Politecnico di Torino

Department of Environment, Land and Infrastructure Engineering

Master of Science in Petroleum Engineering

Using vehicle-induced noise as local seismic source through ML and interferometry

Tarlan Khoveiledy

Supervisors:

Prof. Laura Valentina Socco

Dr. Mohammadkarim Karimpour

Politecnico di Torino 2023

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor, Prof. Laura Valentina Socco, whose guidance, expertise, and mentorship have been invaluable. Your patience, dedication, and belief in my abilities have shaped this thesis and my growth as a researcher. Thank you for your unwavering support and for pushing me to reach my full potential.

I would like to thank Dr. Mohammadkarim Karimpour for his guidance and support, and I am truly grateful for his presence in my journey.

I want to thank my loving family for their endless support and understanding. Your belief in me and constant encouragement have been a source of strength and motivation throughout this endeavor. I am deeply grateful for the love, patience, and sacrifices you have shown me.

I would like to thank my friends who have stood by my side, offering their unwavering support and encouragement. Your presence, laughter, and uplifting words have brought joy to my life and made this journey more meaningful.

I would like to express my gratitude to applied geophysics group, Professor Farbod Khosro Anjom, Professor Chiara Colombero for their support.

I would like to express my gratitude to the Odysseus Unbound Foundation for generously providing funding for the field data acquisition, as well as to all individuals who actively participated in the field work.

LIST OF SYMBOLS AND ACRONYMS

ML	Machine Learning
DC	Dispersion curve
VS	shear wave velocity
MASW	multichannel analysis of surface waves
MAPS	multichannel analysis of passive surface waves
CNN	Convolutional Neural Network
ReLU	Rectified Linear Unit
CCS	Carbon dioxide capture and sequestration
P	Compression
S	Shear
λ	Wavelength
E	Energy
A	Amplitude
k	Wave number
p	Slowness
f	Frequency
v	Phase velocity
x	Space

t	Time
$\mathbb{D}$	Domain
$\partial\mathbb{D}$	Domain boundary
ω	Angular frequency
G	Green's function in time-domain
j	Imaginary unit
ρ	Mass density
$*$	Time convolution
c_P	P-wave propagation velocity
c_S	S-wave propagation velocity
c	Velocity of the medium
S_0	Autocorrelation function
u	Time-reversed wavefield
v_T	Testing phase velocity
φ	Phase
M_i	Normalized misfit
M_{tot}	Total misfit

TABLE OF CONTENTS

1 Introduction ***1***

1.1 Background and Motivation *1*

1.2 Outline *5*

2 Surface wave ***6***

2.1 Surface wave definition *6*

2.2 Surface wave properties *7*

2.3 Surface wave analysis *8*

2.3.1 Data acquisition *8*

2.3.1.1 Active prospecting *8*

2.3.1.2 Passive survey *9*

2.3.2 Processing *9*

2.3.2.1 Virtual shot gathers *9*

2.3.2.2 Phase shift method *11*

2.3.3 Inversion *12*

3 Method ***13***

3.1 Introduction *13*

3.2 Main Processing steps *13*

3.3 Raw noise recording *14*

3.4 Pre-processing 14

3.5 Identifying the time windows with truck passage 15

3.6 Automatic identification of virtual source position 18

3.7 Computation of virtual shot gathers 18

3.8 Dispersion image computation 18

3.8.1 Stacking of the spectra 19

3.8.2 Spatial Window 19

3.9 Dispersion curves 19

3.10 Conclusion 19

4 Data, Processing and Results 21

4.1 Introduction 21

4.2 Data 21

4.3 Processing 22

4.4 Quality control 29

5 Conclusion 31

5.1 Final remark 31

5.2 Suggestions for future development 32

6 References 33

LIST OF FIGURES

Figure 2.1 a) Rayleigh waves cause a rolling motion. B) Love waves cause the ground to shift from side to side (Earle, 2015) 7

Figure 2.2 Geometric dispersion of surface waves in vertically heterogeneous media. λ is the wavelength of the surface wave (Foti et al., 2018) 8

Figure 3.1 The processing steps to retrieve DCs from vehicle-induced noise data 14

Figure 3.2 Illustration of source (moving vehicle) and receiver layout. Δx shows the receivers spacing 14

Figure 3.3 Steps for identifying windows with truck passage 15

Figure 3.4 Data classification steps 16

Figure 3.5 Convolutional neural network. Each neuron of a neural network is connected to a neighborhood of neurons in the previous layer. 16

Figure 3.6 Illustration of spatial windowing with continuously moving window. N represents number of receivers in each spatial segment moving every one receiver 19

Figure 4.1 The aerial view of the field, the maroon line represents the acquisition line 22

Figure 4.2 Eexample of 5 second gather after de-meaning 22

Figure 4.3 a) An example of gather with the truck passage present b) An example of gather with no truck passage 23

Figure 4.4 Training and validation accuracy versus epoch (Total number of iterations of all the training data in one cycle of training) 24

Figure 4.5 The green asterisks are representing the maximum energy picked in each window 25

Figure 4.6 Between two truck passage within the same window the one with higher energy is picked, the maximum energy picked in window is marked with the green asterisks 25

Figure 4.7 a) An example of recording with a truck passage. The red arrow shows the most energetic trace that is used as the virtual source b) The virtual shot gather using seismic interferometry c) An example of active recordings where the position of the active source is near to virtual source in plot(a) 26

Figure 4.8 Example of picked DC as a maxima in the dispersion image corresponding to the spatial window from receiver 144 to 155 using the proposed method 27

Figure 4.9 Computed dispersion image corresponding to the spatial window from receiver 143 to 154 using: a) active data, b) the proposed method. Receiver 100 to 111 using: c) active data, d) the proposed method. Receiver 103 to 114 using: e) active data, f) the proposed method. Black asterisks show the estimated DCs. 28

Figure 4.10 a) Pseudo-section of the DCs for active data along the line. b) Pseudo-section of the DCs for the proposed 29

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND AND MOTIVATION

The utilization of surface wave methods for subsurface characterization is a widely adopted practice that extends from shallow depths to considerable depths. This approach allows researchers to gain valuable insights into the near-surface structures. Acquiring knowledge about the distribution of subsurface properties plays a crucial role in order to effectively plan, design, manage, and assess actions that are influenced by the subsurface.

Understanding the properties of the shallow subsurface is crucial for various projects. Whether it involves constructing buildings, bridges, or infrastructure, the mechanical properties of soils and shallow bedrock play a fundamental role. This knowledge allows engineers to mitigate geological risks during the construction phase as well as throughout the lifespan and decommissioning of structures (Anbazhagan et al., 2013).

By characterizing the shallow subsurface, engineers can determine the soil properties and their variation with depth, which are crucial for foundation design and predicting the dynamic response of the wind turbine system. The main differences between land-based wind turbines and offshore wind turbines are substructures, and the more complex offshore environmental condition bring higher requirements for the foundation design. The foundation, who interacts with the environment directly, plays the key role in the development of offshore

wind industry. Therefore, understandings of current status, adaptabilities, and limitations of offshore wind foundations are necessary for the design and risk control (Pérez-Collazo et al., 2015; Wang et al., 2018; Bhattacharya et al., 2021).

Carbon dioxide capture and sequestration (CCS) in deep geological formations has emerged as an important option for reducing greenhouse emissions. A fundamental understanding of the geologic, hydrologic, geomechanical, and geochemical processes controlling the fate and migration of CO2 in the subsurface is necessary to provide a base for developing methods to characterize storage sites and to select sites with minimal leakage risk. (Benson & Cole, 2008). Monitoring injected CO2 is a regulatory requirement to assure safe storage and is also an important tool to optimize injection operations and confirm storage volumes. One of the monitoring programs is seismic monitoring of the large subsurface area (Furre et al., 2020).

Both passive and active seismic data can be employed for the characterization of these structures. Passive seismic data acquisition relies on ambient vibration wavefields, utilizing the naturally occurring vibrations in the environment. This approach offers several advantages, including its cost-effectiveness, availability, and environmentally friendly nature. To process passive seismic data, seismic interferometry is commonly employed, which involves cross-correlating seismic recordings from different receiver locations. By doing so, Green's functions of the wave field are obtained, representing the response that would be observed at one receiver position if there were an impulsive source at another location (Wapenaar, 2004).

Among the various sources of ambient vibrations, traffic noise has gained significant attention among researchers. Traffic-induced noise is readily available, cost-effective, and environmentally friendly, making it an attractive choice for passive seismic surveys. This literature review aims to explore the current state of research and advancements in the use of traffic-induced noise as passive source method for subsurface characterization. Curtis et al. (2006) explained the operation of seismic interferometry and provides a few examples of its application. Bensen et al. (2007) explained and justified the ambient noise data processing development through salient examples. Halliday et al. (2008) showed results of several different approaches to the interferometric estimation of surface waves in a suburban environment, using both active and passive sources. For the passive case, illustrats that surface wave estimates can be produced using interferometry and that the quality of these estimates can be enhanced with some simple filtering steps. These steps account for the adverse effects of directional bias in the noise and the short recording periods used. To take full advantage of the noise in this setting, concluds that it may be beneficial to account for known sources of noise during survey planning.

Park & Miller (2008) described passive version of the multichannel analysis of surface waves (MASW) method that can be implemented with the conventional linear receiver array deployed alongside a road. They accountd for the cylindrical, instead of planar, nature of surface wave propagation that often occurs due to the proximity of source points, by

considering the distance between a receiver and a possible source point. Nakata et al. (2011) applied the cross-coherence method to the seismic interferometry of traffic noise, which originates from roads and railways, to retrieve both body waves and surface-waves. Their algorithm in the presence of highly variable and strong additive random noise uses cross-coherence, which uses normalization by the spectral amplitude of each of the traces, rather than cross-correlation or deconvolution. However, this algorithm is particularly effective where the relative amplitude among the original traces is highly variable from trace to trace.

Behm & Snieder (2013) showed that surface waves can be efficiently recovered from interferometry applied to local ambient noise. They showed that Love waves in the frequency band of 1.5 to 5 Hz can be obtained from local noise interferometry, and that they are of comparable S/N as Rayleigh waves. Thus, they may also be used to constrain the near-surface structure. Behm et al. (2014) investigated the feasibility of using locally-generated seismic noise in the 2–6 Hz band to obtain a subsurface model via interferometric analysis. They found traffic noise from a nearby state road to be the dominant source of surface waves recorded on the array and observe surface wave arrivals associated with this source up to distances of 5 kms. Cheng et al. (2015) proposed passive seismic method based on seismic interferometry and MASW. They conclude that proposed passive seismic method is an energetic method which combines advantages of both seismic interferometry and MASW. It is non-destructive, non-invasive, effective, and relatively more accurate. Besides, it costs less and explores deeper.

Quiros et al. (2016) applied Seismic interferometry to railroad traffic recorded by an array of vertical component seismographs along a railway within the Rio Grande rift has recovered surface and body waves characteristic of the geology beneath the railway. These results confirm that train-generated vibrations represent a practical source of high-resolution subsurface information, with particular relevance to geotechnical and environmental applications. Cheng et al. (2016) proposed multichannel analysis of passive surface (MAPS) waves based on long noise sequence cross-correlations. A hybrid method of seismic interferometry and the roadside passive MASW using cross-correlation to produce common virtual source gathers from multichannel noise records is introduced. Common virtual source gathers are then used to do dispersion analysis with an active scheme based on phase-shift measurement.

Lehujeur et al. (2017) developed and applied a method for ambient noise surface wave tomography that can deal with noise cross-correlation functions governed to first order by a non-uniform distribution of the ambient seismic noise sources. They Showed that for networks with limited aperture, taking the azimuthal variations of the noise energy into account has significant impact on the surface wave dispersion maps. Dou et al. (2017) reported the first end-to-end study of time-lapse VS imaging that uses traffic noise continuously recorded on linear distributed acoustic sensing (DAS) arrays. This study demonstrats the efficacy of near-surface seismic monitoring using DAS-recorded ambient noise.

