Lawrence Berkeley National Laboratory
Recent Work

Title

Automated Data Selection in the Tau–p Domain: Application to Passive Surface Wave Imaging

Permalink

Journal

ISSN

Authors

Cheng, F
Xia, J
Behm, M
et al.

Publication Date

DOI

Check for
updates

Automated Data Selection in the Tau–*p* Domain: Application to Passive Surface Wave Imaging

Feng Cheng[1,2] · Jianghai Xia[3] · Michael Behm[4] · Yue Hu[1] · Jingyin Pang[1]

Abstract

In the recent decades, passive surface wave methods have gained much attention in the near-surface community due to their ability to retrieve low-frequency surface wave information. Temporal averaging over a sufficiently long period of time is a crucial step in the workflow to fulfill the randomization requirement of the stationary source distribution. Because of logistical constraints, passive seismic acquisition in urban areas is mostly limited to short recording periods. Due to insufficient temporal averaging, contributions from non-stationary sources can smear the stacked dispersion measurements, especially for the low-frequency band. We formulate a criterion in the tau–*p* domain for selective stacking of dispersion measurements from passive surface waves and apply it to high-frequency (> 1 Hz) traffic noise. The criterion is based on the automated detection of input data with a high signal-to-noise ratio in a desired velocity range. Modeling tests demonstrate the ability of the proposed criterion to capture the contributions from the non-stationary sources and classify the passive surface wave data. A real-world application shows that the proposed data selection approach improves the dispersion measurements by extending the frequency band below 5 Hz and attenuating the distortion between 6 and 13 Hz. Our results indicate that significant improvements can be obtained by considering tau–*p*-based data selection in the workflow of passive surface wave processing and interpretation.

✉ Jianghai Xia
   jianghai_xia@yahoo.com; jhxia@zju.edu.cn

1 Subsurface Imaging and Sensing Laboratory, Institute of Geophysics and Geomatics, China University of Geosciences, 388 Rumo Rd, Wuhan 430074, Hubei, China

2 Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

3 School of Earth Sciences, Zhejiang University, 38 Zheda Road, Hangzhou 310027, Zhejiang, China

4 ConocoPhillips School of Geology and Geophysics, University of Oklahoma, Norman, OK 73019, USA

🍨 Springer

1 Introduction

Vertical variation of shear (S)-wave velocity structure can be derived by inverting the dispersive phase velocity of both Rayleigh and Love waves (Dorman and Ewing 1962) since surface wave dispersion is more sensitive to S-wave velocity than P-wave velocity or density for layered earth models (Xia et al. 1999). For geotechnical applications that require small-strain dynamic properties of geomaterials, surface wave methods provide an important tool due to their simplicity in the field, non-destructiveness, low cost, and relatively high vertical resolution (Xia et al. 2009). Several methods exist for estimating near-surface S-wave velocity utilizing the dispersive character of surface waves, and they can be classified in two groups related to the energy source type: active surface wave methods and passive surface wave methods.

Active surface wave methods use hammers, weight drops, electromechanical shakers, and vibrators as seismic sources. Stokoe and Nazarian (1983) and Nazarian et al. (1983) present the SASW method (spectral analysis of surface waves), which analyzes the dispersion curve of Rayleigh waves to produce near-surface S-wave velocity profiles. To improve inherent difficulties in evaluating and distinguishing signal from noise with only a pair of receivers in SASW measurements, the multichannel analysis of surface wave (MASW) method, typically using multiple geophones (i.e., 12–24), was developed (Song et al. 1989; Miller et al. 1999; Park et al. 1999; Xia et al. 1999). A growing trend in near-surface survey is toward the application of MASW for spatially 2D S-wave velocity imaging over the past two decades (e.g., Xia et al. 2003, 2009; Socco et al. 2010; Mi et al. 2017, 2018).

Passive surface wave methods use ambient seismic energy from natural or anthropogenic sources (e.g., earthquakes, ocean–seafloor interaction, traffic noise, industrial activities). The main difference from active methods is that the location and origin timing of the sources do not need to be known. Compared to the active sources, the passive sources extend the source spectrum to lower frequencies which provide the potential for deeper imaging. Aki (1957) introduces a passive surface wave method to derive the S-wave velocity structure by the inversion of spatial autocorrelation (SPAC) curves using microtremors which are composed of mainly Rayleigh wave energy. Okada (2003) presents an overview of the SPAC method and further develops a microtremor array measurement (MAM) in order to improve the flexibility of the receiver configuration and explore deeper S-wave velocity structure. Nakamura (1989) shows that the site response can be estimated from the horizontal to vertical spectral ratio (HVSR) of noise observed at the same site. Louie (2001) proposes the refraction microtremor (ReMi) method, in which passive surface waves are recorded using a linear array. The continuously recorded passive data are transformed to the frequency domain using a tau–p transform. In addition, some dispersion measurement schemes for active surface waves can also be applied on passive surface waves, e.g., phase-shift method (Park et al. 2004). The methods described above directly derive subsurface structure from passive surface waves. Seismic interferometry refers to a group of methods where the Green's function between two receivers is estimated from the correlation of their recordings (Lobkis and Weaver 2001; Campillo and Paul 2003; Shapiro and Campillo 2004). These methods have been successful in retrieving surface waves traveling between receivers based on ambient seismic energy (Sabra et al. 2005; Yao et al. 2006; Ma et al. 2008; Picozzi et al. 2009; Nakata et al. 2011; Draganov et al. 2013; Behm et al. 2014; Nakata 2016). The surface waves can be extracted using interferometric techniques such as cross-correlation (Shapiro and Campillo 2004; Stehly et al. 2006), cross-coherence (Aki 1957; Claerbout 1968; Schuster et al. 2004; Nakata et al. 2011), deconvolution (Vasconcelos and Snieder 2008a, b; Snieder et al. 2009), and multi-dimensional deconvolution

(Wapenaar et al. 2008, 2011; Van Dalen et al. 2015; Weemstra et al. 2017; Cheng et al. 2017). Interferometry techniques can enhance coherent signals from ambient energy and minimize incoherent phase-shift during dispersion measurements because cross-correlation does not reconstruct for unrelated signals (e.g., O'Connell and Turner 2011; Boschi et al. 2013; Le Feuvre et al. 2015). Cheng et al. (2016) proposed a hybrid method, that combines cross-correlation and MASW method, called multichannel analysis of passive surface waves (MAPS).

During the past few decades, considerable progress has been made toward the development of surface wave methods utilizing ambient noise, especially traffic noise (e.g., Okada 2003; Nakata et al. 2011; Behm and Snieder 2013; Chang et al. 2016). Compared to active surface wave methods, passive surface wave methods have the advantages: (1) to explore greater depth due to an extending of the source spectrum with lower frequencies; (2) to save costs associated with active sources in field operations; (3) to monitor the long-term mechanical evolution of structures and grounds for civil engineering purposes. These advantages offer significant complements to active survey methods. In that context, passive surface wave methods are also applied for data acquired with the fiber-optic distributed acoustic sensing (DAS) technique for subsurface seismic monitoring (Daley et al. 2013; Dou et al. 2017; Zeng et al. 2017).

Many passive surface wave studies focus on the mathematical motivation, e.g., the feasibility of retrieving signals from ambient energy, or as it is often referred to, from "noise" (e.g., Campillo and Paul 2003; Snieder 2004; Wapenaar and Fokkema 2006; Stehly et al. 2006; Cheng et al. 2018a), and others focus on the application, e.g., the data processing schemes (e.g., Park et al. 2004; Bensen et al. 2007; Groos et al. 2012; Fichtner 2014; Behm et al. 2014), the comparison with the active surface wave method (e.g., Cheng et al. 2015; Xia et al. 2017), and the limitations and recent improvements (e.g., Wapenaar and Ruigrok 2011; Chávez-García and Kang 2014; Luo et al. 2015, 2016; Cheng et al. 2018b; Zhou et al. 2018). There are also several papers that focus on the noise source mechanisms, e.g., the seismic signature of noise itself or the passive surface waves (e.g., McNamara and Buland 2004; Groos and Ritter 2009; Wang et al. 2014).

