Improving signal-to-noise ratios of ambient noise cross-correlation functions using local attributes

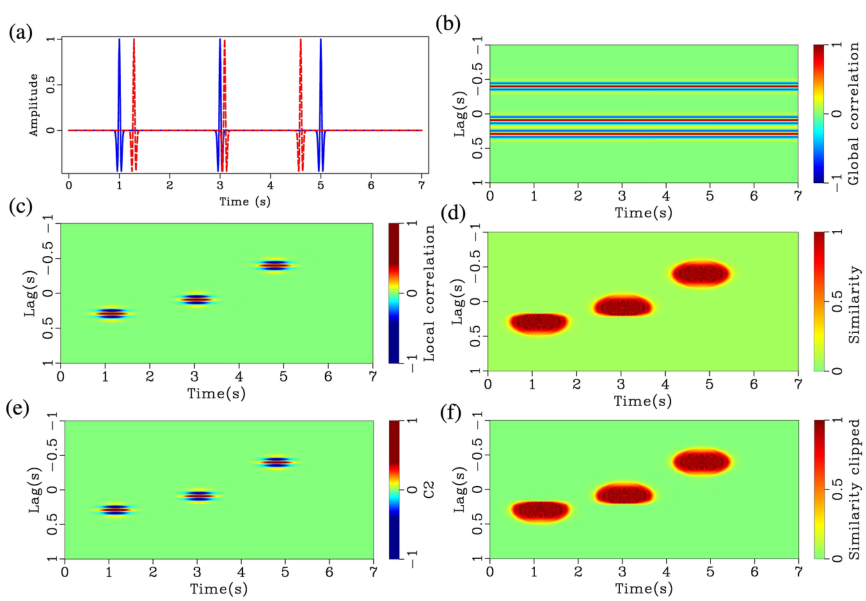
Bin He[1], Hejun Zhu[1], and David E Lumley[1]

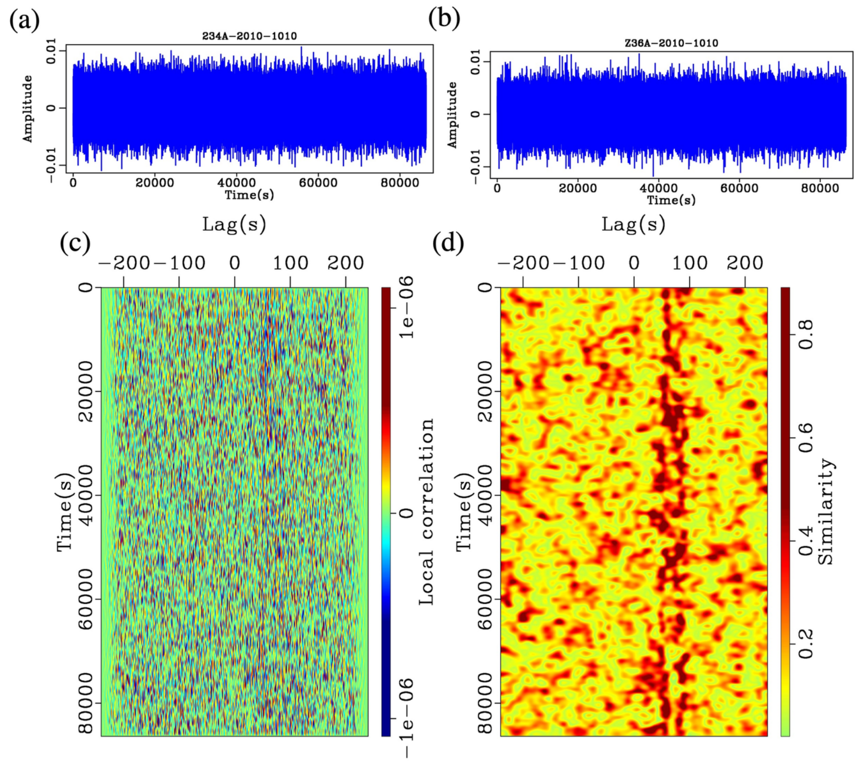
[1]University of Texas at Dallas

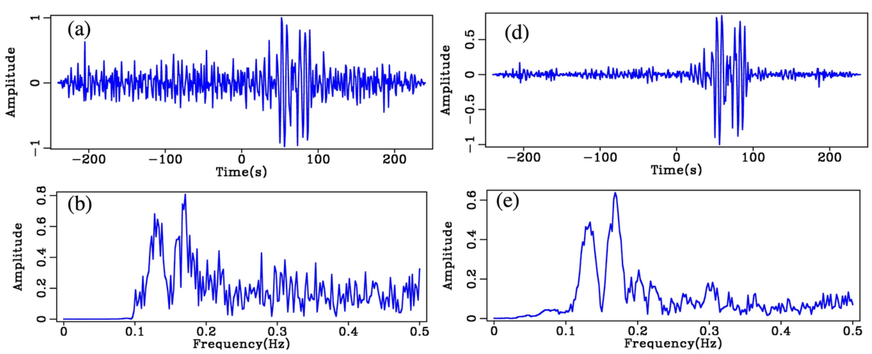
December 1, 2023

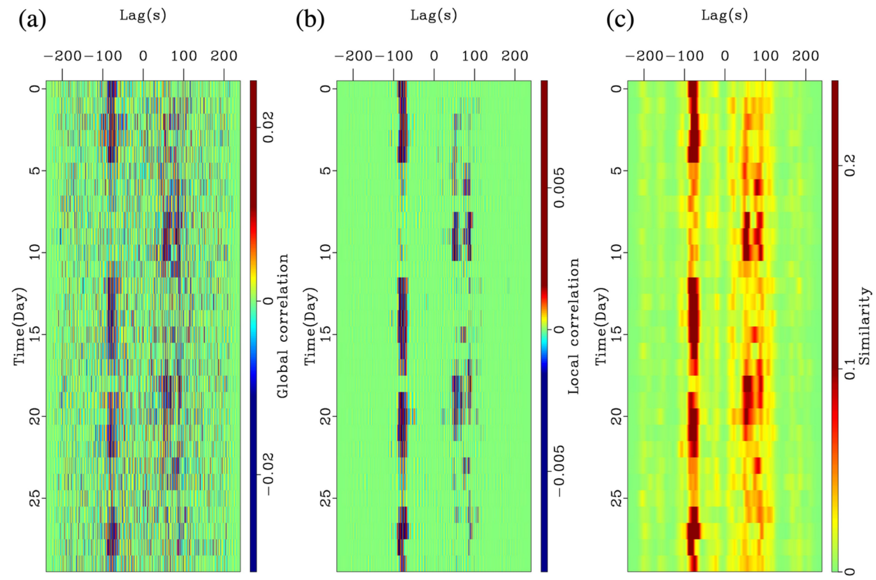
Abstract

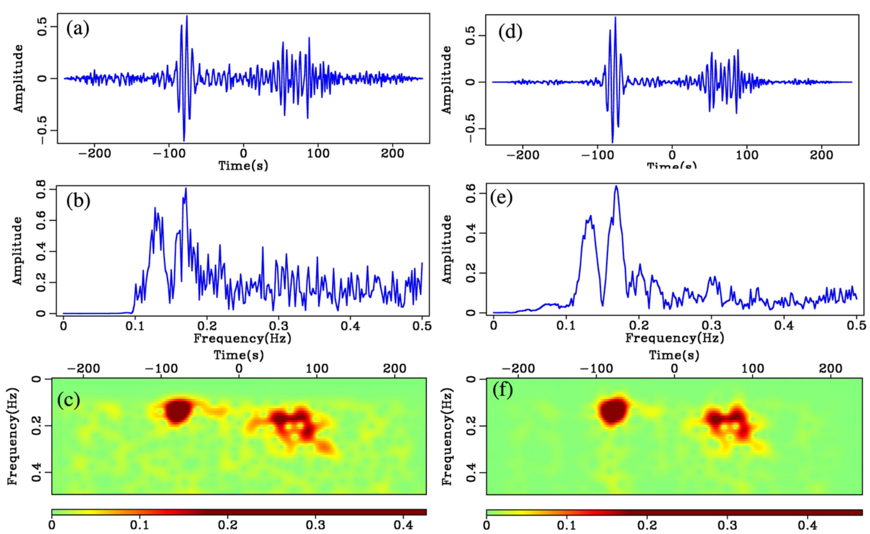
For seismographic stations with short acquisition duration, the signal-to-noise ratios (SNRs) of ambient noise cross-correlation functions (CCFs) are typically low, preventing us from accurately measuring surface wave dispersion curves or waveform characteristics. In addition, with low-quality CCFs, it is difficult to monitor temporal variations of subsurface physical states or extract relatively weak signals such as body waves. In this study, we propose to use local attributes to improve the SNRs of ambient noise CCFs, which allows us to enhance the quality of CCFs for stations with limited acquisition duration. Two local attributes: local cross-correlation and local similarity, are used in this study. The local cross-correlation allows us to extend the dimensionality of daily CCFs with computational costs similar to global cross-correlation. Taking advantage of this extended dimensionality, the local similarity is then used to measure non-stationary similarity between the extended daily CCFs with a reference stacking trace, which enables us to design better stacking weights to enhance coherent features and attenuate incoherent background noises. Ambient noise recorded by several broadband stations from the USArray in North Texas and Oklahoma, the Superior Province Rifting EarthScope Experiment in Minnesota and Wisconsin and a high-frequency nodal array deployed in the San Bernardino basin are used to demonstrate the performance of the proposed approach for improving the SNR of CCFs.

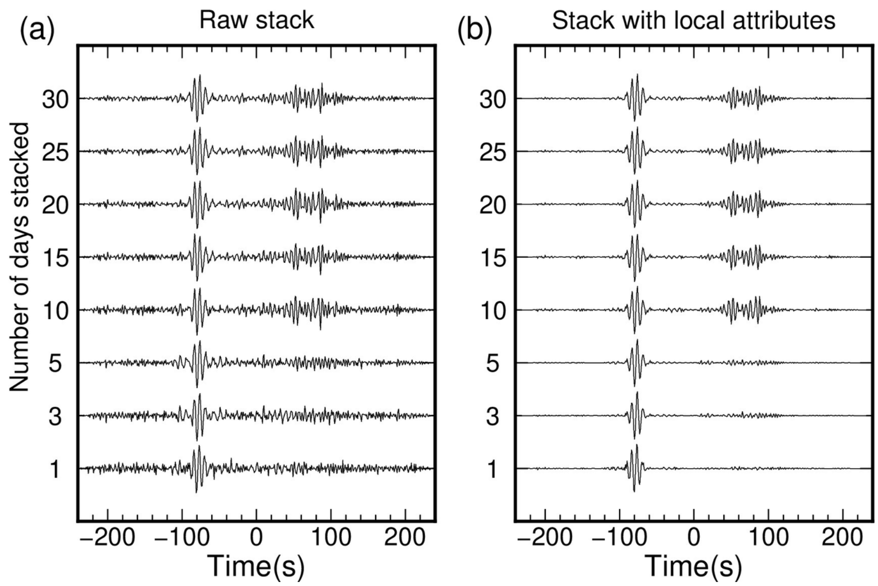



(a)
Z34A−2010−1010
Amplitude
Time(s)
(b)
Z36A−2010−1010
Amplitude
Time(s)
(c)
Lag(s)
Time(s)
Local correlation
(d)
Lag(s)
Time(s)
Similarity

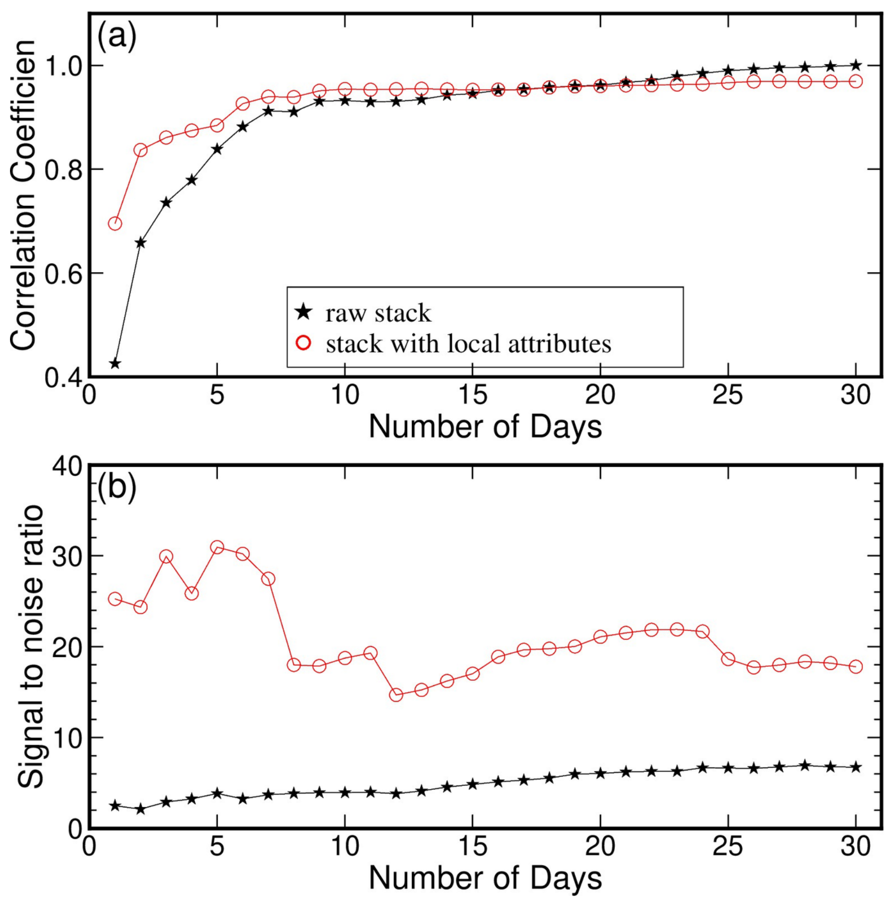

(a)
Amplitude
Time(s)
(b)
Amplitude
Frequency(Hz)
(c)
Time(s)
Frequency(Hz)
(d)
Amplitude
Time(s)
(e)
Amplitude
Frequency(Hz)
(f)
Time(s)
Frequency(Hz)


(a)
(b)
(c)
Lag(s)
Time(Day)
Global correlation
Local correlation
Similarity

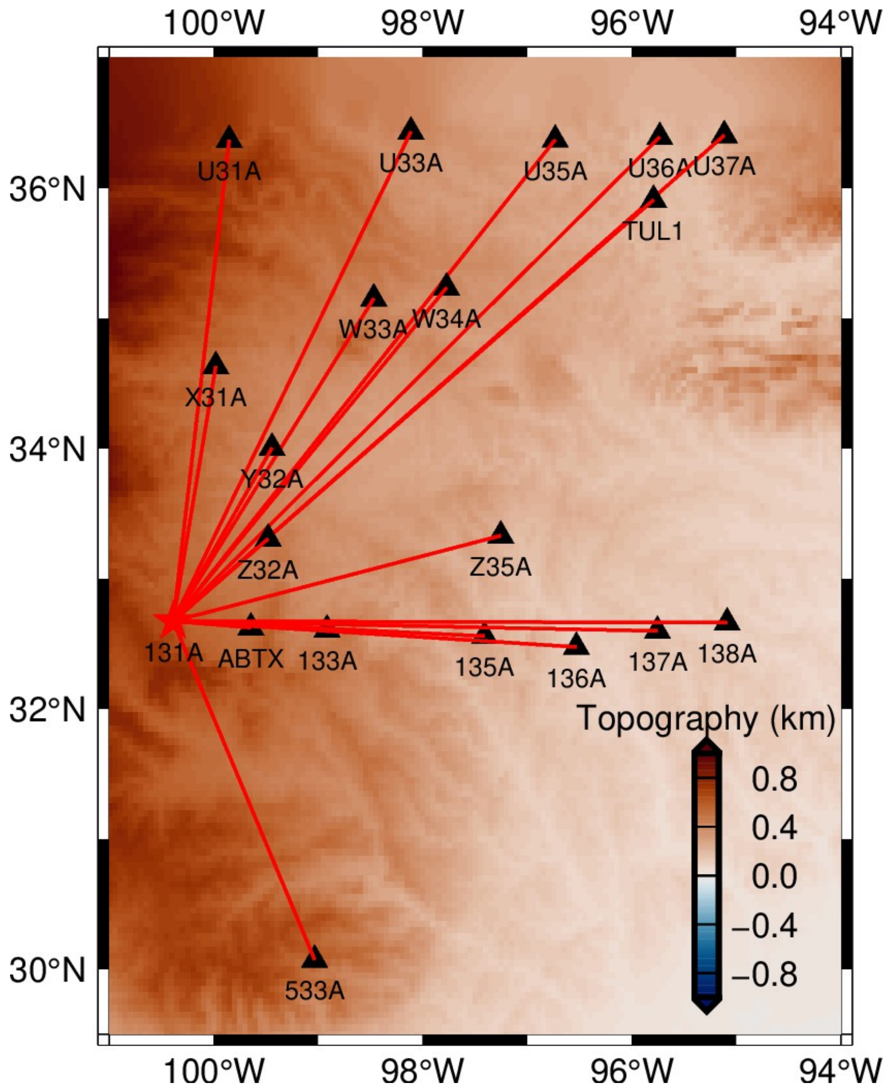

(a)
(b)
(c)
(d)
(e)
(f)
Amplitude
Time(s)
Frequency(Hz)

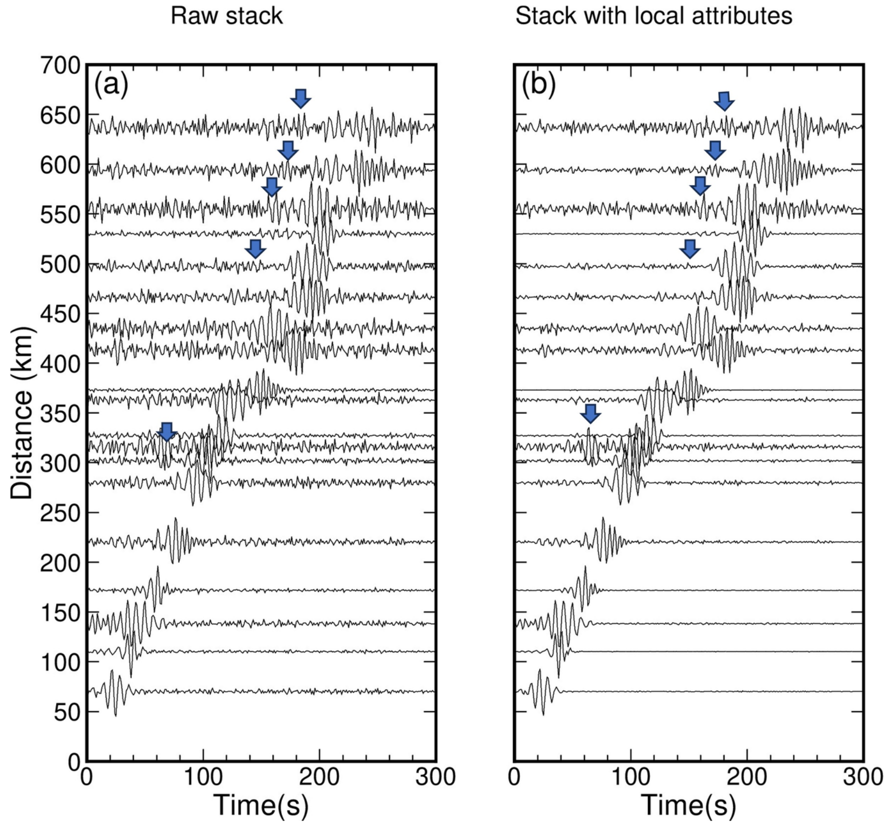


(a)
Raw stack
(b)
Stack with local attributes
Number of days stacked
30
25
20
15
10
5
3
1
−200 −100 0 100 200
Time(s)

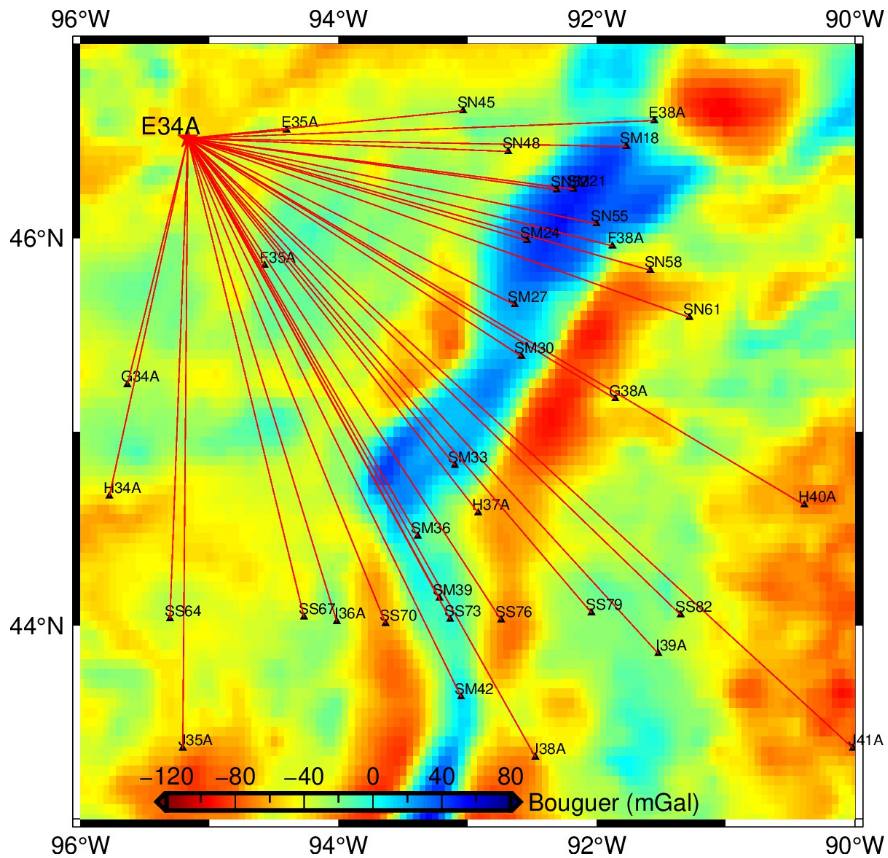


(a)
Correlation Coefficien
1.0
0.8
0.6
0.4
raw stack
stack with local attributes
0 5 10 15 20 25 30
Number of Days
(b)
Signal to noise ratio
40
30
20
10
0
Number of Days

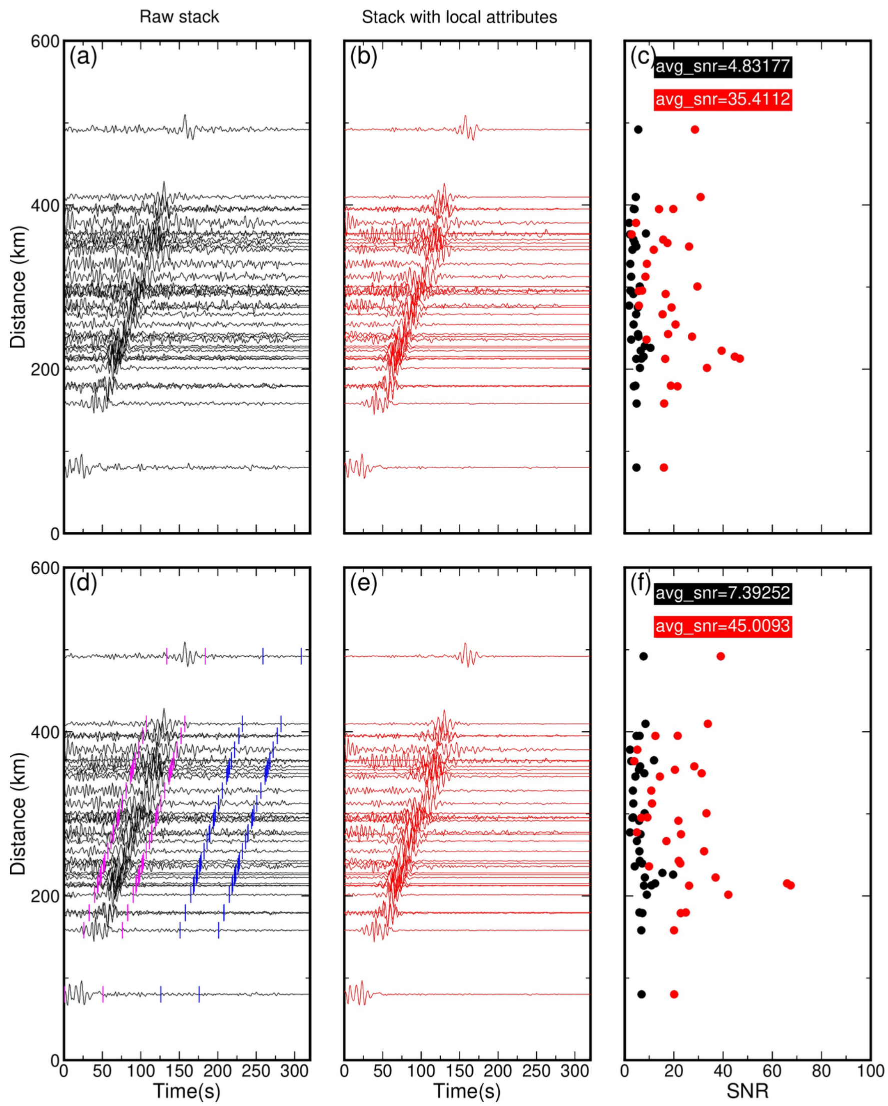

100°W
98°W
96°W
94°W
36°N
34°N
32°N
30°N
U31A
U33A
U35A
U36A
U37A
TUL1
W33A
W34A
X31A
Y32A
Z32A
Z35A
131A
ABTX
133A
135A
136A
137A
138A
533A
Topography (km)
0.8
0.4
0.0
−0.4
−0.8



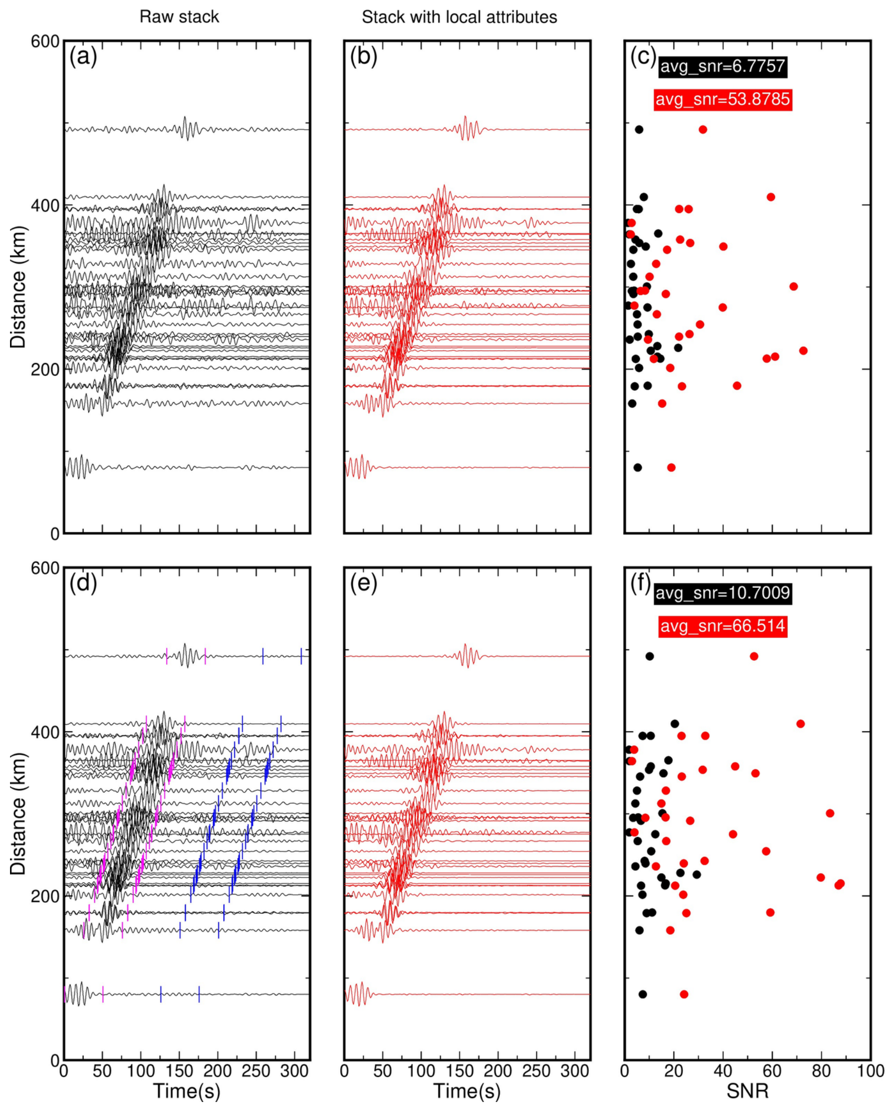
Raw stack
Stack with local attributes
(a)
(b)
Distance (km)
Time(s)
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
0
100
200
300

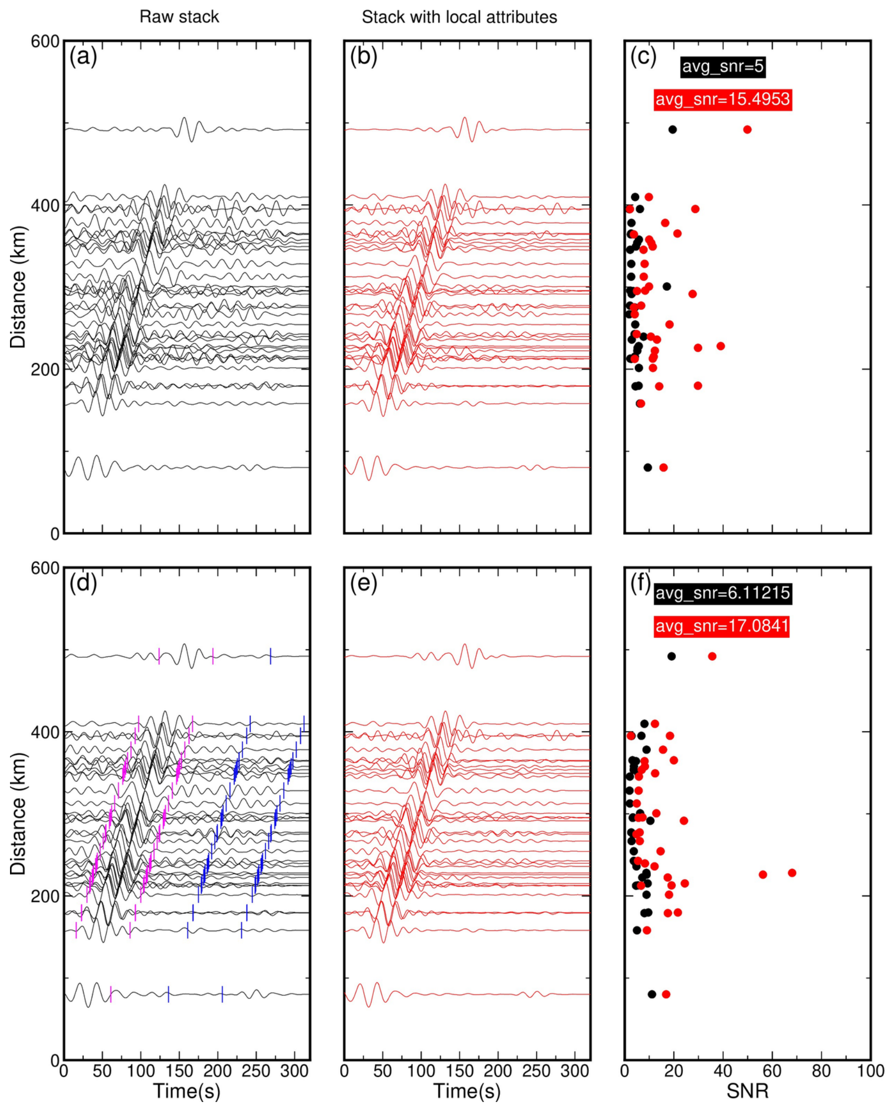


96°W
94°W
92°W
90°W
46°N
44°N
E34A
E35A
SN45
E38A
SN48
SM18
SM21
SN55
SM24
F38A
SN58
F35A
SM27
SN61
SM30
G34A
G38A
SM33
H34A
H40A
H37A
SM36
SM39
SS64
SS67
I36A
SS70
SS73
SS76
SS79
SS82
I39A
SM42
I35A
I38A
I41A
−120
−80
−40
0
40
80
Bouguer (mGal)

