view of the site through the offset-depth sections along the branches of the 'L-shaped' array. The two branches of the array N3 (Fig. 6.15), even when converted into Vs sections, demonstrate the clear heterogeneity of the area, showing marked different velocities along the two directions investigated and across the approximately 50m depth sampled.

Branch A, oriented approximately in a N–S direction, and branch B, which develops mainly in an E–W direction, in fact appear to sample partially different structural contexts, despite showing some common features. In branch B the presence of high-velocity structures is clearly visible at depths between 25 and 50 m, with Vs values reaching and exceeding 250 m/s. This level can be interpreted as a relatively rigid geological bedrock in the context of the study area. This high-velocity horizon is overall continuous at depth but shows a localized lowering between 60 and 80m along the offset, suggesting a gradually reduction in the stiffness of the medium.

A particularly relevant element clearly visible in branch B is the low-velocity anomaly near the zero-offsets position, that extends eastward to intermediate depths up to 20-25m. This structure could be consistent with the presence of anomalies related to surface volcanic activity and could represent a small volume of less consolidated material or a potential subsurface reservoir associated with the central vent of the area.

Branch A, instead, shows a good connection with branch B near the surface. Despite it shows lower velocities in the first few meters and higher values in depth than the lateral portions of branch B, the relative configurations are overall similar between the two branches. In branch A, a relatively continuous low-velocity layer is also recognizable along the entire offset between about 5 and 10 m depth, with Vs values tending to decrease slightly as one proceeds south. Below 20 m depth, the section shows a progressive decrease in velocities towards the south, indicating a lateral variation in elastic properties even in the deeper layers.

Branch A, instead, shows a good agreement with branch B near the surface. Despite it shows lower velocities in the first few meters and higher values in depth than the lateral portions of branch B, the relative configurations are overall similar between the two branches. In branch A, a relatively continuous low-velocity layer is also recognizable along the entire offset between about 5 and 10 m depth, with Vs values tending to slightly decrease proceeding toward south. Below 20 m depth, the section shows a progressive decrease in velocities towards the south, indicating a lateral variation in elastic properties in the deeper layers.

The slight discrepancy observed between the two sections in terms of velocity is not unexpected and can be partly attributed to the acquisition geometry. The geophones of the two branches sample portions of the

subsurface that are influenced by structures present even at a distance along each branch. Consequently, the sections can be assumed to represent integrated responses to partially overlapping structural contexts: branch A reflects the overall lower velocities of the sampled area, while branch B is also impacted by the general sharp contrast approximately at 30m depth.

Overall, although the two branches do not perfectly converge in depth each other, they show good coherence in the surface layers and provide an overall consistent structural picture. The estimated velocities are within the expected ranges based on the literature, and the main low-velocity anomalies at depths are consistent with the main surface expressions linked to the fluid movements.

While the size and geometry of such anomalies should be interpreted with caution, given the simplified inversion approach, the presence of low-velocity structures beneath the main vent sampled by the array N3 represents a significant result and a solid starting point for site interpretation.

Smaller anomalies are also recognizable along the offsets of both sections, manifesting as low-velocity areas embedded in higher-velocity contexts, or vice versa.

It is important to emphasize that the reliability of these minor structures cannot be considered unambiguous. While anomalies that persist along the offset or at depth can be attributed to stable contributions from diverse local dispersion curves, smaller-scale anomalies could instead be influenced by local fluctuations in the dispersion curves, potentially reflecting the effects of regularization terms (both spatial and frequency) introduced into the inversion scheme, and not necessarily correspond to real subsurface heterogeneities.

In a more general perspective, the analysis of N3 array sections provides a useful framework for interpreting local-scale structures. The highest resolution arrays N1 and N2, suitably projected, can be contextualized along the branches of the array N3: the array N1 is close to the terminal portion of branch A, while the array N2 is located in proximity to the final part of branch B. This allows a direct comparison between results obtained at different spatial scales, which will be discussed below.

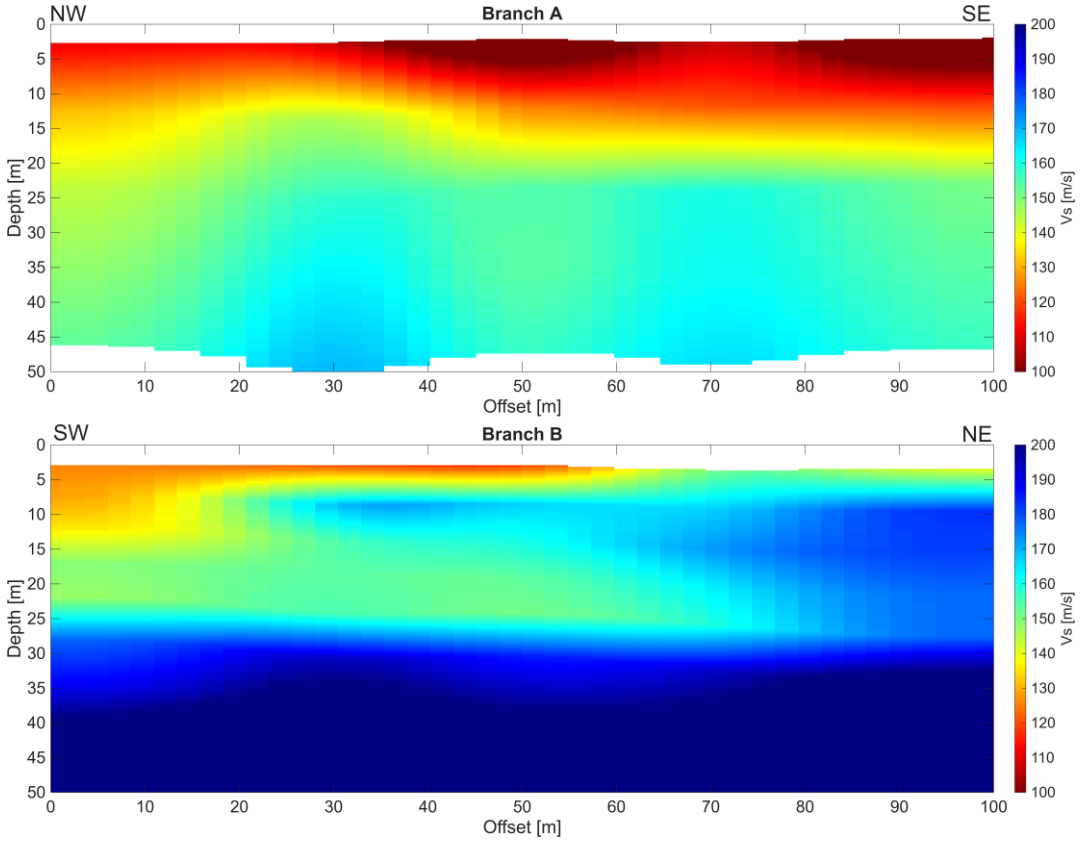

Figure 6.15 - Shear wave velocity (Vs) sections obtained by inverting the array N3 along the two array branches. Branch A (upper panel) oriented approximately N–S, and branch B (lower panel), oriented E–W, show marked lateral variability of velocities in the first ~50 m of depth investigated. The sections, especially branch B, highlight the presence of low-velocity structures in the subsurface layers and the emergence of more stiff structures at depth, with significant differences between the two branches, interpretable in relation to the different local geological context.

Then, attention was focused on arrays N1 and N2, characterised by 2D geometry and greater geophone density. For these arrays, the inversion of local curves allows the reconstruction of sets of Vs profiles distributed in the x-y plane, which can be combined to obtain a 3D model of the near-surface shear wave velocity. These results represent the highest level of detail in the performed analysis and form the basis for discussing lateral subsoil structures on scales of a few tens of metres.

Fig. 6.16 and 6.17 show the shear wave velocity (Vs) models obtained for arrays N1 and N2, hereafter models N1 and N2. The models are represented by horizontal Vs maps extracted at fixed depths (5m, 10m, 15m, 20m, 25m and 30m); the vertical direction is deliberately exaggerated to make it easier to observe velocity variations as the depth increases and to allow a direct comparison between the different investigated subsoil portions.

Considering the Vs models as a whole, the results obtained through the entire workflow led to consistent velocity values across the different arrays, in agreement with what already emerged in the previous steps, from the dispersion curves extraction to the phase velocities inversion. In both arrays, the final models show marked velocity contrasts, while remaining within a relatively narrow velocity interval, with Vs values falling just below 100 m/s and reaching values slightly above 200 m/s. These values are also in good agreement with shear wave velocities identified along the sections of the array N3.

Looking at the two models in more detail, model N2 shows a well-limited low-velocity anomaly located precisely beneath the eruptive vent that progressively extends towards the surface, with minimum Vs values located between approximately 5 and 10 m depth. This restricted geometry is clearly visible both on the velocity maps in Fig. 6.17, and in Fig. 6.18 which represents a section along the y-axis of the N2 model extracted at x = 20 m. High velocities are also observed along the surface edges of the array N2, which appear to laterally close the central anomaly at 5m.

The N1 model, on the other hand, has a rather different configuration and is more difficult to interpret. In this case, the low velocities do not appear concentrated in a small volume attributable to a possible fluid upwelling system but are laterally more continuous within the sampled domain. Again, a vertical section constructed along the y-direction and extracted at x = 20 m (Fig. 6.19) shows how low-velocity zones develop laterally, while towards the surface they tend to overlap high-velocity materials. This configuration suggests the presence of contact between materials with different elastic properties, rather than localized anomalies associated with a conduit or mud-feed structure.

As regards models N1 and N2, the high density of sensors and 2D geometry made it possible to obtain a significantly higher resolution than N3. Array N2 is surely the one that shows the most structured and interpretable results.

Overall, while for array N2 the low-velocity anomaly (approximately 100–120 m/s) appears well-defined and interpretable as a localized structure, probably related to the eruptive vent feeding system, for array N1 the velocity contrasts seem rather to reflect the presence of different in-place materials characterized by their own elastic properties. This interpretation is also consistent with what was observed in section A of the array N3, along which a lateral transition from lower to higher velocities is evident, maintaining a low-velocity surface coverage.

The estimated velocities are nevertheless comparable to those of N2, in accordance with the relative spatial proximity of the two configurations, but without highlighting anomalies directly interpretable as derived from the eruptive and degassing processes. This behaviour is consistent with the absence, in the area covered by N1, of evident surface manifestations of volcanic activity or fluid emissions. However, even in the model N1 there is a consistent shallow velocity reversal (also visible in section in Fig. 6.19) between 5m and 20m possibly related to the underground fluid's movement.

In both the N1 and N2 model contexts, the velocity variations were resolved quite well, highlighting changes in space of a few tens of meters. Despite this, it must also be noted that sharp velocity contrasts were never highlighted. This, although plausible, could be the effect of the regularization strategy operated during tomographic inversion, which might smooth out even abrupt velocity variations within a few meters. Based on this, many structures that appear to progressively change from higher to lower velocities, or vice versa, could be much more marked than the obtained models show.

Some representative results of the inversion of local dispersion curves are given in Fig. A10 (array N1) and Fig. A11 (array N2) in Appendix A . These examples show the comparison between observed data (local curves

obtained from tomographic inversion) and inverted models for several selected cases from array N1 and N2, providing visual evidence of the quality of fit and stability of the obtained effective solutions.

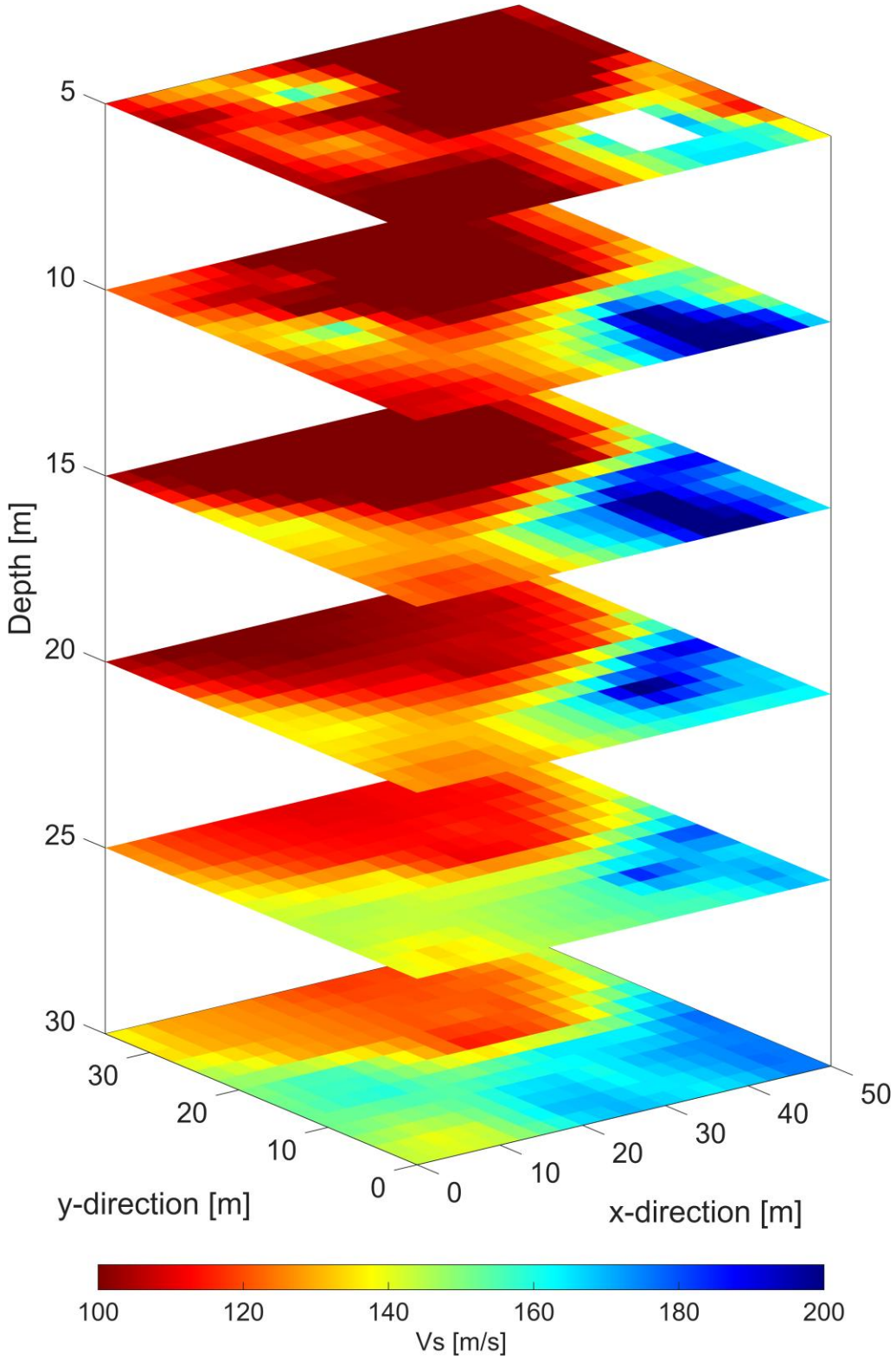

Figure 6.16 - 3D visualization of the shear wave velocity (Vs) model obtained from the local dispersion curves inversion for array N1. A set of horizontal Vs maps are extracted at depths 5m, 10m, 15m, 20m, 25m and 30m, in order to highlight

the lateral and vertical variations in velocities within the investigated domain. The colour scale represents Vs values, with red colour associated with lower velocities and blue with higher velocities.

