Geophys. J. Int. (–) –, –

Optimal Stacking of Noise Cross-Correlation Functions

Xiaotao Yang[1,2⋆], Jared Bryan[3], Kurama Okubo[1,4], Chengxin Jiang[1,5], Timothy Clements[1,6], and Marine A. Denolle[1,7]

[1] *Department of Earth and Planetary Sciences, Harvard University. Cambridge, MA, USA*
[2] *Now at Department of Earth, Atmospheric, and Planetary Sciences, Purdue University. West Lafayette, IN, USA*
[3] *Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology. Cambridge, MA USA*
[4] *Now at National Research Institute for Earth Science and Disaster Resilience, Tsukuba, Japan*
[5] *Now at Research School of Earth Sciences, Australian National University. Acton, ACT Australia*
[6] *Now at Earthquake Science Center, United States Geological Survey. Moffet Field, CA USA*
[7] *Now at Department of Earth and Space Sciences, University of Washington, Seattle, WA USA*

Received –; in original form –

SUMMARY

Cross-correlations of ambient seismic noise are widely used for seismic velocity imaging, monitoring, and ground motion analyses. A typical step in analyzing Noise Cross-correlation Functions (NCFs) is stacking short-term NCFs over longer time periods to increase the signal quality. Spurious NCFs could contaminate the stack, degrade its quality, and limit its use. Many methods have been developed to improve the stacking of coherent waveforms, including earthquake waveforms, receiver functions, and NCFs. This study systematically evaluates and compares the performance of eight stacking methods, including arithmetic mean or linear stacking, robust stacking, selective stacking, cluster stacking, phase-weighted stacking, time-frequency phase-weighted stacking, N^{th}-root stacking, and averaging after applying an adaptive covariance filter. Our results demonstrate that, in most cases, all methods can retrieve clear ballistic or first arrivals. However, they yield significant differences in preserving the phase and amplitude information. This study provides a practical guide for choosing the optimal stacking method for specific research applications in ambient noise seismology. We evaluate the performance using multiple onshore and offshore seismic arrays in the Pacific Northwest region. We compare these stacking methods for NCFs calculated from raw ambient noise (referred to as Raw NCFs) and from ambient noise normalized using a one-bit clipping time normalization method (referred to as One-bit NCFs). We evaluate six metrics, including signal-to-noise ratios, phase dispersion images, convergence rate, temporal changes in the ballistic and coda waves, relative amplitude decays with distance, and computational time. We show that robust stacking

is the best choice for all applications (velocity tomography, monitoring, and attenuation studies) using Raw NCFs. For applications using One-bit NCFs, all methods but phase-weighted, time-frequency phase-weighted, and N^{th}-root stacking are good choices for seismic velocity tomography. Linear, robust, and selective stacking methods are all equally appropriate choices when using One-bit NCFs for monitoring applications. For applications relying on accurate relative amplitudes, both the robust and cluster stacking methods perform well with One-bit NCFs. The evaluations in this study can be generalized to a broad range of time-series analysis that utilizes data coherence to perform ensemble stacking. Another contribution of this study is the accompanying open-source software, which can be used for general purposes in time-series stacking.

Key words: Seismic interferometry, Seismic tomography, Earthquake ground motions, Seismic noise, Time-series analysis, Coda waves

⋆ Corresponding author: xtyang@purdue.edu (X. Yang)

1 INTRODUCTION

Cross-correlations of ambient seismic noise have been widely used to image the Earth's elastic (L. Feng and Ritzwoller 2019; Shapiro et al. 2005; S.-M. Wu et al. 2021; X. Yang and Gao 2018, 2020) and anelastic structure (Prieto et al. 2009), model ground motions (Marine A Denolle et al. 2013, 2014, 2018; Kwak et al. 2017; Viens and Marine A Denolle 2019; Viens et al. 2017), and monitor transient velocity changes in the shallow subsurface (F. Brenguier et al. 2008; Clements and Marine A Denolle 2018; Donaldson et al. 2019; K.-F. Feng et al. 2021; Olivier et al. 2019; Q.-Y. Wang et al. 2017; Z. Yang et al. 2022). The Noise Cross-correlation Functions (NCFs) are typically computed in short-time windows (such as hours or days) after diverse pre-processing on the raw ambient noise waveforms (Bensen et al. 2007; L. Feng and Ritzwoller 2019). Stacking of NCFs over a longer period (such as weeks, months, or years) is a common procedure for most applications utilizing NCFs (e.g., Seats et al. 2012). For seismic tomography, the final stack over the entire time period is used to make waveform or travel-time measurements. Seismic monitoring requires stacking over a subset of time windows, balancing the temporal resolution and the coherence of the NCFs (Hadziioannou et al. 2011). These applications rely on the phase information of ballistic (first arrival) and/or coda waves. Studies of ground motion prediction and attenuation tomography, on the other hand, require accurate relative amplitude measurements between station pairs. Because NCFs consist primarily of dispersive surface waves, preserving the relative amplitude also preserves the spectral content. Therefore, the performance of stacking and the preservation of either phase or amplitude information are important to ensure the robustness of the scientific results.

The most commonly used stacking method is the arithmetic mean of the data over temporal or spatial samples, referred to as Linear stacking. For NCFs, substantial temporal variation may exist across individual short time windows, which could result from the seasonal change of the location and strength of the microseismic sources (Bensen et al. 2007; L. Ermert et al. 2016; L. A. Ermert et al. 2021; Tian and Ritzwoller 2017; X. Yang et al. 2019), disruption from tectonic signals (Baig et al. 2009), or unknown instrumental failures. It is important to use an appropriate stacking method to extract coherent NCFs, particularly when strong variations of the data quality are present. Many more advanced methods have been developed to improve the stacking of NCFs, and time-series data in general and overcome the limitations of Linear stacking. Rückemann (2012) provides a summary of the theoretical background of several stacking methods to improve the common-mid-point gathers in seismic reflection data. A systematical evaluation of the performance of different stacking methods in processing real NCF data is highly desirable and required for optimal processing strategies.

In this paper, we compare eight algorithms used for stacking NCFs and evaluate their performance for canonical research applications. These methods include: robust stacking (Pavlis and Vernon 2010), selective stacking (modified from Liu et al. 2009; Olivier et al. 2015), cluster stacking (a new method), Phase-Weighted Stacking (PWS; Schimmel and Paulssen 1997), time-frequency Phase-Weighted Stacking (tf-PWS; Baig et al. 2009; Li et al. 2017; Schimmel and Gallart 2007; Schimmel et al. 2011; Thurber et al. 2014; Zeng and Thurber 2016), N^{th}-root stacking (Millet et al. 2019; Rückemann 2012), and linear stacking after applying an Adaptive Covariance Filter (ACF; Nakata et al. 2015). We exclude the stacking methods that use

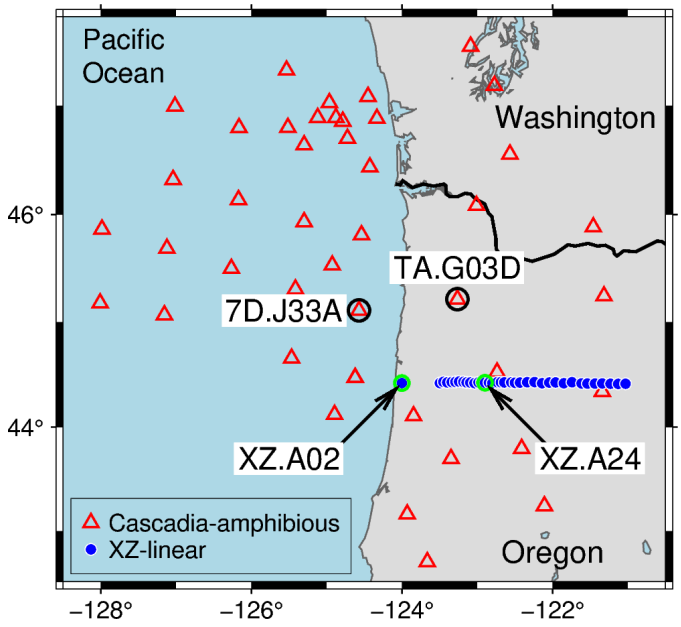

Figure 1. Seismic stations used in this study. The USArray Transportable Array and Cascadia Initiative Ocean Bottom Seismographs (red open triangles) form a composite dataset that is referred to as the "Amphibious" dataset. The 1993-1994 Cascadia broadband XZ linear array (blue dots) is referred to as the "XZ" dataset. Labeled stations are used as virtual sources and receivers in the examples in this paper.

the curvelet transform (Laurent Stehly et al. 2011) and singular value decomposition (Moreau et al. 2017), as these focus specifically on short-term convergence alone. We assess the performance of the eight algorithms using six different metrics relevant to specific research applications. Because the one-bit pre-processing is designed to reduce the influence of earthquake-like transient signals and to improve the quality of individual NCFs (Bensen et al. 2007), we benchmark our tests using both the raw and one-bit normalization pre-processing techniques. The development of new methods and computer codes and the performance tests from this study aim to guide the community in choosing the appropriate stacking method for ambient-noise seismology use cases, though they can also be generalized to process other time-series data.

2 AMBIENT NOISE CROSS-CORRELATIONS

This study uses the Pacific Northwest as a natural laboratory (Fig. 1). We gather data from 29 Cascadia Initiative Ocean Bottom Seismographs (OBS) (network code: 7D; `https://doi.org/10.7914/SN/7D_2011`), 15 US Transportable Array stations (network code: TA; `https://doi.org/10.7914/SN/TA`), and 43 stations from the 1993-1994 Cascadia subduction zone experiment (network code: XZ; Nabelek et al. 1993). We group the data into two datasets: 1) the composite dataset from the Cascadia Initiative OBSs and onshore TA stations is referred to as the "Amphibious" dataset; 2) the data from the East-West XZ linear array in northern Oregon is referred to as the "XZ" dataset. All data are publicly accessible from the seismic data archive of the Incorporate Research Institutions for Seismology Data Management Center (IRIS DMC).

We use the *SeisGo* Python toolbox to download the continuous seismic waveforms and compute the NCFs (X. Yang et al. 2022a). The cross-correlation function in *SeisGo* was modified from *NoisePy* (Jiang and Marine A Denolle 2020). We download

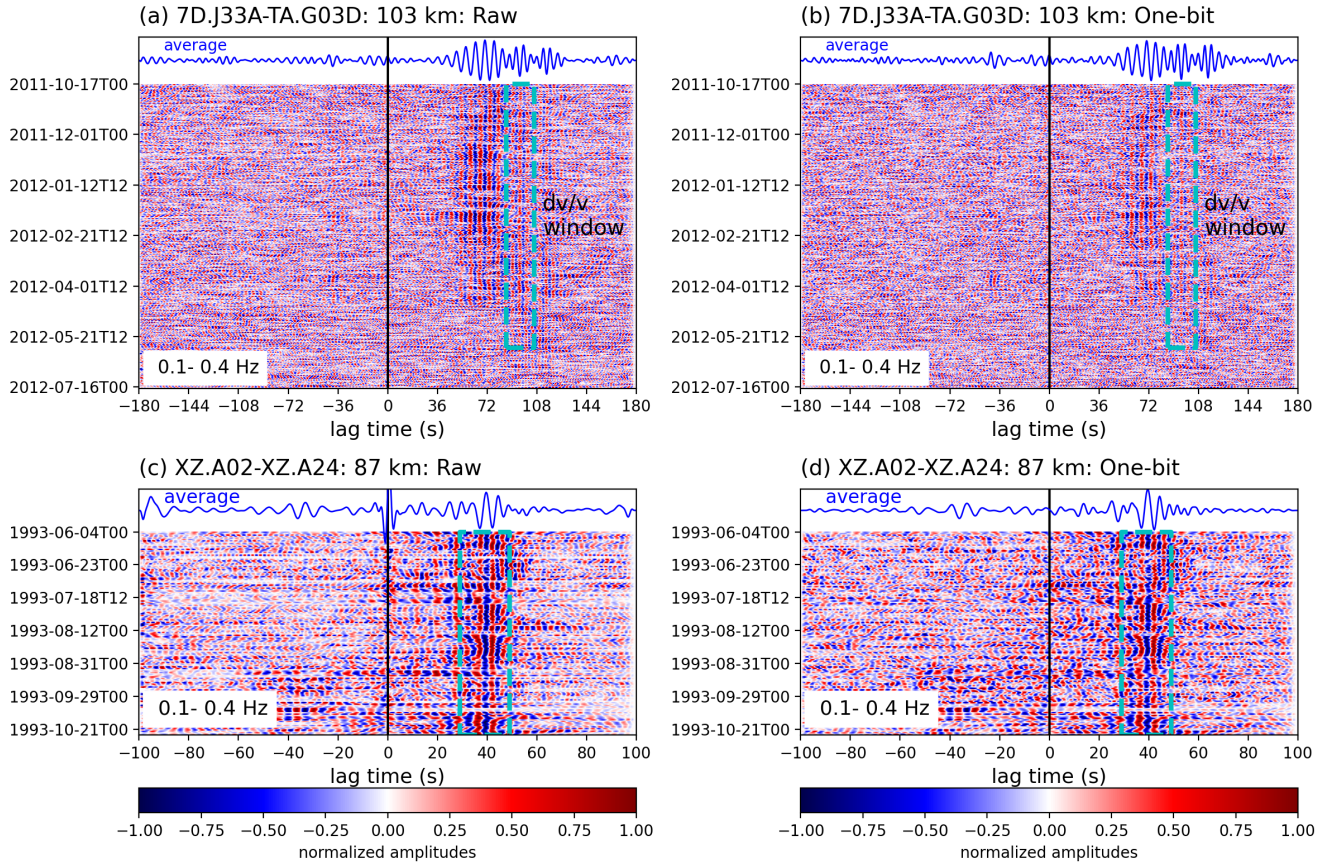

Figure 2. Examples of ambient noise cross-correlation functions (NCFs) between (a-b) OBS station 7D.J33A and onshore station TA.G03D and (c-d) XZ linear array stations XZ.A02 and XZ.A24. The NCFs are filtered at 0.1-0.4 Hz. The left panels are NCFs computed with raw waveforms and the right panels are NCFs computed after applying one-bit time-domain normalization. The dashed cyan boxes mark the windows used for the analysis of the preservation of transient phase changes (Section 4.4) using (a-b) coda waves and (c-d) ballistic phases. On the top of each panel, we show the linear stack as arithmetic averages. The NCFs are color-coded by the normalized amplitudes (red is 1, blue is -1).

the vertical component of the continuous waveforms in 12-hour segments. We select data from 10/1/2011 to 7/31/2012 for the Amphibious dataset and from 6/1/1993 to 5/31/1994 for the XZ dataset. We remove the instrument responses and convert the waveforms to displacements, followed by down-sampling to 5 Hz. We compute the NCFs in 6-hour windows sliding every 3 hours. The data is demeaned and detrended for each 6-hour segment. We attempt to remove transient signals in the ambient noise that will contaminate the noise cross correlations by removing the windows with anomalous amplitudes. For that, we calculate the maximum absolute raw noise amplitude of the segment and the standard deviation of the raw noise amplitudes of all segments. We discard waveform segments with peak amplitudes greater than 10 times the standard deviation, as implemented in *NoisePy*.