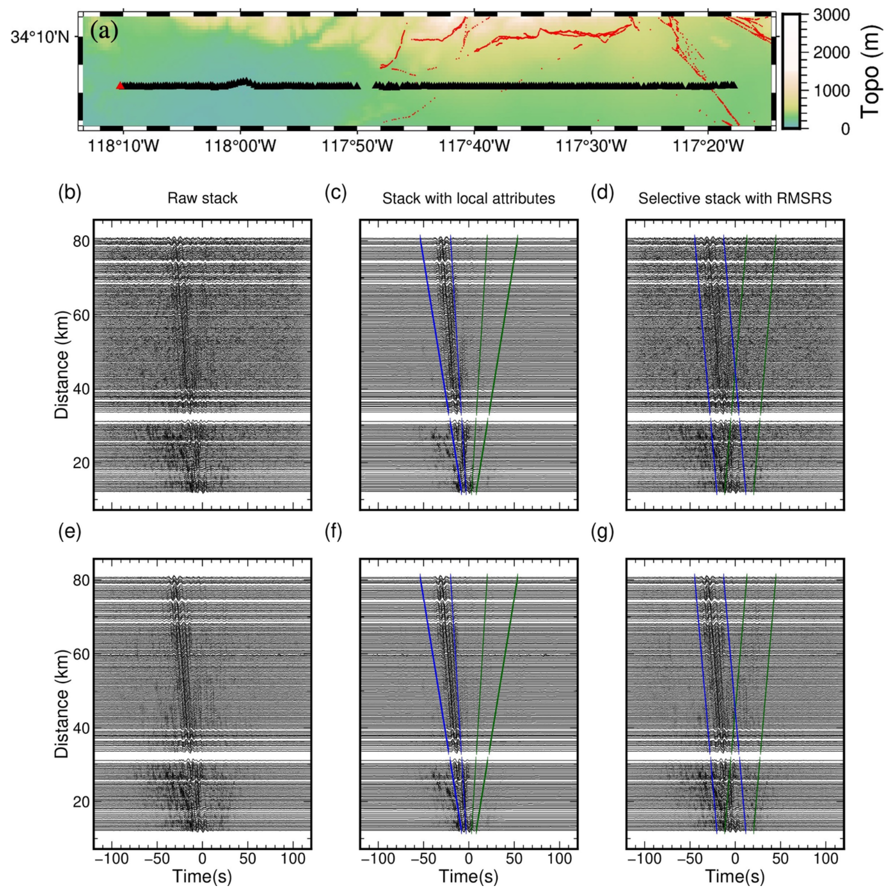


Raw stack
Stack with local attributes
(a)
(b)
(c)
avg_snr=4.83177
avg_snr=35.4112
(d)
(e)
(f)
avg_snr=7.39252
avg_snr=45.0093
Distance (km)
Time(s)
SNR
0
200
400
600
0 50 100 150 200 250 300
0 20 40 60 80 100

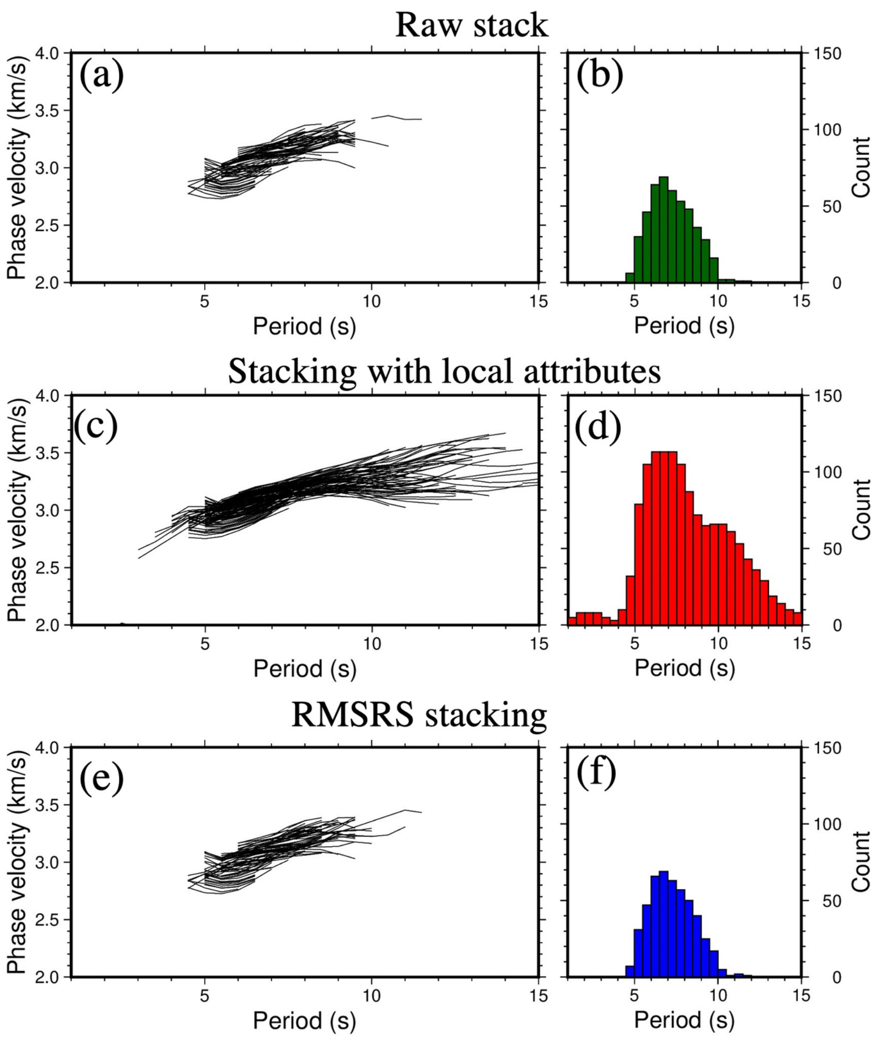

Raw stack
Stack with local attributes
(a)
(b)
(c)
avg_snr=6.7757
avg_snr=53.8785
(d)
(e)
(f)
avg_snr=10.7009
avg_snr=66.514
Distance (km)
Time(s)
SNR
0
200
400
600
0
50
100
150
200
250
300
0
20
40
60
80
100



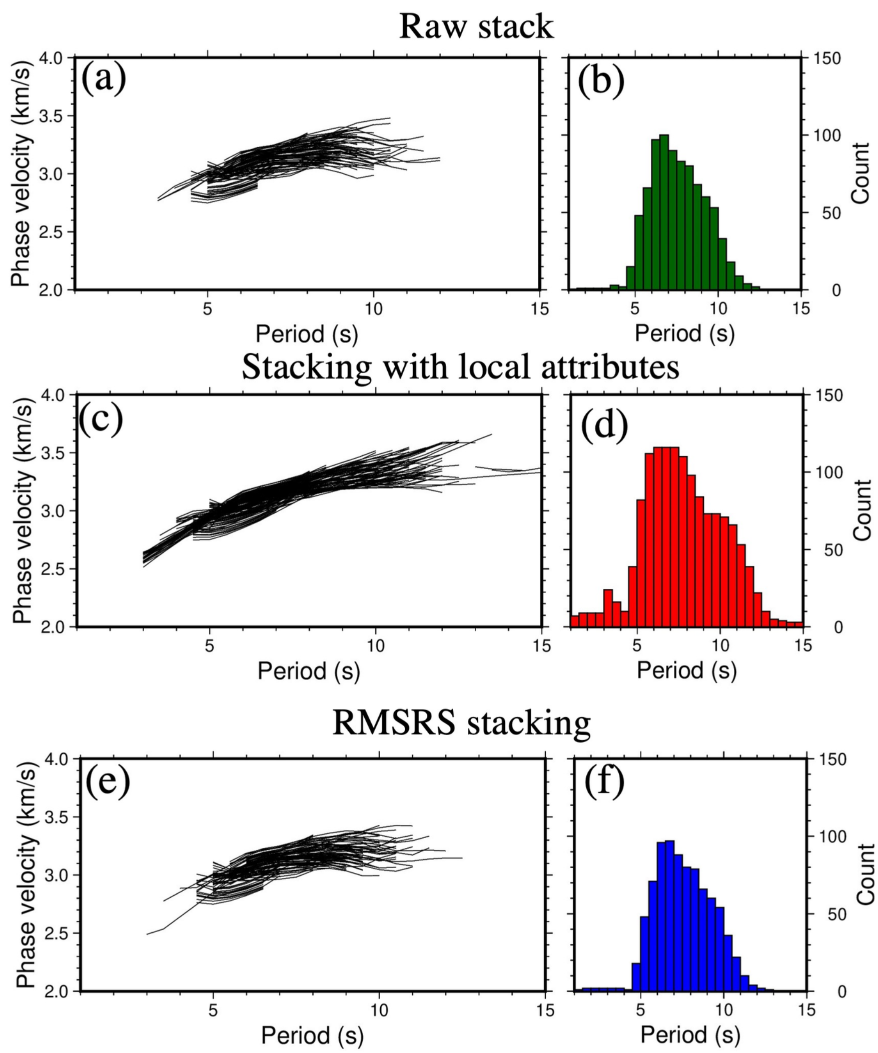
Raw stack
Stack with local attributes
(a)
(b)
(c)
avg_snr=5
avg_snr=15.4953
(d)
(e)
(f)
avg_snr=6.11215
avg_snr=17.0841
Distance (km)
Time(s)
SNR

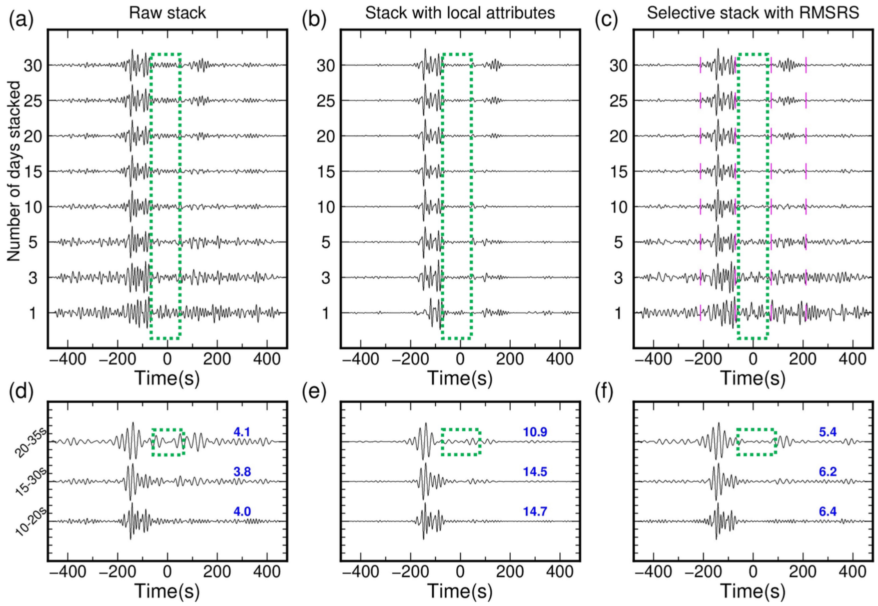


(a)
34°10'N
118°10'W
118°00'W
117°50'W
117°40'W
117°30'W
117°20'W
Topo (m)
3000
2000
1000
0
(b)
Raw stack
(c)
Stack with local attributes
(d)
Selective stack with RMSRS
(e)
(f)
(g)
Distance (km)
80
60
40
20
Time(s)
−100
−50
0
50
100



Raw stack
(a)
(b)
Phase velocity (km/s)
Period (s)
Count
Stacking with local attributes
(c)
(d)
RMSRS stacking
(e)
(f)


Raw stack
(a)
(b)
Phase velocity (km/s)
Period (s)
Count
Stacking with local attributes
(c)
(d)
Phase velocity (km/s)
Period (s)
Count
RMSRS stacking
(e)
(f)
Phase velocity (km/s)
Period (s)
Count



(a)
Raw stack
(b)
Stack with local attributes
(c)
Selective stack with RMSRS
Number of days stacked
30
25
20
15
10
5
3
1
−400 −200 0 200 400
Time(s)
(d)
(e)
(f)
20-35s
15-30s
10-20s
4.1
3.8
4.0
10.9
14.5
14.7
5.4
6.2
6.4

Improving signal-to-noise ratios of ambient noise cross-correlation functions using local attributes

Bin He[1], **Hejun Zhu**[1,2], **David Lumley**[1,2]

[1]Department of Geosciences, The University of Texas at Dallas, Richardson, TX 75080, USA
[2]Department of Physics, the University of Texas at Dallas, Richardson, TX 75080, USA

Key Points:

- Signal-to-noise ratio of ambient noise cross-correlation functions can be improved by using local cross-correlation and local similarity.
- The local cross-correlation function is efficiently implemented by solving the heat equation.
- Applications to broadband and high-frequency nodal arrays validate the effectiveness of the proposed method.

Corresponding author: Hejun Zhu, hejun.zhu@utdallas.edu

Abstract

For seismographic stations with short acquisition duration, the signal-to-noise ratios (SNRs) of ambient noise cross-correlation functions (CCFs) are typically low, preventing us from accurately measuring surface wave dispersion curves or waveform characteristics. In addition, with low-quality CCFs, it is difficult to monitor temporal variations of subsurface physical states or extract relatively weak signals such as body waves. In this study, we propose to use local attributes to improve the SNRs of ambient noise CCFs, which allows us to enhance the quality of CCFs for stations with limited acquisition duration. Two local attributes: local cross-correlation and local similarity, are used in this study. The local cross-correlation allows us to extend the dimensionality of daily CCFs with computational costs similar to global cross-correlation. Taking advantage of this extended dimensionality, the local similarity is then used to measure non-stationary similarity between the extended daily CCFs with a reference stacking trace, which enables us to design better stacking weights to enhance coherent features and attenuate incoherent background noises. Ambient noise recorded by several broadband stations from the USArray in North Texas and Oklahoma, the Superior Province Rifting EarthScope Experiment in Minnesota and Wisconsin and a high-frequency nodal array deployed in the San Bernardino basin are used to demonstrate the performance of the proposed approach for improving the SNR of CCFs.

Plain Language Summary

Seismic ambient noise has been widely used for imaging and monitoring subsurface structures by using cross-correlation functions (CCFs) of continuous recordings between station pairs. Typically, we have to stack a lot of CCFs to enhance signals (e.g., surface and body waves) of CCFs. However, for temporal monitoring purposes or those deployed arrays with short acquisition duration, it is impossible to have a lot of stacking. The lack of enough stacking could result in low signal-to-noise ratios (SNRs) of the CCFs. We proposed a new approach to improve the SNRs of CCFs by using two local attributes: local cross-correlation and local similarity. We first extend the CCFs to a higher dimension by using local cross-correlation, and then the local similarity is used to define a better weighting factor for the final stacking. We use several field data examples to prove the effectiveness of the proposed approach.

1 Introduction

With the assumption of homogeneously distributed noise sources, both theoretical and experimental studies have demonstrated that empirical Green's functions can be retrieved by cross-correlating continuous ambient noise records between two seismographic stations (Aki, 1957; Claerbout, 1968; Buckingham et al., 1992; Lobkis & Weaver, 2001; Weaver & Lobkis, 2001; Shapiro & Campillo, 2004; Wapenaar, 2004; Nakata et al., 2019). Both surface (Campillo & Paul, 2003; Shapiro & Campillo, 2004; Sabra et al., 2005) and body wave signals (Draganov et al., 2009; Zhan et al., 2010; Poli et al., 2012; Lin et al., 2013; Nakata et al., 2015) have been successively extracted from ambient noise recordings. With empirical Green's functions between each pair of stations in a seismic array, we are able to measure surface wave phase or group dispersion curves, and then perform surface wave tomography to estimate seismic properties in the subsurface. Compared with earthquake tomography, ambient noise tomography allows us to image tectonically inactive regions and achieve better cross-path coverages. To date, ambient noise cross-correlation functions (CCFs) have been successively used to investigate velocity as well as anisotropic structures within the crust and uppermost mantle (Shapiro et al., 2005; Y. Yang et al., 2007; Lin et al., 2008; Moschetti et al., 2010; Yao et al., 2010; Huang et al., 2010; Lin et al., 2011; Shen & Ritzwoller, 2016). Recently, ambient noise CCFs are combined with waveform inversion to better constrain subsurface velocity structures (Gao & Shen, 2014; M. Chen et al., 2014; Lee et al., 2014; Y. Liu et al., 2017; Sager et al., 2018; K. Wang et al., 2018; Zhu, 2018; Sager et al., 2020; Fan et al., 2022; Maguire et al., 2022). Taking advantage of continuous records, ambient noise CCFs can also be utilized to monitor temporal evolution of physical states in the subsurface (Brenguier et al., 2008; Nakata & Snieder, 2011; Hadziioannou et al., 2011; Mainsant et al., 2012; De Plaen et al., 2016; Q.-Y. Wang & Yao, 2020; Le Breton et al., 2021; Mao et al., 2022).

Previous studies have demonstrated that we are able to improve the signal-to-noise ratios (SNRs) of ambient noise CCFs as well as reducing the effect of source directionality by increasing the stacking duration of signals (Bensen et al., 2007). To obtain high-quality CCFs, typically we have to use long continuous records and sometimes have to stack data with several years of acquisition. However, for some temporary experiments e.g., Brenguier et al. (2008); Issa et al. (2017); G. Liu et al. (2018); Dougherty et al. (2019); G. Chen et al. (2023); Wu et al. (2023), the SNRs of ambient noise CCFs are typically low, preventing us to measure robust surface wave dispersion curves or waveform characteristics. For these cases, we have to design effective approaches to improve the SNRs of ambient noise CCFs. In addition, the SNRs are crucial for extracting weak coherent signals, such as body waves (Snieder, 2004; Nakata et al., 2015, 2016). There are several studies towards improving the SNR of ambient noise CCFs. For instance, Baig et al. (2009) designed better-stacking weights in the time-frequency domain with discrete orthogonal S transform, which allowed them to measure robust Rayleigh and Love arrivals for stations with long offsets. G. Li et al. (2018) further compared the performances of time and frequency domain inverse S transforms for this stacking procedure. Schimmel et al. (2011) used instantaneous phase coherence to avoid strong amplitude arrivals, such as earthquake signals, and enhance coherent features in noise records. Seats et al. (2012) and Clarke et al. (2011) utilized overlapped moving windows (Welch's method) to improve the convergence of CCFs towards stable Green's functions. In addition, wavelet and curvelet transforms have also been applied to denoise CCFs by Stehly et al. (2011) and Mao et al. (2022). Furthermore, Weaver and Yoritomo (2018) proposed several schemes to choose optimal weights for stacking so that the effective incident intensity distribution is closer to isotropic. Xie et al. (2020) used the root-mean-square ratio to remove those CCFs with low SNR. A systematic evaluation and comparison of the performance of several stacking methods is discussed by X. Yang et al. (2023).

In this study, we propose to use local attributes (Rickett & Lumley, 2001; Hale, 2006; Fomel, 2007a) to denoise CCFs when we have short acquisition durations. In comparison with global attributes, the local attributes enable us to extract non-stationary characteristics in seismic data. There are a variety of local attributes, including local similarity (Fomel, 2007a), local cross-correlation (Hale, 2006), local frequency (Fomel, 2007a) and local skewness (Fomel & van der Baan, 2014), etc. In this study, we use local cross-correlation (Hale, 2006) to extend the dimensionality of daily CCFs, and then use stacking weights measured by local similarity to improve the stacking quality of CCFs. The local cross-correlation can be considered as a natural extension of the Welch's method (Seats et al., 2012) but with longer overlap windows and much lower computational costs. Its overall computational cost is similar to the conventional global cross-correlation. The local similarity has been used to measure non-stationary similarity between time-lapse images (Fomel & Jin, 2009), improve stacking quality of normal moveout data (G. Liu et al., 2009) as well as angle-domain common-imaging-gathers (Lin et al., 2011). All these studies demonstrate the capability of using local similarity to improve the stacking quality.

We first review local cross-correlation and local similarity. Next, they are used to improve the SNRs of ambient noise CCFs with one-day and one-month acquisition durations. Continuous noise records from several broadband USArray stations in North Texas and Oklahoma, the Superior Province Rifting EarthScope Experiment in Minnesota and a high-frequency nodal array in the San Bernardino basin are used to demonstrate the performance of the proposed approach.

2 Method

2.1 Local cross-correlation

The global cross-correlation function $c(\tau)$ between two signals $f(t)$ and $g(t)$ can be defined (Nakata & Snieder, 2011; Harris et al., 2020) as follows:

$$c(\tau) = \int_{-\infty}^{\infty} f(t+\tau)g(t)dt \quad . \tag{1}$$

The result of the global cross-correlation is a function of time lag τ, which cannot capture non-stationary time shifts between two input signals and therefore makes it difficult to separate signals and noises. In order to fulfill this goal, one way is to perform time-windowed cross-correlation, which has much higher computational costs in comparison to Equation 1. In addition, there might be leakage and edge problems due to the selected window functions. To solve these problems, Hale (2006) designed an efficient algorithm to compute the local CCF and measure time-varying correlations between two signals. It was first used to measure non-stationary warpings between time-lapse migration images. Later, the local cross-correlation was applied to measure non-stationary travel time differences between two seismograms, which can be utilized to constrain subsurface velocity structures through full-waveform inversion (Díaz & Sava, 2015).

The idea of local cross-correlation is similar to the moving window cross-correlation for capturing transient time shifts, but with much higher computational efficiency. In addition, it helps us to separate signal and noise in a higher dimension. Taking advantage of several properties of Gaussian windows, such as the product of two Gaussian windows is still a Gaussian function, the computational cost of local cross-correlation can be reduced as similar to that of global cross-correlation. For instance, the Gaussian-windowed versions of signals $f(t)$ and $g(t)$ can be represented as

$$\hat{f}(t, t_0) = f(t)\omega(t_0 - t) \quad ,$$
$$\hat{g}(t, t_0) = g(t)\omega(t_0 - t) \quad . \tag{2}$$

where $\omega(t_0-t)$ is the Gaussian window function located at t_0, i.e., $\omega(t_0-t) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(t_0-t)^2/2\sigma^2}$. Here σ is the standard deviation, which controls the width of the Gaussian window. With these windowed signals, we are able to compute the cross-correlation function at each local time step t_0. Without optimization, the computational cost of this moving window cross-correlation is $O(N_l N_w N_s)$. Here, N_l, N_w and N_s are the numbers of samples for time lags, Gaussian window, and input signals, respectively.

The windowed cross-correlation can be written as

$$\begin{aligned} c(t_0, \tau) &= \int_{-\infty}^{\infty} f(t+\tau, t_0+\tau)g(t, t_0)dt \\ &= \int_{-\infty}^{\infty} f(t+\tau)\omega(t_0-t)g(t)\omega(t_0-t)dt \\ &= \int_{-\infty}^{\infty} f(t+\tau)g(t)W(t, t_0)dt \quad , \end{aligned} \tag{3}$$

where

$$W(t, t_0) = \omega(t_0 - t)\omega(t_0 - t) = \frac{1}{2\pi\sigma^2}e^{-(t_0-t)^2/\sigma^2} \quad . \tag{4}$$

$W(t, t_0)$ is again a Gaussian function with a smaller width $\sigma/\sqrt{2}$ compared to $w(t_0 - t)$. Based on this property, we can first compute the pre-stacked cross-correlation between two signals, i.e., $h(t, \tau) = f(t+\tau)g(t)$. It is then filtered by a Gaussian function $W(t, t_0)$ as

$$c(t_0, \tau) = \int_{-\infty}^{\infty} h(t, \tau)W(t, t_0)dt \quad . \tag{5}$$

Combining with some efficient recursive Gaussian filter algorithms (Deriche, 1993; Alvarez & Mazorra, 1994; Van Vliet et al., 1998), the computational cost of implementing Equation 3 can be similar as the global cross-correlation shown in Equation 1, i.e., $O(N_l N_s)$. This makes it attractive for processing large-scale datasets, such as continuous ambient noise records.

2.2 Recursive Gaussian convolution

How to efficiently compute the convolution of a signal or an image with a Gaussian function is a common problem in data processing. There are several ways to perform this filtering procedure. For instance, by taking advantage of the Fourier transform, Gaussian convolution can be computed efficiently as elemental-wise multiplication in the frequency domain. Another efficient way to implement the Gaussian filter is to consider the result as the solution of the following heat equation:

$$\begin{aligned} \frac{\partial u}{\partial y} - \frac{\partial^2 u}{\partial x^2} &= 0 \quad , \\ u(x,0) &= f(x) \quad . \end{aligned} \tag{6}$$

Here, if we let $y = \sigma^2/2$, then the final solution of the above heat equations $u(x, y)$ is equivalent to the convolution of signal $f(x)$ with a Gaussian function $\omega(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$. This heat equation can be solved by using finite-difference schemes. Alvarez and Mazorra (1994) used this property to derive an efficient procedure to compute the Gaussian convolution. For instance, they discretized the simulation domain with

$$u_i^j = u(i\Delta x, j\Delta y) \quad , \tag{7}$$

Here Δx and Δy are the grid spacings along the spatial and time directions, i and j are the associate indices. Then, the finite-difference solution of the above heat equation can be written as

$$\frac{u_i^{j+1} - u_i^j}{\Delta t} - \frac{u_{i+1}^{j+1} + u_{i-1}^{j+1} - 2u_i^{j+1}}{\Delta x^2} = 0 \quad . \tag{8}$$

Let $\lambda = \Delta y/\Delta x^2$, then

$$(1 + 2\lambda)u_i^{j+1} - \lambda u_{i-1}^{j+1} - \lambda u_{i+1}^{j+1} = u_i^j \quad . \tag{9}$$

We can solve this implicit finite-difference equation using the following filter

$$H(z) = \frac{\nu}{\lambda} \frac{1}{(1 - \nu z^{-1})(1 - \nu z)} \quad , \tag{10}$$

with

$$\nu = \frac{1 + 2\lambda - \sqrt{(1 + 4\lambda)}}{2\lambda} \quad . \tag{11}$$

The above filtering procedure can be efficiently solved by using the following three steps, marching from u_i^j to u_i^{j+1} as

$$\begin{aligned} u_i^{j'} &= u_i^j + \nu u_{i-1}^{j'} \quad , \\ u_i^{j''} &= u_i^{j'} + \nu u_{i+1}^{j''} \quad , \\ u_i^{j+1} &= \frac{\nu}{\lambda} u_i^{j''} \quad . \end{aligned} \tag{12}$$

The first and second steps in the above procedure are the applications of causal and acausal filters, respectively. The third step is just a simple multiplication. Thanks to the unconditional stability of this implicit solution (Alvarez & Mazorra, 1994), only a few iterations (four iterations are used in this study to propagate along the time dimension y) enable us to converge to the solution of the Gaussian filter. Therefore, the time grid can be defined by $\Delta y = \sigma^2/(2N_y)$, where N_y is the total number of iterations. Combining this recursive Gaussian filter with Equation 5, we are able to efficiently compute the local cross-correlation function. For example, for a certain τ, the source term for Equation 6 is the pre-stacked cross-correlation $h(t,\tau)$. Then solution of the corresponding heat equation is the non-stationary CCFs $u(t,\tau)$ at this particular τ.

Figure 1 shows a simple example to illustrate the advantages of local cross-correlation over global cross-correlation. There are non-stationary time shifts for three events at 1, 3 and 5 s between two input signals. The global cross-correlation is shown in Figure 1b. Although it captures the correct time shifts around -0.4 s, 0.1 s and 0.3 s, it does not carry any useful information about the locations of these time shifts. The local cross-correlation is presented in Figure 1c, with three energy spots located at the right locations with correct time lags. This local cross-correlation function enables us to extend the dimensionality of the input signals as well as the cross-correlation function, which can be utilized to design a better strategy to improve stacking quality as discussed below.

2.3 Local similarity

Local similarity is a local attribute that allows us to measure non-stationary similarity between two time series or images (Fomel & Jin, 2009; G. Liu et al., 2009, 2011). First, the global similarity between two signals $f(t)$ and $g(t)$ can be defined as follows

$$s = \frac{\int f(t)g(t)dt}{\sqrt{\int f^2(t)dt}\sqrt{\int g^2(t)dt}} \quad . \tag{13}$$

In comparison with the global cross-correlation in Equation 1, the global similarity can be considered as a zero-lag cross-correlation function normalized by the energies of each signal. Therefore, it is insensitive to the absolute amplitudes of input signals. Its magnitude ranges from -1 to 1. Similar to local cross-correlation, local similarity was designed to measure non-stationary similarity between time-lapse images (Rickett & Lumley, 2001; Fomel & Jin, 2009). Once taking the square of Equation 13 on both sides, we can write its discretized version as

$$s^2 = \frac{(\mathbf{f}^T\mathbf{g})(\mathbf{g}^T\mathbf{f})}{(\mathbf{f}^T\mathbf{f})(\mathbf{g}^T\mathbf{g})} \quad . \tag{14}$$

Fomel (2007a) considered the above equation as the solution to the following two least-squares problems:

$$\begin{aligned} \mathbf{s}_1 &= \arg \min_{\mathbf{s}_1} |\mathbf{F}\mathbf{s}_1 - \mathbf{g}| \quad , \\ \mathbf{s}_2 &= \arg \min_{\mathbf{s}_2} |\mathbf{G}\mathbf{s}_2 - \mathbf{f}| \quad . \end{aligned} \tag{15}$$

Here, $\mathbf{F}$ and $\mathbf{G}$ are the diagonal operators constructed from the elements of input vectors f and g, respectively. Shaping regularization (Fomel, 2007b) and a conjugate gradient method can be used to solve the above least-squares problems. Then, the local similarity can be computed as elemental-wise multiplication between $\mathbf{s}_1$ and $\mathbf{s}_2$ as

$$\mathbf{s} = \sqrt{\mathbf{s}_1\mathbf{s}_2} \quad . \tag{16}$$

As pointed out in Fomel (2007a), the normalized local correlation can be considered as the first iteration from local similarity. Compared to a scalar value for the global similarity in Equation 13, the local similarity is a vector with the same length as the input signals. It can also be utilized to measure the similarity between two 2D or 3D images with high efficiency.

2.4 Stacking strategy with local similarity

From local cross-correlation, we obtain a function $c(t, \tau)$. The simplest way to calculate the CCF $c_1(\tau)$ is to use the following linear stacking

$$c_1(\tau) = \frac{1}{N} \sum_t c(t, \tau) \quad . \tag{17}$$

If we use the above expression to compute noise CCF, its result is similar to the direct implementation of Equation 1. In this study, we choose $c_1(\tau)$ as a reference trace and then compute the local similarity $s(t, \tau)$ between $c_1(\tau)$ and $c(t, \tau)$. Once with the local similarity, we can use the following strategy to improve the stacking quality

$$c_2(\tau) = \frac{1}{N} \sum_t \alpha(t, \tau) c(t, \tau) \quad , \tag{18}$$

with

$$\alpha(t, \tau) = \begin{cases} s(t, \tau) - \alpha_0 \quad , & \text{if } \alpha(t, \tau) > \alpha_0 \\ 0 \quad , & \text{if } \alpha(t, \tau) \leq \alpha_0 \end{cases} \tag{19}$$

Here α_0 is a pre-defined parameter to decide the stacking level. This stacking procedure has been applied to enhance coherent signals for normal moveout corrected data (G. Liu et al., 2009) and angle-domain common-image-gathers (G. Liu et al., 2011). In this study, we use it to improve the stacking quality of noise CCFs. In practice, $\alpha(t,\tau)$ is a soft threshold solved in the seislet domain (Daubechies et al., 2004; G. Liu et al., 2009). Compared with the similarity shown in Figure 1d, the clipped similarity ($\alpha(t,\tau)$ in Eq. 19) in Figure 1f shows zero background values away from signal regions, while keeping almost 1.0 weights around the signal regions. This helps to suppress noises away from the signal regions. Because this synthetic test does not involve noises, the weighted prestack local correlations ($c_2(\tau)$ in Eq. 18) shown in Figure 1e are basically identical to the prestack local correlations shown in Figure 1c. To validate the effectiveness of the proposed method to improve the SNR, we give its definition used in this study as

$$SNR = \sqrt{\frac{\frac{1}{N_s}\int_{s_1}^{s_2} f^2(t)dt}{\frac{1}{N_n}\int_{n_1}^{n_2} f^2(t)dt}} \quad . \tag{20}$$

Where $[s_1,s_2]$ and $[n_1,n_2]$ represent signal and noise windows, N_s and N_n denote the number of time samples for signal and noise, respectively.

