Nondestructive testing of railway embankments by measuring multi-modal dispersion of surface waves induced by high-speed trains with linear geophone arrays

José Cunha Teixeira, Ludovic Bodet, Marine Dangeard, Alexandrine Gesret, Amélie Hallier, Agnès Rivière, Audrey Burzawa, Julio José Cárdenas Chapellín, Marie Fonda, Ramon Sanchez Gonzalez, et al.

doi:10.26443/seismica.v4i1.1150

Nondestructive testing of railway embankments by measuring multi-modal dispersion of surface waves induced by high-speed trains with linear geophone arrays

José Cunha Teixeira * [1,2,3], **Ludovic Bodet** [1], **Marine Dangeard** [2], **Alexandrine Gesret** [3], **Amélie Hallier** [2], **Agnès Rivière** [3], **Audrey Burzawa** [1,2], **Julio José Cárdenas Chapellín** [1], **Marie Fonda** [2], **Ramon Sanchez Gonzalez** [3], **Amine Dhemaied** [2], **Joséphine Boisson Gaboriau** [2]

[1]Sorbonne Universite, CNRS, EPHE, UMR 7619 METIS, 4 place Jussieu, 75252 Paris 05, France, [2]SNCF Réseau, 6 avenue François Mitterrand, 93210 Saint-Denis, France, [3]Geosciences Department, Mines Paris - PSL, PSL University, Paris, France

Author contributions: *Conceptualization*: J. Cunha Teixeira, L. Bodet. *Methodology*: J. Cunha Teixeira. *Software*: J. Cunha Teixeira. *Validation*: J. Cunha Teixeira, L. Bodet. *Formal Analysis*: J. Cunha Teixeira. *Investigation*: J. Cunha Teixeira, L. Bodet, M. Dangeard, A. Burzawa, J. J. Cárdenas Chapellín, M. Fonda, R. Sanchez Gonzalez, A. Dhemaied. *Writing - Original draft*: J. Cunha Teixeira, L. Bodet. *Supervision*: L. Bodet, A. Gesret, A. Rivière, A. Hallier. *Project administration*: L. Bodet, A. Dhemaied, J. Boisson Gaboriau.

Abstract The characterization, imaging, and monitoring of mechanical properties in engineered structures are crucial for ensuring safety and maintaining the reliability of operations and services. This study leverages well-constrained experimental sites to assess the viability of a passive-Multichannel Analysis of Surface Waves technique for the characterization of railway embankments, utilizing seismic waves induced by high-speed train traffic instead of conventional active sources. Seismic waves were recorded from the side of the railway track at three different sites using a uniform linear array of 96 geophones. These recordings, each with a maximum duration of 120 s, were segmented into very short segments that underwent an automatic sorting process based on a Frequency-Wavenumber transform, estimating the direction of wave propagation and the relative position of the source. Only segments where the train is in close proximity to the array were retained and categorized into 'train at the left' and 'train at the right'. This information was then utilized to select between the causal and acausal parts of the cross-correlations when generating virtual shot-gathers through interferometry. A phase-weighted stack was applied to the entire set of virtual shot-gathers, and dispersion images were computed using a phase-shift transform. This process enabled the extraction of various modes of propagation by identifying the maxima in the dispersion images. They show a high multi-modal character which is discussed before exploring array density limits to ensure reliable results. This method enabled a precise characterization of the shear-wave velocity of thin layers with critical importance to the mechanical stability of the railway embankment, highlighting its potential for enhancing the mechanical testing and monitoring of railway embankments.

Production Editor: Gareth Funning
Handling Editor: Suzan van der Lee
Copy & Layout Editor: Théa Ragon

Received: November 11, 2024
Accepted: December 2, 2024
Published: January 13, 2025

Acronyms

V_P	Pressure-wave velocity
V_R	Rayleigh-wave phase velocity
V_S	Shear-wave velocity
DAS	Distributed acoustic sensor
DC	Dispersion curve
DCP	Dynamic cone penetrometer
DI	Dispersion image
ERT	Electrical Resistivity Tomography
FK	Frequency-Wavenumber
GPR	Ground Penetrating Radar
HSL	High-speed line
HST	High-speed train
MASW	Multichannel Analysis of Surface Waves
PWS	Phase-weighted stack
RE	Railway embankment
SASW	Spectral Analysis of Surface Waves
SG	Shot-gather
SW	Surface-wave
T1	Track 1
T2	Track 2
ULA	Uniform linear array
VSG	Virtual shot-gather

*Corresponding author: jose.teixeira@sorbonne-universite.fr

1 Introduction

The continuous expansion of urban areas, combined with the ongoing environmental changes, has increased the vulnerability of those regions to a variety of natural hazards that impact the near-surface proprieties. In this sense, railway companies face a critical need to characterize the stiffness of railway embankments (REs) in order to ensure the safety and reliability of

rail traffic (Rhayma et al., 2013). Traditionally, this assessment involves periodic integrity checks using *in situ* tests, such as boreholes and penetrometers (Okada and Ghataora, 2002; Haddani et al., 2016). More recently, non-destructive geophysical techniques have gained prominence for their ease, speed, and cost-effectiveness in characterizing REs across extensive distances (Artagan et al., 2020). A combination of several methods is typically employed, with the most commonly used techniques being Ground Penetrating Radar (GPR) (Narayanan et al., 2004; Donohue et al., 2011; Fontul et al., 2016), Electrical Resistivity Tomography (ERT) (Barta, 2010; Donohue et al., 2011; Fontul et al., 2016), Microgravimetry (Barta, 2010; Nebieridze, 2011; Fontul et al., 2016), and seismic approaches such as Seismic Refraction (Barta, 2010), Seismic Reflection (Fontul et al., 2016) and surface-wave (SW) methods.

In near-surface applications, SW methods predominantly rely on the study of the propagation of Rayleigh-waves (Rayleigh, 1885), which stand out for their ease of detection within the seismic wavefield thanks to their limited geometric attenuation when compared with other types of seismic waves. SWs are typically manually generated by punctual active sources (e.g. hammer, weight-drop) and are recorded either by a pair of geophones in the Spectral Analysis of Surface Waves (SASW) method (Nazarian et al., 1983) or by a uniform linear array (ULA) of multiple geophones in the Multichannel Analysis of Surface Waves (MASW) method (Park et al., 1999; Xia et al., 1999). The latter involves transforming a shot-gather (SG), initially recorded in the distance-time domain, into a dispersion image (DI) in the frequency-phase velocity domain (McMechan and Yedlin, 1981; Mokhtar et al., 1988; Park et al., 1998). The maxima in the DI correspond to a set of dispersion curves (DCs), representing the variation of the Rayleigh-wave phase velocity (V_R) over frequency, for different modes of propagation. V_R variation over frequency is closely linked to the medium's Shear-wave velocity (V_S) variation over depth, which, in turn, is related to the medium's modulus of rigidity at different depths. Thus, MASW is utilized to infer the 1D V_S of subsurface structures (Gabriels et al., 1987; Jongmans and Demanet, 1993; Park et al., 1999; Lai et al., 2002; O'Neill et al., 2003; Socco and Strobbia, 2004; Socco et al., 2010; Bergamo and Socco, 2016; Foti et al., 2018), or to generate pseudo-2D V_S profiles (Bohlen et al., 2004; Socco and Strobbia, 2004; Socco et al., 2009, 2010; Boiero and Socco, 2010, 2011; Bergamo et al., 2012; Pasquet and Bodet, 2017).

In railway applications, MASW has for instance proven rapid and reliable for assessing underground stability of REs on conventional lines affected by landslides (Donohue et al., 2011), sinkholes (Boisson-Gaboriau et al., 2016), or various climate effects over time (Bergamo et al., 2016). More recently, MASW has been adapted for use in a high-speed lines (HSLs) to obtain a more refined analysis and control of the thin subsurface layers supporting the track (Bodet et al., 2017; Bodet, 2019; Burzawa et al., 2023). However, many investigations face limitations due to operational constraints and access restrictions, making continuous or periodic characterizations of REs difficult to deploy with active sources.

In contrast, passive source SW methods, known as seismic interferometry or Green's function retrieval by crosscorrelation, use coherent ambient seismic noise generated by natural or anthropogenic activities to retrieve SW signals. Firstly theorized by (Aki, 1957), these methods have been widely developed over the last decades for the field of seismology (Derode et al., 2003; Snieder, 2004; Wapenaar, 2004; Weaver and Lobkis, 2004; Bakulin and Calvert, 2006; Wapenaar et al., 2010a,b; Larose et al., 2015; Boué and Paul, 2020), and play a crucial role in tomographic tools for characterizing deep structures in continental-scale applications with large 2D arrays. Seismic noise is typically generated by natural phenomena such as microseisms (Ritzwoller and Levshin, 1998; Barmin et al., 2001; Shapiro and Campillo, 2004; Yao et al., 2006; Bensen et al., 2007, 2008), but researchers have also harnessed anthropogenic sources, such as oil platforms (Mordret et al., 2013), car traffic (Behm et al., 2014), and even heavy freight trains (Brenguier et al., 2019; Sheng et al., 2023).

In various near-surface applications, passive source SWs have been used for subsurface characterization in a scheme of passive-MASW with 1D ULAs. Firstly without interferometry along roadsides (Park et al., 2004; Park and Miller, 2008), and then with interferometry along roadsides (O'Connell and Turner, 2011; Xu et al., 2013; Le Feuvre et al., 2015; Hayashi et al., 2015; Cheng et al., 2015, 2016; Mi et al., 2022; Czarny et al., 2023), as well as along railway tracks (Quiros et al., 2016; Rezaeifar et al., 2023).

In railway applications, passive-MASW has been used on conventional lines for continuous monitoring of sinkhole development beneath tracks. Even though the active method may be advantageous for one-time characterization, as it does not require long-term train recordings or automated data collection, the passive method is better suited for continuous or repeated monitoring since it eliminates the need for field interventions. Applications have involved the use of dense 2D geophone arrays and extended recording periods of ambient noise (Talfumière and Nebieridze, 2008), and of SWs induced by trains (Bardainne and Rondeleux, 2018; Bardainne et al., 2022; Tarnus et al., 2022b,a; Bardainne et al., 2023a,b; Rebert et al., 2024b). More recently, 1D ULAs have also been employed to record SWs induced by trains, with the aim of assessing shallow structures beneath viaducts (You et al., 2023; Mi et al., 2023). However, these applications focus on frequencies ranging from 5 to 50 Hz and do not prioritize the characterization of the very first meters of the RE.

In this study, we explore the feasibility of employing a passive-MASW approach on HSLs, utilizing a compact ULA and brief recordings of seismic waves generated by the passage of high-speed trains (HSTs). The purpose is to characterize the thin subsurface layers supporting the track for a broad band of frequencies (10 to 200 Hz). Tests were operated at three different sites along a french HSL, benefiting from robust constraints provided by previous geotechnical studies and geophysical surveys (Bodet et al., 2017; Bodet, 2019; Burzawa et al., 2023). After introducing the test sites and detail-

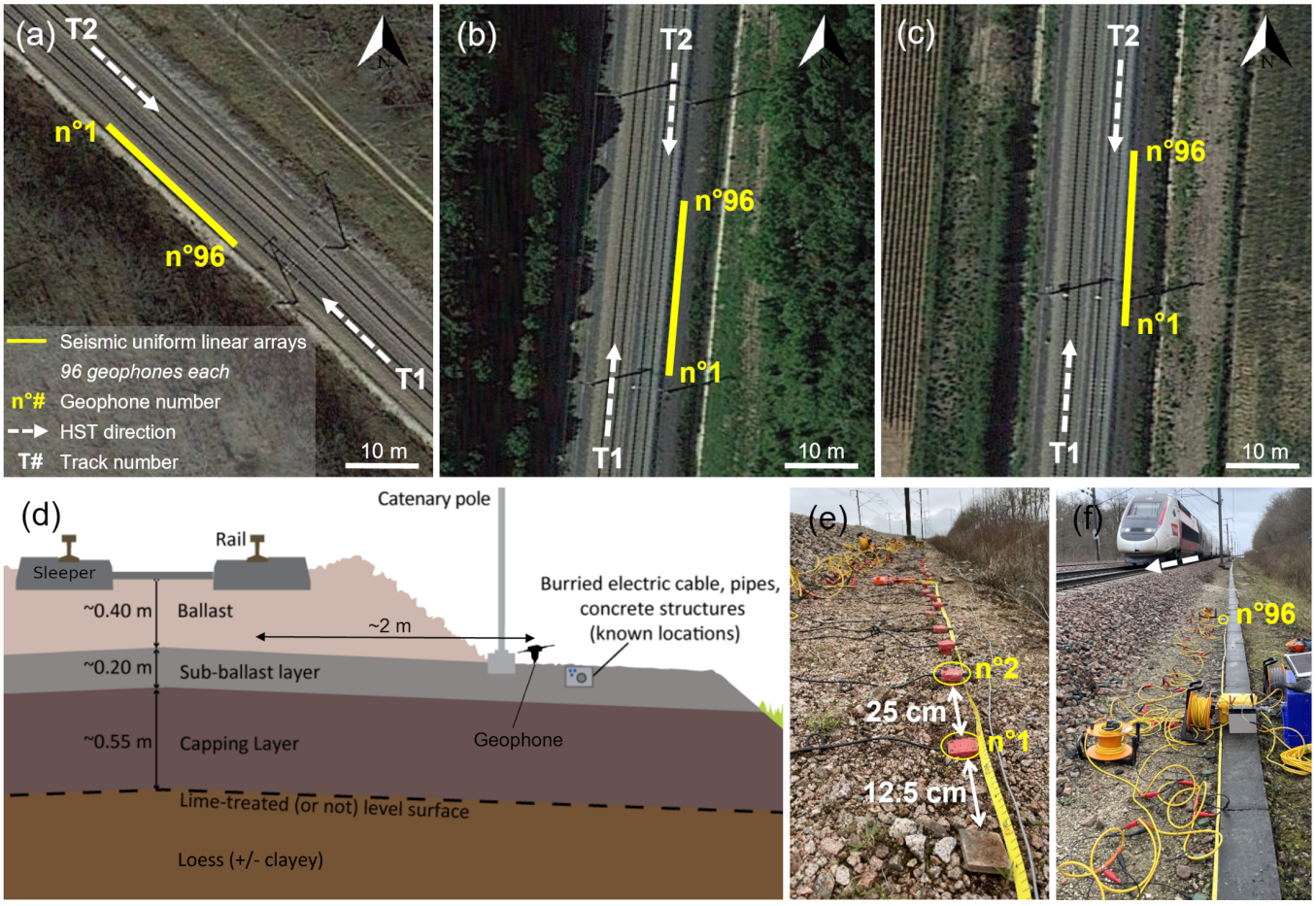

Figure 1 (a), (b), and (c) Schematic maps of sites A, B, and C, respectively. The seismic ULAs are represented by yellow lines with the position of the first and last geophones. Track numbers (T1 and T2) and train directions are also specified. (d) Schematic cross-section [adapted from Bodet et al. (2017) and Bodet (2019)] illustrating the structure of the studied RE, conforming to SNCF Réseau standards (IN32787, IG90260). (e) Setup used for active seismic acquisition, installed along the track side. (f) Similar setup as in (e), employed for recording HST passages, with an HST approaching from the right of the ULA.

ing the seismic survey geometry, we describe the processing method applied to select and categorize HST passage records according to the train travel direction and distance from the array. Subsequently, we present the workflow used to extract SW propagation from HST-induced seismic noise, and to derive DCs. We then present results obtained using this technique and compare them with the DCs obtained through traditional active MASW, acquired concurrently during the same experiment and with the exact same setup. Lastly, we discuss the richness of the passive data, the occurrence of higher modes on DIs and their impact on the inversion procedure, and the influence of the array density on the quality of dispersion retrieval.

2 Field deployment and data acquisition

Passive seismic tests were conducted at three distinct sites (referred to as sites A, B, and C in Figs. 1a-c) along a French HSL, between January and May 2023. These sites have experienced an abnormal increase in the need for maintenance over the past decade that could be attributed to significant variations in soil properties or inadequate water drainage. Geological and geotechnical surveys have shown that the REs exhibit a typical structure as defined by French HSL standards (refer to Fig. 1d), composed with a sub-ballast layer, very compact with an average thickness of 20 cm, a capping layer, less compact with an average thickness of 60 cm, and a loess layer, usually slightly less compact with an average thickness of 5 m. The platform rests on a more compact chalk substratum not represented in Figure 1d. These shallow, thin layers are crucial to the mechanical stability of the railway and require careful monitoring and control. No standard V_S values are defined universally for healthy RE, as these values can vary significantly depending on local geological and engineering conditions. However a decrease over time of V_S for the capping layer is generally considered indicative of potential issues with the stiffness or stability of the RE (Burzawa et al., 2023).