In the presented study, we propose a new approach to analyze the spatial and temporal noise patterns in the tau–p (intercept–slowness) domain. We formulate a criterion to classify and select passive surface wave data prior to the dispersion measurement. Two modeling tests demonstrate the motivation and workflow of the proposed method. Further insights are gained from the application to a traffic noise dataset.

In this paper, we use the term passive surface wave (PSW) for ambient seismic energy (or noise) excited by moving road vehicles and railway trains, which are also referred to as microtremors in other papers (e.g., Aki 1957; Okada 2003). We focus on high-frequency (>1 Hz) PSW because they contribute significantly to urban seismic noise in a broad frequency range from 1 Hz to more than 45 Hz with maximum amplitudes between 1 and 10 Hz (Groos and Ritter 2009). Besides, they can also provide a broad wavelength range from several meters to more than several kilometers, which is the highlighted target region for subsurface surveys for geotechnical applications.

2 Data selection in Passive Surface Wave Survey

Cheng et al. (2018b) outline a workflow for the processing of ambient noise data for the subsurface shear-wave velocity structure including splitting, whitening, dispersion measurement (e.g., roadside passive MASW method in Park et al. 2004, called PMSAW; MAPS method in Cheng et al. 2016; SPAC method in Cho et al. 2004 and

Chávez-García et al. 2006), stacking (over the frequency–velocity (*f*–*v*) domain), and inversion (Fig. 1). The stacking process serves as a temporal averaging function. Temporal averaging over a sufficiently long period of time is a crucial step in every workflow applied to ambient noise data. It aims at randomization of the noise source distribution to fulfill requirements for the stationary phase assumption (Shapiro and Campillo 2004; Snieder 2004; Sabra et al. 2005; Stehly et al. 2006). It implies splitting a sufficiently long period (e.g., months to years) of continuously recorded seismic noise into shorter segments (e.g., days or hours, even minutes) and stacking the obtained interferograms or dispersion spectra for each segment over the entire record. In practice, the record-length is always limited to save costs in field work, especially in urban environments. The limited record-length makes it challenging to provide an effectively stacked dispersion measurement from temporal averaging.

Including contributions from non-stationary sources will degrade the stacked record, if those non-stationary sources are not equally distributed in space. Recording over short time periods decreases the chance of sampling equally distributed non-stationary ambient noise sources. Thus, "the more the better" will not hold for the temporal averaging in real-world applications of high-frequency passive surface wave surveys. In order to avoid non-stationary contributions, it is important to apply data selection (quality control) before dispersion measurement into the regular workflow of the data processing.

We study the spatial and temporal noise patterns in the tau–*p* domain, and propose a simple, but stable, criterion to classify PSW by calculating the signal-to-noise ratio (SNR) of the root-mean-square (RMS) energy of the tau–*p* domain wavefield, which will be demonstrated in the following Sect. 3.

In this study, we use the terms "coherent noise" and "incoherent noise" to demonstrate the "good" and "bad" contribution of the passive surface waves to dispersion measurement, respectively (Cupillard et al. 2011). Directional noise effects or azimuthal adjustment would not be involved due to the linear receiver configuration along roads or railways (e.g., Park et al. 2004; Hayashi et al. 2016; Dou et al. 2017; Zeng et al. 2017). We do not consider the statistical or stochastic properties (e.g., Groos and Ritter 2009; Nakata and Beroza 2015; Zhong et al. 2015) because the conditions based on the
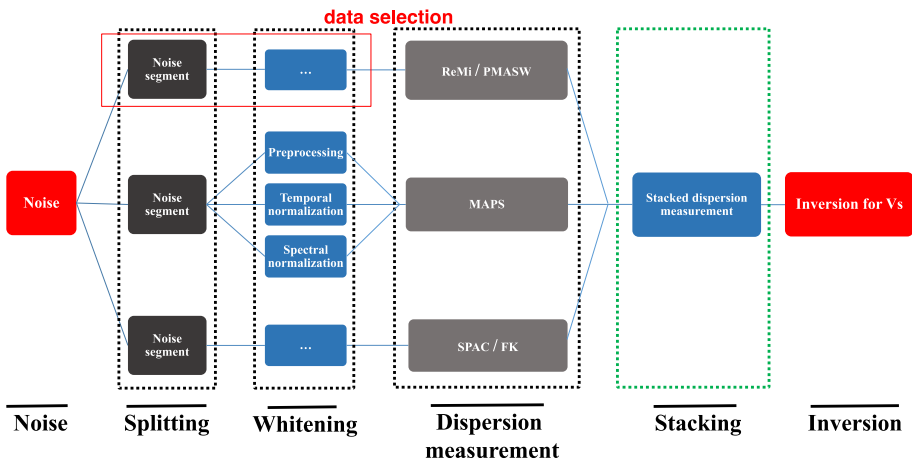

Fig. 1 Passive surface wave data processing scheme. Modified from Cheng et al. (2018b)

statistical rules would not hold under the fact that the length of the time series is limited to seconds or hours (Brockwell and Davis 2013).

3 Numerical Modeling

We carry out two simplified roadside modeling tests to investigate the influence of different source distributions on dispersion measurements and on the characteristics of PSW in the tau–*p* domain. Considering a laterally homogeneous dissipative medium, we generate the passive seismic wavefield by summing of modal components (Aki and Richards 1980). We closely follow the noise simulation procedure described in Cheng et al. (2016). The medium is a two-layer model (Table 1), whose properties are taken from Le Feuvre et al. (2015) and Cheng et al. (2016). Sources are randomly located on the road, such as the dark gray stars shown in Fig. 2a. A 15.0 Hz Ricker wavelet is chosen as the source impulse.

Figure 2b displays 5-s-long PSW in the space–time (*x–t*) domain recorded at 24 vertical-component receivers. The observed wavefield is mostly made up of Rayleigh waves. Several linear arrivals indicate the emergence of Rayleigh waves, e.g., at ~2.0 s, while some arrivals are interfered by the arrivals coming from the opposite direction, e.g., at ~1.7 s. We refer to these opposite direction arrivals as the incoherent noise due to their interference with the other (desired) arrivals. In well-randomized seismic source fields, these interferences will be canceled after sufficient temporal averaging (Snieder 2004; Campillo 2006). In practice, however, they will produce two kinds of aliasing during dispersion measurement: One is the "crossed" aliasing which has been first presented and addressed using the FK-based method in Cheng et al. (2018b); the other is the high-velocity aliasing which will be analyzed in this study. When the forward-propagating and back-propagating surface waves are not well separated, their waveforms overlap and crosstalk between these arrivals with opposite propagation directions will be introduced. In this case, the velocity scanning on the crosstalk will produce the high-velocity aliasing during dispersion measurement. It usually happens in the lower-frequency band, especially below 5 Hz, because longer wavelengths increase the chance for crosstalk.