We construct two sets of cross-correlations using the raw waveforms (Viens et al. 2017) and waveforms normalized in the time domain using the `sign` function (e.g., Bensen et al. 2007; Shen et al. 2012). We denote the two datasets as Raw NCFs and One-bit NCFs, respectively. To avoid spectral leakage, the short 6-hour time windows are tapered using a Tukey window with a cosine fraction of 0.05. We then apply a Fourier transform to the short-window segment and compute the cross-correlation in

the frequency domain using the following equation:

$$X = \mathcal{F}^{-1}\left(\mathcal{F}^{*}(d_s)\mathcal{F}(d_r)\right), \tag{1}$$

where X is the NCF between time series d_s (virtual source) and d_r (receiver), $\mathcal{F}$ is the Fourier transform operator, $\mathcal{F}^{*}$ stands for the complex conjugate of the Fourier transform, and $\mathcal{F}^{-1}$ is the inverse Fourier transform operator.

We show a few examples of the computed NCFs (bandpass filtered at 0.1-0.4 Hz) in Fig. 2 for both pre-processing strategies. The asymmetry of the NCF is a known effect of noise source directionality (Stehly et al. 2006). The first-order observation is that the individual NCFs contain incoherent noise that may overwhelm the coherent signals. The NCFs from the XZ land station pair all show clear ballistic phases over the entire period range of this analysis (Fig. 2c-d). For the amphibious station pair 7D.J33A-TA.G03D, clear ballistic phases are present between 50s and 100s for most of the NCFs between October 2011 and April 2012, for both Raw (Fig. 2a) and One-bit (Fig. 2b) NCFs. However, the ballistic phases are almost invisible for the NCFs after May 2012. These "noisy" NCFs may contaminate the linear stack of the NCFs. The simple time-domain representation in Fig. 2 only provides a qualitative visual assessment. We later quantify the performance using different stacking methods with multiple metrics.

3 STACKING METHODS

This section describes the algorithms beyond the arithmetic mean (or Linear stacking) for optimally stacking NCFs. Most of these methods are adapted from the published literature. All methods can be generalized as weighted stacking with various ways of computing the stacking weight, which could be either trace weight or sample weight. We express the weighted stacking scheme as:

$$\mathbf{b} = \sum_{i=1}^{N} w_i \mathbf{d}_i, \tag{2}$$

where $\mathbf{b}$ is the final stack. For trace-weighted stacking, $\mathbf{d}_i$ is the i^{th} NCF trace, w_i is the stacking weight for the i^{th} NCF trace satisfying $\sum w_i = 1$, and N is the total number of NCF traces. For sample-weighted stacking, $\mathbf{d}_i$ is the i^{th} sample of the arithmetic average or linear stack of all NCF traces, w_i is the stacking weight for the i^{th} sample satisfying $\sum w_i = 1$, and N is the total number of samples in each NCF trace. Under the stacking formulation of Equation 2, Linear stacking is a special case of trace-weighted stacking where all traces have the same weight ($w_i = 1$). In the methods described in this section, the Robust (Section 3.1), Selective (Section 3.2), and Cluster (Section 3.3) stacking methods are all trace-weighted stacking methods. The phase-weighted (Section 3.4), time-frequency phase-weighted (Section 3.5), N^{th}-root (Section 3.6), and adaptive covariance filter (Section 3.7) stacking methods are all sample-weighted stacking methods. In the following subsections, we describe the computation of the stacking weights for each method in detail.

3.1 Robust stacking

Pavlis and Vernon (2010) first introduced Robust stacking to improve the accuracy of picking phase arrival times in seismograms from teleseismic earthquakes. We adapt this generic, iterative, weighted stacking method to process NCFs. The weight in Robust stacking is based on waveform de-coherence and penalizes the NCF that deviates too much from a reference stack. The weight w_i of the i^{th} NCF, $\mathbf{d}_i$, for the j^{th} iteration is calculated as (Pavlis and Vernon 2010):

$$w_i^j = \frac{|\mathbf{b}_j \cdot \mathbf{d}_i|}{||\mathbf{d}_i||||\mathbf{r}_i||}, \tag{3}$$

where $\mathbf{b}_j$ is the reference stack of the j^{th} iteration, $||*||$ is the L2-norm operator, and $\mathbf{r}_i = \mathbf{b}_j - (\mathbf{b}_j \cdot \mathbf{d}_i)\mathbf{d}_i$ is a modified residual quantifying the difference between the current reference stack $\mathbf{b}_j$ and the i^{th} NCF scaled by the dot product of the two time series. The weight w_i penalizes twice the data that does not resemble the reference stack. The first element is the dot product between the NCF and the stack: $\mathbf{b}_j \cdot \mathbf{d}_i|/||\mathbf{d}_i||$. The second is the modified residual r_i: if it is high, the $1/||r_i||$ is small, and w_i is low. The weights can be calculated over either the entire NCF or a specific time window that contains the seismic phase of interest. For the examples in this paper, we compute the weights over the entire NCF, though our software package allows the user to specify the time window. After computing the weights for all NCFs, we normalize them to their sum and use them to produce the updated stack.

The initial reference stack $\mathbf{b}_1$ is chosen as the median NCF, following the observation of Pavlis and Vernon (2010) that the median is more representative of the concentration of data in the presence of outliers. In our implementation, the users may specify another trace as the initial stack, such as the Linear stack. The iterative procedure stops when a convergence criterion is satisfied. We adopt the criterion proposed by Pavlis and Vernon (2010):

$$\frac{|\mathbf{b}_j - \mathbf{b}_{j-1}|}{||\mathbf{d}_i||M} < \epsilon, \tag{4}$$

where M is the number of samples in each NCF time-series and ϵ is a small number, with a default value of 10^{-5}.

3.2 Selective stacking

In the presence of strong noise, it could be effective to stack only a subset of the NCF ensemble that exceeds a quality threshold and to ignore the low-quality NCFs. This concept, called selective stacking, has been implemented in previous studies (e.g., Olivier et al. 2015; Thangraj and Pulliam 2021). The signal-to-noise ratio (SNR) or the correlation to the reference can be used to evaluate the quality of each NCF. Olivier et al. (2015) used the SNR around the expected S-wave arrival time to reconstruct S-waves in the stacked NCF. Liu et al. (2009) proposed the weighted stack of common-midpoint gather using the local correlation within the moving window to improve the SNR and suppress the random noise in the stacked trace. To enhance the coherence of the signal, we use a criterion based on the similarity, quantified as the Pearson correlation coefficient in our implementation, between each NCF and the reference NCF. In our implementation, by default, the reference NCF is

initialized as the Linear stack of the entire ensemble (similar to Liu et al. 2009). As in Robust stacking, the user can specify any trace, such as the median NCF, as the reference stack and the window used to estimate the similarity. Selective stacking sets the weights of low-similarity NCFs to zero, computed as:

$$w_i = \begin{cases} 0, & \rho_{\mathbf{b},\mathbf{d}_i} \leq \rho_t \\ 1, & \rho_{\mathbf{b},\mathbf{d}_i} > \rho_t \end{cases} . \tag{5}$$

where $\rho_{\mathbf{b},\mathbf{d}_i}$ is the Pearson correlation coefficient between the individual trace and the reference NCF and ρ_t is the user-defined threshold. Note that the criterion used in this study is equivalent to the global correlation described in Liu et al. (2009) rather than the local correlation for the sake of simplicity in thresholding. A similarity threshold allows us to discriminate the set of NCFs to keep or discard (weight is set to zero). We then compute the Linear stack over the subset of NCFs with a Pearson correlation coefficient above the threshold and discard the NCFs below the threshold. Similar to robust stacking, we implement the selective stacking method in an iterative scheme, with the same convergence criterion as in Equation 4. In our examples, we use a correlation coefficient threshold of 0.

3.3 Cluster stacking

An alternative method to ranking waveforms according to a specific similarity metric is to group them into clusters and perform the stack for each cluster. Viens and Iwata (2020) proposed the idea of clustering the NCFs using a dimensionality reduction technique (principal component analysis), as exploited by Toghramadjian et al. (2021). In our implementation, we use the *Tslearn* Python toolkit (Tavenard et al. 2020) for k-means clustering of the NCF waveforms based on the Euclidean distance. We impose two clusters to separate NCFs with higher quality from those with lower quality.

After clustering the NCFs into two clusters, the final stack **b** is computed as a weighted stack of the two cluster centers

$$\mathbf{b} = \sum_{i=1,2} w_i C_i, \tag{6}$$

where C_i (i=1, 2) are the centers (sample averages or linear stacks) of the two clusters and w_i (i=1, 2) are the weights of the cluster centers. The center stacking weights, w_i (i=1, 2), are computed as

$$w_{i=1,2} = \begin{cases} \begin{cases} 0, & p_i < max(p_1, p_2) \\ 1, & p_i >= max(p_1, p_2) \end{cases} , & cc < h \\ p_i/(p_1 + p_2), & cc \geq h \end{cases} , \tag{7}$$

where cc is the correlation coefficient between the two cluster centers C_1 and C_2, h is the similarity threshold specified by the user, and p_i (i=1, 2) is the relative peak amplitude of the cluster center, computed as the ratio of the maximum and root-mean-square of the absolute amplitudes. The relative peak amplitude we use here is an equivalent measure to the signal-to-noise ratio of the cluster center. In our examples, p_1 and p_2 are computed over the entire NCF. Our implementation also allows the user to specify the window of interest to compute the maximum of the absolute amplitudes. According to Equation 7, the two centers

are weighted by the relative peak amplitude when their correlation coefficient exceeds the threshold (h=0.75 for our examples). Otherwise, we use the center with a higher peak amplitude as the final stack.

3.4 Phase-weighted stacking

Phase-weighted stacking (PWS) computes the weight for each trace by the coherency of the instantaneous phases (Schimmel and Paulssen 1997). Our implementation of the PWS method follows the description in Schimmel and Paulssen (1997). Readers are referred to Schimmel and Paulssen (1997) for detailed equations and procedures of the PWS method. The final stack is constructed as a non-linear weighted stack where each sample is weighted by the instantaneous phase. The key parameter for PWS is a harshness parameter that quantifies the fall-off of the weight of each NCF with decreasing similarity. We set the harshness to 2, as in Schimmel and Paulssen (1997).

3.5 Time-frequency phase-weighted stacking

Schimmel and Gallart (2007) and Schimmel et al. (2011) proposed an improved phase-weighted stacking method that computes the weight in both the time and frequency domains. This method, referred to as tf-PWS, projects each seismic trace into the time-frequency domain through the Stockwell transform (S-transform; Stockwell et al. 1996). Baig et al. (2009) adapted the stacking method based on a more efficient discrete orthogonal S-transform that is introduced by Stockwell (2007). The implementation of tf-PWS in this study follows the description by Schimmel et al. (2011) using the original S-transform (Equations 4-7 therein). Readers are referred to Schimmel and Gallart (2007) and Schimmel et al. (2011) for detailed formulation of the tf-PWS stacking method. Similar to PWS, the tf-PWS is primarily controlled by a harshness parameter that determines the sharpness of the transition between phase similarity and dissimilarity. We set the harshness to 2 for our examples, as in Schimmel et al. (2011).

3.6 N^{th}-root stacking

N^{th}-root stacking method is commonly employed in array seismology (Kanasewich et al. 1973; Millet et al. 2019; Muirhead 1968; Rost and Thomas 2002; Rückemann 2012; Schimmel and Paulssen 1997). It is a nonlinear stacking method that takes the N^{th}-root of the absolute amplitudes of each individual trace and sums them together. The summation is then raised to the power of N to assign the sample weight. The polarity of each sample is recovered with a sign function of the summation. Our implementation follows the description in Millet et al. (2019) (their equation 18) and can be summarized as:

$$\mathbf{b} = sign(\mathbf{r})|\mathbf{r}|^n, \tag{8}$$

where $\mathbf{b}$ is the final stack, n is the specified order of root, $\mathbf{r} = \frac{1}{N}\sum_{i=1}^{N} sign(\mathbf{d}_i)|\mathbf{d}_i|^{1/n}$, $\mathbf{d}_i$ is the i^{th} NCF trace, and N is the total number of NCF traces. We use the square root (i.e., n=2) in our examples.

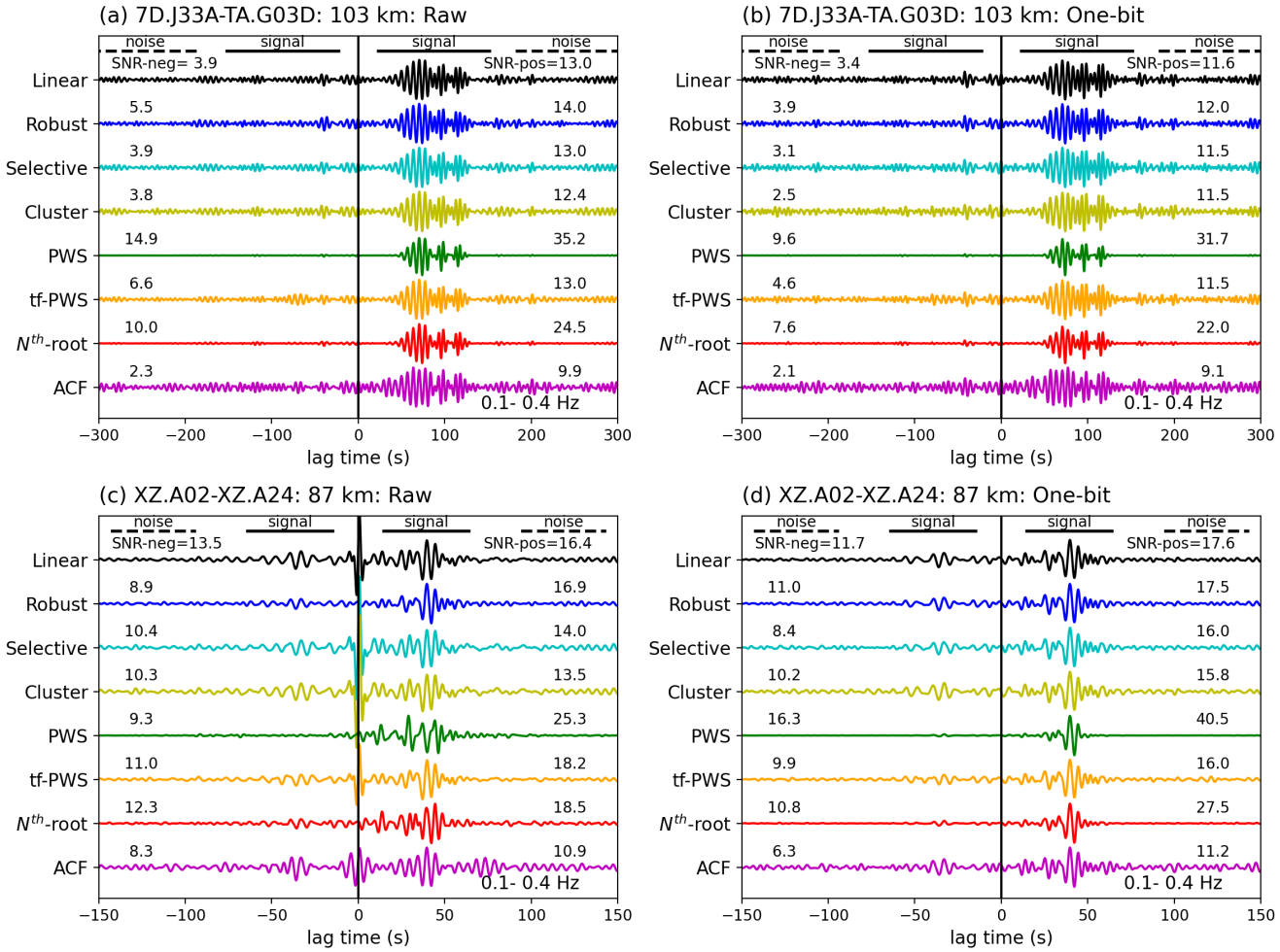

Figure 3. Examples of the stacked NCFs between station pairs (a-b) 7D.J33A-TA.G03D and (c-d) XZ.A02-XZ.A24, using different methods. The left panels are the stacking results for Raw NCFs and the right panels are the stacking results for One-bit NCFs. All NCFs are filtered at 0.1-0.4 Hz after stacking. The stacking algorithm is labeled on the left of each panel. The SNR in decibel (Equation 10) is indicated above each trace on each side of the correlations. The thick solid and dashed lines mark the starts and ends of the signal and noise windows, respectively.