Pang et al. (2019) proposed a data selection technique in time domain for selective stacking of cross-correlations and applied it to improve the MAPS method. Zhang et al. (2020) indicates that traffic noise can be created using vehicles or vibrators to capture surface waves within a reliable frequency band of 2–25 Hz if no vehicles are moving along the survey line. They showed that the virtual shot gathers obtained by cross-correlation appear less noisy than those obtained by cross-coherence. (Mi et al., 2023) Proposed direct-wave interferometry with correlated sources that stacking the interferometric responses from high-speed trains with different speeds and demonstrated it reduces the spurious imprint of source self-correlations, but it cannot remediate the problem. They eliminated the influence of the source wavelet by using cross coherence.

Mi et al. (2022) developed an application of using traffic noise with seismic interferometry to investigate near-surface structures. Noise data were recorded by dense linear arrays with approximately 5 m spacing deployed along two crossing roads. They extracted phase-velocity dispersion curves (DCs) from virtual shot gathers using MASWand the result coincides with the results from active seismic data, which supports the deployment of receivers along roads for efficiency and convenience in highly populated urban areas. The results demonstrate the accuracy and efficiency of delineating near-surface structures from traffic-induced noise, which has great potential for monitoring subsurface changes in urban areas.

You et al. (2023) proposed a processing scheme to retrieve high-quality surface wave signals from vibrations generated by high-speed trains moving on the track built on viaducts. They used a signal selection method to attenuate the directional noise effect of linear-array based passive surface wave measurement, which automatically and directly selects signals generated by the high-speed train when it runs in the stationary phase zone.

Since the size of ambient noise data are enormous application of ML in surface wave analysis utilizing ambient noise has gained attention among researchers. Huot et al. (2017) showed a variety of methods, mixing traditional signal processing and machine learning, to automatically assist geophysicists in analyzing the ambient noise recorded and selecting non-ideal noises. They demonstrated that different types of noise and specific events can be identified using ML. Martin et al. (2018) analyzed 37 days of ambient noise; introduced an exploratory data analysis tool that clusters a week of noise to quickly find coherent, repeating noises that inhibit reliable extraction of useful signals. Huot & Biondi (2018) built a neural network for detecting traffic noise. They pointed out that while training neural networks can be computationally involved, they perform remarkably fast at run-time and are hence well-suited for event detection problems.

Most studies utilizing traffic noise do not precisely locate the position of individual vehicles. Instead, they assume that the noise originates from the surrounding seismic deployment or a particular direction relative to it. To overcome this limitation, a seismic line can be deployed along a road with regular truck passage, enabling each vehicle to be treated as a distinct source. However, one of the challenges in this approach is automatically identifying

the time windows that contain vehicle passages. To address this challenge in the present work, we employ machine learning (ML) algorithm to detect the presence of vehicles in the seismic recordings. Once the vehicles are identified, we apply seismic interferometry to generate virtual seismic gathers, treating the trucks as local active sources. We compute dispersion images using phase shift method (Park et al., 1999) and stack them to increase quality. These dispersion images represent the variation of surface wave velocities across the surveyed area. By stacking multiple dispersion images, the overall quality of the results can be improved. Subsequently, surface wave DCs can be estimated, which provide valuable information about the subsurface properties.

In summary, this thesis aims to utilize surface wave methods for subsurface characterization by leveraging vehicle-induced noise. By employing a ML algorithm to detect vehicle presence, applying seismic interferometry to generate virtual seismic gathers. Through the computation of dispersion images and estimation of surface wave DCs, a comprehensive understanding of the subsurface structures can be achieved. This research approach offers significant potential for enhancing the efficiency and effectiveness of subsurface characterization techniques.

1.2 OUTLINE

The second chapter of the thesis, presents the basics of surface wave methods and propagation. We discuss the main steps for the surface analysis: Seismic data acquisition, processing and inversion. We explain both active and passive data acquisition briefly.

In the third chapter, we explain processing procedure from raw passive recording to retrieval of the DCs. In this work we utilized traffic induced noise as a passive source by considering the direction of the noise. We explain the processing steps needed for the locating the truck in the recording, seismic interferometry, and procedure for retrieval of DCs in this chapter.

In the fourth chapter, we apply the method proposed in the third chapter, on the 2 days recording in Kefalonia, Greece. We present the results and the DCs extracted from passive data are compared to the active data to assess the quality of the proposed method.

In the final chapter we critically analyze the outcomes of this work and provide recommendations for future work.

CHAPTER 2

SURFACE WAVE

Traditional seismic techniques rely on body-wave propagation, specifically P-wave reflection. Surface waves, are considered unwanted noise that needs to be eliminated or reduced. However, surface waves can be interpreted to characterize the shallow near surface. Surface waves have attracted researchers from diverse disciplines like solid-state physics, engineering, seismology, and geophysics (Shapiro & Campillo, 2004; Boiero et al., 2011; Bussat & Kugler, 2011; Garofalo et al., 2016; Kästle et al., 2018; Jokar et al., 2019). These waves offer noninvasive methods to characterize materials at different scales, from small defects to large-scale Earth structure investigations. By exploiting the geometric dispersion of surface waves, valuable information about the medium's properties can be extracted.

2.1 SURFACE WAVE DEFINITION

Seismic surface waves are a type of seismic wave that propagates along or near the Earth's surface. Seismic surface waves consist of: Rayleigh waves, Love waves, and Scholte waves. Rayleigh waves, are a combination of both compressional (P) and shear (S) waves. They cause elliptical motion of particles as they propagate, with the largest amplitudes at the Earth's surface (Figure 2.1.a). Love waves, are horizontally polarized shear waves that move

in a horizontal, side-to-side motion perpendicular to the direction of wave propagation (Figure 2.1.b). Scholte waves are equal to Rayleigh waves but instead of an interface between air and solid we have an interface between liquid and solid. Typically they generate along the seabed when the source is close enough to the water bottom.

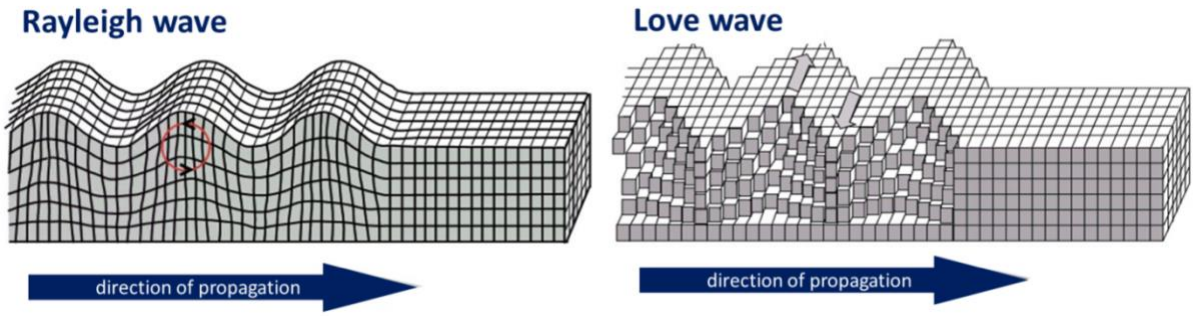


Figure 2.1 a) Rayleigh waves cause a rolling motion. B) Love waves cause the ground to shift from side to side (Earle, 2015)

2.2 SURFACE WAVE PROPERTIES

Surface waves are seismic waves that propagate parallel to the earth's surface. Their amplitude decreases exponentially with depth, and most of the energy is contained within one wavelength (λ) from the surface. Their propagation depends only on the mechanical and geometric properties of the subsurface. Their propagation depth depends on the wavelength, higher wavelengths (low frequencies) propagate to higher depth compare to the lower wavelengths (high frequencies), i.e., geometric dispersion. Therefore, in vertically heterogeneous medium, the harmonics of the propagating surface wave with high frequencies propagate in top layers of subsurface thus their velocity depends on properties of shallow subsurface, and the harmonics of the propagating surface wave that have lower frequencies propagate in deeper layers of subsurface thus their velocity depends on properties of deeper subsurface. Surface waves, compared to body waves, are more energetic and experience less energy loss due to geometric spreading therefore they are dominant in seismic recording (Figure 2.2).

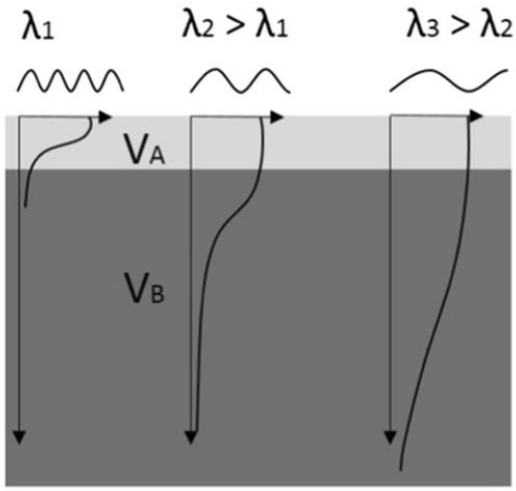


Figure 2.2 Geometric dispersion of surface waves in vertically heterogeneous media. λ is the wavelength of the surface wave (Foti et al., 2018)

2.3 SURFACE WAVE ANALYSIS

Surface wave analysis is conducted to estimate the profile of seismic shear wave velocity (VS) by solving an inverse problem that involves identifying model parameters through the analysis of experimental DCs. This analysis typically follows a three-step process: seismic data acquisition, processing, and inversion.

2.3.1 DATA ACQUISITION

There are two types of surface wave measurements: active and passive. Active measurements involve using a seismic source, while passive measurements utilize ambient vibration wavefields. The choice of array geometry depends on the measurement type. For active-source prospecting, receivers are typically placed in-line with the seismic source. For passive tests, 2D arrays of sensors deployed on the ground surface are preferred for developing robust VS profiles. Whenever it is feasible, the combination of active-source and passive data is beneficial for obtaining a well-constrained VS model that extends from the surface to large depths.

2.3.1.1 Active prospecting

The MASW (Park et al., 1999; Xia et al., 1999) method is the most commonly used acquisition layout for surface wave analysis, where vertical receivers are evenly spaced and aligned with the seismic source. To ensure an adequate signal-to-noise ratio over the desired frequency range for the target investigation depth, the energy provided by the seismic source

must be carefully considered. Preliminary assumptions about the expected velocity range are necessary to define the required frequency band of the source.

The choice of source depends on factors such as expected velocity range, presence of velocity contrasts, and cost considerations. Active prospecting options include vertically operated shakers, impact sources, weight-drop systems, explosive sources, and sledgehammers. Multiple sources may be used to cover a wider frequency range, and signal-to-noise ratio can be improved by stacking records from multiple shots. Vertical geophones are commonly used as receivers for Rayleigh wave data acquisition in surface wave analysis(Foti et al., 2018).

2.3.1.2 Passive survey

The penetration depth of a surface wave is approximately equal to its wavelength (Richart et al., 1970) .When small sources (e.g., hammer, small weight drop) are used is would be rare to generate energy at frequencies less than 8–10 Hz. Conversely, passive surveys allow measuring dispersion from low to intermediate frequencies, and 2D sensor arrays are recommended for capturing vibrations from different directions (Foti et al., 2018). Passive surface-wave analysis records ambient vibrations without an artificial seismic source. Ambient vibrations arise from natural and human activities. Assessing the quality of passive surveys is complex, and careful differentiation between "signal" and "noise" is necessary. Noise can come from various sources, while the signal consists of Rayleigh (and possibly Love) waves from distant sources. The reliability of results varies with the site's ambient vibration level. In the following, we explain the processing technique designed for passive surveys, focusing on the analysis and processing of vehicle-induced noise.

2.3.2 PROCESSING

In the second step, the field data undergo processing to extract an experimental DC. Various processing techniques, particularly in the spectral domain, can be employed for seismic data analysis. One of the commonly used processing methods for extraction of phase-velocity DCs is Multichannel method. Multichannel is a transform-based method, converts space-time (x-t) domain into frequency-wavenumber (f-k), frequency-slowness (f-p), and frequency-phase velocity (f-v) domains. In this transformed domain, the data form a dispersion image, that shows the energy distribution, allows for the identification and picking of DCs by locating the spectral maxima.