However, the crosstalk is not always directly visible in the *x–t* domain. In order to reveal the temporal distribution and variation of the velocity information, we turn to the tau–*p* domain. We apply a slant-stacking operator (Diebold and Stoffa 1981; McMechan and Yedlin 1981) on the PSW wavefield *s*(*x*, *t*) to transform the wavefield from the *x–t* domain to the tau–*p* domain:

$$u(p, \tau) = \sum_{i=1}^{N} s\left(x_i, t = \tau + px_i\right) \tag{1}$$

where *p* is the slowness and τ is the intercept on the time axis. As for each τ, $u(p, \tau)$ will reach a maximum if the scanning slowness *p* is close to the real slowness of PSW. For convenience, we convert slowness (*p*) into velocity (*v* = *1/p*). Crosstalk in the *x–t* domain (the cyan boxes highlighted in Fig. 2b) maps the associated high-velocity aliasing (from

Table 1 Earth model parameters

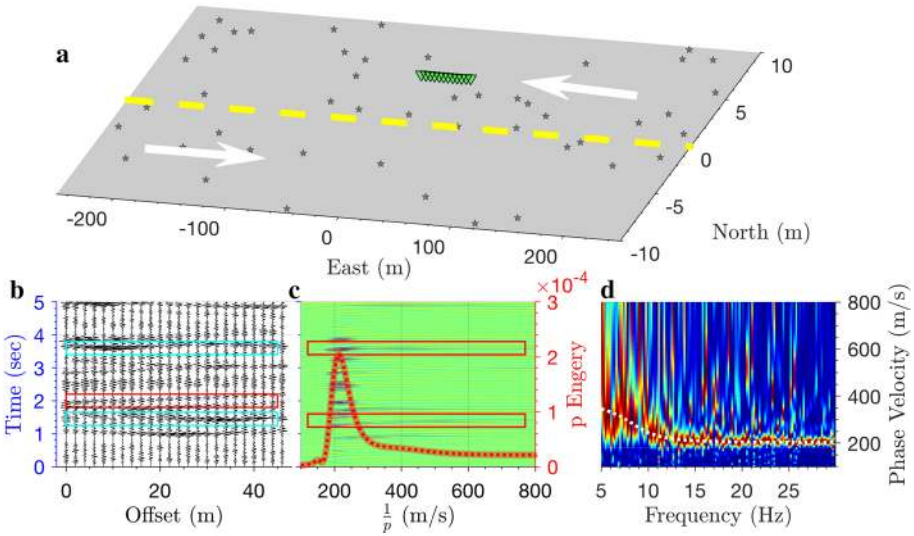| Layer	V_S (m/s)	V_P (m/s)	ρ (g/cm^3)	h (m)
1	220	380	2.0	10
2	440	760	2.0	Infinite

Fig. 2 **a** Configuration of sources and receivers. The dark gray stars denote the random source locations; the green triangles denote the linear array with a 2-m spatial interval and a 200 Hz sampling rate. The white arrows indicate the driving direction. **b** Observed noise wavefield in the *x–t* domain. The red box indicates the linear Rayleigh waves arrival. Two cyan boxes indicate the crosstalk between arrivals with opposite propagation directions. **c** The measured wavefield in the tau–*p* domain. The red boxes shows the p energy curve, the RMS-value of the wavefield $u(p, \tau)$ for each *p*. **d** The measured dispersion energy image in the *f–v* domain. The white dashed line indicates the synthetic dispersion curve

500 to 800 m/s) in the tau–*p* domain (the red boxes highlighted in Fig. 2c) at around 1.4 s and 3.6 s. The coherent noise at around 2.0 s indicates an expected apparent velocity range between 150 and 400 m/s, and this velocity range is considered as the "signal window." Next, we calculate the root-mean-square (RMS) value of the wavefield $u(p, \tau)$ along the p direction.

$$\Delta(p) = \sqrt{\frac{\sum_i^N u(p, \tau_i)^2}{N}} \tag{2}$$

$\Delta(p)$ reflects the effective wavefield energy for each slowness *p*, and refer to it as *p* energy curve (the red dashed line in Fig. 2c). The dominant peak of *p* occurs at the signal window, which means the modeled PSW window is dominated by coherent noise. This is consistent with the dispersion energy image in Fig. 2d, which is measured using the roadside passive MASW method (Park et al. 2004).

In comparison, we add interferences (red stars in Fig. 3a) into the same model with some near-field sources around the receivers (Fig. 3a). It is hard to tell any significant changes between the wavefield in the *x–t* (Figs. 2b, 3b), but the wavefield in the tau–*p* domain (Fig. 3c) does present a distinct increase in the high-velocity aliasing, which indicates the quality of the modeled PSW decreases with increasing of crosstalk. Although the peak of *p* energy is still located at the signal window (the red dashed line in Fig. 3c), the dispersion measurement is less clear compared to that measured from the wavefield without near-field sources (Fig. 3d). We introduce the concept of
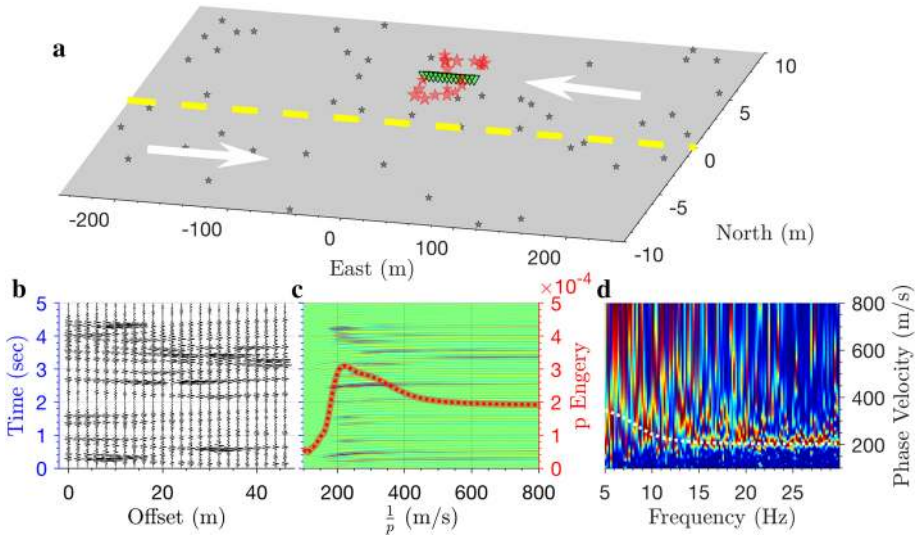
Fig. 3 **a** Configuration of sources and receivers. **b** Observed noise wavefield in the *x–t* domain. **c** The measured wavefield in the tau–*p* domain. **d** The measured dispersion energy image in the *f–v* domain

signal-to-noise ratio (SNR) to quantify the quality of PSW by calculating SNR of p energy curve, called factor ϕ:

$$\phi = \frac{\max\left\{\Delta(p)|_{v_1 < 1/p < v_2}\right\}}{\text{rms}\left\{\Delta(p)|1/p < v_1, 1/p > v_2\right\}} \tag{3}$$

where v_1 and v_2 are used to define the signal window (the coherent noise with $v_1 < v < v_2$).

Figure 4 displays the normalized p energy curves from the first model (red line) and from the second model (black line), and the signal window between 150 and 400 m/s (two dashed lines). This signal window indicates the most common phase velocity range for the near-surface structure in urban areas. In practice, we suggest to draw all the p energy curves together to locate the main p energy peak v_0 and define the signal window with $[v_0 - 100, v_0 + 100]$. Using Eq. 3, we calculate SNR of each p energy curve with $\phi_1 = 8.78$

Fig. 4 The normalized p energy curves determined from the first and the second models. Two black vertical dashed lines indicate the signal window, 150–400 m/s, used for the calculation of factor ϕ
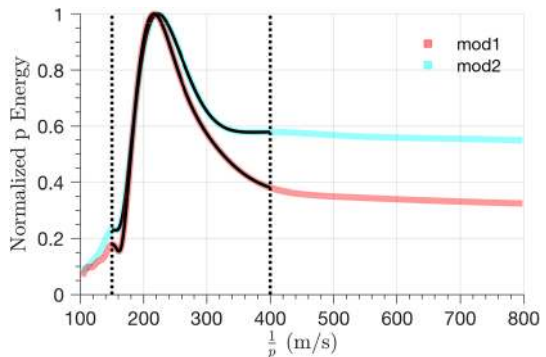
and $\phi_2 = 1.62$, which is also consistent with the corresponding dispersion measurements (Figs. 2d, 3d). It shows that we can use a threshold method to select the segments that satisfy our expectations on dispersion measurement (e.g., imaging resolution, reliability, accuracy). More details about the threshold method for factor ϕ will be described in the following section.