2.5 Workflow

In summary, the proposed method based on two local attributes can be implemented through the following six steps:

1. Download and preprocess continuous ambient noise data for each station, and cut them into daily recordings.
2. Calculate the pre-stacked cross-correlation function $h(t,\tau) = f(t+\tau)g(t)$ as shown in Equation 3 between daily recordings from station pairs.
3. Stack all daily global stacking CCFs and extend it along the time axis to obtain two-dimensional data as a reference.
4. Solve the heat equation in Equation 6 to obtain a daily local CCF. Note here, the pre-stacked CCF is the source term $f(x)$ in Equation 6, and the local CCF is also two-dimensional data.
5. Calculate the daily local similarity between the reference CCF (step 3) and the daily local CCF(step 4).
6. Stack all daily local CCFs weighted by the local similarity according to Equation 18 for each station pair.

3 Numerical examples

We download continuous ambient noise data, which are then processed by using the NoisePy package (Jiang & Denolle, 2020) following standard procedures from Bensen et al. (2007). It includes deconvolution of instrument responses, bandpass filter (2–10 s), time domain normalization and spectral whitening. The preprocessed data are used to retrieve the CCFs and demonstrate the advantages of the proposed method.

3.1 One-day records

Figures 2a and b are vertical component displacements for the USArray stations TA.234A and TA.Z36A recorded on October 10th, 2010. The inter-station between this station pair is 213 km. As shown in Figure 3a, the SNR of the global cross-correlation is relatively low because we only use daily records. Only causal Rayleigh waves can be clearly observed in this CCF. Both amplitude and time-frequency spectra of the global cross-correlation function are presented in Figures 3b and c, respectively. Besides the dominant Rayleigh wave arrival, there are many background noises, especially from 0.2 to 0.5 Hz. The local cross-correlation between Figures 2a and b is shown in Figure 2c. With the local cross-correlation, we are able to extend the 1D CCF to a 2D image, which can be used to measure local similarity and design a better strategy to improve stacking quality. Although there are many incoherent background noises in Figure 2c, it includes a vertical coherent feature around time lags between 60 to 90 s. If we use a simple linear stacking for Figure 2c along the time axis (Equation 17), we will obtain the same signal as the global cross-correlation shown in Figure 3a. Next, we compute the local similarity between Figure 2c and the raw stacked CCF shown in Figure 3a. In the local similarity result (Figure 2d), the coherent signals between 60 to 90 s stand out in comparison to the background noise level. With the stacking weights based on the local similarity (Equation 18), the new CCF and its amplitude/time-frequency spectra are presented in Figures 3d–f. Compared with the global CCF in Figures 3a–c, the incoherent noises are significantly reduced without changing the dominant Rayleigh wave arrival. In the meanwhile, the overall amplitude spectrum does not change before and after applying the stacking weights (Figures 3b and e). Now, it is much easier for us to measure the phase or group dispersion curves of Rayleigh waves for the newly stacked signal (Figure 3d).

3.2 One-month records

Figures 4a and b present ambient noise CCFs for daily records (October 2010) by using global cross-correlation and local attributes, respectively. With these local attributes, incoherent background noises are attenuated while surface wave signals are preserved and much easier for us to measure. We observe alternative causal and acausal surface wave signals, which might come from the changes in ambient noise source distributions. Since there is data redundancy for Figure 4b, instead of directly stacking it over the time axis, we calculate local similarity again between the raw stacked CCF (Figure 5a) with Figure 4b, which is shown in Figure 4c. Again, coherent signals, such as surface wave packages, stand out in comparison with background noises. Figures 5a and d compare the raw stacked CCF and the new signal with local attributes. We observe significant improvement of the new signal in terms of SNR. Based on the time-frequency analysis shown in Figures 5c and f, the dispersive characteristics of Rayleigh waves are preserved while background noises from 0.1 to 0.3 Hz are attenuated. Similar to the previous daily example, the overall amplitude spectra do not change too much between these two CCFs.

Next, we compare the convergence of the raw global cross-correlation and the new result based on local attributes. Figure 6a shows the convergence of the raw daily global CCFs over one month. With the increasing stacking duration, the SNR is improved and coherent Rayleigh wave signals stand out gradually. However, the convergence and improvement of SNR are relatively low (Figure 7b). Here we use a one-month stacked CCF as the reference trace to compute the similarity between two signals (Equation 13). In this test, the signal window is [-100, -50] s and [50, 100] s, while the noise window ranges from [-200, -150] s and [150, 200] s. Improvements for the local attribute stacking CCF are shown in Figure 6b. In addition, another causal signal appears around 50 s after 10 days of stacking, which could be body waves or higher mode surface waves with an apparent velocity about 4.3 km/s. But those incoherent events (e.g., 20 s) are relatively well suppressed. We also speculate that the improvement of SNR with our method is significant with only several days of stacking (Figure 7b). By stacking with more than 10 days, both the correlation coefficient and SNR gradually become stable (Figure 7). It is intriguing that the first 5 days CCFs show strong acausal signals (e.g., -80 s in Figures 6a-b) and our improved SNR shows a bump around 5 days (Figure 7b). This might indicate the noise sources were mainly from station Z36A to station 234A before the first 5 days. Then more complex noise sources from stations 234A to station Z36A appeared and resulted in stronger causal signals (e.g. 50-80 s) and complex noises in CCFs. As a result, the SNR goes lower from 5-12 day stacking and gradually turns high with longer stacking.

We further test the proposed approach for 20 USArray stations deployed in North Texas and Oklahoma(Figure 8). One-month data (October 2010) are used for both raw stacking and the new approach. Comparisons of these two results can be found in Figure 9. We observe the SNRs of the new approach are much higher in comparison with the classical stacking approach. Except for the dominant Rayleigh waves, we also notice that there might be additional earlier arriving weaker events emerging from 500 km distance (highlighted by dark blue arrows). For example, the early arrival at about 640 km has an apparent velocity of 3.5 km/s, which is higher than the group velocity ($\sim$2.6 km/s) of the dominant Rayleigh waves. It could be a candidate for head/diving body waves or higher mode surface waves contaminated with noises. Even they are weak signals, our method can retain them as long as they are coherent.

Next, we test the proposed approach for several stations from three dense arrays and surrounding stations deployed in Minnesota and Wisconsin around the Midcontinent Rift (Figure 10; Wolin et al., 2015). Figure 11 shows the comparisons of one-month stacked CCFs based on the conventional approach and the proposed procedure. It is obvious the SNRs are improved quite a lot for most traces shown in Figure 11c with the proposed method. The overall average of the SNR from the traditional stacking method (4.8) is much smaller than the one (35.4) from the proposed method. Although the noise level between 6-15 s period is lower, the SNRs are improved from 6.7 to 53.8 by using the proposed stacking method as shown in Figure 12. It is interesting that SNRs for data in 15-30 s period band (Figure 13) are smaller compared to those in the periods of 6-15 s, and the proposed method helps us to improve the SNRs as expected. Panels d-f show the improved SNRs from two-month stacking compared with the monthly stacking (panels a-c) of Figures 11-13. We speculate that with a two-month stacking duration, the improvements of SNRs using conventional linear stacking are larger than the proposed method, but the SNRs from one-month stacking of the proposed method are still higher than those from two-month stacking with the conventional method. This is important for monitoring time-lapse changes of near-surface velocity changes with temporary arrays which usually have quite short acquisition durations (Nakata et al., 2016; Issa et al., 2017; Mordret et al., 2020; Zhang et al., 2022).

3.3 Applications to a nodal array

Between 2017 and 2019, 10 linear dense Distribution of Basin Amplification Seismic Investigation (BASIN) nodal arrays (SG1–SG4, and SB1–SB6) were deployed in the San Gabriel and San Bernardino basins for Fine characterization of basin shapes and depths (Y. Li et al., 2022; X. Wang et al., 2021). We apply our method to the SB1 array to validate its performance. The SB1 array (Figure. 14a) consists of 239 Fairfield ZLand nodes with standard 5 Hz 3-component geophones with spatial sampling of $\sim$250 m. It was deployed for approximately one month. The basin depth beneath the SB1 array is about 0-3.0 km (Y. Li et al., 2022), therefore, it is essential to have more measurements at lower period bands, such as 1.0-5.0 s. The data downloading and preprocessing are similar to previous tests except the continuous noises were down-sampled with a sampling frequency of 4 Hz and then bandpass filtered between 1-20 s.

Taking the first station as the master station, the corresponding CCFs are arranged according to their offset and displayed in Figure 14. As expected, with 5 days of stacking (Figure 14b), the CCFs from traditional linear stacking have strong noises for all station pairs. The 30-day stacking clearly improves the data quality so that we can observe the dominant acausal Rayleigh waves. Compared with CCFs from linear stacking, the proposed method helps us to enhance coherent signals even with only 5 days of stacking. The 30-day stacking further improves the data quality. Next, we measure phase dispersion curves for the stacked CCFs, which are the input for surface wave tomography (Yao et al., 2006; Fang et al., 2015). There are three criteria to make sure the measured dispersion curves are stable. (1) The SNR is larger than 5.0; (2) The inter-station distance is larger than 1.5 wavelength at corresponding periods (Bensen et al., 2007; G. Chen et al., 2023). (3) The picked phase velocity is within $\pm$ 12% of a 3D reference phase velocity model. Here the 3D reference phase velocity model is constructed from a local tomographic model, CVM-S 4.26 (Lee et al., 2014), by conducting a forward modeling based on the fast-marching method (Rawlinson & Sambridge, 2004; Fang et al., 2015). Thanks to the improvement of SNR from our method, more high-quality dispersion curves pass the selection criteria with 5-day stacking compared with those measured from linear stacking as illustrated in Figure 15. More importantly, the number of measurements for 1-5 s and 10-15 s period bands are also increased. These are further improved for 30-day stacked CCFs using the proposed method.

4 Discussions

Commonly used ambient noise processing procedures (Bensen et al., 2007) require stacking over long, continuous records to enhance coherent signals, also require the assumption of evenly distributed ambient sources in order to obtain good estimates of Green's functions. However, both conditions impose constraints on the application of temporary arrays with short acquisition duration.

To improve the SNR of CCFs stacked with a short duration, Xie et al. (2020) proposed a root-mean-square-ratio selective (RMSRS) stacking procedure to remove those CCFs that negatively contribute to the SNR of the final stacked CCF. It is realized by comparing the root-mean-square ratio of signals and noises for each CCF and the stacked CCF. Therefore, its effectiveness depends on an accurate definition of signal/noise windows. Here, we compare the conventional linear stacking, proposed procedure and RMSRS stacking for the station pair TA.A12A-TA.A18A at 10-35 s period band. The signal window (Figure 17c) is defined with the reference time t_{ref} and maximum period of interest T_{max} as $[t_{ref} - 2T_{max}, t_{ref} + 2T_{max}]$. Here, the reference time $t_{ref} = d/v_{ref}$ is defined by the inter-station distance d (425 km) and a reference group velocity v_{ref} (3.0 km/s). The noise window is defined from 0 to the signal windows and $4T_{max}$ out of the signal windows (Xie et al., 2020). Compared with linear stacking, the RMSRS stacking successfully suppresses the noises within the defined noise windows and therefore improves the SNR as highlighted in Figures 17d and f. Similarly, our approach also successfully attenuates those incoherent noises and significantly improves the SNR, which is about two times the other two methods. However, the coherent signal at about 75 seconds with an apparent velocity of 5.6 km/s is retained by our approach due to the similarity between local CCFs and the stacked CCFs. The main reason that the RMSRS-based method helps to suppress this signal is because they are selected as noises. Such selection seems to be challenging to deal with low-SNR CCFs at shorter period bands (3-16 s) obtained from the high-frequency nodal array (Figure 14). As illustrated in Figures 14d, the improvement of selective stacking is limited compared with linear stacking. On the contrary, our approach significantly improves the stacking quality and helps us to obtain more high-quality dispersion curves. We note here that, for fair comparisons, the RMSRS-based method is implemented on the raw stacked CCFs. For better performance, Xie et al. (2020) suggested using RMSRS stacking at several narrow period bands so that it could define a better selection window, which is out of the scope of this study.

Although the proposed approach helps us to improve the SNR of CCFs with only several days of stacking, the non-causality and asymmetry, which mainly arise from the uneven distribution of ambient noise sources, remain challenging. For tomography purposes, theoretical works indicate that phase velocities can be estimated from the empirical Green's functions, which are obtained by taking the negative time derivative of the symmetric cross-correlation under the assumption of a spatially homogeneous ambient-noise source distribution (Lobkis & Weaver, 2001; Sabra et al., 2005; Snieder, 2004; Yao et al., 2006; Lin et al., 2008). As suggested by Yao et al. (2006), inhomogeneous source distribution may contribute to 1–3 percent inconsistency between phase velocity measurements and the traditional earthquake-based two-station method between periods of 20–30 s. Therefore, the "symmetric" CCF is usually taken by the average of the cross-correlation at positive and negative correlation lag times (Yao et al., 2006; Lin et al., 2008). However, how much such averaged CCFs affect tomography results remains unknown. On the other hand, several studies suggested measuring full-waveform differences of CCFs and source location simultaneously (Tromp et al., 2010; Sager et al., 2018; Datta et al., 2023), which naturally mitigates the uncertainty caused by the source distributions. In both cases, it is important for us to obtain high SNR CCFs, especially for short-duration nodal arrays. In addition, because the high-quality phase velocity measurements obtained from 5-day stacking (Figure 15c) by the proposed approach is comparable with those obtained from the 30-day linear stacking, the surface wave-based monitoring (Durand et al., 2011; Brenguier et al., 2020) seems to be possible, albeit coda waves are mostly used (Mao et al., 2019, 2020; Luo et al., 2023). Another potential application for our approach could be weak coherent signal extraction, such as body waves (Zhan et al., 2010; Poli et al., 2012; Nakata et al., 2015, 2016; Mao et al., 2020). Body waves extracted from ambient noise CCFs have been proven to be capable of improving the imaging resolution compared to surface wave tomography (Nakata et al., 2015). As illustrated in Figures 6 and 14, those coherent signals with an apparent velocity larger than 4.0 km/s might be good candidates for body waves, although we are not able to rule out the possibility of higher-order surface waves.

The last factor we need to consider is the computational and memory cost for our approach. Taking station pair TA.A12A and TA.A18A as an example, we have 30 days of recordings and 86400-time samples for each day. The cross-correlation time lag varies from -480 s to 480 s for every 1 s. We calculate the global cross-correlations every 60 min-utes with 75% overlaps, which yields 91 subset CCFs for each day. They are then normalized and stacked to obtain the dayily CCF. Finally, 30 days of global cross-correlation and linear stacking take 47.9 s and 0.02 s, respectively, while the local cross-correlation, local similarity and weighted stacking take 196.7 s, 12,150 s and 78.3 s. We note here, that the local cross-correlation solved with our method actually contains 86,400 CCFs for each day. It helps us to extend the dimension of CCFs dramatically (534 times the number of global CCFs), but with only four times the computational cost compared to global stacking. Whereas, such a high dimension, in turn, greatly decreases the efficiency of our approach. To mitigate this problem, the local CCFs are downsampled 60 times (from 1.0 s to 60 s) by taking the average CCFs for every 60 samples, and then the computational cost for local similarity and weighted stacking is decreased to 167.1 and 10.0 s, respectively. The final stacked CCFs by using these local CCFs are almost the same as the original local CCFs. Overall, the computational cost for the proposed approach is 7.8 times the computational cost compared to the hourly global stacking, but with 15.8 times more CCFs. In addition, increasing the cross-correlation time window (e.g., from one hour to three hours) does not affect the efficiency of our approach, but will increase the computational time for global cross-correlation. Finally, we compare the computation cost using the nodal array. For each station, we have 345,600 time samples per day with a 4 Hz sampling frequency. The cross-correlation time lag varies from -120 s to 120 s for every 0.25 s. The global and local CCFs are calculated the same as in the previous example. To save computational costs, we downsample the local CCFs 60 times (from 0.25 s to 15 s). Then all station-pairs are distributed to 72 CPU cores for parallel computations. Our approach takes 5792.0 s and the traditional linear stacking takes 792.0 s. The final stacked CCFs are compared in section 3.3. In conclusion, our approach significantly helps us to improve the SNR of CCFs, but with about 7.5 times the computation cost compared to traditional linear stacking. Such extra computational costs are bearable compared to the following computational costs tomography (Zhu, 2018; Wu et al., 2023; G. Chen et al., 2023).

5 Conclusion

In this study, by taking advantage of local attributes, we present a new approach to increase the SNRs of ambient noise CCFs. Two local attributes are used in this study: local cross-correlation and local similarity. The local cross-correlation is employed to extend the dimensionality of daily CCFs, and the local similarity is used to design better-stacking weights to enhance coherent signals and attenuate incoherent background noises. Applications to ambient noise records from several broadband stations and a high-frequency nodal array demonstrate the performance of the proposed approach. With higher SNRs, we are able to extract more high-quality dispersion curves, which are important for surface wave tomography. In addition, 5-day stacking by our approach can produce CCFs comparable to 30-day linear stacking in terms of SNRs, demonstrating its potential applications for time-lapse monitoring. In addition, extracting coherent weak signals, such as body waves, could be another application of the proposed approach.

Open Research

All seismic data used in this study can be obtained from the IRIS Data Management Center (`https://ds.iris.edu/ds`) under the network codes TA and XI. We use the Noisepy (`https://noise-python.readthedocs.io/en/latest/`) for parallel data downloading and preprocessing (Jiang & Denolle, 2020). The open software Madagascar (Fomel et al., 2013) download from (`http://www.ahay.org`) is used to calculate local similarity and plot figures. We also use PyGMT (Wessel et al., 2019) downloaded from (`https://www.pygmt.org/latest/`) for plotting figures.

Author Contributions

The authors confirm their contribution to the paper as follows: study conception and design: Hejun Zhu; data collection: Hejun Zhu and Bin He; analysis and interpretation of results: Hejun Zhu, Bin He and David Lumley; draft manuscript preparation: Hejun Zhu, Bin He and David Lumley. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments

This paper is contribution no. *** from the Department of Geosciences at the University of Texas at Dallas. The numerical results are computed through the Optane nodes on the UTD Seismology Group HPC clusters. This research is partially supported by the sponsors of the UT Dallas 3D + 4D Seismic Full Waveform Inversion research consortium.

References

Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bulletin of the Earthquake Research Institute*, *35*, 415–456.

Alvarez, L., & Mazorra, L. (1994). Signal and image restoration using shock filters and anisotropic diffusion. *SIAM journal on numerical analysis*, *31*(2), 590–605.

Baig, A. M., Campillo, M., & Brenguier, F. (2009). Denoising seismic noise cross correlations. *Journal of Geophysical Research: Solid Earth*, *114*(B8).

Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A. L., Lin, F., Moschetti, M., ... Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical journal international*, *169*(3), 1239–1260.

Brenguier, F., Courbis, R., Mordret, A., Campman, X., Boué, P., Chmiel, M., ... others (2020). Noise-based ballistic wave passive seismic monitoring. part 1: body waves. *Geophysical Journal International*, *221*(1), 683–691.

Brenguier, F., Shapiro, N. M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., & Nercessian, A. (2008). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, *1*(2), 126–130.

Buckingham, M. J., Berknout, B. V., & Glegg, S. A. (1992). Imaging the ocean with ambient noise. *Nature*, *356*(6367), 327–329.

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

Chen, G., Chen, J., Tape, C., Wu, H., & Tong, P. (2023). Double-difference adjoint tomography of the crust and uppermost mantle beneath alaska. *Journal of Geophysical Research: Solid Earth*, *128*(1), e2022JB025168.

Chen, M., Huang, H., Yao, H., van der Hilst, R., & Niu, F. (2014). Low wave speed zones in the crust beneath se tibet revealed by ambient noise adjoint tomography. *Geophysical Research Letters*, *41*(2), 334–340.

Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, *33*(2), 264–269.

Clarke, D., Zaccarelli, L., Shapiro, N., & Brenguier, F. (2011). Assessment of resolution and accuracy of the moving window cross spectral technique for monitoring crustal temporal variations using ambient seismic noise. *Geophysical Journal International*, *186*(2), 867–882.

Datta, A., Shekar, B., & Kumar, P. L. (2023). Acoustic full waveform inversion for 2-d ambient noise source imaging. *Geophysical Journal International*, *234*(3), 1628–1639.

Daubechies, I., Defrise, M., & De Mol, C. (2004). An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, *57*(11), 1413–1457.

De Plaen, R. S., Lecocq, T., Caudron, C., Ferrazzini, V., & Francis, O. (2016). Single-station monitoring of volcanoes using seismic ambient noise. *Geophysical Research Letters*, *43*(16), 8511–8518.

Deriche, R. (1993). *Recursively implementating the gaussian and its derivatives* (Unpublished doctoral dissertation). INRIA.

Díaz, E., & Sava, P. (2015). Data domain wavefield tomography using local correlation functions. In *2015 seg annual meeting.*

Dougherty, S. L., Cochran, E. S., & Harrington, R. M. (2019). The large-n seismic survey in oklahoma (lasso) experiment. *Seismological Research Letters*, *90*(5), 2051–2057.

Draganov, D., Campman, X., Thorbecke, J., Verdel, A., & Wapenaar, K. (2009). Reflection images from ambient seismic noise. *Geophysics*, *74*(5), A63–A67.

Durand, S., Montagner, J., Roux, P., Brenguier, F., Nadeau, R., & Ricard, Y. (2011). Passive monitoring of anisotropy change associated with the parkfield 2004 earthquake. *Geophysical Research Letters*, *38*(13).

Fan, X., Guo, Z., Zhao, Y., & Chen, Q.-F. (2022). Crust and uppermost mantle magma plumbing system beneath changbaishan intraplate volcano, china/north korea, revealed by ambient noise adjoint tomography. *Geophysical Research Letters*, *49*(12), e2022GL098308.

Fang, H., Yao, H., Zhang, H., Huang, Y.-C., & van der Hilst, R. D. (2015). Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: methodology and application. *Geophysical Journal International*, *201*(3), 1251–1263.

Fomel, S. (2007a). Local seismic attributes. *Geophysics*, *72*(3), A29–A33.

Fomel, S. (2007b). Shaping regularization in geophysical-estimation problems. *Geophysics*, *72*(2), R29–R36.

Fomel, S., & Jin, L. (2009). Time-lapse image registration using the local similarity attribute. *Geophysics*, *74*(2), A7–A11.

Fomel, S., Sava, P., Vlad, I., Liu, Y., & Bashkardin, V. (2013). Madagascar: Open-source software project for multidimensional data analysis and reproducible computational experiments. *Journal of Open Research Software*, *1*(1).

Fomel, S., & van der Baan, M. (2014). Local skewness attribute as a seismic phase detector. *Interpretation*, *2*(1), SA49–SA56.

Gao, H., & Shen, Y. (2014). Upper mantle structure of the cascades from full-wave ambient noise tomography: Evidence for 3d mantle upwelling in the back-arc. *Earth and Planetary Science Letters*, *390*, 222–233.

Hadziioannou, C., Larose, E., Baig, A., Roux, P., & Campillo, M. (2011). Improving temporal resolution in ambient noise monitoring of seismic wave speed. *Journal of Geophysical Research: Solid Earth*, *116*(B7).

Hale, D. (2006). Fast local cross-correlations of images. In *Seg technical program expanded abstracts 2006* (pp. 3160–3164). Society of Exploration Geophysicists.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., . . . Oliphant, T. E. (2020, September). Array programming with NumPy. *Nature*, *585*(7825), 357–362. Retrieved from `https://doi.org/10.1038/s41586-020-2649-2` doi: 10.1038/s41586-020-2649-2

Huang, H., Yao, H., & van der Hilst, R. D. (2010). Radial anisotropy in the crust of se tibet and sw china from ambient noise interferometry. *Geophysical Research Letters*, *37*(21).

Issa, N. A., Lumley, D., & Pevzner, R. (2017). Passive seismic imaging at reservoir depths using ambient seismic noise recorded at the otway co2 geological storage research facility. *Geophysical Journal International*, *209*(3), 1622–1628.

Jiang, C., & Denolle, M. A. (2020). Noisepy: A new high-performance python tool for ambient-noise seismology. *Seismological Research Letters*, *91*(3), 1853–1866.

Le Breton, M., Bontemps, N., Guillemot, A., Baillet, L., & Larose, É. (2021). Landslide monitoring using seismic ambient noise correlation: challenges and applications. *Earth-Science Reviews*, *216*, 103518.

Lee, E.-J., Chen, P., Jordan, T. H., Maechling, P. B., Denolle, M. A., & Beroza, G. C. (2014). Full-3-d tomography for crustal structure in southern california based on the scattering-integral and the adjoint-wavefield methods. *Journal of Geophysical Research: Solid Earth*, *119*(8), 6421–6451.

Li, G., Niu, F., Yang, Y., & Xie, J. (2018). An investigation of time-frequency domain phase-weighted stacking and its application to phase-velocity extraction from ambient noise's empirical green's functions. *Geophysical Journal International*, *212*(2), 1143–1156.

Li, Y., Villa, V., Clayton, R., & Persaud, P. (2022). Shear wave velocities in the san gabriel and san bernardino basins, california. *Authorea Preprints*.

Lin, F.-C., Moschetti, M. P., & Ritzwoller, M. H. (2008). Surface wave tomography of the western united states from ambient seismic noise: Rayleigh and love wave phase velocity maps. *Geophysical Journal International*, *173*(1), 281–298.

Lin, F.-C., Ritzwoller, M. H., Yang, Y., Moschetti, M. P., & Fouch, M. J. (2011). Complex and variable crustal and uppermost mantle seismic anisotropy in the western united states. *Nature Geoscience*, *4*(1), 55–61.

Lin, F.-C., Tsai, V. C., Schmandt, B., Duputel, Z., & Zhan, Z. (2013). Extracting seismic core phases with array interferometry. *Geophysical Research Letters*, *40*(6), 1049–1053.

Liu, G., Fomel, S., & Chen, X. (2011). Stacking angle-domain common-image gathers for normalization of illumination. *Geophysical Prospecting*, *59*(2), 244–255.

Liu, G., Fomel, S., Jin, L., & Chen, X. (2009). Stacking seismic data using local correlation. *Geophysics*, *74*(3), V43–V48.

Liu, G., Persaud, P., & Clayton, R. W. (2018). Structure of the northern los angeles basins revealed in teleseismic receiver functions from short-term nodal seismic arrays. *Seismological Research Letters*, *89*(5), 1680–1689.

Liu, Y., Niu, F., Chen, M., & Yang, W. (2017). 3-d crustal and uppermost mantle structure beneath ne china revealed by ambient noise adjoint tomography. *Earth and Planetary Science Letters*, *461*, 20–29.

Lobkis, O. I., & Weaver, R. L. (2001). On the emergence of the green's function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, *110*(6), 3011–3017.

Luo, B., Zhang, S., & Zhu, H. (2023). Monitoring seasonal fluctuation and long-term trends for the greenland ice sheet using seismic noise auto-correlations. *Geophysical Research Letters*, *50*(7), e2022GL102146.

Maguire, R., Schmandt, B., Li, J., Jiang, C., Li, G., Wilgus, J., & Chen, M. (2022). Magma accumulation at depths of prior rhyolite storage beneath yellowstone caldera. *Science*, *378*(6623), 1001–1004.

Mainsant, G., Larose, E., Brönnimann, C., Jongmans, D., Michoud, C., & Jaboyedoff, M. (2012). Ambient seismic noise monitoring of a clay landslide: Toward failure prediction. *Journal of Geophysical Research: Earth Surface*, *117*(F1).

Mao, S., Campillo, M., van Der Hilst, R. D., Brenguier, F., Stehly, L., & Hillers, G. (2019). High temporal resolution monitoring of small variations in crustal strain by dense seismic arrays. *Geophysical Research Letters*, *46*(1), 128–137.

Mao, S., Lecointre, A., van der Hilst, R. D., & Campillo, M. (2022). Space-time monitoring of groundwater fluctuations with passive seismic interferometry. *Nature communications*, *13*(1), 1–9.

Mao, S., Mordret, A., Campillo, M., Fang, H., & van der Hilst, R. D. (2020). On the measurement of seismic traveltime changes in the time–frequency domain with wavelet cross-spectrum analysis. *Geophysical Journal International*, *221*(1), 550–568.

Mordret, A., Courbis, R., Brenguier, F., Chmiel, M., Garambois, S., Mao, S., ... others (2020). Noise-based ballistic wave passive seismic monitoring–part 2: surface waves. *Geophysical Journal International*, *221*(1), 692–705.

Moschetti, M., Ritzwoller, M., Lin, F., & Yang, Y. (2010). Seismic evidence for widespread western-us deep-crustal deformation caused by extension. *Nature*, *464*(7290), 885–889.

Nakata, N., Boué, P., Brenguier, F., Roux, P., Ferrazzini, V., & Campillo, M. (2016). Body and surface wave reconstruction from seismic noise correlations between arrays at piton de la fournaise volcano. *Geophysical Research Letters*, *43*(3), 1047–1054.

Nakata, N., Chang, J. P., Lawrence, J. F., & Boué, P. (2015). Body wave extraction and tomography at long beach, california, with ambient-noise interferometry. *Journal of Geophysical Research: Solid Earth*, *120*(2), 1159–1173.

Nakata, N., Gualtieri, L., & Fichtner, A. (2019). *Seismic ambient noise*. Cambridge

University Press.

Nakata, N., & Snieder, R. (2011). Near-surface weakening in japan after the 2011 tohoku-oki earthquake. *Geophysical Research Letters*, *38*(17).

Poli, P., Pedersen, H., & Campillo, M. (2012). Emergence of body waves from cross-correlation of short period seismic noise. *Geophysical Journal International*, *188*(2), 549–558.

Rawlinson, N., & Sambridge, M. (2004). Wave front evolution in strongly heterogeneous layered media using the fast marching method. *Geophysical Journal International*, *156*(3), 631–647.

Rickett, J., & Lumley, D. (2001). Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the gulf of mexico. *Geophysics*, *66*(4), 1015–1025.

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W., & Fehler, M. C. (2005). Extracting time-domain green's function estimates from ambient seismic noise. *Geophysical research letters*, *32*(3).

Sager, K., Boehm, C., Ermert, L., Krischer, L., & Fichtner, A. (2020). Global-scale full-waveform ambient noise inversion. *Journal of Geophysical Research: Solid Earth*, *125*(4), e2019JB018644.

Sager, K., Ermert, L., Boehm, C., & Fichtner, A. (2018). Towards full waveform ambient noise inversion. *Geophysical Journal International*, *212*(1), 566–590.

Schimmel, M., Stutzmann, E., & Gallart, J. (2011). Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale. *Geophysical Journal International*, *184*(1), 494–506.

Seats, K. J., Lawrence, J. F., & Prieto, G. A. (2012). Improved ambient noise correlation functions using welch s method. *Geophysical Journal International*, *188*(2), 513–523.

Shapiro, N. M., & Campillo, M. (2004). Emergence of broadband rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, *31*(7).

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*, *307*(5715), 1615–1618.

Shen, W., & Ritzwoller, M. H. (2016). Crustal and uppermost mantle structure beneath the united states. *Journal of Geophysical Research: Solid Earth*, *121*(6), 4306–4342.

Snieder, R. (2004). Extracting the green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical review E*, *69*(4), 046610.

Stehly, L., Cupillard, P., & Romanowicz, B. (2011). Towards improving ambient noise tomography using curvelet denoising filters and sem simulations of seismic ambient noise simultaneously. *Comptes Rendus Geoscience*, *343*(8-9), 591–599.

Tromp, J., Luo, Y., Hanasoge, S., & Peter, D. (2010). Noise cross-correlation sensitivity kernels. *Geophysical Journal International*, *183*(2), 791–819.

Van Vliet, L. J., Young, I. T., & Verbeek, P. W. (1998). Recursive gaussian derivative filters. In *Proceedings. fourteenth international conference on pattern recognition (cat. no. 98ex170)* (Vol. 1, pp. 509–514).