On each site, a seismic ULA with a length of 23.75 m and composed by 96 geophones spaced at intervals of 25 cm was installed. The ULAs were positioned on the the track side along the railway, at approximately 2 m from the rails, on Track 1 (T1), for site A, and Track 2 (T2), for sites B and C. It is important to note

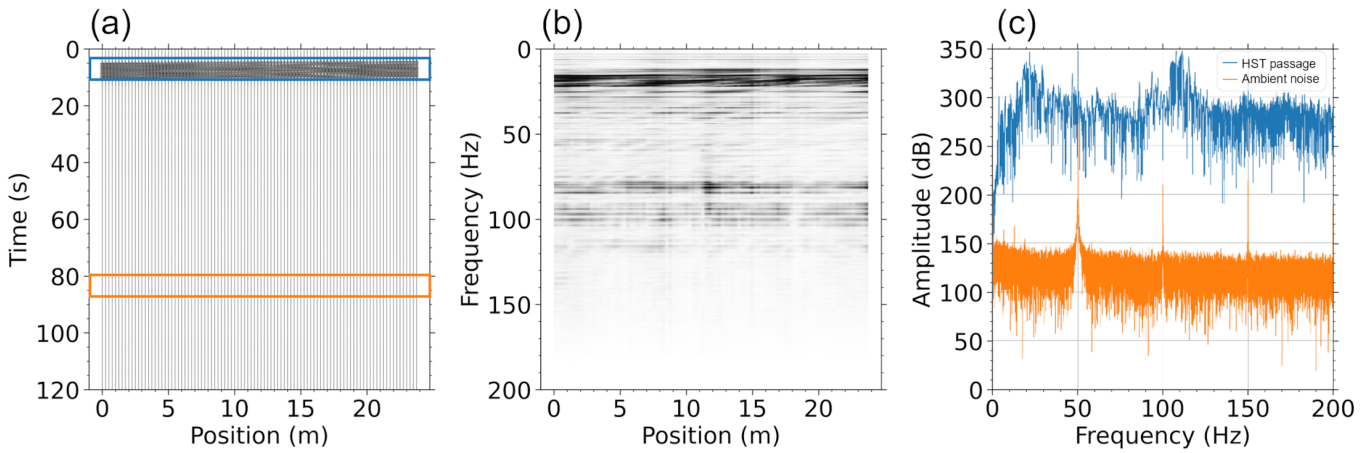

Figure 2 (a) Raw HST record. The blue rectangle corresponds to the time segments where the HST is passing. The orange rectangle corresponds to a calm segment with only ambient noise. (b) Normalized by frequency spectrogram of the same HST record. (c) Frequency power spectrum of the first geophone for the same selected periods during the HST passage (blue) and during a calm segment (orange). French HSLs are electrified with a 25 kV, 50 Hz alternating current. Therefore, the peaks observed at harmonics of 50 Hz correspond to stray currents originating from the traction system, which travel along the rails and through the soil, serving as the return conductor.

that the sites in question have no curvature and a zero slope. Active seismic data was acquired by vertically striking a metal plate placed in-line, 12.5 cm away from the first (direct shot) and last (inverse shot) geophones (see Fig. 1e). The impact was achieved with a hammer (1.5 kg) and waves propagation along the ULA was recorded for 2 s with a sampling interval of 0.5 ms (i.e. sampling rate of 2 kHz) and a pre-triggering delay of -0.02 s. Six direct and inverse shots were stacked in the time domain to enhance signal-to-noise ratio.

HST passages on tracks 1 and 2 were also recorded with the exact same ULAs for durations of 120 s (see Table 1) with a sampling interval of 2 ms (i.e. sampling rate of 500 Hz). The recordings were initiated manually upon visual confirmation of approaching trains at the site. Seismic waves generated by trains contain high energy for frequencies ranging from 10 to 200 Hz with an estimated signal-to-noise ratio of 150 dB (see Fig. 2).

3 Automatic data selection

Before any data transformation, each HST record is split into segments of 5 s by using a Tukey sliding window with a moving step of 1 s. However, not all segments are suitable for use in a passive-MASW scheme because interferometry relies on waves that travel in-line from a source and pass through all geophones to reconstruct an accurate impulse response of the propagation medium between pairs of geophones. Misalignment leads to incorrect V_R estimation (Park and Miller, 2008) and produces artifacts in DIs (Cheng et al., 2023). Configurations with coherent waves, such as train-induced waves, traveling simultaneously in opposite directions must also be avoided to not distort interferometry results. See Section 4.1 for more details about interferometry.

In this context, we established the 1st and 96th geophones as the left and right extremities of the ULA, respectively, with an increasing distance axis from left to right (see Figs. 1a-c). Thus, there are four distinct configurations between the train position and the ULA, each of which must be distinguished on every segment. Figure 3 shows HST records at sites B and C with some

Table 1 HST records on tracks 1 and 2 (T1 and T2) for sites A, B, and C with the sampling interval dt and record length Dt.

Record (#)	**Track**	**dt** (s)	**Dt** (s)
A-1	T1, T2	2	120
A-2	T1	2	120
A-3	T1, T2	2	120
A-4	T2, T1	2	120
A-5	T2	2	120
A-6	T1, T2	2	120
B-1	T1	2	120
B-2	T1	2	120
B-3	T2	2	120
B-4	T1	2	120
B-5	T1	2	120
B-6	T1, T2	2	120
B-7	No train recorded	2	120
B-8	T2	2	120
B-9	T2	2	120
B-10	T2, T1	2	120
C-1	T1	2	120
C-2	T1	2	120
C-3	T2	2	120
C-4	T2	2	120
C-5	T2	2	120
C-6	T1	2	120
C-7	T1	2	120
C-8	T2	2	120
C-9	T2	2	120
C-10	T2	2	120

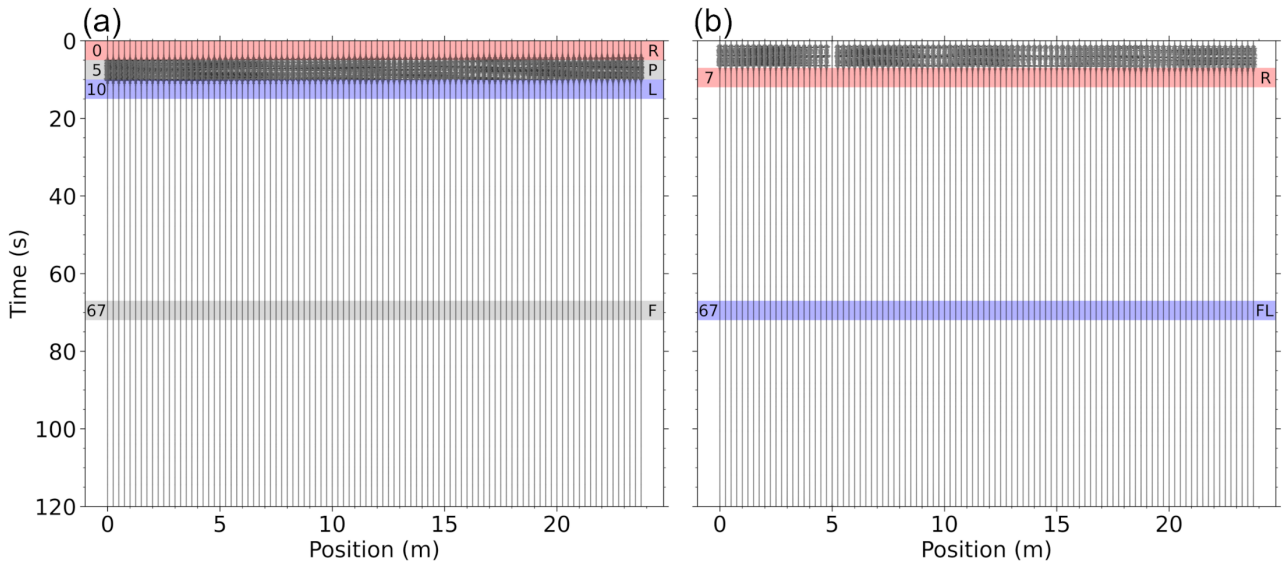

Figure 3 (a) 9th HST passage recorded on site B (record B-9 on Table 1) on T2 passing from the right to the left of the ULA. Four segments of 5 s at 0, 5, 10 and 67 s are shown and labeled according to the train's position (R for right, P for passage, L for left and F for far). (b) 7th HST passage recorded on site C (record C-7 on Table 1) on T1 passing from the left to the right of the ULA. Two segments of 5 s at 7 and 67 s are depicted and labeled according to the HST's position (R for right, and FL for far-left with an unknown signal detected coming from the left as shown in Fig. 4).

highlighted segments for each possible case:

- Case F: Train is far from the ULA. No SW signal is recorded by the geophones (this threshold distance depends on the types of sites and trains and is not estimated. Only the amount of energy recorded the the ULA's position is calculated).
- Case P: Train is passing by the ULA. The source and the geophones are not aligned in an unique direction.
- Case L: Train is close and at the left of the ULA. The source is aligned with all geophones. Waves have a rightward propagation from geophone n°1 to geophone n°96.
- Case R: Train close and at the right of the ULA. The source is aligned with all geophones. Waves have a leftward propagation from geophone n°96 to geophone n°1.

An anomalous case FL is also highlighted for a segment when the train is far yet a strong unknown signal is detected coming from the left.

In the literature, various techniques have been employed as automatic segment selection tools. Examples include Frequency-Wavenumber (FK)-Based data selection (Cheng et al., 2018), data selection in the Tau-p domain (Cheng et al., 2019), beamforming techniques (Ning et al., 2022), and covariance matrix methods for train localization in time and space (Rezaeifar et al., 2023). Other approaches utilize a single geophone positioned at a certain distance from both sides of the ULA to detect train arrivals and estimate Fresnel angles (You et al., 2023). In our study, we use an FK-based data selection technique adapted from Cheng et al. (2018) that suits well with a 1D linear geometry and is computationally efficient for segment classification.

To determine the correct propagation direction of the incoming waves, we estimate the FK spectrum in the negative and positive wavenumber domains. We then calculate the sum of each side between 0 and 200 Hz, -2 and 0 m^{-1} for the negative wavenumber part, and 0 and 2 m^{-1} for the positive wavenumber part. A negative part (K^-) greater than the positive part (K^+) indicates a preferential leftward propagation, suggesting that the source is situated at the right of the ULA (case R). Conversely, a positive part greater than the negative part indicates a preferential rightward propagation, suggesting that the source is situated at the left of the ULA (case L). However, when the source is placed along the ULA or no source is nearby (cases P and F), the negative and positive parts can be similar, showing no specific direction of propagation. In accordance with (Cheng et al., 2018), we define a threshold Λ for the ratio of the negative and positive parts as follows:

- If $K^- > K^+$ and $-\frac{K^-}{K^+} + 1 < -\Lambda$ then, the source is estimated to be at the right of the ULA (case R).
- If $K^+ > K^-$ and $\frac{K^+}{K^-} - 1 > \Lambda$ then, the source is estimated to be at the left of the ULA (case L).
- else, the segment is eliminated (cases P and F).

The threshold value Λ is however site-specific. After several trials, we determined that a value of 0.3 for site A and 0.6 for sites B and C yields the optimal results. Site A appears to exhibit lower attenuation for noise generated on the left side of the ULA compared to sites B and C. This could be explained by a different nature of materials despite a similar HSL structure enabling the recording of the noise even when the HST is further. On aver-

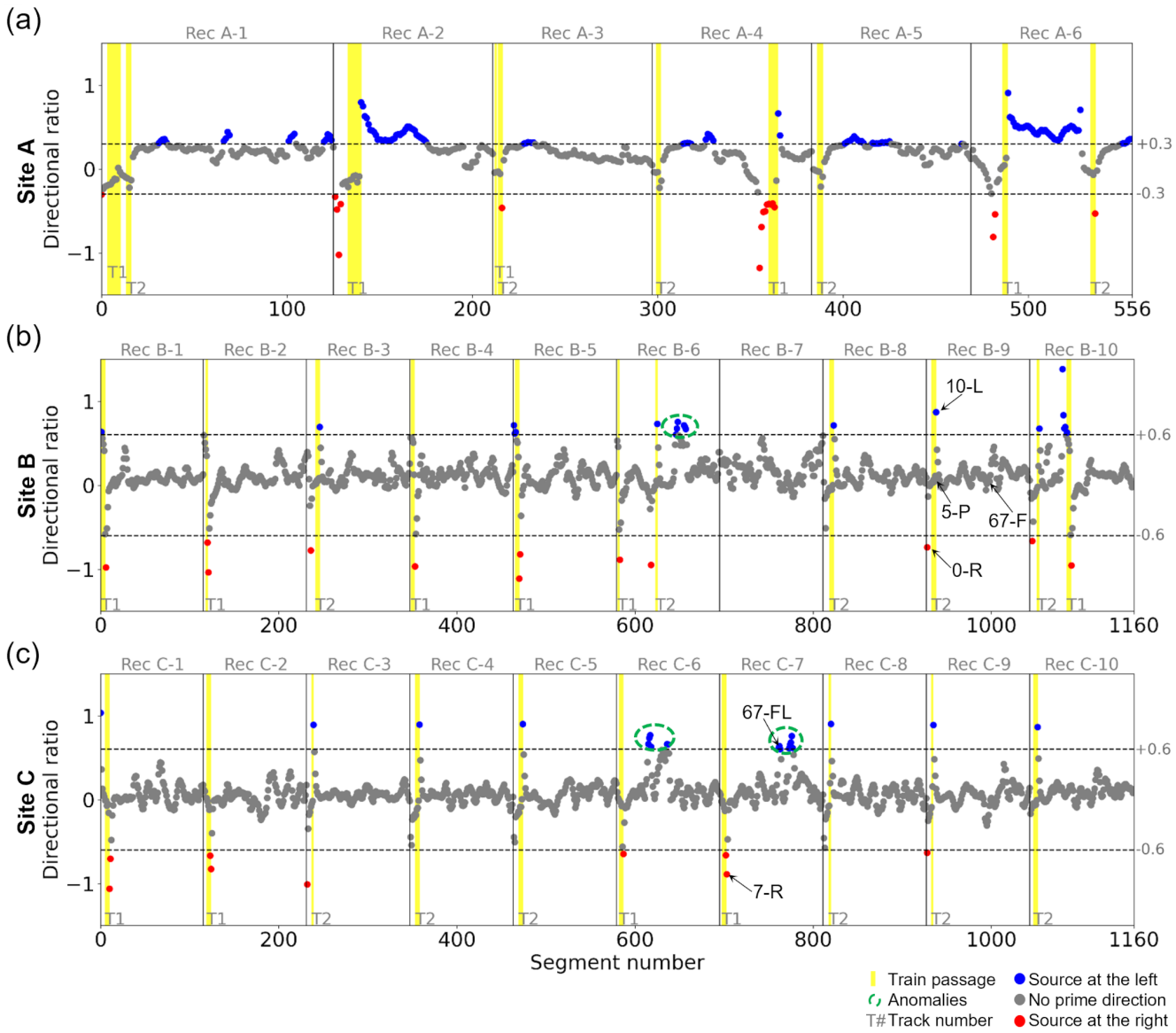

Figure 4 (a), (b), and (c) FK segment selection for every record of sites A, B, and C, respectively (see Fig. 1 for sites location). Red, blue, and grey dots indicate segments with detected seismic waves coming from the right, left and not showing distinct direction, respectively. Yellow vertical lines represent time segments of HST passages on T1 and T2 (see Table 1). Highlighted segments in records B-9 and C-7 are the same shown in Figure 3. Green circles denote anomalies where waves were detected coming from the left but not corresponding to any HST passages.

age, a Λ value around 0.4 proves to be effective for the HST records.

Figure 4 illustrates the directional ratio estimation for each segment. Notably, only segments immediately before or after HST passages are retained (in blue for HSTs at the left and red for HSTs at the right of the ULA). Segments occurring during train passages or when the HST is at a considerable distance from the ULA, exhibiting directional ratios between -Λ and +Λ, are excluded (in grey). Furthermore, the estimated source directions align with the HST directions of movement (see Fig. 1 and Table 1). For site A, HSTs travel from right to left on T1 and from left to right on T2. For sites B and C, HSTs travel from left to right on T1 and from right to left on T2. Source direction estimations from the same six segments in records B-9 and C-7, as shown in Figure 3, are correctly identified. However, segment 67-FL of record C-7, which is far from the HST passage and is expected to contain minimal seismic signal, appears to exhibit a significant unknown source strongly located on the left side of the ULA. This anomaly is observed in records B-6, C-6, and C-7 at sites B and C, which are geographically close to each other.