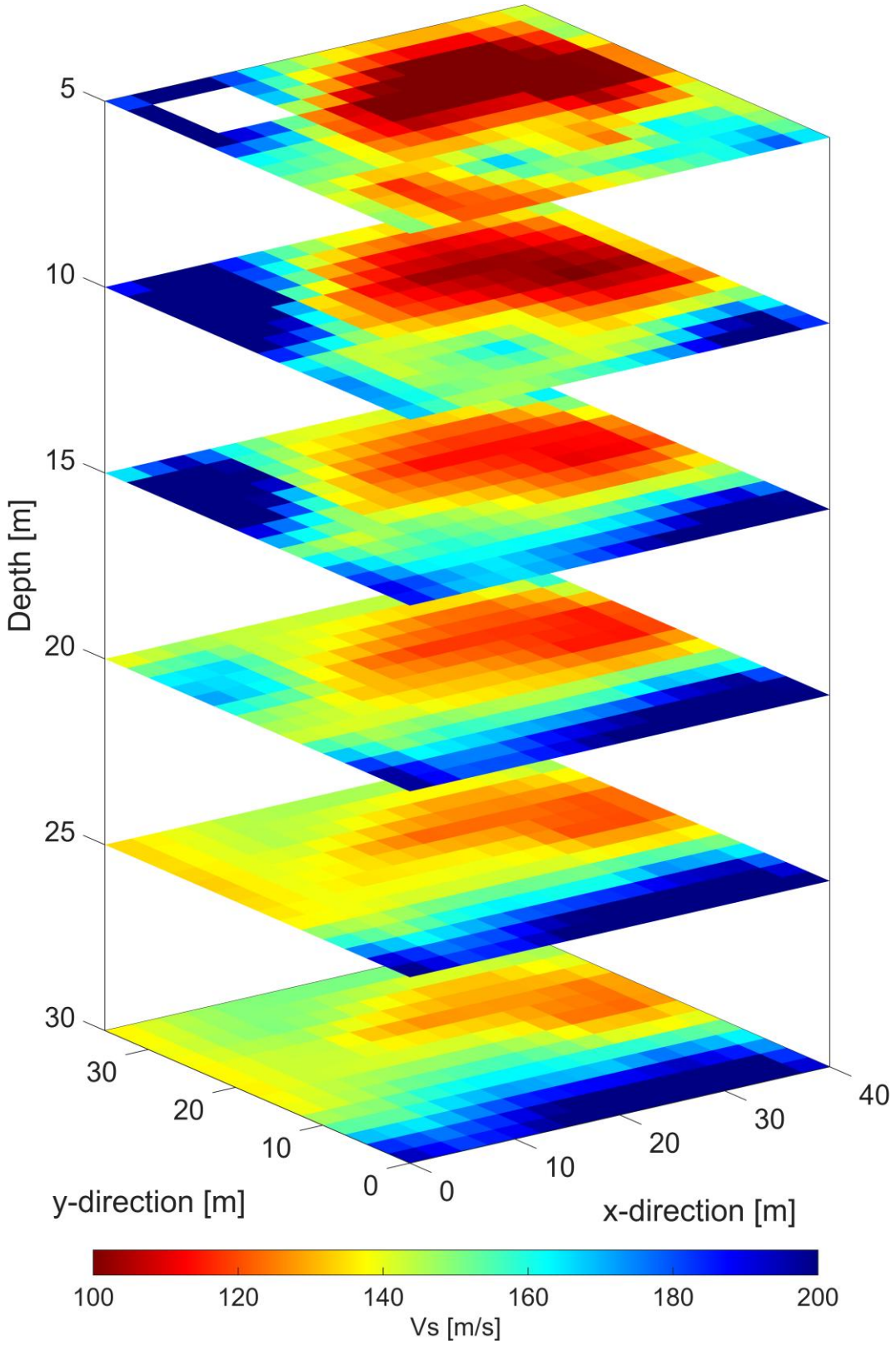

Figure 6.17 - 3D plot of the shear wave velocity (Vs) model obtained from array N2. As with the array N1, the model is represented by a set Vs maps at depths 5m 10m, 15m, 20m, 25m and 30m, which allow to visualize the geometry of the main structures at different velocities and their spatial variability in depth and laterally.

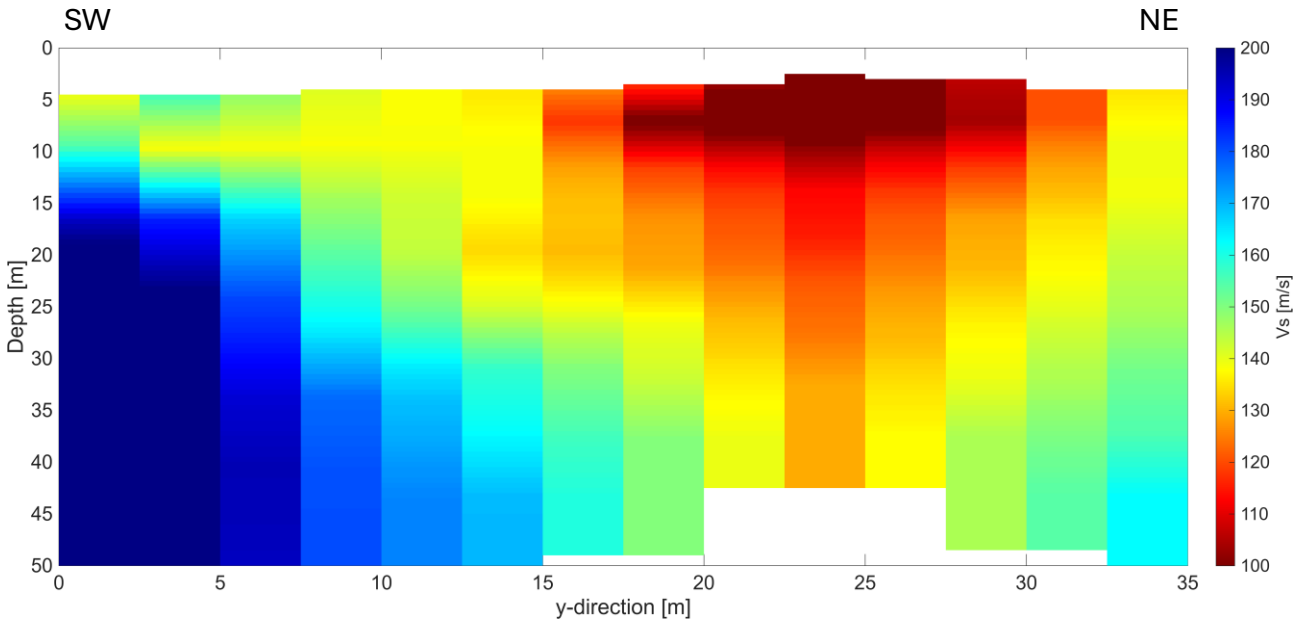

Figure 6.18 - Vertical section of the shear wave velocity (Vs) model obtained for the array N2, extracted along the y-direction at fixed x=20 m coordinate. The section shows the depth velocity distribution along an approximately eastward direction (from x=0 to x=35m). The presence of a low-velocity anomaly concentrated in the central part of the profile is evident, which tends to extend towards the surface resulting laterally delimited by materials characterised by higher velocities.

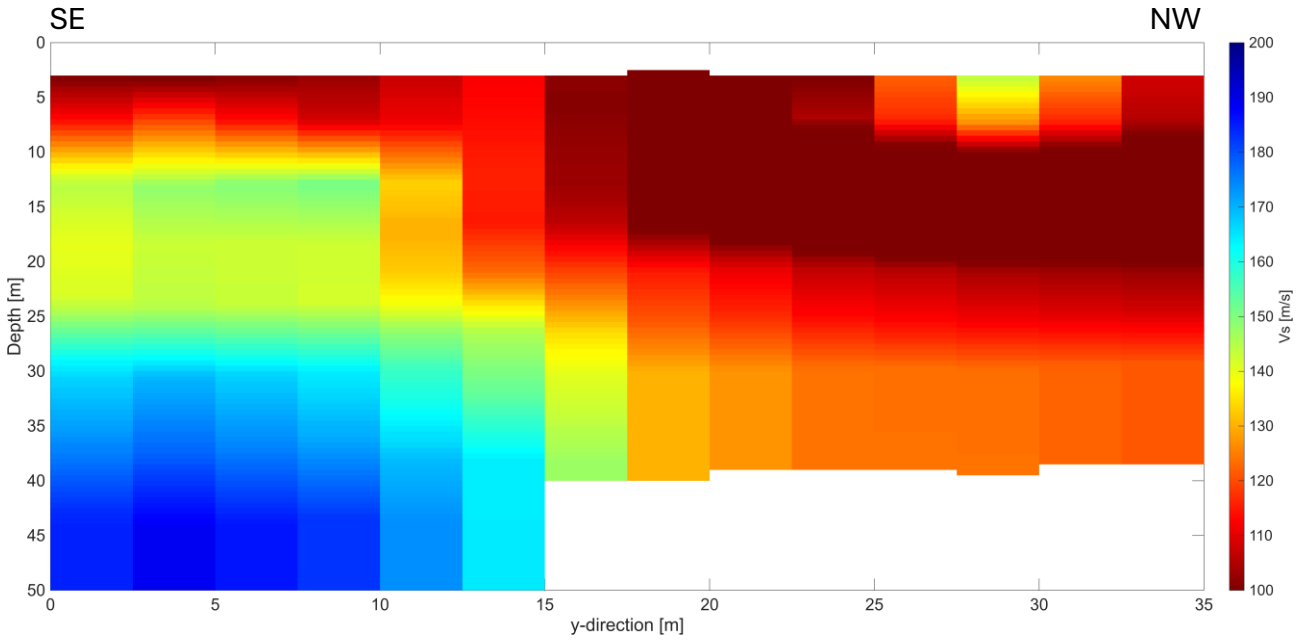

Figure 6.19 – Vertical section of the shear wave velocity (Vs) model obtained for array N1, extracted along the y-direction at fixed x=20m coordinate. The section, oriented approximately in the north direction (from x=0 to x=35m) shows a Vs distribution characterised by marked lateral contrasts. The low-velocity zones are laterally continuous at depth and tend to overlap, towards the surface, with higher velocity materials, suggesting the presence of in-place materials with different elastic properties rather than a localized anomaly.

6.4 Models Interpretation

The interpretation of shear wave velocity models is addressed here by integrating the results obtained with observations from independent studies conducted in the same area. We start with the array N3, which provides an overall and reference view of the investigated subsurface.

The two 100m-long sections reconstructed from the array N3 clearly highlight the marked heterogeneities within the study area, both laterally and in depth. In particular, section A, oriented approximately in the NW-SE direction, shows a laterally continuous low-velocity surface coverage with a thickness between approximately 5 and 10 m, characterized by Vs values of the order of 100m/s (Branch A in Fig. 6.15). This configuration is in good agreement with the geoelectric profiles 2 and 3 produced by Lupi et al. (2016), in which the first 10 m of subsurface are characterised by the lowest resistivity values. These values can be interpreted as indicative of the strong presence of fluids beneath the surface, also in relation to the surface activity of the mud volcanoes, within which resistivities values are comparable to those observed in the uppermost levels of the profiles (Lupi et al., 2016).

Several studies (e.g., Sciarra et al., 2017; Carfagna et al., 2024) have documented diffuse gas seepage across the entire investigated area, not exclusively concentrated near the eruptive vents, suggesting that gas and fluid accumulation in the uppermost subsurface is a pervasive phenomenon extending over the first ~10 m of depth. Although a direct causal relationship between gas presence and Vs reduction cannot be established, since gas concentration is known to primarily affect Vp rather than Vs (Lee, 2004), the persistent interaction between rising fluids and the shallow geological matrix may contribute to a progressive mechanical weakening of the material, which in turn could explain the low-velocity horizon observed in section A.

Section A also contains a zone with moderately low velocities along the firsts 30-40m of the offset, even below 20 m depth; however, it is in section B, approximately oriented NE-SW, that the geometry of a potential anomaly is most clearly interpretable. In this section, in fact, a low-velocity area rising to a depth of about 15 m near the cone can be observed. This low-velocity structure tends to deepen slightly towards northeast. At the same time, a low-velocity surface layer persists up to half of the investigated offset (~50m).

Again, the low-velocity distribution is consistent with the low-resistivity anomalies identified in the geoelectric investigations. Looking at profile 1 of Lupi et al. (2016), the subsurface portion between close to profile 3 shows low resistivity features mainly located between 10 and 30 m depth, in good agreement with the low-velocity structures highlighted by seismic models of section B.

Partial disagreement between the two branches of array N3 is observed at depth, particularly near the zero offset, which coincides with the position of the central eruptive vent of the area. In section B a high-velocity basement is clearly recognisable, which appears relatively continuous along the section, and which can be interpreted as a more rigid horizon, similar to a geological bedrock. The Vs values associated with this level locally reach up to 250–300 m/s and are consistent, both in terms of velocity and depth, with surface-waves punctual measurements available in literature (see Vs profiles from Antunes et al., 2022).

Beyond the local discrepancy observed at the zero-offset of both branches, it is possible to make some more general structural considerations. In particular, proceeding towards the NE the subsoil is expected to be characterised by stiffer materials, as suggested by the results of the gravimetric investigations (see Figure 2 and 3 from Nespoli et al., 2023), in response of the lack of shallow loose material produced by the fluid reservoirs. This transition can be easily recognized across section B. The low-velocity anomaly that propagates from zero-offset in the NE direction inserts into rigid material ever closer to the surface, which precisely towards NE reaches a few meters of depth. This velocity configuration is consistent with the presence of a structurally more competent (geological) bedrock in the NE portion of the site, as also indicated by the density anomalies identified by gravimetric investigations.

This potential structural setting is partially confirmed by some two-station dispersion analysis computed beyond the road in the NE direction (see Fig. 6.3). Dispersion curve extracted and then inverted provided valuable insight about the sharp changes in the subsoil stratigraphy only a few tens of meters away. In Fig. 6.20 some of the inverted Vs profiles are shown. All profiles show a normally dispersive behaviour, with Vs values already reaching approximately 200 m/s or above within the first 10 m of depth, values that are never observed in section A and are typically reached only below 30 m along section B.

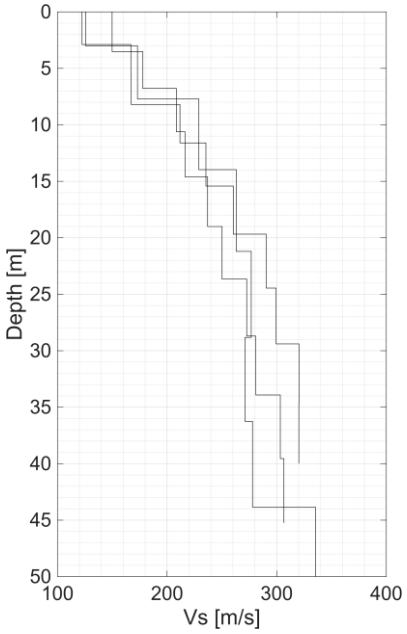

Figure 6.20 – Shear wave velocity (Vs) profiles obtained by inverting dispersion curves extracted from pairs of stations, located in the northeastern portion of the study area, beyond the road. The profiles show a progressive increase in velocities with depth and indicate the presence of stiffer and more competent materials that tend to approach the surface proceeding towards NE, in agreement with what was observed in the sections and 2D models previously discussed.

Taken together, the results obtained from array N3 suggest the presence of a heterogeneous subsurface, with a low-velocity surface layer in the uppermost ~0–10 m likely related to mud and gas accumulations. Deeper velocity anomalies are also partially consistent with the fluid distribution delineated by geoelectric investigations, further supporting the overall interpretation. This consistency between previous electrical resistivity investigations and the interpretation of Vs models anyway, should be interpreted with caution: inter-station distance along the N3 branches (20m) necessarily tends to suppress local velocity fluctuations in favour of more smoothed lateral profiles, while compared geo-electrical measurements were conducted with denser electrodes distribution (2 or 5m inter-distance between the electrodes). Despite this, the obtained Vs sections can be interpreted as representative of the overall subsurface configurations and provides a reference for the models N1 and N2.