4 Application to Field Data

The passive surface wave survey was carried out in the city of Yueyang, China (Fig. 5). A linear array with 24 receivers (4.5 Hz vertical component) was deployed along the Beijing–Guangzhou railway (white dot-dashed line in Fig. 5), which is one of the busiest railways in China for both passenger trains and freight trains. The trace interval was fixed to 10.0 m, and the sampling interval was 2.0 ms. Thirty-min-long continuous PSW data were recorded from local time 12:15–12:45 on May 6, 2016. After this, an active shot was also recorded at the center of the same array with a 20 kg rock as shown in the red box in Fig. 5.

We apply the following workflow for automatic data selection of PSW: ① Divide the consecutive PSW into a series of 5-s-long noise segments overlapping by 50%; ② calculate the factor ϕ for each segment; ③ filter the segments with a threshold method; ④ disperse measurements on the selected segments with higher ϕ; ⑤ stack all the selected dispersion measurements and pick the final dispersion curve; ⑥ invert for 1D Vs profile; ⑦ construct a pseudo-2D Vs map. PMASW method (Park et al. 2004) is implemented for dispersion measurement. We use 12-trace-shot bins for dispersion imaging, and roll along the line with 1-trace interval.

Data examples from segment 431 (left panel of Fig. 6) and 412 (right panel of Fig. 6) illustrate how we calculate the factor ϕ. First, the x–t domain (Fig. 6a1, a2) is transformed

Fig. 5 An aerial photograph of the survey along the Beijing–Guangzhou Railway, Yueyang, China. The white dashed line indicates the railway line, and the red dashed line indicates the survey line

to the tau–p domain (Fig. 6b1, b2). In the next step, we plot the p energy curves using Eq. 2 (Fig. 7). Figure 7a displays the p energy distribution for the whole segments and indicates a main energy peak at $v_0 = 300$ m/s. Thus, we define a signal window, 200–400 m/s, as the dashed lines show in Fig. 7a and b. Finally, the factor ϕ is calculated using Eq. 3 ($\phi_{412} = 12.74$; $\phi_{431} = 1.30$). The measured ϕ-values are also consistent with the measured dispersion images (Fig. 6c1, c2). If we do not reject the segment 431, it will deteriorate the stacked dispersion measurement by introducing high-velocity aliasing below 5 Hz and by distorting the dispersion energy within the medial frequency band (6–13 Hz).

To identify the frequency- and time-dependent behavior of source processes, we apply short-time Fourier transform (STFT), with a 5-s time window, on the 30-min recording (Fig. 8a). The obtained spectrogram in Fig. 8b reveals a regular pattern of higher power spectral density (PSD) while trains passed by and lower PSD during quiet time. Traffic-induced seismic energy contributes significantly to the observed PSW in a broad frequency range from ∼2 Hz to more than 50 Hz. The maximum amplitudes occur at around 10 Hz and tend to slowly decrease with the increasing frequency. Long lasting
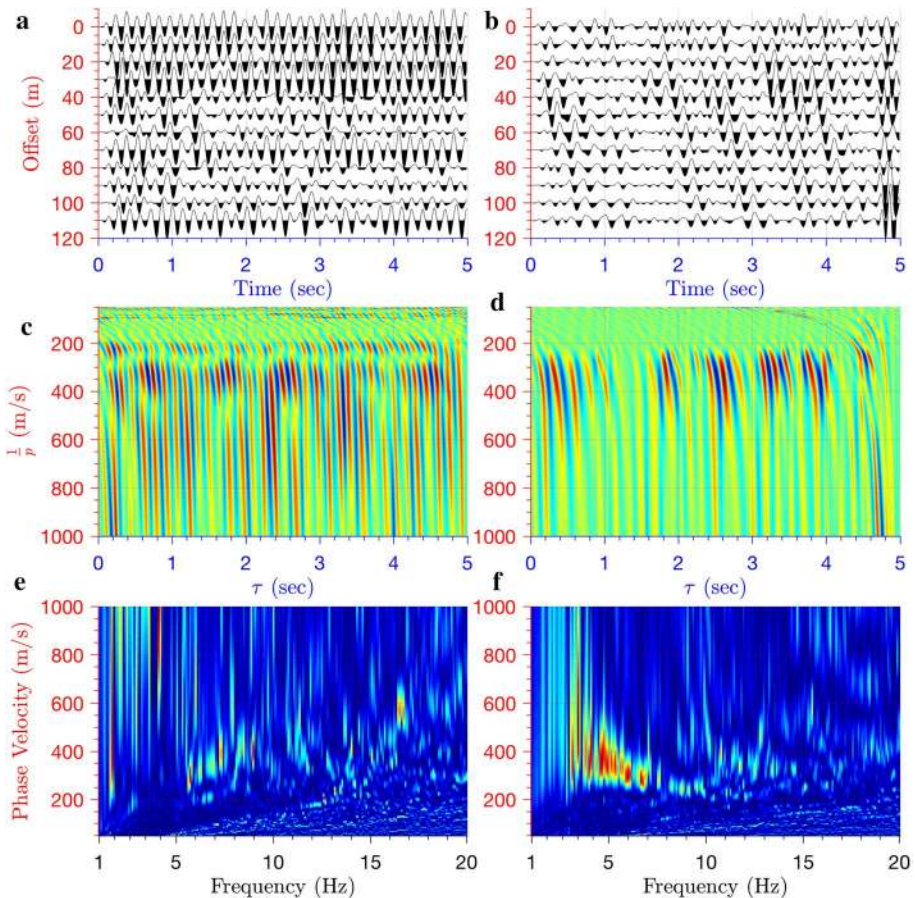

Fig. 6 Wavefield images in the x–t domain (**a**, **b**), the tau–p domain (**c**, **d**) and the f–v domain (**e**, **f**) for segments 431 (left panel) and 412 (right panel)
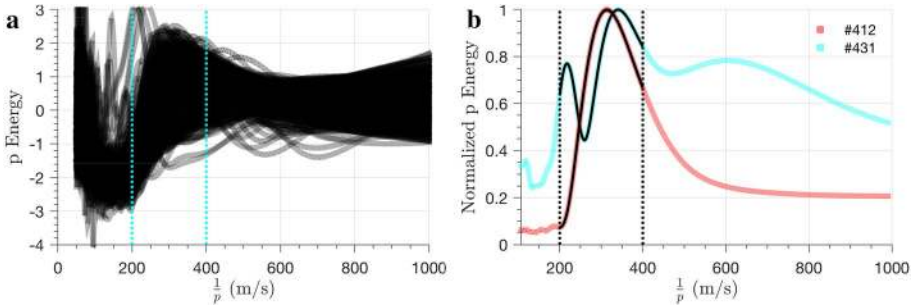
Fig. 7 a *p* energy curves distribution for the whole segments. Two cyan dashed lines define the signal window where the main energy peak locates at 300 m/s. **b** The normalized p energy curves determined from segments 431 and 412. Two black dashed lines indicate the signal window, 200–400 m/s
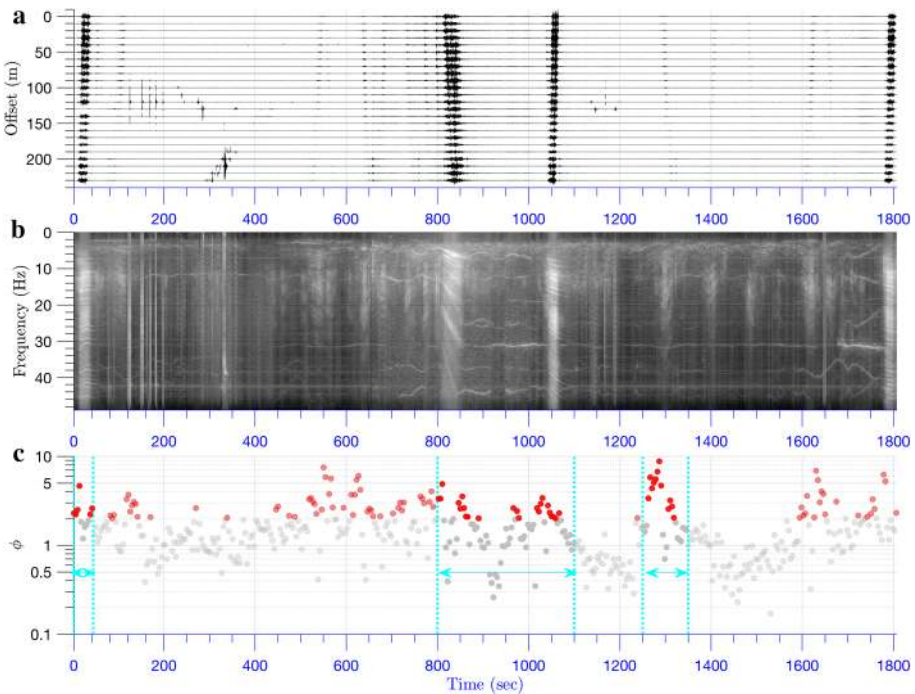