Wang, K., Yang, Y., Basini, P., Tong, P., Tape, C., & Liu, Q. (2018). Refined crustal and uppermost mantle structure of southern california by ambient noise adjoint tomography. *Geophysical Journal International*, *215*(2), 844–863.

Wang, Q.-Y., & Yao, H. (2020). Monitoring of velocity changes based on seismic ambient noise: A brief review and perspective. *Earth and Planetary Physics*, *4*(5), 532–542.

Wang, X., Zhan, Z., Zhong, M., Persaud, P., & Clayton, R. W. (2021). Urban basin structure imaging based on dense arrays and bayesian array-based coherent receiver functions. *Journal of Geophysical Research: Solid Earth*, *126*(9), e2021JB022279.

Wapenaar, K. (2004). Retrieving the elastodynamic green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical review letters*, *93*(25), 254301.

Weaver, R. L., & Lobkis, O. I. (2001). Ultrasonics without a source: Thermal fluctuation correlations at mhz frequencies. *Physical Review Letters*, *87*(13), 134301.

Weaver, R. L., & Yoritomo, J. Y. (2018). Temporally weighting a time varying noise field to improve green function retrieval. *The Journal of the Acoustical Society of America*, *143*(6), 3706–3719.

Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556–5564.

Wolin, E., van der Lee, S., Bollmann, T. A., Wiens, D. A., Revenaugh, J., Darbyshire, F. A., ... Wysession, M. E. (2015). Seasonal and diurnal variations in long-period noise at spree stations: The influence of soil characteristics on shallow stations' performance. *Bulletin of the Seismological Society of America*, *105*(5), 2433–2452.

Woollard, G. (1965). The bouguer gravity anomaly map of the united states. *Eos, Transactions American Geophysical Union*, *46*(1), 197–202.

Wu, S.-M., Huang, H.-H., Lin, F.-C., Farrell, J., & Schmandt, B. (2023). Extreme seismic anisotropy indicates shallow accumulation of magmatic sills beneath yellowstone caldera. *Earth and Planetary Science Letters*, *616*, 118244.

Xie, J., Yang, Y., & Luo, Y. (2020). Improving cross-correlations of ambient noise using an rms-ratio selection stacking method. *Geophysical Journal International*, *222*(2), 989–1002.

Yang, X., Bryan, J., Okubo, K., Jiang, C., Clements, T., & Denolle, M. A. (2023). Optimal stacking of noise cross-correlation functions. *Geophysical Journal International*, *232*(3), 1600–1618.

Yang, Y., Ritzwoller, M. H., Levshin, A. L., & Shapiro, N. M. (2007). Ambient noise rayleigh wave tomography across europe. *Geophysical Journal International*, *168*(1), 259–274.

Yao, H., van Der Hilst, R. D., & De Hoop, M. V. (2006). Surface-wave array tomography in se tibet from ambient seismic noise and two-station analysis—i. phase velocity maps. *Geophysical Journal International*, *166*(2), 732–744.

Yao, H., Van Der Hilst, R. D., & Montagner, J.-P. (2010). Heterogeneity and anisotropy of the lithosphere of se tibet from surface wave array tomography. *Journal of Geophysical Research: Solid Earth*, *115*(B12).

Zhan, Z., Ni, S., Helmberger, D. V., & Clayton, R. W. (2010). Retrieval of moho-reflected shear wave arrivals from ambient seismic noise. *Geophysical Journal International*, *182*(1), 408–420.

Zhang, Z., Nakata, N., Karplus, M., Kaip, G., & Yi, J. (2022). Shallow ice-sheet composite structure revealed by seismic imaging near the west antarctic ice

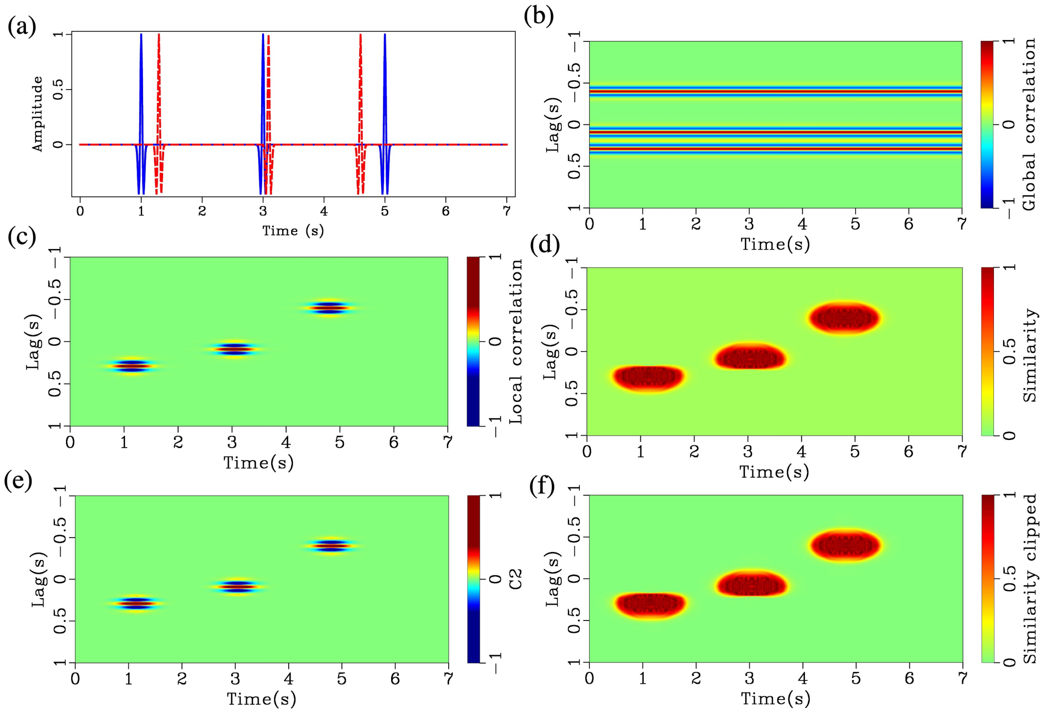


Figure 1. Comparison of local and global cross-correlations. Panel(a) shows two signals with non-stationary time shifts. Three events with time shifts of -0.4 s, 0.1 s and 0.3 s are used. Panels (b) and (c) present global and local cross-correlations, respectively. Here, the 1D global cross-correlations are extended along the time dimension for better comparisons with local cross-correlations. σ is set to 0.2 s so that it is small enough to capture the non-stationary property of these two signals. Panels (d) and (f) show the similarity ($\mathbf{s}$ in Equation 16) between local and global cross-correlations before and after applying a threshold ($\alpha(t,\tau)$ in Equation 19), respectively. Panel e shows the prestack local cross-correlations weighted by local similarity (c_2 in Equation 18)

.

sheet (wais) divide camp. *Journal of Geophysical Research: Earth Surface*, e2022JF006777.

Zhu, H. (2018). Crustal wave speed structure of north texas and oklahoma based on ambient noise cross-correlation functions and adjoint tomography. *Geophysical Journal International*, *214*(1), 716–730.

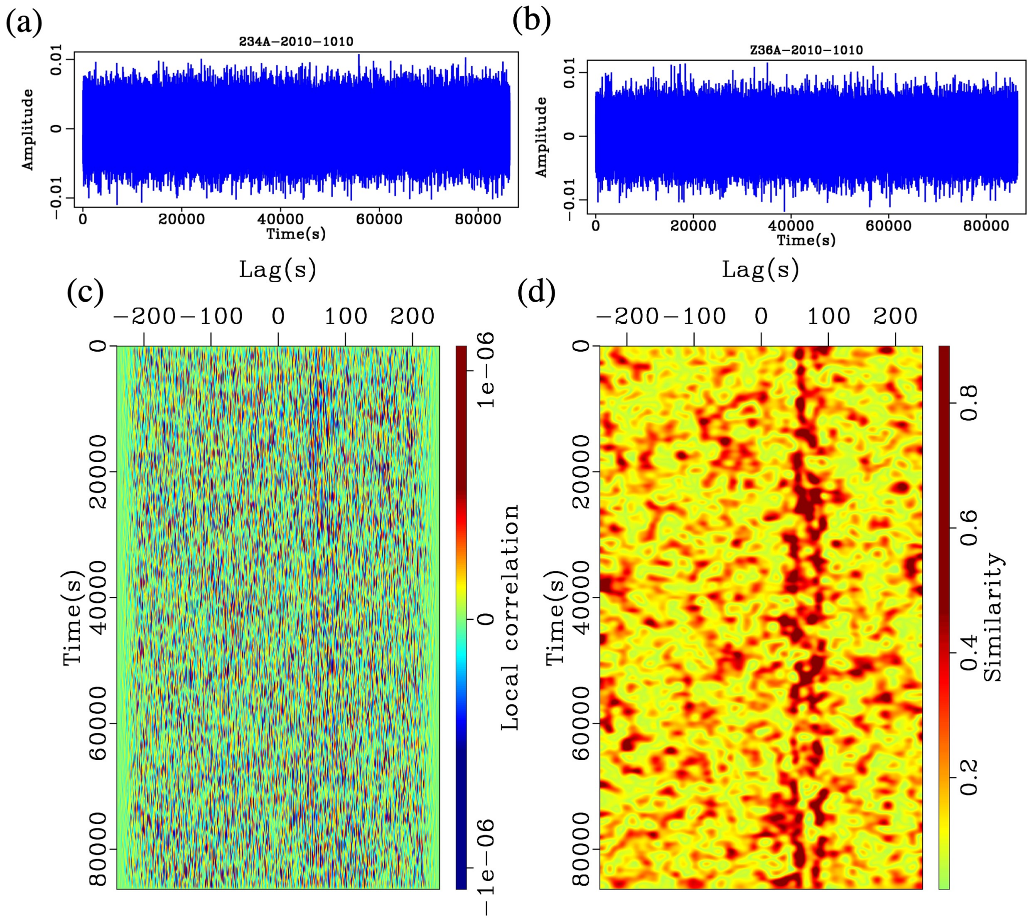


Figure 2. Local cross-correlation and local similarity for one-day CCFs between stations TA.234A and TA.Z36A on October 10, 2010. Panels (a) and (b) show one-day records for stations TA.234A and TA.Z36A, respectively. Panel (c) presents local cross-correlation between these two signals. Panel (d) shows the local similarity between panel (c) and the linearly stacked signal shown in Figure 3a. The causal Rayleigh waves between 60-80 s can be clearly observed in panel (d).

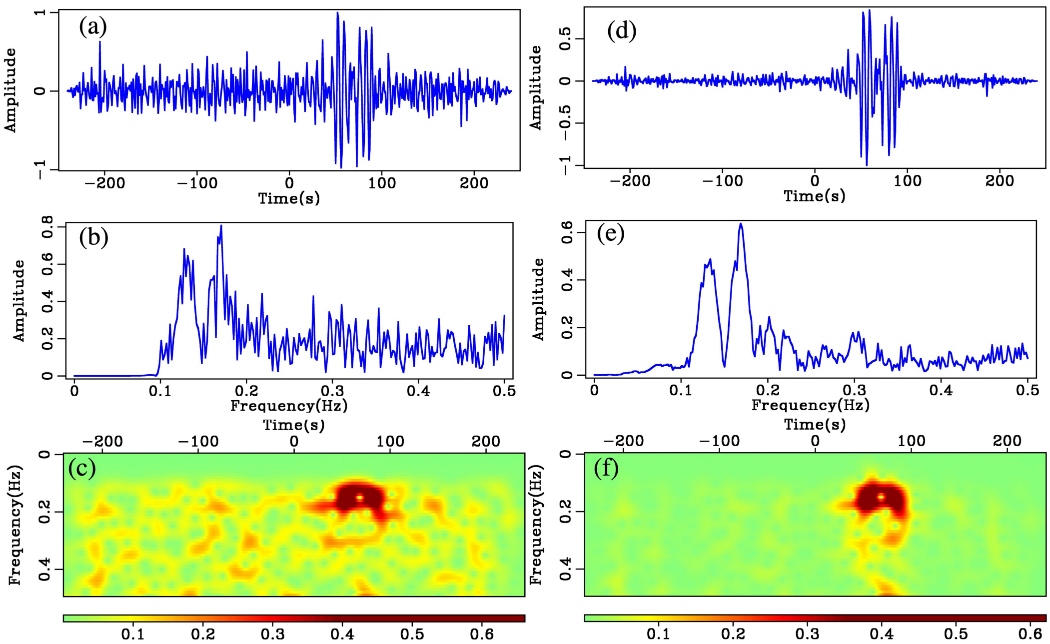


Figure 3. Comparisons of one-day CCFs from simple stacking and stacking with local attributes. Panel (a) shows the CCF with simple linear stacking (Equation 1). Panels (b) and (c) are the amplitude and time-frequency spectra of panel (a), respectively. Panel (d) is the CCF stacked with local similarity. Panels (e) and (f) are the amplitude and time-frequency spectra of panel (d).

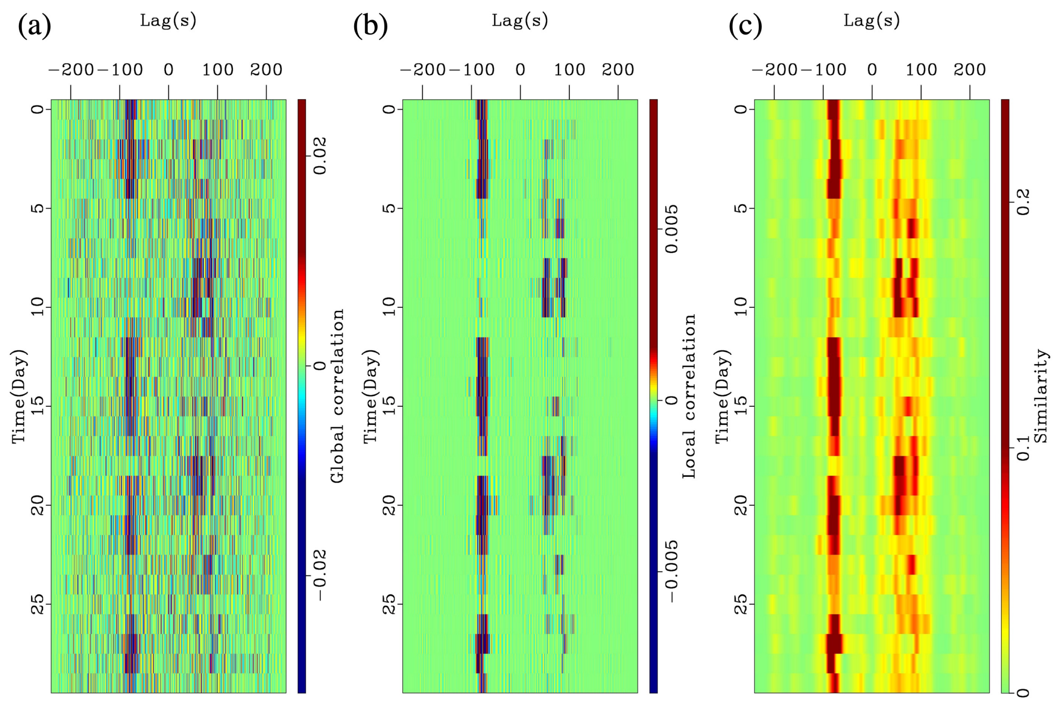


Figure 4. Comparisons of daily CCFs within one month (October 2010) for simple stacking (a) and stacking with local attribute (b). Panel (c) shows the local similarity between panel (b) and the linear stacked result shown in Figure 5a.

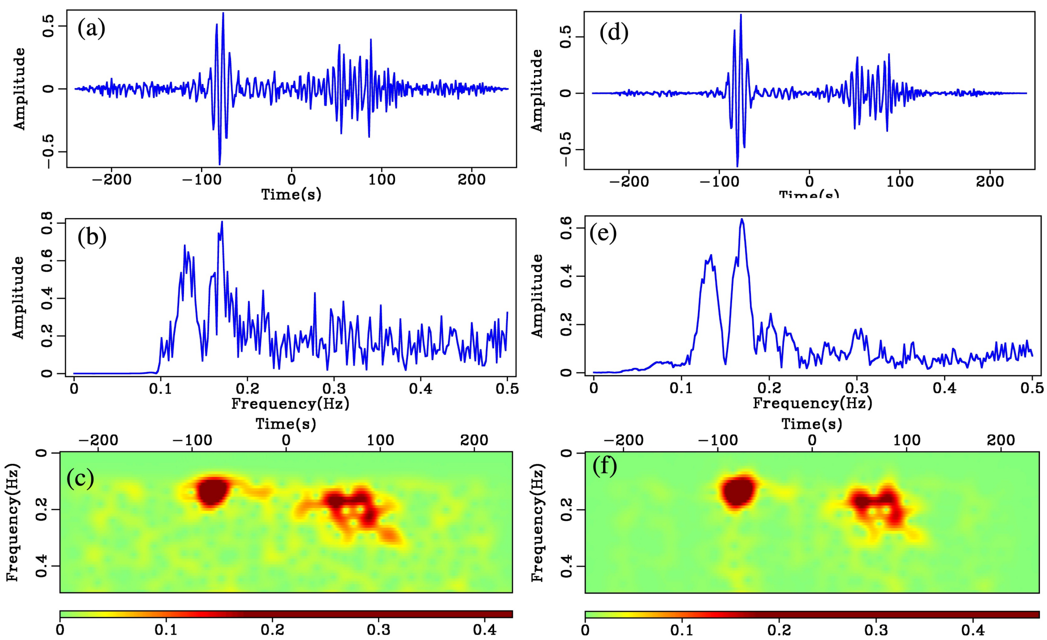


Figure 5. The same setting as Figure 3 except for monthly stacked results on October 2010.

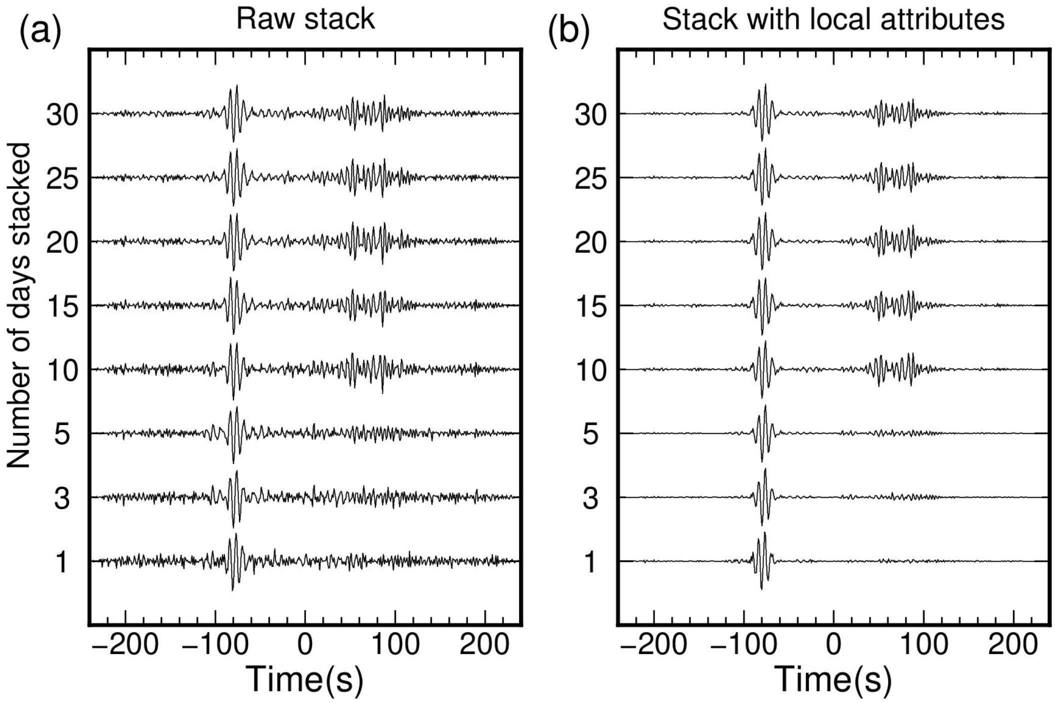


Figure 6. Comparisons of convergence for conventional (a) and the proposed procedure (b). The causal and acasual signals could be clearly identified around ±80 s with an apparent group velocity of 2.7 km/s. After 10 days of stacking, we observe a causal coherent signal appear around 50 s with an apparent velocity of 4.3 km/s, possibly representing body waves.

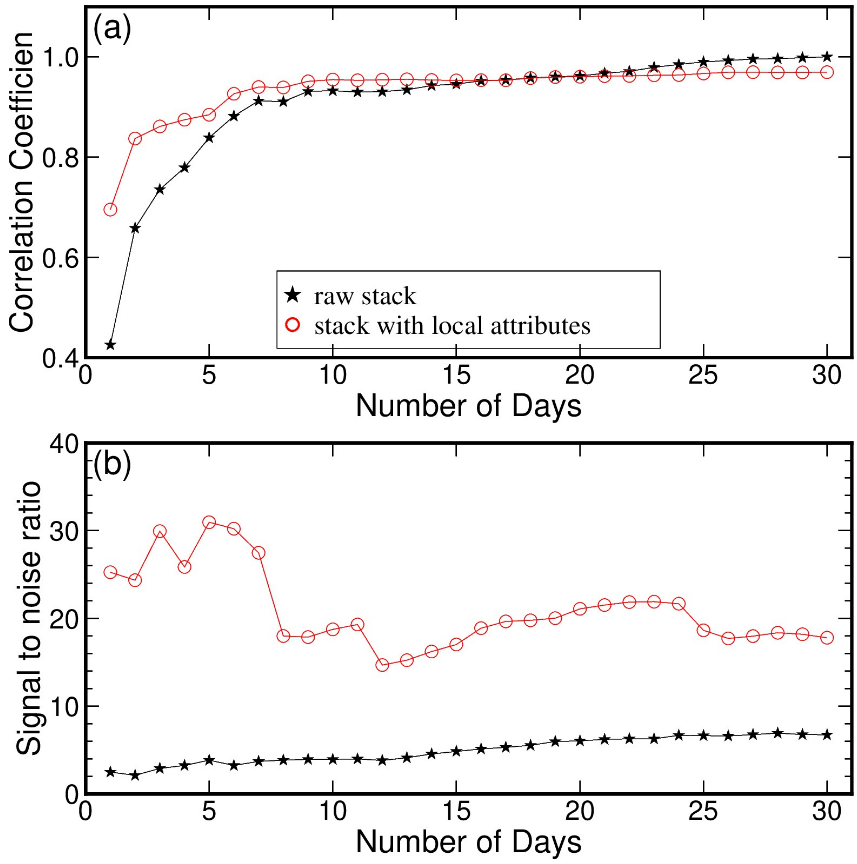


Figure 7. Improvements of correlation coefficients (a) and SNRs (b) for one-month stacking results. Black stars and open red circles are the results of the conventional stacking procedure and the proposed approach. The reference trace for calculating the correlation coefficients is the 30-day raw stack.

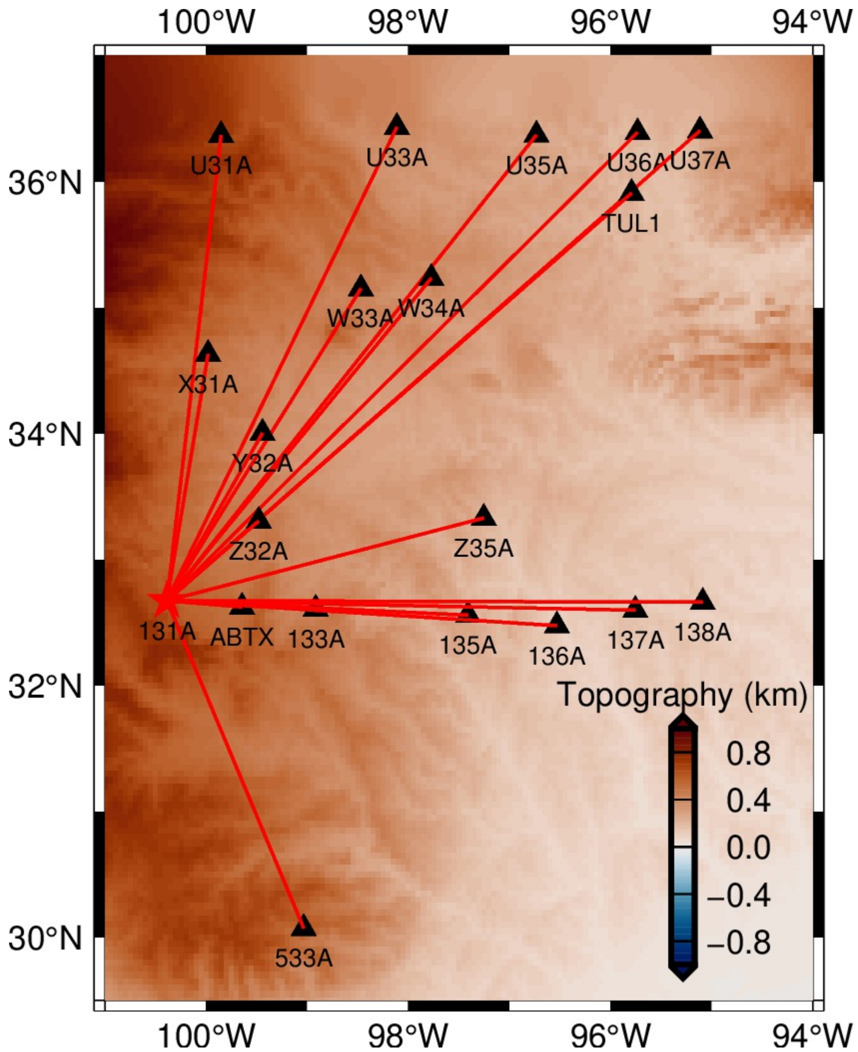


Figure 8. Distributions of the USArray stations in north Texas and Oklahoma used for comparisons in Figure 9.

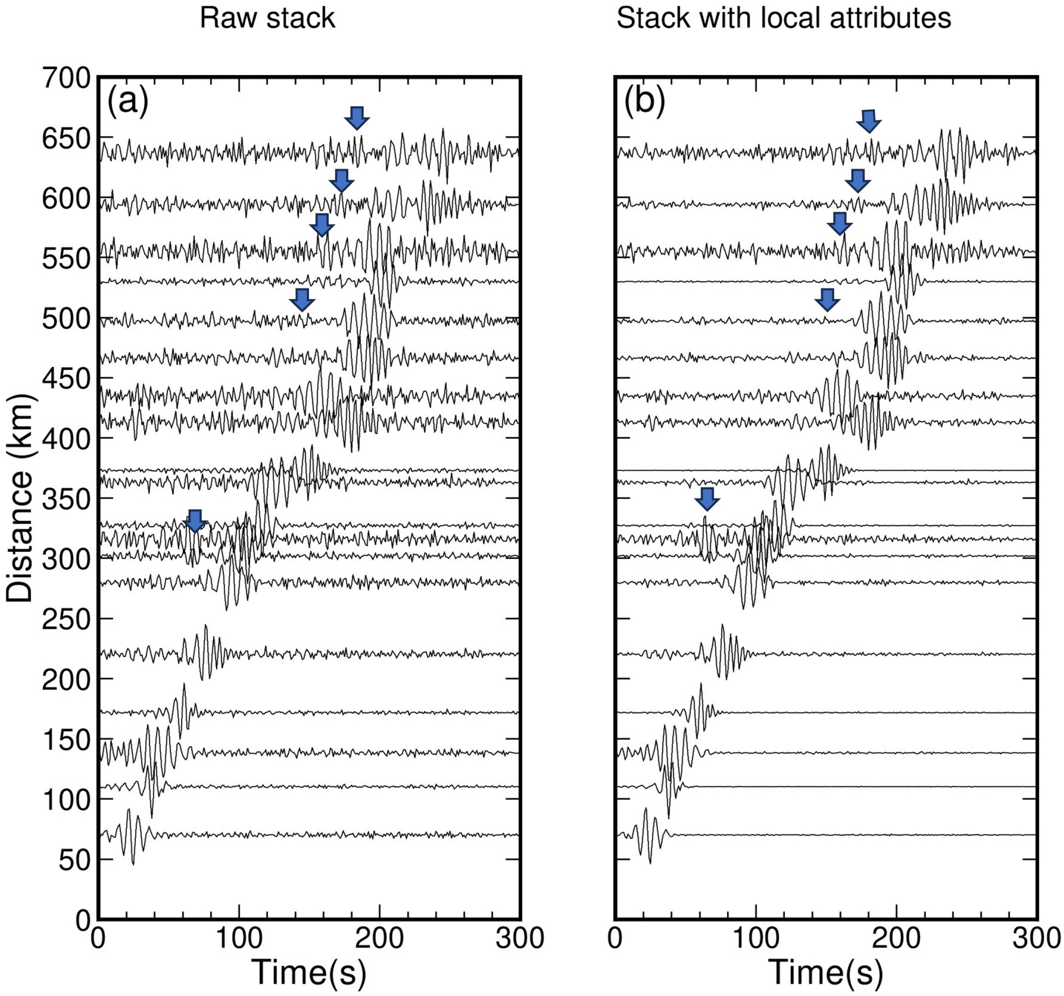


Figure 9. Comparisons of one-month stacked CCFs averaged from negative and positive lags from conventional approach (a) and the proposed procedure (b) for station pairs shown in Figure 8.

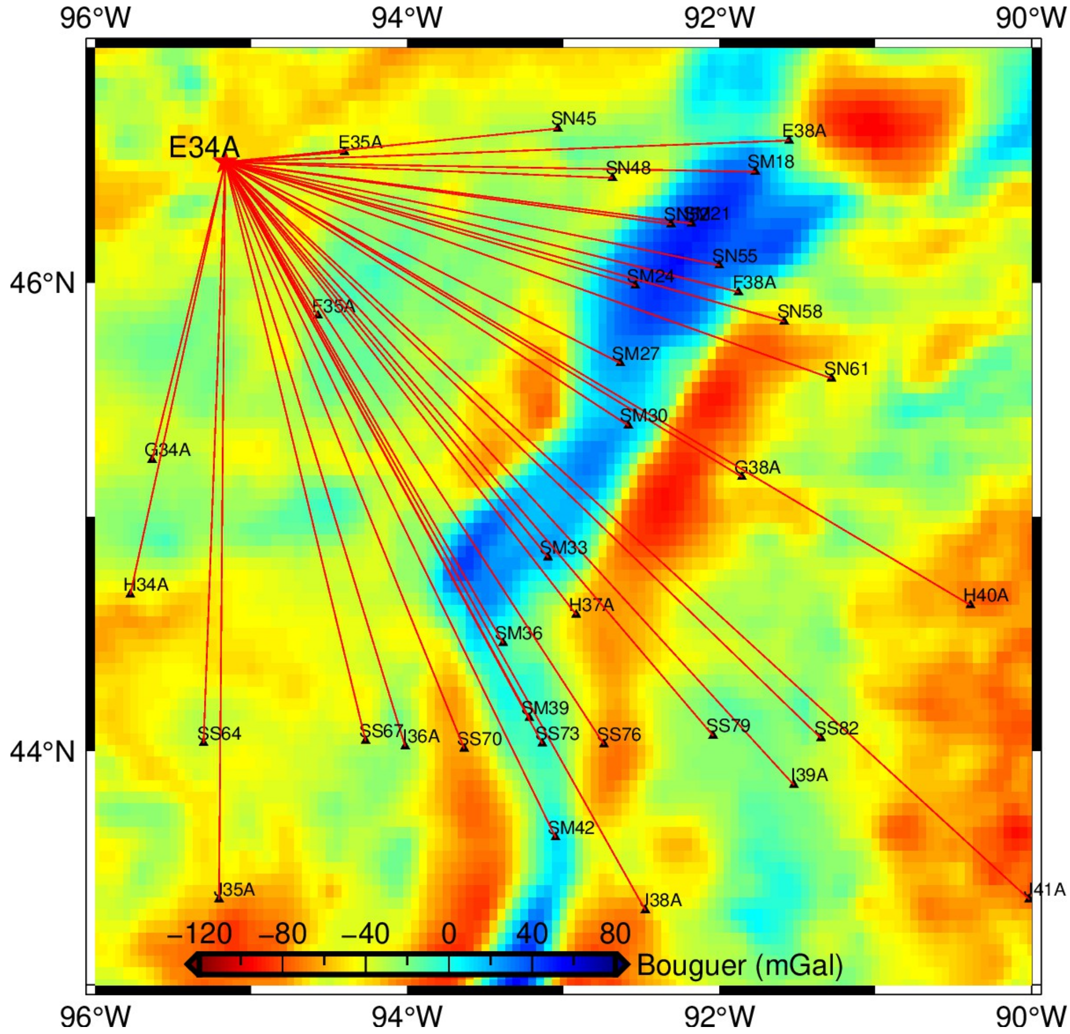


Figure 10. 35 seismic stations used for comparisons shown in Figures 11-13. The SM, SN and SS stations are parts of the Superior Province Rifting Earthscope Experiment (Wolin et al., 2015). Other stations come from the USArray Transportable Array. The background color is the Bouguer gravity anomaly (Woollard, 1965), where the linear feature with positive (blue) values highlights the extension of the Mid-continent Rift.