The interpretation of the N1 and N2 models is now set into the framework provided by the sections of the array N3, which allow to constrain the main structural characteristics of the subsoil on a larger scale, based on reliable comparison with previous studies conducted. These models (N1 and N2) allow for a detailed analysis of local variations in seismic velocities within smaller portions of the study area. Unlike sections A and B, here the spatial resolution is higher, this helps to capture more abrupt velocity variations even in a few meters with a good approximation and with less risk of achieve smoothed and flat models from tomographic inversion.

Fig. 6.21 shows a set of horizontal slices of the shear wave velocity (Vs) models N1 and N2 at different depths. The representation is superimposed on a satellite basemap and allows to direct compare the lateral distribution of velocities in the two domains, analysing their evolution with depth.

Although the studied domains are spatially close, they show only partial agreement with each other: while some features are consistent, others velocity patterns show significant differences, suggesting that the main structures do not extend continuously between the models.

Focusing initially on the model N2, this can be projected approximately into the final portion of section B previously discussed. From this point of view, the relatively high velocities shown in model N2 are consistent with those observed along section B, which however shows no clear evidence of low-velocity anomalies in that area. This discrepancy is not surprising, considering that the anomaly identified at the centre of the model N2 shows maximum diameter estimated at approximately 15–20 m, making it difficult to be resolved adequately within section B.

Model N2, as already anticipated in Fig. 6.18, is instead particularly well constrained. It is characterized by a significant low-velocity anomaly located in its central part, which, inserted within an interpretative framework, can be now associated with a conduit of rising fluids toward the surface.

The shallow subsoil beneath the vent in question, as well as the surrounding area, has been subject of previous studies already cited (e.g., Carfagna et al., 2024; Antunes et al., 2022), focused on the emission of seismic signals related to volcanic activity, particularly in the vicinity of the two cones located to the northeast. Carfagna et al. (2024), through the localization of the events responsible for the volcanic tremor of the area, identified two possible conduits located below the area between the two volcanic vents.

The low-velocity anomaly highlighted below the vent is not entirely consistent with the sources localized by Carfagna et al. (2024), but it still supports the hypothesis that the northeastern area beneath the two vents represents a preferential escape path to the surface for mud and gases accumulated in the first subsoil. Moreover, looking in detail at Fig. 6.21, deeper Vs maps (20-40m) show a contact zone of also low velocities in the direction of the northern cone, highlighting the possible presence of adjacent structure that act as underground path for the fluid movements beneath the vents.

Introducing model N1, it emerges that, despite the differences highlighted above, the models show partial agreement both at greater depth (~30-40m) and in the shallower portions (0-10m), suggesting the existence of some common structural elements.

At depth (e.g., 40 m slice in Fig. 6.21), although the Vs reconstruction does not cover entirely the 2D domain, the estimated velocities are fully consistent between the models. In both cases, high velocity values, above 200 m/s, are observed, which tend to decrease as one move towards the conduit beneath the vent in model N2, while in N1 low velocities become predominant as one move to the centre of the site. This setting is particularly evident looking at the N1 section in Fig. 6.19, where proceeding in NW direction, soft material becomes prevalent even at 30-35m depth. This result is in part consistent with what observed along section A, where a continuous, low-velocity layer with a rather constant thickness (between 15-20 m) was identified.

However, the bootstrap uncertainty analysis discussed in Section 6.3.2 (Fig. A6) suggests that the high-velocity area identified in model N1 should be interpreted with some caution. The standard deviation maps show that this portion of the investigated domain corresponds to the least constrained area of the model, exhibiting relatively higher uncertainty compared to the surrounding lower-velocity zones. This pattern may indicate that the high-velocity values are partially influenced by a limited number of specific station pairs, and that the actual velocity contrast in those areas could be less pronounced than what the tomographic solution suggests.

Overall, the area investigated by models N1 and N2 appears to represent a complex transitional zone between quite different subsurface contexts. The eastern–northeastern (E–NE) part is characterized by relatively high velocities already in the first meters of subsoil, as previously discussed, while proceeding towards the centre of the volcanic field, the subsoil appears strongly influenced by the continuous ejection of fluid mud, which has progressively filled the depression in which the main volcanic vents arise.

In this area, the low-velocity patterns observed in models N1 and N2 cannot be simply attributed to continuous horizons but rather reflect the presence of conduits and pathways that play an active role in the fluid dynamics of the volcanic system, acting as temporary accumulation and redistribution areas for gas and mud, ultimately controlling the surface volcanic activity and outgassing phenomena.

In such a complex setting, where lateral heterogeneities reflect not only thickness variations but also abrupt changes in material properties and fluid content, the subsoil reconstruction can be extremely challenging and the obtained models N1 and N2 have well highlighted this aspect.

Moving further E–SE, the first 10–20 m of subsoil is instead almost entirely characterised by low velocities, attributable to the filling material continuously expelled to the surface through the cones, pools and small gryphons spread over the site. These shallow levels correspond, according to Lupi et al. (2016), with near-surface low resistivity material, interpreted as layers that constantly feed the flow, acting as shallow reservoirs of gas and mud. Overall, the integrated information provided by models N1, N2, and sections A and B from the array N3 produced a consistent picture of the subsurface in the study area, highlighting a heterogeneous system beneath the surface composed of fluid rising structures and laterally continuous low-velocity surface layers. This observed spatial variability also explains the role of a multi-scale approach in interpreting velocity models. While array N3 provides the overall structural framework of the study area, the higher-resolution models obtained from arrays N1 and N2 are essential to resolve localized structures associated with fluid pathways, which tend to be smoothed out in larger-scale reconstructions. Among the two, model N2 lends itself to a more straightforward geological interpretation, as its central low-velocity anomaly can be directly associated with a conduit of rising fluids beneath the overlying eruptive vent. Model N1, located in a structurally more complex area, is more difficult to interpret unambiguously; nevertheless, the Vs sections suggest the presence of a shallow low-velocity horizon with velocities comparable to those observed within the conduit anomaly of N2, consistent with the hypothesis of a diffuse shallow reservoir feeding the widespread gas seepage observed across the site.

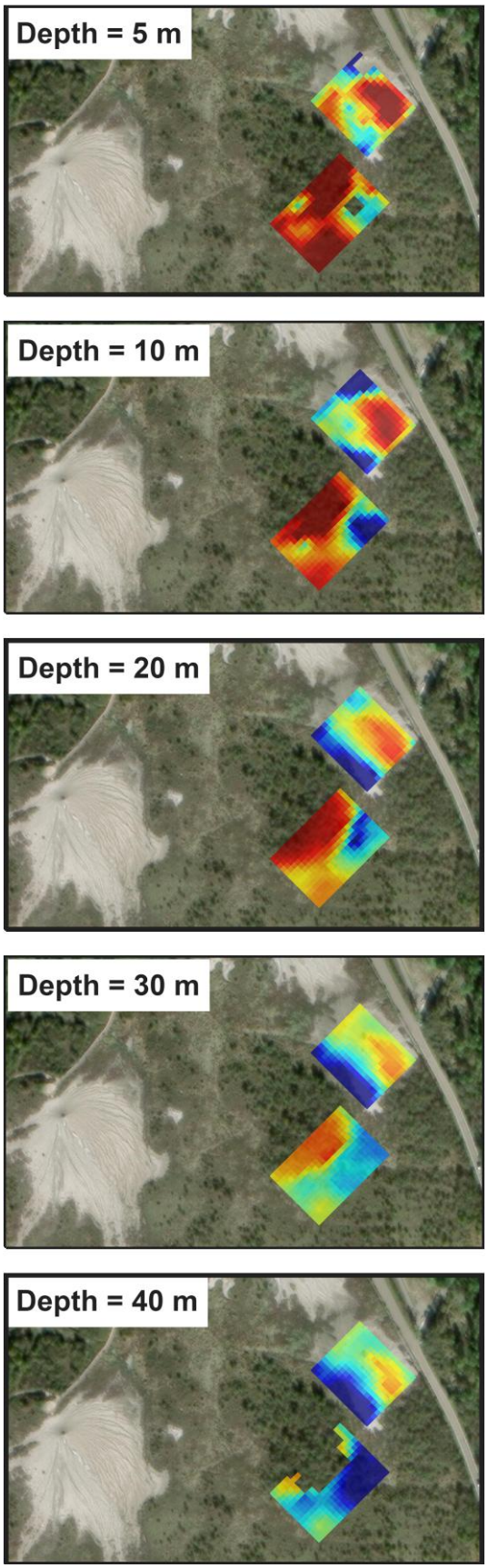

Figure 6.21 - Shear wave velocity (Vs) maps obtained for the models N1 and N2 at different depths (5, 10, 20, 30 and 40 m). The maps are placed above the satellite base map in order to contextualize the lateral distribution of velocities. The representation allows for a direct comparison between the two domains across different depths.

Chapter VII.
Conclusions

In this thesis, a methodology for seismic subsurface characterization was developed, implemented, and applied starting from the analysis of ambient seismic noise, with particular attention to the estimation of Rayleigh wave dispersion curves using acquisition configurations based on two-station approaches. The main objective of the work was to evaluate the information provided by minimal applications, both in terms of the number of sensors (16) and the duration of acquisitions (~1h), in complex and heterogeneous contexts, where the use of extended arrays or active sources is sometimes impractical.

Although ambient noise has long been an established tool for regional or large-scale studies, its use in near-surface investigations is more critical. In these contexts, in fact, the need for high frequencies often makes it preferable active seismic approaches, in which the use of repeated sources allows for a significant improvement in signal quality and a more direct reconstruction of the subsurface structure along sections. The lack of controlled sources, typical of analyses based on ambient noise, therefore produces intrinsic limitations that must be considered over the whole workflow.

The thesis starts with a detailed introduction focused on the main aspect about surface waves tomography strategies, with particular attention to the fundamental assumptions that govern their application. In this frame, the intrinsic limitations of these approaches are discussed, and several examples of their use are presented, mainly at large scales, but also including applications in smaller contexts, to highlight their potential and criticalities depending on the scale of investigation.

In Chapters II, III and IV, the foundations of seismic noise-based methods were recalled, with reference to the seismic interferometry and spatial autocorrelation methods. In this context, the link between correlation functions and the dispersive properties of surface waves was discussed, highlighting the main intrinsic limitations of these approaches, especially regarding the assumptions of isotropy of the noise wavefield.

The codes and algorithms implemented during the work are extensively discussed in Chapter V, with particular attention to the procedures developed for the extraction of phase dispersion curves from pairs of stations. An innovative element is represented by the picking algorithm, which jointly exploits the information contained in the group and phase velocity curves to select the physically correct phase curve, allowing to overcome the ambiguities typical of methods based exclusively on two sensors with a fixed inter-station distance.

Validation tests and synthetic experiments showed that the extraction of phase dispersion curves based on comparison with group velocity curves is fully adequate for the purpose. In particular, it emerged that relatively short-duration recordings are sufficient to obtain significant phase curves, confirming that, in this context, it is not strictly necessary to have very long acquisition times, as is generally required for large-scale applications. In both the test cases and the real case study, it was also observed that even pairs of relatively nearby stations were able to provide useful insight, managing to capture, at least in part, the dispersive behaviour of surface waves during propagation in the medium, with an inevitable increase in the uncertainty associated with the estimates.

The picking algorithm has been shown to perform effectively in the presence of registrations of good overall quality. It allowed to overcome local criticalities such as the sudden loss of coherence at specific frequencies, allowing to reconstruct reliable phase dispersion curves when analysed as a whole and appropriately compared.

In the case study, even in the presence of local irregularities, the curves obtained are overall coherent and physically consistent.

Furthermore, linear tomographic inversion of phase travel-times is illustrated and implemented in a single step by simultaneously considering all the analysed frequencies. The tomographic inversion approach adopted, although based on a linear model that assumes a straight path between the stations, which is a strong assumption, was still able to capture significant variations within the investigated domain. The use of a frequency correlation for simultaneously invert all analysed frequencies proved to be particularly useful from a physical point of view, since the expected local dispersion curves, due to the nature of the propagation process, should exhibit at least partially smooth trend, without sharp jumps in velocity between adjacent frequencies. This approach inevitably results in less robust data fit than fully independent solutions for each frequency but ensures greater conceptual robustness and better physical consistency of the result.

Finally, a novel dispersion curve inversion algorithm to efficiently handle many curves is described. It is based on the association between frequency and investigated depth. This approach, while not involving extensive sampling of the model space, allows us to obtain useful solutions by effectively reproducing the observed data. The adopted algorithm is deliberately simple and has higher level of approximation. However, the main goal was not to find a unique or optimal solution in an absolute sense, but rather to reliably reproduce the local data for a large number of available curves. In this sense, the approach has proven effectiveness, allowing for good fit between observed data and inverted models. The algorithm has proven particularly effective for handling a large number of curves. It is therefore important to underline that the proposed solution represents one of the possible interpretations compatible with the data and not a single description of the shallow subsoil.

Despite being a relatively simple approach, it was able to reach good agreement with the observed data in a few iterations, with residuals generally less than 1%. In the case of curves characterized by normally dispersive behaviour, convergence was found to be stable and rapid (few tens of iterations). Conversely, where tomographic inversion produced poorly constrained curves, such as near the domain edges or at extreme frequencies, where subsampling becomes more relevant, the algorithm showed obvious limitations. In these cases, by not exploring the model space extensively, the solution tends to remain trapped near the local configuration, being sensitive to any discontinuities or irregularities in the experimental curve.

In such cases, reducing the number of layers can improve the stability of the solution, but carries the risk of oversimplification, with the possibility of not adequately reproducing the real local complexity. It is however important to underline that the risk of obtaining excessively complex curves with rapid jumps over the frequencies is partly mitigated by the contribution of the frequency correlation parameter in the tomographic inversion, which helps to suppress unrealistic oscillations and make local curves more physically consistent before the final 1D depth inversion.

Codes and algorithms described in Chapter V were applied in a context characterized by expected heterogeneities in the uppermost subsoil, with the aim of evaluating their reliability along the entire processing chain, from the extraction of dispersion curves to the construction of final shear wave velocity (Vs) models. To this end, the field application consisted of three seismic arrays of different sizes and geometries. The smaller

arrays N1 and N2 were designed to investigate limited portions of the site with higher spatial resolution, adopting an inter-station distance of 5 m along the branches, in order to compare two distinct areas where significant variations in velocity structure were expected. The array N3, consisting of two branches 100 m long and characterized by an inter-station distance of 20 m, was instead deployed with the aim of looking at the general context of the study area and to evaluate, in a more approximate way, variations in stratigraphy at larger scales, including the possible presence of low-velocity shallow reservoirs beneath the surface.

Overall, all implemented codes have been shown to perform satisfactorily, returning consistent results in terms of errors across the different analysis steps and leading to Vs models and velocity sections that, while representing an indirect result produced by a chain of different approximations and steps, are plausible and interpretable in the considered geological context.

The first results obtained show that, despite the geometric simplicity of the adopted configurations, the two-station approach is able to return stable and physically consistent dispersion curves over a wide frequency range, on condition that appropriate selection and quality control criteria are applied.

An element of particular interest that emerged from this work is the application of a tomographic-like approach to obtain 2D velocity sections from data entirely acquired from ambient noise, where typically active investigations are used with this aim. The arrangement of geophones along a straight line indeed does not ensure the achievement of the isotropy conditions required by theoretical assumptions and therefore constitutes a critical point of the methodology. In such configurations, the directional distribution of the wavefield can be strongly anisotropic, introducing potential systematic biases into the initial dispersion curves. In particular, if the isotropy conditions are not completely satisfied for a single pair of stations, it is plausible that the same will happen for the entire set of analysed stations, introducing potential systematic deviations into the procedure. Despite this, the results obtained suggest that, at least in the case study considered, the approach is robust enough to return consistent models compared to the available literature.