2.3.2.1 Virtual shot gathers

Passive seismic data can be processed using seismic interferometry. Seismic interferometry refers to the principle of generating new seismic responses by cross-correlating

seismic observations at different receiver locations. The result is Green's function of the wave field that would be observed at one of these receiver positions if there were an impulsive source at the other (Wapenaar, 2004).

The Green's function $\hat{G}(\mathbf{x}, \mathbf{x}_A, \omega)$ is the Fourier transform of the causal time-domain Green's function $G(\mathbf{x}, \mathbf{x}_A, t)$, which represents an impulse response observed at $\mathbf{x}$, due to a source at $\mathbf{x}_A$, this is also applied for $\hat{G}(\mathbf{x}, \mathbf{x}_B, \omega)$. The basic representation for seismic interferometry is:

$$\hat{G}(\mathbf{x}_B, \mathbf{x}_A, \omega) + \hat{G}^*(\mathbf{x}_B, \mathbf{x}_A, \omega) = \oint_{\partial\mathbb{D}} \frac{-1}{j\omega\rho(\mathbf{x})} \left(\hat{G}^*(\mathbf{x}_A, \mathbf{x}, \omega)\, \partial_i \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega) - \left(\partial_i \hat{G}^*(\mathbf{x}_A, \mathbf{x}, \omega)\right) \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega)\right) n_i \mathrm{d}^2\mathbf{x} \quad 2.1$$

where $\mathbb{D}$ is domain, $\partial\mathbb{D}$ is the boundary of domain, j is the imaginary unit, ω the angular frequency, ρ mass density, $*$ denotes time convolution, $\hat{G}$ and $\partial_i \hat{G} n_i$ under the integral in the right-hand side of equation represent responses of monopole and dipole sources at $\mathbf{x}$ on ∂D. The products $\hat{G}^*\, \partial_i \hat{G} n_i$, etc., correspond to cross-correlations in the time domain (van Manen et al., 2005; Wapenaar & Fokkema, 2006). By assuming that the medium at and outside $\partial\mathbb{D}$ is homogeneous and isotropic, with P-wave propagation velocity c_P, and S-wave propagation velocities c_S. In the The Eq. 2.1 can be rewritten as:

$$\hat{G}(\mathbf{x}_B, \mathbf{x}_A, \omega) + \hat{G}^*(\mathbf{x}_B, \mathbf{x}_A, \omega) \approx -\frac{2}{j\omega\rho} \oint_{\partial\mathrm{D}} \left(n_i\, \partial_i \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega)\right) \hat{G}^*(\mathbf{x}_A, \mathbf{x}, \omega) \mathrm{d}^2\mathbf{x} \quad 2.2$$

(Wapenaar & Fokkema, 2006). In passive interferometry, the positions of the sources are unknown and $\partial\mathbb{D}$ can be very irregular. In that case, the operation $n_i\, \partial_i$ is replaced by a factor $-jk$, where $k = \omega/c$, and c is the velocity of the medium which leads to:

$$\hat{G}(\mathbf{x}_B, \mathbf{x}_A, \omega) + \hat{G}^*(\mathbf{x}_B, \mathbf{x}_A, \omega) \& \approx \frac{2}{\rho c} \oint_{\partial\mathrm{D}} \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega) \hat{G}^*(\mathbf{x}_A, \mathbf{x}, \omega) \mathrm{d}^2\mathbf{x} \quad 2.3$$

Transforming both sides of Eq. 2.3 back to the time domain yields:

$$G(\mathbf{x}_B, \mathbf{x}_A, t) + G(\mathbf{x}_B, \mathbf{x}_A, -t) \approx \frac{2}{\rho c} \oint_{\partial\mathrm{D}} G(\mathbf{x}_B, \mathbf{x}, t) * G(\mathbf{x}_A, \mathbf{x}, -t) \mathrm{d}^2\mathbf{x} \quad 2.4$$

(Wapenaar, Slob, et al., 2010). For sources acting separately in time, and having an equal autocorrelation function $S_0(t)$, located at positions $\mathbf{x}$ along an enclosing source boundary $\partial\mathbb{D}$, one may use the seismic interferometry cross-correlation expression:

$$\Re\{G(\mathbf{x}_B, \mathbf{x}_A, t)\} * S_0(t) \approx \frac{1}{\rho c} \oint_{\partial\mathbb{D}} u^{\mathrm{obs}}(\mathbf{x}_A, \mathbf{x}, -t) * u^{\mathrm{obs}}(\mathbf{x}_B, \mathbf{x}, t) \mathrm{d}\mathbf{x} \quad 2.5$$

where $\Re$ denotes the real part, and $u^{\text{obs}}(\mathbf{x}_A, \mathbf{x}, -t)$ is the time-reversed wavefield observed at $\mathbf{x}_A$ due to a transient source at $\mathbf{x}$.

For seismic interferometry with uncorrelated noise sources, a relation to retrieve the Green's function $G(\mathbf{x}_B, \mathbf{x}_A, t)$ derived as:

$$\Re\{G(\mathbf{x}_B, \mathbf{x}_A, t)\} * S_0(t) \approx \langle u^{\text{obs}}(\mathbf{x}_A, -t) * u^{\text{obs}}(\mathbf{x}_B, t) \rangle \quad 2.6$$

(Wapenaar & Fokkema, 2006) where the noise sources are assumed to have the same autocorrelation function $S_0(t), u^{\text{obs}}(\mathbf{x}_A, -t)$ stands for the time-reversed total recorded noise at $\mathbf{x}_A$ due to all the noise sources and $\langle \cdot \rangle$ denotes ensemble average. For field applications, the ensemble average is exchanged for averaging over long recording times. As the long-time recordings are stored in time windows with certain length, the time averaging is exchanged for summation over all i time windows:

$$\Re\{G(\mathbf{x}_B, \mathbf{x}_A, t)\} * S_0(t) \approx \sum_i \left(u^{\text{obs}}(\mathbf{x}_A, -t) * u^{\text{obs}}(\mathbf{x}_B, t) \right)_i \quad 2.7$$

(Vidal et al., 2014). According to Eq. 2.7, traces recorded at two different locations, are cross-correlated. Then all the cross-correlated responses over all source locations are summed to reconstruct a virtual trace received at B with a virtual source at A (Wapenaar, Draganov, et al., 2010). Once surface waves have been reconstructed through seismic interferometry, the next step is to analyze their dispersion characteristics.

2.3.2.2 Phase shift method

The Phase Shift method (Park et al., 1999) analyzes the seismic wave data in the frequency-phase velocity (f-v) domain to estimate the phase velocity DC. It involves representing the spectrum of the seismic data and extracting the phase information. The phase of a wave represents the fraction of a cycle that a given point has completed at a specific time. It is defined as:

$$\varphi_i(f) = \frac{2\pi f x_i}{v(f)} \quad 2.8$$

Where φ_i is the phase at x_i, f is the frequency of the wave, x_i is the position in space v is the phase velocity of the wave at x_i. The relationship between the phase and waveform can be described as:

$$U(x_i, f) = A_i e^{-j\varphi_i(f)} \quad 2.9$$

Where $U(x_i, f)$ is the seismic wave data at position x_i and frequency f, A_i is the amplitude of the wave at x_i. Applying the Fourier transform to the waveform we get:

$$U(x_i, f) = \sum_{i=1}^{N} A_i e^{-j\varphi_i(f)} \quad 2.10$$

Assuming various testing phase velocities v_T, a phase shift is calculated for each trace and frequency as:

$$\varphi(f)_{Ti} = \frac{2\pi f x_i}{v_T} \quad 2.11$$

Substituting the phase definition and slant stacking along all v_T values, by averaging the phase-shifted seismic wave data $U(x_i, f)$ over all receiver positions x_i at each frequency f, gives us the estimation of the summed amplitude:

$$\widehat{U}(f, v_T) = \sum_{i=1}^{N} U(x_i, f) e^{-j\varphi(f)_{\mathrm{T}i}} \quad 2.12$$

The resulting spectrum in the phase velocity domain provides information about the dispersion characteristics of the seismic waves.

2.3.3 INVERSION

The primary objective of the inversion process is to identify the optimal subsurface model that closely matches the experimental data. In the inversion process, the experimental DCs are utilized as the reference to determine model parameters. The subsurface is typically represented as a layered elastic medium, with emphasis placed on layer thickness and shear-wave velocities due to their influence on surface waves. The objective is to identify the shear-wave velocity profile that offers the closest match to the experimental data.

CHAPTER 3

METHOD

3.1 INTRODUCTION

In this chapter, we present a method for utilization of vehicle-induced noise by locating the position of the individual vehicle using ML algorithm and energy computation. We explain processing data using seismic interferometry, by cross-correlating seismic recording at different receiver locations to generate virtual source gathers (Wapenaar, 2004), computing dispersion images using spatial windowing and phase shift method (Park et al., 1999), and extracting the DCs and assessing the quality of the extracted DCs.

3.2 MAIN PROCESSING STEPS

In Figure 3.1, we provide a sketch of the main processing steps. We then describe in detail each of them.

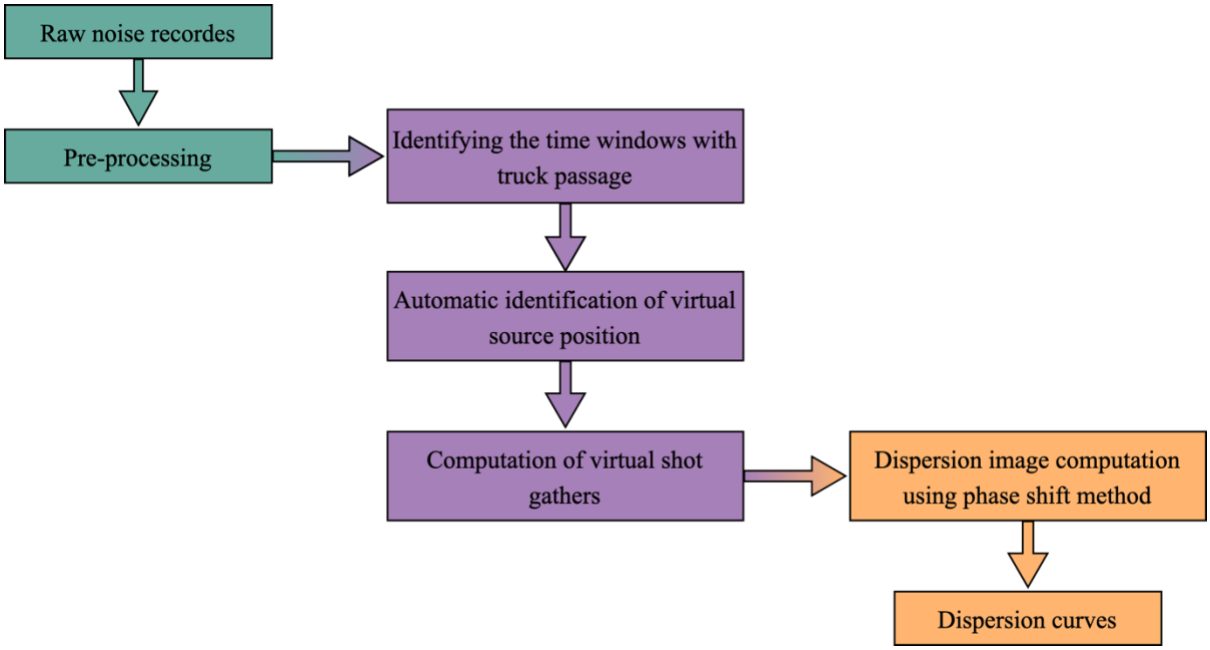


Figure 3.1 The processing steps to retrieve DCs from vehicle-induced noise data

3.3 RAW NOISE RECORDING

We acquire the data along a road where trucks were passing by, employing vertical receivers. Figure 3.2 shows a vehicle is moving from left to right direction. There is a linear receiver array parallel to the road that record vehicle-induced noise.