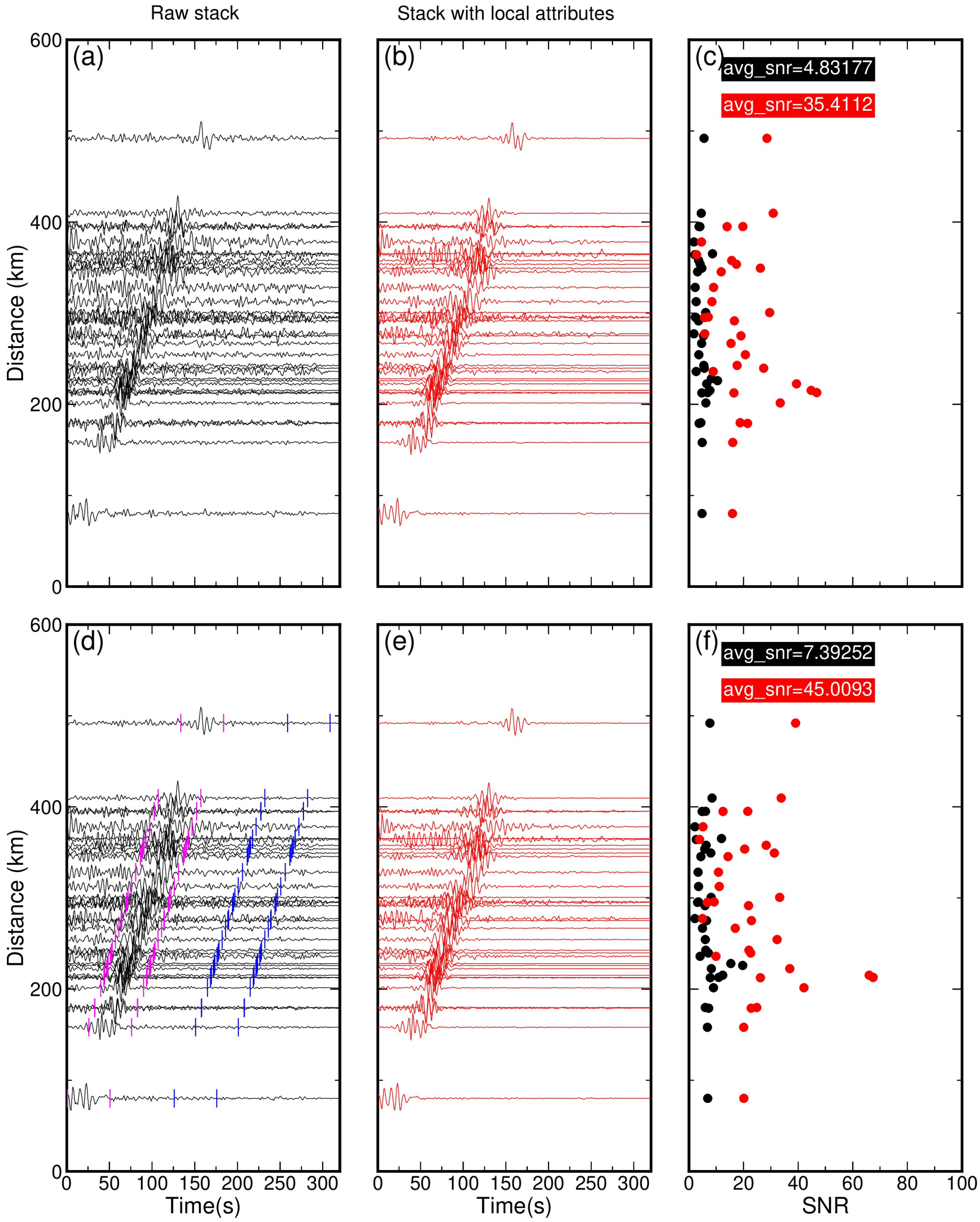


Figure 11. Comparisons of one-month stacked CCFs averaged from negative and positive lags (without bandpass filter) from conventional approach (a) and proposed procedure (b). The improvement of averaged SNR for each trace is shown in panel (c) with black (conventional) and red (proposed) dots. Panels d-f are the same as panels a-c except for two-month stacking. The short magenta and blue solid lines in panel (d) represent the signal and noise windows used to calculate the SNRs shown in panels c and f.

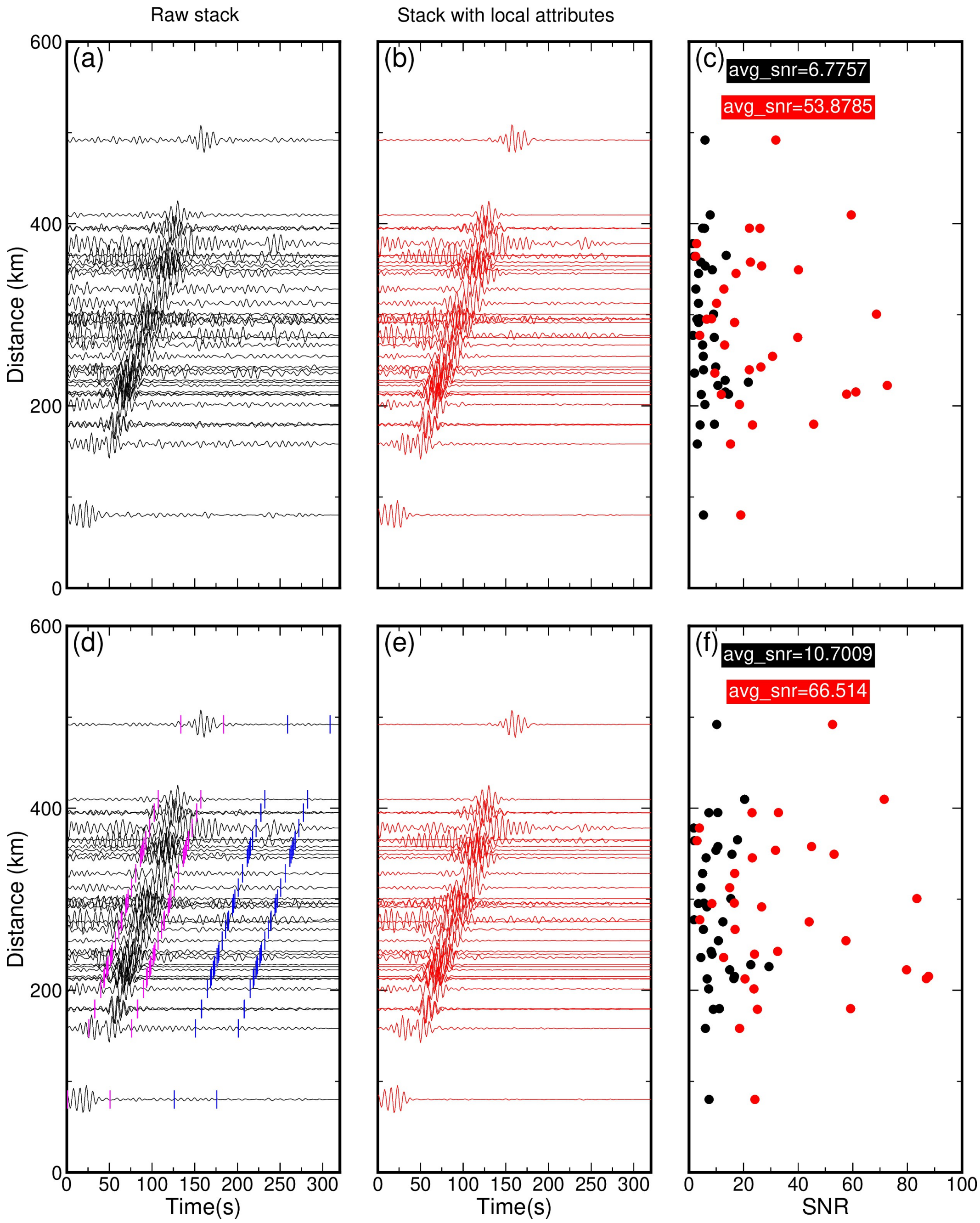


Figure 12. The same setting as Figure 11, but for the stacked data bandpass filtered between 6-15 s.

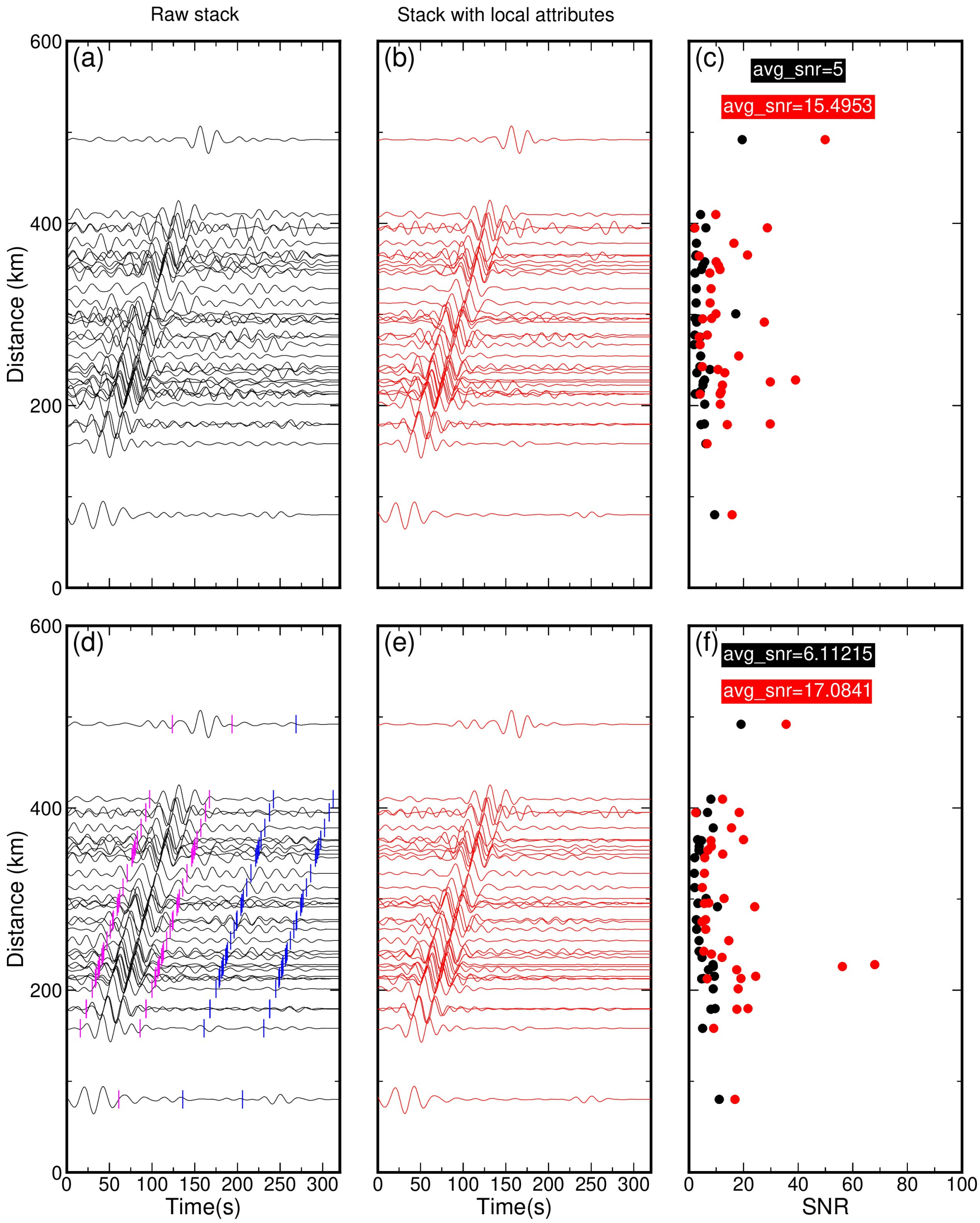


Figure 13. Same as Figure 11 but from the stacked data bandpass filtered between 15-30 s.

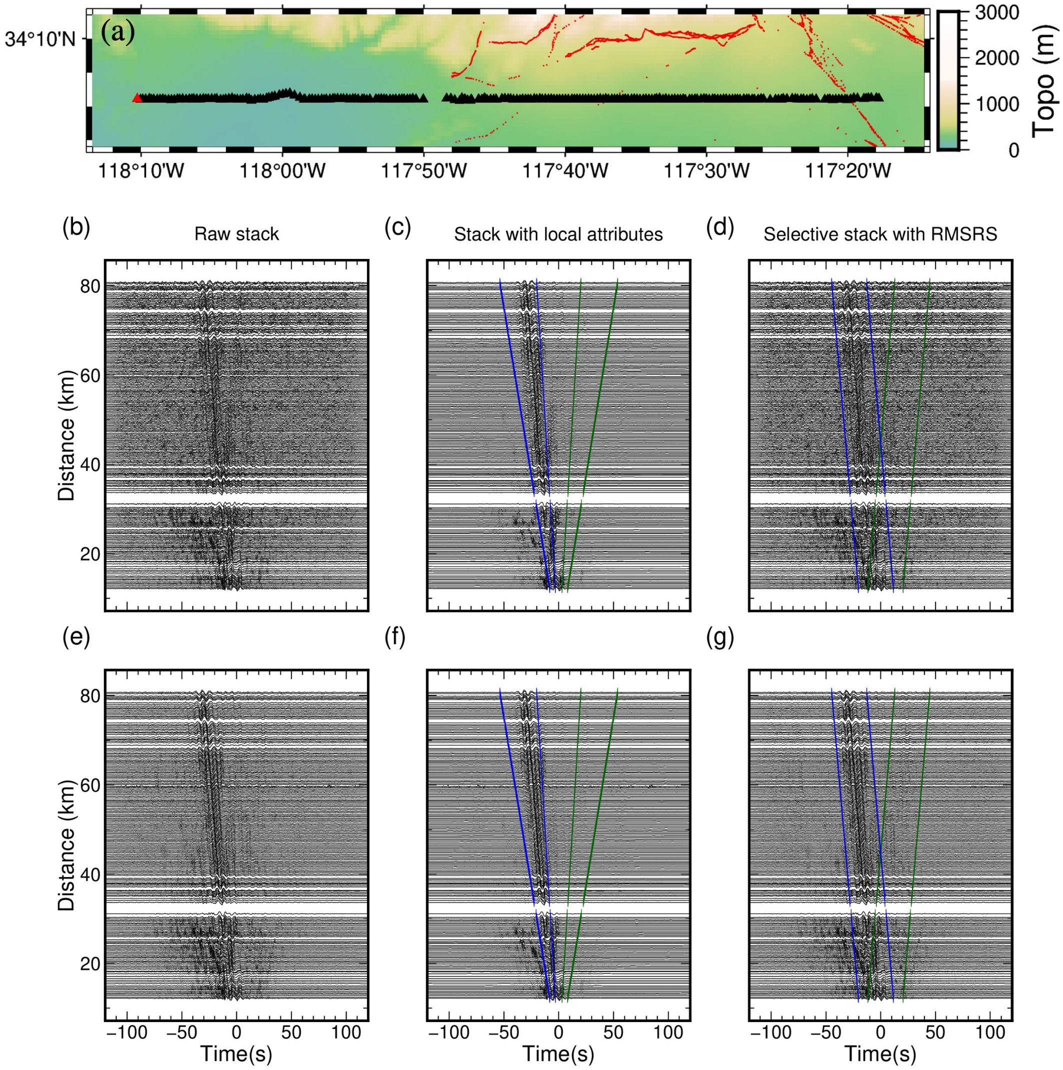


Figure 14. (a) Station distribution of one dense array deployed in the San Bernardino basin. The red triangle denotes the master station while the other black ones are stations used to calculate the CCFs. The red dots represent faults from the U.S. Geological Surveys. Shot gathers of 5 days (b) linear, (e) proposed local and (d) root-mean-square ratio based selective (RMSRS) (Xie et al., 2020) stacking. The blue and dark green lines in panels c and e are acausal and causal arrival times with a group velocity of 1.5 km/s and 3.5 km/s, respectively. They are used to highlight the potential ranges of Rayleigh wave arrivals. The blue and dark green lines in panels d and g are used to define the signal windows for root-mean-square ratio calculation. Panels (e-g) display the corresponding 30-day stacking results.

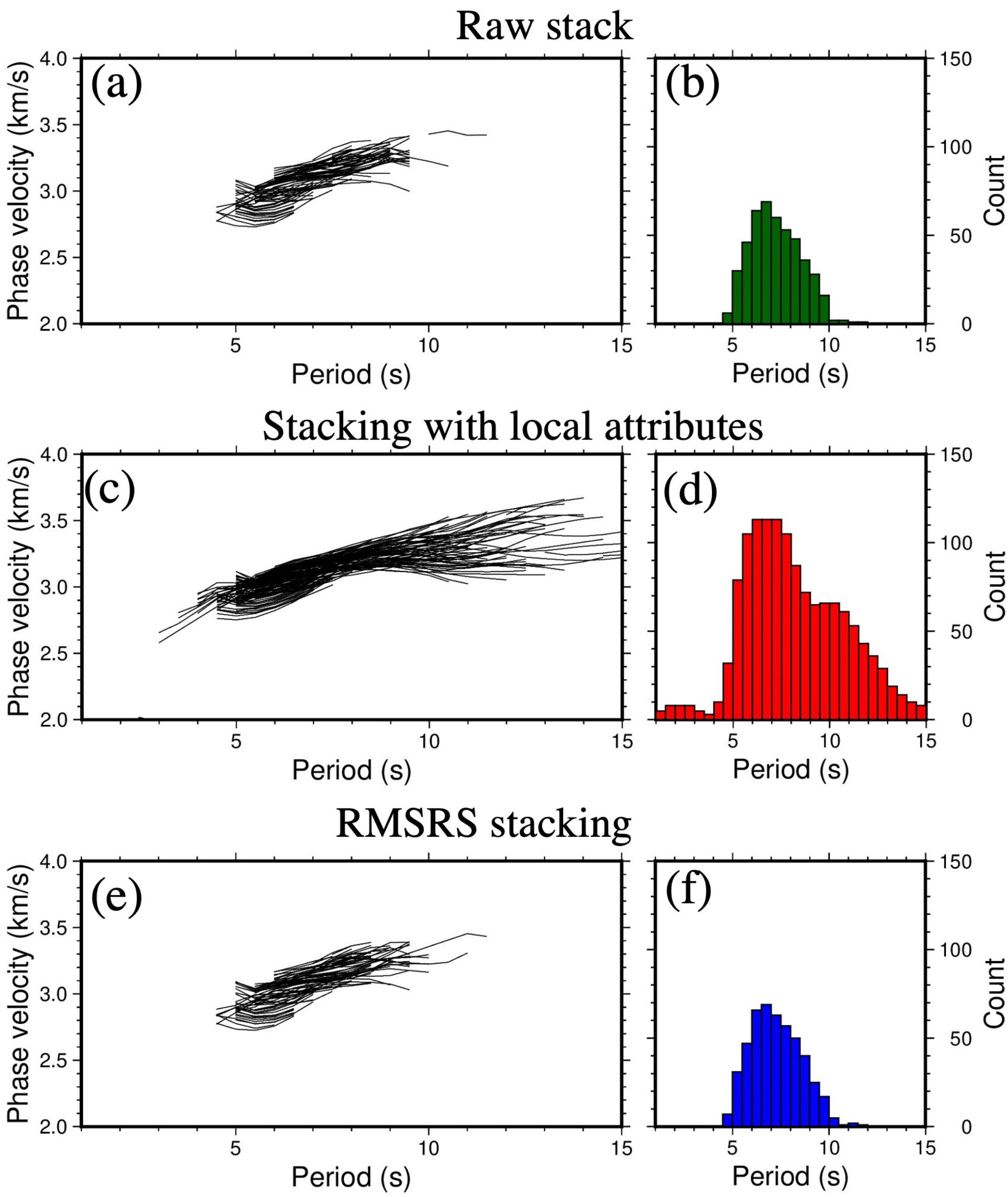


Figure 15. Comparison of selected phase dispersion curves from 5-day raw (a), proposed (b) and RMSRS (c) stacking CCFs displayed in Figures 14b-d. Panels (b), (d) and (f) show the number of measurements for every 0.5 s from 1.0 s to 15.0 s.

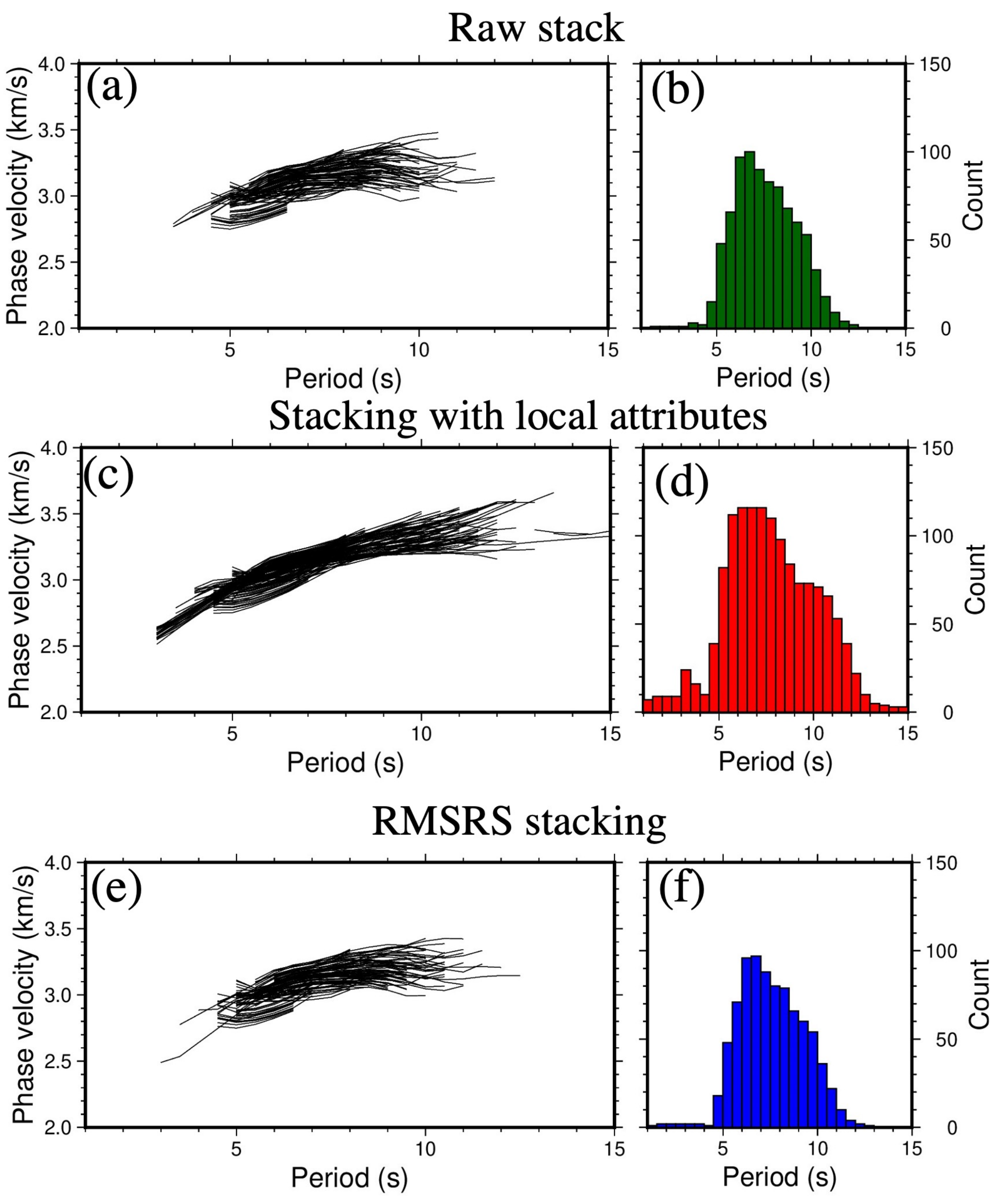


Figure 16. Similar to Figure 15 but from 30-day stacking CCFs displayed in Figures 14e-g

.

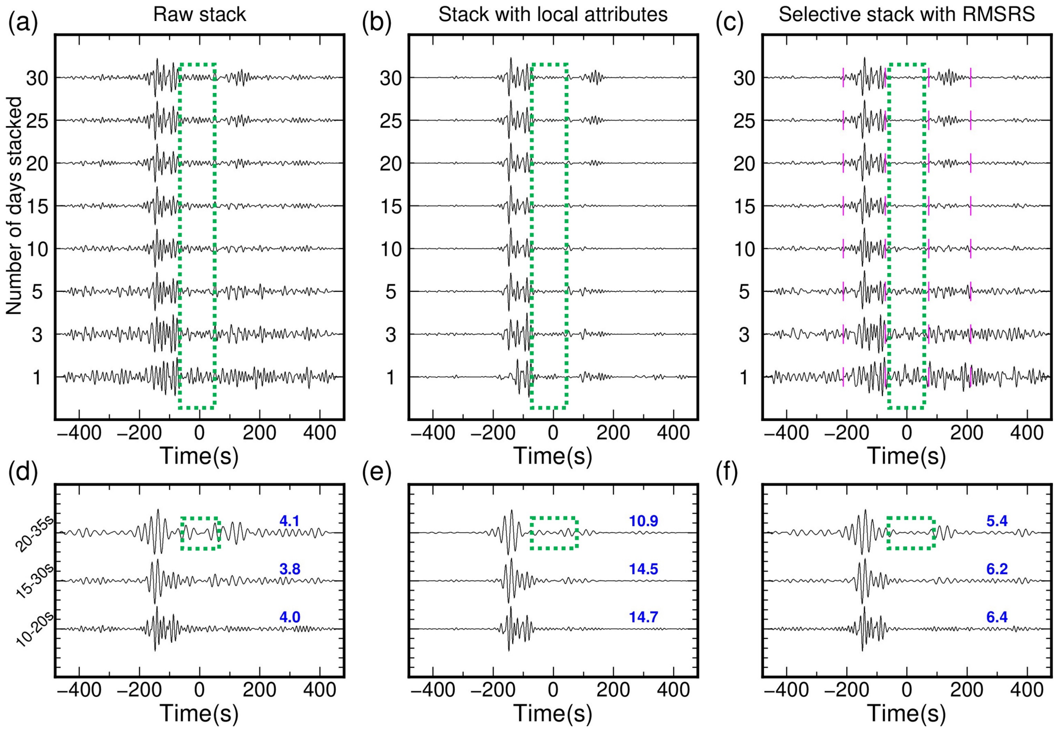


Figure 17. Comparisons of convergence for conventional (a), the proposed procedure (b) and root-mean-square-ratio selective (RMSRS) (Xie et al., 2020) stacking of the TA.A12A-TA.A18A station pair. Panels (d-e) compare 15-day stacked CCFs of the three stacking methods at 10-20, 15-30 and 20-35 s period bands. The blue numbers are the SNR for the corresponding trace. To calculate the SNR, we choose 57-221 s as the signal window and 221-357 s as the noise window. The magenta short lines (57-221 s) are used to define the signal windows for RMSRS calculation. The green dashed rectangles are used to highlight the improvement of the RMSS stacking method. Both RMSRS stacking and our proposed approaches help us to improve the SNR at different period bands.

Figure 1.

(a)
Amplitude
Time (s)
(b)
Lag(s)
Time(s)
Global correlation
(c)
Lag(s)
Time(s)
Local correlation
(d)
Lag(s)
Time(s)
Similarity
(e)
Lag(s)
Time(s)
C2
(f)
Lag(s)
Time(s)
Similarity clipped

Figure 2.

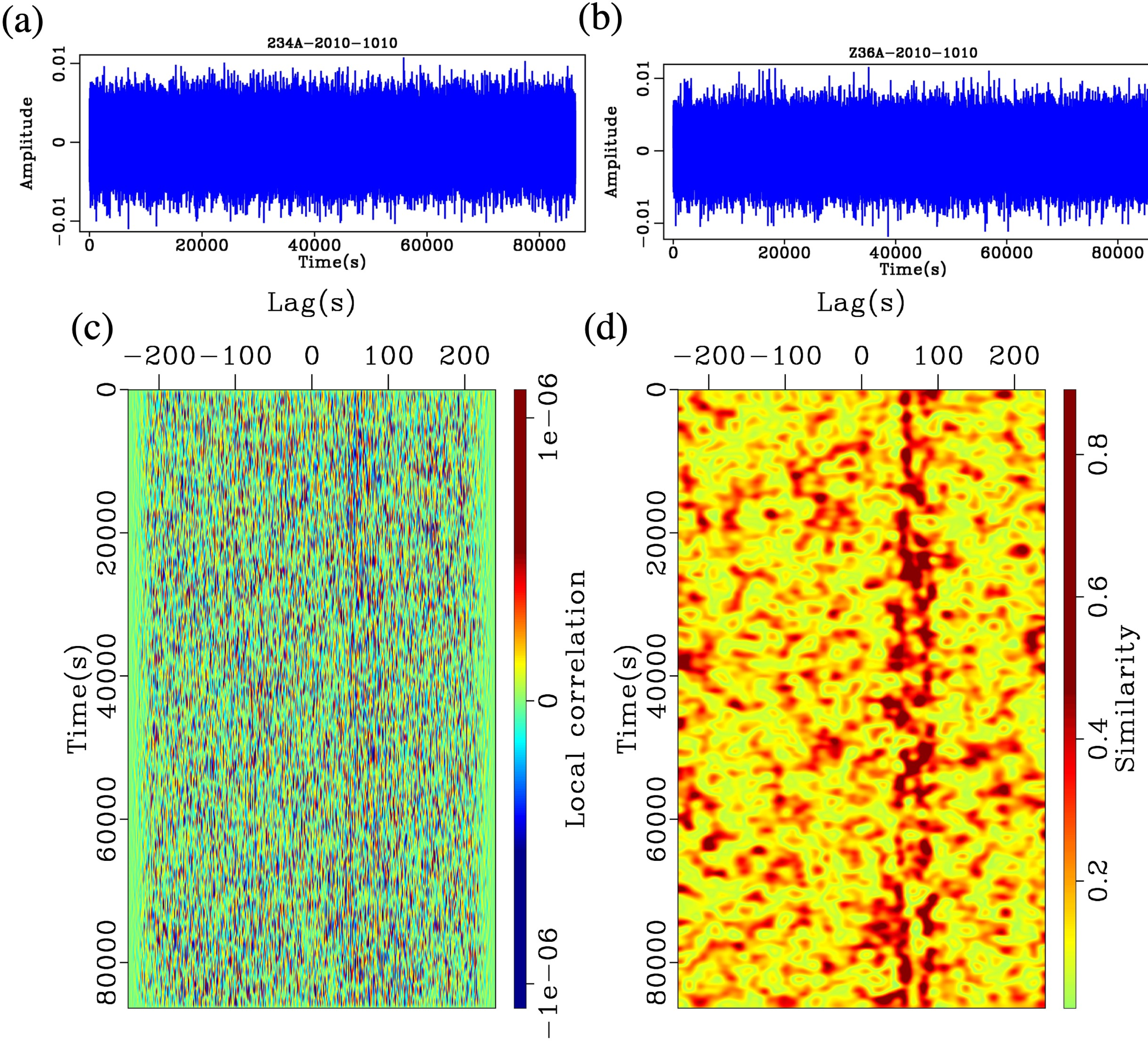

(a)
234A−2010−1010
Amplitude
Time(s)
(b)
Z36A−2010−1010
Amplitude
Time(s)
(c)
Lag(s)
Time(s)
Local correlation
(d)
Lag(s)
Time(s)
Similarity

Figure 3.