A relevant contribution is represented by the introduction of regularization terms within tomographic inversion. In the case study, these terms proved to be essential to compensate for poorly sampled parameters within the domain, allowing for more stable and overall reliable models to be obtained even under subsampling conditions. Intermediate tomographic inversion of the entire process therefore plays a central role, as it allows us to transform the dispersion curves into a set of more direct observations, in terms of phase velocity maps, at distinct frequencies and to provide the local curves with a form that does not depend exclusively on the raw data, but also incorporates physical realism criteria related to the considered model.

The reliability of the tomographic models was further assessed through a bootstrap analysis, which confirmed the overall robustness of the recovered velocity patterns across all arrays, with standard deviation values remaining within acceptable limits in the well-sampled portions of the investigated domains.

A further methodological aspect concerning the choice of regularization parameters adopted in tomographic inversion is that the spatial correlation coefficients, frequency correlation terms and prior variance amplitude (σ^2) were assumed constant for all frequencies considered within the one-step inversion. This choice ensured simplicity of implementation but inevitably represents a simplification of the physical problem. In fact, low

frequencies (e.g. 3 Hz) are sensitive to larger and smoother structural variations, while higher frequencies (20–25 Hz) respond to near-surface heterogeneities. It is therefore plausible that the regularization parameters can also be frequency-dependent, for example by modulating the spatial correlation lengths and the prior variance amplitude as a function of wavelength. The introduction of a frequency-dependent regularization scheme therefore represents a possible future development of the methodology, aimed at a more physically consistent representation of the multi-scale nature of the medium.

In the case study, the models N1 and N2 allowed to highlight significant heterogeneities within the site even at very small scales, compatible with the extension of the arrays themselves. Beyond the real quantitative reliability of the individual values, which nevertheless show limited final errors overall, the results clearly indicate how the area is extremely complex, especially in the first 10–20 m, characterized by the presence of conduits and preferential paths through which fluids and gases can move within the shallow subsoil and then allow them to rise towards the surface. In particular, the N2 model highlighted a low-velocity zone located immediately below one of the main volcanic vents, with possible deeper connections to other adjacent ducts. At greater depths, the complexity of the structural setting tends to decrease, and in this sense the contribution of the Vs sections obtained from array N3 has proven particularly useful. Indeed, this array provided important insights into larger-scale variations in subsurface structure, highlighting how volcanic activity influences the site not only at shallower levels but also at greater depths.

A further important element concerns the application context chosen for the validation of the methodology. Many near-surface ANT applications have been developed in geological settings characterized by relatively gradual variations, such as layer deepening or progressive thickness variations, or simple-geometry sedimentary basin structures. In such situations, the lateral evolution of elastic properties is often regular and predictable. The site under investigation here, on the contrary, is characterized by an extremely articulated structure and localized heterogeneities, linked to the presence of conduits and fluid migration paths. In such a context, it was by no means expected that an approach based on minimal configurations and simplified tomographic inversion would lead to consistent patterns. The case of the N2 array is particularly significant: the identification of the low-velocity anomaly, extending for about 10–20 m and associable with a fluid conduit, suggests that the methodology is also able to capture small-scale lateral variations. In this sense, the introduction of regularization terms within tomographic inversion probably played a determining role, helping to stabilize the solution in the presence of non-ideal sampling and to return physically plausible models even in a structurally complex context. Although a direct quantitative validation through dedicated measurements, such as borehole logs or co-located active seismic surveys, was beyond the scope of the present study, the consistency of the obtained Vs models with results from previous geophysical investigations conducted in the same area provides a consistent and coherent picture of the subsurface, supporting the plausibility of the obtained models.

The proposed methodology proves particularly interesting as it allows obtaining 2D and 3D models of the subsurface even in narrow and highly constrained contexts, without requiring significant operational efforts.

The entire procedure is in fact based exclusively on passive seismic acquisitions, without the use of controlled sources, and on a largely automated processing chain, which requires only a post-validation phase of the results. Despite these simplifications, the approach proved to be reliable overall.

Tomographic inversion, despite its simple linear formulation compared to more refined methods, benefits from single-step implementation across all frequencies, facilitating the entire process both computationally and conceptually. Even more relevant is the inversion of local dispersion curves, which has proven surprisingly effective: the ability to handle a large number of curves quickly and stably makes this approach particularly useful in situations where large volumes of data need to be analysed or preliminary results obtained in a short time. Overall, the procedure is configured as a solid and flexible tool, capable of providing plausible and interpretable Vs models even under non-ideal operating conditions. For these reasons, the proposed approach appears easily replicable in numerous other contexts, both urban and natural, representing an effective solution for local seismic characterization studies where speed, operational simplicity, and robustness of the results are fundamental requirements.

References

Accaino, F., Bratus, A., Conti, S., Fontana, D., & Tinivella, U. (2007). Fluid seepage in mud volcanoes of the northern Apennines: An integrated geophysical and geological study. Journal of Applied Geophysics, 63(2), 90–101. https://doi.org/10.1016/j.jappgeo.2007.06.002

Achenbach, J. D. (1999). Wave propagation in elastic solids. Elsevier.

Aki, K. (1957). Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors. Bulletin of Earthquake Research Institute, 35, 415–457.

Aki, K. (1965). A note on the use of microseisms in determining the shallow structure of the Earth's crust. Geophysics, 30(4), 665–666. https://doi.org/10.1190/1.1439640

Aki, K., & Lee, W. H. K. (1976). Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model. Journal of Geophysical Research, 81(23), 4381–4399. https://doi.org/10.1029/JB081i023p04381

Aki, K., & Richards, P. G. (1980). Quantitative Seismology: I and II (University Science Books, Ed.).

Albarello, D., Palo, M., & Martinelli, G. (2012). Monitoring methane emission of mud volcanoes by seismic tremor measurements: A pilot study. Natural Hazards and Earth System Science, 12(12), 3617–3629. https://doi.org/10.5194/nhess-12-3617-2012

Antunes, V., Planès, T., Obermann, A., Panzera, F., D'Amico, S., Mazzini, A., Sciarra, A., Ricci, T., & Lupi, M. (2022). Insights into the dynamics of the Nirano Mud Volcano through seismic characterization of drumbeat signals and V/H analysis. Journal of Volcanology and Geothermal Research, 431. https://doi.org/10.1016/j.jvolgeores.2022.107619

Arai, H., & Tokimatsu, K. (1998). Evaluation of local site effects based on microtremor H/V spectra. 2nd International Conference on the Effects of Surface Geology on Seismic Motion, 673–680.

Arai, H., & Tokimatsu, K. (2005). S-wave velocity profiling by joint inversion of microtremor dispersion curve and horizontal-to-vertical (H/V) spectrum. Bulletin of the Seismological Society of America, 95(5), 1766–1778. https://doi.org/10.1785/0120040243

Asten, M. W. (1978). Geological control on the three-component spectra of Rayleigh-wave. Bulletin of the Seismological Society of America, 1623–1636.

Asten, M. W. (1983). Seismic array noise studies at Roosevelt Hot Springs, Utah geothermal area; discussion. Geophysics, 48(11), 1560–1562. https://doi.org/10.1190/1.1441439

Asten, M. W. (2001). The Spatial Auto-Correlation Method for Phase Velocity of Microseisms-Another Method for Characterisation of Sedimentary Overburden. Proceedings of the 2001 Conference of the Australian Earthquake Engineering Society: Earthquake Codes in the Real World.

Asten, M. W. (2003). Lessons from alternative array design used for high-frequency microtremor array studies. Proceedings of the Annual Conference of the Australian Earthquake Engineering Society: Earthquake Risk Mitigation.

Asten, M. W. (2004). Passive seismic methods using the microtremor wave field for engineering and earthquake site zonation.

Asten, M. W. (2006). On bias and noise in passive seismic data from finite circular array data processed using SPAC methods. Geophysics, 71(6). https://doi.org/10.1190/1.2345054

Asten, M. W., & Hayashi, K. (2018). Application of the Spatial Auto-Correlation Method for Shear-Wave Velocity Studies Using Ambient Noise. In Surveys in Geophysics (Vol. 39, Number 4, pp. 633–659). Springer Netherlands. https://doi.org/10.1007/s10712-018-9474-2

Asten, M. W., & Henstridge, J. D. (1984). Array estimators and the use of microseisms for reconnaissance of sedimentary basins. http://library.seg.org/

Aster, R. c., Borchers, B., & Thurber, C. H. (2018). PARAMETER ESTIMATION AND INVERSE PROBLEMS (Third). Elsevier.

Bard, P.-Y. (1998). Microtremor measurements: A tool for site effect estimation? Proceedings of the 2nd International Symposium on the Effects of Surface Geology on Seismic Motion, 1251–1279.

Barone, I., Boaga, J., Carrera, A., Flores-Orozco, A., & Cassiani, G. (2021). Tackling Lateral Variability Using Surface Waves: A Tomography-Like Approach. In Surveys in Geophysics (Vol. 42, Number 2, pp. 317–338). Springer Science and Business Media B.V. https://doi.org/10.1007/s10712-021-09631-x

Beaty, K. S., Schmitt, D. R., & Sacchi, M. (2002). Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure. Geophysical Journal International, 151(2), 622–631. https://doi.org/10.1046/j.1365-246X.2002.01809.x

Bendat, J. S., & Piersol, A. G. (2010). Random Data: Analysis and Measurement Procedures (4th ed.). John Wiley & Sons.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., & Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. Geophysical Journal International, 169(3), 1239–1260. https://doi.org/10.1111/j.1365-246X.2007.03374.x

Bergamo, P., Boiero, D., & Socco, L. V. (2012). Retrieving 2D structures from surface-wave data by means of space-varying spatial windowing. Geophysics, 77(4), EN39–EN51. https://doi.org/10.1190/geo2012-0031.1

Bertelli, T. (1872). Osservazioni sui piccoli movimenti dei pendoli in relazione ad alcuni fenomeni meteorologici. Bullettino Meteorologico Dell'Osservatorio Del Collegio Romano.

Boiero, D., & Socco, L. V. (2010). Retrieving lateral variations from surface wave dispersion curves. Geophysical Prospecting, 58(6), 977–996. https://doi.org/10.1111/j.1365-2478.2010.00877.x

Bonini, M. (2007). Interrelations of mud volcanism, fluid venting, and thrust-anticline folding: Examples from the external northern Apennines (Emilia-Romagna, Italy). Journal of Geophysical Research: Solid Earth, 112(B8). https://doi.org/10.1029/2006JB004859

Bonini, M. (2008). Elliptical mud volcano caldera as stress indicator in an active compressional setting (Nirano, Pede-Apennine margin, northern Italy). Geology, 36(2), 131–134. https://doi.org/10.1130/G24158A.1

Bonini, M. (2009). Mud volcano eruptions and earthquakes in the Northern Apennines and Sicily, Italy. Tectonophysics, 474(3–4), 723–735. https://doi.org/10.1016/j.tecto.2009.05.018

Bonnefoy-Claudet, S., Cornou, C., Bard, P.-Y., & Cotton, F. (2004). Nature of noise wavefield, SESAME report.

Bonnefoy-Claudet, S., Cotton, F., & Bard, P. Y. (2006). The nature of noise wavefield and its applications for site effects studies. A literature review. Earth-Science Reviews, 79(3–4), 205–227. https://doi.org/10.1016/j.earscirev.2006.07.004

Boschi, L., & Ekström, G. (2002). New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies. Journal of Geophysical Research: Solid Earth, 107(4). https://doi.org/10.1029/2000jb000059

Boschi, L., Weemstra, C., Verbeke, J., Ekström, G., Zunino, A., & Giardini, D. (2013). On measuring surface wave phase velocity from station-station cross-correlation of ambient signal. Geophysical Journal International, 192(1), 346–358. https://doi.org/10.1093/gji/ggs023

Boxberger, T., Picozzi, M., & Parolai, S. (2011). Shallow geology characterization using Rayleigh and Love wave dispersion curves derived from seismic noise array measurements. Journal of Applied Geophysics, 75(2), 345–354. https://doi.org/10.1016/j.jappgeo.2011.06.032

Brillouin, L. (1960). WAVE ·PROPAGATION. AND GROUP VELOCITY ACADEMIC PRESS New York and London. Academic Press.

Brindisi, A., Carfagna, N., Paolucci, E., Salleolini, M., & Albarello, D. (2024). The fine structure of seismic emissions from the Nirano mud volcanoes (northern Apennines, Italy): a phenomenological study. Bulletin of Geophysics and Oceanography, 65(2), 165–176. https://doi.org/10.4430/bgo00437

Brune, J. N., Benioff, H., & Ewing, M. (1961). Long-Period Surface Waves from the Chilean Earthquake of May 22, 1960, Recorded on Linear Strain Seismographs. In JOURNAL OF GEOPIIYSICAL RESEARCH (Vol. 66, Number 9).

Buchen, P. W., & Ben-Hador, R. (1996). Free-mode surface-wave computations. Geophysical Journal International, 124(3), 869–887. https://doi.org/10.1111/j.1365-246X.1996.tb05642.x

Campillo, M. (2006). Phase and Correlation in "Random" Seismic Fields and the Reconstruction of the Green Function. Pure and Applied Geophysics, 163, 475–502.

Campillo, M., & Paul, A. (2003). Long range correlations in the diffuse seismic coda. Science, 299(5606), 547–549. https://doi.org/10.1126/science.1078551

Capon, J. (1969). High-Resolution Frequency-Wavenumber Spectrum Analysis. IEEE, 1408–1419.

Capozzi, R., & Picotti, V. (2010). Spontaneous fluid emissions in the Northern Apennines: Geochemistry, structures and implications for the petroleum system. In G. P. Geoffrey, J. Craig, T. 4 Needham, & R. Scott (Eds.), Hydrocarbons in Contractional Belts (Vol. 348, pp. 115–135). Geological Society of London.

Carfagna, N., Brindisi, A., Paolucci, E., & Albarello, D. (2024). Seismic monitoring of gas emissions at mud volcanoes: The case of Nirano (northern Italy). Journal of Volcanology and Geothermal Research, 446. https://doi.org/10.1016/j.jvolgeores.2023.107993

Cho, I., Tada, T., & Shinozaki, Y. (2004). A new method to determine phase velocities of Rayleigh waves from microseisms. Geophysics, 69(6), 1535–1551. https://doi.org/10.1190/1.1836827

Claprood, M., Asten, M. W., & Kristek, J. (2011). Using the SPAC microtremor method to identify 2D effects and evaluate 1D shear-wave velocity profile in Valleys. Bulletin of the Seismological Society of America, 101(2), 826–847. https://doi.org/10.1785/0120090232

Cornou, C. (2002). Traitement d'antenne et imagerie sismique dans l'agglomération grenobloise. Implications pour le risque sismique [PhD thesis]. Université Joseph Fourier.

Cox, H. (1973). Spatial correlation in arbitrary noise fields with application to ambient sea noise. The Journal of the Acoustical Society of America, 55(4), 1289–1301. https://doi.org/10.1121/1.1914426

Dimitrov, L. I. (2002). Mud volcanoes—the most important pathway for degassing deeply buried sediments. Earth-Science Reviews, 59(1–4), 49–76. https://doi.org/10.1016/S0012-8252(02)00069-7

Dunkin, J. W. (1965). Computation of modal solutions in layered, elastic media at high frequencies. Bulletin of the Seismological Society of America, 55(2), 335–358. https://doi.org/10.1785/BSSA0550020335

Duvall, T. L., Jeffferies, S. M., Harvey, J. W., & Pomerantz, M. A. (1993). Time–distance helioseismology. Nature, 362(6419), 430–432. https://doi.org/10.1038/362430a0

Dziewonski, A., Bloch, S., & Landisman, M. (1969). A technique for the analysis of transient seismic signals. Bulletin of the Seismological Society of America, 59(1), 427–444. https://doi.org/10.1785/BSSA0590010427

Dziewonski, A. M. (1984). Mapping the lower mantle: determination of lateral heterogeneity in P velocity up to degree and order 6. Journal of Geophysical Research, 89(B7), 5929–5952. https://doi.org/10.1029/JB089iB07p05929

Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model *. In Preliminary reference Earth modeL Phys. Earth Planet. Inter (Vol. 25).