Fig. 8 a Observed noise wavefield in the *x–t* domain with a 30 min duration. **b** STFT spectrogram of the continuous record with a 5-s window. **c** The curve determined from the continuous record. The red dots indicate the "high quality" segments above the threshold 2, while the gray dots indicate the "low quality" segments below the threshold. The cyan lines highlight the time windows for further analysis

and very narrow-band signals above 2 Hz, recognized as horizontal lines of increased PSD in the spectrograms (e.g., 30.7 Hz, 42.1 Hz), are sinusoidal-type seismic waves most likely excited by rotating machinery such as train engines (Kar and Mohanty 2006; Groos and Ritter 2009; Nedilko 2016).

We calculate the factor ϕ for each noise segment and select the noise segments with high ϕ-values for the further dispersion measurement. Figure 8c displays the measured factor ϕ for all noise segments along the time direction, which reflects changes in the quality of PSW in terms of coherency. In general, the higher factor ϕ indicates the higher quality PSW, which would contribute to the dispersion measurement, and vice versa. However, the total number of segments is limited due to the limited recording length. In this case, using only selected segments for dispersion stacking will decrease with the increasing threshold value due to lack of data. Therefore, we need a trade-off solution between the accuracy and the stability of dispersion measurement stacking. We set a threshold value $\phi_0 = 2.0$, from which the factor ϕ increases rapidly as indicated by the red dots peak between 1250 and 1400 s (Fig. 8c). Only segments with higher ϕ-value (red dots in Fig. 8c) are kept for further dispersion measurement.

Compared to the raw stacked dispersion image with all of the segments involved (Fig. 9a1), the selective stacked result is significantly improved with much clearer and more continuous dispersion energy trend (Fig. 9b1). The information retrieved below 5 Hz is important for V_s model construction at larger depths. Besides, our approach also improves higher mode coherency which will contribute to the stability and accuracy of the subsequent inversion process (Xia et al. 2000, 2003; Luo et al. 2007). We also present the active surface wave inversion result here. Active data acquisition is challenged by the ever-present traffic noise (Fig. 9a2), and the dispersion measurements are affected by a distorted fundamental mode and the lack of higher modes (Fig. 9b2). For comparison, Fig. 10a exhibits all these picked dispersion curves together, as well as the inverted one. The blue dots indicate the biased dispersion curve from the raw result (Fig. 9a1) with a higher fundamental mode below 5 Hz and an abnormal bump around 12 Hz. Compared to the black dots in
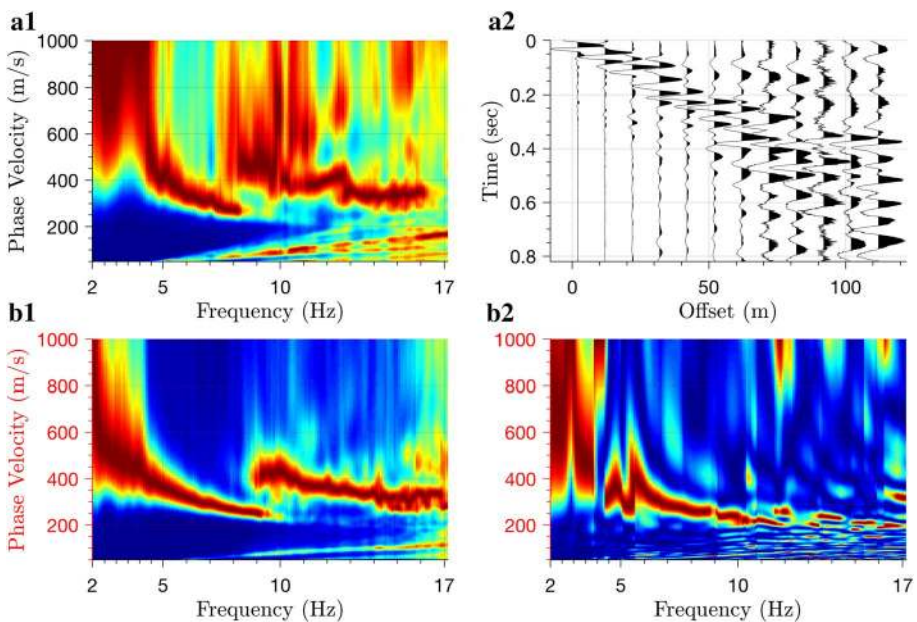

Fig. 9 a1 Dispersion measurement before data selection. **b1** Dispersion measurement after data selection. **a2** Wavefield of the active surface wave in the *x–t* domain. **b2** Dispersion measurement of the active surface wave
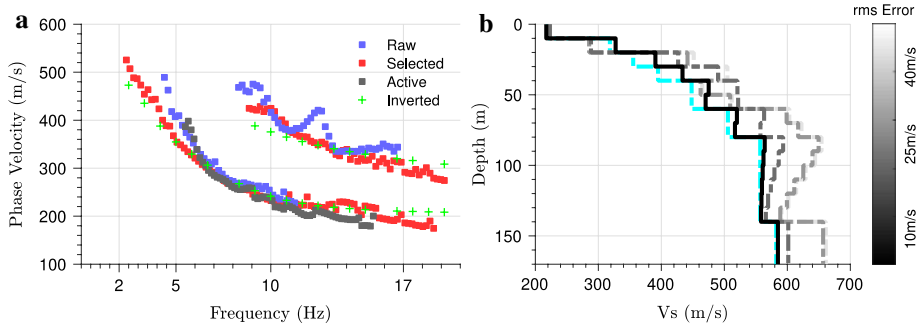
Fig. 10 a The dispersion curves picked from the dispersion images in Fig. 9. The blue dots denote the dispersion measurement before data selection. The red dots denote the result after data selection. The gray dots denote the result of active surface wave. The green crosses denote the forward dispersion curves from the last iteration of the inversion. **b** The inverted 1D V_s models from the picked dispersion curves after data selection (the red curves in a). The cyan dashed line indicates the initial V_s model, the gray dashed lines indicate the inverted V_s model for each iteration, and the black solid line indicates the final V_s model

the active result, the red dots extend the critical lower-frequency band from 5 to 2 Hz and retrieve more realistic higher mode information. Both the fundamental and higher mode data from the selected stacked result (the red curves) were used for the final inversion using the damped least-squares method and the singular-value decomposition (SVD) technique (Xia et al. 1999, 2003). The black line in Fig. 10b shows the final inverted 1D Vs model when the RMS error reaches the error tolerance (10 m/s) after three iterations. The cyan line indicates the initial V_s model created by the algorithm presented in Xia et al. (1999).