(a)
Amplitude
Time(s)
(b)
Amplitude
Frequency(Hz)
(c)
Time(s)
Frequency(Hz)
(d)
Amplitude
Time(s)
(e)
Amplitude
Frequency(Hz)
(f)
Time(s)
Frequency(Hz)

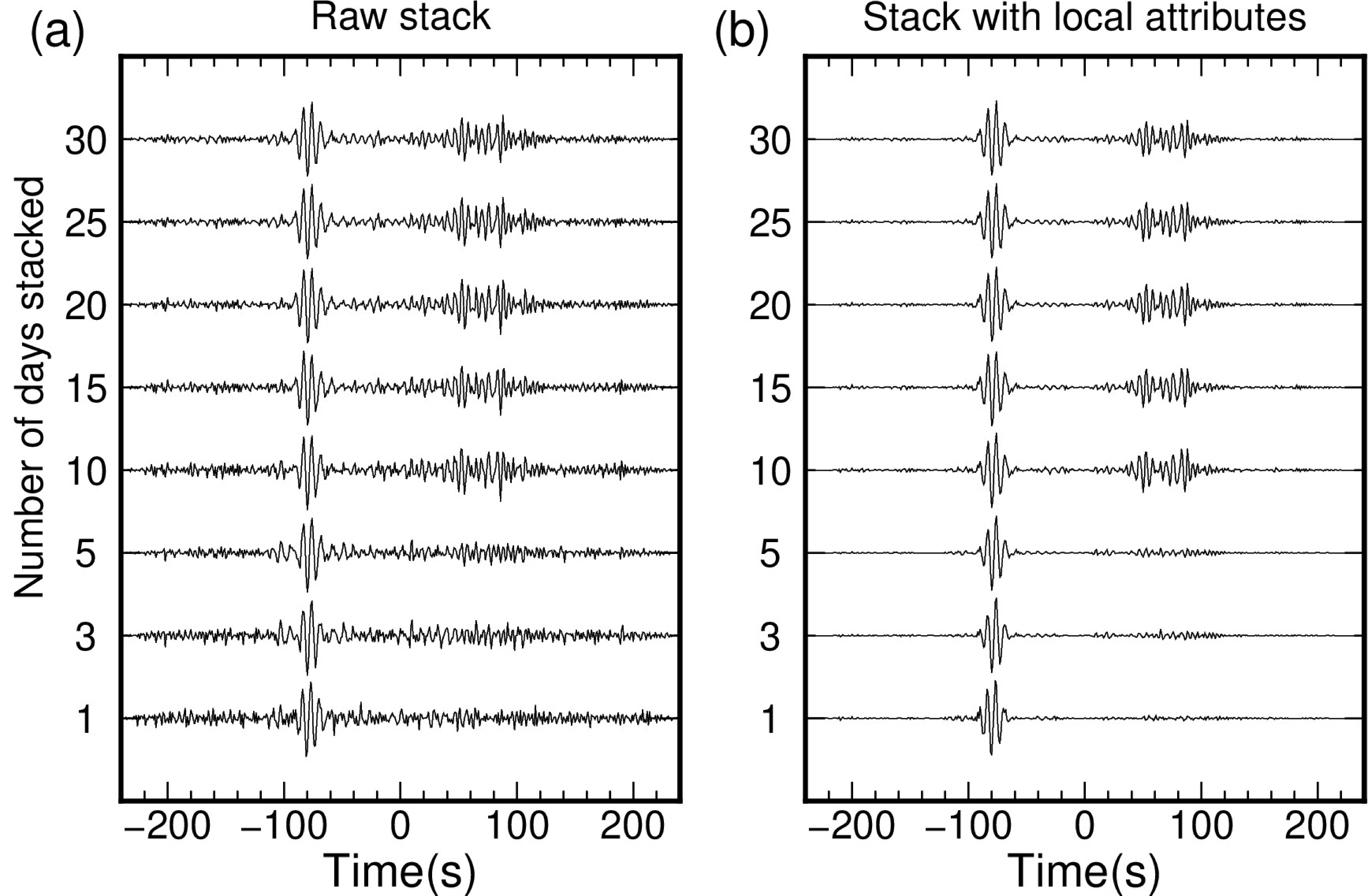
Figure 4.

(a)
Lag(s)
Time(Day)
Global correlation
(b)
Lag(s)
Time(Day)
Local correlation
(c)
Lag(s)
Time(Day)
Similarity

Figure 5.


(a)
Amplitude
Time(s)
(b)
Amplitude
Frequency(Hz)
(c)
Time(s)
Frequency(Hz)
(d)
Amplitude
(e)
Amplitude
Frequency(Hz)
(f)
Time(s)
Frequency(Hz)

Figure 6.



(a)
Raw stack
(b)
Stack with local attributes
Number of days stacked
30
25
20
15
10
5
3
1
−200
−100
0
100
200
Time(s)

Figure 7.

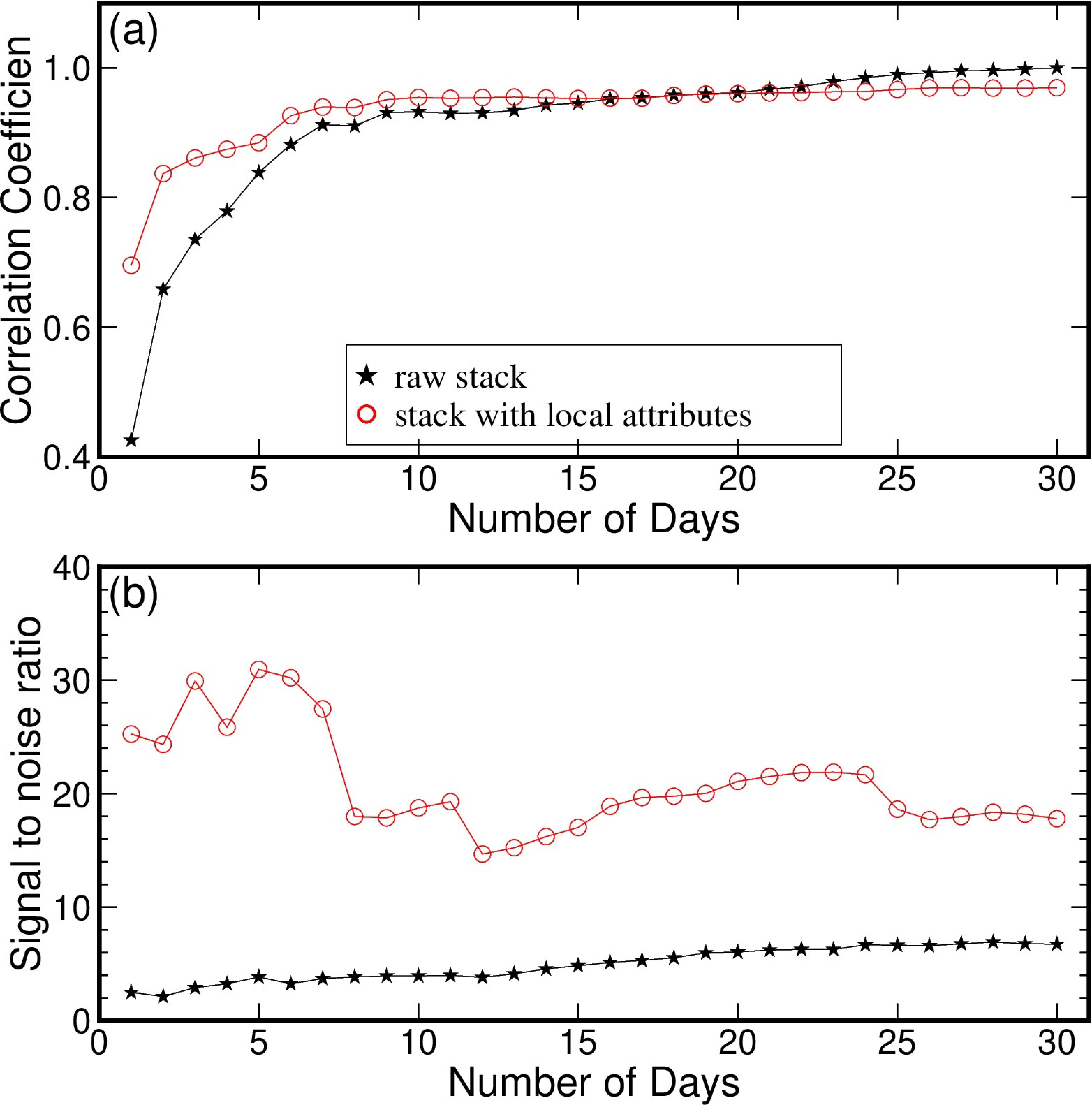

(a)
Correlation Coefficien
raw stack
stack with local attributes
Number of Days
(b)
Signal to noise ratio
Number of Days

Figure 8.

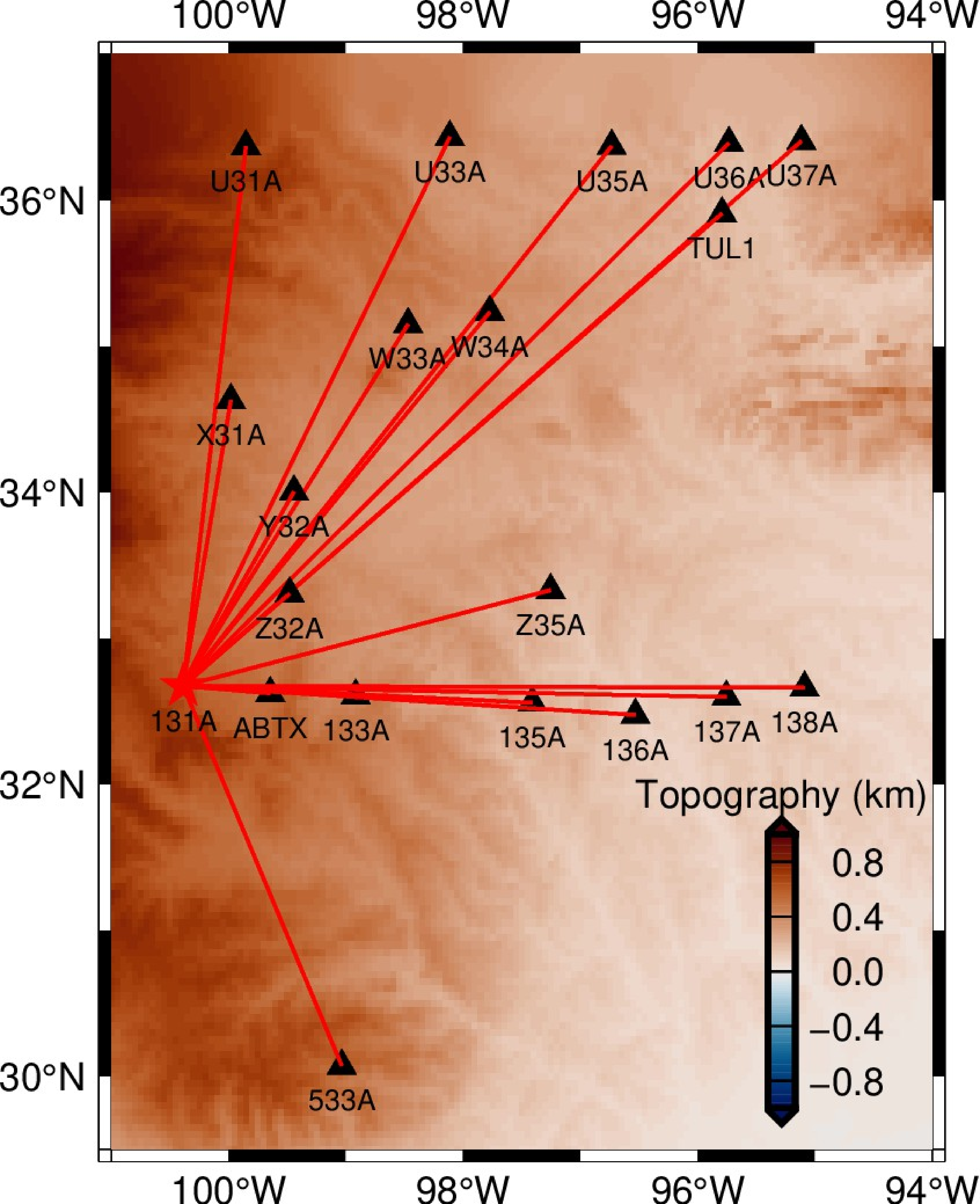
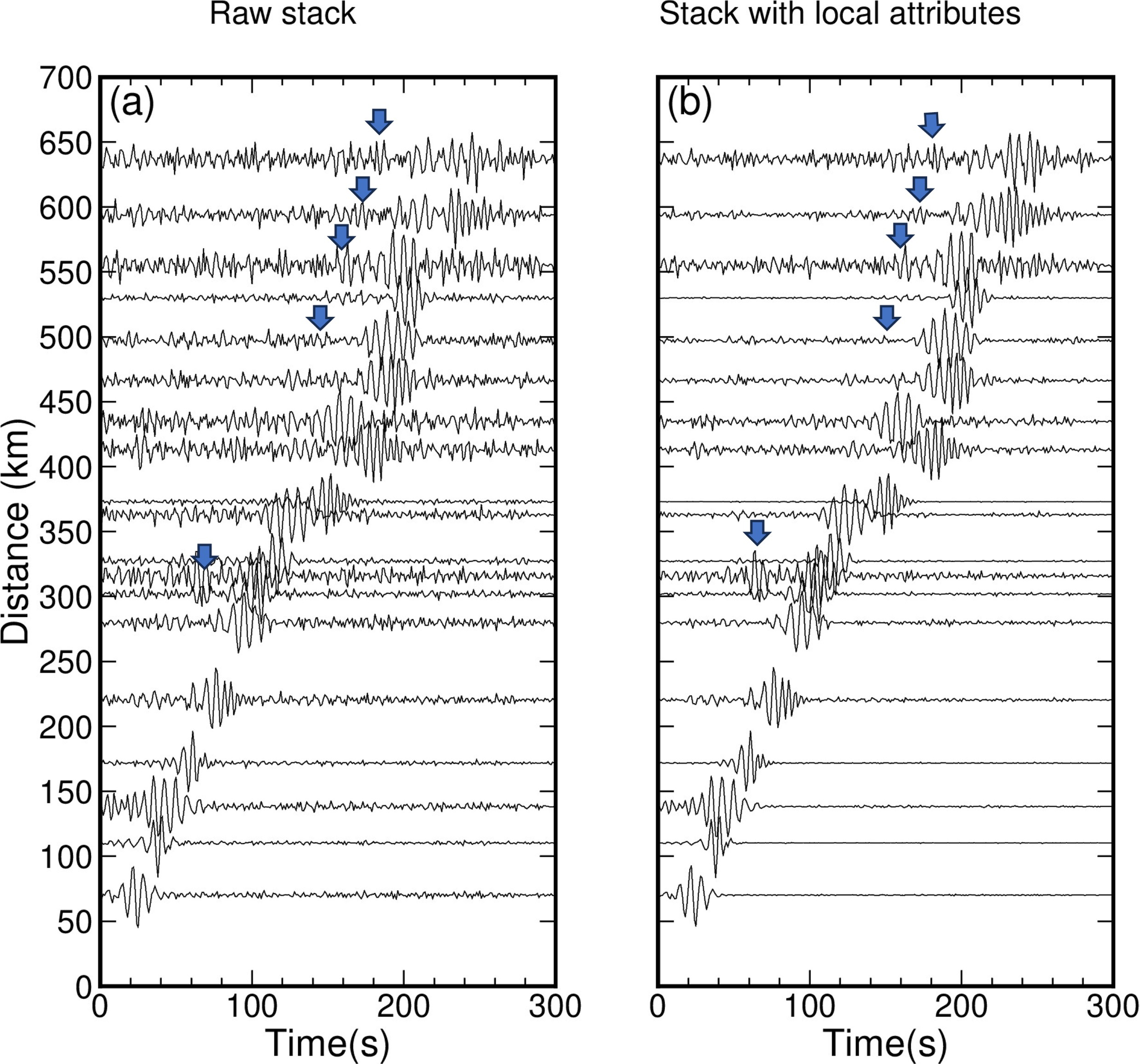

100°W
98°W
96°W
94°W
36°N
34°N
32°N
30°N
U31A
U33A
U35A
U36A
U37A
TUL1
W33A
W34A
X31A
Y32A
Z32A
Z35A
131A
ABTX
133A
135A
136A
137A
138A
533A
Topography (km)
0.8
0.4
0.0
−0.4
−0.8

Figure 9.



Raw stack
Stack with local attributes
(a)
(b)
Distance (km)
Time(s)
700
650
600
550
500
450
400
350
300
250
200
150
100
50
0
0
100
200
300

Figure 10.

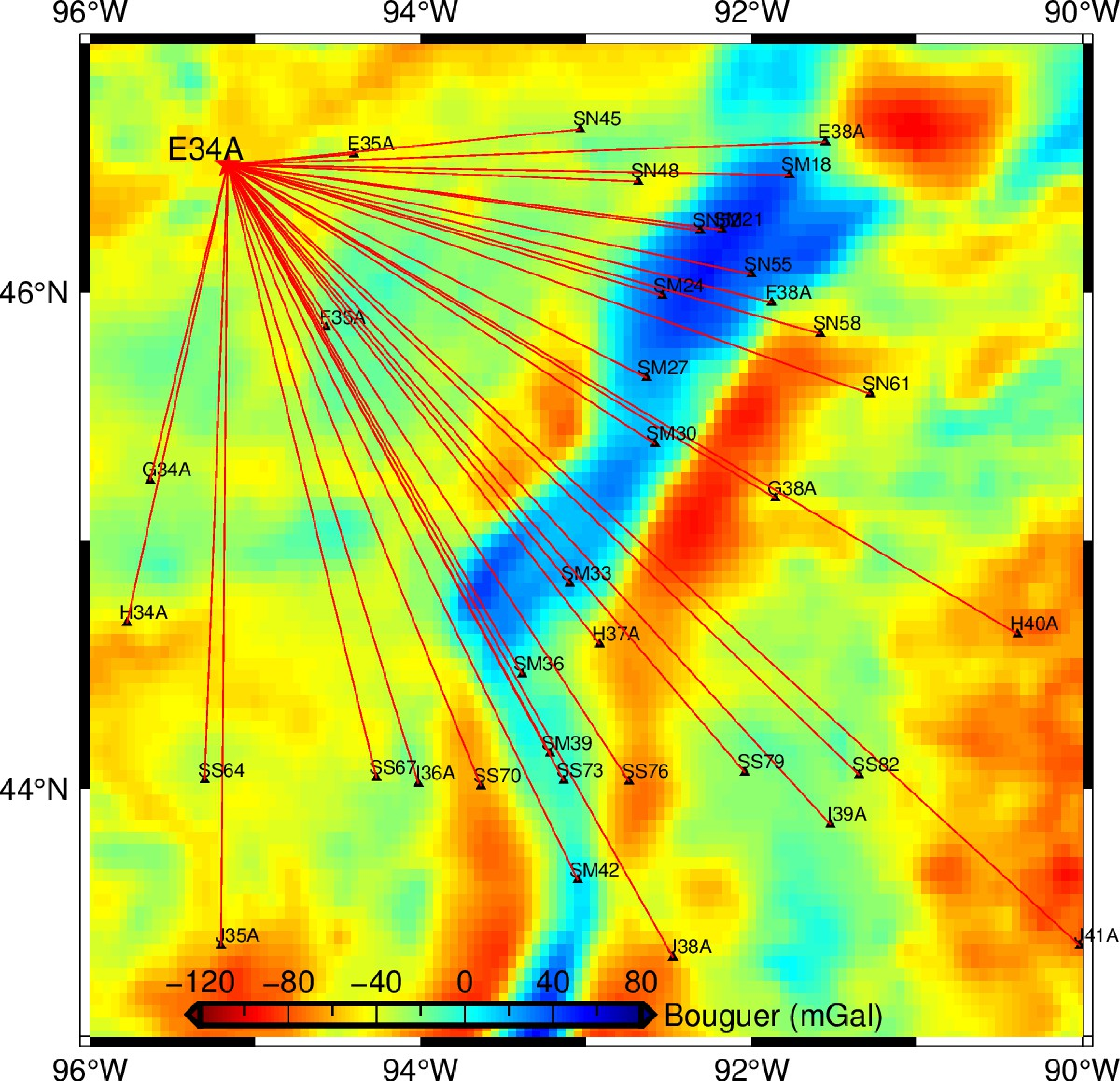
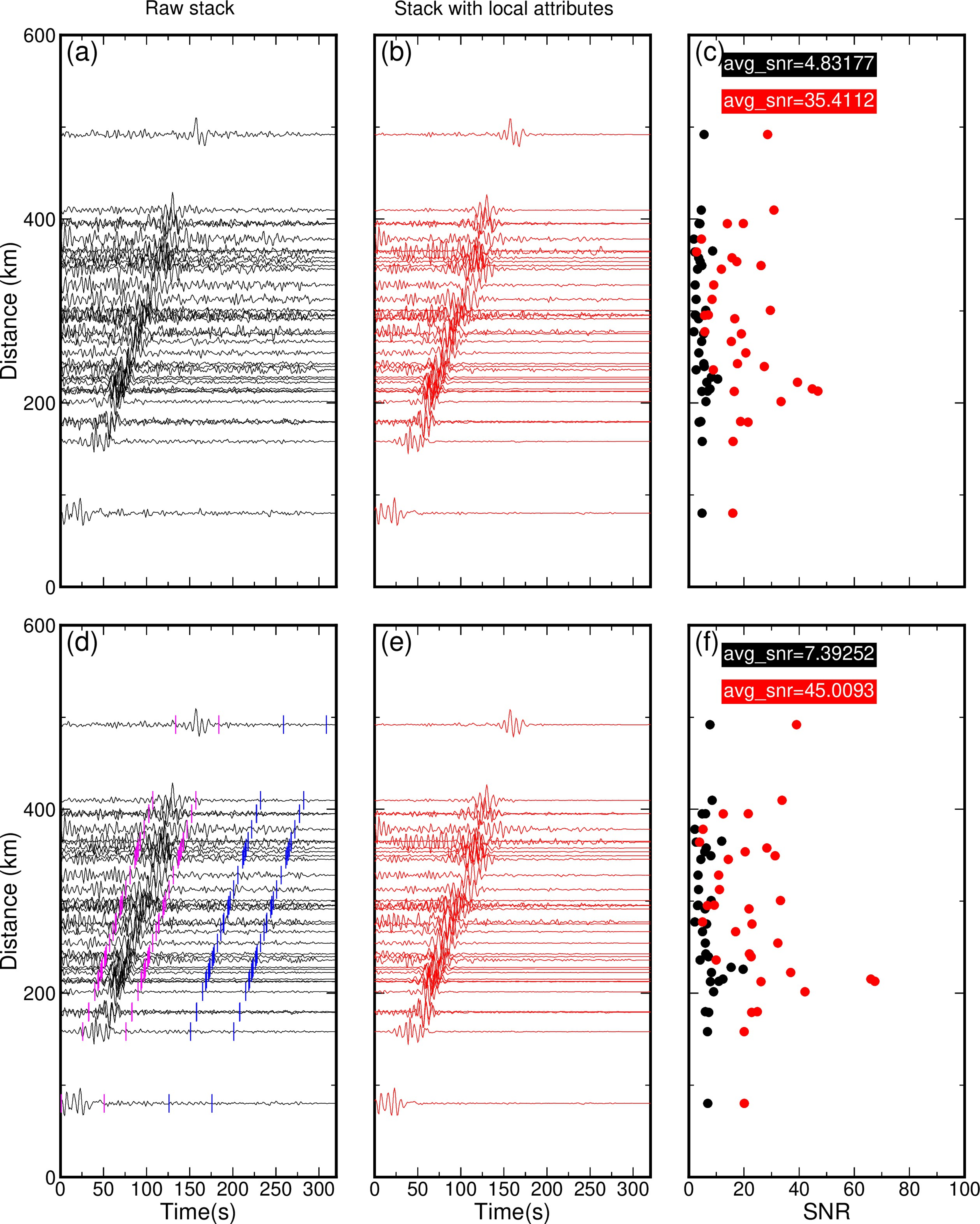
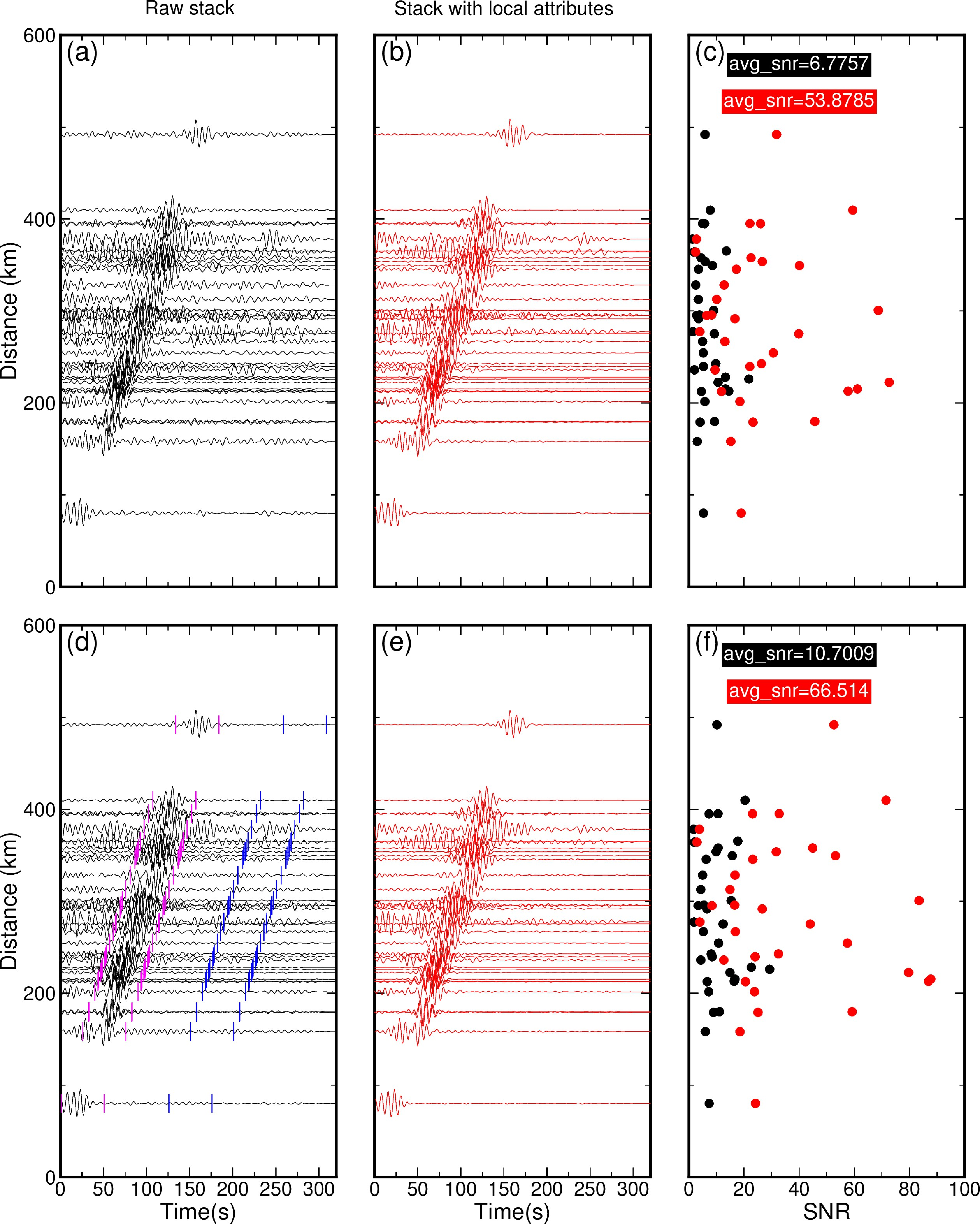

96°W
94°W
92°W
90°W
46°N
44°N
E34A
SN45
E35A
E38A
SN48
SM18
SN55
SM24
F38A
F35A
SN58
SM27
SN61
SM30
G34A
G38A
SM33
H34A
H40A
H37A
SM36
SM39
SS64
SS67
I36A
SS70
SS73
SS76
SS79
SS82
I39A
SM42
I35A
I38A
I41A
−120
−80
−40
0
40
80
Bouguer (mGal)

Figure 11.



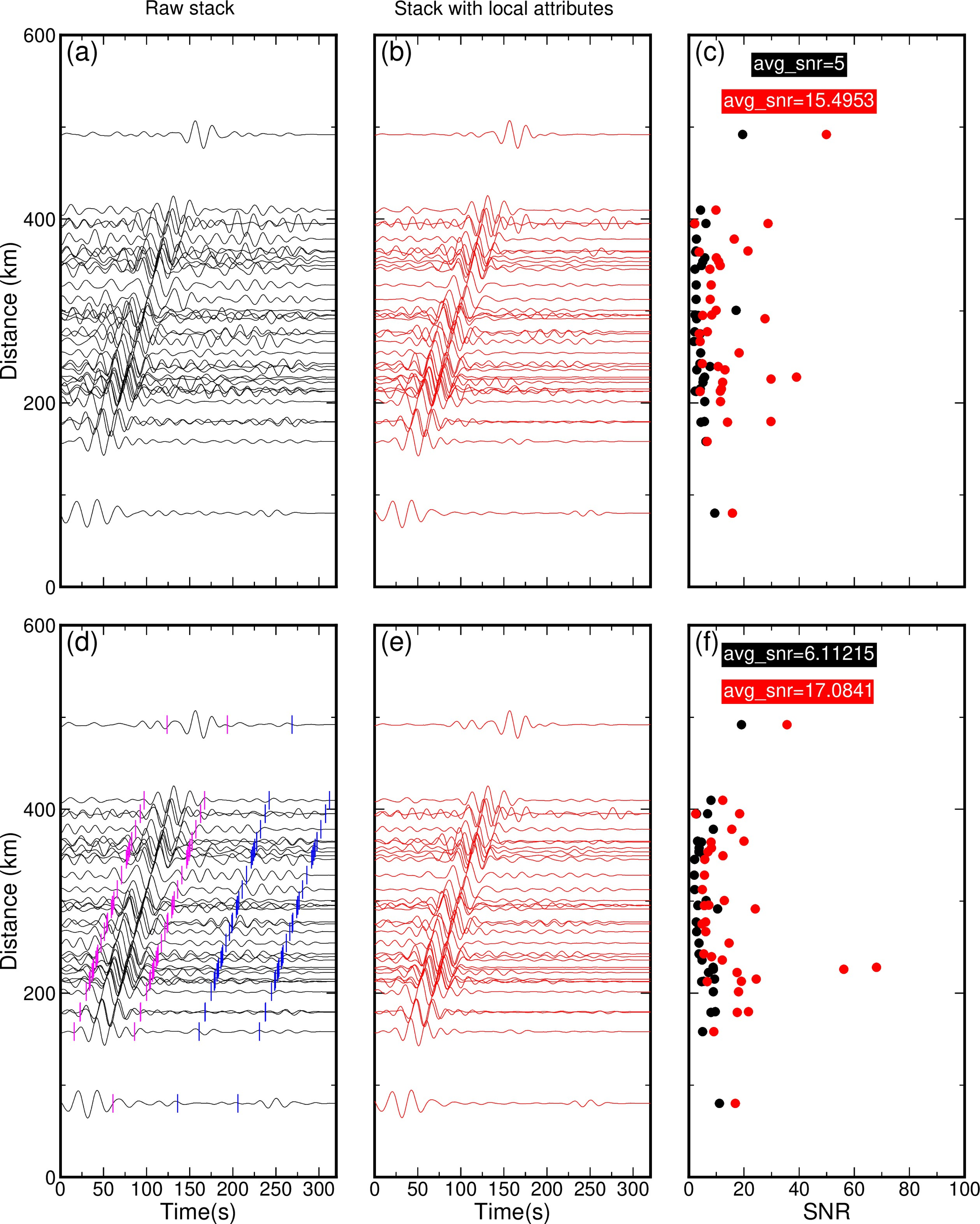
Raw stack
Stack with local attributes
(a)
(b)
(c)
avg_snr=4.83177
avg_snr=35.4112
(d)
(e)
(f)
avg_snr=7.39252
avg_snr=45.0093
Distance (km)
Time(s)
SNR

Figure 12.