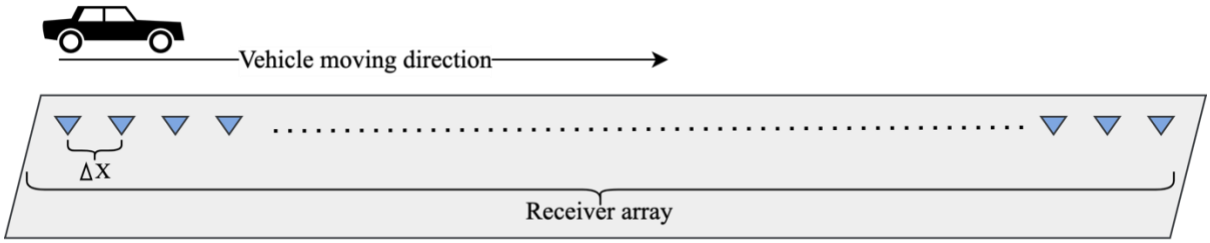


Figure 3.2 Illustration of source (moving vehicle) and receiver layout. Δx shows the receivers spacing

3.4 PRE-PROCESSING

The first step of data processing consists of pre-processing of raw data recording. For processing noise data firstly, we divide the raw data into short time window gathers. This helps to segment the data into smaller, manageable units for further processing. Secondly, the de-meaning and low-pass filtering are utilized to these gathers.

3.5 IDENTIFYING THE TIME WINDOWS WITH TRUCK PASSAGE

We employ supervised ML algorithms in Python to recognize the gathers in which the recorded data correspond to the passage of the vehicles. In Figure 3.3, we provide a sketch of steps for identifying windows with truck passage. We then describe in detail each of them.

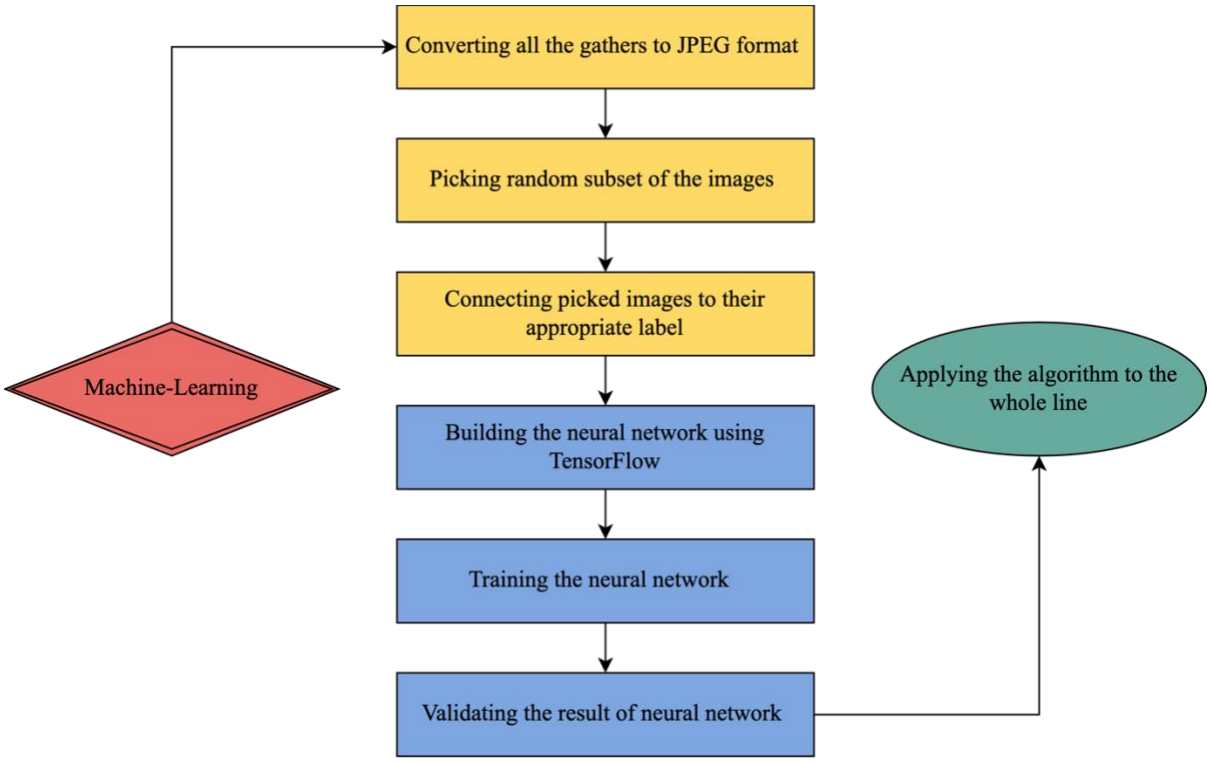


Figure 3.3 Steps for identifying windows with truck passage

To apply ML, we convert the gathers into images in JPEG format to reduce the size of the field data. We randomly pick a small subset of these images, and we divide it into two classes with appropriate labels. We then subdivide each class into training and validation subdirectories (Figure 3.4).

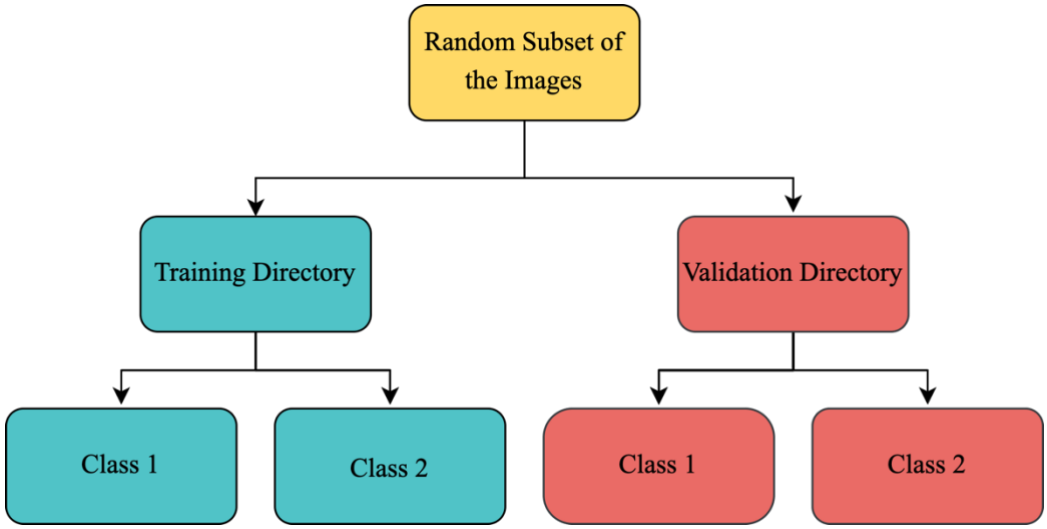


Figure 3.4 Data classification steps

We define a neural network to match the data to their appropriate labels. We use TensorFlow (Abadi et al., 2016) to create a Convolutional Neural Network (CNN) utilizing Keras sequential layers. The series consist of convolutional, pooling, flattening, and dense layers (Figure 3.5).

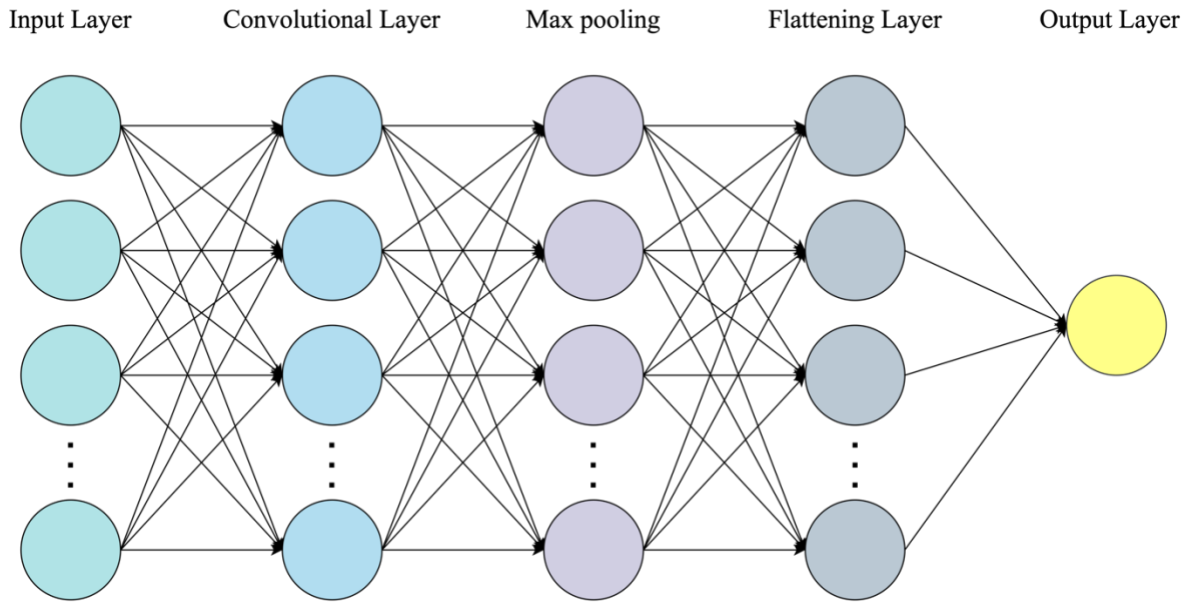


Figure 3.5 Convolutional neural network. Each neuron of a neural network is connected to a neighborhood of neurons in the previous layer.

The object of the convolutional layer is to extract the common features of the images that belong to the same directory. It learns the best filter that matches these images to their labels. The next layer is max pooling, which reduces the size of the image to make the process faster. Pooling is a fundamental concept in CNNs that plays a crucial role in reducing the computational burden and controlling the number of connections between convolutional layers.

In CNNs, pooling is typically applied after convolutional layers to down sample the feature maps while retaining the most relevant information. The most common type of pooling is max pooling. Max Pooling layer is a layer that computes the max of each filter output within adjacent windows (Ranzato et al., 2007). This is important because the size of the data in passive seismic data processing can be very large. The maxed pooling function can be formulated as follows:

$$y_{kij} = \max_{(p,q)\in\mathcal{R}_{ij}} x_{kpq} \quad 3.1$$

where y_{kij} is the output of the pooling operator related to position (i, j) in k^{th} feature map, and x_{kpq} is the element at (p, q)(Gu et al., 2018). The flatten layer is typically used after the convolutional and pooling layers in a CNN architecture. We utilize a flattening layer to transform the data into a one-dimensional array, facilitating processing by the dense layer.

We employ dense layers to classify images based on output from convolutional layers. The initial dense layer consists of neurons with a Rectified Linear Unit (ReLU) activation function. The ReLU activation function is a widely used non-linear activation function in neural networks. It introduces non-linearity to the network, allowing it to learn complex relationships between inputs and outputs. The ReLU activation function is defined as follows:

$$y_i = max(0, z_i) \quad 3.2$$

where z_i is the input of i^{th} channel. ReLU is a piecewise linear function which prunes the negative part to zero and retains the positive part (Nair & Hinton, 2010). Neural networks can be effectively trained using the ReLU activation function without the need for pre-training (Krizhevsky et al., 2017). Given our aim to perform binary classification, the output dense layer comprises a single neuron with a sigmoid activation function. The sigmoid activation function is commonly used in binary classification tasks. It maps the input values to a range between 0 and 1, which can be interpreted as the probability of the input belonging to the positive class. The sigmoid function is defined as:

$$f(x) = 1/(1 + \exp(-x)) \quad 3.3$$

where x is the input to the function and $f(x)$ is the output. In binary classification, the output of the sigmoid function can be interpreted as the probability of the input belonging to the positive class, while the probability of belonging to the negative class is simply $1 - f(x)$. We compile the model with a binary-crossentropy loss function and accuracy metrics. It measures the dissimilarity between the predicted probabilities and the true class labels. The formula for binary cross-entropy loss is:

$$L = -[y * log\ (p) + (1 - y) * log\ (1 - p)] \quad 3.4$$

Where *L represents loss,* y represents the true class label, and p represents the predicted probability of the positive class obtained using the sigmoid activation function. The goal is to minimize this loss function during training. In order to mitigate overfitting, which can result in inaccurate labelling of new data, we expand the training dataset using image augmentation techniques such as width and height shifting, zooming, and horizontal flipping.