4 Interferometry with high-speed trains

4.1 Importance of signal direction

A seismic signal, originating from any source and recorded by a distant receiver, carries both the information related to the source (energy, duration, directionality, etc.) and the information about the path traveled by the waves between the source and receiver. The elas-

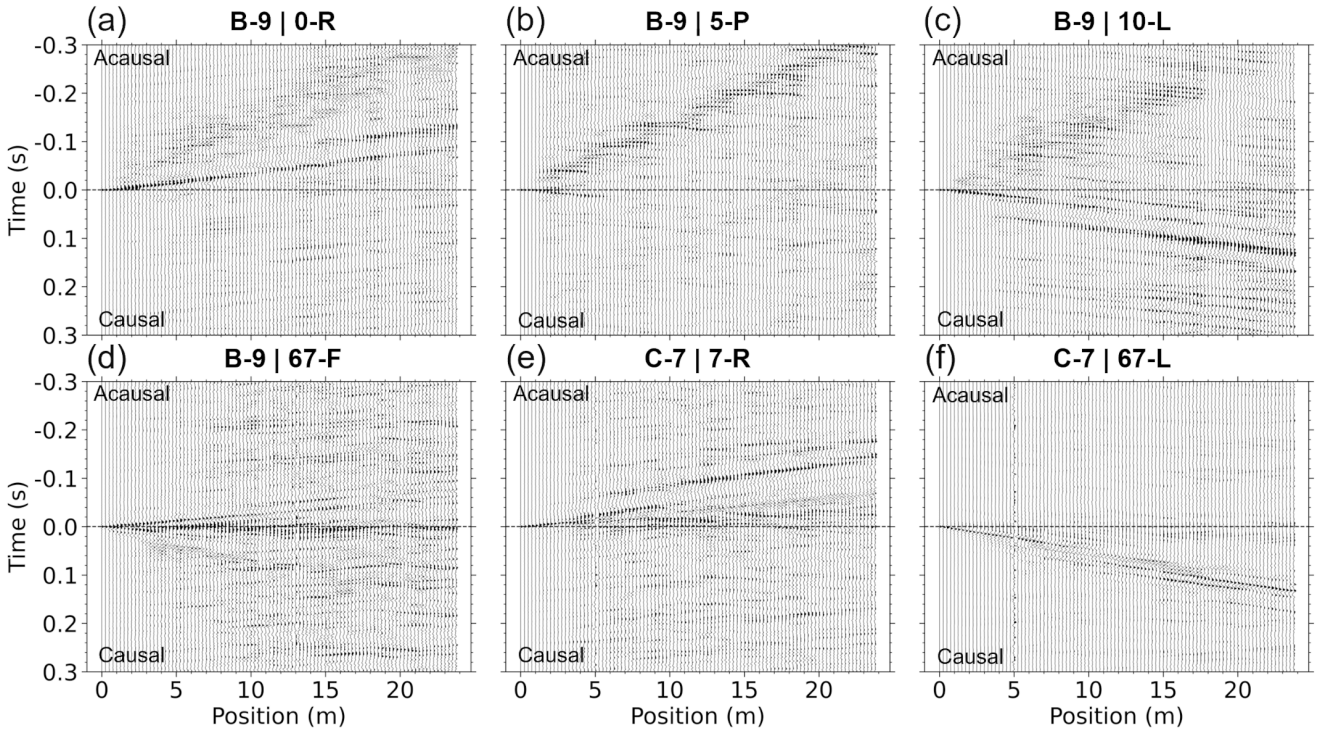

Figure 5 Direct VSGs (i.e. source at geophone n°1) with causal and acausal parts, obtained from segments highlighted in Figures 3 and 4. The propagation of SW is retrieved in the acausal part for (a) and (e) when the source is at the right, in the causal part for (c) and (f) when the source is at the left of the ULA, and is not conclusive for (b) and (d) when the HST is passing or too far.

todynamic transfer function of the medium is referred as the Green's function. It represents the medium response to a punctual excitation (i.e. a Dirac distribution) in time and space, which fully describes the propagation of elastic waves between the source and receiver. If a source function is defined by some wavelet $s(x_S,t)$ generated at the position x_S, then the wavefields recorded by two different geophones at x_A and x_B can be defined as:

$$\begin{cases} u\left(x_A,s,t\right) = s\left(x_S,t\right) * G\left(x_A,x_S,t\right) \\ u\left(x_B,s,t\right) = s\left(x_S,t\right) * G\left(x_B,x_S,t\right) \end{cases} , \tag{1}$$

where the asterisk denotes the temporal convolution and, $G(x_A,x_S,t)$ and $G(x_B,x_S,t)$ are the Green's functions observed at x_A and x_B, respectively. Wapenaar et al. (2010a,b) and Boué and Paul (2020) have shown that the Green's function between x_A and x_B can be estimated by crosscorrelation of the wavefields recorded at both points. The crosscorelation of the wavefields recorded at x_A and x_B is defined as

$$C_{AB} = u\left(x_A,s,t\right) * u\left(x_B,s,-t\right). \tag{2}$$

Let $S(x_S,t) = s\left(x_S,t\right) * s\left(x_S,-t\right)$ is the autocorrelation of the source wavelet, therefore,

$$C_{AB} = S\left(x_S,t\right) * G\left(x_A,x_S,t\right) * G\left(x_B,x_S,-t\right). \tag{3}$$

In a 1D case, when x_S, x_A and x_B are aligned along a x-axis, if the wave is traveling rightward (source at the left) along x_S, x_A and x_B, then the cross correlation is

$$C_{AB} = S\left(x_S,t\right) * G\left(x_B,x_A,t\right) , \tag{4}$$

and, if the wave is traveling leftward (source at the right) along x_A, x_B and x_S, then the cross correlation is

$$C_{AB} = S\left(x_S,t\right) * G\left(x_B,x_A,-t\right) . \tag{5}$$

Because $S(x_S,t)$ approximates a delta function, we can write equations 4 and 5 as the an approximation of the Green's function:

$$C_{AB} \approx G\left(x_B,x_A,t\right) , \tag{6}$$

and

$$C_{AB} \approx G\left(x_B,x_A,-t\right) . \tag{7}$$

$G\left(x_B,x_A,t\right)$ and $G\left(x_B,x_A,-t\right)$ represent the causal and acausal responses, respectively, of an impulse generated at $t = 0$ s, providing the response at x_B if the source was located at x_A. This formulates that the cross-correlation of wavefields recorded at two different geophones gives the theoretical wavefield at one geophone if a source, close to an impulse, was at the other geophone.

In a one-dimensional, non-diffuse wavefield scenario, this approximation relies on the precise alignment of the geophones and the source, as well as the sequential propagation of waves from the source through each geophone. It is also essential to note that when correlated sources are simultaneously generated from both sides, the cross-correlation can contain cross-terms that lack physical meaning (Wapenaar et al., 2010a). Conversely, if both sources are uncorrelated, the causal and acausal components tend to be symmetric. Thus, using HSTs as seismic sources in a 1D setup requires information about the train's direction and relative position. The segment selection technique outlined in Section 3 enables to overcome those requirements. One

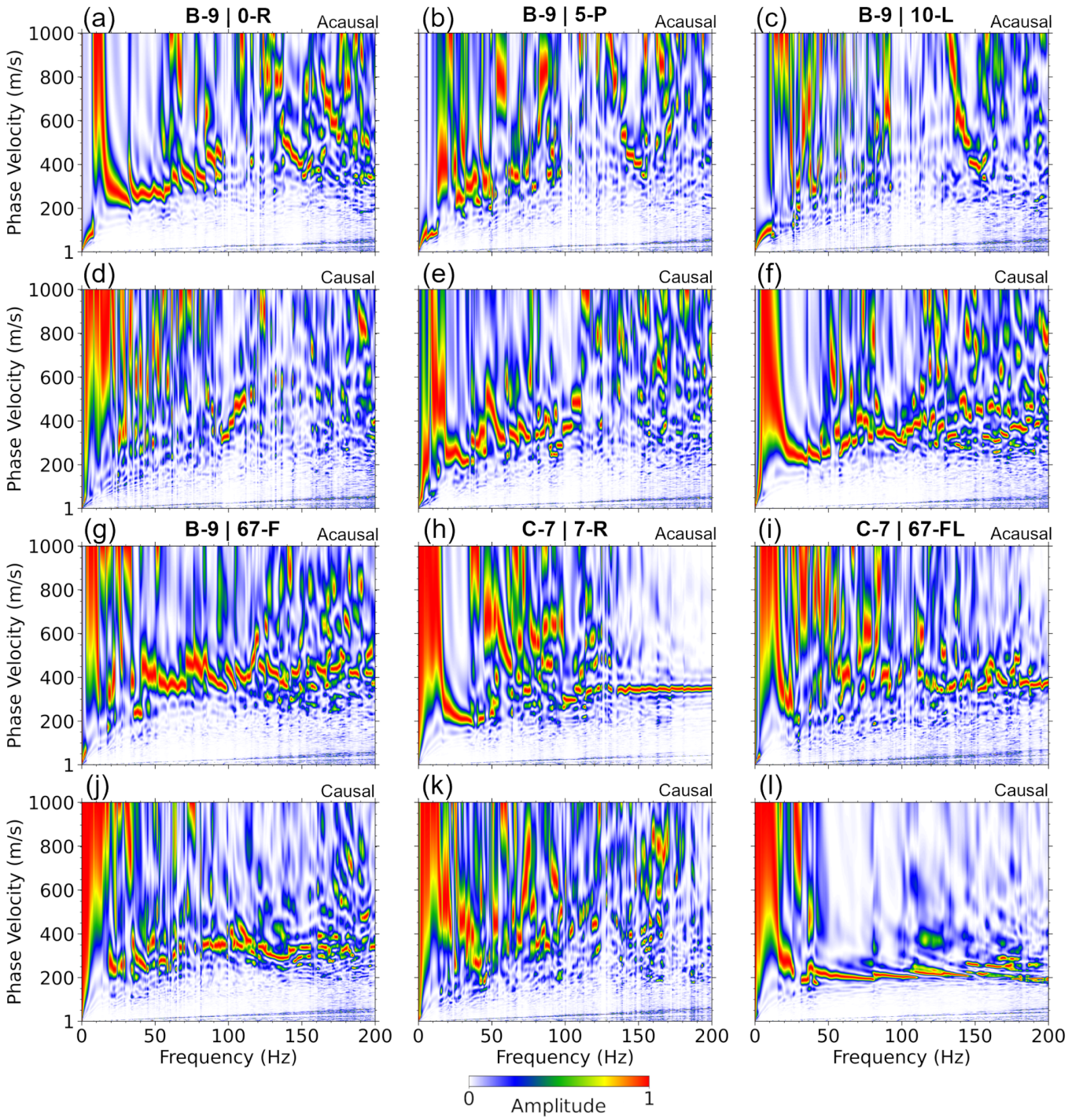

Figure 6 DIs normalized by frequency, obtained from the direct VSGs (i.e. source at geophone n°1) from Figure 5. (a), (c), (e), (g), (i) and (k) were obtained with the acausal part of the cross-correlations and (b), (d), (f), (h), (j), and (l) with the causal part.

must also note that, although the signal induced by each HST is unique, varying with factors such as speed, weight, and number of cars, it does not fundamentally alter the outcomes of the cross-correlation, since the Green's function only represents the medium response to a punctual excitation. However because $S(x_S, t)$ only approximates a delta function, in practice the frequency range of the source, and consequently of the recorded signal, is limited by the type and speed of the train, rather than being infinite.

4.2 Preprocessing

Before applying interferometry, each segment undergoes a sequence of preprocessing steps. We follow the approach outlined by Bensen et al. (2007), which involves processing each passive seismic trace individually. This processing includes demeaning, detrending, temporal normalization, and spectral whitening. The initial three steps—demeaning, detrending, and temporal normalization—are utilized to mitigate the effects of instrumental irregularities. Bensen et al. (2007) consider five different temporal normalization methods

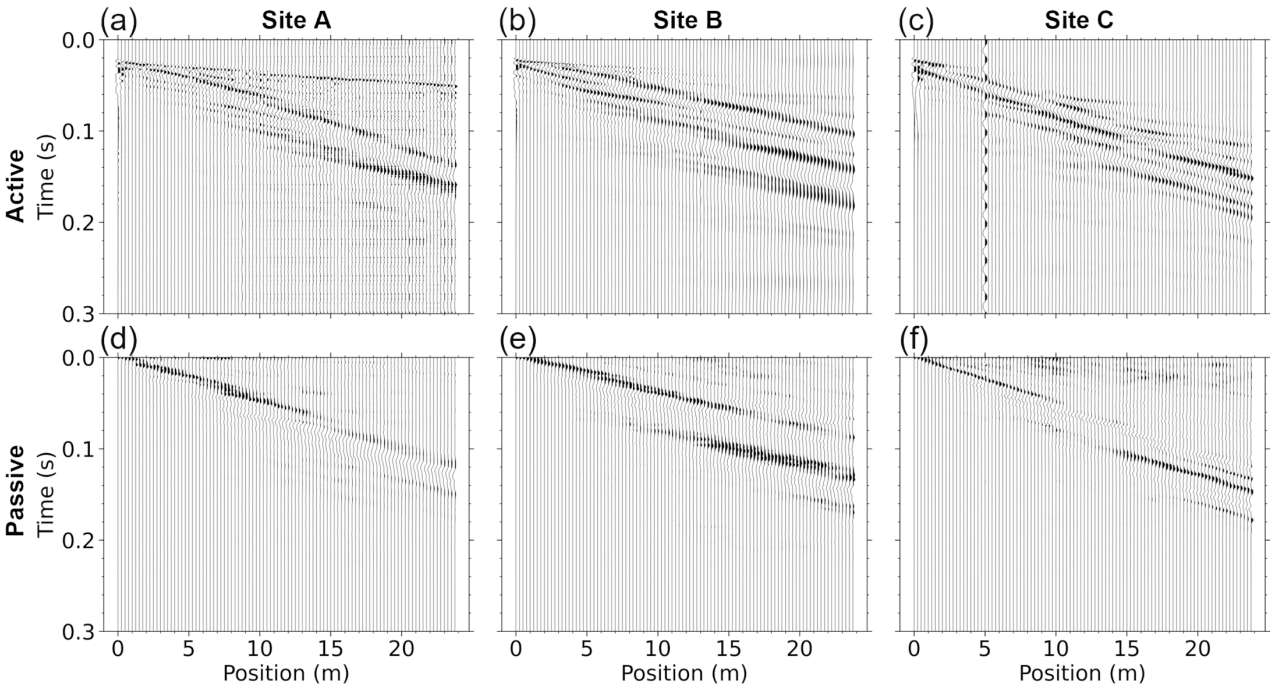

Figure 7 (a), (b), and (c) Final normalized direct SGs acquired through classical active seismic for sites A, B, and C. (d), (e), and (f) Normalized VSGs obtained by passive seismic for sites A, B, and C after FK selection, preprocessing and squared PWS ($\nu = 2$).

but recommend the running-absolute-mean technique, which entails a sliding averaging of the absolute value of the waveform within a temporal window with a fixed width equal to half the maximum period of the signal. However, through several trials, we found that temporal normalization did not yield notable quality improvements in the data. This is expected, given that we use a strong source in close proximity to our geophones. As a result, this step is considered optional. Spectral whitening, on the other hand, is applied to the entire spectrum of each trace to correct for high-frequency attenuation and make the spectrum more uniform.

4.3 Virtual shot-gathers

We apply cross-correlations between the first geophone (i.e. n°1) and all geophones, as well as between the last geophone (i.e. n°96) and all geophones, resulting in a direct and inverse virtual shot-gather (VSG). These VSGs illustrate the propagation along the ULA of an impulse that would be generated at the first and last geophones. However, it's crucial to ensure that the propagation of train-induced waves is properly aligned with all the geophones. Additionally, we need to distinguish whether the train is located to the left or the right of the ULA. This distinction is crucial to select between the causal and acausal parts of the cross-correlation:

- For the direct VSG, if the train is at the left the causal part is conserved and if the train is at the right the acausal part is conserved.
- For the inverse VSG, conversely, if the train is at the left the acausal part is conserved and if the train is at the right the causal part is conserved.