Figure 11 maps the final pseudo-2D V_s section under the survey line. Although the resolution of the V_s section is slightly low, a flat bedrock interface is delineated with confidence at 65 m deep (the dashed line) and low velocity zones (LVZs) are outlined with the contour line for $V_s = 400$ m/s. These LVZs were supposed to be the indicator of the permeable fractures (e.g., Saito et al. 2004), and these sites need to be monitored considering the long-term safety of the railway transportation system. The overburden is mainly made of clays, but the south side is more loose and wet and used for planting vegetables, and the north side is more solid and dry as shown in the photographs at the bottom of Fig. 11. These soil conditions are consistent with the inverted pseudo-2D Vs model.

5 Discussion

The proposed tau–p-based method aims to establish a classification and selection criterion for PSW. Data segments with minimum influence from non-stationary sources distribution are identified and used for further dispersion measurement. This is a general data selection processor that is also applicable to other surface wave methods such as ReMi (Louie 2001), SPAC (Chávez-García et al. 2006) or hybrid methods based on interferometry technique like MAPS (Cheng et al. 2016; Pang et al. 2019). In order to apply the proposed method on interferometry methods like MAPS, data selection should be performed after cross-correlation because cross-correlation could contribute to the retrieval of coherent surface waves, and it would avoid the potential waste of the limited data.

Figure 8 presents a different data segment of train-induced seismic noise. We recorded the timeline for four slow-speed (80–120 km/h) trains that passed by at around
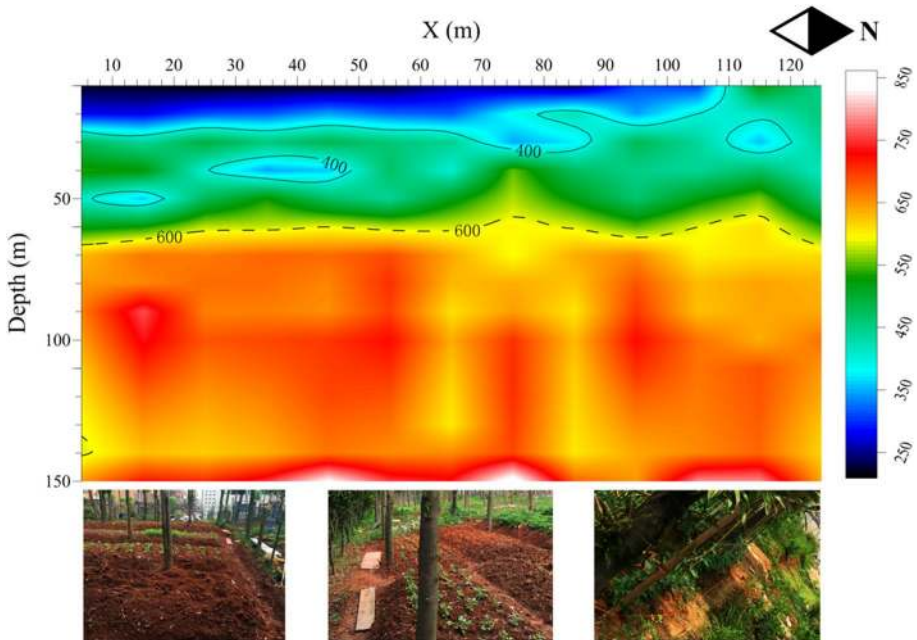
Fig. 11 The final pseudo-2D Vs section. The dash line is inferred as surface of the bedrock. The solid lines indicate the potential "dangerous" areas, the low velocity zone (LVZ). The bottom panel displays the real-site pictures for reference at the south end, the middle, the north end, respectively

$t1 = 20$ s, $t2 = 820$ s, $t3 = 1040$ s, $t4 = 1780$ s, respectively (Fig. 12), to connect the seismic signature with the associated train event. We did not have time to take photographs for the China Railways High-speed (CRH) trains whose design speed is about 380 km/h, but our sensors did capture their footprints, e.g., at around 640 s and 1290 s (Fig. 8a).

It takes about 10 s for passenger train A, as well as passenger train B and C, to pass through the line from the end side, but about 15 s for the freight train (Fig. 8a). The freight train moves slower than passenger trains. Two parallel events are detected on the time window from 820 to 860 s (Fig. 8a) with different loads on the wagons (highlighted by the yellow box in Fig. 12). The first half of the wagons were loaded with heavy stone for construction usage, while the last half of wagons were loaded with light oil.

In order to further understand the seismic signature of moving trains, three time windows with different train events (indicated by the cyan dashed line in Fig. 8c) are presented in Fig. 13. Considering only 50-s time window for trains arriving and departing, we get an M-shaped change for the $\phi - t$ curves (Fig. 13). Table 2 presents the time windows for the M-shaped change for the $\phi - t$ curves of four trains. We suppose this phenomenon is caused by the transfer of energy between ballistic waves and coda waves as the distance changes. In the far distance, the recorded wavefields are mostly weak coda waves. In the intermediate distance, the energy of the coda waves increases with the trains arriving or departing. In the short distance, there is a transfer of energy from coda waves to ballistic waves as the distance decreases. It is a common practice in seismology to use the seismic coda of earthquakes for coda wave imaging (Aki 1969; Aki and Chouet 1975; Campillo and Paul 2003; Snieder 2004, 2006; Paul et al. 2005). In
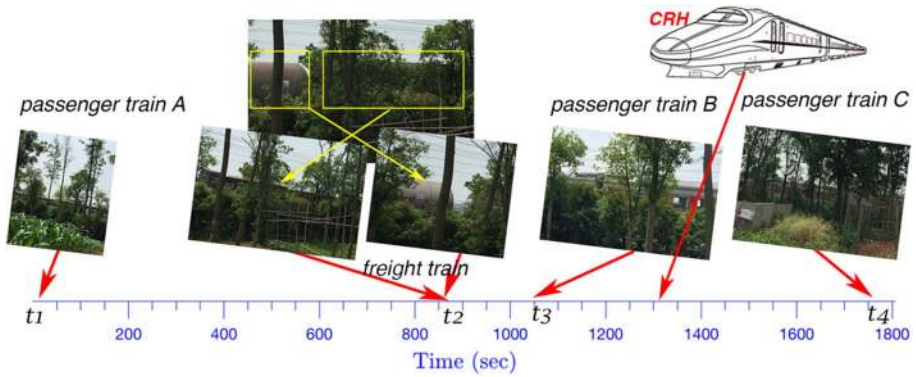
Fig. 12 Timeline for four trains that passed through our survey line at $t1 = 20$ s, $t2 = 820$ s, $t3 = 1040$ s, $t4 = 1780$ s, respectively. The yellow boxes show the mixed wagon train, with the first half loading heavy stone and the last half loading light oil. We did not capture the CRH trains' photograph due to their high speed
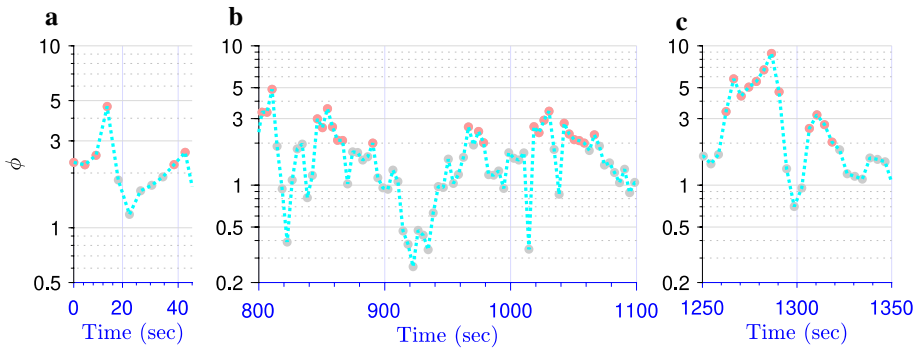

Fig. 13 The zoomed in view of the highlighted windows in Fig. 8c

Table 2 Time windows for each process change in ϕ in Fig. 8

ϕ change	Train A (s)	Freight train (s)	Train B (s)	CRH (s)
↑	0–16	800–810	1020–1035	1265–1280
↓	16–20	810–820	1035–1040	1280–1300
↑	20–40	820–840	1040–1045	1300–1315
↓	40–45	840–850	1045–1060	1315–1330

seismology, the scattered waves of the coda are usually considered as random, as well as seismic noise. The time window between which two face-to-face trains get close to each other shows a V-shaped change in factor ϕ (e.g., 860–960 s), because the left-going and right-going waves are out of phase with each other and the coupled PSW will undergo destructive interference.