We created identification number for each gather. The result of ML classification gives us the identification number of gathers with truck passage present. With these identification number we find the signal gathers with truck passage.

3.6 AUTOMATIC IDENTIFICATION OF VIRTUAL SOURCE POSITION

After the classification of the data and finding the signal gathers with truck passage using the ML algorithm. We track the position of the truck in time and space in each gather. We assume that the trace closest to the vehicle at each time has the highest energy and we identify it as:

$$E_i = \sum |A_i|^2 \tag{3.5}$$

where E_i represent the energy of the ith trace, and A_i is the amplitude of the seismic signal of the ith trace (Colombero et al., 2019).

3.7 COMPUTATION OF VIRTUAL SHOT GATHERS

The highest energy trace computed by Eq. 3.5 is used as virtual source and cross-correlated using Eq. 2.7 with all other traces, resulting in a virtual shot gather. This process is repeated for all the identified virtual sources, leading to the production of a large collection of virtual shot gathers.

3.8 DISPERSION IMAGE COMPUTATION

The virtual shot gathers are then processed to retrieve surface wave DCs along the seismic line by spanning the line with a spatial window and using phase shift method. For each window position we stack the computed dispersion images of several shot positions to increase the quality of the final dispersion image.

3.8.1 STACKING OF THE SPECTRA

To reduce the experimental uncertainty, the quality of an experimental DC can be improved by employing a stacking technique in the frequency domain. The resulting dispersion image from stacking in frequency domain has higher resolution and S/N ratio (Grandjean & Bitri, 2006; Neducza, 2007).

3.8.2 SPATIAL WINDOW

We extract several local DCs, each of them referring to a different subsurface portion, along the survey line using a spatial windowing, the window continuously moves along the common receiver gather. Spatial windowing of the seismic traces, mitigate the effect of lateral variations to focus the investigation within a certain spatial range. In this way, the DC becomes a local property of the subsurface beneath the receivers whose traces are weighted more.

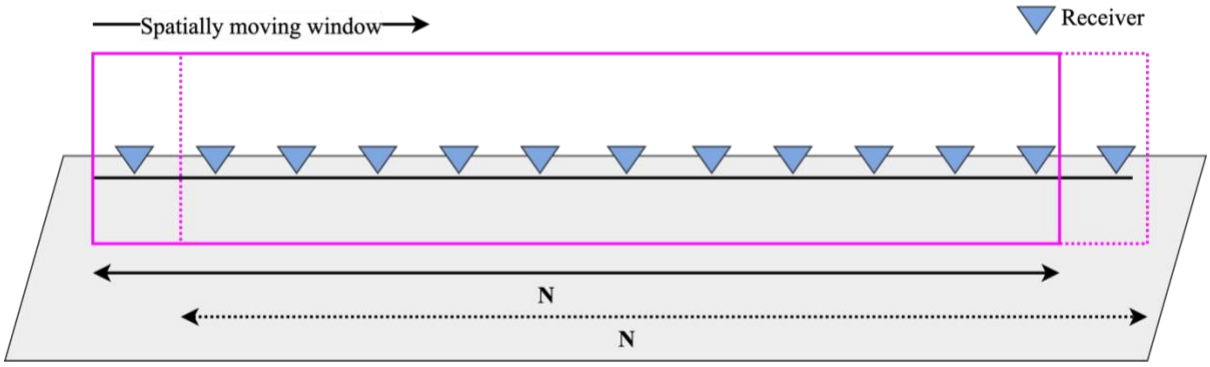


Figure 3.6 Illustration of spatial windowing with continuously moving window. N represents number of receivers in each spatial segment moving every one receiver

3.9 DISPERSION CURVES

The data in f-v_T domain shows the energy distribution in form of dispersion image. We extract the DC as the maxima of the stacked dispersion image. By repeating this procedure for every window position, we obtain a set of DCs along the receiver line.

3.10 CONCLUSION

We have presented workflow for the estimation of the surface DCs from vehicle-induced noise seismic data using ML and seismic interferometry (Wapenaar, 2004). We described method for processing of the retrieved virtual gathers using phase shift method (Park

et al., 1999) to compute dispersion image and extract DCs. In the following chapter the application of the workflow on the vehicle noise recording from Kefalonia, Greece is presented.

CHAPTER 4

DATA, PROCESSING AND RESULTS

4.1 INTRODUCTION

In this chapter, we apply the method to the vehicle-induced noise dataset from from Kefalonia, Greece. We pre-processed the data to get small time windows we searched for the vehicle passage in the time window and we find vehicles position in windows by energy computation. We use seismic interferometry to generate virtual gathers. From the generated virtual gathers using phase shift method we compute the dispersion image, we extract the DCs from the computed dispersion image. We assess the quality of the DCs extracted from passive data by comparing them with extracted DCs from the active data.

4.2 DATA

Field experiment was carried out in Kefalonia, Greece, data where continuously recorded for a period of 2 days. The data were acquired along a road where trucks were passing by, employing 234 vertical 4.5 Hz receivers with a 5 m receiver spacing (Figure 4.1).

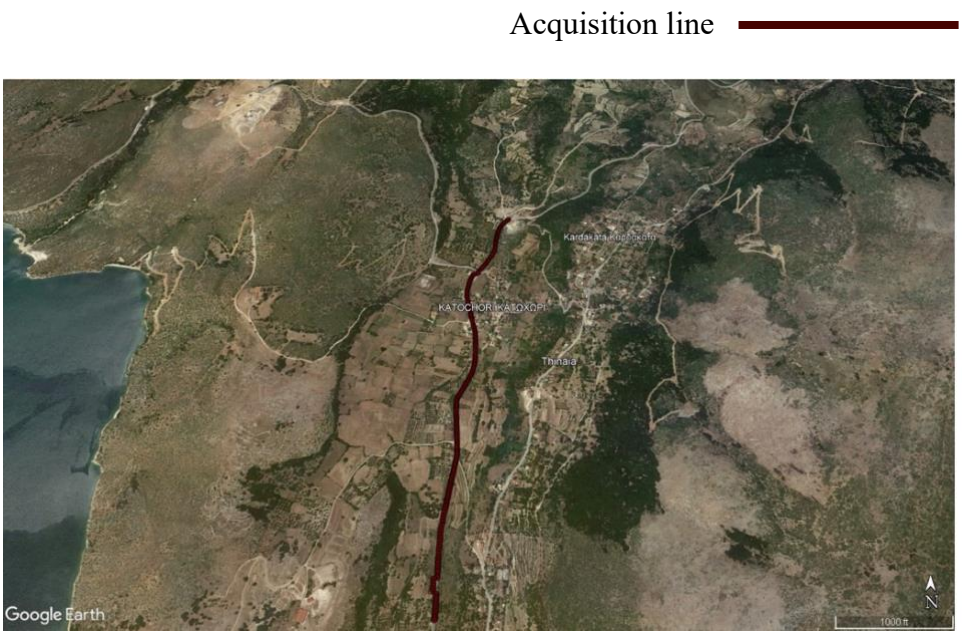


Figure 4.1 The aerial view of the field, the maroon line represents the acquisition line

4.3 PROCESSING

First, we divided continuous recordings into 5 second window gathers. We applied de-meaning and low-pass (100 Hz) frequency filtering to all the gathers (Figure 4.2). De-meaning involves subtracting the mean value from each data point within the time window gather. A low-pass filter attenuates or reduces high-frequency components in the data, allowing only the lower frequency components to pass through. This is useful for removing high-frequency noise or unwanted signal variations that may be present in the raw data. Low-pass filtering at 100 Hz is the filter is designed to pass frequencies below 100 Hz and attenuate frequencies above that.

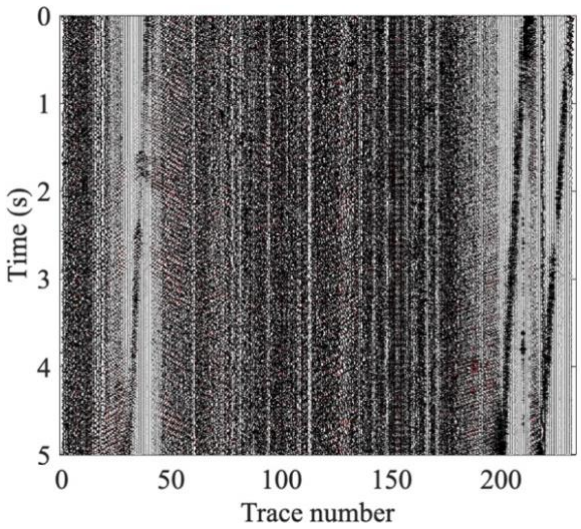


Figure 4.2 Eexample of 5 second gather after de-meaning

Then we searched for patterns indicative of vehicle passage using ML. To avoid analysing a large number of (approximately 25,000) gathers, we selected 400 samples to train the neural network algorithm. We classified the selected 400 figures into two categories, labelled as "Truck" and "No Truck" (Figure 4.3.a and Figure 4.3.b respectively).

a)

b)

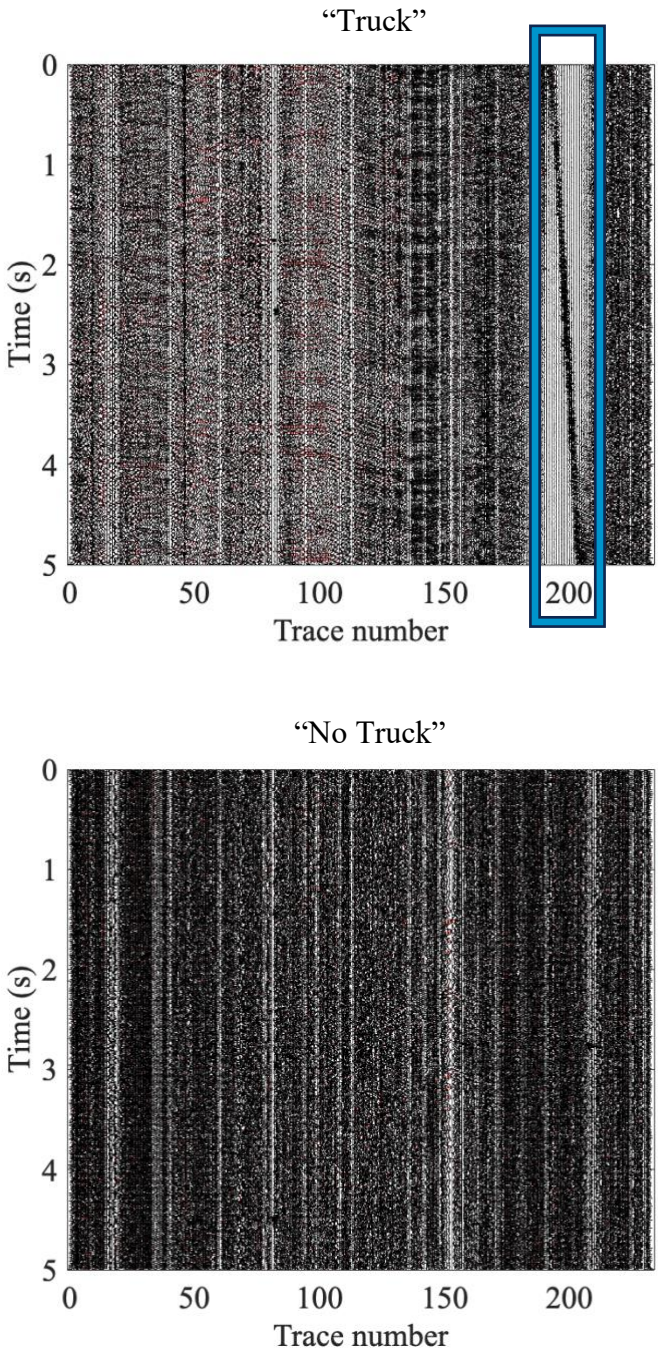


Figure 4.3 a) An example of gather with the truck passage present b) An example of gather with no truck passage

We divided 80% of these data as the training dataset and the remaining 20% as the validation dataset. We achieved 99% accuracy in training dataset and 97% accuracy in validation dataset, after 20 epochs of training on our training and validation datasets (Figure 4.4).