Additionally, given that sources generated by trains closely resemble correlated impulses produced each time a wheel passes over a sleeper (Lavoué et al., 2020; Rebert et al., 2023; Rebert et al., 2024a), wavefields recorded when trains are simultaneously traveling in both tracks must be avoided in order to prevent cross-terms. The segment selection method described in Section 3 allows us to verify all these requirements.

Figures 5 and 6 show the direct VSG and the DIs obtained from the same segments highlighted in Figures 3 and 4 for both the causal and acausal parts. The DIs were generated using a phase-shift transform described in Appendix B. We verify the assertion that when the train is at the left, the causal part contains the medium's response (see Figs. 5c,f and 6f,l) and conversely, when the HST is at the right, the acausal part contains the medium's response (see Figs. 5a,e and 6a,i). However, when the HST is far from the ULA or passes by it, both causal and acausal parts are hardly exploitable (see Figs. 5b,d and 6c,d,g,h). Moreover, Figure 6l provides more insights into the unknown source affecting records B-6, C-6, and C-7. It appears that the source has a high-frequency spectrum and reveals higher modes in the DIs.

5 Time-domain stacking

A stacking operation is applied to the entire set of VSGs, without differentiating direct and inverse VSGs, with the primary aim of enhancing the signal-to-noise ratio. Several stacking techniques are commonly used in seismology, Yang et al. (2023) provided a comprehensive list and evaluation. The most well-known methods are the standard linear stack, the nth-root stack (Kanasewich et al.,

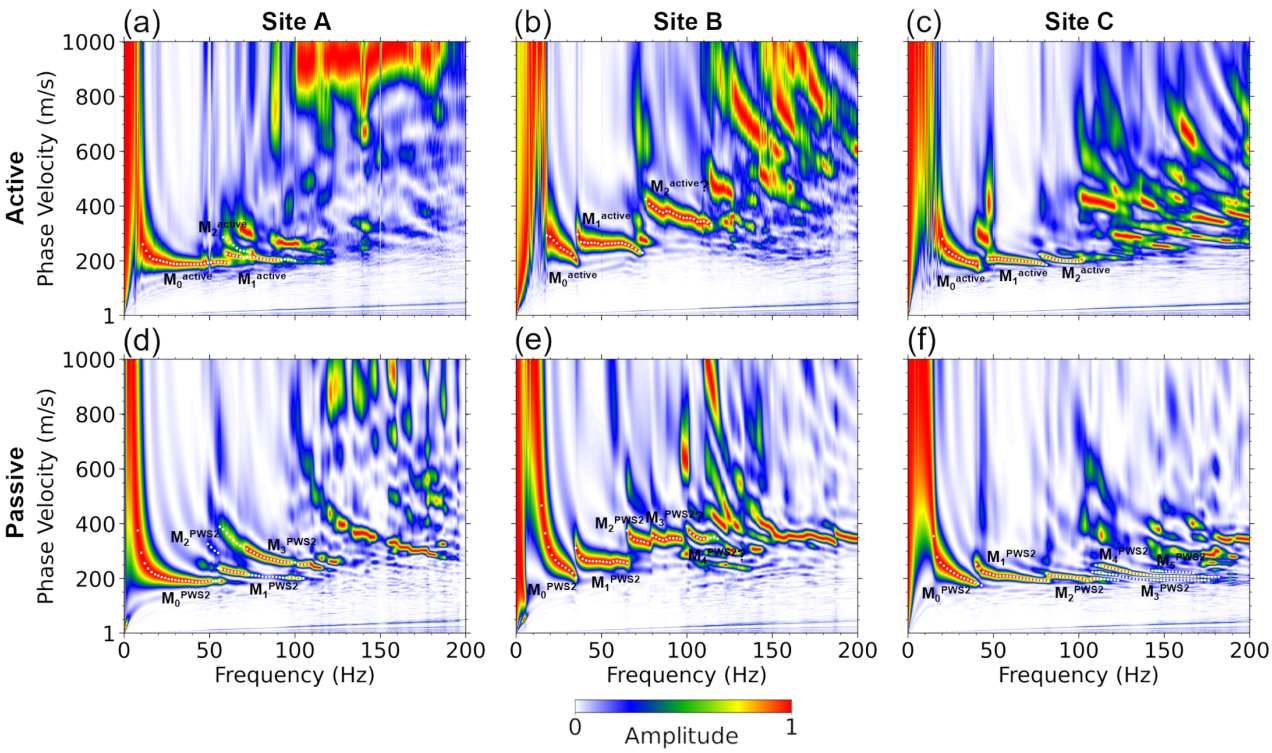

Figure 8 (a), (b), and (c) DIs normalized by frequency, obtained by classical active seismic for sites A, B, and C. (d), (e), and (f) DIs normalized by frequency, obtained by passive for seismic after FK selection, preprocessing and squared PWS ($\nu = 2$) for sites A, B, and C. Picked DCs (white dots) reveal several modes of propagation in the 0-100 Hz and 0-400 m/s ranges that were previously observed in other railway contexts.

1973), and the phase-weighted stack (PWS) (Schimmel and Paulssen, 1997). In our study, we utilize the PWS method due to its demonstrated efficacy in passive seismic applications (Thurber et al., 2014; Ning et al., 2022; You et al., 2023).

In accordance with Schimmel and Paulssen (1997), each trace $s_j(t)$ (real part) can be combined with its Hilbert transform $\mathrm{H}[s_j(t)]$ to form the complex trace (analytic signal) $S_j(t)$,

$$S_j(t) = s_j(t) + i\mathrm{H}[s_j(t)] \ , \tag{8}$$

which can be rewritten in terms of amplitude $A_j(t)$ and instantaneous phase $\Phi_j(t)$:

$$S_j(t) = A_j(t)\exp[i\Phi_j(t)] \ . \tag{9}$$

The phase stack $c(t)$ is defined as:

$$c(t) = \frac{1}{N}\left|\sum_{j=1}^{N}\exp[i\Phi_j(t)]\right| \ , \tag{10}$$

where N represents the number of traces being stacked. The phase stack is a time-dependent function ranging between 0 and 1, serving as a direct measure of coherency. If the instantaneous phases of the signals at a particular time are coherent, this function equals one, and zero if the summed signals are completely destructive. Notably, it's worth mentioning that the amplitudes of the stacked signals are not utilized and have no impact on the phase stack. The PWS, denoted as $g(t)$, is then obtained by using the phase stack as a time-varying weight function for the linear stack:

$$g(t) = \frac{1}{N}\sum_{j=1}^{N} s_j(t)\left|\frac{1}{N}\sum_{k=1}^{N}\exp[i\Phi_k(t)]\right|^{\nu} \ , \tag{11}$$

with ν the power of the stack. The PWS method enhances the standard linear stack by assigning lower weights to components of the traces that don't share the same instantaneous phase. It acts as a filter that amplifies weak coherent signals while reducing incoherent noise. Its sharpness is controlled by the power parameter ν and the linear stack can be retrieved with $\nu = 0$. Notice that the PWS is a non-linear stack, which can lead to waveform distortion, particularly for incoherent signals. This distortion is similar to the one generated by the nth-root stack, but to a much lesser extent, and appears to narrow wavelet widths, especially when high stacking powers are applied (McFadden et al., 1986; Schimmel and Paulssen, 1997).

Figures 7d-f display the final VSGs obtained after applying the PWS with $\nu = 2$, for sites A, B, and C. The comparison with the SGs recorded with the active seismic method (see Figs. 7a-c) reveals that the same SW signatures are successfully retrieved. However, P-waves, which are clearly visible in SGs from active seismic, are more challenging to recover using passive VSGs but can still be discerned in Figure 7d.

6 Results

6.1 Active versus Passive

Figure 8 displays the DIs obtained by phase-shift transform (see Appendix B), with both the traditional ac-

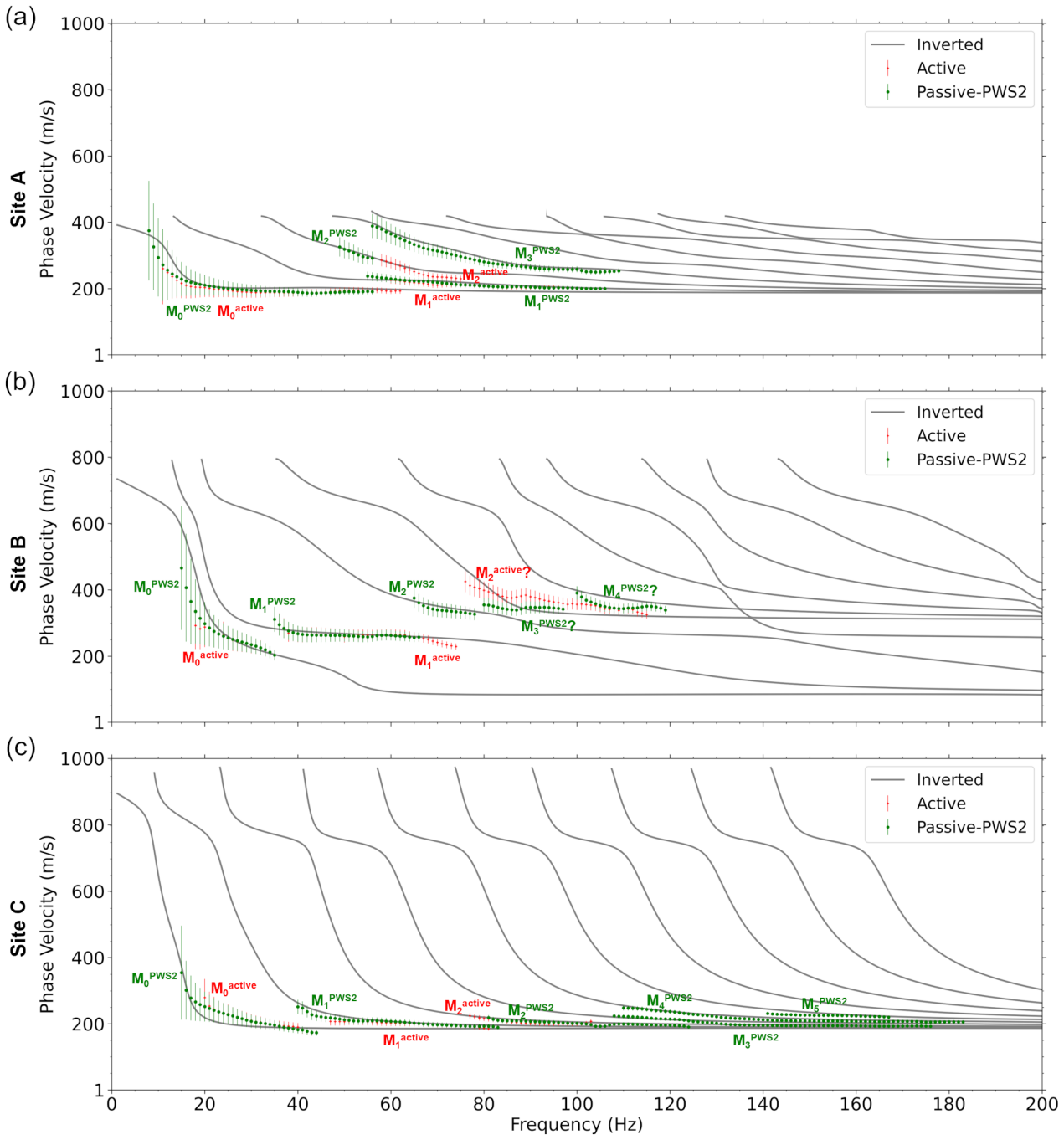

Figure 9 (a), (b), and (c) Comparison of picked DCs, with error-bars, from active and passive ($\nu = 2$) seismic methods along with the DCs obtained by forward modeling from the average velocity models inverted with M_0^{PWS2}, M_1^{PWS2}, and M_2^{PWS2}, for sites A, B, and C.

tive MASW method, and the passive-MASW after FK segment selection, preprocessing, interferometry, and squared PWS ($\nu = 2$), for sites A, B and C. See Appendix A for more details about the active MASW processing workflow. Both techniques result in highly similar images but with particularities to each method. On one hand, passive DIs exhibit overall great richness of information and feature coherent content at higher frequencies than active results.

DCs for several modes were semi-automatically picked on the DIs. For a given mode i, picked active DCs are referred to as M_i^{active} and picked passive DCs, obtained using a PWS power $\nu = 2$, are referred to as M_i^{PWS2}. Both techniques allow the recognition of, at least, the fundamental mode (M_0), as well as the first (M_1) and second (M_2) higher modes. However, for site A, a third passive higher mode (M_3^{PWS2}) is retrieved, but is difficult to discern on the active DI. For site B, the second active higher mode (M_2^{active}) seems to be a superposition of several modes that are more disentangled in passive results (M_2^{PWS2}, M_3^{PWS2}, and M_4^{PWS2}). Lastly, for site C, a glimpse of three additional modes is retrieved in the passive results in comparison with the active seismic data.

6.2 Multi-modal dispersion curves

Figure 9 overlays the experimental active (M_i^{active} in red) and passive (M_i^{PWS2} in green) DCs, with their theoretical uncertainty $\delta_c(f)$ represented by error-bars and calculated in accordance with O'Neill (2003):

$$\delta_c(f) = 10^{-a}\left|\frac{1}{\frac{1}{V_R(f)} - \frac{1}{2fN_x\Delta_x}} - \frac{1}{\frac{1}{V_R(f)} + \frac{1}{2fN_x\Delta_x}}\right|, \tag{12}$$

with a being the logarithmic reduction factor, usually 0.5 (O'Neill, 2003), N_x the number of geophones used to compute the DI, and Δ_x the space interval between geophones. Note that the uncertainty is higher at low velocities but that the global uncertainty across all frequencies can be reduced by increasing the length of the geophone array $N_x\Delta_x$. This reflects the typical compromise that has to be done between investigation depth (usually captured by lower frequencies), resolution, and the 1D approximation required for the inversion of dispersion data. On the one hand, lengthening the array naturally improves the resolution of the DI. Since uncertainties are dependent on resolution, this helps reduce potential errors at low frequencies and thus enhances the interpretation of deeper layers. On the other hand, extending the array too much may introduce significant lateral variations, invalidating the 1D assumption. This type of error has been highlighted by O'Neill (2003) and others in the field of near-surface geophysics. Furthermore, large uncertainties at low frequencies can help minimize the influence of near-offsets effects (O'Neill, 2003; Zywicki and Rix, 2005; Bodet et al., 2009).

The DCs obtained from both methods exhibit very good coherence, with most curves falling within the uncertainty range. Passive seismic analysis appears to offer more detailed information regarding the fundamental mode at lower frequencies. However, it is crucial to approach the results with caution, as the uncertainty is higher at these low frequencies. In the case of site A, M_2^{PWS2} seems to be a continuation of M_2^{active} at lower frequencies. For site B, we more precisely observe that M_2^{active} appears to be a composite of different modes on the DIs. In fact, the passive seismic method seems to be able to separate it into three distinct modes (M_2^{PWS2}, M_3^{PWS2}, and M_4^{PWS2}).

For each site, the Rayleigh-wave DCs corresponding to M_0^{PWS2}, M_1^{PWS2}, and M_2^{PWS2}, were inverted to obtain a 1D V_S versus depth ground model (more details about the inversion process are described in Appendix C). Up to ten DC modes were then recomputed from the resulting average model by forward modeling using a Thomson-Haskell propagation matrix, and plotted on the background of Figure 9 (in grey). Not surprisingly, we observe that the DCs obtained after inversion align closely with experimental M_0^{PWS2}, M_1^{PWS2}, and M_2^{PWS2}, which were used for the inversion. Nonetheless, it is also interesting to observe that, even higher modes, such as M_3^{PWS2} for site A, and M_3^{PWS2}, M_4^{PWS2}, and M_5^{PWS2} for site C, almost perfectly match. However, for site B, even if M_3^{PWS2} and M_4^{PWS2} still coherent with the modeled DCs, we were not able to obtain a satisfying inverted model with realistic velocities for the second layer (see Table 2 and Appendix C for detailed inversion results), which corresponds to the capping layer (see Fig. 1d). This indicates that neither the active nor the passive approach using a squared PWS provided sufficient resolution to accurately characterize this critical thin layer.

7 Discussion

7.1 Enhancing multi-modal characterization of extracted dispersion

Figures 10a-c present the VSGs obtained using a PWS power $\nu = 10$. As expected, the wavelet widths are narrowed, and only coherent strong signals are retained. Again, DCs for several modes were semi-automatically picked on the new corresponding DIs (see Figs. 10d-f). For a given mode i, picked passive DCs are referred to as M_i^{PWS10}. For site A, we retrieve the same modes (M_0^{PWS10} to M_3^{PWS10}) with improved continuity, and four additional modes are now visible. For site C, we recover the same five modes, but with increased clarity. Surprisingly, for site B, the new DI reveals a total of eleven modes that can be picked. This adjustment in the PWS power leads to notable changes in the visibility and differentiation of the modes in the DIs.