Ekström, G. (2014). Love and Rayleigh phase-velocity maps, 5-40 s, of the western and central USA from USArray data. Earth and Planetary Science Letters, 402(C), 42–49. https://doi.org/10.1016/j.epsl.2013.11.022

Ekström, G., Abers, G. A., & Webb, S. C. (2009). Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation. Geophysical Research Letters, 36(18). https://doi.org/10.1029/2009GL039131

Fichtner, A. (2021). Lecture Notes on Inverse Theory.

Foti, S., Hollender, F., Garofalo, F., Albarello, D., Asten, M., Bard, P. Y., Comina, C., Cornou, C., Cox, B., Di Giulio, G., Forbriger, T., Hayashi, K., Lunedei, E., Martin, A., Mercerat, D., Ohrnberger, M., Poggi, V., Renalier, F., Sicilia, D., & Socco, V. (2018). Guidelines for the good practice of surface wave analysis: a product of the InterPACIFIC project. Bulletin of Earthquake Engineering, 16(6), 2367–2420. https://doi.org/10.1007/s10518-017-0206-7

Foti, S., Lai, C. G., Rix, G. J., & Strobbia, C. (2014). Surface Wave Methods for Near-Surface Site Characterization.

Foti, S., Parolai, S., Albarello, D., & Picozzi, M. (2011). Application of Surface-wave methods for seismic site characterization. Surveys in Geophysics, 32(6), 777–825. https://doi.org/10.1007/s10712-011-9134-2

Frantti, G. (1963). The nature of high-frequency earth noise spectra. Geophysics, 28(4), 547–562.

Frantti, G., Willis, D. E., & Wilson, J. T. (1962). The spectrum of seismic noise. Bulletin of the Seismological Society of America, 52(1), 113–121.

Friederich, W., Wielandt, E., & Stange, S. (1993). Multiple forward scattering of surface waves: comparison with an exact solution. Geophysical Journal International, 112(2), 264–275.

García-Jerez, A., Piña-Flores, J., Sánchez-Sesma, F. J., Luzón, F., & Perton, M. (2016). A computer code for forward calculation and inversion of the H/V spectral ratio under the diffuse field assumption. Computers & Geosciences, 97, 67–78. https://doi.org/10.1016/j.cageo.2016.06.016

Garofalo, F., Foti, S., Hollender, F., Bard, P. Y., Cornou, C., Cox, B. R., Ohrnberger, M., Sicilia, D., Asten, M., Di Giulio, G., Forbriger, T., Guillier, B., Hayashi, K., Martin, A., Matsushima, S., Mercerat, D., Poggi, V., & Yamanaka, H. (2016). InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part I: Intra-comparison of surface wave methods. Soil Dynamics and Earthquake Engineering, 82, 222–240. https://doi.org/10.1016/j.soildyn.2015.12.010

Gasperi, G., Bettelli, G., Panini, F., Pizziolo, M., Bonazzi, U., Fioroni, C., Fregni, P., & Vaiani, S. C. (2005). Note illustrative della Carta Geologica d'Italia alla scala 1:50.000, Foglio 219 – Sassuolo.

Giambastiani, B. M. S., Antonellini, M., Nespoli, M., Bacchetti, M., Calafato, A., Conventi, M., Dadomo, A., Martinelli, G., Morena, M., Venturoli, S., & Piombo, A. (2022). Mud flow dynamics at gas seeps - Nirano Salse, Italy -. Environmental Earth Sciences, 81(19), 480. https://doi.org/10.1007/s12665-022-10615-2

Gilbert, F., & Backus, G. E. (1966). Propagator matrices in elastic wave and vibration problems. Geophysics, 31, 326–332.

Giovani, L., Vallocchia, M., Antunes, V., Lupi, M., Obermann, A., Mazzini, A., Sciarra, A., Ricci, T., & Moretti, M. (2017). Esperimento di sismica passiva per lo studio di dettaglio dei vulcani di fango nella Riserva Naturale Regionale delle Salse di Nirano (Modena). https://doi.org/10.13140/RG.2.2.20068.48008

Gouédard, P., Stehly, L., Brenguier, F., Campillo, M., Colin De Verdière, Y., Larose, E., Margerin, L., Roux, P., Sánchez-sesma, F. J., Shapiro, N. M., & Weaver, R. L. (2008). Cross-correlation of random fields: Mathematical approach and applications. Geophysical Prospecting, 56(3), 375–393. https://doi.org/10.1111/j.1365-2478.2007.00684.x

Gutenberg, B. (1911). Über die periodischen Bodenbewegungen. Gerlands Beiträge Zur Geophysik, 11.

Gutenberg, B. (1958). MICROSEISMS.

Hadamard, J. (1923). Lectures on Cauchy's Problem in Linear Partial Differential Equations. Yale University Press.

Haskell, N. A. (1953). The dispersion of surface waves on multilayered media*. Bulletin of the Seismological Society of America, 43(1), 17–34. https://doi.org/10.1785/BSSA0430010017

Hayashi, K., & Suzuki, H. (2004). CMP cross-correlation analysis of multi-channel surface-wave data. Exploration Geophysics, 35(1), 7–13. https://doi.org/10.1071/EG04007

Herak, M. (2008). ModelHVSR-A Matlab® tool to model horizontal-to-vertical spectral ratio of ambient noise. Computers and Geosciences, 34(11), 1514–1526. https://doi.org/10.1016/j.cageo.2007.07.009

Herrmann, R. B. (2013). Computer Programs in Seismology: An Evolving Tool for Instruction and Research. Seismological Research Letters, 84(6), 1081–1088. https://doi.org/10.1785/0220110096

Hestenes, M. R., & Stiefel, E. (1952). Methods of conjugate gradient for solving linear systems. Journal of Research of the National Bureau of Standards, 49, 409–435.

Horike, M. (1985). INVERSION OF PHASE VELOCITY OF LONG-PERIOD MICROTREMORS TO THE S-WAVE-VELOCITY STRUCTURE DOWN TO THE BASEMENT IN URBANIZED AREAS. In J. Phys. Earth (Vol. 33).

Ikeda, T., Tsuji, T., Konishi, C., & Saito, H. (2021). Spatial autocorrelation method for reliable measurements of two-station dispersion curves in heterogeneous ambient noise wavefields. Geophysical Journal International, 226(2), 1130–1147. https://doi.org/10.1093/gji/ggab150

Iverson, R. M., Dzurisin, D., Gardner, C. A., Gerlach, T. M., LaHusen, R. G., Lisowski, M., Major, J. J., Malone, S. D., Messerich, J. A., Moran, S. C., Pallister, J. S., Qamar, A. I., Schilling, S. P., & Vallance, J. W. (2006). Dynamics of seismogenic volcanic extrusion at Mount St Helens in 2004-05. Nature, 444(7118), 439–443. https://doi.org/10.1038/nature05322

Jackson, D. D. (1979). The use of a priori data to resolve nonuniqueness in linear inversion. Geophysical Journal of the Royal Astronomical Society, 57, 137–157.

Judd, A. (2005). GAS EMISSIONS FROM MUD VOLCANOES Significance to Global Climate Change.

Kästle, E. D., Soomro, R., Weemstra, C., Boschi, L., & Meier, T. (2016). Two-receiver measurements of phase velocity: Cross-validation of ambient-noise and earthquake-based observations. Geophysical Journal International, 207(3), 1493–1512. https://doi.org/10.1093/gji/ggw341

Keilis-Borok, V. I., & Levshin, A. L. (1989). Seismic Surface Waves in a Laterally Inhomogeneous Earth (1st ed.). Kluwer Academic Publishers.

Kendrick, J. E., Lavallée, Y., Hirose, T., Di Toro, G., Hornby, A. J., De Angelis, S., & Dingwell, D. B. (2014). Volcanic drumbeat seismicity caused by stick-slip motion and magmatic frictional melting. Nature Geoscience, 7(6), 438–442. https://doi.org/10.1038/ngeo2146

Kennett, B. L. N. (1974). Reflections, rays, and reverberations. Bulletin of the Seismological Society of America, 64(6), 1685–1696. https://doi.org/10.1785/BSSA0640061685

Kennett, B. L. N., & Kerry, N. J. (1979). Seismic waves in a stratified half space. In Geophys. J. R. astr. Soc (Vol. 57). https://academic.oup.com/gji/article/57/3/557/681409

Kirkpatrick, S., Gelatt, ; C D, & Vecchi, ; M P. (1983). Optimization by Simulated Annealing. In New Series (Vol. 220, Number 4598).

Knopoff, L. (1964). A matrix method for elastic wave problems. Bulletin of the Seismological Society of America, 54(1), 431–438. https://doi.org/10.1785/BSSA0540010431

Komatitsch, D., & Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation. Geophysical Journal International, 139(3), 806–822. https://doi.org/10.1046/j.1365-246x.1999.00967.x

Konno, K., & Ohmachi, T. (1998). Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. Bulletin of the Seismological Society of America, 88(1), 228–241. https://doi.org/10.1785/BSSA0880010228

Kopf, A. J. (2002). Significance of mud volcanism. Reviews of Geophysics, 40(2), 2-1-2–52. https://doi.org/10.1029/2000RG000093

La Rocca, M., Galluzzo, D., Nardone, L., Gaudiosi, G., & Di Luccio, F. (2023). Hydrothermal Seismic Tremor in a Wide Frequency Band: The Nonvolcanic CO2 Degassing Site of Mefite d'Ansanto, Italy. Bulletin of the Seismological Society of America, 113(3), 1102–1114. https://doi.org/10.1785/0120220243

Lacoss, R. T., Kelly, E. J., & Toksöz, M. N. (1969). ESTIMATION OF SEISMIC NOISE STRUCTURE USING ARRAYSt. In GEOPHYSICS (Vol. 34, Number 1). http://library.seg.org/

Lai, C. G., & Rix, G. J. (1998). Simultaneous Inversion of Rayleigh Phase Velocity and Attenuation for Near-Surface Site Characterization School of Civil and Environmental Engineering. Georgia Institute of Technology.

Lay, T., & Wallace, T. C. (1995). Modern Global Seismology. Academic Press.

Lee, M. W. (2004). Elastic velocities of partially gas-saturated unconsolidated sediments. Marine and Petroleum Geology, 21(6), 641–650. https://doi.org/10.1016/j.marpetgeo.2004.03.002

Lin, F.-C., Moschetti, M. P., & Ritzwoller, M. H. (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. Geophysical Journal International, 173(1), 281–298. https://doi.org/10.1111/j.1365-246X.2008.03720.x

Lin, F.-C., Ritzwoller, M. H., Townend, J., Bannister, S., & Savage, M. K. (2007). Ambient noise Rayleigh wave tomography of New Zealand. Geophysical Journal International, 170(2), 649–666. https://doi.org/10.1111/j.1365-246X.2007.03414.x

Ling, S., & Okada, H. (1993). An Extended Use of the Spatial Autocorrelation Method for the Estimation of Geological Structure Using Microtremors. Proceedings of the SEGJ Conference, 44–48.

Lobkis, O. I., & Weaver, R. L. (2001). On the emergence of the Green's function in the correlations of a diffuse field. The Journal of the Acoustical Society of America, 110(6), 3011–3017. https://doi.org/10.1121/1.1417528

Lupi, M., Ricci, B. S., Kenkel, J., Ricci, T., Fuchs, F., Miller, S. A., & Kemna, A. (2016). Subsurface fluid distribution and possible seismic precursory signal at the Salse di Nirano mud volcanic field, Italy. Geophysical Journal International, 204(2), 907–917. https://doi.org/10.1093/gji/ggv454

Magrini, F., Diaferia, G., Boschi, L., & Cammarano, F. (2020). Arrival-angle effects on two-receiver measurements of phase velocity. Geophysical Journal International, 220(3), 1838–1844. https://doi.org/10.1093/gji/ggz560

Magrini, F., Lauro, S., Kästle, E., & Boschi, L. (2022). Surface-wave tomography using SeisLib: A Python package for multiscale seismic imaging. Geophysical Journal International, 231(2), 1011–1030. https://doi.org/10.1093/gji/ggac236

Malagnini, L., Rovelli, A., Hough, S. E., & Seeber, L. (1993). SITE AMPLIFICATION ESTIMATES IN THE GARIGLIANO VALLEY, CENTRAL ITALY, BASED ON DENSE ARRAY MEASUREMENTS OF AMBIENT NOISE. In Bulletin of the Seismological Society of America (Vol. 83, Number 6).

Marquering, H., Sniederl, R., & Nolet2, G. (1996). Waveform inversions and the significance of surface-wave mode coupling. In Geophys. J. Int (Vol. 124).

Martinelli, G., & Judd, A. (2004). Mud volcanoes of Italy. Geological Journal, 39(1), 49–61. https://doi.org/10.1002/gj.943

Martinelli, G., & Panahi, B. (2003). Mud Volcanoes, Geodynamics and Seismicity. http://www.nato.int/science

Maupin, V. (1988). Surface waves across 2-D structures: a method based on coupled local modes. Geophysical Journal, 93(1), 173–185. https://doi.org/10.1111/j.1365-246X.1988.tb01396.x

Mazzini, A., & Etiope, G. (2017). Mud volcanism: An updated review. Earth-Science Reviews, 168, 81–112. https://doi.org/10.1016/j.earscirev.2017.03.001

Mazzini, A., Nermoen, A., Krotkiewski, M., Podladchikov, Y., Planke, S., & Svensen, H. (2009). Strike-slip faulting as a trigger mechanism for overpressure release through piercement structures. Implications for the Lusi mud volcano, Indonesia. Marine and Petroleum Geology, 26(9), 1751–1765. https://doi.org/10.1016/j.marpetgeo.2009.03.001

McMechan, G. A., & Yedlin, M. J. (1981). Analysis of dispersive by wave field transformation. Geophysics, 46(6), 869–874. https://doi.org/10.1190/1.1441225

Menke, W. (2012). Geophysical Data Analysis: Discrete Inverse Theory. In Geophysical Data Analysis: Discrete Inverse Theory (Third). Academic Press. https://doi.org/10.1016/b978-0-12-397160-9.00001-1

Milkov, A. V. (2000). Worldwide distribution of submarine mud volcanoes and associated gas hydrates. Marine Geology, 167(1–2), 29–42. https://doi.org/10.1016/S0025-3227(00)00022-0

Mohammadi, N., Gholami, A., Rahimi, H., & Aoudia, A. (2020). Simultaneous tomography of all periods in surface wave analysis. Physics of the Earth and Planetary Interiors, 298. https://doi.org/10.1016/j.pepi.2019.106338

Nakamura, Y. (1989). A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface.

Nakanishi, I., & Anderson, D. L. (1982). WORLDWIDE DISTRIBUTION OF GROUP VELOCITY OF MANTLE RAYLEIGH WAVES AS DETERMINED BY SPHERICAL HARMONIC INVERSION. In Bulletin of the Seismological Society of America (Vol. 72, Number 4).

Nakata, N., Gualtieri, L., & Fichtner, A. (2019). Seismic ambient noise. In Seismic Ambient Noise. Cambridge University Press. https://doi.org/10.1017/9781108264808

Nazarian, S., & Stokoe, K. H. (1984). In situ shear wave velocities from spectral analysis of surface waves. Proceedings of the 8th World Conference on Earthquake Engineering.