3.7 Adaptive covariance filter stacking

Nakata et al. (2015) introduced an adaptive covariance filter (ACF) to suppress incoherent noise in seismic data based on the adaptive polarization filter (Du et al. 2000; Samson and Olson 1981). We implement the ACF stacking method following Nakata et al. (2015), which the readers are referred to for detailed formulations of the method. The final stack is the Linear stack after applying the ACF. We use 1 as the harshness of the filter.

4 EVALUATION AND COMPARISON OF STACKING ALGORITHMS

We implement the stacking algorithms described in Section 3 as a standalone Python package *StackMaster* (X. Yang et al. 2022b). The package is available as a repository on Python Package Index (PyPI). It can be installed with: `pip install stackmaster`. In this paper, we install *StackMaster* with the *SeisGo* toolbox (X. Yang et al. 2022a) under

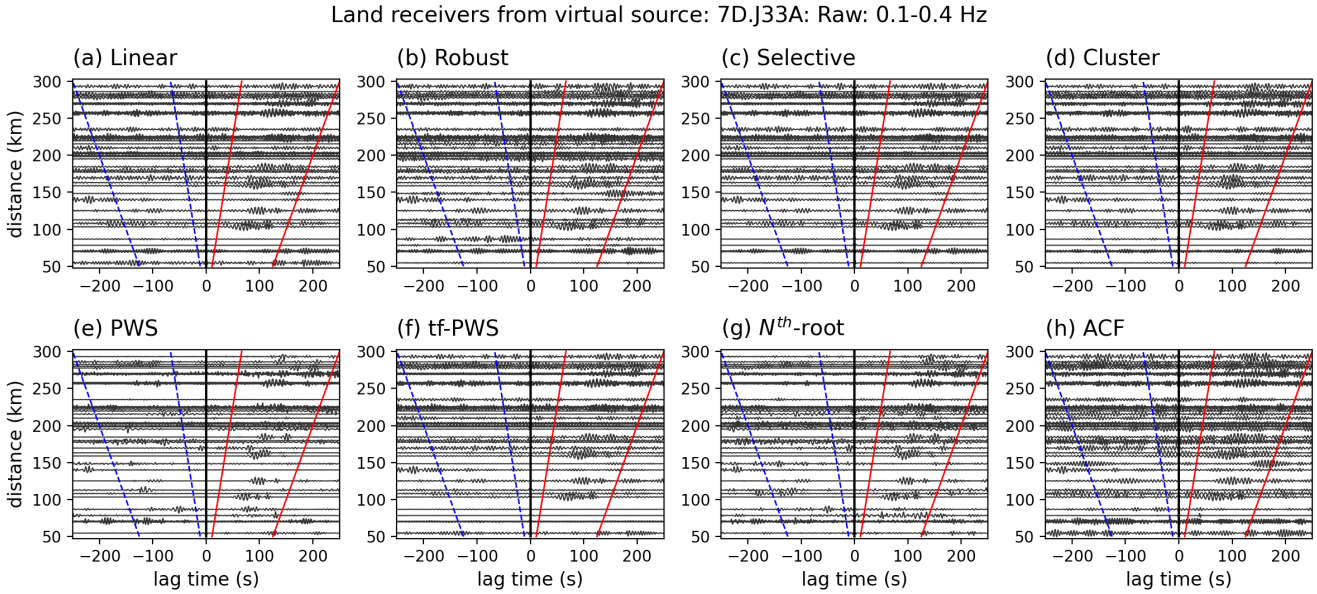

Figure 4. Stacked Raw NCFs of the Cascadia amphibious array from 7D.J33A to other land receivers using different stacking methods, filtered at 0.1-0.4 Hz. (a-h) The results using Linear, Robust, Selective, Cluster, PWS, tf-PWS, N^{th}-root, and ACF stacking methods, respectively. The red solid lines and the blue dashed lines outline the positive-lag signal window and the negative-lag signal window, respectively, used to compute the SNR in Fig. 7. The signal and noise windows are determined with the same method as in Fig. 3a-b. See Fig. 1 for station locations.

the same Anaconda environment. We apply the stacking techniques to the Amphibious and the XZ datasets. Fig. 3 shows the stacking results for the two station pairs in Fig. 2. Figs 4-6 and S1-S3 in the supplement are the stacking results of all station pairs from the virtual sources at 7D.J33A (Figs 4-5 and S1-S2) and at XZ.A02 (Figs 6 and S3).

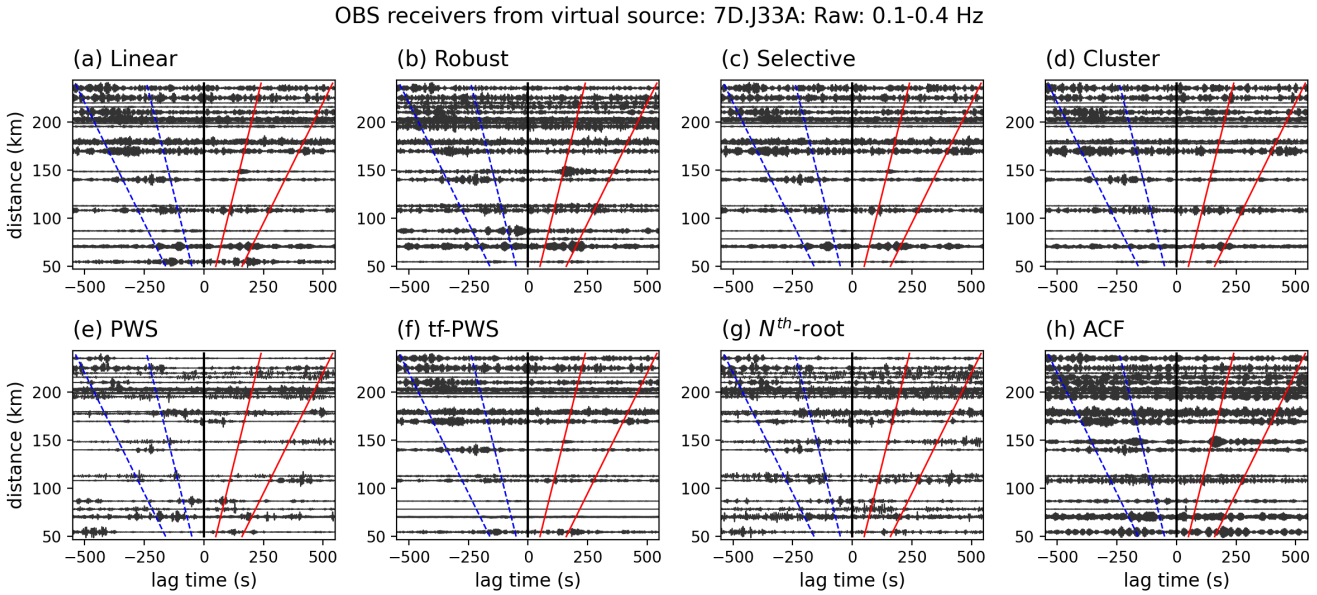

Figure 5. Same as Fig. 4 but for NCFs from 7D.J33A to other OBS receivers. To contain the visually identified ballistic phases from these OBS station pairs, we use a different velocity range (0.5-1.0 km/s) here to predict the signal window of the weakly coherent signals. We extend the window for an additional 60 s after the latest predicted arrival. See Fig. 1 for locations of the OBS receivers.

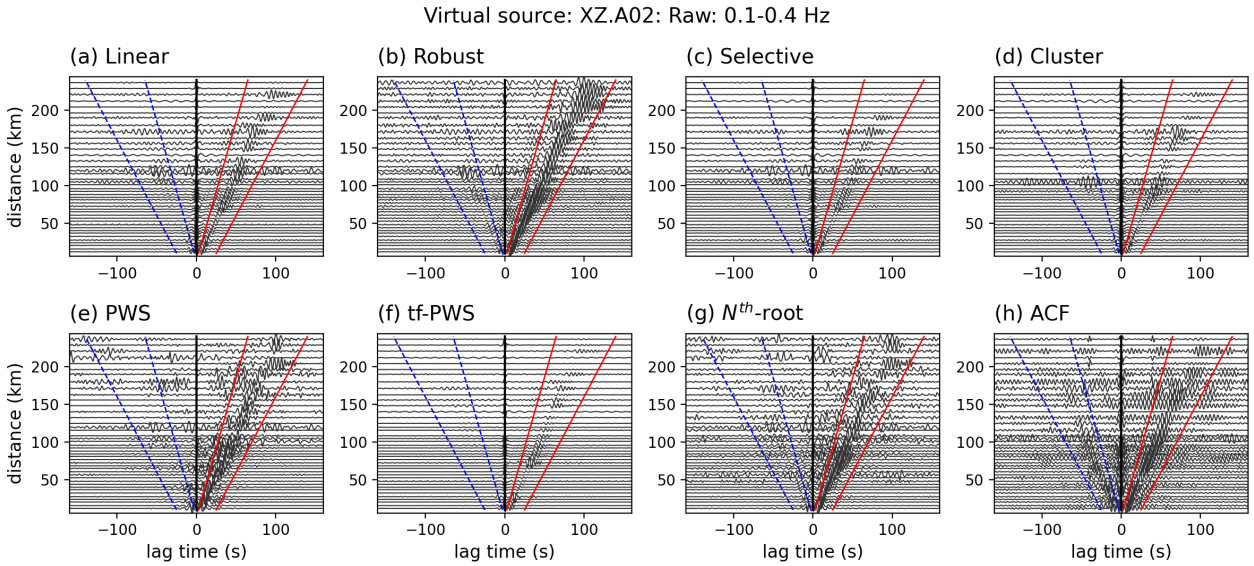

Figure 6. Stacked Raw NCFs between the XZ linear array stations from XZ.A02 to other receivers, filtered at 0.1-0.4 Hz. (a-h) The results using Linear, Robust, Selective, Cluster, PWS, tf-PWS, N^{th}-root, and ACF stacking methods, respectively. The red solid lines and the blue dashed lines outline the positive signal window and the negative signal window, respectively, used to compute the SNR in Fig. 7. The signal and noise windows are determined with the same method as in Fig. 3c-d. See Fig. 1 for station locations.

We then evaluate and compare the performance of these stacking methods. Recognizing that NCFs are used in multiple applications (e.g., seismic velocity tomography, monitoring, and attenuation/ground motion analysis), we evaluate the performance of the algorithm over a range of metrics: 1) signal-to-noise ratios, 2) surface wave dispersion, 3) convergence of short-term stacks to the long-term stack, 4) transient phase changes, 5) peak amplitudes of ballistic phases, and 6) computational expense. We compare the methods below according to each metric. It is worth noting that there are multiple parameters that can be tuned for each stacking method, though we only focus on the most commonly used or suggested parameters in this study. In Section 5, we discuss the choice of stacking methods for different categories of applications using NCFs.

4.1 Signal-to-noise ratios

We use the SNR of the ballistic phase as a proxy for the quality of the NCF stack. In this study, we define the SNR, R, as

$$R = \left[\frac{rms(|A_{signal}|)}{rms(|A_{noise}|)}\right]^2, \tag{9}$$

where $rms()$ is the root-mean-square operator, $|A_{signal}|$ are the absolute amplitudes within the signal window, and $|A_{noise}|$ are the absolute amplitudes within the noise window. For the NCFs from amphibious station pairs (Figs 3a-b, 4, and S1), the signal window is defined by a fast wave traveling at 4.5 km/s, a slow wave traveling at 2 km/s, and we add 100 s to the end of the window. For OBS station pairs (Figs 5 and S2), we use a velocity range of 0.5-1.0 km/s to predict the signal window and add 60 s after the latest predicted arrival time. For the XZ dataset (Figs 3c-d, 6, and S3), the signal window is defined by waves traveling between 2 km/s and 3.7 km/s, extending for an additional 20 s. For both datasets, the noise window has the same

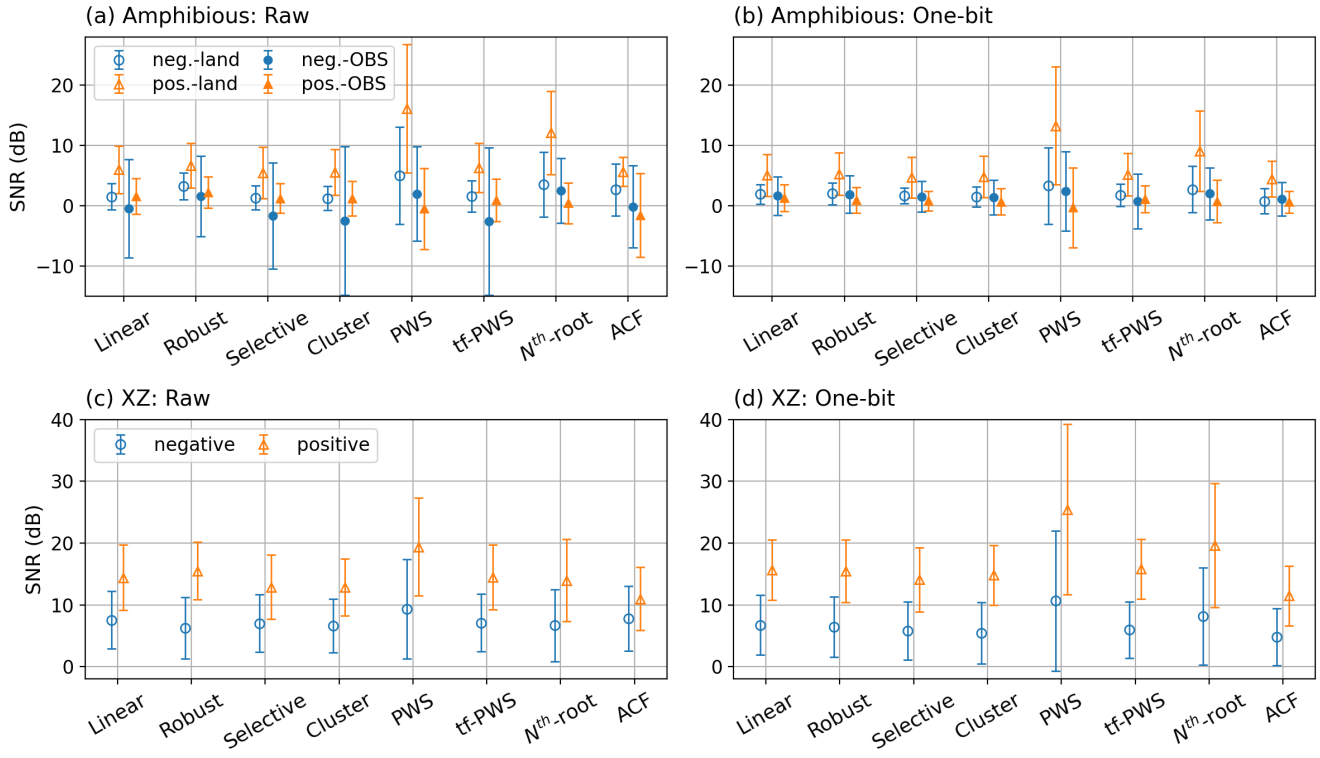

Figure 7. Comparison between the mean signal-to-noise ratios (SNRs) for the NCFs shown in Figs 4-6 and S1-S3. (a-b) Mean SNRs for the Cascadia amphibious array Raw and One-bit NCFs, respectively. The SNRs for the land (open circles and triangles) and OBS (dots and filled triangles) stations are plotted separately. (c-d) Results for the XZ linear array NCFs from Raw and One-bit NCFs, respectively. The error bars show the standard deviations of the SNRs.

length as the signal window with an offset of 30 s after the end of the signal window. Considering the relatively large dynamic range of the computed SNRs, we convert the ratios to decibel scales for all examples in this paper to assist the comparison and visualization, with

$$SNR = 10log_{10}R. \tag{10}$$

It is worth noting that the SNR is unitless and the conversion in Equation (10) is mainly for scaling.