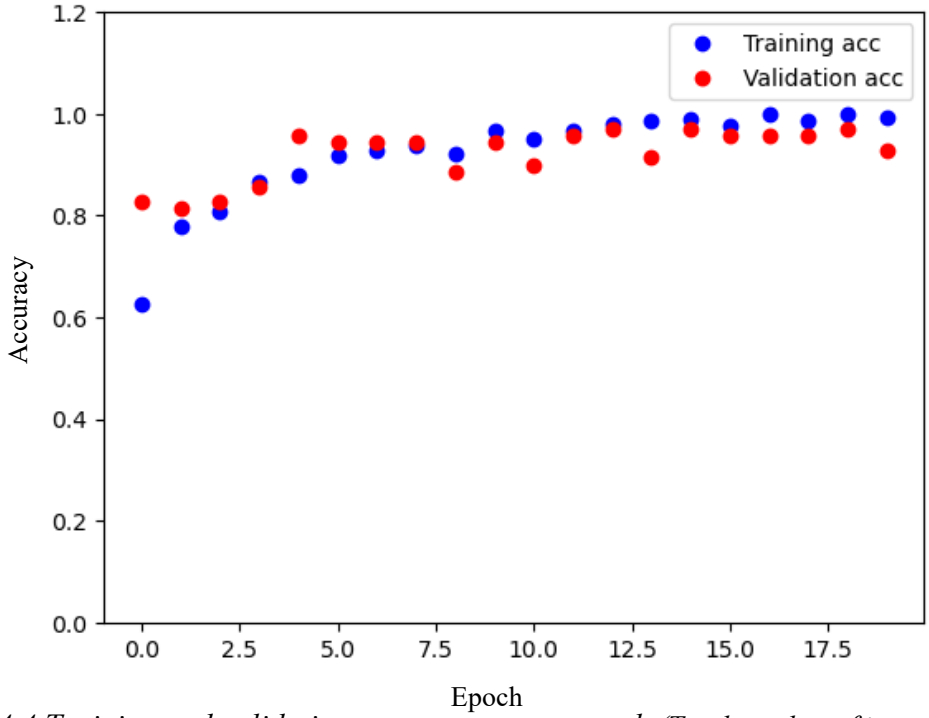
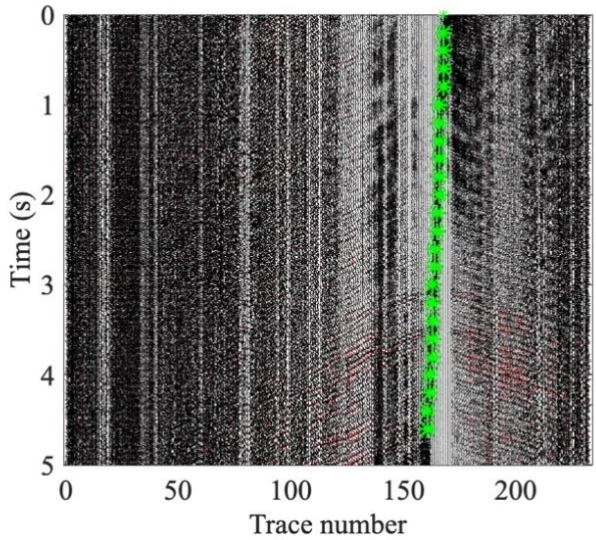


Figure 4.4 Training and validation accuracy versus epoch (Total number of iterations of all the training data in one cycle of training)

We then applied the implemented algorithm to the 24,600 remaining gathers, and identified 5676 gathers with the presence of truck passage. For each of these gathers we computed energy using Eq.1 every 0.1 seconds and we identified maximum energy (Figure 4.5), which corresponds to the truck position.



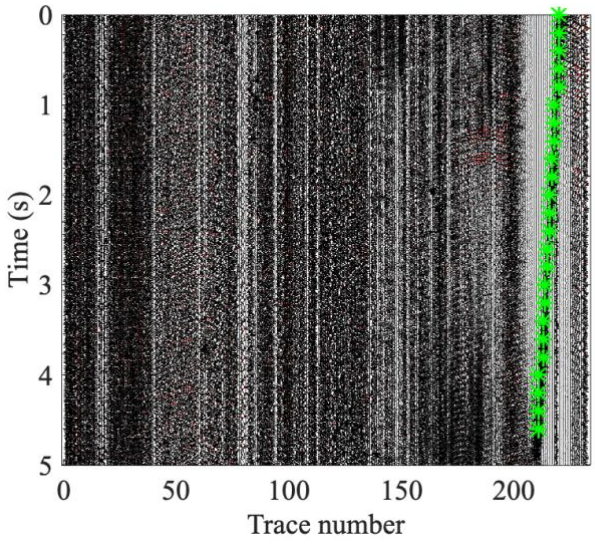


Figure 4.5 The green asterisks are representing the maximum energy picked in each window

Figure 4.6 shows that in the case that several truck passage are present within same window, we can correctly identify the event with the highest energy.

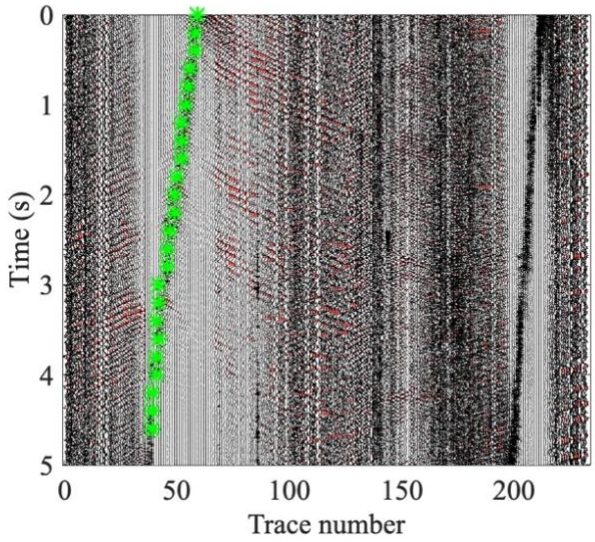


Figure 4.6 Between two truck passage within the same window the one with higher energy is picked, the maximum energy picked in window is marked with the green asterisks

The closest receiver to the location of the truck was used as the virtual source. Then, the trace corresponding to the virtual source was cross-correlated with all other receivers resulting in a virtual shot gather. In Figure 4.7.a we show an example of a record with the truck passing by, where we can see the sign of the truck passage between traces 164 and 171. The virtual gather obtained from interferometry using the 171st trace as a virtual source and its corresponding active shot gather are shown in Figure 4.7.b and Figure 4.7.c.

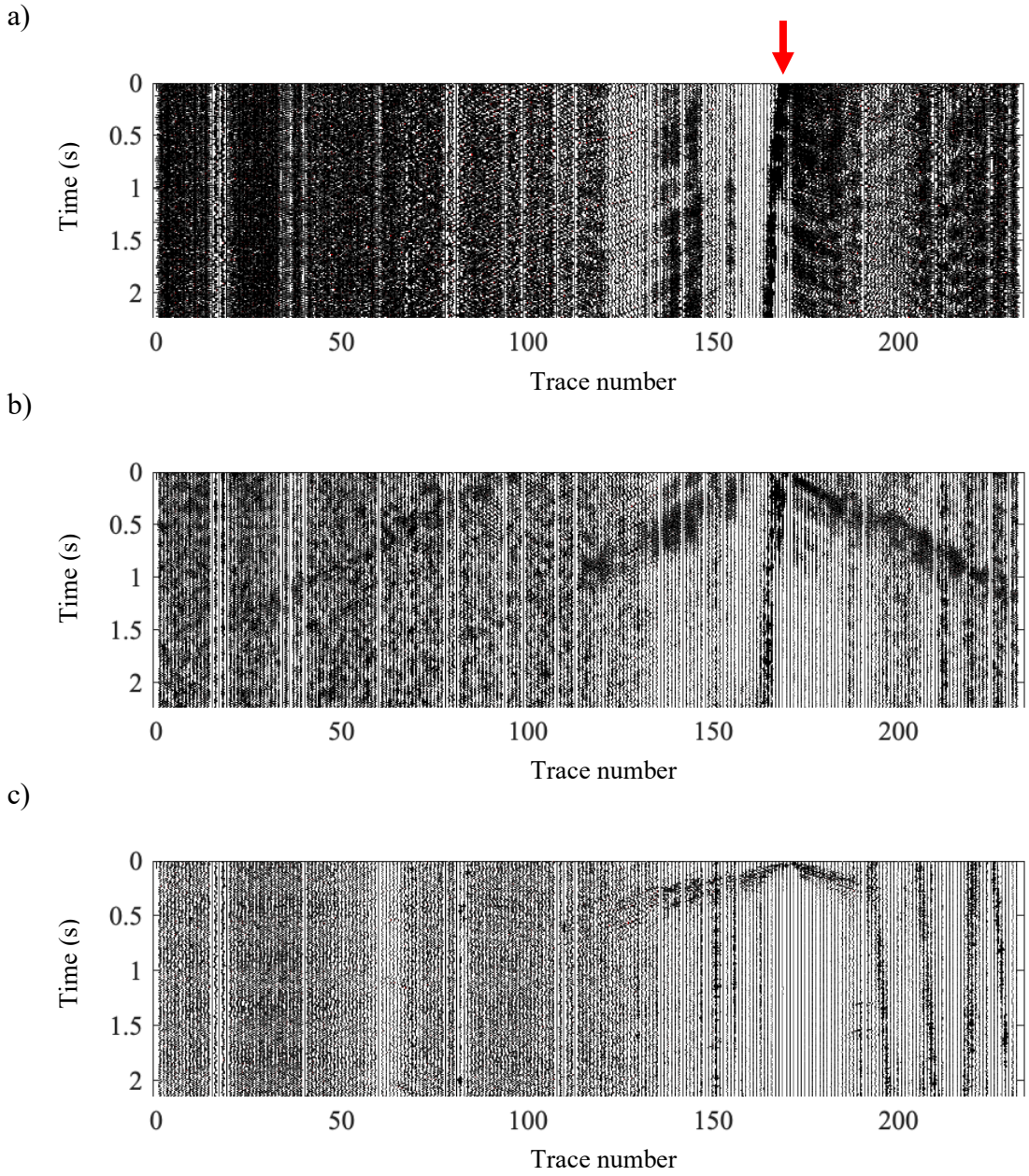


Figure 4.7 a) An example of recording with a truck passage. The red arrow shows the most energetic trace that is used as the virtual source b) The virtual shot gather using seismic interferometry c) An example of active recordings where the position of the active source is near to virtual source in plot(a)

We defined spatially moving window of 12 receivers, moving every 1 receiver along the line. For each spatial window, 30 closest virtual shots were chosen, and the dispersion image related to each shot was computed using the phase shift method. The individual dispersion images were stacked and the DC was picked as maxima in the dispersion image (Figure 4.8). Then, we repeated this process for all 221 spatial windows (resulting to the computation of 6630 virtual shot gathers).