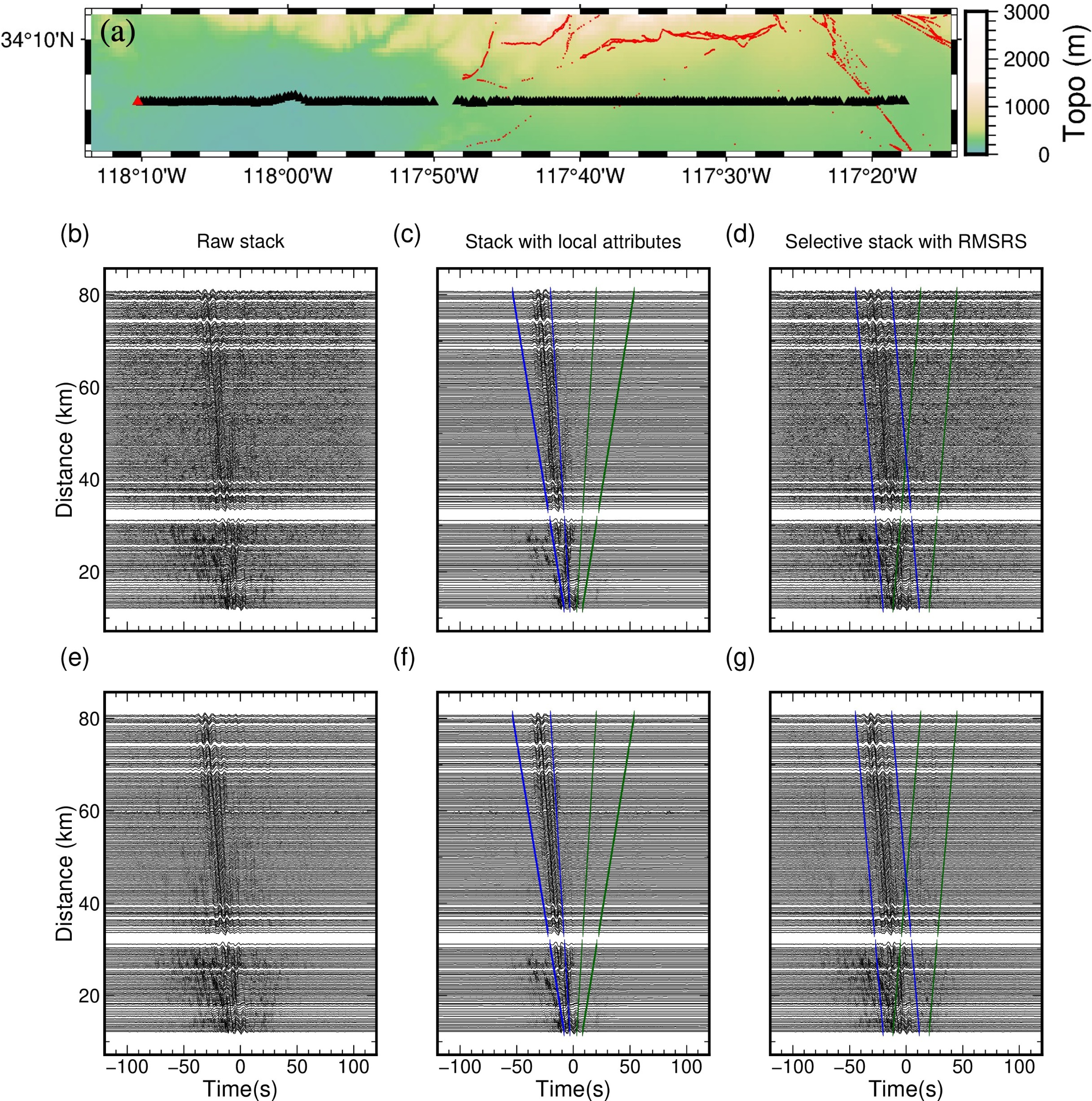
Raw stack
Stack with local attributes
(a)
(b)
(c)
avg_snr=6.7757
avg_snr=53.8785
(d)
(e)
(f)
avg_snr=10.7009
avg_snr=66.514
Distance (km)
Time(s)
SNR
0
200
400
600
50
100
150
250
300
20
40
60
80

Figure 13.



Raw stack
Stack with local attributes
(a)
(b)
(c)
avg_snr=5
avg_snr=15.4953
(d)
(e)
(f)
avg_snr=6.11215
avg_snr=17.0841
Distance (km)
Time(s)
SNR

Figure 14.



(a)
34°10'N
118°10'W
118°00'W
117°50'W
117°40'W
117°30'W
117°20'W
Topo (m)
3000
2000
1000
0
(b)
Raw stack
(c)
Stack with local attributes
(d)
Selective stack with RMSRS
(e)
(f)
(g)
Distance (km)
80
60
40
20
Time(s)
−100
−50
0
50
100

Figure 15.

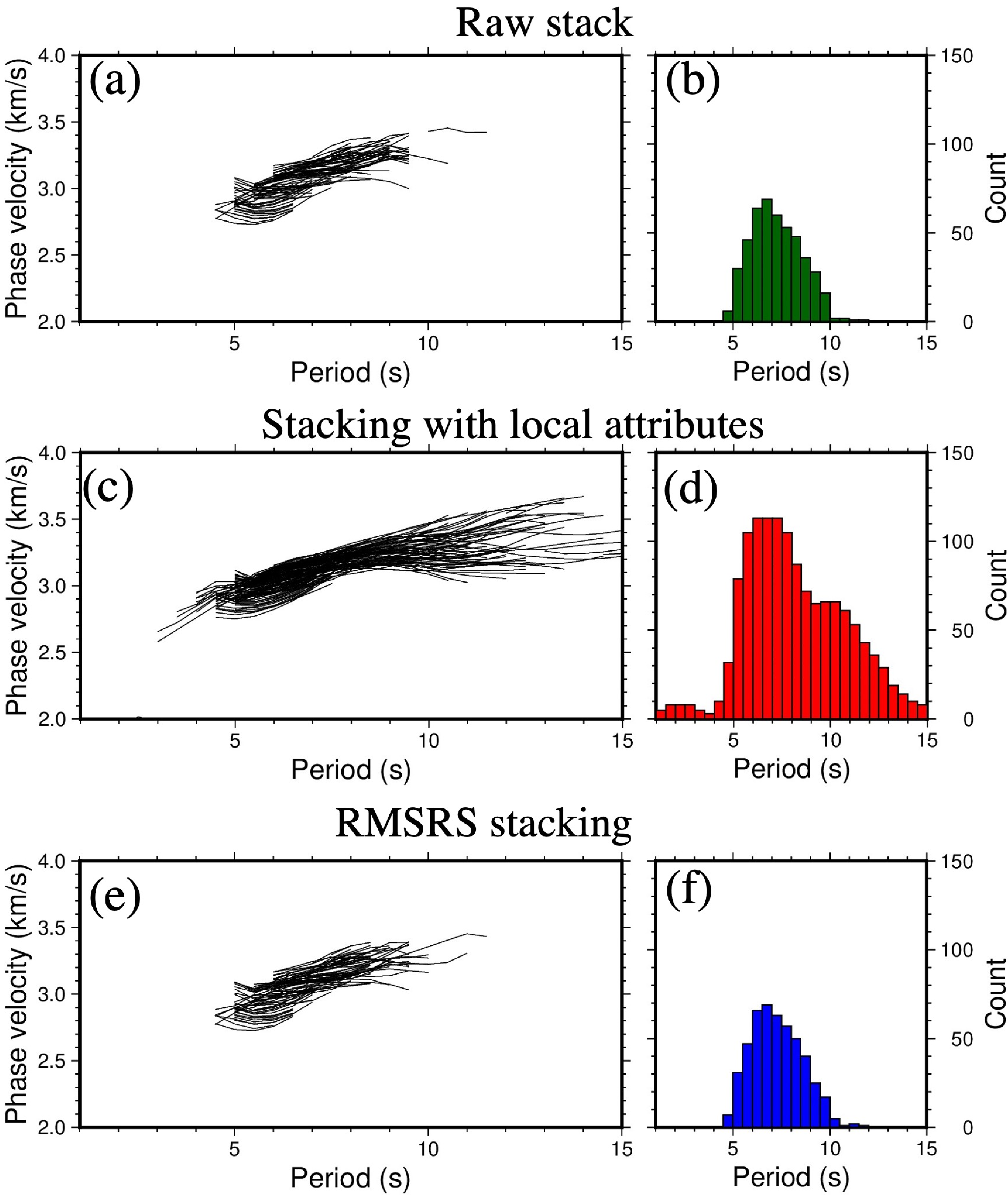

Raw stack
(a)
(b)
Phase velocity (km/s)
4.0
3.5
3.0
2.5
2.0
5
10
15
Period (s)
Count
150
100
50
0
Stacking with local attributes
(c)
(d)
RMSRS stacking
(e)
(f)

Figure 16.

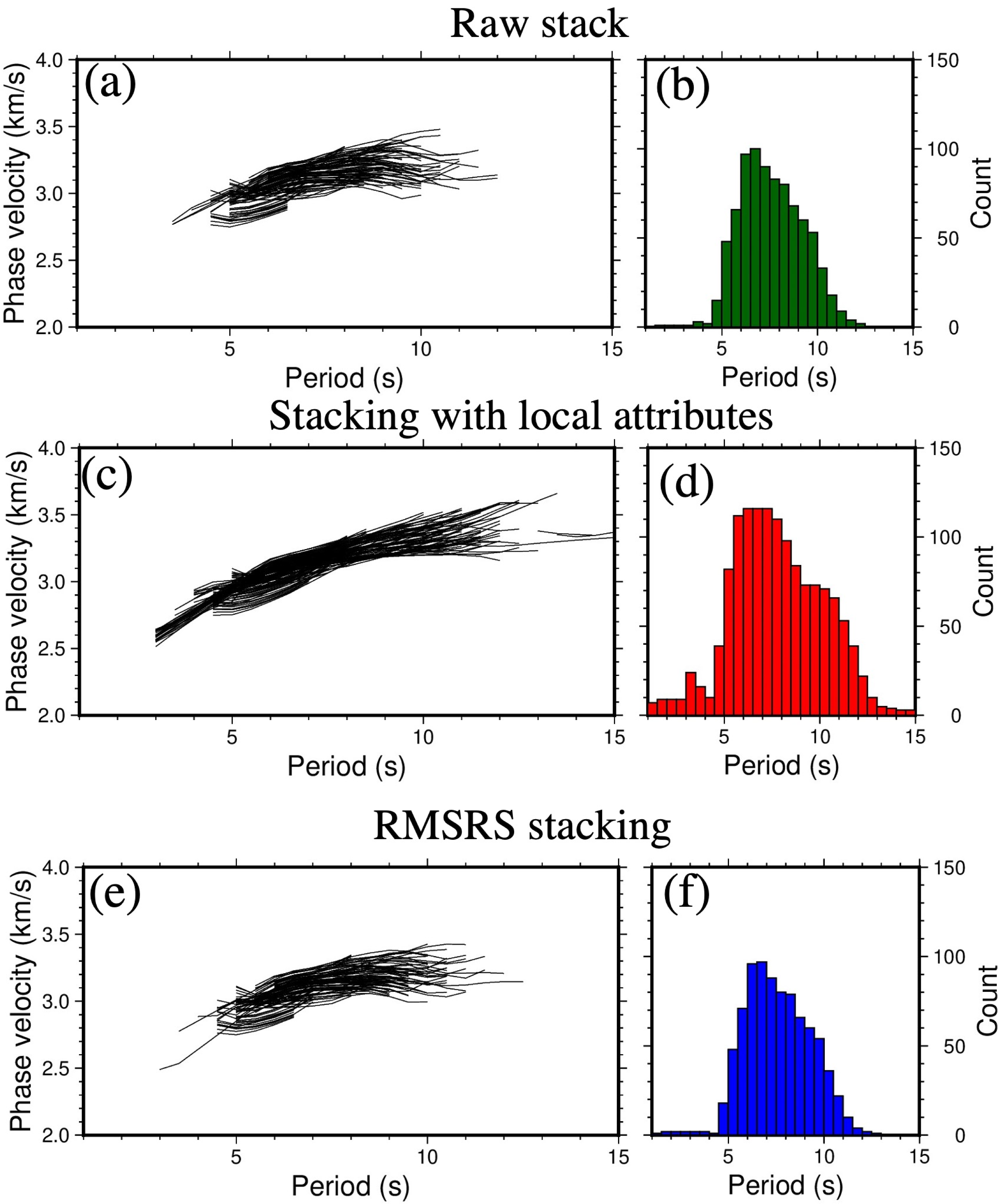

Raw stack
(a)
(b)
Phase velocity (km/s)
Count
Period (s)
Stacking with local attributes
(c)
(d)
RMSRS stacking
(e)
(f)
4.0
3.5
3.0
2.5
2.0
5
10
15
150
100
50
0

Figure 17.


(a)
Raw stack
(b)
Stack with local attributes
(c)
Selective stack with RMSRS
Number of days stacked
30
25
20
15
10
5
3
1
−400
−200
0
200
400
Time(s)
(d)
(e)
(f)
20-35s
15-30s
10-20s
4.1
3.8
4.0
10.9
14.5
14.7
5.4
6.2
6.4

Improving signal-to-noise ratios of ambient noise cross-correlation functions using local attributes

Bin He[1], **Hejun Zhu**[1,2], **David Lumley**[1,2]

[1]Department of Geosciences, The University of Texas at Dallas, Richardson, TX 75080, USA
[2]Department of Physics, the University of Texas at Dallas, Richardson, TX 75080, USA

Key Points:

- Signal-to-noise ratio of ambient noise cross-correlation functions can be improved by using local cross-correlation and local similarity.
- The local cross-correlation function is efficiently implemented by solving the heat equation.
- Applications to broadband and high-frequency nodal arrays validate the effectiveness of the proposed method.

Corresponding author: Hejun Zhu, `hejun.zhu@utdallas.edu`

Abstract

For seismographic stations with short acquisition duration, the signal-to-noise ratios (SNRs) of ambient noise cross-correlation functions (CCFs) are typically low, preventing us from accurately measuring surface wave dispersion curves or waveform characteristics. In addition, with low-quality CCFs, it is difficult to monitor temporal variations of subsurface physical states or extract relatively weak signals such as body waves. In this study, we propose to use local attributes to improve the SNRs of ambient noise CCFs, which allows us to enhance the quality of CCFs for stations with limited acquisition duration. Two local attributes: local cross-correlation and local similarity, are used in this study. The local cross-correlation allows us to extend the dimensionality of daily CCFs with computational costs similar to global cross-correlation. Taking advantage of this extended dimensionality, the local similarity is then used to measure non-stationary similarity between the extended daily CCFs with a reference stacking trace, which enables us to design better stacking weights to enhance coherent features and attenuate incoherent background noises. Ambient noise recorded by several broadband stations from the USArray in North Texas and Oklahoma, the Superior Province Rifting EarthScope Experiment in Minnesota and Wisconsin and a high-frequency nodal array deployed in the San Bernardino basin are used to demonstrate the performance of the proposed approach for improving the SNR of CCFs.

Plain Language Summary

Seismic ambient noise has been widely used for imaging and monitoring subsurface structures by using cross-correlation functions (CCFs) of continuous recordings between station pairs. Typically, we have to stack a lot of CCFs to enhance signals (e.g., surface and body waves) of CCFs. However, for temporal monitoring purposes or those deployed arrays with short acquisition duration, it is impossible to have a lot of stacking. The lack of enough stacking could result in low signal-to-noise ratios (SNRs) of the CCFs. We proposed a new approach to improve the SNRs of CCFs by using two local attributes: local cross-correlation and local similarity. We first extend the CCFs to a higher dimension by using local cross-correlation, and then the local similarity is used to define a better weighting factor for the final stacking. We use several field data examples to prove the effectiveness of the proposed approach.

1 Introduction

With the assumption of homogeneously distributed noise sources, both theoretical and experimental studies have demonstrated that empirical Green's functions can be retrieved by cross-correlating continuous ambient noise records between two seismographic stations (Aki, 1957; Claerbout, 1968; Buckingham et al., 1992; Lobkis & Weaver, 2001; Weaver & Lobkis, 2001; Shapiro & Campillo, 2004; Wapenaar, 2004; Nakata et al., 2019). Both surface (Campillo & Paul, 2003; Shapiro & Campillo, 2004; Sabra et al., 2005) and body wave signals (Draganov et al., 2009; Zhan et al., 2010; Poli et al., 2012; Lin et al., 2013; Nakata et al., 2015) have been successively extracted from ambient noise recordings. With empirical Green's functions between each pair of stations in a seismic array, we are able to measure surface wave phase or group dispersion curves, and then perform surface wave tomography to estimate seismic properties in the subsurface. Compared with earthquake tomography, ambient noise tomography allows us to image tectonically inactive regions and achieve better cross-path coverages. To date, ambient noise cross-correlation functions (CCFs) have been successively used to investigate velocity as well as anisotropic structures within the crust and uppermost mantle (Shapiro et al., 2005; Y. Yang et al., 2007; Lin et al., 2008; Moschetti et al., 2010; Yao et al., 2010; Huang et al., 2010; Lin et al., 2011; Shen & Ritzwoller, 2016). Recently, ambient noise CCFs are combined with waveform inversion to better constrain subsurface velocity structures (Gao & Shen, 2014; M. Chen et al., 2014; Lee et al., 2014; Y. Liu et al., 2017; Sager et al., 2018; K. Wang et al., 2018; Zhu, 2018; Sager et al., 2020; Fan et al., 2022; Maguire et al., 2022). Taking advantage of continuous records, ambient noise CCFs can also be utilized to monitor temporal evolution of physical states in the subsurface (Brenguier et al., 2008; Nakata & Snieder, 2011; Hadziioannou et al., 2011; Mainsant et al., 2012; De Plaen et al., 2016; Q.-Y. Wang & Yao, 2020; Le Breton et al., 2021; Mao et al., 2022).

Previous studies have demonstrated that we are able to improve the signal-to-noise ratios (SNRs) of ambient noise CCFs as well as reducing the effect of source directionality by increasing the stacking duration of signals (Bensen et al., 2007). To obtain high-quality CCFs, typically we have to use long continuous records and sometimes have to stack data with several years of acquisition. However, for some temporary experiments e.g., Brenguier et al. (2008); Issa et al. (2017); G. Liu et al. (2018); Dougherty et al. (2019); G. Chen et al. (2023); Wu et al. (2023), the SNRs of ambient noise CCFs are typically low, preventing us to measure robust surface wave dispersion curves or waveform characteristics. For these cases, we have to design effective approaches to improve the SNRs of ambient noise CCFs. In addition, the SNRs are crucial for extracting weak coherent signals, such as body waves (Snieder, 2004; Nakata et al., 2015, 2016). There are several studies towards improving the SNR of ambient noise CCFs. For instance, Baig et al. (2009) designed better-stacking weights in the time-frequency domain with discrete orthogonal S transform, which allowed them to measure robust Rayleigh and Love arrivals for stations with long offsets. G. Li et al. (2018) further compared the performances of time and frequency domain inverse S transforms for this stacking procedure. Schimmel et al. (2011) used instantaneous phase coherence to avoid strong amplitude arrivals, such as earthquake signals, and enhance coherent features in noise records. Seats et al. (2012) and Clarke et al. (2011) utilized overlapped moving windows (Welch's method) to improve the convergence of CCFs towards stable Green's functions. In addition, wavelet and curvelet transforms have also been applied to denoise CCFs by Stehly et al. (2011) and Mao et al. (2022). Furthermore, Weaver and Yoritomo (2018) proposed several schemes to choose optimal weights for stacking so that the effective incident intensity distribution is closer to isotropic. Xie et al. (2020) used the root-mean-square ratio to remove those CCFs with low SNR. A systematic evaluation and comparison of the performance of several stacking methods is discussed by X. Yang et al. (2023).

In this study, we propose to use local attributes (Rickett & Lumley, 2001; Hale, 2006; Fomel, 2007a) to denoise CCFs when we have short acquisition durations. In comparison with global attributes, the local attributes enable us to extract non-stationary characteristics in seismic data. There are a variety of local attributes, including local similarity (Fomel, 2007a), local cross-correlation (Hale, 2006), local frequency (Fomel, 2007a) and local skewness (Fomel & van der Baan, 2014), etc. In this study, we use local cross-correlation (Hale, 2006) to extend the dimensionality of daily CCFs, and then use stacking weights measured by local similarity to improve the stacking quality of CCFs. The local cross-correlation can be considered as a natural extension of the Welch's method (Seats et al., 2012) but with longer overlap windows and much lower computational costs. Its overall computational cost is similar to the conventional global cross-correlation. The local similarity has been used to measure non-stationary similarity between time-lapse images (Fomel & Jin, 2009), improve stacking quality of normal moveout data (G. Liu et al., 2009) as well as angle-domain common-imaging-gathers (Lin et al., 2011). All these studies demonstrate the capability of using local similarity to improve the stacking quality.

We first review local cross-correlation and local similarity. Next, they are used to improve the SNRs of ambient noise CCFs with one-day and one-month acquisition durations. Continuous noise records from several broadband USArray stations in North Texas and Oklahoma, the Superior Province Rifting EarthScope Experiment in Minnesota and a high-frequency nodal array in the San Bernardino basin are used to demonstrate the performance of the proposed approach.

2 Method

2.1 Local cross-correlation

The global cross-correlation function $c(\tau)$ between two signals $f(t)$ and $g(t)$ can be defined (Nakata & Snieder, 2011; Harris et al., 2020) as follows:

$$c(\tau) = \int_{-\infty}^{\infty} f(t+\tau)g(t)dt \quad . \tag{1}$$

The result of the global cross-correlation is a function of time lag τ, which cannot capture non-stationary time shifts between two input signals and therefore makes it difficult to separate signals and noises. In order to fulfill this goal, one way is to perform time-windowed cross-correlation, which has much higher computational costs in comparison to Equation 1. In addition, there might be leakage and edge problems due to the selected window functions. To solve these problems, Hale (2006) designed an efficient algorithm to compute the local CCF and measure time-varying correlations between two signals. It was first used to measure non-stationary warpings between time-lapse migration images. Later, the local cross-correlation was applied to measure non-stationary travel time differences between two seismograms, which can be utilized to constrain subsurface velocity structures through full-waveform inversion (Díaz & Sava, 2015).

The idea of local cross-correlation is similar to the moving window cross-correlation for capturing transient time shifts, but with much higher computational efficiency. In addition, it helps us to separate signal and noise in a higher dimension. Taking advantage of several properties of Gaussian windows, such as the product of two Gaussian windows is still a Gaussian function, the computational cost of local cross-correlation can be reduced as similar to that of global cross-correlation. For instance, the Gaussian-windowed versions of signals $f(t)$ and $g(t)$ can be represented as

$$\hat{f}(t,t_0) = f(t)\omega(t_0 - t) \quad ,$$
$$\hat{g}(t,t_0) = g(t)\omega(t_0 - t) \quad . \tag{2}$$

where $\omega(t_0 - t)$ is the Gaussian window function located at t_0, i.e., $\omega(t_0 - t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(t_0 - t)^2/2\sigma^2}$. Here σ is the standard deviation, which controls the width of the Gaussian window. With these windowed signals, we are able to compute the cross-correlation function at each local time step t_0. Without optimization, the computational cost of this moving window cross-correlation is $O(N_l N_w N_s)$. Here, N_l, N_w and N_s are the numbers of samples for time lags, Gaussian window, and input signals, respectively.

The windowed cross-correlation can be written as

$$\begin{aligned} c(t_0,\tau) &= \int_{-\infty}^{\infty} f(t+\tau, t_0+\tau) g(t,t_0) dt \\ &= \int_{-\infty}^{\infty} f(t+\tau)\omega(t_0 - t) g(t) \omega(t_0 - t) dt \\ &= \int_{-\infty}^{\infty} f(t+\tau) g(t) W(t,t_0) dt \quad , \end{aligned} \tag{3}$$

where

$$W(t,t_0) = \omega(t_0 - t)\omega(t_0 - t) = \frac{1}{2\pi\sigma^2} e^{-(t_0-t)^2/\sigma^2} \quad . \tag{4}$$

$W(t,t_0)$ is again a Gaussian function with a smaller width $\sigma/\sqrt{2}$ compared to $w(t_0 - t)$. Based on this property, we can first compute the pre-stacked cross-correlation between two signals, i.e., $h(t,\tau) = f(t+\tau)g(t)$. It is then filtered by a Gaussian function $W(t,t_0)$ as

$$c(t_0,\tau) = \int_{-\infty}^{\infty} h(t,\tau) W(t,t_0) dt \quad . \tag{5}$$

Combining with some efficient recursive Gaussian filter algorithms (Deriche, 1993; Alvarez & Mazorra, 1994; Van Vliet et al., 1998), the computational cost of implementing Equation 3 can be similar as the global cross-correlation shown in Equation 1, i.e., $O(N_l N_s)$. This makes it attractive for processing large-scale datasets, such as continuous ambient noise records.

2.2 Recursive Gaussian convolution

How to efficiently compute the convolution of a signal or an image with a Gaussian function is a common problem in data processing. There are several ways to perform this filtering procedure. For instance, by taking advantage of the Fourier transform, Gaussian convolution can be computed efficiently as elemental-wise multiplication in the frequency domain. Another efficient way to implement the Gaussian filter is to consider the result as the solution of the following heat equation:

$$\begin{aligned} \frac{\partial u}{\partial y} - \frac{\partial^2 u}{\partial x^2} &= 0 \quad , \\ u(x,0) &= f(x) \quad . \end{aligned} \tag{6}$$

Here, if we let $y = \sigma^2/2$, then the final solution of the above heat equations $u(x,y)$ is equivalent to the convolution of signal $f(x)$ with a Gaussian function $\omega(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$. This heat equation can be solved by using finite-difference schemes. Alvarez and Mazorra (1994) used this property to derive an efficient procedure to compute the Gaussian convolution. For instance, they discretized the simulation domain with

$$u_i^j = u(i\Delta x, j\Delta y) \quad , \tag{7}$$

Here Δx and Δy are the grid spacings along the spatial and time directions, i and j are the associate indices. Then, the finite-difference solution of the above heat equation can be written as

$$\frac{u_i^{j+1} - u_i^j}{\Delta t} - \frac{u_{i+1}^{j+1} + u_{i-1}^{j+1} - 2u_i^{j+1}}{\Delta x^2} = 0 \quad . \tag{8}$$

Let $\lambda = \Delta y/\Delta x^2$, then

$$(1+2\lambda)u_i^{j+1} - \lambda u_{i-1}^{j+1} - \lambda u_{i+1}^{j+1} = u_i^j \quad . \tag{9}$$

We can solve this implicit finite-difference equation using the following filter

$$H(z) = \frac{\nu}{\lambda} \frac{1}{(1-\nu z^{-1})(1-\nu z)} \quad , \tag{10}$$

with

$$\nu = \frac{1 + 2\lambda - \sqrt{(1+4\lambda)}}{2\lambda} \quad . \tag{11}$$

The above filtering procedure can be efficiently solved by using the following three steps, marching from u_i^j to u_i^{j+1} as

$$\begin{aligned} u_i^{j'} &= u_i^j + \nu u_{i-1}^{j'} \quad , \\ u_i^{j''} &= u_i^{j'} + \nu u_{i+1}^{j''} \quad , \\ u_i^{j+1} &= \frac{\nu}{\lambda} u_i^{j''} \quad . \end{aligned} \tag{12}$$

The first and second steps in the above procedure are the applications of causal and acausal filters, respectively. The third step is just a simple multiplication. Thanks to the unconditional stability of this implicit solution (Alvarez & Mazorra, 1994), only a few iterations (four iterations are used in this study to propagate along the time dimension y) enable us to converge to the solution of the Gaussian filter. Therefore, the time grid can be defined by $\Delta y = \sigma^2/(2N_y)$, where N_y is the total number of iterations. Combining this recursive Gaussian filter with Equation 5, we are able to efficiently compute the local cross-correlation function. For example, for a certain τ, the source term for Equation 6 is the pre-stacked cross-correlation $h(t, \tau)$. Then solution of the corresponding heat equation is the non-stationary CCFs $u(t, \tau)$ at this particular τ.

Figure 1 shows a simple example to illustrate the advantages of local cross-correlation over global cross-correlation. There are non-stationary time shifts for three events at 1, 3 and 5 s between two input signals. The global cross-correlation is shown in Figure 1b. Although it captures the correct time shifts around -0.4 s, 0.1 s and 0.3 s, it does not carry any useful information about the locations of these time shifts. The local cross-correlation is presented in Figure 1c, with three energy spots located at the right locations with correct time lags. This local cross-correlation function enables us to extend the dimensionality of the input signals as well as the cross-correlation function, which can be utilized to design a better strategy to improve stacking quality as discussed below.

2.3 Local similarity

Local similarity is a local attribute that allows us to measure non-stationary similarity between two time series or images (Fomel & Jin, 2009; G. Liu et al., 2009, 2011). First, the global similarity between two signals $f(t)$ and $g(t)$ can be defined as follows

$$s = \frac{\int f(t)g(t)dt}{\sqrt{\int f^2(t)dt}\sqrt{\int g^2(t)dt}} \quad . \tag{13}$$

In comparison with the global cross-correlation in Equation 1, the global similarity can be considered as a zero-lag cross-correlation function normalized by the energies of each signal. Therefore, it is insensitive to the absolute amplitudes of input signals. Its magnitude ranges from -1 to 1. Similar to local cross-correlation, local similarity was designed to measure non-stationary similarity between time-lapse images (Rickett & Lumley, 2001; Fomel & Jin, 2009). Once taking the square of Equation 13 on both sides, we can write its discretized version as

$$s^2 = \frac{(\mathbf{f}^T\mathbf{g})(\mathbf{g}^T\mathbf{f})}{(\mathbf{f}^T\mathbf{f})(\mathbf{g}^T\mathbf{g})} \quad . \tag{14}$$

Fomel (2007a) considered the above equation as the solution to the following two least-squares problems:

$$\begin{aligned} \mathbf{s}_1 &= \arg\min_{\mathbf{s}_1} |\mathbf{F}\mathbf{s}_1 - \mathbf{g}| \quad , \\ \mathbf{s}_2 &= \arg\min_{\mathbf{s}_2} |\mathbf{G}\mathbf{s}_2 - \mathbf{f}| \quad . \end{aligned} \tag{15}$$

Here, $\mathbf{F}$ and $\mathbf{G}$ are the diagonal operators constructed from the elements of input vectors f and g, respectively. Shaping regularization (Fomel, 2007b) and a conjugate gradient method can be used to solve the above least-squares problems. Then, the local similarity can be computed as elemental-wise multiplication between $\mathbf{s}_1$ and $\mathbf{s}_2$ as

$$\mathbf{s} = \sqrt{\mathbf{s}_1\mathbf{s}_2} \quad . \tag{16}$$

As pointed out in Fomel (2007a), the normalized local correlation can be considered as the first iteration from local similarity. Compared to a scalar value for the global similarity in Equation 13, the local similarity is a vector with the same length as the input signals. It can also be utilized to measure the similarity between two 2D or 3D images with high efficiency.

2.4 Stacking strategy with local similarity

From local cross-correlation, we obtain a function $c(t,\tau)$. The simplest way to calculate the CCF $c_1(\tau)$ is to use the following linear stacking

$$c_1(\tau) = \frac{1}{N}\sum_t c(t,\tau) \quad . \tag{17}$$

If we use the above expression to compute noise CCF, its result is similar to the direct implementation of Equation 1. In this study, we choose $c_1(\tau)$ as a reference trace and then compute the local similarity $s(t,\tau)$ between $c_1(\tau)$ and $c(t,\tau)$. Once with the local similarity, we can use the following strategy to improve the stacking quality

$$c_2(\tau) = \frac{1}{N}\sum_t \alpha(t,\tau)c(t,\tau) \quad , \tag{18}$$

with

$$\alpha(t,\tau) = \begin{cases} s(t,\tau) - \alpha_0 \quad , & \text{if } \alpha(t,\tau) > \alpha_0 \\ 0 \quad , & \text{if } \alpha(t,\tau) \le \alpha_0 \end{cases} \tag{19}$$

Here α_0 is a pre-defined parameter to decide the stacking level. This stacking procedure has been applied to enhance coherent signals for normal moveout corrected data (G. Liu et al., 2009) and angle-domain common-image-gathers (G. Liu et al., 2011). In this study, we use it to improve the stacking quality of noise CCFs. In practice, $\alpha(t,\tau)$ is a soft threshold solved in the seislet domain (Daubechies et al., 2004; G. Liu et al., 2009). Compared with the similarity shown in Figure 1d, the clipped similarity ($\alpha(t,\tau)$ in Eq. 19) in Figure 1f shows zero background values away from signal regions, while keeping almost 1.0 weights around the signal regions. This helps to suppress noises away from the signal regions. Because this synthetic test does not involve noises, the weighted prestack local correlations ($c_2(\tau)$ in Eq. 18) shown in Figure 1e are basically identical to the prestack local correlations shown in Figure 1c. To validate the effectiveness of the proposed method to improve the SNR, we give its definition used in this study as

$$SNR = \sqrt{\frac{\frac{1}{N_s}\int_{s_1}^{s_2} f^2(t)dt}{\frac{1}{N_n}\int_{n_1}^{n_2} f^2(t)dt}} \quad . \tag{20}$$

Where $[s_1,s_2]$ and $[n_1,n_2]$ represent signal and noise windows, N_s and N_n denote the number of time samples for signal and noise, respectively.