We observe notable variations in the SNRs of the ballistic phases retrieved using different stacking methods. From the single-pair examples (Fig. 3), the ballistic phases are most prominent on the positive lags from all stacking results, with relatively weaker phases on the negative lags, shown as lower SNRs overall. For the Cascadia amphibious station pair, the SNRs range from 9.9 to 35.2 for the Raw NCFs and 9.1-31.7 for the One-bit NCFs (Fig. 3a-b). For the XZ station pair, the SNRs range from 10.9 to 25.3 for the Raw NCFs and 11.2-40.5 for the One-bit NCFs (Fig. 3c-d). For all datasets, the highest SNRs are achieved with the PWS method, while the ACF method produces the lowest SNRs. The stacking results using the N^{th}-root method also show relatively high SNRs. The SNRs of other stacking results are at a comparable level.

The stacking results of all station pairs from the same virtual source provide a more holistic comparison of the performance of different methods (Figs 4-6 and S1-S3). We use the mean SNRs across all station pairs to quantify the comparison of stacks

with different methods (Fig. 7). From the visual inspection of the move-out plots (time-lags vs. inter-station distances), the stacks of Raw (Figs 4 and 6) and One-bit (Figs S1 and S3) NCFs have comparable quality overall for both the Amphibious and the XZ datasets. This is also evident from the comparable mean SNR values of the two pre-processing methods (Fig. 7). For the Cascadia Amphibious dataset, most of the NCFs from the OBS receivers are relatively noisy (Figs 5 and S2), with mean SNRs lower than those from the onshore receivers (Fig. 7a-b). This may result from the contamination of tilt and compliance noise at most OBS stations, a well-known problem in offshore data (Tian and Ritzwoller 2017). Most of the stacked XZ NCFs show clear ballistic phases (Figs 6 and S3), with much higher mean SNRs than those of the Amphibious dataset (Fig. 7). In all examples with land receivers, the ballistic phases are dominantly visible at positive lag times, representing surface waves propagating away from the virtual sources, which we focus on for the following description. For all four datasets (land receivers only for the Amphibious NCFs), the PWS method produces the highest SNRs, while the standard deviation of the SNRs is also the largest. For the Cascadia Amphibious NCFs (Fig. 7a-b), N^{th}-root stacking yields the second highest SNRs, while the SNRs of other stacking results are at a comparable level of 0-10 for both the Raw and One-bit NCFs. For the XZ NCFs (Fig. 7c-d), the lowest SNRs are observed with ACF stacking, though the ballistic phases are still clearly retrieved (Figs 6h and S3h).

4.2 Surface wave dispersion

Surface waves are dispersive, which means that their wavespeed depends on the frequency of the wave. Lower frequency (longer period) surface waves generally travel faster than higher frequency (shorter period) waves and are more sensitive to greater depths. This frequency-depth characteristic makes surface waves a popular choice for elastic and anelastic seismic tomography of the subsurface. Preserving the surface wave dispersion in the stacked NCFs is, therefore, one of the critical metrics to evaluate the performance of different stacking methods. We assess this performance by extracting and comparing the phase-velocity dispersion images. We use the Raw and One-bit NCFs from the XZ linear array between the virtual source at XZ.A02 and all other receivers (Fig. 6) to evaluate this metric.

We follow the phase-shift dispersion analysis procedure described in Park et al. (1998, 1999) and Ryden et al. (2004). We implement the procedure in the time domain through narrow-band filters. We narrow bandpass the seismograms using a Butterworth filter progressively between 1 and 15 seconds period, with a moving period band of 2 seconds and a step size of 0.1 s. We then shift the seismogram with a phase shift of r/v, where r is the inter-station distance and v is the phase velocity that varies between 1.5 to 5.5 km/s, with a step of 0.05 km/s. We trim the data over an adaptive window length that scales with the central period, computed as

$$L_i = a_i T_i, \tag{11}$$

where L_i is the window length for the i^{th} period T_i. a_i is the scaling factor, which is determined by

$$a_i = a_{min} + (i-1)\frac{a_{max} - a_{min}}{N_T}, \tag{12}$$

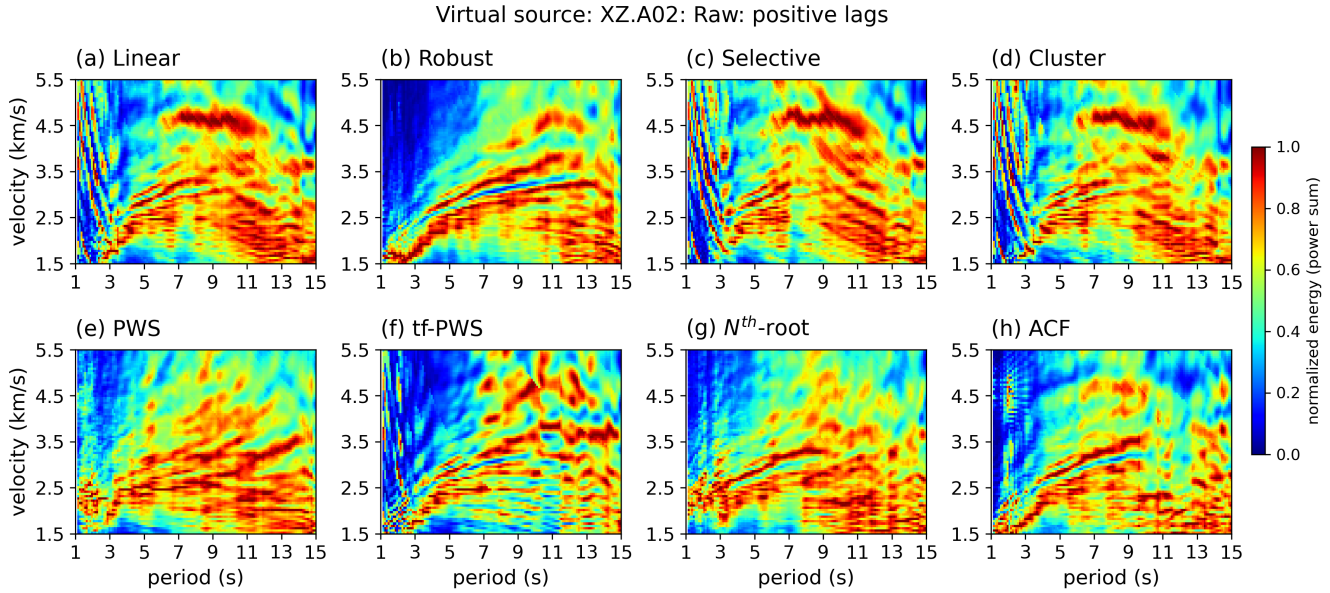

Figure 8. Dispersion images extracted from the Raw NCFs between the virtual source at XZ.A02 and other receivers using the time-domain phase-shift method through multichannel analysis of surface waves (e.g., Park et al. 1998, 1999; Ryden et al. 2004). (a-h) Results from the stacks using Linear, Robust, Selective, Cluster, PWS, tf-PWS, N^{th}-root, and ACF stacking methods, respectively. The color scale shows the power sum normalized at each velocity value. The NCFs are shown in Fig. 6

where a_{min} and a_{max} are the minimum and maximum of the scaling range and N_T is the number of period steps. For our examples, a_{min} = 1 and a_{max} = 2. Finally, to extract the dispersion image, we calculate the energy (sum of squared amplitude) of the windowed, filtered, shifted, and stacked seismograms. The images are shown in Figs 8 and 9.

We apply this procedure to all station pairs with the common virtual source XZ.A02, with only results from the positive

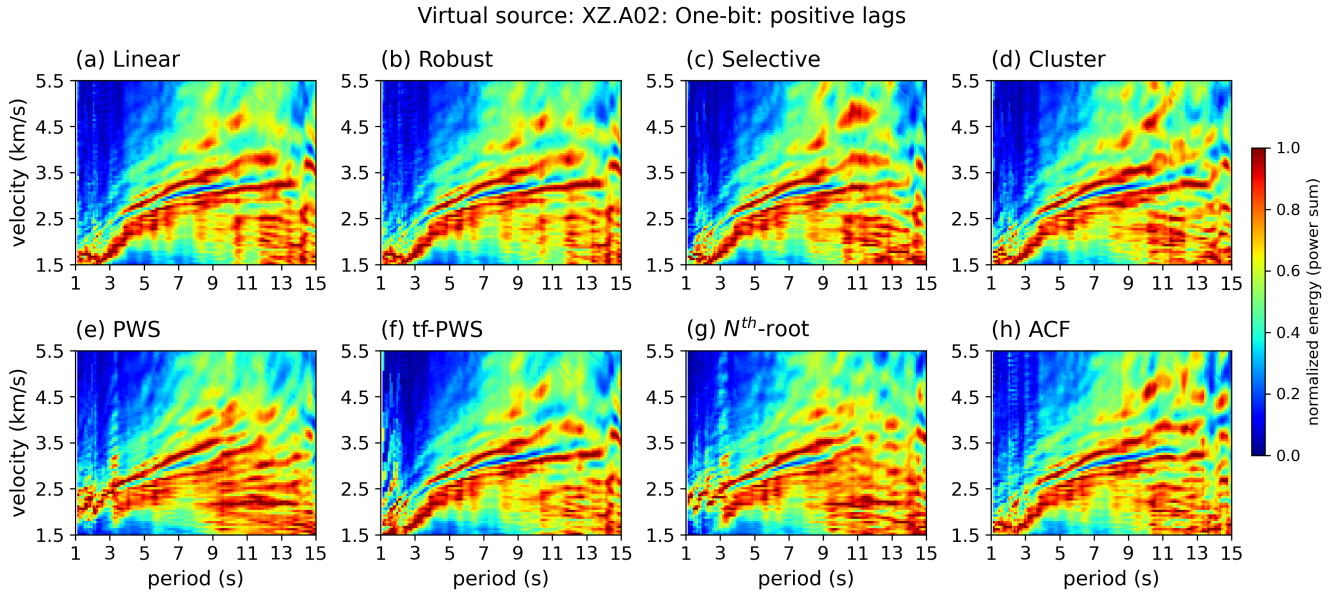

Figure 9. Same as Fig. 8 but for stacked One-bit NCFs as shown in Fig. S3 in the supplement.

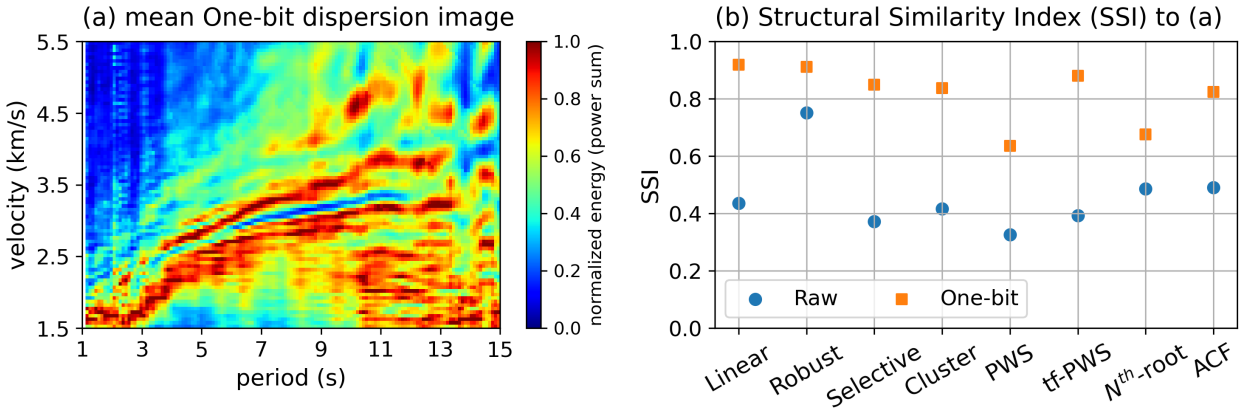

Figure 10. Comparison of the dispersion images in Figs 8 and 9. (a) The average dispersion image of the One-bit NCF stacks in Fig. 9, which is used as the reference for the comparison. (b) Structural Similarity Index (SSI) between individual dispersion images for Raw (dots) and One-bit (squares) NCF stacks and the reference dispersion image as in (a). The dispersion images are shown in Figs 8 and 9.

lags of the NCFs being shown here. We only analyze the results from station pairs with inter-station distances of at least 1.5 times the wavelength for each velocity-period pair to remain in the far-field regime.

The dispersion images from the One-bit stacks are largely similar to each other (Fig. 9). This contrasts with the great variability of the dispersion images from the Raw NCF stacks (Fig. 8). We use the average dispersion image of the One-bit results (Fig. 10a) as a reference. We compute the Structural Similarity Index (SSI; Z. Wang et al. 2004) between the reference (Fig. 10a) and individual dispersion images from the NCF stacks using different methods (Figs 8 and 9). The SSI is widely used in the analysis of image degradation or alteration, with 1 meaning the two images are identical and 0 meaning they are completely different. The SSI results over the entire period range of 1-15 s for both the Raw and One-bit dispersion images are shown in Fig. 10b. To examine the performance of different stacking methods at different period bands, we also compute the frequency-dependent SSI, as shown in Fig. 11.

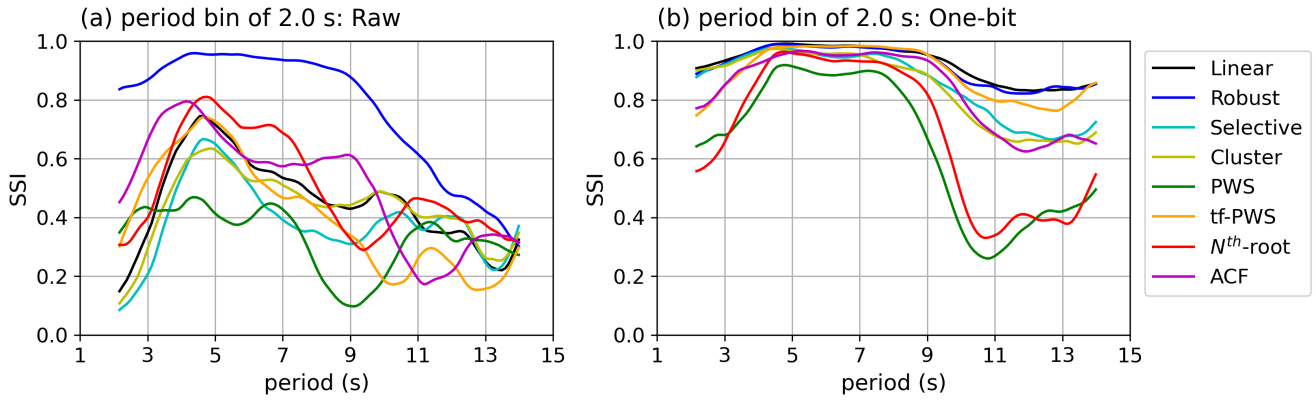

Figure 11. Frequency-dependent comparison of the dispersion images with the average dispersion image of the One-bit NCF stacks (Fig. 10a). (a) Structural Similarity Index (SSI) between individual dispersion images for Raw NCF stacks with a period bin of 2 s and a step of 0.1 s. The x-axis shows the central period of each bin. (b) Same as (a) but for One-bit NCF stacks. The dispersion images are shown in Figs 8 and 9.