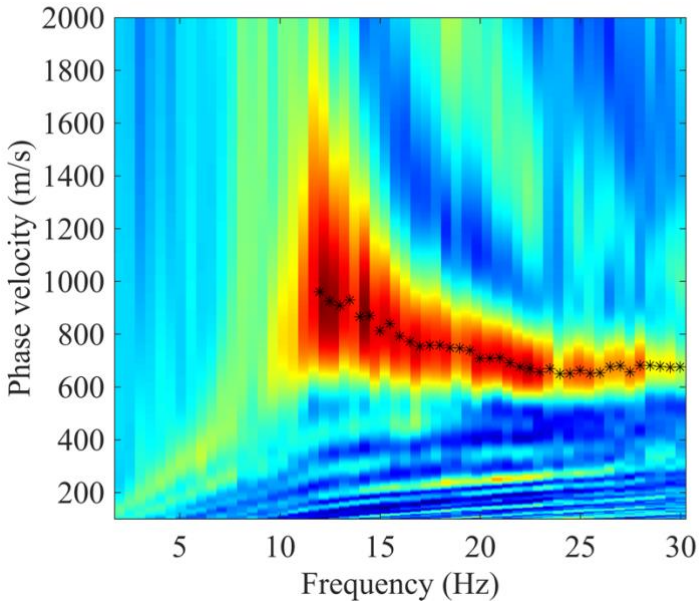
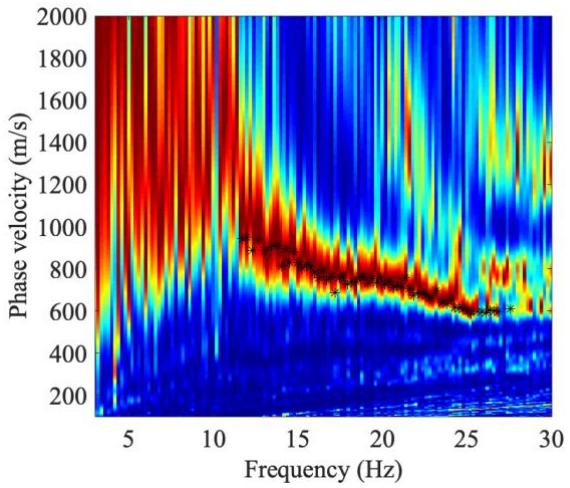


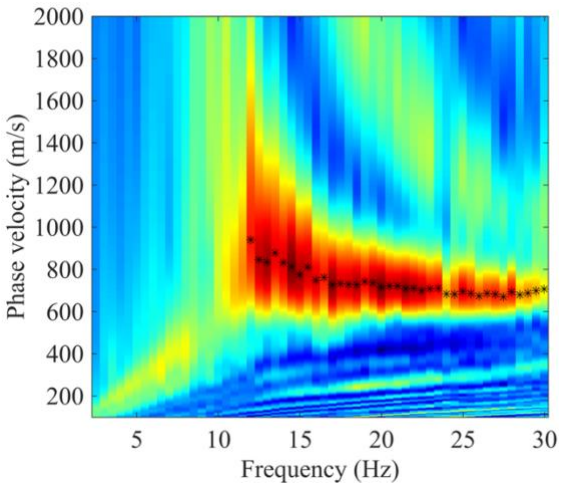
Figure 4.8 Example of picked DC as a maxima in the dispersion image corresponding to the spatial window from receiver 144 to 155 using the proposed method

To evaluate the quality of the estimated DCs from the proposed method, we compared them with the DCs from the active data that have been picked independently. The source was a weight drop for active acquisition, with source spacing equal to 10 m. (Figure 4.9.a, Figure 4.9.b, Figure 4.9.c, Figure 4.9.d, Figure 4.9.e, and Figure 4.9.f).

a)



b)



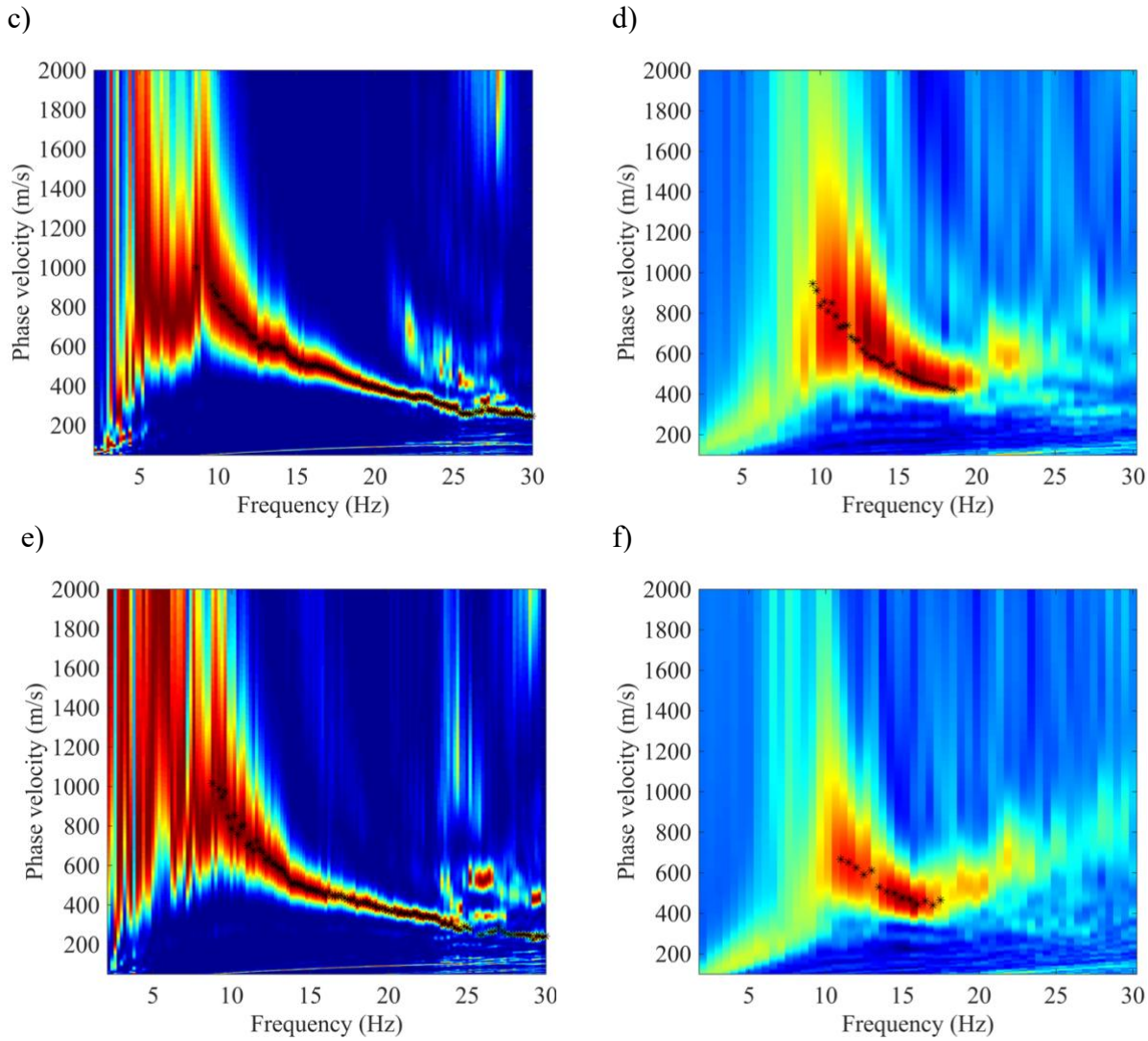
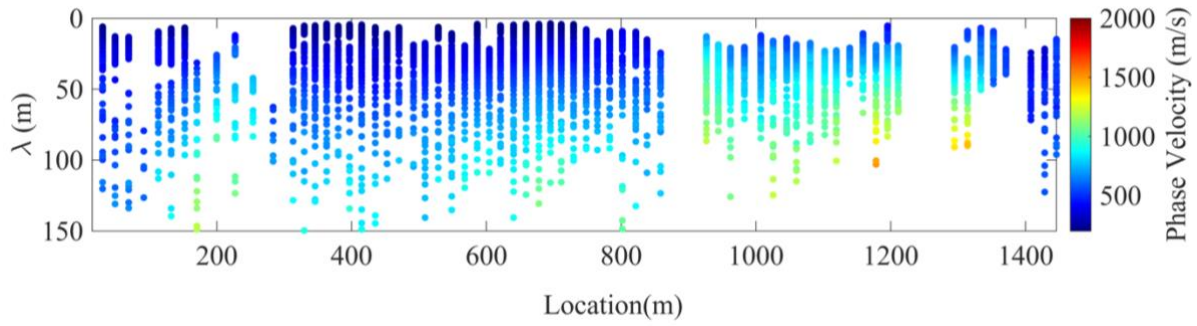


Figure 4.9 Computed dispersion image corresponding to the spatial window from receiver 143 to 154 using: a) active data, b) the proposed method. Receiver 100 to 111 using: c) active data, d) the proposed method. Receiver 103 to 114 using: e) active data, f) the proposed method. Black asterisks show the estimated DCs.

In Figure 4.10.a and Figure 4.10.b, we show the pseudo-sections of DCs obtained from active data, acquired along the same line using a weight drop source, with the proposed method. Each DC is plotted according to the location of the center of the corresponding window, the vertical axis shows the wavelength, the horizontal axis is the location, and the phase velocity is given by colormap.

a)



b)

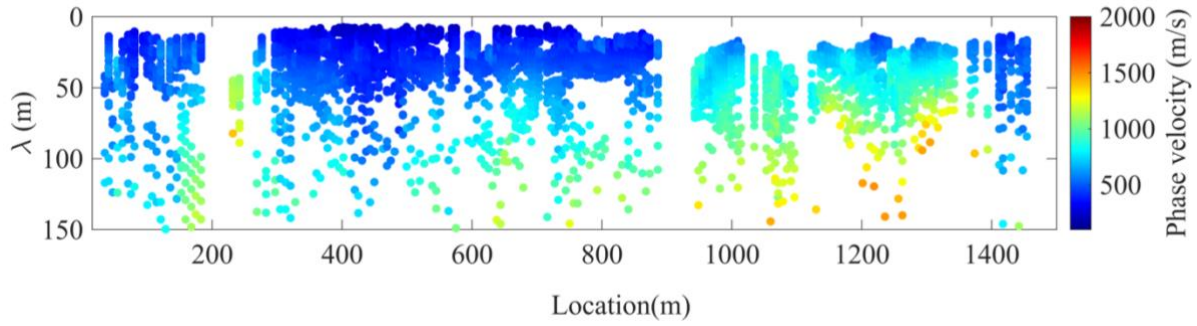


Figure 4.10 a) Pseudo-section of the DCs for active data along the line. b) Pseudo-section of the DCs for the proposed

As we can see in Figure 4.10.a and Figure 4.10.b, the DCs are generally consistent, however, there are some differences. The number of DCs retrieved from active data is 67 compared to 184 DCs retrieved from the proposed method. In certain zones of the line (see for instance 750 to 830 m), the DCs obtained from interferometry have narrower bandwidth at low frequency, while from 900 to 1500 m, they are broader band with respect to the DCs obtained from active data and from about 1230 to 1300, the passive data fill a gap in active data .

4.4 QUALITY CONTROL

To asses the quality of the picked DCs from passive data, we compared DCs from the passive data with their corresponding active data quantitatively. We computed normalized misfit (M_i) between ith passive and active DCs as:

$$M_i = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{V_{p_j} - V_{a_j}}{V_{a_j}}\right| \qquad 4.1$$

where n represents the number of common frequencies between the ith active and passive DCs, V_{p_j} and V_{a_j} are the velocity of jth frequency in the ith passive and the active DC respectively. We calculated the normalized total misfit (M_{tot}) existing between all the DCs as:

$$M_{tot} = \frac{1}{k}\sum_{i=1}^{k} M_i \qquad 4.2$$

where k is the number of locations that DCs are available for both active and passive data. We calculated the normalized total misfit (M_{tot}) existing between all the DCs as using Eq. 4.2. The calculated normalized total misfit is 17.7%. Vehicles are not always in line with the receiver array, this leads to phase velocities overestimation that could be a reason for this relatively high misfit (You et al., 2023).

CHAPTER 5

CONCLUSION

5.1 FINAL REMARK

In our research, we successfully demonstrated that vehicles moving along the seismic line can serve as seismic sources, providing valuable data for subsurface characterization. The data were acquired along a road where trucks were passing by in Kefalonia, Greece for a period of 2 days. We implemented a ML algorithm capable of accurately identifying each time window that contains truck passage. This algorithm efficiently tracks vehicles positions along the seismic line (Khoveiledy et al., 2023).

Using the closest receiver to the source as a virtual source, we effectively simulated the presence of an impulsive source at the location of each identified vehicle. This enabled us to reconstruct the wavefield response as if it were generated by an active seismic source. To generate virtual shot gathers, we employed seismic interferometry techniques.