Table 2 Final average velocity ground models obtained after inversion of M_0^{PWS2}, M_1^{PWS2}, and M_2^{PWS2}, for sites A, B, and C.

Site	Layer	Thickness	V_P	V_S	Density ρ
	(#)	(m)	(m/s)	(m/s)	(kg/m^3)
A	1	0.4	1375	512	2204
	2	0.2	1092	369	2234
	3	5.0	340	186	2297
	½-space	∞	1232	419	2229
B	1	0.5	812	99	2304
	2	0.1	810	34	2265
	3	6.3	1044	309	2260
	½-space	∞	1620	795	2243
C	1	0.1	1743	677	2316
	2	0.8	1033	181	2050
	3	4.9	1539	187	2415
	½-space	∞	1835	973	2217

For each site, the new passive M_0^{PWS10}, M_1^{PWS10}, and M_2^{PWS10}, were as well inverted. For sites A and C, the resulting velocity models are very similar to the previously obtained in Section 6.2. However, for site B, the velocity model shows now more realistic velocities, allowing for better differentiation of the capping layer (second layer), which now exhibits the expected thickness of 60 cm (see Table 3 and Appendix C for more detailed inversion results).

Again, up to ten DC modes are recomputed from the resulting average models by forward modeling in order to compare them with the experimental DCs. In Figure 11, we overlay the experimental active DCs (M_i^{active} in red), the experimental passive DCs obtained using

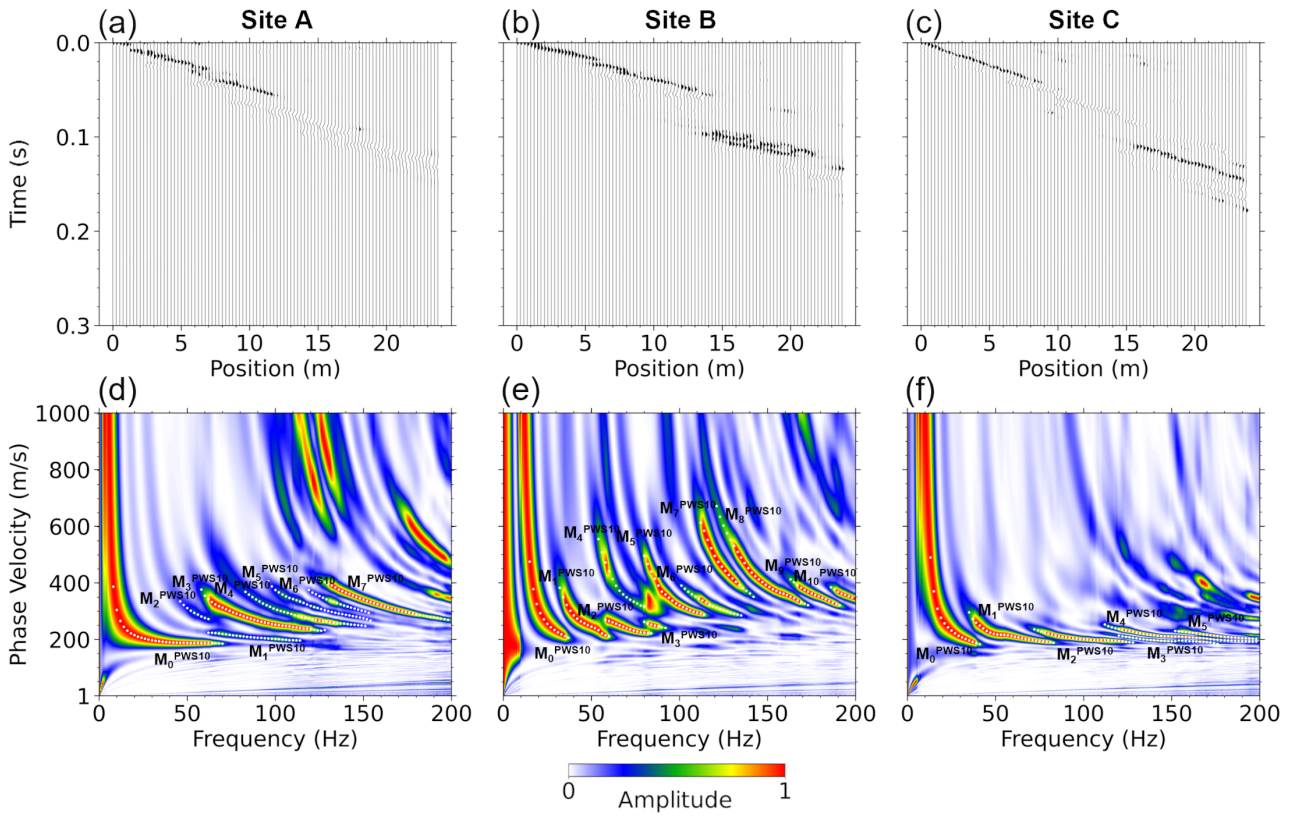

Figure 10 (a), (b), and (c) Normalized VSGs obtained through passive seismic after the FK selection, preprocessing, and PWS ($\nu = 10$) for sites A, B, and C. (d), (e), and (f) DIs normalized by frequency, obtained through passive seismic after FK selection, preprocessing, and PWS ($\nu = 10$) for sites A, B, and C.

$\nu = 2$ (M_i^{PWS2} in green), the experimental passive DCs obtained using $\nu = 10$ (M_i^{PWS10} in blue), and the DCs obtained by forward modeling (in grey). For sites A and C, the new DCs obtained after inversion exhibit minimal change in comparison with Figure 9, and every single mode aligns well with the inverted model. However, the results for site B are markedly different. It appears that the previous modes have been split. Specifically, M_1^{active} and M_1^{PWS2} seem to have been separated into two or three distinct modes (M_1^{PWS10}, M_2^{PWS10}, and M_3^{PWS10}). We note that M_3^{PWS10} seems actually be a part of M_2^{PWS10}. The same effect can be seen for M_2^{active}, which appears to represent an effective dispersion of three different modes (M_4^{PWS10}, M_5^{PWS10}, and M_6^{PWS10}), as assumed in Section 6. Additionally, all new higher modes also fit well with the inverted model.

This comparison further highlights the impact of the chosen PWS power ν on the DCs and subsequent interpretation. Increasing the PWS power ν appears to enhance the retrieval of more higher modes for sites A and B. However, this increase in ν also leads to DIs that are less identical to those obtained through active seismic methods, and showing smoother modes of propagation. Essentially, if the objective is to achieve results similar to those obtained with active seismic, then using $\nu = 2$ is more suitable. However, when the goal is to obtain an inversion that is well-constrained at high frequencies or to monitor V_R for several modes, a higher ν can provide additional information. This research highlights the importance of considering the trade-off between the precision of low-frequency modes and the inclusion of higher modes when selecting the appropriate PWS power. The effects of using a high ν on DIs have been relatively underexplored in the literature. We show that almost all higher modes align with the inversion results, which were computed using only the first three modes. However, further comprehensive studies are necessary to thoroughly validate these higher modes as actual Rayleigh-wave modes.

Table 3 Final average velocity ground models obtained after inversion of M_0^{PWS10}, M_1^{PWS10}, and M_2^{PWS10}, for sites A, B, and C.

Site	Layer (#)	Thickness (m)	V_P (m/s)	V_S (m/s)	Density ρ (kg/m^3)
A	1	0.3	1108	344	2255
	2	0.2	996	246	2240
	3	5.2	713	193	2326
	½-space	∞	1062	387	2237
B	1	0.2	955	207	2214
	2	0.6	747	118	2156
	3	5.1	1296	215	2344
	½-space	∞	1669	856	2258
C	1	0.1	1458	574	2167
	2	0.7	1128	149	2062
	3	4.3	743	183	2447
	½-space	∞	1613	745	2298

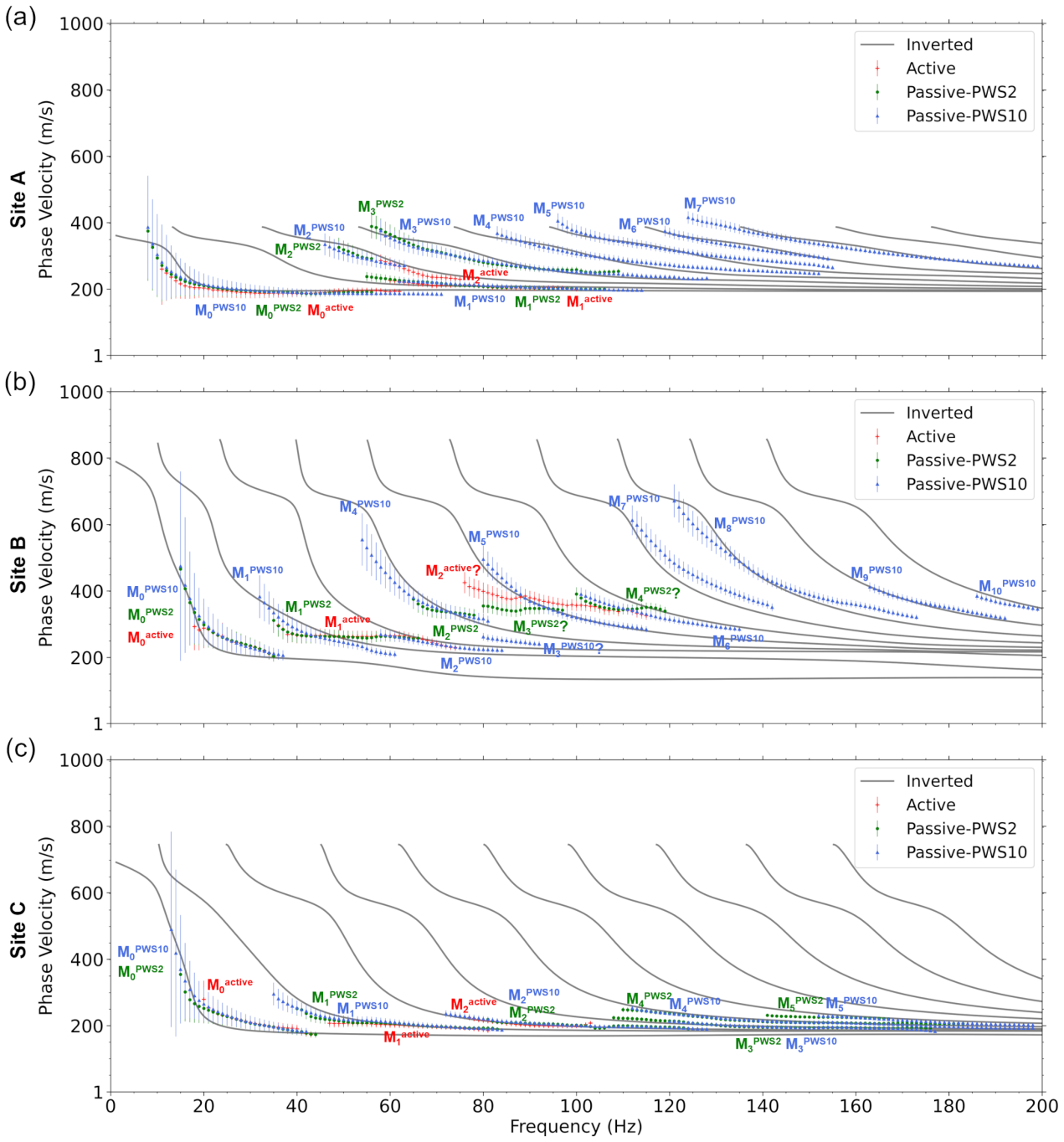

Figure 11 (a), (b), and (c) Comparison of picked DCs, with error-bars, from active and passive ($\nu = 10$) seismic methods along with the DCs obtained by forward modeling from the average velocity models inverted with M_0^{PWS10}, M_1^{PWS10}, and M_2^{PWS10}, for sites A, B, and C.

7.2 Most modest setup for satisfying results

Figures 12 and 13 present the DIs for sites A, B, and C, obtained using less dense ULAs with 48, 24, and 12 geophones over the same initial distance of 23.75 m, and considering two PWS power values, $\nu = 2$ and $\nu = 10$. We observe that even with an array density reduced by a factor of four (24 geophones), the distinction between the various modes of propagation is still feasible (see Figs. 12d-f and 13d-f). However, when using only 12 geophones over the same distance, multiple modes tend logically to merge together, leading to the observation of an effective dispersion rather than individual modes. It is also important to note that as the array density decreases, the influence of the PWS power ν on the resulting DIs becomes less significant (see Figs. 12g-i and 13g-i). This observation underscores the critical role of array density in determining the resolution and interpretability of the DIs in conjunction with the chosen PWS power.

Using a ULA with 24 geophones is a economical option for RE diagnostic, providing results hat are comparable to those obtained with a more dense array of 96 geophones. However, if the primary goal is to globally estimate or monitor V_R variations, a 12-geophone array could also suffice, with a focus on extracting velocity ranges instead of modal responses. It is impor-

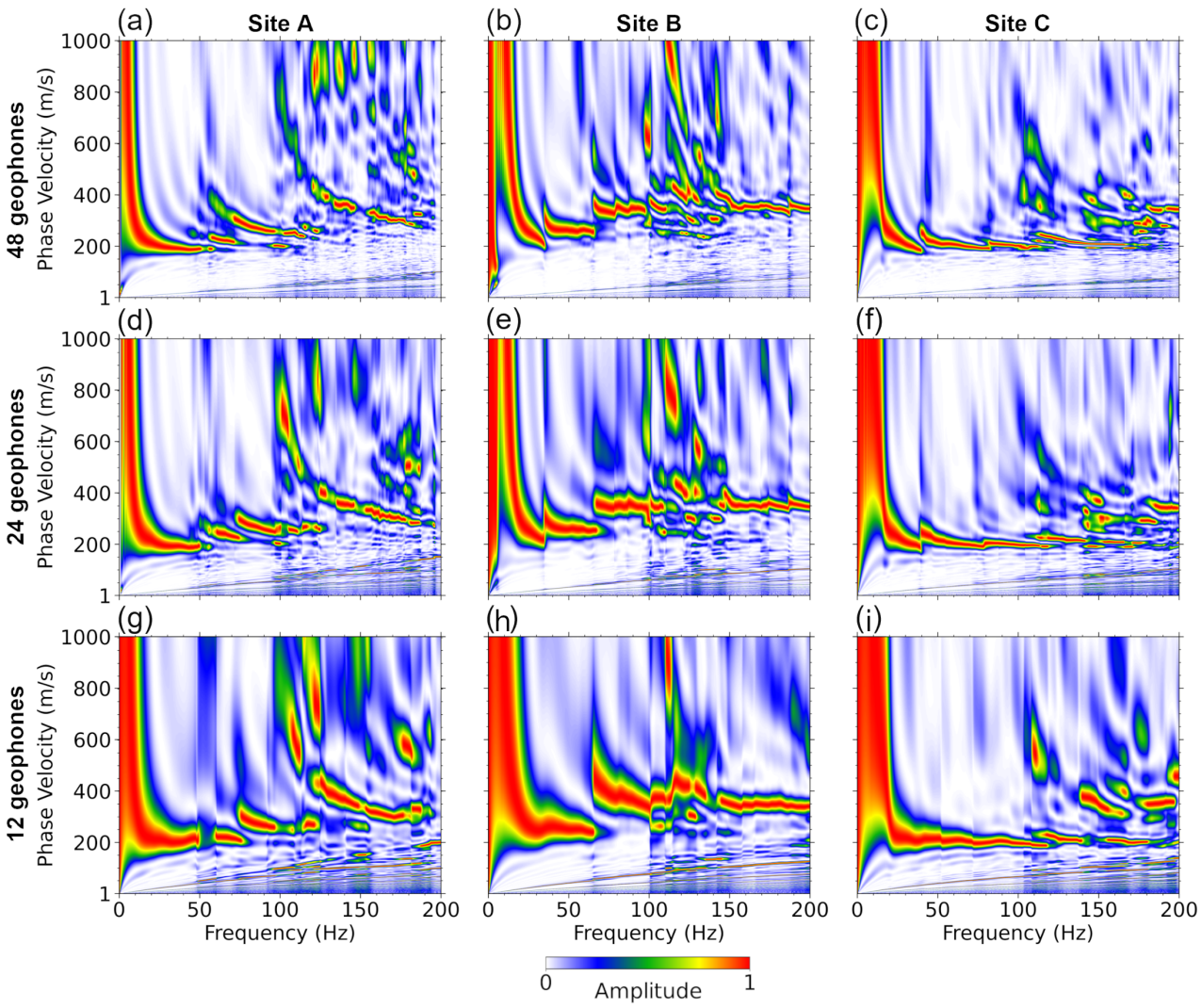

Figure 12 DIs normalized by frequency, obtained by passive seismic after FK selection, preprocessing and squared PWS ($\nu = 2$) for sites A, B, and C using a ULA of (a, b, and c) 48 geophones spaced by 50 cm, (d, e, and f) 24 geophones spaced by 75 cm, and (g, h, and i) 12 geophones spaced by 1 m.

tant to consider the maximum wavenumber that can be identified, which depends on the receiver spacing dx. Energy associated with wavenumbers above π/dx get aliased, leading to the loss of information from short wavelengths (i.e. high frequencies) that typically offer insights into the uppermost layers. However, limitations on the highest detectable frequencies are often primarily related to attenuation and site response, and receiver spacing is critical only at very low-velocity sites (Socco and Strobbia, 2004). Therefore, selecting the appropriate receiver spacing is crucial in optimizing the trade-off between economical considerations and the desired resolution of the SW velocity estimation. A baseline investigation should be performed before any permanent survey, to determine the best configuration.