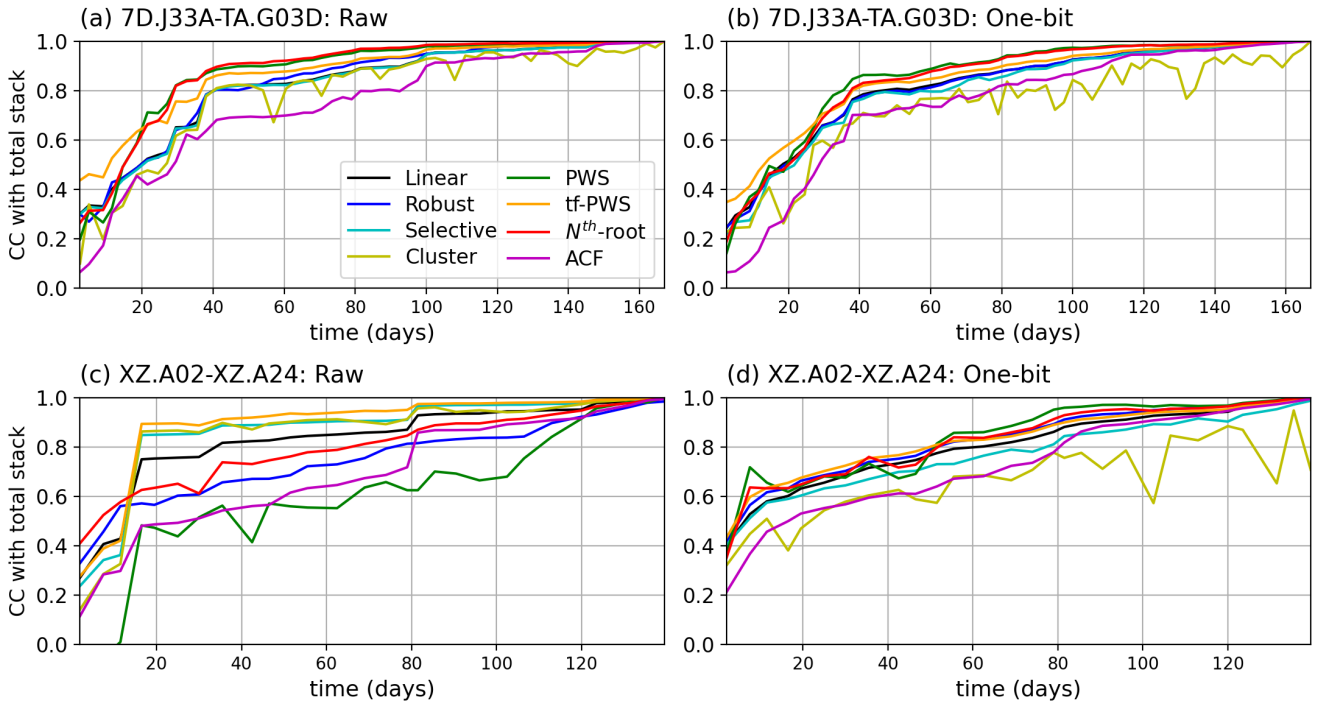

Figure 12. Convergence to long-term stacks with the increasing number of NCFs, shown as the equivalent time span of NCFs in days. (a-b) Correlation coefficients (CC) between the progressive stacking results and the stacks over the entire time period for Raw and One-bit NCFs between station pair 7D.J33A-TA.G03D. The pre-stack NCFs are shown in Fig. 2a-b. We only use NCFs before 4/1/2012 for the convergence analysis to exclude the much noisier data in later dates. All stacks are filtered at 0.1-0.4 Hz. (c-d) Same as (a) and (b) but for the XZ station pair XZ.A02-XZ.A24. The pre-stack NCFs are shown in Fig. 2c-d. We use all NCFs for this station pair.

The reference dispersion image (Fig. 10a) resembles major features of the One-bit dispersion images using different stacking methods (Fig. 9). Over the entire period range of 1-15 s, the dispersion images using the PWS and N^{th}-root methods for One-bit NCFs differ the most from the reference, with SSI below 0.8 (Fig. 10b). All other stacking methods show higher SSI values. In contrast with the One-bit results, most of the dispersion images for the Raw NCFs possess relatively lower SSI values (< 0.6), except for the Robust stacking method, which shows a SSI of close to 0.8. The frequency-dependent results (Fig. 11) show that all One-bit NCF stacks (Fig. 11b) and the Robust stack of the Raw NCFs (Fig. 11a) retrieve the most consistent dispersion images at the period range of 4-9 s. Outside this period range, the SSI values for the PWS and N^{th}-root stacks decrease dramatically. All stacking results of Raw NCFs, except for the PWS result, show SSI values of > 0.6 in a narrow band of 3-6 s (Fig. 11a). In summary, the Robust stacking method preserves the dispersion information the best for the Raw NCFs, with a dispersion image closest to the average of One-bit NCF stacks. For the One-bit NCFs, the Linear, Robust and tf-PWS stacks perform comparably well in preserving the dispersion information over the entire examined period range from 1 to 15 s.

4.3 Convergence of short-term stacks to the long-term stack

In measuring the temporal changes associated with the elastic properties of the crust, we often stack a subset of the short-term NCFs over an intermediate time period (e.g., F. Brenguier et al. 2008; Clements and Marine A Denolle 2018; Donaldson et al. 2019; Hadziioannou et al. 2011; Olivier et al. 2019; Seats et al. 2012; Yuan et al. 2021). The time needed to reach convergence limits the temporal resolution of the monitoring of the change in elastic properties. Here, we quantify the *rate of convergence* as the time duration of NCFs needed to achieve a certain threshold of the correlation coefficient between the stack of the subset NCFs and the reference stack of NCFs over the entire period. We compute the correlation coefficient between the reference stack and the stack of a progressively growing subset of NCFs ordered by date and time with an increment of 5 NCFs (Fig. 12).

The rate of convergence varies among methods (Fig. 12). Overall, the PWS, tf-PWS, and N^{th}-root stacking results have the fastest convergence to the reference stack (Fig. 12a, b, and d), except for the Raw NCFs from the XZ station pair (Fig. 12c) where the PWS method converges the slowest. The ACF stacking result converges the slowest for the Raw NCFs of the Cascadia amphibious station pair (Fig. 12a). The Cluster stacking method converges the slowest for the One-bit NCFs from both station pairs (Fig. 12b and d), though it shows a fast convergence rate for the Raw NCFs of the XZ station par (Fig. 12c). The strong fluctuation of the correlation coefficient using Cluster stacking may be due to the fact that the k-means clustering method involves some random processes in assigning the clusters. Most stacking results achieve high ($\geq$ 0.7) correlation coefficients with more than about 35 days (Fig. 12a-b). For the XZ NCFs to achieve a similar correlation coefficient, it takes about 50 days (Fig. 12c-d). From the results shown here, the rate of convergence depends strongly on the specific dataset and the quality of individual short-term NCFs.

4.4 Transient phase changes

Seismic monitoring using ambient noise interferometry relies on the phase difference of seismic waves measured at different lag times (e.g., F. Brenguier et al. 2008; Clements and Marine A Denolle 2018; Donaldson et al. 2019; Hadziioannou et al. 2011; Olivier et al. 2019; Seats et al. 2012; Yuan et al. 2021). We analyze how these phase changes are preserved with different stacking methods by estimating the velocity changes (dv/v) over time. We utilize the trace stretching method (e.g., Florent Brenguier et al. 2008; Lobkis and Weaver 2003; A. Obermann et al. 2014; Yuan et al. 2021) to measure the dv/v between the two example station pairs: 7D.J33A-TA.G03D and XZ.A02-XZ.A24 (Fig. 13). The trace stretching method estimates dv/v by maximizing the correlation coefficient between the individual NCF coda with the reference coda through linearly stretching or squeezing the waveform within the specified time window. The Python function for trace stretching is built in the *SeisGo* interface (X. Yang et al. 2022a). For all results shown in Fig. 13, we measure the velocity changes in the frequency range of 0.1-0.4 Hz. For all datasets, we use a sub-stacking window length of 96 hours (4 days). The measuring windows are 86-106 s for 7D.J33A-TA.G03D to capture the coda waves and 29-49 s for XZ.A02-XZ.A24 containing the ballistic phases. The dv/v measuring windows for both datasets are marked in Fig. 2.

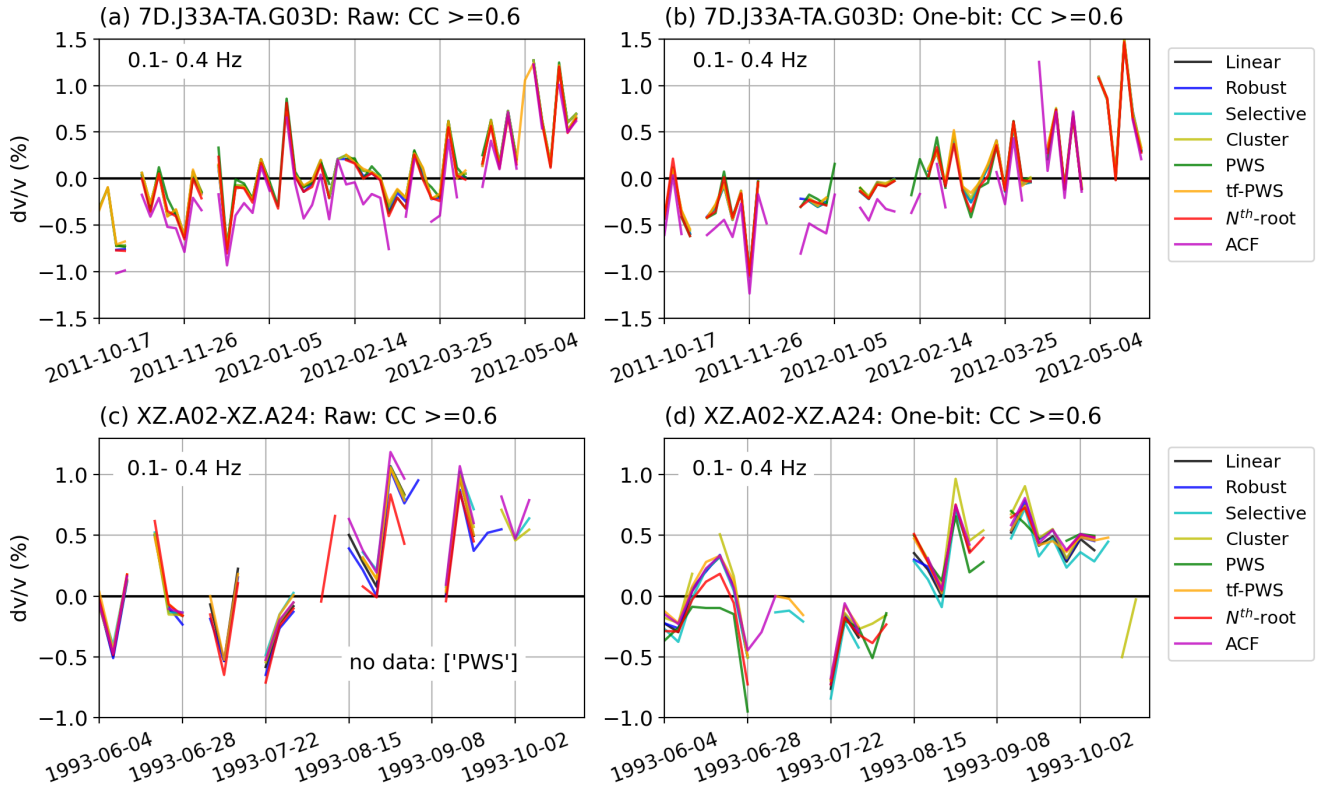

Figure 13. Transient velocity changes (dv/v) using the trace stretching method measured from NCFs from (a) 7D.J33A-TA.G03D and (b) XZ.A02-XZ.A24. The pre-stack short-term NCFs and the measuring windows are shown in Fig. 2. We use coda waves for the amphibious station pair (7D.J33A-TA.G03D) and ballistic phases for the XZ station pair (XZ.A02-XZ.A24). The coda waves for the XZ station pair (both Raw and One-bit NCFs) are not coherent and produce unstable dv/v measurements, regardless of the stacking method used. For 7D.J33A-TA.G03D, we only analyze the NCFs before 6/2/2012, during which we have clear coda phases on most of the NCFs. We sub-stack the NCFs over 96 hours (4 days) for all station pairs. We only plot dv/v results with the post-stretching correlation coefficient (CC) of ≥ 0.6.

For each station pair, different stacking methods produce very similar patterns of changes in dv/v over the observational periods (Fig. 13). An exception is the dv/v measurements from the PWS stacking result of the Raw XZ NCFs, with stronger variations and much lower after-stretching correlation coefficient compared to other stacking results. After applying a threshold of geq 0.6 for the post-stretching correlation coefficient in plotting Fig. 13, there is no PWS measurement left. For 7D.J33A-TA.G03D, a general increase in seismic velocity from -0.5% to 0.5% is observed from both Raw and One-bit NCFs using all stacking methods (Fig. 13a-b). For XZ.A02-XZ.A24, the dv/v fluctuates between -0.5% and 0.5% from June to July of 1993 (Fig. 13c-d). From August 1993 to the end of the observational period, the dv/v varies between 0% and 1% (Fig. 13c-d). While we do not intend to interpret these dv/v measurements, we note the great similarity of the measurements across methods. All stacking methods show comparable results in preserving the phase differences for coda waves from both Raw and One-bit NCFs and the ballistic phases from One-bit NCFs. The PWS method fails with this metric for ballistic phases from the XZ Raw NCFs.