Nogoshi, M., & Igarashi, T. (1971). On the Amplitude Characteristics of Microtremor (Part 2). Journal of Seismological Society of Japan, 24, 26–40.

Ohori, M., Nobata, A., & Wakamatsu, K. (2002). A Comparison of ESAC and FK Methods of Estimating Phase Velocity Using Arbitrarily Shaped Microtremor Arrays. In Bulletin of the Seismological Society of America (Vol. 92, Number 6).

Okada, Hiroshi. (2003). The microtremor survey method. Society of Exploration Geophysicists with the cooperation of Society of Exploration Geophysicists of Japan [and] Australian Society of Exploration Geophysicists.

Oppo, D. (2011). Studio dei vulcani di fango per la definizione della migrazione dei fluidi profondi. Università di Bologna.

Paige, C. C., & Saunders, M. A. (1982). LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. ACM Transactions on Mathematical Software, 8, 43–71.

Papadopoulou, M., Anjom, F. K., Karimpour, M. K., & Socco, V. L. (2021). Advances in surface-wave tomography for near-surface applications. Leading Edge, 40(8), 567–575. https://doi.org/10.1190/tle40080567.1

Park, C. B., Miller, R. D., & Xia, J. (1999). Multichannel analysis of surface waves. Geophysics, 64(3), 800–808. https://doi.org/10.1190/1.1444590

Parolai, S., Picozzi, M., Richwalski, S. M., & 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. Geophysical Research Letters, 32(1). https://doi.org/10.1029/2004GL021115

Parolai, S., Richwalski, S. M., Milkereit, C., & Fäh, D. (2006). S-wave velocity profiles for earthquake engineering purposes for the Cologne area (Germany). Bulletin of Earthquake Engineering (Vol. 4, Number 1, pp. 65–94). https://doi.org/10.1007/s10518-005-5758-2

Pestel, Eduard., & Leckie, F. A. (1963). Matrix methods in elastomechanics. McGraw-Hill.

Peterson, J. (1993). Observations and modeling of seismic background noise.

Picozzi, M., & Albarello, D. (2007). Combining genetic and linearized algorithms for a two-step joint inversion of Rayleigh wave dispersion and H/V spectral ratio curves. Geophysical Journal International, 169(1), 189–200. https://doi.org/10.1111/j.1365-246X.2006.03282.x

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. https://doi.org/10.1111/j.1365-246X.2008.03966.x

Picozzi, M., Parolai, S., & Richwalski, S. M. (2005). Joint inversion of H/V ratios and dispersion curves from seismic noise: Estimating the S-wave velocity of bedrock. Geophysical Research Letters, 32(11). https://doi.org/10.1029/2005GL022878

Pilz, M., Parolai, S., Picozzi, M., & Bindi, D. (2012). Three-dimensional shear wave velocity imaging by ambient seismic noise tomography. Geophysical Journal International, 189(1), 501–512. https://doi.org/10.1111/j.1365-246X.2011.05340.x

Poli, P., Boaga, J., Molinari, I., Cascone, V., & Boschi, L. (2020). The 2020 coronavirus lockdown and seismic monitoring of anthropic activities in Northern Italy. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-66368-0

Priestley, M. B. (1981). Spectral Analysis and Time Series. Academic Press.

Pujol, J., & Smalley, R. (1990). A PRELIMINARY EARTHQUAKE LOCATION METHOD BASED ON A HYPERBOLIC APPROXIMATION TO TRAVEL TIMES. In Bulletin of the Seismological Society of America (Vol. 80, Number 6).

Ritzwoller, M. H., Lin, F.-C., & Shen, W. (2011). Ambient noise tomography with a large seismic array. Comptes Rendus. Géoscience, 343(8–9), 558–570. https://doi.org/10.1016/j.crte.2011.03.007

Rothman, D. H. (1985). Nonlinear inversion, statistical mechanics, and residual statics estimation. Geophysics, 50(12), 2784–2796. https://doi.org/10.1190/1.1441899

Rothman, D. H. (1986). Automatic estimation of large residual statics corrections. Geophysics, 51(2), 332–346. https://doi.org/10.1190/1.1442092

Roux, P., Sabra, K. G., Kuperman, W. A., & Roux, A. (2005). Ambient noise cross correlation in free space: Theoretical approach. The Journal of the Acoustical Society of America, 117(1), 79–84. https://doi.org/10.1121/1.1830673

Sager, K., Ermert, L., Boehm, C., & Fichtner, A. (2018). Towards full waveform ambient noise inversion. Geophysical Journal International, 212(1), 566–590. https://doi.org/10.1093/gji/ggx429

Sambridge, M. (1999a). Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space. Geophysical Journal International, 138(2), 479–494. https://doi.org/10.1046/j.1365-246X.1999.00876.x

Sambridge, M. (1999b). Geophysical inversion with a neighbourhood algorithm—II. Appraising the ensemble. Geophysical Journal International, 138(3), 727–746. https://doi.org/10.1046/j.1365-246x.1999.00900.x

Sambridge, M., & Drijkoningen, G. (1992). Genetic algorithms in seismic waveform inversion. Geophysical Journal International, 109(2), 323–342. https://doi.org/10.1111/j.1365-246X.1992.tb00100.x

Sambridge, M., & Mosegaard, K. (2002). MONTE CARLO METHODS IN GEOPHYSICAL INVERSE PROBLEMS. Rev. Geophys, 40(3), 1009. https://doi.org/10.1029/2000RG00089

Sciarra, A., Cantucci, B., Ricci, T., Tomonaga, Y., & Mazzini, A. (2019). Geochemical characterization of the Nirano mud volcano, Italy. Applied Geochemistry, 102, 77–87. https://doi.org/10.1016/j.apgeochem.2019.01.006

Sen, M. K., & Stoffa, P. L. (1995). Global Optimization Methods in Geophysical Inversion. Elsevier. https://doi.org/10.1017/CBO9780511997570

Sethian, J. A. (1996). A fast-marching level set method for monotonically advancing fronts. Proceedings of the National Academy of Sciences, 93(4), 1591–1595. https://doi.org/10.1073/pnas.93.4.1591

Shapiro, N. M., & Campillo, M. (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. Geophysical Research Letters, 31(7). https://doi.org/10.1029/2004GL019491

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H. (2005). High-Resolution Surface-Wave Tomography from Ambient Seismic Noise. SCIENCE, 307, 1615–1618.

Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, 69(4), 8. https://doi.org/10.1103/PhysRevE.69.046610

Socco*, L. V., Boiero, D., Bergamo, P., Garofalo, F., Yao, H., Hilst, R. D. van der, & Col, F. Da. (2014). Surface wave tomography to retrieve near surface velocity models. SEG Technical Program Expanded Abstracts 2014, 2013–2018. https://doi.org/10.1190/segam2014-1278.1

Socco, L. V., & Comina, C. (2017). Time-average velocity estimation through surface-wave analysis: Part 2 - P-wave velocity. Geophysics, 82(3), U61–U73. https://doi.org/10.1190/GEO2016-0368.1

Socco, L. V., Comina, C., & Anjom, F. K. (2017). Time-average velocity estimation through surface-wave analysis: Part 1-s-wave velocity. Geophysics, 82(3), U49–U59. https://doi.org/10.1190/GEO2016-0367.1

Socco, L. V., Foti, S., & Boiero, D. (2010). Surface-wave analysis for building near-surface velocity models - Established approaches and new perspectives. Geophysics, 75(5). https://doi.org/10.1190/1.3479491

Sorrells, G. G., McDonald, J. A., Der, Z. A., & Herrin, E. (1971). Earth Motion Caused by Local Atmospheric Pressure Changes. Geophysical Journal of the Royal Astronomical Society, 26(1–4), 83–98. https://doi.org/10.1111/j.1365-246X.1971.tb03384.x

Stokoe, K. H. I., Wright, S. G., Bay, J., & Roesset, J. (1994). Characterization of geotechnical sites by SASW method.

Szanyi, G., Gráczer, Z., Győri, E., Kaláb, Z., & Lednická, M. (2016). Ambient Seismic Noise Tomography of a Loess High Bank at Dunaszekcső (Hungary). Pure and Applied Geophysics, 173(8), 2913–2928. https://doi.org/10.1007/s00024-016-1304-1

Takeuchi, H., & Saito, M. (1972). Seismic Surface Waves (pp. 217–295). https://doi.org/10.1016/B978-0-12-460811-5.50010-6

Tarantola, A. (2005). Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM.

Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. Reviews of Geophysics, 20(2), 219–232. https://doi.org/10.1029/RG020i002p00219

Thomson, W. T. (1950). Transmission of elastic waves through a stratified solid medium. Journal of Applied Physics, 21(2), 89–93. https://doi.org/10.1063/1.1699629

Thrasher, J., Fleet, A. J., Hay, S. J., Hovland, M., & Düppenbecker, S. (1996). Understanding geology as the key to using seepage in exploration: the spectrum of seepage styles. In Hydrocarbon Migration and Its Near-Surface Expression. American Association of Petroleum Geologists. https://doi.org/10.1306/m66606c17

Tokimatsu, K. (1995). Geotechnical site characterisation using surface waves. 1st International Conference on Earthquake Geotechnical Engineering (IS–Tokyo '95).

Tokimatsu, K., Tamura, S., & Kojima, H. (1992). Effects of Multiple Modes on Rayleigh Wave Dispersion Characteristics. Journal of Geotechnical Engineering, 118(10), 1529–1543. https://doi.org/10.1061/(ASCE)0733-9410(1992)118:10(1529)

Toksöz, M. N., & Anderson, D. L. (1966). Phase velocities of long-period surface waves and structure of the upper mantle: 1. Great-Circle Love and Rayleigh wave data. Journal of Geophysical Research, 71(6), 1649–1658. https://doi.org/10.1029/JZ071i006p01649

Toksöz, M. N., & Lacoss, R. T. (1968). Microseisms: Mode structure and excitation mechanisms. Geophysical Journal of the Royal Astronomical Society, 16(1), 1–13.

Tromp, J., Luo, Y., Hanasoge, S., & Peter, D. (2010). Noise cross-correlation sensitivity kernels. Geophysical Journal International, 183(2), 791–819. https://doi.org/10.1111/j.1365-246X.2010.04721.x

Tsai, V. C. (2009). On establishing the accuracy of noise tomography travel-time measurements in a realistic medium. Geophysical Journal International, 178(3), 1555–1564. https://doi.org/10.1111/j.1365-246X.2009.04239.x

Tsai, V. C. (2011). Understanding the amplitudes of noise correlation measurements. Journal of Geophysical Research, 116(B9), B09311. https://doi.org/10.1029/2011JB008483

Tsai, V. C., & Moschetti, M. P. (2010). An explicit relationship between time-domain noise correlation and spatial autocorrelation (SPAC) results. Geophysical Journal International, 182(1), 454–460. https://doi.org/10.1111/j.1365-246X.2010.04633.x

Verbeke, J., Boschi, L., Stehly, L., Kissling, E., & Michelini, A. (2012). High-resolution Rayleigh-wave velocity maps of central Europe from a dense ambient-noise data set. Geophysical Journal International, 188(3), 1173–1187. https://doi.org/10.1111/j.1365-246X.2011.05308.x

Vidale, J. (1988). FINITE-DIFFERENCE CALCULATION OF TRAVEL TIMES. In Bulletin of the Seismological Society of America (Vol. 78, Number 6).

Vidale, J. E. (1990). Finite-difference calculation of traveltimes in three dimensions. Geophysics, 55(5), 521–526.

Wathelet, M., Chatelain, J.-L., Cornou, C., Giulio, G. Di, Guillier, B., Ohrnberger, M., & Savvaidis, A. (2020). Geopsy: A User-Friendly Open-Source Tool Set for Ambient Vibration Processing. Seismological Research Letters, 91(3), 1878–1889. https://doi.org/10.1785/0220190360

Wathelet, M., Jongmans, D., & Ohrnberger, M. (2005). Direct inversion of spatial autocorrelation curves with the neighbourhood algorithm. Bulletin of the Seismological Society of America, 95(5), 1787–1800. https://doi.org/10.1785/0120040220

Woodhouse, J. H. (1974). Surface Waves in a Laterally Varying Layered Structure. Geophysical Journal of the Royal Astronomical Society, 37(3), 461–490. https://doi.org/10.1111/j.1365-246X.1974.tb04098.x

Woodhouse, J. H., & Dziewonski, A. M. (1984). Mapping the upper mantle: three-dimensional modeling of Earth structure by inversion of seismic waveforms. Journal of Geophysical Research, 89(B7), 5953–5986. https://doi.org/10.1029/JB089iB07p05953

Yamamoto, H. (2000). "Estimation of shallow S-wave velocity structure from phase velocities of Love- and Rayleigh waves in microtremors. 12th World Conference of Earthquake Engineering.

Yamanaka, H., Dravinski, M., & Kagami, H. (1993). Continuous measurements of microtremors on sediments and basement in Los Angeles, California. Bulletin of the Seismological Society of America, 83(5), 1595–1609. https://doi.org/10.1785/BSSA0830051595

Yao, H., van der Hilst, R. D., & de Hoop, M. V. (2006). Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - I. Phase velocity maps. Geophysical Journal International, 166(2), 732–744. https://doi.org/10.1111/j.1365-246X.2006.03028.x

Appendix A

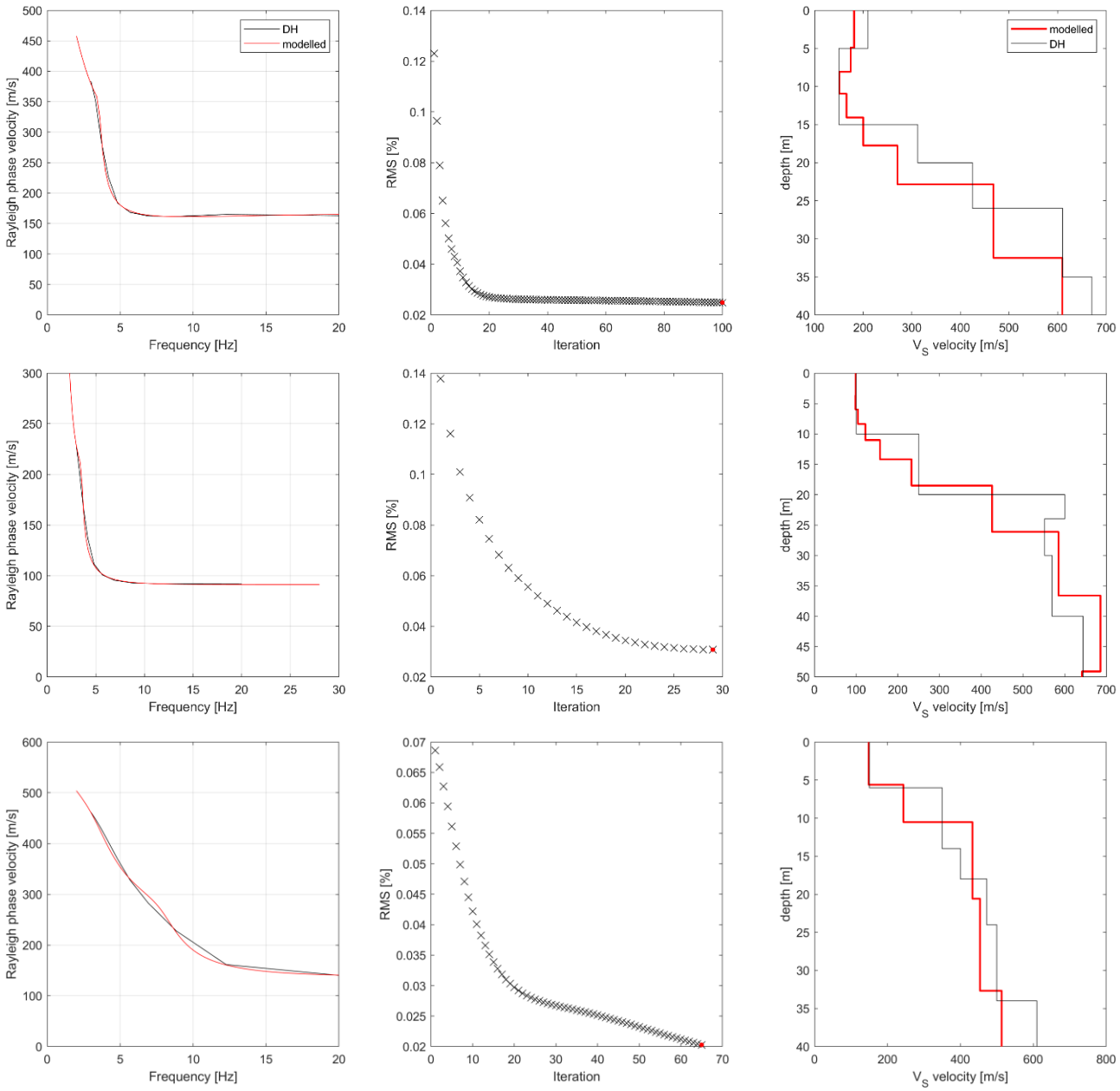

Figure A1 - Validation tests of the inversion code using three examples based on Vs profiles derived from downhole tests. For each panel, the following are shown: (left) the comparison between phase dispersion curves from the downhole profile (black) and modelled (red); (centre) the trend of the RMS (%) as a function of iterations; (right) the comparison between the target Vs profile and the profile obtained from the inversion. The inversion satisfactorily reproduces the general trend and vertical behaviour of the Vs in the surface levels, while at the greater depths differences are observed attributable to the reduction in vertical resolution due to the progressive increase in layer thicknesses of the adopted model.