It is also worth noting that if a geophone is not functioning correctly, the impact on overall data quality is minimal. This is because the phase-shift transform method used to generate DIs leverages the redundancy of data collected along the geophone line. This robustness is one of the key strengths of the MASW method, ensuring reliable data processing even when minor issues occur during field deployment.

8 Conclusions

The comparison between DIs generated by passive-MASW, utilizing HST induced waves, and DIs generated by classical MASW, using a hammer as an active source, reveals a significant coherence between the two methods. The superposition of the DCs from both methods is nearly perfect, but it appears possible to identify more modes of propagation and to achieve better mode continuity using passive-MASW.

On two out of the three test sites, increasing the power of the PWS appears to facilitate the identification of additional modes of propagation. Strong *a priori* information about the investigated site made it possible to perform forward computation in order to verify the relevance of these additional modes. We demonstrated that being able to correctly discern higher modes with the passive method, enabled a better characterization the capping layer with realistic thicknesses and mechanical properties, making it possible to effectively monitor this layer over time.

We also showed that reducing the array density by a factor of 4 (array of 24 geophones) remains an econom-

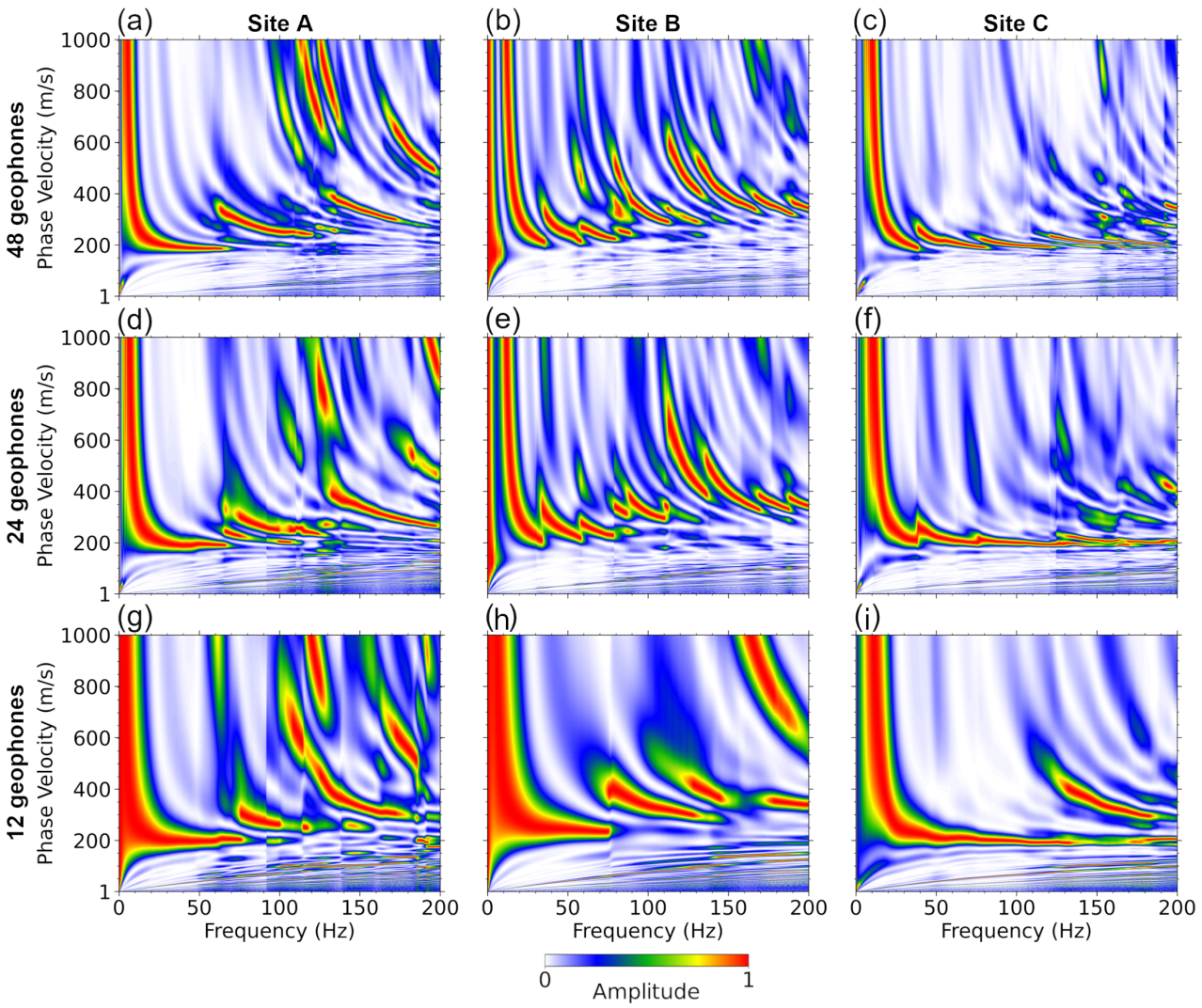

Figure 13 DIs normalized by frequency, obtained by passive seismic after FK selection, preprocessing and squared PWS ($\nu = 10$) for sites A, B, and C using a ULA of (a, b, and c) 48 geophones spaced by 50 cm, (d, e, and f) 24 geophones spaced by 75 cm, and (g, h, and i) 12 geophones spaced by 1 m.

ical option that still allows mode distinction for such high-resolution near-surface applications. Under this threshold, modes logically tend to merge, leading to the observation of only an effective dispersion that can still be used for monitoring an effective dispersion variation of the site.

This study serves as a compelling proof of concept for utilizing induced HST seismic waves as a source for 1D passive-MASW. Future work could build on this by exploring the potential of this technique for continuous monitoring of critical infrastructures, by either employing geophones or distributed acoustic sensing (DAS) with optical fibers (Bardainne et al., 2023c). Such advancements could enable real-time assessment of the mechanical stability of REs, improving safety and maintenance efficiency. Additionally, integrating dynamic cone penetrometer (DCP) data could further validate the characterization of subsurface layers, further proving the accuracy and reliability of this method (Burzawa et al., 2023).

Our findings suggest that this method could be adapted and applied to a wide range of site types, including outside of railway environments, making it a versatile and light tool for subsurface characterization in various geological and traffic conditions.

9 Acknowledgments

This research was made possible through funding from SNCF Réseau, CNRS, Sorbonne University, Mines Paris-PSL research contract, and the ANRT/Cifre-SNCF Réseau n°2021/1552 convention. The geophysical equipment was provided by the METIS laboratory at Sorbonne University. The authors express gratitude to the local teams from SNCF Réseau for their invaluable assistance and support during the fieldwork phase, and to the editor and the anonymous reviewer for their constructive comments and suggestions, which have greatly contributed to improving the quality and clarity of this paper.

Raw seismic data was read using the open-source Obspy package (docs.obspy.org), and dispersion curve inversion was conducted using the open-source software package Dinver (geopsy.org) implemented by Wathelet (2008).

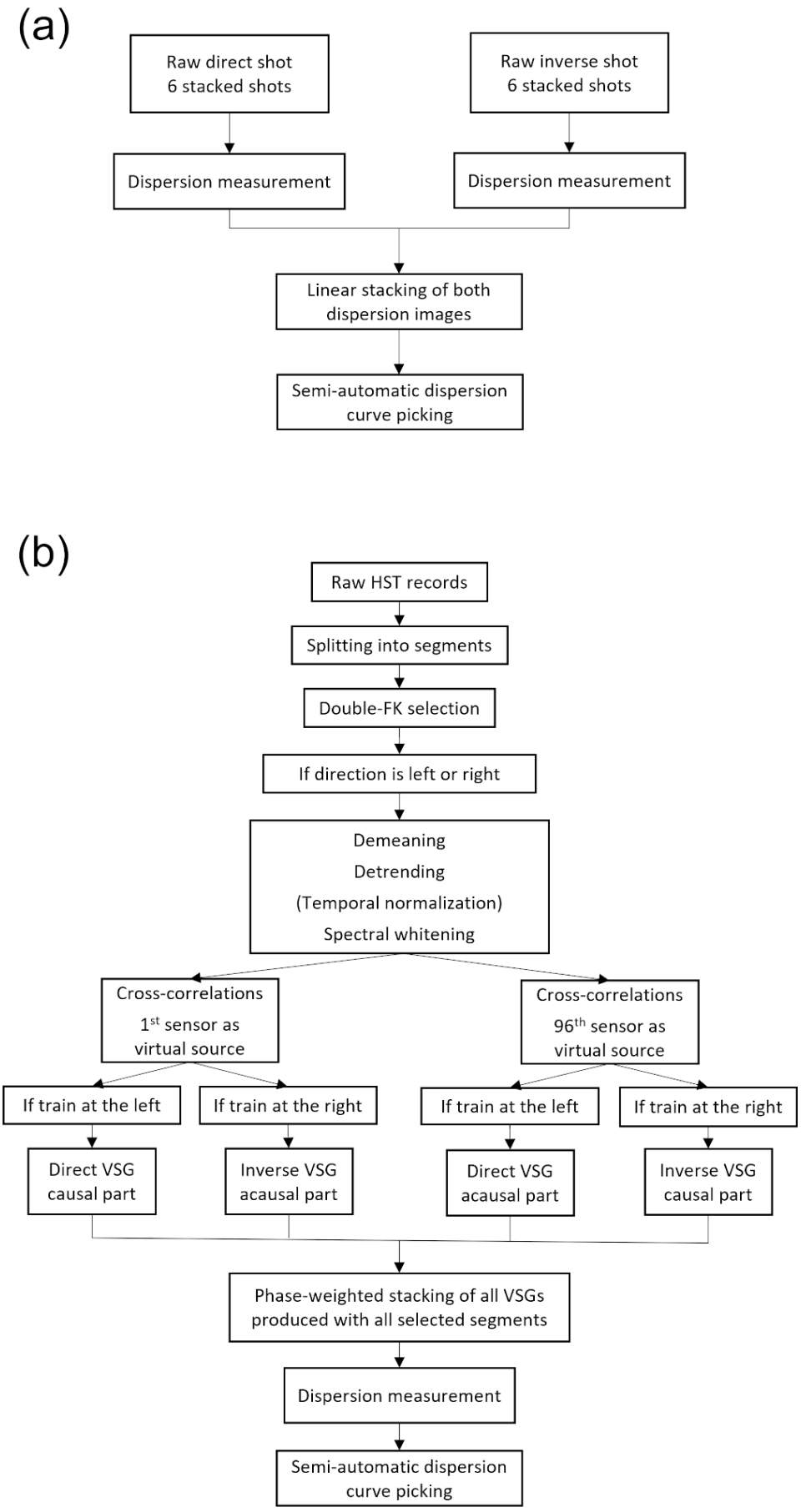

Figure A1 Schematic representation of the data processing steps for traditional active MASW (a) and passive-MASW (b).

10 Data and code availability

Raw seismic records in SEGY format and all the processed data can be accessed in Zenodo at https://doi.org/10.5281/zenodo.10245089. Processing codes are available at https://github.com/JoseCunhaTeixeira/passive-MASW.

11 Appendix

A Processing workflows and parameters

Figure A1 shows the workflows used to process active and passive data. All processing parameters used for the passive-MASW are summarized in Table A1.

Table A1 Passive-MASW data processing parameters.

Parameter	Value
HST record total number	6 (Site A)
	10 (Sites B and C)
HST recording lengths	120 s
Sampling rate	500 Hz
Segment length	5 s
Segment step	1 s
Direction ratio threshold (Λ)	0.3 (Site A)
	0.6 (Sites B and C)
PWS power (ν)	2 or 10

B Dispersion extraction

SG are transformed from the distance-time (x,t) domain into DIs in the frequency-phase velocity (f, c_f) domain using a phase-shift transform (Park et al., 1998) over the 96 geophones. Aki and Richards (1981) described a SG as a superposition of an infinite number of modes of propagation that can be written as:

$$u(x,t) = a_m(x,t)e^{i[2\pi ft - k_m(f)x]} \ , \tag{13}$$

where f is the frequency in Hz, $a_m(x,t)$ the amplitude and $k_m(f)$ the wavenumber for the mode of propagation m. Its Fourier transform is defined as:

$$\begin{aligned} U(x,f) &= \int_{-\infty}^{+\infty} u(x,t)e^{-i2\pi ft}dt \\ &= \int_{-\infty}^{+\infty} a_m(x,t)e^{i[2\pi ft - k_m(f)x]}e^{-i2\pi ft}dt \\ &= A_m(x,f)\int_{-\infty}^{+\infty} e^{-ik_m(f)x}dt \ , \end{aligned} \tag{14}$$

with $A_m(x,f)$ the spectrum amplitude of the signal. The phase-shift transformation of a SG $u(x,t)$ into the (f, c_f) domain is defined as:

$$U(f,c_f) = \int_{-\infty}^{+\infty} \frac{U(x,f)e^{i\phi x}}{|U(x,f)|}dx \ , \tag{15}$$

with $\phi = 2\pi f/c_f$ the phase-shift determined for covered phase velocity c_f and frequency f values. By injecting Equation 14 in Equation 15 we obtain:

$$U(f,c_f) = \int_{-\infty}^{+\infty} \frac{A_m(x,f)e^{-i(k_m(f)-\phi)x}}{|A_m(x,f)|}dx \ . \tag{16}$$

This last equation illustrates that the maxima of $U(f,c_f)$ will occur when $k_m(f) = \phi = f/c_f$ for each mode m. These maxima correspond to DCs that describe the variation of phase velocity with frequency for a specific mode.

Table A2 Inversion parameter space.

Layer (#)	Thickness (m)	V_P (m/s)	V_S (m/s)	ρ (kg/m^3)	ν
1	0.1-1	200-2000	10-750	2000-2500	0.1-0.5
2	0.1-1	200-2000	10-750	2000-2500	0.1-0.5
3	1-10	200-2000	10-750	2000-2500	0.1-0.5
½-space	∞	200-2000	10-1500	2000-2500	0.1-0.5

C Inversion

For each site, the DCs corresponding to the fundamental mode (M_0), the first higher mode (M_1), and the second higher mode (M_2), obtained by passive-MASW, were resampled in wavelength and inverted to generate a SW velocity model. This operation was made for DCs resulting from passive-MASW using a PWS weight $\nu = 2$ and $\nu = 10$. We use the software *Dinver* that utilizes a neighbourhood algorithm developed by Sambridge (1999) and implemented by Wathelet (2008) to solve the inverse problem parameterized with three layers above a half-space, in accordance with French HSL standards (see Fig. 1d). This method involves a stochastic exploration of a parameter space in order to search for a minimum misfit between measured and simulated DCs. The selected parameter space (see Table A2), comprising the layer thickness, Pressure-wave velocity (V_P), V_S, density ρ, and Poisson's ratio ν, aligns with previous studies (Bodet, 2019; Burzawa et al., 2023) and relies on geothechnical *a priori* information.