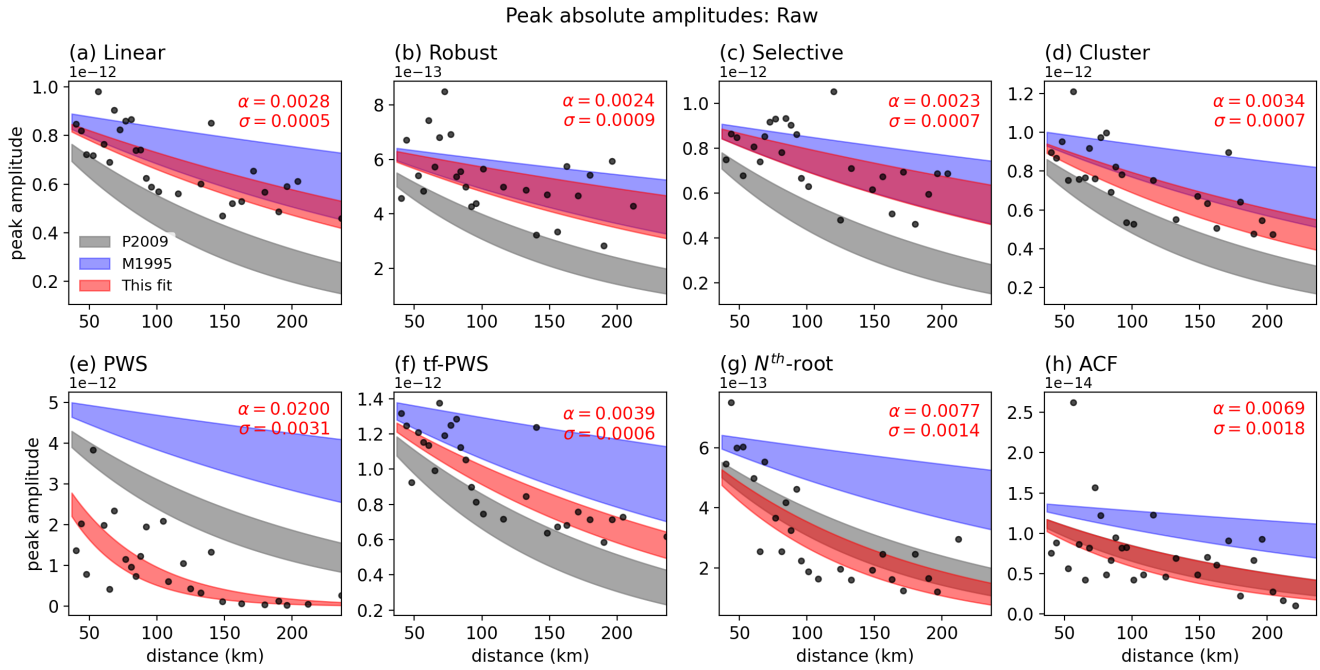

Figure 14. Peak absolute amplitudes of the positive lags of the XZ Raw NCFs within the predicted ballistic arrival windows computed using a velocity range of 2-3.7 km/s and corrected by $\sqrt{D}$, D is the inter-station distance. In each panel, we only show the attenuation parameter, α, and its standard deviation, σ. The exponential fit from this study is shown as the red shaded area. For reference, we also show the amplitude decay estimated based on the attenuation parameters from Prieto et al. (2009) (P2009; α=0.0064$\pm$0.0013; gray shaded area) and Mitchell (1995) (M1995; α=0.002$\pm$0.001; blue shaded area). See Fig. 6 for the moveout plots of the stacked NCFs.

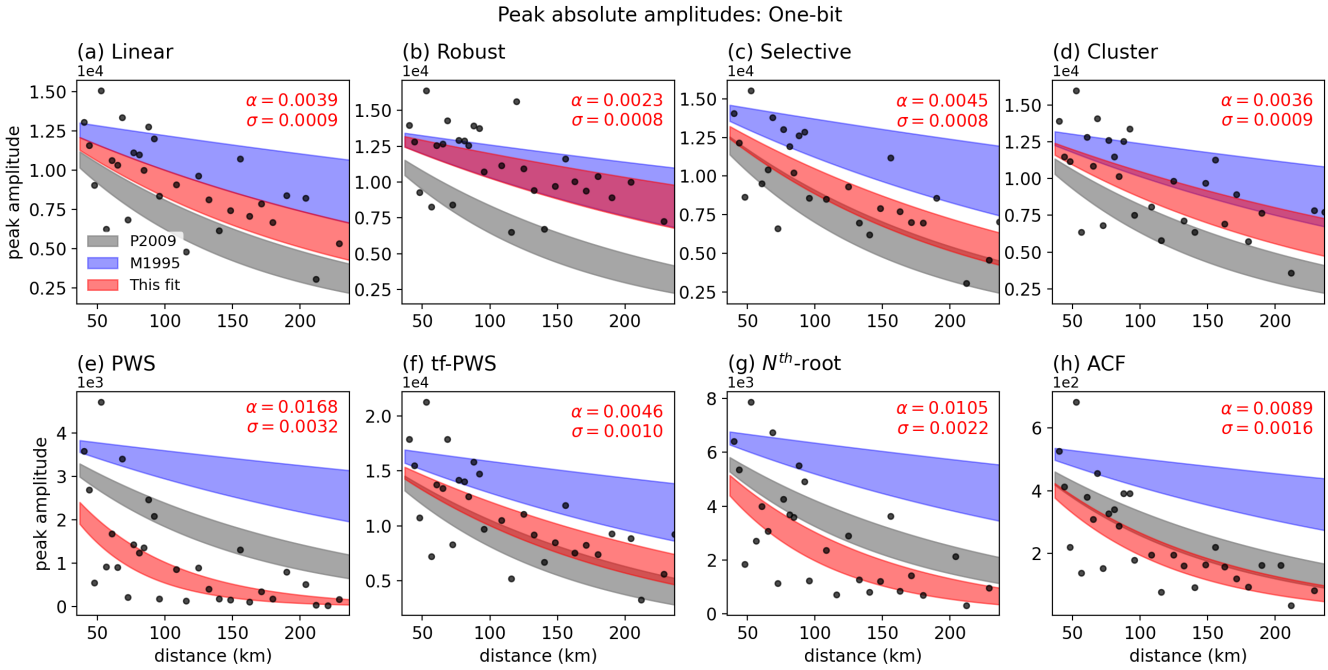

Figure 15. Same as Fig. 14 but for measurements from the XZ One-bit NCFs. See Fig. S3 in the supplement for the moveout plots of the stacked NCFs.

4.5 Peak amplitudes of ballistic phases

Important information about the Earth's anelastic structure can be retrieved from the relative amplitudes of NCFsPrieto et al. 2009. Here, we focus on the amplitude decay with distance and ignore the spatial variations in relative amplitudes that are subject to additional uncertainties. The NCFs chosen in this example are the same as in Fig 6 and supplementary Fig S3, filtered at 0.1-0.4 Hz. The NCFs at each station pair are stacked over the entire duration of data availability. We then measure the peak absolute amplitudes of the predicted ballistic waves for each station pair NCF stack within the time windows computed with a velocity range of 2-3.7 km/s (Figs 6 and S3). The XZ broadband network is an East-West trending linear array in central Oregon, United States (Fig. 1). It is located at the active Cascadia convergent margin, spanning from the Cascadia Volcanic Arc in the East to the coastal mountains in the West. Considering that noise is dominantly generated by the oceanic microseisms in this frequency range (e.g., Webb 1998; Y. Yang and Ritzwoller 2008), the XZ linear array provides an appropriate dataset for analyzing the attenuation of surface waves extracted from NCFs, minimizing the azimuthal dependence of noise sources. In this work, we only aim to compare the stability of the relative amplitude information with different stacking methods and do not intend to advocate for an estimate of attenuation, which might still depend on the distribution of noise sources (Laurent Stehly and Boué 2017). The microseismic noise is generated by the ocean and dominantly propagates eastward. Therefore, XZ.A02 is an appropriate choice for a virtual source (Fig. 1). We only measure the attenuation from the positive lags of the NCFs.

Assuming that the main signals of the NCF are fundamental-mode surface waves, the maximum absolute amplitudes of the stacked NCFs decay with distance following the relation:

$$A(D) = A_0/\sqrt{D}\exp(-\alpha D), \tag{13}$$

where D is the inter-station distance, A_0 is a reference amplitude at the virtual source, α is a measure of "attenuation" that could be attributed to intrinsic and scattering attenuation that further reduces the ground motion. To fit α, we correct for the geometrical spreading by scaling the amplitudes with a factor of $\sqrt{D}$. The peak absolute amplitudes are then fit to an exponential function $y = e^{-\alpha x}$. We estimate the best fit through least square linear regression of the peak amplitudes in a natural logarithmic space using the following steps: 1) calculate the natural log peak absolute amplitudes, 2) fit all data points using a least square linear regression (`scipy.stats.linregress`), 3) correct the data with the best fit model and measure misfit, 4) compute the mean and standard deviation of the misfit, 5) remove the outliers (data points that are more than one standard deviation away from the mean), and 6) use the remaining subset of the data to repeat step-2 to get the final linear fit parameters. The slope parameter in the linear regression from Step-2 is the attenuation factor α. The error in slope estimate given by `scipy.stats.linregress` is then used as the uncertainty of α, as shown in Figs 14 and 15.

We only use subsets of the NCFs in Figs 6 and S3 that have inter-station distances between 37 km, which is the wavelength of a 0.1 Hz wave traveling at a velocity of 3.7 km/s, and 240 km. The exponential fit and the parameters the uncertainties are shown as the red shaded areas in Figs 14 and 15. For comparison, we also plot the independent attenuation measurements

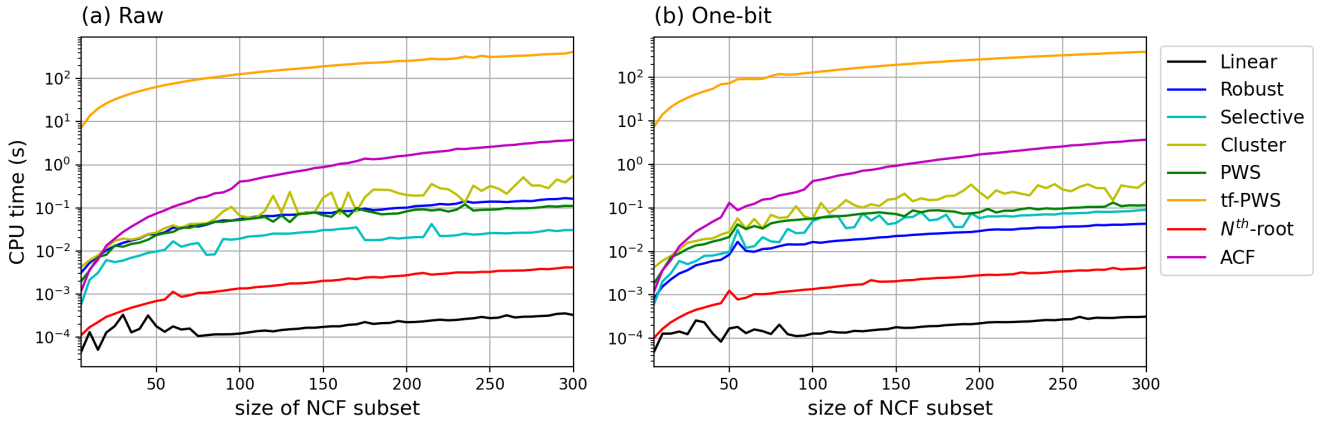

Figure 16. Computing times spent on stacking as a function of the number of NCFs to stack in a 5-step increment for (a) Raw and (b) One-bit NCFs from station pair 7D.J33A-TA.G03D.

from Prieto et al. (2009) (α=0.0064$\pm$0.0013), who used spectral methods to measure attenuation in the Los Angeles basin in southern California using ambient noise NCF filtered at 0.2 Hz, and Mitchell (1995) (α=0.002$\pm$0.001) for active tectonic regions measured at 0.1-0.2 Hz using a global compilation of earthquake surface waves.

The decay coefficient α can be measured with all methods and for all pre-processing techniques. However, there is great variability in the model parameter among the methods (0.0023-0.02 for Raw NCFs and 0.0023-0.0168 for One-bit NCFs; Figs 14-15). For the Raw NCFs (Fig. 14), the attenuation parameters measured from the stacking results using Linear, Robust, Selective, and Cluster methods are similar to the values by Mitchell (1995). The coefficients using the N^{th}-root and ACF methods are comparable to that found by Prieto et al. (2009). The coefficient from the tf-PWS method falls in-between the values provided by Mitchell (1995) and Prieto et al. (2009). For the estimates from One-bit NCFs (Fig. 15), the result using the Robust method overlaps the most with the value proposed by Mitchell (1995). The Linear, Selective, Cluster, and tf-PWS stacking results all fall in-between the values by Mitchell (1995) and Prieto et al. (2009). The estimates of α from the N^{th}-root and ACF stacks are similar to that used by Prieto et al. (2009). Compared to the attenuation measurements from the Raw NCFs, the attenuation measurements from the One-bit NCFs seem to be more sensitive to the choice of stacking methods, shown as larger variations among different methods. For both datasets, the PWS method tends to overestimate the attenuation factors, α, with much higher values than other methods.

4.6 Computational expenses

The computational efficiency of different stacking methods varies significantly among different stacking methods but does not change between Raw and One-bit NCFs. The compute times are estimated from a single core 3.6 GHz Intel Core i9 CPU. We compare the compute time spent stacking subsets of the NCFs (with an increment of 5 NCFs) using the different stacking methods (Fig. 16). The tf-PWS stacking method uses the most CPU time, while Linear stacking is the fastest method (Fig. 16). Most of the time in tf-PWS is spent on computing the Stockwell transform on individual NCFs. While our implementation

Table 1. Our recommendations of stacking methods for major research applications using NCFs based on the evaluation metrics in Section 4. For methods denoted with (*), see the appropriate sections for additional discussion.

Applications	**Evaluation metrics**	**Recommendations**
Velocity tomography (Section 5.1)	Signal-to-noise ratio (Section 4.1), surface wave dispersion (Section 4.2), and computational expenses (Section 4.6)	Raw: Robust One-bit: Linear, Robust, Selective, ACF, Cluster
Monitoring (Section 5.2)	Signal-to-noise ratio (Section 4.1), surface-wave dispersion (Section 4.2), convergence of short-term stacks to the long-term stack (Section 4.3), transient phase changes (Section 4.4), and computational expenses (Section 4.6)	Raw: Robust One-bit: Linear, Robust, Selective, tf-PWS(*)
Anelastic properties (Section 5.3)	Signal-to-noise ratio (Section 4.1), surface wave dispersion (Section 4.2), peak amplitudes of ballistic phases (Section 4.5), and computational expenses (Section 4.6)	Raw: Robust One-bit: Robust, Cluster

follows that of the original Stockwell transform, Baig et al. (2009) found a more efficient algorithm to compute the Stockwell transform. The total time to perform tf-PWS is about six orders of magnitudes longer than that of the Linear stacking method. For example, tf-PWS takes about 400 seconds to stack 300 NCFs, while Linear stacking takes 0.0005 seconds. The ACF stacking method is the second most expensive method in terms of computing time, with about 4 seconds to stack 300 NCFs. This is because the ACF method needs to compute the spectrum of each individual trace as well as the cross-spectrum in moving windows. The Robust, Selective, and Cluster methods take about 0.02-0.4 seconds to stack 300 NCFs. The N^{th}-root method is the second fastest method following the Linear stack. These methods all scale nearly linearly with the number of NCFs to stack.

5 CHOICE OF STACKING METHOD

In this section, we discuss the choice of stacking methods for different applications that utilize NCFs. We focus on the following major applications: 1) tomography of velocity structures, 2) monitoring of transient velocity changes, and 3) characterization of anelastic properties. These three research applications extract different information from the NCFs: 1) seismic velocity tomography using surface waves requires the dispersion information of the ballistic surface-wave phases, 2) seismic monitoring requires dispersion and slight variations in phase information of the ballistic phases and/or the coda waves, and 3) the imaging of seismic attenuation and prediction of strong ground motion relies on relative amplitudes of the ballistic phases. Although the performance could vary with different datasets, we provide here the advantages and limitations of the stacking methods for each application in Table 1.

5.1 Stacking for tomography of velocity structures

Ambient noise tomography typically requires extracting the surface wave dispersion information from the stack of all NCFs for each station pair (e.g., Bensen et al. 2007). The fundamental mode surface waves are most commonly visible in the data, though there are some studies using higher modes for imaging (e.g., Jiang and Marine A. Denolle 2022; G. Wu et al. 2020). Our recommendation for seismic velocity tomography is based on the following metrics: signal-to-noise ratio (Section 4.1), surface wave dispersion (Section 4.2), and computational expense (Section 4.6).