By utilizing the phase shift method, we calculated dispersion images that captured the variation of surface wave velocities across the surveyed area. DCs were picked as maxima in the stacked dispersion image. These DCs provide crucial information about the subsurface properties and allow us to infer the structural characteristics beneath the seismic line.

Comparing the estimated DCs derived from the computed dispersion images of the vehicle-induced noise data with those obtained from active seismic sources, we established a consistent pattern. This finding indicated that the information obtained from the vehicle-induced noise data was reliable and comparable to that obtained through traditional active seismic surveys.

5.2 SUGGESTIONS FOR FUTURE DEVELOPMENT

To decrease the misfit between the active and passive data directional noise effect could be attenuated. Only signals could be used that are in the stationary phase zone.

In this work the DCs are picked manually, to make the processing steps fully automatic, ML-Based DC Picking can also be used. Utilizing ML techniques to automatically pick the DCs from the computed dispersion images. Training a supervised ML model using a large dataset of manually picked DCs. The model can learn the patterns and features indicative of DCs and then automatically pick them from new datasets.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., & Devin, M. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *ArXiv Preprint ArXiv:1603.04467*.

Anbazhagan, P., Sheikh, M. N., & Parihar, A. (2013). Influence of Rock Depth on Seismic Site Classification for Shallow Bedrock Regions. *Natural Hazards Review*, *14*(2), 108–121. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000088

Behm, M., Leahy, G. M., & Snieder, R. (2014). Retrieval of local surface wave velocities from traffic noise - an example from the La Barge basin (Wyoming). *Geophysical Prospecting*, *62*(2), 223–243. https://doi.org/10.1111/1365-2478.12080

Behm, M., & Snieder, R. (2013). Love waves from local traffic noise interferometry. *The Leading Edge*, *32*(6), 628–632. https://doi.org/10.1190/tle32060628.1

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

Benson, S. M., & Cole, D. R. (2008). CO2 Sequestration in Deep Sedimentary Formations. *Elements*, *4*(5), 325–331. https://doi.org/10.2113/gselements.4.5.325

Bhattacharya, S., De Risi, R., Lombardi, D., Ali, A., Demirci, H. E., & Haldar, S. (2021). On the seismic analysis and design of offshore wind turbines. *Soil Dynamics and Earthquake Engineering*, *145*, 106692. https://doi.org/10.1016/j.soildyn.2021.106692

Cheng, F., Xia, J., Luo, Y., Xu, Z., Wang, L., Shen, C., Liu, R., Pan, Y., Mi, B., & Hu, Y. (2016). Multichannel analysis of passive surface waves based on crosscorrelations. *GEOPHYSICS*, *81*(5), EN57–EN66. https://doi.org/10.1190/geo2015-0505.1

Cheng, F., Xia, J., Xu, Y., Xu, Z., & Pan, Y. (2015). A new passive seismic method based on seismic interferometry and multichannel analysis of surface waves. *Journal of Applied Geophysics*, *117*, 126–135. https://doi.org/10.1016/j.jappgeo.2015.04.005

References

Colombero, C., Comina, C., & Valentina Socco, L. (2019). Imaging near-surface sharp lateral variations with surface-wave methods — Part 1: Detection and location. *GEOPHYSICS*, *84*(6), EN93–EN111. https://doi.org/10.1190/geo2019-0149.1

Curtis, A., Gerstoft, P., Sato, H., Snieder, R., & Wapenaar, K. (2006). Seismic interferometry—turning noise into signal. *The Leading Edge*, *25*(9), 1082–1092. https://doi.org/10.1190/1.2349814

Dou, S., Lindsey, N., Wagner, A. M., Daley, T. M., Freifeld, B., Robertson, M., Peterson, J., Ulrich, C., Martin, E. R., & Ajo-Franklin, J. B. (2017). Distributed Acoustic Sensing for Seismic Monitoring of The Near Surface: A Traffic-Noise Interferometry Case Study. *Scientific Reports*, *7*(1), 11620. https://doi.org/10.1038/s41598-017-11986-4

Earle, S. (2015). *Physical geology*. BCcampus.

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

Furre, A.-K., Meneguolo, R., Pinturier, L., & Bakke, K. (2020). Planning deep subsurface CO2 storage monitoring for the Norwegian full-scale CCS project. *First Break*, *38*(10), 55–60. https://doi.org/10.3997/1365-2397.fb2020074

Grandjean, G., & Bitri, A. (2006). 2M-SASW: Multifold multichannel seismic inversion of local dispersion of Rayleigh waves in laterally heterogeneous subsurfaces: application to the Super-Sauze earthflow, France. *Near Surface Geophysics*, *4*(6), 367–375. https://doi.org/10.3997/1873-0604.2006010

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, *77*, 354–377. https://doi.org/10.1016/j.patcog.2017.10.013

Halliday, D., Curtis, A., & Kragh, E. (2008). Seismic surface waves in a suburban environment: Active and passive interferometric methods. *The Leading Edge*, *27*(2), 210–218. https://doi.org/10.1190/1.2840369

Huot, F., & Biondi, B. (2018). 'Machine learning algorithms for automated seismic ambient noise processing applied to DASacquisition'. *SEG Technical Program Expanded Abstracts 2018*, 5501–5505. https://doi.org/10.1190/segam2018-w20-03.1

Huot, F., Ma, Y., Cieplicki, R., Martin, E., & Biondi, B. (2017). Automatic noise exploration in urban areas. *SEG Technical Program Expanded Abstracts 2017*, 5027–5032. https://doi.org/10.1190/segam2017-17774369.1

Khoveiledy, T., Karimpour, M., & Socco, V. (2023). Using vehicle-induced noise as local seismic source through ML and interferometry. *NSG Conference & Exhibition* .

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Lehujeur, M., Vergne, J., Maggi, A., & Schmittbuhl, J. (2017). Ambient noise tomography with non-uniform noise sources and low aperture networks: case study of deep geothermal reservoirs in northern Alsace, France. *Geophysical Journal International*, *208*(1), 193–210. https://doi.org/10.1093/gji/ggw373

Martin, E. R., Huot, F., Ma, Y., Cieplicki, R., Cole, S., Karrenbach, M., & Biondi, B. L. (2018). A Seismic Shift in Scalable Acquisition Demands New Processing: Fiber-Optic Seismic Signal Retrieval in Urban Areas with Unsupervised Learning for Coherent Noise Removal. *IEEE Signal Processing Magazine*, *35*(2), 31–40. https://doi.org/10.1109/MSP.2017.2783381

Mi, B., Xia, J., Tian, G., Shi, Z., Xing, H., Chang, X., Xi, C., Liu, Y., Ning, L., Dai, T., Pang, J., Chen, X., Zhou, C., & Zhang, H. (2022). Near-surface imaging from traffic-induced surface waves with dense linear arrays: An application in the urban area of Hangzhou, China. *GEOPHYSICS*, *87*(2), B145–B158. https://doi.org/10.1190/geo2021-0184.1

Mi, B., Xia, J., Xu, Y., You, B., & Chen, Y. (2023). Retrieval of surface waves from high-speed-train-induced vibrations using seismic interferometry. *GEOPHYSICS*, *88*(5), KS113–KS126. https://doi.org/10.1190/geo2022-0603.1

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 807–814.

Nakata, N., Snieder, R., Tsuji, T., Larner, K., & Matsuoka, T. (2011). Shear wave imaging from traffic noise using seismic interferometry by cross-coherence. *Geophysics*, *76*(6), SA97–SA106.

Nasseri-Moghaddam, A., Cascante, G., & Hutchinson, J. (2005). A New Quantitative Procedure to Determine the Location and Embedment Depth of a Void Using Surface Waves. *Journal of Environmental and Engineering Geophysics*, *10*(1), 51–64. https://doi.org/10.2113/JEEG10.1.51

References

Neducza, B. (2007). Stacking of surface waves. *GEOPHYSICS*, *72*(2), V51–V58. https://doi.org/10.1190/1.2431635

Pang, J., Cheng, F., Shen, C., Dai, T., Ning, L., & Zhang, K. (2019). Automatic passive data selection in time domain for imaging near-surface surface waves. *Journal of Applied Geophysics*, *162*, 108–117. https://doi.org/10.1016/j.jappgeo.2018.12.018

Park, C. B., & Miller, R. D. (2008). Roadside Passive Multichannel Analysis of Surface Waves (MASW). *Journal of Environmental and Engineering Geophysics*, *13*(1), 1–11. https://doi.org/10.2113/JEEG13.1.1

Park, C. B., Miller, R. D., & Xia, J. (1999). Multichannel analysis of surface waves. *Geophysics*, *64*(3), 800–808.

Pérez-Collazo, C., Greaves, D., & Iglesias, G. (2015). A review of combined wave and offshore wind energy. *Renewable and Sustainable Energy Reviews*, *42*, 141–153. https://doi.org/10.1016/j.rser.2014.09.032

Quiros, D. A., Brown, L. D., & Kim, D. (2016). Seismic interferometry of railroad induced ground motions: body and surface wave imaging. *Geophysical Journal International*, *205*(1), 301–313. https://doi.org/10.1093/gji/ggw033

Ranzato, M., Huang, F. J., Boureau, Y.-L., & LeCun, Y. (2007). Unsupervised Learning of Invariant Feature Hierarchies with Applications to Object Recognition. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 1–8. https://doi.org/10.1109/CVPR.2007.383157

Richart, F. E., Hall, J. R., & Woods, R. D. (1970). *Vibrations of soils and foundations*.

van Manen, D.-J., Robertsson, J. O. A., & Curtis, A. (2005). Modeling of Wave Propagation in Inhomogeneous Media. *Physical Review Letters*, *94*(16), 164301. https://doi.org/10.1103/PhysRevLett.94.164301

Vidal, C. A., Draganov, D., van der Neut, J., Drijkoningen, G., & Wapenaar, K. (2014). Retrieval of reflections from ambient noise using illumination diagnosis. *Geophysical Journal International*, *198*(3), 1572–1584. https://doi.org/10.1093/gji/ggu164

Wang, X., Zeng, X., Li, J., Yang, X., & Wang, H. (2018). A review on recent advancements of substructures for offshore wind turbines. *Energy Conversion and Management*, *158*, 103–119. https://doi.org/10.1016/j.enconman.2017.12.061

Wapenaar, K. (2004). Retrieving the Elastodynamic Green's Function of an Arbitrary Inhomogeneous Medium by Cross Correlation. *Physical Review Letters*, *93*(25), 254301. https://doi.org/10.1103/PhysRevLett.93.254301

References

Wapenaar, K., Draganov, D., Snieder, R., Campman, X., & Verdel, A. (2010). Tutorial on seismic interferometry: Part 1 — Basic principles and applications. *GEOPHYSICS*, *75*(5), 75A195-75A209. https://doi.org/10.1190/1.3457445

Wapenaar, K., & Fokkema, J. (2006). Green's function representations for seismic interferometry. *GEOPHYSICS*, *71*(4), SI33–SI46. https://doi.org/10.1190/1.2213955

Wapenaar, K., Slob, E., Snieder, R., & Curtis, A. (2010). Tutorial on seismic interferometry: Part 2 — Underlying theory and new advances. *GEOPHYSICS*, *75*(5), 75A211-75A227. https://doi.org/10.1190/1.3463440

Xia, J., Miller, R. D., & Park, C. B. (1999). Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *GEOPHYSICS*, *64*(3), 691–700. https://doi.org/10.1190/1.1444578

You, B., Mi, B., Guan, B., Zhang, H., & Liu, Y. (2023). High-quality surface wave retrieval from vibrations generated by high-speed trains moving on viaducts. *Journal of Applied Geophysics*, *212*, 105005. https://doi.org/10.1016/j.jappgeo.2023.105005

Zhang, K., Li, H., Wang, X., & Wang, K. (2020). Retrieval of shallow S-wave profiles from seismic reflection surveying and traffic-induced noise. *GEOPHYSICS*, *85*(6), EN105–EN117. https://doi.org/10.1190/geo2019-0845.1