Out of a total 100,000 simulated models, the best 0.1% models are averaged to generate final average velocity models (see Tables 2 and 3). Figures A2a,c,e and A2b,d,f show all the simulated velocity models and corresponding DCs resulting from the inversion of M_0^{PWS2}, M_1^{PWS2}, and M_2^{PWS2}. Figures A3a,c,e and A3b,d,f show all the simulated velocity models and corresponding DCs resulting from the inversion of M_0^{PWS10}, M_1^{PWS10}, and M_2^{PWS10}. Each DC and velocity model is represented with a color depending on the misfit value MF between the experimental data (white dots and error-bars) and the simulated dispersion defined as:

$$MF = \sqrt{\sum_{i=1}^{N_f} \frac{(V_{sim_i} - V_{exp_i})^2}{N_f \sigma_i^2}} \ , \tag{17}$$

with V_{sim_i} and V_{exp_i} being the simulated and experimental phase velocities at each frequency f_i, N_f the number of frequency samples, and σ_i the phase-velocity measurement uncertainty (error-bars) at each frequency f_i.

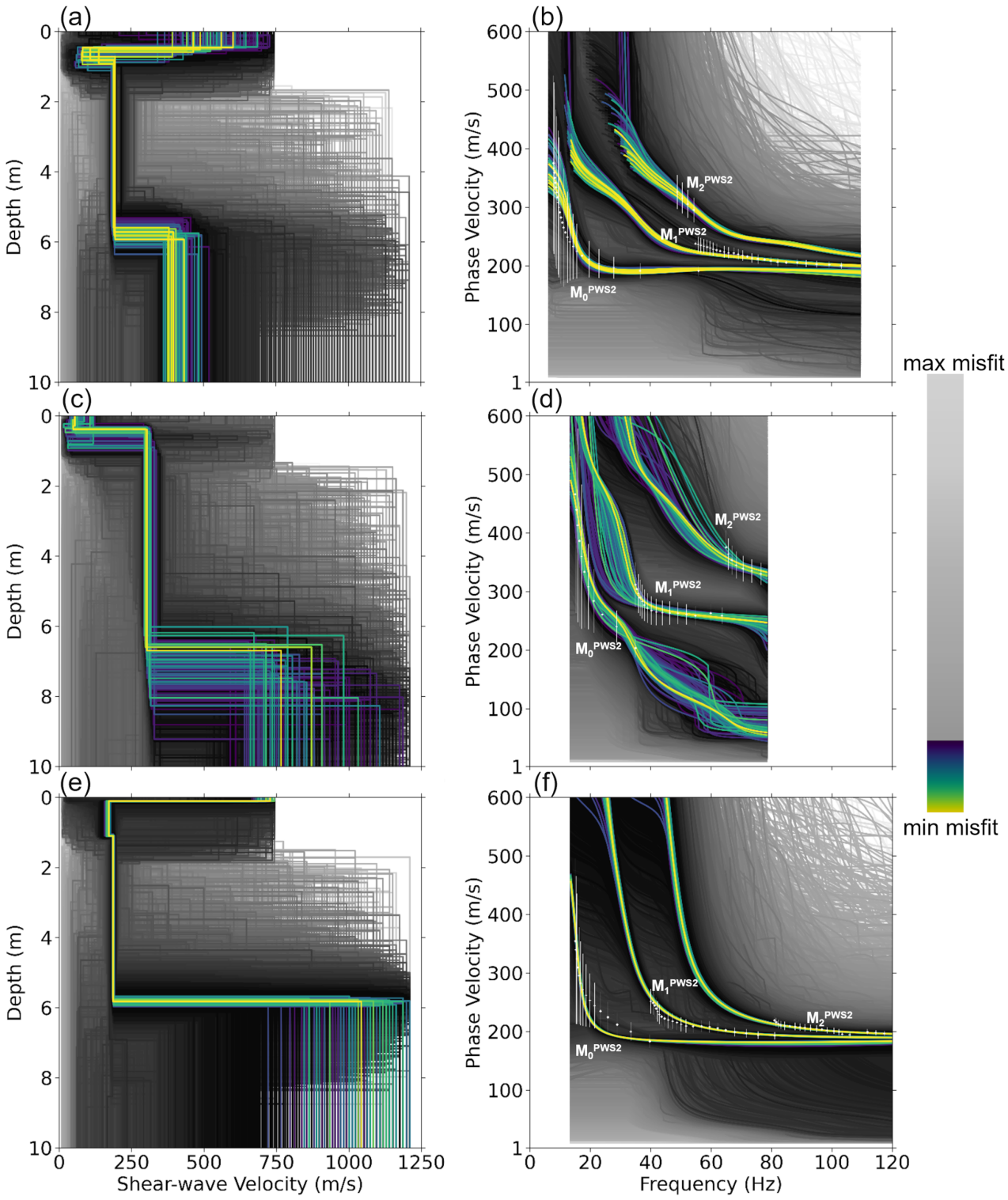

Figure A2 Neighborhood algorithm search results for the inversion of DCs obtained using passive-MASW with a PWS weight $\nu = 2$: at the top for site A (a,b), at the middle for site B (c,d), and a the bottom for site C (e,f). (a), (c) and (e) represent the modeled velocity models. (b), (d) and (f) show the modeled DCs, with error-bars, for the fundamental mode (M_0^{PWS2}), and the first and second higher modes (M_1^{PWS2} and M_2^{PWS2}). Each DC and velocity model is represented with a color depending on the misfit value between the modeled and experimental DCs (white crosses and error-bars). The best 0.1% models, in terms of misfit, are plotted in color and the rest are in gray scale.

Please note that due to the DCs providing less information at lower frequencies for sites B and C, the inversion process did not converge to a stable minimal misfit at greater depths. This variability is evident in Figures A2e,f and A3e,f, where the grey shading indicates a wider range of models far from the best fits for the deepest layer. This layer has limited influence on the dispersion curves, resulting in a wider range of potential models that might have a good fitting. Consequently, the inversion process tends to yield best-fit solutions that are

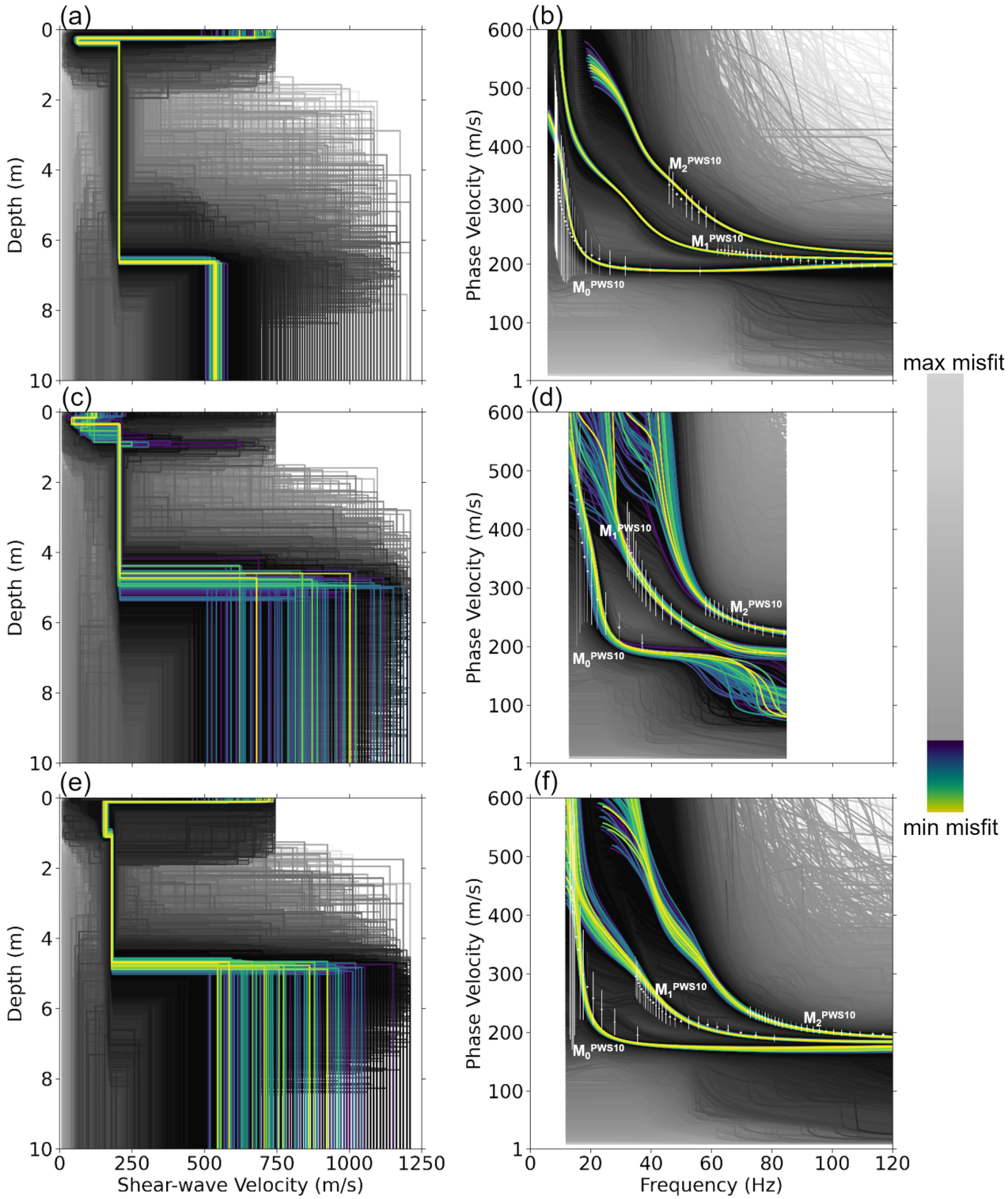

Figure A3 Neighborhood algorithm search results for the inversion of DCs obtained using passive-MASW with a PWS weight $\nu = 10$: at the top for site A (a,b), at the middle for site B (c,d), and a the bottom for site C (e,f). (a), (c) and (e) represent the modeled velocity models. (b), (d) and (f) show the modeled DCs, with error-bars, for the fundamental mode (M_0^{PWS10}), and the first and second higher modes (M_1^{PWS10} and M_2^{PWS10}). Each DC and velocity model is represented with a color depending on the misfit value between the modeled and experimental DCs (white crosses and error-bars). The best 0.1% models, in terms of misfit, are plotted in color and the rest are in gray scale.

less centered within all tested models during the inversion, as these models reflect the uncertainty rather than a more definitive central tendency. The results in terms of velocity models therefore show many equivalences at depth. However, inferred velocities are very consistent in shallow layers, thanks to the strong a priori information available on the geometry of the RE.

References

Aki, K. Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors. *Bulletin of the Earthquake Research Institute*, 35:415–456, 1957.

Aki, K. and Richards, P. G. Quantitative Seismology, Theory and Methods. *Geological Magazine*, 118(2):208, 1981. doi: 10.1017/S0016756800034439.

Artagan, S. S., Bianchini Ciampoli, L., D'Amico, F., Calvi, A., and Tosti, F. Non-destructive Assessment and Health Monitoring of Railway Infrastructures. *Surveys in Geophysics*, 41(3):447–483, May 2020. doi: 10.1007/s10712-019-09544-w.

Bakulin, A. and Calvert, R. The virtual source method: Theory and case study. *Geophysics*, 71(4):SI139–SI150, Jul 2006. doi: 10.1190/1.2216190.

Bardainne, T. and Rondeleux, B. Method and device for monitoring the subsoil of the earth under a target zone, 2018. https://patents.google.com/patent/WO2020021177A1/en?inventor=Bardainne&oq=Bardainne.

Bardainne, T., Tarnus, R., Deladerriere, N., Hallier, A., Boisson-Gaboriau, J., Valentin, J., Cai, C., and Bouhdiche, O. Permanent passive seismic monitoring of the near-surface ground beneath railways using trains as sources. volume 1, page 14. Fifth International Conference on Railway Technology: Research, Development and Maintenance, 2022. doi: doi:10.4203/ccc.1.10.14.

Bardainne, T., Cai, C., Rebert, T., Tarnus, R., and Allemand, T. Passive Seismic Monitoring Using Trains as Sources to Characterize Near-Surface and Prevent Sinkholes. volume 2023, pages 1–5. European Association of Geoscientists and Engineers, 2023a. doi: 10.3997/2214-4609.2023101262.

Bardainne, T., Tarnus, R., Vivin, L., Cai, C., Rebert, T., Allemand, T., and Toubiana, H. Stimulated Noise and Surface Wave Interferometry Processing for Hybrid Seismic Imaging. volume 2023, pages 1–5. European Association of Geoscientists and Engineers, 2023b. doi: 10.3997/2214-4609.202320206.

Bardainne, T., Vivin, L., and Tarnus, R. Railway Near-Surface Passive Seismic Using Trains as Sources and Fiber Optic Monitoring. volume 2023, pages 1–5. European Association of Geoscientists and Engineers, 2023c. doi: 10.3997/2214-4609.202320093.

Barmin, M. P., Ritzwoller, M. H., and Levshin, A. L. A Fast and Reliable Method for Surface Wave Tomography. *Pure and Applied Geophysics*, 158(8):1351–1375, Aug 2001. doi: 10.1007/PL00001225.

Barta, J. A methodology for geophysical investigation of track defects. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 224(4):237–244, Jul 2010. doi: 10.1243/09544097JRRT331.

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

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., and Yang, Y. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169(3):1239–1260, Jun 2007. doi: 10.1111/j.1365-246X.2007.03374.x.

Bensen, G. D., Ritzwoller, M. H., and Shapiro, N. M. Broadband ambient noise surface wave tomography across the United States. *Journal of Geophysical Research: Solid Earth*, 113(B5), 2008. doi: 10.1029/2007JB005248.

Bergamo, P. and Socco, L. V. P- and S-wave velocity models of shallow dry sand formations from surface wave multimodal inversion. *Geophysics*, 81(4):R197–R209, Jun 2016. doi: 10.1190/geo2015-0542.1.

Bergamo, P., Boiero, D., and Socco, L. V. Retrieving 2D structures from surface-wave data by means of space-varying spatial windowing. *Geophysics*, 77(4):EN39–EN51, Jul 2012. doi: 10.1190/geo2012-0031.1.

Bergamo, P., Dashwood, B., Uhlemann, S., Swift, R., Chambers, J. E., Gunn, D. A., and Donohue, S. Time-lapse monitoring of climate effects on earthworks using surface waves. *Gephysics*, 81(2):EN1–EN15, Mar 2016. doi: 10.1190/geo2015-0275.1.

Bodet, L. Surface waves modelling and analysis in media of increasing degrees of complexity. HDR Thesis, 2019.

Bodet, L., Abraham, O., and Clorennec, D. Near-offset effects on Rayleigh-wave dispersion measurements: Physical modeling. *Journal of Applied Geophysics*, 68(1):95–103, 2009. doi: https://-doi.org/10.1016/j.jappgeo.2009.02.012.

Bodet, L., Rahmania, I., Kyrkou, K., Wacquier, L., Dangeard, M., Pasquet, S., Vitale, Q., Dhemaied, A., Boisson-Gaboriau, J., and Nebieridze, S. Estimation in situ des variations de propriétés mécaniques de la plateforme ferroviaire par ondes sismiques de surface. SNCF Réseau - CNRS - UPMC Internal report, 2017.

Bohlen, T., Kugler, S., Klein, G., and Theilen, F. 1.5D inversion of lateral variation of Scholte-wave dispersion. *Geophysics*, 69(22):330–344, 2004. doi: 10.1190/1.1707052.

Boiero, D. and Socco, L. V. Retrieving lateral variations from surface wave dispersion curves. *Geophysical Prospecting*, Mar 2010. doi: 10.1111/j.1365-2478.2010.00877.x.

Boiero, D. and Socco, L. V. The meaning of surface wave dispersion curves in weakly laterally varying structures. *Near Surface Geophysics*, 9(6):561–570, 2011. doi: 10.3997/1873-0604.2011042.

Boisson-Gaboriau, J., Dhemaied, A., Rahmania, I., Bodet, L., Pasquet, S., Brand, L., Calon, N., Terpereau, J. M., and Kahil, S. Etude géophysique, géotechnique te historique d'une voie ferrée sur sol tourbeux. Nancy, 2016. Journées Nationales de Géotechnique et de Géologie de l'Ingénieur.

Boué, P. and Paul, A. Imagerie sismique par corrélation de bruit ambiant : du laboratoire à l'échelle globale. *Reflets de la physique*, 64:12–16, Jan 2020. doi: 10.1051/refdp/202064012.

Brenguier, F., Boué, P., Ben-Zion, Y., Vernon, F., Johnson, C., Mordret, A., Coutant, O., Share, P.-E., Beaucé, E., Hollis, D., and Lecocq, T. Train Traffic as a Powerful Noise Source for Monitoring Active Faults With Seismic Interferometry. *Geophysical Research Letters*, 46(16):9529–9536, 2019. doi: 10.1029/2019GL083438.