Based on our examples in Section 4.1, except for the OBS-OBS station pairs, the PWS method produces stacks with the highest average SNRs, though with a large variance (see Figs 4-7 and S1-S3 in the supplement). The stacks using the N^{th}-root method also have relatively high SNRs. Other stacking methods perform at a comparable level in terms of the SNRs of the ballistic phases. The phase dispersion is better recovered using the Robust method than others for Raw NCFs over the entire examined period of 1-15 s (Fig. 10b), especially in the period range of 4-9 s (Fig. 11b). When using One-bit NCFs, all but the PWS and N^{th}-root methods perform well overall in the dispersion analysis (Figs 9, 10b, and 11b). However, the high computational cost of the tf-PWS method makes it impractical for processing large data sets for tomographic imaging (Fig. 16). In summary, our recommendations for tomographic imaging are the Robust stacking for Raw NCFs and the Linear, Robust, Selective, Cluster, and ACF methods for One-bit NCFs.

5.2 Stacking for monitoring of transient velocity changes

Seismic monitoring uses ballistic or coda wave interferometry to infer small changes in the subsurface from short-time stacks of NCFs (e.g., F. Brenguier et al. 2008; Clements and Marine A Denolle 2018; Donaldson et al. 2019; Lobkis and Weaver 2003; Anne Obermann and Hillers 2019). A faster convergence of the NCF would lead to a higher temporal resolution in seismic monitoring. The convergence is often hindered and thus limits the temporal resolution (Hadziioannou et al. 2011). It is therefore important to find the optimal length of data that yields a reasonable convergence of the NCF stack. Meanwhile, time-lapse imaging requires that the dispersion of surface waves remains stable through time (e.g., Bergamo et al. 2016). Therefore, our recommendation for seismic monitoring using ambient noise interferometry is based on the following additional metrics on top of those for tomography (Sections 4.1, 4.2, and 4.6): convergence of short-term stack to long-term stack (Section 4.3) and transient phase changes (Section 4.4).

Regarding the convergence metric, all stacking methods show comparable performance (Fig. 12). The relatively large variability in the rate of convergence using the Cluster stacking method may lead to artifacts in monitoring. The ACF method is also relatively slow in converging to the long-term stack. It is worth noting again that the rate/time of convergence also depends on the quality of specific datasets. The dispersion analyses are most stable using the Robust method for Raw NCFs. For One-bit NCFs, all but the PWS and N^{th}-root methods are appropriate choices (Figs 8-11).

For transient phase changes, all but the PWS method perform comparably well for all of the four examined examples (Fig.

13). The tf-PWS method may be used to process the NCFs when long computational times are acceptable to the user, though its computational efficiency needs to be improved to process large datasets (Fig. 16), or when using Graphic Processing Units (Zeng and Thurber 2016). In summary, our recommendations for seismic monitoring are the Robust method for Raw NCFs and the Linear, Robust, and Selective methods for One-bit NCFs.

5.3 Stacking for characterization of anelastic properties

The relative amplitudes of NCFs have been used to characterize the seismic attenuation properties of the Earth's lithosphere (Prieto et al. 2009) and the ground motion patterns (e.g., Marine A Denolle et al. 2013, 2014; Viens et al. 2017). Therefore, in addition to the metrics pertinent to tomography (Sections 4.1, 4.2, and 4.6), our recommendation for the characterization of anelastic properties using NCFs also evaluates the stability of peak absolute amplitudes of ballistic phases (Section 4.5).

Mitchell (1995) measured attenuation from a compilation of global earthquake records from a variety of tectonic settings, including active margins, which are similar to the setting for the XZ linear array at the Cascadia margin. Because it is measured from earthquake surface-wave data, we argue that the attenuation parameter by Mitchell (1995) can serve as an independent benchmark in our evaluation. With this criterion, the Linear, Robust, Cluster, and ACF stacking methods are ranked as the top four among all of the methods for Raw NCFs (Fig. 14). However, the requirement of a reliable measurement of frequency-dependent attenuation and dispersion (Figs 8, 10a, and 11a) would narrow the choice down to Robust stacking for Raw NCFs. For One-bit NCFs, the Robust and Cluster stacking methods rank the best among all methods. Our recommendations for attenuation and ground motion studies, or any applications utilizing relative amplitudes, are the Robust method for Raw NCFs and the Robust and Cluster methods for One-bit NCFs.

6 CONCLUSIONS

NCFs are widely used in seismic velocity and attenuation imaging, monitoring, and ground motion analyses. The stacking of NCFs over longer time periods is needed for most applications utilizing NCFs to increase the strength of the coherent signals. There have been many temporal stacking methods developed to improve the stacking of NCFs. We compare eight temporal stacking methods, including Linear, Robust, Selective, Cluster, Phase-Weighted, time-frequency Phase-Weighted, N^{th}-root, and Adaptive Covariance Filter stacking methods, to investigate their performance to enhance the quality of the stacked NCF. We examine the performance of these methods with six metrics, including signal-to-noise ratios, surface-wave phase velocity dispersion, the convergence of short-term stacks to the long-term stack, wavespeed changes, peak amplitudes, and computational expenses. Our analyses demonstrate that although all methods are able to retrieve clear ballistic phases, their spectral contents and peak amplitudes vary strongly across methods. Based on multiple evaluation metrics, we recommend the Robust method for all three categories of applications using Raw NCFs, including velocity tomography, monitoring, and attenuation studies. For tomography using One-bit NCFs, all methods except for Phase-Weighted, time-frequency phase-weighted, and

N^{th}-root stacking methods are good choices. For monitoring using One-bit NCFs, the Linear, Robust, and Selective stacking methods are preferred choices, with the possibility of using time-frequency Phase-Weighted stacking for processing small datasets. For applications utilizing One-bit NCFs to extract relative amplitude information, both the Robust and Cluster stacking methods perform well. The findings in this study provide a practical guideline for choosing the appropriate stacking method for major applications utilizing NCFs. This work did not address the validity and effects of pre-processing techniques on our various use cases, but future work might address this (e.g., Fichtner et al. 2020). The open-source computer codes produced in this study can also be used for general time series stacking analyses.

ACKNOWLEDGMENTS

This research has been funded by the startup funding of Purdue University for Xiaotao Yang, ARC grant DE220100907 to Chengxin Jiang, and NSF grant EAR1749556 to Marine Denolle. Data from the TA network were made freely available as part of the EarthScope USArray facility, operated by Incorporated Research Institutions for Seismology (IRIS) and supported by the National Science Foundation, under Cooperative Agreements EAR-1261681. ***Author Contributions:*** X. Yang contributed to conceptualization, supervision, funding acquisition, data curation, methodology, formal analysis, software development, and writing of the original draft. M. Denolle contributed to the conceptualization, supervision, and funding acquisition of the project. J. Bryan and K. Okubo contributed to methodology development. C. Jiang, M. Denolle, and T. Clements contributed to the software development. All authors contributed to reviewing and editing of the manuscript.

Data Availability

The seismic waveforms were downloaded from IRIS Data Management Center and processed using *SeisGo* (`https://doi.org/10.5281/zenodo.5873724`), built upon *ObsPy* (Beyreuther et al. 2010) and *NoisePy* (Jiang and Marine A Denolle 2020). The computer codes for the stacking methods are available as a standalone Python package, *StackMaster*, which is archived on Zenodo (`https://doi.org/10.5281/zenodo.5951013`) and is accessible on Python Package Index (`https://pypi.org/project/stackmaster`). All Python scripts and Jupyter notebooks used in the processing and visualizing of the data and stacking results are available at `https://github.com/xtyangpsp/Paper_OptimalStacking` and are archived on Zenodo (`https://doi.org/10.5281/zenodo.6363585`).

References

Baig, A. M., M. Campillo, and F. Brenguier (2009). "Denoising seismic noise cross correlations". In: *Journal of Geophysical Research: Solid Earth* 114.8, pp. 1–12. ISSN: 21699356. DOI: 10.1029/2008JB006085.

Bensen, G. D. et al. (2007). "Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements". In: *Geophysical Journal International* 169.3, pp. 1239–1260. ISSN: 0956-540X. DOI: 10.1111/j.1365-246X.2007.03374.x. eprint: http://oup.prod.sis.lan/gji/article-pdf/169/3/1239/6082124/169-3-1239.pdf. URL: https://doi.org/10.1111/j.1365-246X.2007.03374.x.

Bergamo, Paolo et al. (2016). "Time-lapse monitoring of climate effects on earthworks using surface wavesTime-lapse seismic monitoring with SW". In: *Geophysics* 81.2, EN1–EN15.

Beyreuther, Moritz et al. (2010). "ObsPy: A Python Toolbox for Seismology". In: *Seismological Research Letters* 81.3, pp. 530–533. ISSN: 0895-0695. DOI: 10.1785/gssrl.81.3.530. eprint: https://pubs.geoscienceworld.org/ssa/srl/article-pdf/81/3/530/2762059/530.pdf. URL: https://doi.org/10.1785/gssrl.81.3.530.

Brenguier, F. et al. (2008). "Postseismic Relaxation Along the San Andreas Fault at Parkfield from Continuous Seismological Observations". In: *Science* 321.5895, pp. 1478–1481. DOI: 10.1126/science.1160943.

Brenguier, Florent et al. (2008). "Towards forecasting volcanic eruptions using seismic noise". In: *Nature Geoscience* 1.2, pp. 126–130. ISSN: 1752-0908. DOI: 10.1038/NGEO104. arXiv: 0706.1935. URL: https://www-nature-com.ezproxy.lib.purdue.edu/articles/ngeo104.

Clements, Timothy and Marine A Denolle (2018). "Tracking Groundwater Levels Using the Ambient Seismic Field". In: *Geophysical Research Letters* 45.13, pp. 6459–6465. ISSN: 19448007. DOI: 10.1029/2018GL077706. URL: https://doi.org/10.1029/2018GL077706.

Denolle, Marine A et al. (2013). "Ground motion prediction of realistic earthquake sources using the ambient seismic field". In: *JOURNAL OF GEOPHYSICAL RESEARCH: SOLID EARTH* 118, pp. 2102–2118. DOI: 10.1029/2012JB009603. URL: www.data.scec.org.

— (2014). "Strong Ground Motion Prediction Using Virtual Earthquakes". In: *Science* 343.6169, pp. 399–403. DOI: 10.1126/science.1245678. eprint: https://www.science.org/doi/pdf/10.1126/science.1245678. URL: https://www.science.org/doi/abs/10.1126/science.1245678.

Denolle, Marine A et al. (2018). "Strong shaking predicted in Tokyo from an expected M7+ Itoigawa-Shizuoka earthquake". In: *Journal of Geophysical Research: Solid Earth* 123.5, pp. 3968–3992.

Donaldson, C. et al. (2019). "Crustal seismic velocity responds to a magmatic intrusion and seasonal loading in Iceland's Northern Volcanic Zone". In: *Science Advances* 5.11. ISSN: 23752548. DOI: 10.1126/sciadv.aax6642.

Du, Zhijun, G. R. Foulger, and Weijian Mao (2000). "Noise reduction for broad-band, three-component seismograms using data-adaptive polarization filters". In: *Geophysical Journal International* 141.3, pp. 820–828. ISSN: 0956-540X. DOI: 10.1046/j.1365-246x.2000.00156.x. eprint: https://academic.oup.com/gji/article-pdf/141/3/820/1687588/141-3-820.pdf. URL: https://doi.org/10.1046/j.1365-246x.2000.00156.x.

Ermert, Laura, Antonio Villasenor, and Andreas Fichtner (2016). "Cross-correlation imaging of ambient noise sources". In: *Geophysical Journal International* 204.1, pp. 347–364.

Ermert, Laura A et al. (2021). "Multifrequency inversion of global ambient seismic sources". In: *Geophysical Journal International* 225.3, pp. 1616–1623.

Feng, Kuan-Fu et al. (2021). "Controls on Seasonal Variations of Crustal Seismic Velocity in Taiwan Using Single-Station Cross-Component Analysis of Ambient Noise Interferometry". In: *Journal of Geophysical Research: Solid Earth* 126.11, e2021JB022650.

Feng, Lili and Michael H. Ritzwoller (2019). "A 3-D Shear Velocity Model of the Crust and Uppermost Mantle Beneath Alaska Including Apparent Radial Anisotropy". In: *Journal of Geophysical Research: Solid Earth* 124.10, pp. 10468–10497. ISSN: 21699356. DOI: 10.1029/2019JB018122.

Fichtner, Andreas, Daniel Bowden, and Laura Ermert (2020). "Optimal processing for seismic noise correlations". In: *Geophysical Journal International* 223.3, pp. 1548–1564. ISSN: 0956-540X. DOI: 10.1093/gji/ggaa390. eprint: https://academic.oup.com/gji/article-pdf/223/3/1548/33762033/ggaa390.pdf. URL: https://doi.org/10.1093/gji/ggaa390.

Hadziioannou, Céline et al. (2011). "Improving temporal resolution in ambient noise monitoring of seismic wave speed". In: *Journal of Geophysical Research: Solid Earth* 116.B7.

Jiang, Chengxin and Marine A Denolle (2020). "NoisePy: A New High-Performance Python Tool for Ambient-Noise Seismology". In: *Seismological Research Letters*. ISSN: 0895-0695. DOI: 10.1785/0220190364. eprint: https://pubs.geoscienceworld.org/srl/article-pdf/doi/10.1785/0220190364/4974556/srl-2019364.1.pdf. URL: https://doi.org/10.1785/0220190364.

— (2022). "Pronounced seismic anisotropy in Kanto sedimentary basin: A case study of using dense arrays, ambient noise seismology, and multi-modal surface-wave imaging". In: *Journal of Geophysical Research: Solid Earth*, e2022JB024613. ISSN: 2169-9356. DOI: 10.1029/2022JB024613. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2022JB024613%20https://onlinelibrary.wiley.com/doi/abs/10.1029/2022JB024613%20https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022JB024613.

Kanasewich, E. R., C. D. Hemmings, and T. Alpaslan (1973). "N^{th}-Root Stack Nonlinear Multichannel Filter". In: *http://dx.doi.org/10.1190/1.1440343* 38 (2), pp. 327–338. ISSN: 0016-8033. DOI: 10.1190/1.1440343. URL: https://library.seg.org/doi/abs/10.1190/1.1440343.

Kwak, Sangmin et al. (2017). "Investigating the capability to extract impulse response functions from ambient seismic noise using a mine collapse event". In: *Geophysical Research Letters* 44.19, pp. 9653–9662.

Li, Guoliang et al. (2017). "An investigation of time–frequency domain phase-weighted stacking and its application to phase-velocity extraction from ambient noise's empirical Green's functions". In: *Geophysical Journal International* 212.2, pp. 1143–1156. ISSN: 0956-540X. DOI: 10.1093/gji/ggx448. eprint: https://academic.oup.com/gji/article-pdf/212/2/1143/22297175/ggx448.pdf. URL: https://doi.org/10.1093/gji/ggx448.

Liu, Guochang et al. (2009). "Stacking seismic data using local correlation". In: *Geophysics* 74.3, pp. V43–V48. DOI: 10.1190/1.3085643.

Lobkis, Oleg I. and Richard L. Weaver (2003). "Coda-Wave Interferometry in Finite Solids: Recovery of P-to-S Conversion Rates in an Elastodynamic Billiard". In: *Phys. Rev. Lett.* 90 (25), p. 254302. DOI: 10.1103/PhysRevLett.90.254302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.90.254302.