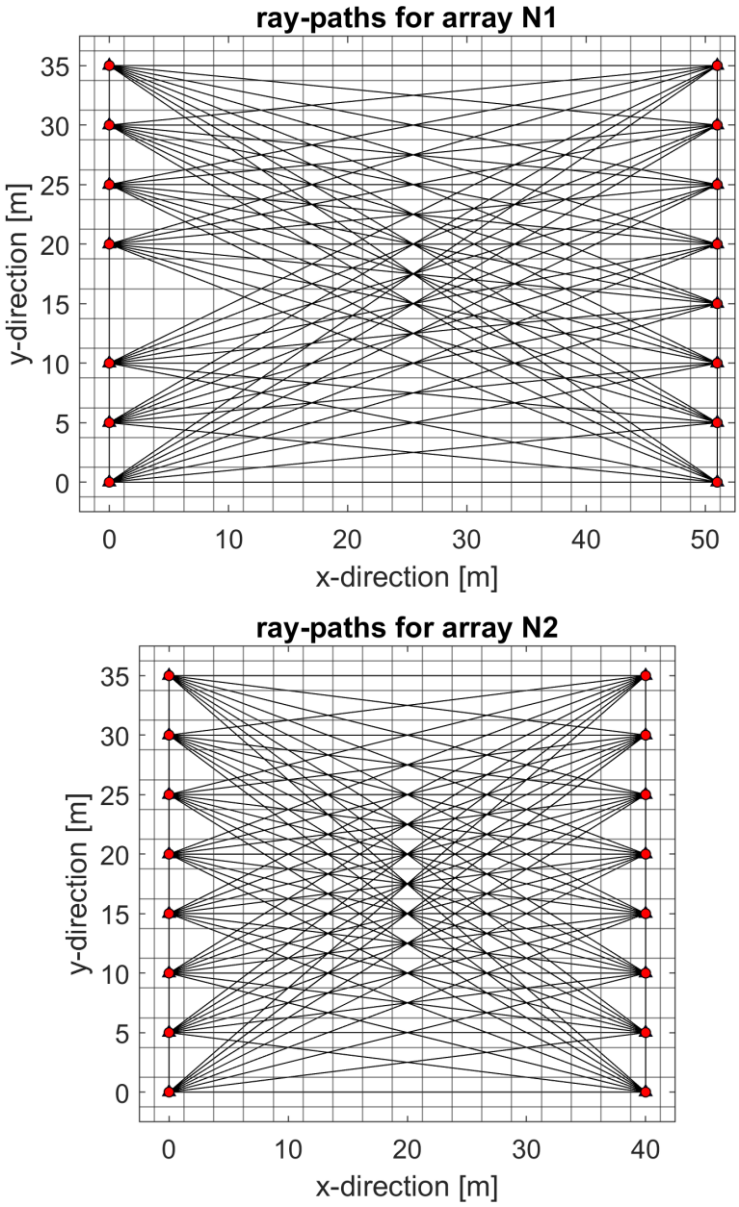

Figure A2 - Representation of ray paths between station pairs from arrays N1 (top) and N2 (bottom). The lines indicate the considered trajectories at a representative frequency, while the red triangles show the position of the geophones. Although ray paths are associated with a specific frequency, they can be considered indicative of the general spatial coverage of the arrays over the frequency range.

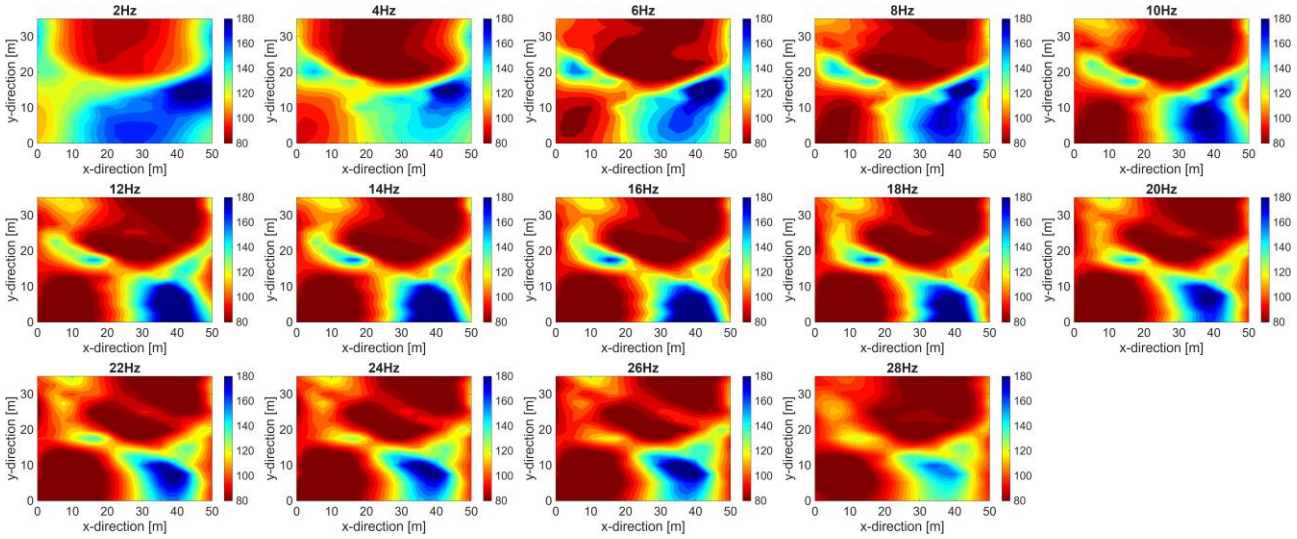

Figure A3 - Phase velocity maps of Rayleigh waves obtained by inversion for the N1 array, in the frequency range considered (starting at 2 Hz, with discretization at 2 Hz intervals). The maps highlight the spatial distribution of lateral variations in phase velocity at the different analysed frequencies and provide a picture of the dispersive response of the subsurface to the different depths investigated.

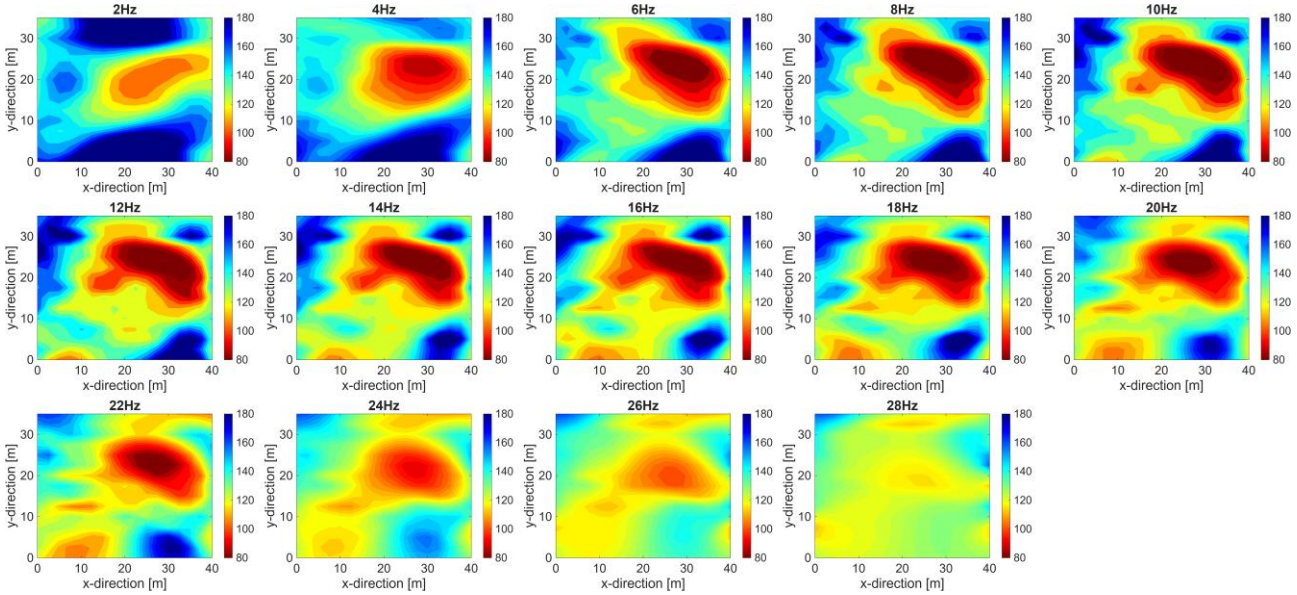

Figure A4 - Phase velocity maps of Rayleigh waves obtained by inversion for the N2 array, in the same frequency range (from 2 Hz with a step of 2 Hz). The maps show the spatial evolution of the phase velocity as function of the frequency, allowing to highlight main lateral heterogeneities within the investigated domain.

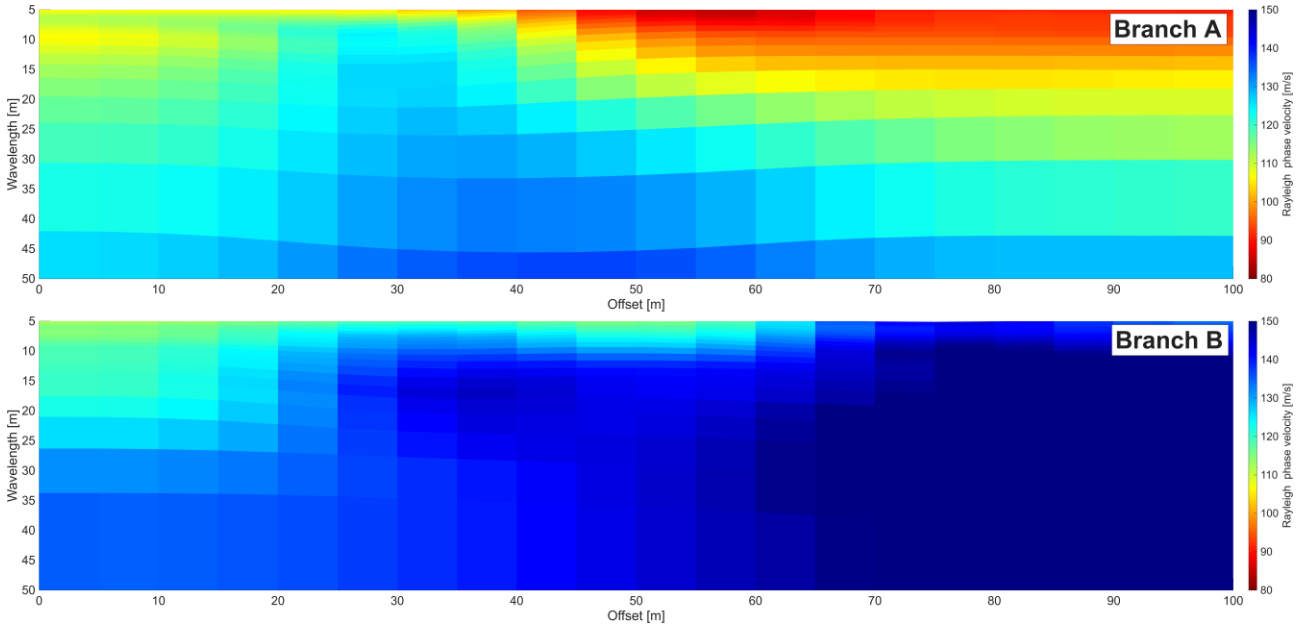

Figure A5 - Phase velocity sections as a function of the considered wavelength along the offset for branches A (top) and B (bottom) of array N3. The colours represent the Rayleigh phase velocities, evidencing the differences between the two branches.

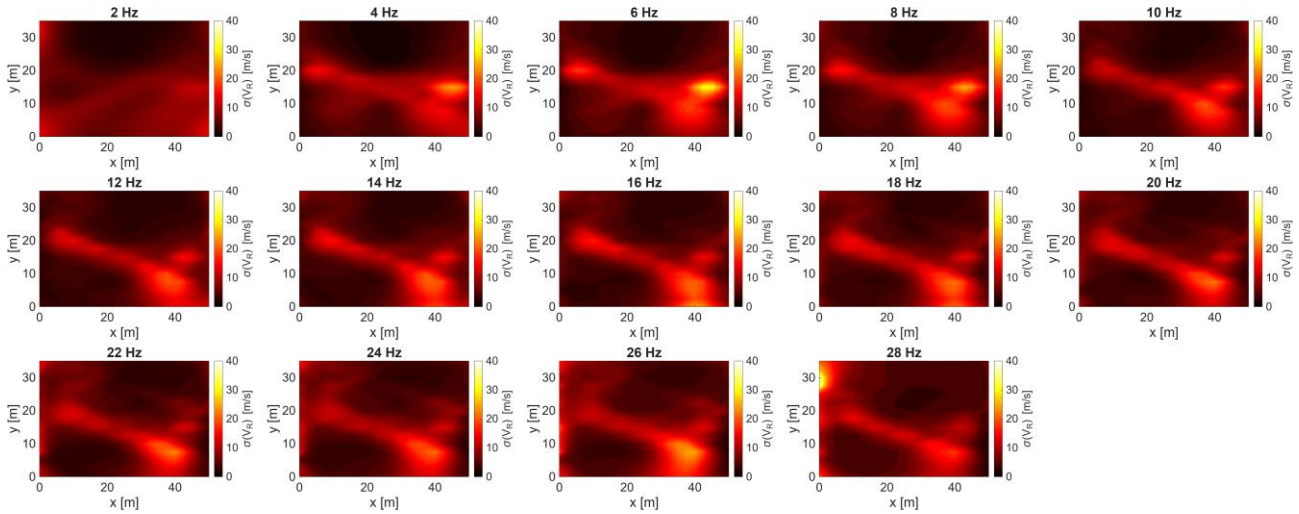

Figure A6 - Bootstrap standard deviation maps for array N1, shown for frequencies ranging from 2 to 28 Hz every 2 Hz. Standard deviation values remain below 40 m/s across the entire domain, with a zone of relatively higher uncertainty (~20% of the estimated velocity) persisting across the frequency range along a geometry consistent with the ray paths between specific station pairs.