6 Conclusions

We established a simple, but stable, criterion in the tau–p domain to select data segments with minimum contribution from non-stationary sources. Modeling tests and real-world application show that the proposed data selection technique in the tau–p domain performs well in improving the dispersion measurement of the passive surface waves. The real-world application also indicates that the $\phi - t$ curve, a by-product of the data selection, provides a novel view to understand and track the time-variant processes for seismic signatures of traffic noise.

In particular, the approach improves the dispersion measurements for datasets of limited recording length. This provides a possibility for very short (tens of seconds to several minutes) passive surface wave surveys, which will significantly reduce costs in field work.

References

Aki K (1957) Space and time spectra of stationary stochastic waves with spectral reference to microtremors. Bull Earthq Res Inst 35:415–456

Aki K (1969) Analysis of the seismic coda of local earthquakes as scattered waves. J Geophys Res 74:615–618

Aki K, Chouet B (1975) Origin of coda waves: source, attenuation, and scattering effects. J Geophys Res 80:3322

Aki K, Richards PG (1980) Quantitative seismology: theory and methods, vol 2. Freeman, S. Francisco, New York

Behm M, Snieder R (2013) Love waves from local traffic noise interferometry. Lead Edge 32(6):628–632

Behm M, Leahy M, Snieder R (2014) Retrieval of local surface wave velocities from traffic noise: an example from the La Barge basin (Wyoming). Geophys Prospect 62:223–243

Bensen GD, Ritzwoller MH, Barmin MP, Levshin AL, Lin F, Moschetti MP, Shapiro NM, Yang Y (2007) Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. Geophys J Int 169:1239–1260

Boschi L, Weemstra C, Verbeke J, Ekström G, Zunino A, Giardini D (2013) On measuring surface wave phase velocity from station-station cross-correlation of ambient signal. Geophys J Int 192(1):346–358

Brockwell PJ, Davis RA (2013) Time series: theory and methods. Springer, Berlin

Campillo M (2006) Phase and correlation in "random" seismic fields and the reconstruction of the green function. Pure Appl Geophys 163:475–502

Campillo M, Paul A (2003) Long-range correlations in the diffuse seismic coda. Science 299(5606):547–549

Chang JP, De Ridder SAL, Biondi BL (2016) Case History High-frequency Rayleigh-wave tomography using traffic noise from Long Beach, California. Geophysics 81(2):B43–B53

Chávez-García FJ, Kang TS (2014) Lateral heterogeneities and microtremors: limitations of HVSR and SPAC based studies for site response. Eng Geol 174:1–10

Chávez-García FJ, Rodríguez M, Stephenson WR (2006) Subsoil structure using SPAC measurements along a line. Bull Seismol Soc Am 96:729–736

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. J Appl Geophys 117:126–135

Cheng F, Xia J, Luo Y, Xu Z, Wang L, Shen C, Liu R, Pan Y, Mi B, Hu Y (2016) Multi-channel analysis of passive surface waves based on cross-correlations. Geophysics 81(5):EN57–EN66

Cheng F, Draganov D, Xia J, Behm M, Hu Y (2017) Deblurring directional-source effects for passive surface-wave surveys using multidimensional deconvolution. In: AGU fall meeting abstracts, p #S21C-0736

Cheng F, Draganov D, Xia J, Hu Y, Liu J (2018a) Q-estimation using seismic interferometry from vertical well data. J Appl Geophys 159:16–22

Cheng F, Xia J, Xu Z, Hu Y, Mi B (2018b) FK-based data selection in high-frequency passive surface wave survey. Surv Geophys 39:661–682

Cho I, Tada T, Shinozaki Y (2004) A new method to determine phase velocities of Rayleigh waves from microseisms. Geophysics 69(6):1535–1551. https://doi.org/10.1190/1.1836827

Claerbout JF (1968) Synthesis of a layered medium from its acoustic transmission response. Geophysics 33:264–269

Cupillard P, Stehly L, Romanowicz B (2011) The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise. Geophys J Int 184(3):1397–1414

Daley TM, Freifeld BM, Ajo-Franklin J, Dou S, Pevzner R, Shulakova V, Kashikar S, Miller DE, Goetz J, Henningers J, Lueth S (2013) Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring. Lead Edge 32(6):699–706

Diebold JB, Stoffa PL (1981) The travel time equation, tau–p mapping, and inversion of common midpoint data. Geophysics 46(3):238–254

Dorman J, Ewing M (1962) Numerical inversion of seismic surface wave dispersion data and crust-mantle structure in the New York-Pennsylvania Area. J Geophys Res 67(9):3554

Dou S, Lindsey N, Wagner AM, Daley TM, Freifeld B, Robertson M, Peterson J, Ulrich C, Martin ER, Ajo-Franklin JB (2017) Distributed acoustic sensing for seismic monitoring of the near surface: a traffic-noise interferometry case study. Sci Rep 7(1):11620

Draganov D, Campman X, Thorbecke J, Verdel A, Wapenaar K (2013) Seismic exploration-scale velocities and structure from ambient seismic noise (> 1 Hz). J Geophys Res Solid Earth 118:4345–4360

Fichtner A (2014) Source and processing effects on noise correlations. Geophys J Int 197(3):1527–1531

Groos JC, Bussat S, Ritter JRR (2012) Performance of different processing schemes in seismic noise cross-correlations. Geophys J Int 188(2):498–512

Groos JC, Ritter JRR (2009) Time domain classification and quantification of seismic noise in an urban environment. Geophys J Int 179(2):1213–1231. https://doi.org/10.1111/j.1365-246X.2009.04343.x

Hayashi K, Cakir R, Walsh TJ, State W (2016) Comparison of dispersion curves and velocity models obtained by active and passive surface wave methods. In: 2016 SEG annual meeting. society of exploration geophysicists, pp 4983–4988

Kar C, Mohanty AR (2006) Monitoring gear vibrations through motor current signature analysis and wavelet transform. Mech Syst Signal Process 20:158–187

Le Feuvre M, Joubert A, Leparoux D, Côte P (2015) Passive multi-channel analysis of surface waves with cross-correlations and beamforming. Application to a sea dike. J Appl Geophys 114:36–51

Lobkis OI, Weaver RL (2001) On the emergence of the Green's function in the correlations of a diffuse field. J Acoust Soc Am 110:3011–3017

Louie J (2001) Faster, better: shear-wave velocity to 100 meters depth from refraction microtremor arrays. Bull Seismol Soc Am 91:347–364

Luo Y, Xia J, Liu J, Liu Q, Xu S (2007) Joint inversion of high-frequency surface waves with fundamental and higher modes. J Appl Geophys 62:375–384

Luo Y, Yang Y, Xu Y, Xu H, Zhao K, Wang K (2015) On the limitations of interstation distances in ambient noise tomography. Geophys J Int 201(2):652–661

Luo S, Luo Y, Zhu L, Xu Y (2016) On the reliability and limitations of the SPAC method with a directional wave field. J Appl Geophys 126:172–182