Burzawa, A., Bodet, L., Dhemaied, A., Dangeard, M., Pasquet, S., Vitale, Q., Boisson-Gaboriau, J., and Cui, Y. Detecting mechanical property anomalies along railway earthworks by Bayesian appraisal of MASW data. *Construction and Building Materials*, 404:133224, 2023. doi: 10.1016/j.conbuildmat.2023.133224.

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

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

Cheng, F., Xia, J., Xu, Z., Hu, Y., and Mi, B. Frequency–Wavenumber (FK)-Based Data Selection in High-Frequency Passive Surface Wave Survey. *Surveys in Geophysics*, 39(4):661–682, Jul 2018. doi: 10.1007/s10712-018-9473-3.

Cheng, F., Xia, J., Behm, M., Hu, Y., and Pang, J. Automated Data Selection in the Tau–p Domain: Application to Passive Sur-

face Wave Imaging. *Surveys in Geophysics*, 40, 09 2019. doi: 10.1007/s10712-019-09530-2.

Cheng, F., Xia, J., and Xi, C. Artifacts in High-Frequency Passive Surface Wave Dispersion Imaging: Toward the Linear Receiver Array. *Surveys in Geophysics.*, Feb 2023. doi: 10.1007/s10712-023-09772-1.

Czarny, R., Zhu, T., and Shen, J. Spatiotemporal evaluation of Rayleigh surface wave estimated from roadside dark fiber DAS array and traffic noise. *Seismica*, 2(22), Aug. 2023. doi: 10.26443/seismica.v2i2.247.

Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M., and Fink, M. Recovering the Green's function from field-field correlations in an open scattering medium (L). *The Journal of the Acoustical Society of America*, 113(6):2973–2976, Jun 2003. doi: 10.1121/1.1570436.

Donohue, S., Gavin, K., and Tolooiyan, A. Geophysical and geotechnical assessment of a railway embankment failure. *Near Surface Geophysics*, 9(1):33–44, 2011. doi: 10.3997/1873-0604.2010040.

Fontul, S., Mínguez, R., Solla, M., and Santos-Assunçao, S. *The Use of Geophysics for the Condition Assessment of Railway Infrastructure*, page 195–212. Apr 2016. doi: 10.1201/b19024-13. journalAbbreviation: Non-Destructive Techniques for the Evaluation of Structures and Infrastructure.

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., and Socco, V. Guidelines for the good practice of surface wave analysis: a product of the InterPACIFIC project. *Bulletin of Earthquake Engineering*, 16(6): 2367–2420, Jun 2018. doi: 10.1007/s10518-017-0206-7.

Gabriels, P., Snieder, R., and Nolet, G. In Situ Measurements of Shear-Wave Velocity in Sediments with Higher-Mode Rayleigh Waves. *Geophysical Prospecting*, 35(2):187–196, 1987. doi: 10.1111/j.1365-2478.1987.tb00812.x.

Haddani, Y., Breul, P., Saussine, G., Navarrete, M. A. B., Ranvier, F., and Gourvès, R. Trackbed Mechanical and Physical Characterization using PANDA®/Geoendoscopy Coupling. *Procedia Engineering*, 143:1201–1209, Jan 2016. doi: 10.1016/j.proeng.2016.06.118.

Hayashi, K., Craig, M., Kita, T., and Inazaki, T. *CMP spatial autocorrelation analysis of multichannel passive surface-wave data*, pages 2200–2204. 2015. doi: 10.1190/segam2015-5913708.1.

Jongmans, D. and Demanet, D. The importance of surface waves in vibration study and the use of Rayleigh waves for estimating the dynamic characteristics of soils. *Engineering Geology*, 34(1): 105–113, May 1993. doi: 10.1016/0013-7952(93)90046-F.

Kanasewich, E. R., Hemmings, C. D., and Alpaslan, T. Nth-root stack nonlinear multichannel filter. *Geophysics*, 38(2):327–338, Apr 1973. doi: 10.1190/1.1440343.

Lai, C. G., Rix, G. J., Foti, S., and Roma, V. Simultaneous measurement and inversion of surface wave dispersion and attenuation curves. *Soil Dynamics and Earthquake Engineering*, 22 (9):923–930, Oct 2002. doi: 10.1016/S0267-7261(02)00116-1.

Larose, E., Carrière, S., Voisin, C., Bottelin, P., Baillet, L., Guéguen, P., Walter, F., Jongmans, D., Guillier, B., Garambois, S., Gimbert, F., and Massey, C. Environmental seismology: What can we learn on earth surface processes with ambient noise? *Journal of Applied Geophysics*, 116:62–74, May 2015. doi: 10.1016/j.jappgeo.2015.02.001.

Lavoué, F., Coutant, O., Boué, P., Pinzon-Rincon, L., Brenguier, F., Brossier, R., Dales, P., Rezaeifar, M., and Bean, C. J. Understanding Seismic Waves Generated by Train Traffic via Modeling: Implications for Seismic Imaging and Monitoring. *Seismological Research Letters*, 92(1):287–300, 10 2020. doi: 10.1785/0220200133.

Le Feuvre, M., Joubert, A., Leparoux, D., and Cote, P. Passive multi-channel analysis of surface waves with cross-correlations and beamforming. Application to a sea dike. *Journal of Applied Geophysics*, 114:pp.36–51, 2015. doi: 10.1016/j.jappgeo.2014.12.014.

McFadden, P. L., Drummond, B. J., and Kravis, S. The Nth-root stack: Theory, applications, and examples. *Geophysics*, 51(10): 1879–1892, Oct 1986. doi: 10.1190/1.1442045.

McMechan, G. A. and Yedlin, M. J. Analysis of dispersive waves by wave field transformation. *Geophysics*, 46(6):869–874, Jun 1981. doi: 10.1190/1.1441225.

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., and Zhang, H. 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, Mar 2022. doi: 10.1190/geo2021-0184.1.

Mi, B., Xia, J., Xu, Y., You, B., and Chen, Y. Retrieval of surface waves from high-speed-train-induced vibrations using seismic interferometry. *Geophysics.*, May 2023. doi: 10.1190/geo2022-0603.1.

Mokhtar, T. A., Herrmann, R. B., and Russell, D. R. Seismic velocity and Q model for the shallow structure of the Arabian shield from short-period Rayleigh waves. *Geophysics*, 53(11):1379–1387, Nov 1988. doi: 10.1190/1.1442417.

Mordret, A., Landès, M., Shapiro, N. M., Singh, S. C., Roux, P., and Barkved, O. I. Near-surface study at the Valhall oil field from ambient noise surface wave tomography. *Geophysical Journal International*, 193(3):1627–1643, Jun 2013. doi: 10.1093/gji/ggt061.

Narayanan, R. M., Jakub, J. W., Li, D., and Elias, S. E. G. Railroad track modulus estimation using ground penetrating radar measurements. *NDT & E International*, 37(2):141–151, Mar 2004. doi: 10.1016/j.ndteint.2003.05.003.

Nazarian, S., Stokoe, K., and Hudson, W. Use of Spectral Analysis of Surface Waves Method for Determination of Moduli and Thicknesses of Pavement Systems. *Transportation Research Record*, 1983.

Nebieridze, S. L'apport de la géophysique à la géotechnique – L'expérience de la station d'essais géophysiques sur la LGV Nord. *Revue Française de géotechnique*, 134–135, 2011. doi: 10.1051/geotech/2011134037.

Ning, L., Xia, J., Dai, T., Liu, Y., Zhang, H., and Xi, C. High-Frequency Surface-Wave Imaging from Traffic-Induced Noise by Selecting In-line Sources. *Surveys in Geophysics*, 43(6):1873–1899, Dec 2022. doi: 10.1007/s10712-022-09723-2.

Okada, K. and Ghataora, G. S. Use of cyclic penetration test to estimate the stiffness of railway subgrade. *NDT & E International*, 35(2):65–74, Mar 2002. doi: 10.1016/S0963-8695(01)00036-6.

O'Connell, D. R. H. and Turner, J. P. Interferometric Multichannel Analysis of Surface Waves (IMASW). *Bulletin of the Seismological Society of America*, 101(5):2122–2141, 2011. doi: 10.1785/0120100230.

O'Neill, A. *Full-waveform reflectivity for modeling, inversion and appraisal of seismic surface wave dispersion in shallow site investigations*. PhD thesis, University of Western Australia, 2003.

O'Neill, A., Dentith, M., and List, R. Full-waveform P-SV reflectivity inversion of surface waves for shallow engineering applications. *Exploration Geophysics*, 34, Jun 2003. doi: 10.1071/EG03158.

Park, C., Miller, R., Laflen, D., Neb, C., Ivanov, J., Bennett, B., and Huggins, R. *Imaging dispersion curves of passive surface waves*,

pages 1357–1360. 2004. doi: 10.1190/1.1851112.

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

Park, C. B., Miller, R. D., and Xia, J. Imaging dispersion curves of surface waves on multi-channel record. *Seg Technical Program Expanded Abstracts*, pages 1377–1380, 1998. https://api.semanticscholar.org/CorpusID:7965198.

Park, C. B., Miller, R. D., and Xia, J. Multichannel analysis of surface waves. *Geophysics*, 64(3):800–808, May 1999. doi: 10.1190/1.1444590.

Pasquet, S. and Bodet, L. SWIP: An integrated workflow for surface-wave dispersion inversion and profiling. *Geophysics*, 82(6): WB47–WB61, Nov 2017. doi: 10.1190/geo2016-0625.1.

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

Rayleigh, L. On Waves Propagated along the Plane Surface of an Elastic Solid. *Proceedings of the London Mathematical Society*, s1-17(1):4–11, 1885. doi: 10.1112/plms/s1-17.1.4.

Rebert, T., Allemand, T., Bardainne, T., Cai, C., and Chauris, H. Seismic emissions from a passing train: turning ambient noise into a controlled source. In *EGU General Assembly Conference Abstracts*, EGU General Assembly Conference Abstracts, pages EGU–7289, 2023. doi: 10.5194/egusphere-egu23-7289.

Rebert, T., Bardainne, T., Allemand, T., Cai, C., and Chauris, H. Characterization of train kinematics and source wavelets from near-field seismic data. *Geophysical Journal International*, 237(2): 697–715, 2024a. doi: 10.1093/gji/ggae067.

Rebert, T., Bardainne, T., Cai, C., Allemand, T., and Chauris, H. Matched Field Processing of train vibrations for opportunistic surface wave tomography. In *EGU General Assembly Conference Abstracts*, EGU General Assembly Conference Abstracts, page 18405, 2024b. doi: 10.5194/egusphere-egu24-18405.

Rezaeifar, M., Lavoué, F., Maggio, G., Xu, Y., Bean, C. J., Pinzon-Rincon, L., Lebedev, S., and Brenguier, F. Imaging shallow structures using interferometry of seismic body waves generated by train traffic. *Geophysical Journal International*, 233(2):964–977, May 2023. doi: 10.1093/gji/ggac507.

Rhayma, N., Bressolette, P., Breul, P., Fogli, M., and Saussine, G. Reliability analysis of maintenance operations for railway tracks. *Reliability Engineering & System Safety*, 114:12–25, Jun 2013. doi: 10.1016/j.ress.2012.12.007.

Ritzwoller, M. H. and Levshin, A. L. Eurasian surface wave tomography: Group velocities. *Journal of Geophysical Research: Solid Earth*, 103(B3):4839–4878, 1998. doi: 10.1029/97JB02622.

Sambridge, M. Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space. *Geophysical Journal International*, 138(2):479–494, Aug 1999. doi: 10.1046/j.1365-246X.1999.00876.x.

Schimmel, M. and Paulssen, H. Noise reduction and detection of weak, coherent signals through phase-weighted stacks. *Geophysical Journal International*, 130(2):497–505, Aug 1997. doi: 10.1111/j.1365-246X.1997.tb05664.x.

Shapiro, N. and Campillo, M. Emergence of broadband Rayleigh waves from correlations of the ambient noise. *Geophysical Research Letters*, 8180, Apr 2004. doi: 10.1029/2004GL019491.

Sheng, Y., Mordret, A., Brenguier, F., Boué, P., Vernon, F., Takeda, T., Aoki, Y., Taira, T., and Ben-Zion, Y. Seeking Repeating Anthropogenic Seismic Sources: Implications for Seismic Velocity Monitoring at Fault Zones. *Journal of Geophysical Research: Solid Earth*, 128, Dec 2023. doi: 10.1029/2022JB024725.

Snieder, R. Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, 69(4):046610, Apr 2004. doi: 10.1103/PhysRevE.69.046610.

Socco, L. and Strobbia, C. Surface-wave method for near-surface characterization: a tutorial. *Near Surface Geophysics*, 2(4): 165–185, 2004. doi: 10.3997/1873-0604.2004015.

Socco, L. V., Boiero, D., Foti, S., and Wisén, R. Laterally constrained inversion of ground roll from seismic reflection records. *Geophysics*, 74(6):G35–G45, Nov 2009. doi: 10.1190/1.3223636.

Socco, L. V., Foti, S., and Boiero, D. Surface-wave analysis for building near-surface velocity models — Established approaches and new perspectives. *Geophysics*, 75(5):75A83–75A102, Sep 2010. doi: 10.1190/1.3479491.

Talfumière, V. and Nebieridze, S. Utilisation du bruit ambiant comme source sismique pour détecter des cavités - Gare de l'Est. In *Journées Nationales de Géotechnique et de Géologie de l'Ingénieur JNGG'08*, Nantes, 2008. https://www.cfmr-roches.org/sites/default/files/jngg/JNGG%202008%20pp%20369-376%20Talfumiere.pdf.

Tarnus, R., Bardainne, T., Michel, L., Deladerriere, N., Cai, C., Hallier, A., and Boisson-Gaboriau, J. A Case Study for Underground Imaging Using Trains as Seismic Signal to Investigate Subsidence Phenomena. volume 2022, pages 1–5. European Association of Geoscientists and Engineers, 2022a. doi: 10.3997/2214-4609.202220033.

Tarnus, R., Bardainne, T., Michel, L., Deladerrière, N., and Cai, C. A case study for railway underground imaging using trains as seismic signal for sinkhole and subsidence phenomena prevention. volume 1, page 13. Fifth International Conference on Railway Technology: Research, Development and Maintenance, 2022b. doi: 10.4203/ccc.1.10.13.

Thurber, C., Zeng, X., Thomas, A., and Audet, P. Phase-Weighted Stacking Applied to Low-Frequency Earthquakes. *Bulletin of the Seismological Society of America*, 104:2567–2572, Oct 2014. doi: 10.1785/0120140077.

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

Wapenaar, K., Draganov, D., Snieder, R., Campman, X., and Verdel, A. Tutorial on seismic interferometry. Part I: Basic principles and applications. *Geophysics*, 75:75A195–75209, Sep 2010a. doi: 10.1190/1.3457445.

Wapenaar, K., Slob, E., Snieder, R., and Curtis, A. Tutorial on seismic interferometry: Part 2 — Underlying theory and new advances. *Geophysics*, 75:75A211–75227, Sep 2010b. doi: 10.1190/1.3463440.

Wathelet, M. An improved neighborhood algorithm: Parameter conditions and dynamic scaling. *Geophysical Research Letters*, 35(9), 2008. doi: 10.1029/2008GL033256.

Weaver, R. L. and Lobkis, O. I. Diffuse fields in open systems and the emergence of the Green's function (L). *The Journal of the Acoustical Society of America*, 116(5):2731–2734, Nov 2004. doi: 10.1121/1.1810232.

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

Xu, Y., Zhang, B., Luo, Y., and Xia, J. Surface-wave observations after integrating active and passive source data. *The Leading Edge*, 32(6):634–637, 2013. doi: 10.1190/tle32060634.1.

Yang, X., Bryan, J., Okubo, K., Jiang, C., Clements, T., and Denolle,

M. A. Optimal stacking of noise cross-correlation functions. *Geophysical Journal International*, 232(3):1600–1618, Mar 2023. doi: 10.1093/gji/ggac410.

Yao, H., van Der Hilst, R. D., and de Hoop, M. V. 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, Aug 2006. doi: 10.1111/j.1365-246X.2006.03028.x.

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

Zywicki, D. J. and Rix, G. J. Mitigation of Near-Field Effects for Seismic Surface Wave Velocity Estimation with Cylindrical Beamformers. *Journal of Geotechnical and Geoenvironmental Engineering*, 131(8):970–977, 2005. doi: 10.1061/(ASCE)1090-0241(2005)131:8(970).