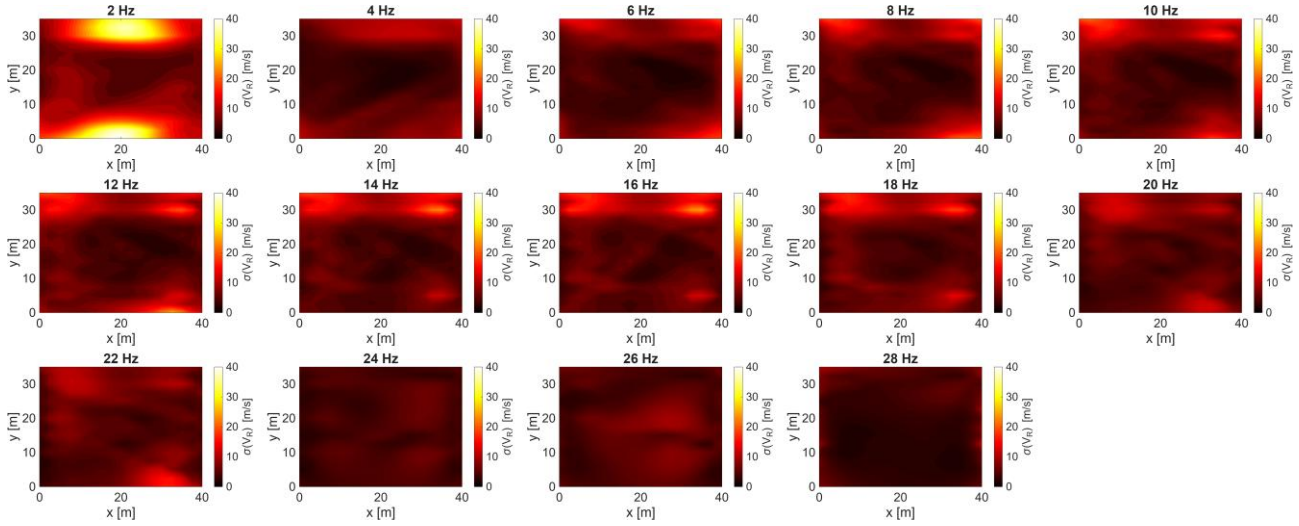

Figure A7 - Bootstrap standard deviation maps for array N2, shown for frequencies ranging from 2 to 28 Hz every 2 Hz. Standard deviation values remain within 10 m/s in the central domain (~5% of the estimated velocity) and below 20 m/s in the peripheral areas (~10%). An exception is observed at 2 Hz, where σ locally reaches ~40 m/s along the upper and lower edges of the domain.

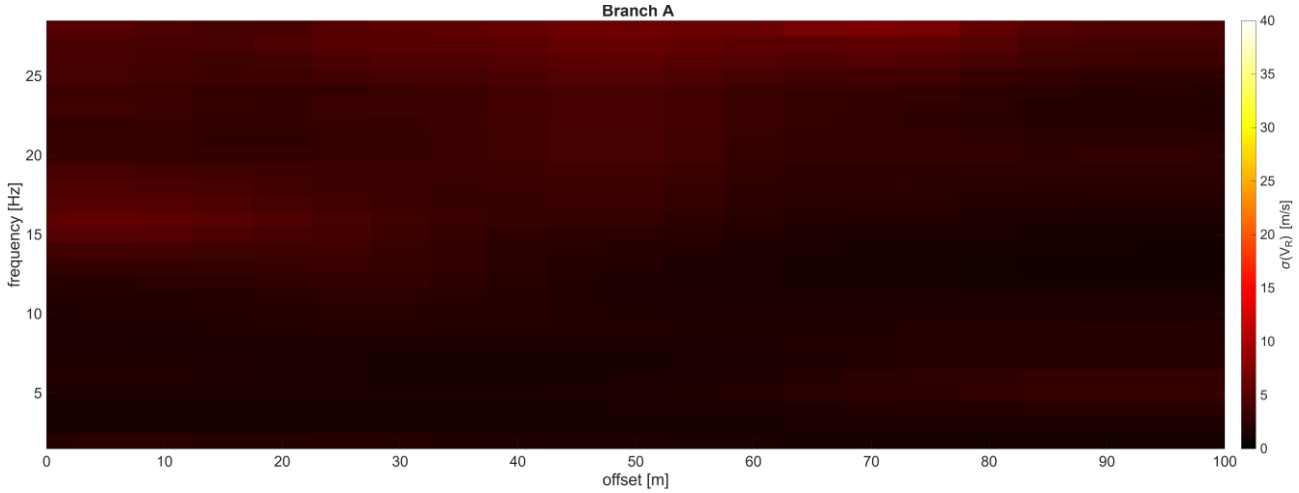

Figure A8 - Bootstrap standard deviation section in the offset-frequency domain for array N3, Branch A. Standard deviation values almost always remain below 10 m/s across the entire domain, with slightly elevated uncertainty observed at high frequencies in the proximal part of the profile.

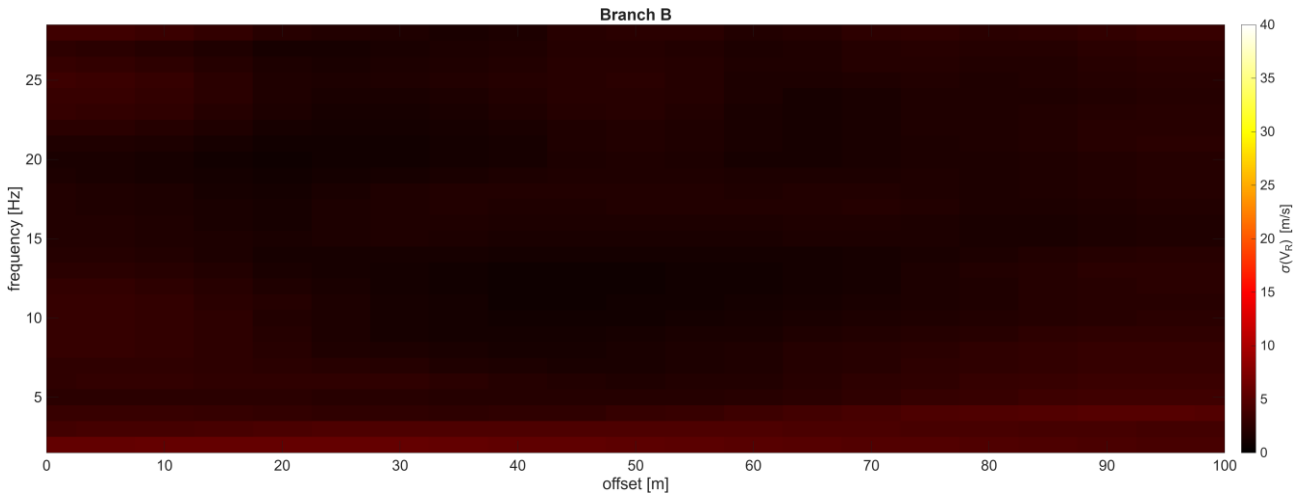

Figure A8 - Bootstrap standard deviation section in the offset-frequency domain for array N3, Branch B. Standard deviation values almost always remain below 10 m/s across the entire domain, with slightly elevated uncertainty observed at low frequencies in the proximal part of the profile.

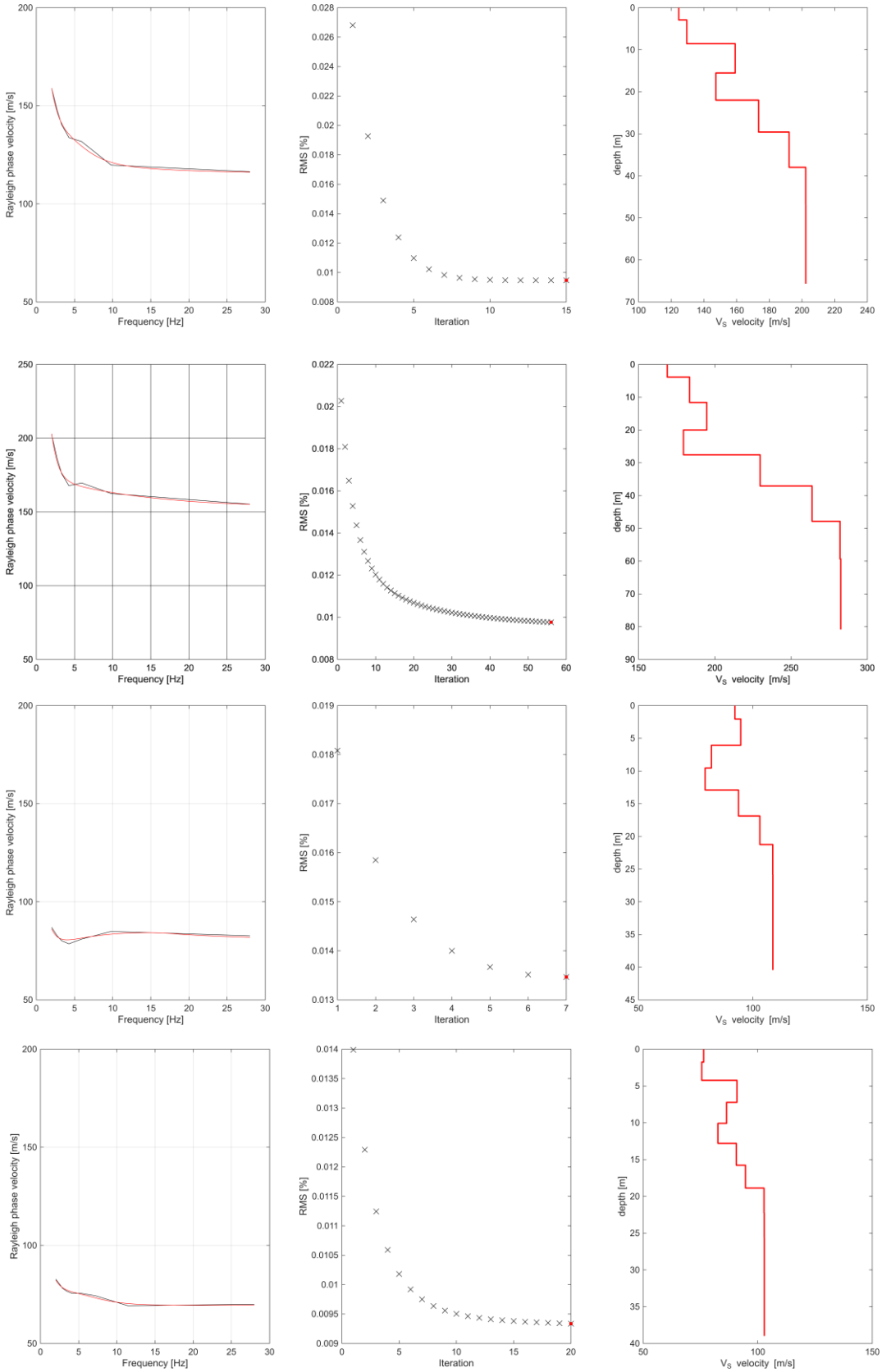

Figure A10 - Example of multiple local dispersion curve inversions from array N1. On the left, each panel shows a comparison between the phase velocity curve obtained from tomographic inversion (black) and the modelled one (red); in the centre there is the trend of the RMS error (%) as a function of the iterations to highlight the convergence towards a

minimum; on the right the shear velocity profile (Vs) obtained from the inversion is illustrated. The result shows a good ability of the model to reproduce the observed data, with stable convergence and satisfactory fit of the experimental curve.

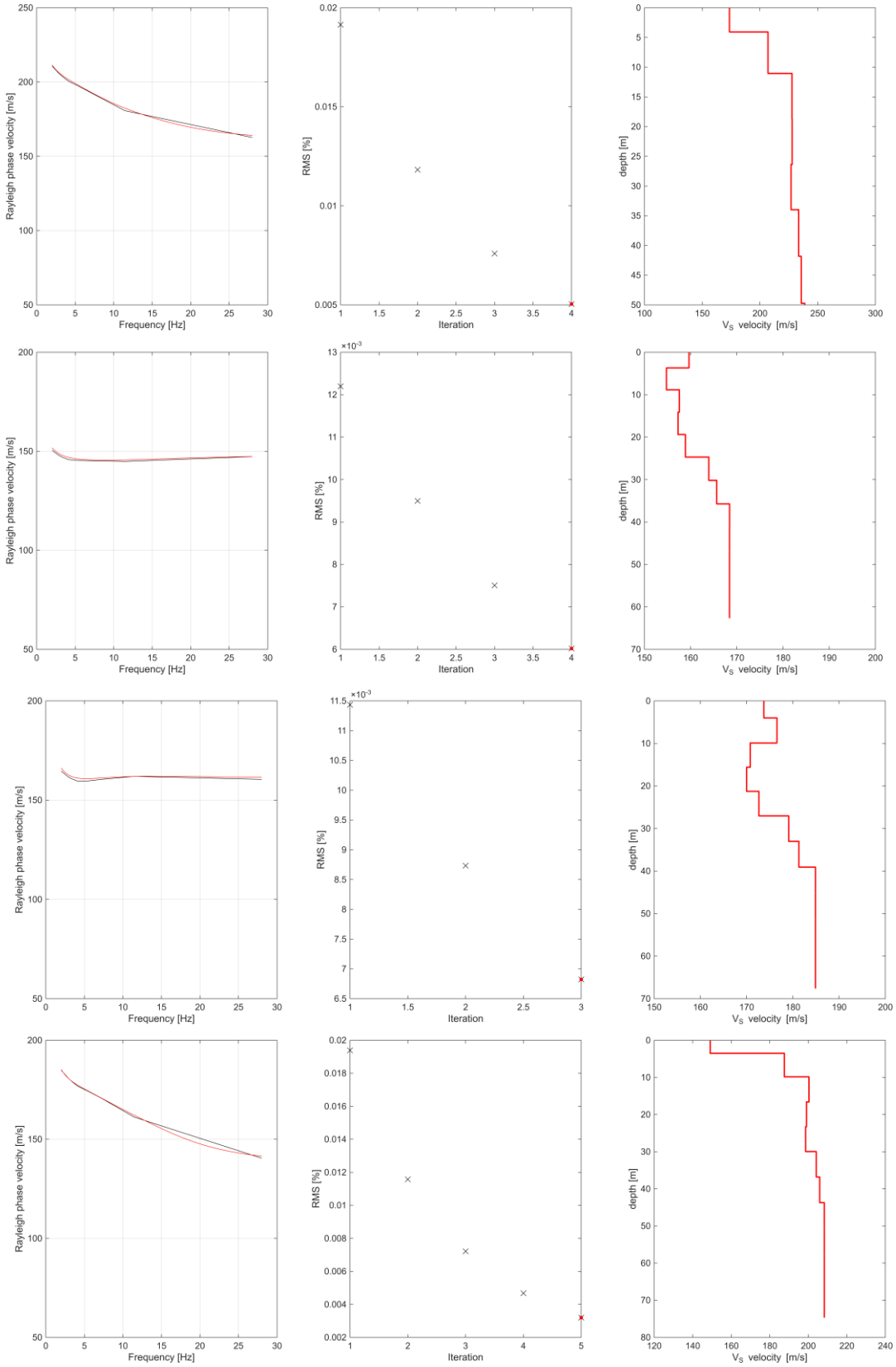

Figure A11 - Example of multiple local dispersion curve inversions from array N2. On the left, each panel shows a comparison between the phase velocity curve obtained from tomographic inversion (black) and the modelled one (red); in the centre there is the trend of the RMS error (%) as a function of the iterations to highlight the convergence towards a minimum; on the right the shear velocity profile (Vs) obtained from the inversion is illustrated. The result shows a good ability of the model to reproduce the observed data, with stable convergence and satisfactory fit of the experimental curve.