562 Millet, F., T. Bodin, and S. Rondenay (2019). "Multimode 3-D Kirchhoff Migration of Receiver Functions at Continental Scale". In:
563 *Journal of Geophysical Research: Solid Earth* 124.8, pp. 8953–8980. DOI: 10.1029/2018JB017288. eprint: https://agupubs.
564 onlinelibrary.wiley.com/doi/pdf/10.1029/2018JB017288. URL: https://agupubs.onlinelibrary.wiley.com/doi/
565 abs/10.1029/2018JB017288.
566 Mitchell, Brian J (1995). "Anelastic structure and evolution of the continentalcrust and upper mantle from seismic surface wave attenuation".
567 In: *Reviews of Geophysics* 33.4, pp. 441–462.
568 Moreau, L et al. (2017). "Improving ambient noise correlation functions with an SVD-based Wiener filter". In: *Geophys. J. Int.* 211.1,
569 pp. 418–426. DOI: 10.1093/gji/ggx306.
570 Muirhead, K. J. (1968). "Eliminating False Alarms when detecting Seismic Events Automatically". In: *Nature 1968 217:5128* 217 (5128),
571 pp. 533–534. ISSN: 1476-4687. DOI: 10.1038/217533a0. URL: https://www.nature.com/articles/217533a0.
572 Nabelek, J. et al. (1993). "A high-resolution image of the Cascadia subduction zone from teleseismic converted phases recorded by a
573 broadband seismic array". In: *Eos Transaction* 74.43, p. 431.
574 Nakata, Nori et al. (2015). "Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry". In: *J.*
575 *Geophys. Res.* 120.2, pp. 1159–1173. DOI: 10.1002/2015JB011870.
576 Obermann, A. et al. (2014). "Seismic noise correlations to image structural and mechanical changes associated with the Mw 7.9 2008
577 Wenchuan earthquake". In: *Journal of Geophysical Research: Solid Earth* 119.4, pp. 3155–3168. ISSN: 21699356. DOI: 10.1002/
578 2013JB010932.
579 Obermann, Anne and Gregor Hillers (2019). "Chapter Two - Seismic time-lapse interferometry across scales". In: *Recent Advances in*
580 *Seismology*. Ed. by Cedric Schmelzbach. Vol. 60. Advances in Geophysics. Elsevier, pp. 65–143. DOI: https://doi.org/10.1016/
581 bs.agph.2019.06.001. URL: https://www.sciencedirect.com/science/article/pii/S0065268719300020.
582 Olivier, G. et al. (2015). "Body-wave reconstruction from ambient seismic noise correlations in an underground mine". In: *GEOPHYSICS*
583 80.3, KS11–KS25. DOI: 10.1190/geo2014-0299.1. eprint: https://doi.org/10.1190/geo2014-0299.1. URL: https:
584 //doi.org/10.1190/geo2014-0299.1.
585 Olivier, G. et al. (2019). "Decrease in Seismic Velocity Observed Prior to the 2018 Eruption of Kīlauea Volcano With Ambient Seismic
586 Noise Interferometry". In: *Geophysical Research Letters* 46.7, pp. 3734–3744. ISSN: 19448007. DOI: 10.1029/2018GL081609.
587 Park, Choon B, Richard D. Miller, and Jianghai Xia (1998). "Imaging dispersion curves of surface waves on multi-channel record". In:
588 Society of Exploration Geophysicists. DOI: 10.1190/1.1820161. URL: https://library.seg.org/doi/10.1190/1.1820161.
589 — (1999). "Multichannel analysis of surface waves". In: *Geophysics* 64 (3), pp. 800–808. ISSN: 00168033. DOI: 10.1190/1.1444590.
590 URL: http://library.seg.org/page/policies/terms.
591 Pavlis, Gary L. and Frank L. Vernon (2010). "Array processing of teleseismic body waves with the USArray". In: *Computers & Geo-*
592 *sciences* 36.7, pp. 910–920. ISSN: 0098-3004. DOI: https://doi.org/10.1016/j.cageo.2009.10.008. URL: http:
593 //www.sciencedirect.com/science/article/pii/S0098300410001159.
594 Prieto, G. A., J. F. Lawrence, and G. C. Beroza (2009). "Anelastic Earth structure from the coherency of the ambient seismic field". In:
595 *Journal of Geophysical Research: Solid Earth* 114.7, pp. 1–15. ISSN: 21699356. DOI: 10.1029/2008JB006067.
596 Rost, Sebastian and Christine Thomas (2002). "Array seismology: Methods and applications". In: *Reviews of geophysics* 40.3, pp. 2–1.

Rückemann, Claus-Peter CP (2012). "Comparison of Stacking Methods Regarding Processing and Computing of Geoscientific Depth Data". In: *In Proc. Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2012)* c, pp. 35–40. URL: http://www.thinkmind.org/index.php?view=article%7B%5C&%7Darticleid=geoprocessing%7B%5C_%7D2012%7B%5C_%7D2%7B%5C_%7D10%7B%5C_%7D30050.

Ryden, Nils et al. (2004). "Multimodal Approach to Seismic Pavement Testing". In: *Journal of Geotechnical and Geoenvironmental Engineering*. DOI: 10.1061/ASCE1090-02412004130:6636.

Samson, J. C. and J. V. Olson (1981). "Data-adaptive polarization filters for multichannel geophysical data". In: *GEOPHYSICS* 46.10, pp. 1423–1431. DOI: 10.1190/1.1441149. eprint: https://doi.org/10.1190/1.1441149. URL: https://doi.org/10.1190/1.1441149.

Schimmel, M. and J. Gallart (2007). "Frequency-dependent phase coherence for noise suppression in seismic array data". In: *Journal of Geophysical Research: Solid Earth* 112.4, pp. 1–14. ISSN: 21699356. DOI: 10.1029/2006JB004680.

Schimmel, M. and H. Paulssen (1997). "Noise reduction and detection of weak, coherent signals through phase-weighted stacks". In: *Geophysical Journal International* 130.2, pp. 497–505. ISSN: 0956540X. DOI: 10.1111/j.1365-246X.1997.tb05664.x.

Schimmel, M., E. Stutzmann, and J. Gallart (2011). "Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale". In: *Geophysical Journal International* 184.1, pp. 494–506. ISSN: 0956540X. DOI: 10.1111/j.1365-246X.2010.04861.x.

Seats, Kevin J, Jesse F Lawrence, and German A Prieto (2012). "Improved ambient noise correlation functions using Welch' s method". In: *Geophysical Journal International* 188.2, pp. 513–523.

Shapiro, Nikolai M. et al. (2005). "High-resolution surface-wave tomography from ambient seismic noise". In: *Science* 307 (5715), pp. 1615–1618. ISSN: 00368075. DOI: 10.1126/SCIENCE.1108339. URL: www.iris.iris.edu/USArray.

Shen, Yang et al. (2012). "An Improved Method to Extract Very-Broadband Empirical Green's Functions from Ambient Seismic Noise". In: *Bulletin of the Seismological Society of America* 102.4, pp. 1872–1877. ISSN: 0037-1106. DOI: 10.1785/0120120023. eprint: https://pubs.geoscienceworld.org/bssa/article-pdf/102/4/1872/2667786/1872.pdf. URL: https://doi.org/10.1785/0120120023.

Stehly, L, Michel Campillo, and NM Shapiro (2006). "A study of the seismic noise from its long-range correlation properties". In: *Journal of Geophysical Research: Solid Earth* 111.B10.

Stehly, Laurent and Pierre Boué (2017). "On the interpretation of the amplitude decay of noise correlations computed along a line of receivers". In: *Geophysical Journal International* 209.1, pp. 358–372. ISSN: 0956-540X. DOI: 10.1093/gji/ggx021. eprint: https://academic.oup.com/gji/article-pdf/209/1/358/10763137/ggx021.pdf. URL: https://doi.org/10.1093/gji/ggx021.

Stehly, Laurent, Paul Cupillard, and Barbara Romanowicz (2011). "Towards improving ambient noise tomography using simultaneously curvelet denoising filters and SEM simulations of seismic ambient noise". In: *Comptes Rendus Geoscience* 343.8-9, pp. 591–599.

Stockwell, R. G. (2007). "A basis for efficient representation of the S-transform". In: *Digital Signal Processing: A Review Journal* 17.1, pp. 371–393. ISSN: 10512004. DOI: 10.1016/j.dsp.2006.04.006.

Stockwell, R. G., L. Mansinha, and R. P. Lowe (1996). "Localization of the complex spectrum: The S transform". In: *IEEE Transactions on Signal Processing* 44.4, pp. 998–1001. ISSN: 1053587X. DOI: 10.1109/78.492555.

633 Tavenard, Romain et al. (2020). "Tslearn, A Machine Learning Toolkit for Time Series Data". In: *Journal of Machine Learning Research*
634 21.118, pp. 1–6. URL: http://jmlr.org/papers/v21/20-091.html.
635 Thangraj, J. S. and J. Pulliam (2021). "Towards real-time assessment of convergence criteria in seismic interferometry: Selective stacking
636 of cross-correlations at the San Emidio geothermal field". In: *Journal of Applied Geophysics* 193, p. 104426. ISSN: 0926-9851. DOI:
637 https://doi.org/10.1016/j.jappgeo.2021.104426. URL: https://www.sciencedirect.com/science/article/pii/
638 S0926985121001737.
639 Thurber, Clifford H. et al. (2014). "Phase-weighted stacking applied to low-frequency Earthquakes". In: *Bulletin of the Seismological*
640 *Society of America* 104.5, pp. 2567–2572. ISSN: 19433573. DOI: 10.1785/0120140077.
641 Tian, Ye and Michael H. Ritzwoller (2017). "Improving ambient noise cross-correlations in the noisy ocean bottom environment of the Juan
642 de Fuca plate". In: *Geophysical Journal International* 210.3, pp. 1787–1805. ISSN: 0956-540X. DOI: 10.1093/gji/ggx281. eprint:
643 https://academic.oup.com/gji/article-pdf/210/3/1787/19194924/ggx281.pdf. URL: https://doi.org/10.1093/
644 gji/ggx281.
645 Toghramadjian, N., L. A. Ermert, and M. A. Denolle (2021). "A tale of urban seismology: ambient seismic noise, machine learning methods,
646 and seismic hazard analysis at the Seattle basin edge". In: SCEC. URL: https://www.scec.org/meetings/2021/am/poster/214.
647 Viens, Loïc and Marine A Denolle (2019). "Long-period ground motions from past and virtual megathrust earthquakes along the Nankai
648 Trough, Japan". In: *Bulletin of the Seismological Society of America* 109.4, pp. 1312–1330.
649 Viens, Loïc and Tomotaka Iwata (2020). "Improving the Retrieval of Offshore-Onshore Correlation Functions With Machine Learning". In:
650 *Journal of Geophysical Research: Solid Earth* 125.8, pp. 1–19. ISSN: 21699356. DOI: 10.1029/2020JB019730.
651 Viens, Loïc et al. (2017). "Retrieving impulse response function amplitudes from the ambient seismic field". In: *Geophysical Journal*
652 *International* 210.1, pp. 210–222. ISSN: 1365246X. DOI: 10.1093/gji/ggx155.
653 Wang, Qing-Yu et al. (2017). "Seasonal crustal seismic velocity changes throughout Japan". In: *Journal of Geophysical Research: Solid*
654 *Earth* 122.10, pp. 7987–8002.
655 Wang, Zhou et al. (2004). "Image Quality Assessment: From Error Visibility to Structural Similarity". In: *IEEE TRANSACTIONS ON*
656 *IMAGE PROCESSING* 13 (4). DOI: 10.1109/TIP.2003.819861. URL: http://www.cns.nyu.edu/~lcv/ssim/..
657 Webb, Spahr C. (1998). "Broadband seismology and noise under the ocean". In: *Reviews of Geophysics* 36.1, pp. 105–142. ISSN: 87551209.
658 DOI: 10.1029/97RG02287.
659 Wu, Gaoxiong et al. (2020). "Shear Velocity Inversion Using Multimodal Dispersion Curves From Ambient Seismic Noise Data of US-
660 Array Transportable Array". In: *Journal of Geophysical Research: Solid Earth* 125.1, pp. 1–14. ISSN: 2169-9313. DOI: 10.1029/
661 2019jb018213.
662 Wu, Sin-Mei et al. (2021). "Imaging the Subsurface Plumbing Complex of Steamboat Geyser and Cistern Spring With Hydrothermal
663 Tremor Migration Using Seismic Interferometry". In: *Journal of Geophysical Research: Solid Earth* 126 (4). ISSN: 21699356. DOI:
664 10.1029/2020JB021128.
665 Yang, Xiaotao and Haiying Gao (2018). "Full-Wave Seismic Tomography in the Northeastern United States: New Insights Into the Uplift
666 Mechanism of the Adirondack Mountains". In: *Geophysical Research Letters* 45.12. ISSN: 19448007. DOI: 10.1029/2018GL078438.

Yang, Xiaotao and Haiying Gao (2020). "Segmentation of the Aleutian-Alaska Subduction Zone Revealed by Full-Wave Ambient Noise Tomography: Implications for the Along-Strike Variation of Volcanism". In: *Journal of Geophysical Research: Solid Earth* 125.11, pp. 1–20. ISSN: 21699356. DOI: 10.1029/2020JB019677.

Yang, Xiaotao et al. (2019). "A Comprehensive Quality Analysis of Empirical Green's Functions at Ocean-Bottom Seismometers in Cascadia". In: *Seismological Research Letters* 90.2A, pp. 744–753. ISSN: 0895-0695. DOI: 10.1785/0220180273. eprint: https://pubs.geoscienceworld.org/srl/article-pdf/90/2A/744/4655388/srl-2018273.1.pdf. URL: https://doi.org/10.1785/0220180273.

Yang, Xiaotao et al. (2022a). *SeisGo: A ready-to-go Python toolbox for seismic data analysis*. Computer Software. Version v0.7.0. DOI: 10.5281/zenodo.5873724. URL: https://doi.org/10.5281/zenodo.5873724.

Yang, Xiaotao et al. (2022b). *StackMaster: A Collection of Methods for Data Stacking*. Computer Software. Version v1.0.3. This replaces StackPy/StackMaster v¡1.0.3. DOI: 10.5281/zenodo.5951013. URL: https://doi.org/10.5281/zenodo.5951013.

Yang, Yingjie and Michael H. Ritzwoller (2008). "Characteristics of ambient seismic noise as a source for surface wave tomography". In: *Geochemistry, Geophysics, Geosystems* 9.2. ISSN: 15252027. DOI: 10.1029/2007GC001814.

Yang, Zhuo, Congcong Yuan, and Marine A Denolle (2022). "Detecting Elevated Pore Pressure due to Wastewater Injection Using Ambient Noise Monitoring". In: *The Seismic Record* 2.1, pp. 38–49.

Yuan, Congcong, Jared Bryan, and Marine A Denolle (2021). "Numerical comparison of time-, frequency-, and wavelet-domain methods for coda wave interferometry". In: *Geophysical Journal International*, pp. 828–846. ISSN: 0956-540X. DOI: 10.1093/gji/ggab140.

Zeng, Xiangfang and Clifford H. Thurber (2016). "A graphics processing unit implementation for time–frequency Phase-Weighted Stacking". In: *Seismological Research Letters* 87.2A, pp. 358–362. ISSN: 0895-0695. DOI: 10.1785/0220150192. eprint: https://pubs.geoscienceworld.org/srl/article-pdf/87/2A/358/2771201/358.pdf. URL: https://doi.org/10.1785/0220150192.