Ma S, Prieto GA, Beroza GC (2008) Testing community velocity models for southern California using the ambient seismic field. Bull Seismol Soc Am 98:2694–2714

McMechan GA, Yedlin MJ (1981) Analysis of dispersive waves by wave field transformation. Geophysics 46:869–874

McNamara DE, Buland RP (2004) Ambient noise levels in the continental United States. Bull Seismol Soc Am 94(4):1517–1527

Mi B, Xia J, Shen C, Wang L, Hu Y, Cheng F (2017) Horizontal resolution of multichannel analysis of surface waves. Geophysics 82(3):EN51–EN66

Mi B, Xia J, Shen C, Wang L (2018) Dispersion energy analysis of Rayleigh and Love waves in the presence of low-velocity layers in near-surface seismic surveys. Surv Geophys 39:271–288

Miller RD, Xia J, Park CB, Ivanov J (1999) Multichannel analysis of surface waves to map bedrock. Lead Edge 18:1392–1396

Nakamura Y (1989) A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. Q Rep RTRI 30:25–33

Nakata N (2016) Near-surface S-wave velocities estimated from traffic-induced Love waves using seismic interferometry with double beamforming. Interpretation 4:SQ23–SQ31

Nakata N, Beroza GC (2015) Stochastic characterization of mesoscale seismic velocity heterogeneity in Long Beach, California. Geophys J Int 203(3):2049–2054

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

Nazarian S, Stokoe KH II, Hudson WR (1983) Use of spectral analysis of surface waves method for determination of moduli and thicknesses of pavement systems. Transp Res Rec 930:38–45

Nedilko B (2016) Seismic detection of rockfalls on railway lines. Doctoral dissertation, University of British Columbia

O'Connell DRH, Turner JP (2011) Interferometric multichannel analysis of surface waves (IMASW). Bull Seismol Soc Am 101(5):2122–2141

Okada H (2003) The microtremor survey method. SEG, Tulsa

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. J Appl Geophys. https://doi.org/10.1016/j.jappgeo.2018.12.018

Park CB, Miller RD, Xia J (1999) Multi-channel analysis of surface waves (MASW). Geophysics 64:800–808

Park CB, Miller RD, Xia J, Ivanov J (2004) Imaging dispersion curves of passive surface waves. In: 74th annual international meeting, SEG, expanded abstracts, pp 1357–1360

Paul A, Campillo M, Margerin L, Larose E (2005) Empirical synthesis of time-asymmetrical Green functions from the correlation of coda waves. J Geophys Res 110(B8):1–13

Picozzi M, Parolai S, Bindi D, Strollo A (2009) Characterization of shallow geology by high-frequency seismic noise tomography. Geophys J Int 176:164–174

Sabra KG, Gerstoft P, Roux P, Kuperman WA, Fehler MC (2005) Surface wave tomography from microseisms in southern California. Geophys Res Lett 32:L14311

Saito H, Hayashi K, Iikura Y (2004) Detection of formation boundaries and permeable fractures based on frequency-domain Stoneley wave logs. Explor Geophys 35:45–50

Schuster GT, Yu J, Sheng J, Rickett J (2004) Interferometric/daylight seismic imaging. Geophys J Int 157:838–852

Shapiro NM, Campillo M (2004) Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. Geophys Res Lett 31(7):L07614

Snieder R (2004) Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase. Phys Rev E 69:046610

Snieder R (2006) The theory of coda wave interferometry. Pure Appl Geophys 163(2–3):455–473

Snieder R, Miyazawa M, Slob E, Vasconcelos I, Wapenaar K (2009) A comparison of strategies for seismic interferometry. Surv Geophys 30(4–5):503–523

Socco LV, Foti S, Boiero D (2010) Surface-wave analysis for building near-surface velocity models-Established approaches and new perspectives. Geophysics 75(5):75A83–75A102

Song YY, Castagna JP, Black RA, Knapp RW (1989) Sensitivity of near-surface shear-wave velocity determination from Rayleigh and Love waves. In: 59th annual international meeting, SEG, expanded abstracts, pp 509–512

Stehly L, Campillo M, Shapiro NM (2006) A study of the seismic noise from its long-range correlation properties. J Geophys Res Solid Earth 111:1–12

Stokoe KH II, Nazarian S (1983) Effectiveness of ground improvement from spectral analysis of surface waves. In: Proceeding of the eighth European conference on soil mechanics and foundation engineering, vol. 1, pp 91–95

Van Dalen KN, Mikesell TD, Ruigrok EN, Wapenaar K (2015) Retrieving surface waves from ambient seismic noise using seismic interferometry by multidimensional deconvolution. J Geophys Res 120:944–961

Vasconcelos I, Snieder R (2008a) Interferometry by deconvolution, part 1-theory for acoustic waves and numerical examples. Geophysics 73:S115–S128

Vasconcelos I, Snieder R (2008b) Interferometry by deconvolution: part 2-theory for elastic waves and application to drill-bit seismic imaging. Geophysics 73:S129–S141

Wang H, Quan W, Wang Y, Miller GR (2014) Dual roadside seismic sensor for moving road vehicle detection and characterization. Sensors (Switzerland) 14(2):2892–2910

Wapenaar K, Fokkema J (2006) Green's function representations for seismic interferometry. Geophysics 71(4):SI33–SI46

Wapenaar K, Ruigrok E (2011) Improved surface-wave retrieval from ambient seismic noise by multidimensional deconvolution. Geophys Res Lett 38:1–5

Wapenaar K, van der Neut J, Ruigrok E (2008) Passive seismic interferometry by multidimensional deconvolution. Geophysics 73(6):A51–A56

Wapenaar K, van der Neut J, Ruigrok E, Draganov D, Hunziker J, Slob E et al (2011) Seismic interferometry by crosscorrelation and by multidimensional deconvolution: a systematic comparison. Geophys J Int 185(3):1335–1364. https://doi.org/10.1111/j.1365-246X.2011.05007.x

Weemstra C, Draganov D, Ruigrok EN, Hunziker J, Gomez M, Wapenaar K (2017) Application of seismic interferometry by multidimensional deconvolution to ambient seismic noise recorded in Malargüe, Argentina. Geophys J Int 60(4):802–823

Xia J, Miller RD, Park CB (1999) Estimation of near-surface shear-wave velocity by inversion of Rayleigh wave. Geophysics 64:691–700

Xia J, Miller RD, Park CB, Survey KG (2000) Advantages of calculating shear-wave velocity from surface waves with higher modes. In: SEG expanded abstracts, pp 1285–1298

Xia J, Miller RD, Park CB, Tian G (2003) Inversion of high frequency surface waves with fundamental and higher modes. J Appl Geophys 52:45–57

Xia J, Miller RD, Xu Y, Luo Y, Chen C, Liu J, Ivanov J, Zeng C (2009) High-frequency Rayleigh-wave method. J Earth Sci 20:563–579

Xia J, Cheng F, Xu Z, Shen C, Liu R (2017) Advantages of multi-channel analysis of passive surface waves (MAPS). In: International conference on engineering geophysics, Al Ain, United Arab Emirates, 9–12 October 2017, pp 94–97

Yao H, van der Hilst RD, de Hoop MV (2006) Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis-I. Phase velocity maps. Geophys J Int 166:732–744

Zeng X, Lancelle C, Thurber C, Fratta D, Wang H, Lord N, Chalari A, Clarke A (2017) Properties of noise cross-correlation functions obtained from a distributed acoustic sensing array at Garner Valley, California. Bull Seismol Soc Am 107(2):603–610

Zhong T, Li Y, Wu N, Nie P, Yang B (2015) Statistical properties of the random noise in seismic data. J Appl Geophys 118:84–91

Zhou C, Xi C, Pang J, Liu Y (2018) Ambient noise data selection based on the asymmetry of cross-correlation functions for near surface applications. J Appl Geophys 159:803–813

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.