2.5 Workflow

In summary, the proposed method based on two local attributes can be implemented through the following six steps:

1. Download and preprocess continuous ambient noise data for each station, and cut them into daily recordings.
2. Calculate the pre-stacked cross-correlation function $h(t,\tau) = f(t+\tau)g(t)$ as shown in Equation 3 between daily recordings from station pairs.
3. Stack all daily global stacking CCFs and extend it along the time axis to obtain two-dimensional data as a reference.
4. Solve the heat equation in Equation 6 to obtain a daily local CCF. Note here, the pre-stacked CCF is the source term $f(x)$ in Equation 6, and the local CCF is also two-dimensional data.
5. Calculate the daily local similarity between the reference CCF (step 3) and the daily local CCF(step 4).
6. Stack all daily local CCFs weighted by the local similarity according to Equation 18 for each station pair.

3 Numerical examples

We download continuous ambient noise data, which are then processed by using the NoisePy package (Jiang & Denolle, 2020) following standard procedures from Bensen et al. (2007). It includes deconvolution of instrument responses, bandpass filter (2–10 s), time domain normalization and spectral whitening. The preprocessed data are used to retrieve the CCFs and demonstrate the advantages of the proposed method.

3.1 One-day records

Figures 2a and b are vertical component displacements for the USArray stations TA.234A and TA.Z36A recorded on October 10th, 2010. The inter-station between this station pair is 213 km. As shown in Figure 3a, the SNR of the global cross-correlation is relatively low because we only use daily records. Only causal Rayleigh waves can be clearly observed in this CCF. Both amplitude and time-frequency spectra of the global cross-correlation function are presented in Figures 3b and c, respectively. Besides the dominant Rayleigh wave arrival, there are many background noises, especially from 0.2 to 0.5 Hz. The local cross-correlation between Figures 2a and b is shown in Figure 2c. With the local cross-correlation, we are able to extend the 1D CCF to a 2D image, which can be used to measure local similarity and design a better strategy to improve stacking quality. Although there are many incoherent background noises in Figure 2c, it includes a vertical coherent feature around time lags between 60 to 90 s. If we use a simple linear stacking for Figure 2c along the time axis (Equation 17), we will obtain the same signal as the global cross-correlation shown in Figure 3a. Next, we compute the local similarity between Figure 2c and the raw stacked CCF shown in Figure 3a. In the local similarity result (Figure 2d), the coherent signals between 60 to 90 s stand out in comparison to the background noise level. With the stacking weights based on the local similarity (Equation 18), the new CCF and its amplitude/time-frequency spectra are presented in Figures 3d–f. Compared with the global CCF in Figures 3a–c, the incoherent noises are significantly reduced without changing the dominant Rayleigh wave arrival. In the meanwhile, the overall amplitude spectrum does not change before and after applying the stacking weights (Figures 3b and e). Now, it is much easier for us to measure the phase or group dispersion curves of Rayleigh waves for the newly stacked signal (Figure 3d).

3.2 One-month records

Figures 4a and b present ambient noise CCFs for daily records (October 2010) by using global cross-correlation and local attributes, respectively. With these local attributes, incoherent background noises are attenuated while surface wave signals are preserved and much easier for us to measure. We observe alternative causal and acausal surface wave signals, which might come from the changes in ambient noise source distributions. Since there is data redundancy for Figure 4b, instead of directly stacking it over the time axis, we calculate local similarity again between the raw stacked CCF (Figure 5a) with Figure 4b, which is shown in Figure 4c. Again, coherent signals, such as surface wave packages, stand out in comparison with background noises. Figures 5a and d compare the raw stacked CCF and the new signal with local attributes. We observe significant improvement of the new signal in terms of SNR. Based on the time-frequency analysis shown in Figures 5c and f, the dispersive characteristics of Rayleigh waves are preserved while background noises from 0.1 to 0.3 Hz are attenuated. Similar to the previous daily example, the overall amplitude spectra do not change too much between these two CCFs.

Next, we compare the convergence of the raw global cross-correlation and the new result based on local attributes. Figure 6a shows the convergence of the raw daily global CCFs over one month. With the increasing stacking duration, the SNR is improved and coherent Rayleigh wave signals stand out gradually. However, the convergence and improvement of SNR are relatively low (Figure 7b). Here we use a one-month stacked CCF as the reference trace to compute the similarity between two signals (Equation 13). In this test, the signal window is [-100, -50] s and [50, 100] s, while the noise window ranges from [-200, -150] s and [150, 200] s. Improvements for the local attribute stacking CCF are shown in Figure 6b. In addition, another causal signal appears around 50 s after 10 days of stacking, which could be body waves or higher mode surface waves with an apparent velocity about 4.3 km/s. But those incoherent events (e.g., 20 s) are relatively well suppressed. We also speculate that the improvement of SNR with our method is significant with only several days of stacking (Figure 7b). By stacking with more than 10 days, both the correlation coefficient and SNR gradually become stable (Figure 7). It is intriguing that the first 5 days CCFs show strong acausal signals (e.g., -80 s in Figures 6a-b) and our improved SNR shows a bump around 5 days (Figure 7b). This might indicate the noise sources were mainly from station Z36A to station 234A before the first 5 days. Then more complex noise sources from stations 234A to station Z36A appeared and resulted in stronger causal signals (e.g. 50-80 s) and complex noises in CCFs. As a result, the SNR goes lower from 5-12 day stacking and gradually turns high with longer stacking.

We further test the proposed approach for 20 USArray stations deployed in North Texas and Oklahoma(Figure 8). One-month data (October 2010) are used for both raw stacking and the new approach. Comparisons of these two results can be found in Figure 9. We observe the SNRs of the new approach are much higher in comparison with the classical stacking approach. Except for the dominant Rayleigh waves, we also notice that there might be additional earlier arriving weaker events emerging from 500 km distance (highlighted by dark blue arrows). For example, the early arrival at about 640 km has an apparent velocity of 3.5 km/s, which is higher than the group velocity ($\sim$2.6 km/s) of the dominant Rayleigh waves. It could be a candidate for head/diving body waves or higher mode surface waves contaminated with noises. Even they are weak signals, our method can retain them as long as they are coherent.

Next, we test the proposed approach for several stations from three dense arrays and surrounding stations deployed in Minnesota and Wisconsin around the Midcontinent Rift (Figure 10; Wolin et al., 2015). Figure 11 shows the comparisons of one-month stacked CCFs based on the conventional approach and the proposed procedure. It is obvious the SNRs are improved quite a lot for most traces shown in Figure 11c with the proposed method. The overall average of the SNR from the traditional stacking method (4.8) is much smaller than the one (35.4) from the proposed method. Although the noise level between 6-15 s period is lower, the SNRs are improved from 6.7 to 53.8 by using the proposed stacking method as shown in Figure 12. It is interesting that SNRs for data in 15-30 s period band (Figure 13) are smaller compared to those in the periods of 6-15 s, and the proposed method helps us to improve the SNRs as expected. Panels d-f show the improved SNRs from two-month stacking compared with the monthly stacking (panels a-c) of Figures 11-13. We speculate that with a two-month stacking duration, the improvements of SNRs using conventional linear stacking are larger than the proposed method, but the SNRs from one-month stacking of the proposed method are still higher than those from two-month stacking with the conventional method. This is important for monitoring time-lapse changes of near-surface velocity changes with temporary arrays which usually have quite short acquisition durations (Nakata et al., 2016; Issa et al., 2017; Mordret et al., 2020; Zhang et al., 2022).

3.3 Applications to a nodal array

Between 2017 and 2019, 10 linear dense Distribution of Basin Amplification Seismic Investigation (BASIN) nodal arrays (SG1–SG4, and SB1–SB6) were deployed in the San Gabriel and San Bernardino basins for Fine characterization of basin shapes and depths (Y. Li et al., 2022; X. Wang et al., 2021). We apply our method to the SB1 array to validate its performance. The SB1 array (Figure. 14a) consists of 239 Fairfield ZLand nodes with standard 5 Hz 3-component geophones with spatial sampling of $\sim$250 m. It was deployed for approximately one month. The basin depth beneath the SB1 array is about 0-3.0 km (Y. Li et al., 2022), therefore, it is essential to have more measurements at lower period bands, such as 1.0-5.0 s. The data downloading and preprocessing are similar to previous tests except the continuous noises were down-sampled with a sampling frequency of 4 Hz and then bandpass filtered between 1-20 s.

Taking the first station as the master station, the corresponding CCFs are arranged according to their offset and displayed in Figure 14. As expected, with 5 days of stacking (Figure 14b), the CCFs from traditional linear stacking have strong noises for all station pairs. The 30-day stacking clearly improves the data quality so that we can observe the dominant acausal Rayleigh waves. Compared with CCFs from linear stacking, the proposed method helps us to enhance coherent signals even with only 5 days of stacking. The 30-day stacking further improves the data quality. Next, we measure phase dispersion curves for the stacked CCFs, which are the input for surface wave tomography (Yao et al., 2006; Fang et al., 2015). There are three criteria to make sure the measured dispersion curves are stable. (1) The SNR is larger than 5.0; (2) The inter-station distance is larger than 1.5 wavelength at corresponding periods (Bensen et al., 2007; G. Chen et al., 2023). (3) The picked phase velocity is within $\pm$ 12% of a 3D reference phase velocity model. Here the 3D reference phase velocity model is constructed from a local tomographic model, CVM-S 4.26 (Lee et al., 2014), by conducting a forward modeling based on the fast-marching method (Rawlinson & Sambridge, 2004; Fang et al., 2015). Thanks to the improvement of SNR from our method, more high-quality dispersion curves pass the selection criteria with 5-day stacking compared with those measured from linear stacking as illustrated in Figure 15. More importantly, the number of measurements for 1-5 s and 10-15 s period bands are also increased. These are further improved for 30-day stacked CCFs using the proposed method.

4 Discussions

Commonly used ambient noise processing procedures (Bensen et al., 2007) require stacking over long, continuous records to enhance coherent signals, also require the assumption of evenly distributed ambient sources in order to obtain good estimates of Green's functions. However, both conditions impose constraints on the application of temporary arrays with short acquisition duration.

To improve the SNR of CCFs stacked with a short duration, Xie et al. (2020) proposed a root-mean-square-ratio selective (RMSRS) stacking procedure to remove those CCFs that negatively contribute to the SNR of the final stacked CCF. It is realized by comparing the root-mean-square ratio of signals and noises for each CCF and the stacked CCF. Therefore, its effectiveness depends on an accurate definition of signal/noise windows. Here, we compare the conventional linear stacking, proposed procedure and RMSRS stacking for the station pair TA.A12A-TA.A18A at 10-35 s period band. The signal window (Figure 17c) is defined with the reference time t_{ref} and maximum period of interest T_{max} as $[t_{ref} - 2T_{max}, t_{ref} + 2T_{max}]$. Here, the reference time $t_{ref} = d/v_{ref}$ is defined by the inter-station distance d (425 km) and a reference group velocity v_{ref} (3.0 km/s). The noise window is defined from 0 to the signal windows and $4T_{max}$ out of the signal windows (Xie et al., 2020). Compared with linear stacking, the RMSRS stacking successfully suppresses the noises within the defined noise windows and therefore improves the SNR as highlighted in Figures 17d and f. Similarly, our approach also successfully attenuates those incoherent noises and significantly improves the SNR, which is about two times the other two methods. However, the coherent signal at about 75 seconds with an apparent velocity of 5.6 km/s is retained by our approach due to the similarity between local CCFs and the stacked CCFs. The main reason that the RMSRS-based method helps to suppress this signal is because they are selected as noises. Such selection seems to be challenging to deal with low-SNR CCFs at shorter period bands (3-16 s) obtained from the high-frequency nodal array (Figure 14). As illustrated in Figures 14d, the improvement of selective stacking is limited compared with linear stacking. On the contrary, our approach significantly improves the stacking quality and helps us to obtain more high-quality dispersion curves. We note here that, for fair comparisons, the RMSRS-based method is implemented on the raw stacked CCFs. For better performance, Xie et al. (2020) suggested using RMSRS stacking at several narrow period bands so that it could define a better selection window, which is out of the scope of this study.

Although the proposed approach helps us to improve the SNR of CCFs with only several days of stacking, the non-causality and asymmetry, which mainly arise from the uneven distribution of ambient noise sources, remain challenging. For tomography purposes, theoretical works indicate that phase velocities can be estimated from the empirical Green's functions, which are obtained by taking the negative time derivative of the symmetric cross-correlation under the assumption of a spatially homogeneous ambient-noise source distribution (Lobkis & Weaver, 2001; Sabra et al., 2005; Snieder, 2004; Yao et al., 2006; Lin et al., 2008). As suggested by Yao et al. (2006), inhomogeneous source distribution may contribute to 1–3 percent inconsistency between phase velocity measurements and the traditional earthquake-based two-station method between periods of 20–30 s. Therefore, the "symmetric" CCF is usually taken by the average of the cross-correlation at positive and negative correlation lag times (Yao et al., 2006; Lin et al., 2008). However, how much such averaged CCFs affect tomography results remains unknown. On the other hand, several studies suggested measuring full-waveform differences of CCFs and source location simultaneously (Tromp et al., 2010; Sager et al., 2018; Datta et al., 2023), which naturally mitigates the uncertainty caused by the source distributions. In both cases, it is important for us to obtain high SNR CCFs, especially for short-duration nodal arrays. In addition, because the high-quality phase velocity measurements obtained from 5-day stacking (Figure 15c) by the proposed approach is comparable with those obtained from the 30-day linear stacking, the surface wave-based monitoring (Durand et al., 2011; Brenguier et al., 2020) seems to be possible, albeit coda waves are mostly used (Mao et al., 2019, 2020; Luo et al., 2023). Another potential application for our approach could be weak coherent signal extraction, such as body waves (Zhan et al., 2010; Poli et al., 2012; Nakata et al., 2015, 2016; Mao et al., 2020). Body waves extracted from ambient noise CCFs have been proven to be capable of improving the imaging resolution compared to surface wave tomography (Nakata et al., 2015). As illustrated in Figures 6 and 14, those coherent signals with an apparent velocity larger than 4.0 km/s might be good candidates for body waves, although we are not able to rule out the possibility of higher-order surface waves.

The last factor we need to consider is the computational and memory cost for our approach. Taking station pair TA.A12A and TA.A18A as an example, we have 30 days of recordings and 86400-time samples for each day. The cross-correlation time lag varies from -480 s to 480 s for every 1 s. We calculate the global cross-correlations every 60 min-utes with 75% overlaps, which yields 91 subset CCFs for each day. They are then normalized and stacked to obtain the dayily CCF. Finally, 30 days of global cross-correlation and linear stacking take 47.9 s and 0.02 s, respectively, while the local cross-correlation, local similarity and weighted stacking take 196.7 s, 12,150 s and 78.3 s. We note here, that the local cross-correlation solved with our method actually contains 86,400 CCFs for each day. It helps us to extend the dimension of CCFs dramatically (534 times the number of global CCFs), but with only four times the computational cost compared to global stacking. Whereas, such a high dimension, in turn, greatly decreases the efficiency of our approach. To mitigate this problem, the local CCFs are downsampled 60 times (from 1.0 s to 60 s) by taking the average CCFs for every 60 samples, and then the computational cost for local similarity and weighted stacking is decreased to 167.1 and 10.0 s, respectively. The final stacked CCFs by using these local CCFs are almost the same as the original local CCFs. Overall, the computational cost for the proposed approach is 7.8 times the computational cost compared to the hourly global stacking, but with 15.8 times more CCFs. In addition, increasing the cross-correlation time window (e.g., from one hour to three hours) does not affect the efficiency of our approach, but will increase the computational time for global cross-correlation. Finally, we compare the computation cost using the nodal array. For each station, we have 345,600 time samples per day with a 4 Hz sampling frequency. The cross-correlation time lag varies from -120 s to 120 s for every 0.25 s. The global and local CCFs are calculated the same as in the previous example. To save computational costs, we downsample the local CCFs 60 times (from 0.25 s to 15 s). Then all station-pairs are distributed to 72 CPU cores for parallel computations. Our approach takes 5792.0 s and the traditional linear stacking takes 792.0 s. The final stacked CCFs are compared in section 3.3. In conclusion, our approach significantly helps us to improve the SNR of CCFs, but with about 7.5 times the computation cost compared to traditional linear stacking. Such extra computational costs are bearable compared to the following computational costs tomography (Zhu, 2018; Wu et al., 2023; G. Chen et al., 2023).

5 Conclusion

In this study, by taking advantage of local attributes, we present a new approach to increase the SNRs of ambient noise CCFs. Two local attributes are used in this study: local cross-correlation and local similarity. The local cross-correlation is employed to extend the dimensionality of daily CCFs, and the local similarity is used to design better-stacking weights to enhance coherent signals and attenuate incoherent background noises. Applications to ambient noise records from several broadband stations and a high-frequency nodal array demonstrate the performance of the proposed approach. With higher SNRs, we are able to extract more high-quality dispersion curves, which are important for surface wave tomography. In addition, 5-day stacking by our approach can produce CCFs comparable to 30-day linear stacking in terms of SNRs, demonstrating its potential applications for time-lapse monitoring. In addition, extracting coherent weak signals, such as body waves, could be another application of the proposed approach.

Open Research

All seismic data used in this study can be obtained from the IRIS Data Management Center (`https://ds.iris.edu/ds`) under the network codes TA and XI. We use the Noisepy (`https://noise-python.readthedocs.io/en/latest/`) for parallel data downloading and preprocessing (Jiang & Denolle, 2020). The open software Madagascar (Fomel et al., 2013) download from (`http://www.ahay.org`) is used to calculate local similarity and plot figures. We also use PyGMT (Wessel et al., 2019) downloaded from (`https://www.pygmt.org/latest/`) for plotting figures.

Author Contributions

The authors confirm their contribution to the paper as follows: study conception and design: Hejun Zhu; data collection: Hejun Zhu and Bin He; analysis and interpretation of results: Hejun Zhu, Bin He and David Lumley; draft manuscript preparation: Hejun Zhu, Bin He and David Lumley. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments

This paper is contribution no. *** from the Department of Geosciences at the University of Texas at Dallas. The numerical results are computed through the Optane nodes on the UTD Seismology Group HPC clusters. This research is partially supported by the sponsors of the UT Dallas 3D + 4D Seismic Full Waveform Inversion research consortium.

References

Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bulletin of the Earthquake Research Institute*, *35*, 415–456.

Alvarez, L., & Mazorra, L. (1994). Signal and image restoration using shock filters and anisotropic diffusion. *SIAM journal on numerical analysis*, *31*(2), 590–605.

Baig, A. M., Campillo, M., & Brenguier, F. (2009). Denoising seismic noise cross correlations. *Journal of Geophysical Research: Solid Earth*, *114*(B8).

Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A. L., Lin, F., Moschetti, M., ... Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical journal international*, *169*(3), 1239–1260.

Brenguier, F., Courbis, R., Mordret, A., Campman, X., Boué, P., Chmiel, M., ... others (2020). Noise-based ballistic wave passive seismic monitoring. part 1: body waves. *Geophysical Journal International*, *221*(1), 683–691.

Brenguier, F., Shapiro, N. M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., & Nercessian, A. (2008). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, *1*(2), 126–130.

Buckingham, M. J., Berknout, B. V., & Glegg, S. A. (1992). Imaging the ocean with ambient noise. *Nature*, *356*(6367), 327–329.

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

Chen, G., Chen, J., Tape, C., Wu, H., & Tong, P. (2023). Double-difference adjoint tomography of the crust and uppermost mantle beneath alaska. *Journal of Geophysical Research: Solid Earth*, *128*(1), e2022JB025168.

Chen, M., Huang, H., Yao, H., van der Hilst, R., & Niu, F. (2014). Low wave speed zones in the crust beneath se tibet revealed by ambient noise adjoint tomography. *Geophysical Research Letters*, *41*(2), 334–340.

Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, *33*(2), 264–269.

Clarke, D., Zaccarelli, L., Shapiro, N., & Brenguier, F. (2011). Assessment of resolution and accuracy of the moving window cross spectral technique for monitoring crustal temporal variations using ambient seismic noise. *Geophysical Journal International*, *186*(2), 867–882.

Datta, A., Shekar, B., & Kumar, P. L. (2023). Acoustic full waveform inversion for 2-d ambient noise source imaging. *Geophysical Journal International*, *234*(3), 1628–1639.

Daubechies, I., Defrise, M., & De Mol, C. (2004). An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, *57*(11), 1413–1457.

De Plaen, R. S., Lecocq, T., Caudron, C., Ferrazzini, V., & Francis, O. (2016). Single-station monitoring of volcanoes using seismic ambient noise. *Geophysical Research Letters*, *43*(16), 8511–8518.

Deriche, R. (1993). *Recursively implementating the gaussian and its derivatives* (Unpublished doctoral dissertation). INRIA.

Díaz, E., & Sava, P. (2015). Data domain wavefield tomography using local correlation functions. In *2015 seg annual meeting.*

Dougherty, S. L., Cochran, E. S., & Harrington, R. M. (2019). The large-n seismic survey in oklahoma (lasso) experiment. *Seismological Research Letters*, *90*(5), 2051–2057.

Draganov, D., Campman, X., Thorbecke, J., Verdel, A., & Wapenaar, K. (2009). Reflection images from ambient seismic noise. *Geophysics*, *74*(5), A63–A67.

Durand, S., Montagner, J., Roux, P., Brenguier, F., Nadeau, R., & Ricard, Y. (2011). Passive monitoring of anisotropy change associated with the parkfield 2004 earthquake. *Geophysical Research Letters*, *38*(13).

Fan, X., Guo, Z., Zhao, Y., & Chen, Q.-F. (2022). Crust and uppermost mantle magma plumbing system beneath changbaishan intraplate volcano, china/north korea, revealed by ambient noise adjoint tomography. *Geophysical Research Letters*, *49*(12), e2022GL098308.

Fang, H., Yao, H., Zhang, H., Huang, Y.-C., & van der Hilst, R. D. (2015). Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: methodology and application. *Geophysical Journal International*, *201*(3), 1251–1263.

Fomel, S. (2007a). Local seismic attributes. *Geophysics*, *72*(3), A29–A33.

Fomel, S. (2007b). Shaping regularization in geophysical-estimation problems. *Geophysics*, *72*(2), R29–R36.

Fomel, S., & Jin, L. (2009). Time-lapse image registration using the local similarity attribute. *Geophysics*, *74*(2), A7–A11.

Fomel, S., Sava, P., Vlad, I., Liu, Y., & Bashkardin, V. (2013). Madagascar: Open-source software project for multidimensional data analysis and reproducible computational experiments. *Journal of Open Research Software*, *1*(1).

Fomel, S., & van der Baan, M. (2014). Local skewness attribute as a seismic phase detector. *Interpretation*, *2*(1), SA49–SA56.

Gao, H., & Shen, Y. (2014). Upper mantle structure of the cascades from full-wave ambient noise tomography: Evidence for 3d mantle upwelling in the back-arc. *Earth and Planetary Science Letters*, *390*, 222–233.

Hadziioannou, C., Larose, E., Baig, A., Roux, P., & Campillo, M. (2011). Improving temporal resolution in ambient noise monitoring of seismic wave speed. *Journal of Geophysical Research: Solid Earth*, *116*(B7).

Hale, D. (2006). Fast local cross-correlations of images. In *Seg technical program expanded abstracts 2006* (pp. 3160–3164). Society of Exploration Geophysicists.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... Oliphant, T. E. (2020, September). Array programming with NumPy. *Nature*, *585*(7825), 357–362. Retrieved from `https://doi.org/10.1038/s41586-020-2649-2` doi: 10.1038/s41586-020-2649-2

Huang, H., Yao, H., & van der Hilst, R. D. (2010). Radial anisotropy in the crust of se tibet and sw china from ambient noise interferometry. *Geophysical Research Letters*, *37*(21).

Issa, N. A., Lumley, D., & Pevzner, R. (2017). Passive seismic imaging at reservoir depths using ambient seismic noise recorded at the otway co2 geological storage research facility. *Geophysical Journal International*, *209*(3), 1622–1628.

Jiang, C., & Denolle, M. A. (2020). Noisepy: A new high-performance python tool for ambient-noise seismology. *Seismological Research Letters*, *91*(3), 1853–1866.

Le Breton, M., Bontemps, N., Guillemot, A., Baillet, L., & Larose, É. (2021). Landslide monitoring using seismic ambient noise correlation: challenges and applications. *Earth-Science Reviews*, *216*, 103518.

Lee, E.-J., Chen, P., Jordan, T. H., Maechling, P. B., Denolle, M. A., & Beroza, G. C. (2014). Full-3-d tomography for crustal structure in southern california based on the scattering-integral and the adjoint-wavefield methods. *Journal of Geophysical Research: Solid Earth*, *119*(8), 6421–6451.

Li, G., Niu, F., Yang, Y., & Xie, J. (2018). An investigation of time-frequency domain phase-weighted stacking and its application to phase-velocity extraction from ambient noise's empirical green's functions. *Geophysical Journal International*, *212*(2), 1143–1156.

Li, Y., Villa, V., Clayton, R., & Persaud, P. (2022). Shear wave velocities in the san gabriel and san bernardino basins, california. *Authorea Preprints*.

Lin, F.-C., Moschetti, M. P., & Ritzwoller, M. H. (2008). Surface wave tomography of the western united states from ambient seismic noise: Rayleigh and love wave phase velocity maps. *Geophysical Journal International*, *173*(1), 281–298.

Lin, F.-C., Ritzwoller, M. H., Yang, Y., Moschetti, M. P., & Fouch, M. J. (2011). Complex and variable crustal and uppermost mantle seismic anisotropy in the western united states. *Nature Geoscience*, *4*(1), 55–61.

Lin, F.-C., Tsai, V. C., Schmandt, B., Duputel, Z., & Zhan, Z. (2013). Extracting seismic core phases with array interferometry. *Geophysical Research Letters*, *40*(6), 1049–1053.

Liu, G., Fomel, S., & Chen, X. (2011). Stacking angle-domain common-image gathers for normalization of illumination. *Geophysical Prospecting*, *59*(2), 244–255.

Liu, G., Fomel, S., Jin, L., & Chen, X. (2009). Stacking seismic data using local correlation. *Geophysics*, *74*(3), V43–V48.

Liu, G., Persaud, P., & Clayton, R. W. (2018). Structure of the northern los angeles basins revealed in teleseismic receiver functions from short-term nodal seismic arrays. *Seismological Research Letters*, *89*(5), 1680–1689.

Liu, Y., Niu, F., Chen, M., & Yang, W. (2017). 3-d crustal and uppermost mantle structure beneath ne china revealed by ambient noise adjoint tomography. *Earth and Planetary Science Letters*, *461*, 20–29.

Lobkis, O. I., & Weaver, R. L. (2001). On the emergence of the green's function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, *110*(6), 3011–3017.

Luo, B., Zhang, S., & Zhu, H. (2023). Monitoring seasonal fluctuation and long-term trends for the greenland ice sheet using seismic noise auto-correlations. *Geophysical Research Letters*, *50*(7), e2022GL102146.

Maguire, R., Schmandt, B., Li, J., Jiang, C., Li, G., Wilgus, J., & Chen, M. (2022). Magma accumulation at depths of prior rhyolite storage beneath yellowstone caldera. *Science*, *378*(6623), 1001–1004.

Mainsant, G., Larose, E., Brönnimann, C., Jongmans, D., Michoud, C., & Jaboyedoff, M. (2012). Ambient seismic noise monitoring of a clay landslide: Toward failure prediction. *Journal of Geophysical Research: Earth Surface*, *117*(F1).

Mao, S., Campillo, M., van Der Hilst, R. D., Brenguier, F., Stehly, L., & Hillers, G. (2019). High temporal resolution monitoring of small variations in crustal strain by dense seismic arrays. *Geophysical Research Letters*, *46*(1), 128–137.

Mao, S., Lecointre, A., van der Hilst, R. D., & Campillo, M. (2022). Space-time monitoring of groundwater fluctuations with passive seismic interferometry. *Nature communications*, *13*(1), 1–9.

Mao, S., Mordret, A., Campillo, M., Fang, H., & van der Hilst, R. D. (2020). On the measurement of seismic traveltime changes in the time–frequency domain with wavelet cross-spectrum analysis. *Geophysical Journal International*, *221*(1), 550–568.

Mordret, A., Courbis, R., Brenguier, F., Chmiel, M., Garambois, S., Mao, S., ... others (2020). Noise-based ballistic wave passive seismic monitoring–part 2: surface waves. *Geophysical Journal International*, *221*(1), 692–705.

Moschetti, M., Ritzwoller, M., Lin, F., & Yang, Y. (2010). Seismic evidence for widespread western-us deep-crustal deformation caused by extension. *Nature*, *464*(7290), 885–889.

Nakata, N., Boué, P., Brenguier, F., Roux, P., Ferrazzini, V., & Campillo, M. (2016). Body and surface wave reconstruction from seismic noise correlations between arrays at piton de la fournaise volcano. *Geophysical Research Letters*, *43*(3), 1047–1054.

Nakata, N., Chang, J. P., Lawrence, J. F., & Boué, P. (2015). Body wave extraction and tomography at long beach, california, with ambient-noise interferometry. *Journal of Geophysical Research: Solid Earth*, *120*(2), 1159–1173.

Nakata, N., Gualtieri, L., & Fichtner, A. (2019). *Seismic ambient noise*. Cambridge

University Press.

Nakata, N., & Snieder, R. (2011). Near-surface weakening in japan after the 2011 tohoku-oki earthquake. *Geophysical Research Letters*, *38*(17).

Poli, P., Pedersen, H., & Campillo, M. (2012). Emergence of body waves from cross-correlation of short period seismic noise. *Geophysical Journal International*, *188*(2), 549–558.

Rawlinson, N., & Sambridge, M. (2004). Wave front evolution in strongly heterogeneous layered media using the fast marching method. *Geophysical Journal International*, *156*(3), 631–647.

Rickett, J., & Lumley, D. (2001). Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the gulf of mexico. *Geophysics*, *66*(4), 1015–1025.

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W., & Fehler, M. C. (2005). Extracting time-domain green's function estimates from ambient seismic noise. *Geophysical research letters*, *32*(3).

Sager, K., Boehm, C., Ermert, L., Krischer, L., & Fichtner, A. (2020). Global-scale full-waveform ambient noise inversion. *Journal of Geophysical Research: Solid Earth*, *125*(4), e2019JB018644.

Sager, K., Ermert, L., Boehm, C., & Fichtner, A. (2018). Towards full waveform ambient noise inversion. *Geophysical Journal International*, *212*(1), 566–590.

Schimmel, M., Stutzmann, E., & Gallart, J. (2011). Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale. *Geophysical Journal International*, *184*(1), 494–506.

Seats, K. J., Lawrence, J. F., & Prieto, G. A. (2012). Improved ambient noise correlation functions using welch s method. *Geophysical Journal International*, *188*(2), 513–523.

Shapiro, N. M., & Campillo, M. (2004). Emergence of broadband rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, *31*(7).

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*, *307*(5715), 1615–1618.

Shen, W., & Ritzwoller, M. H. (2016). Crustal and uppermost mantle structure beneath the united states. *Journal of Geophysical Research: Solid Earth*, *121*(6), 4306–4342.

Snieder, R. (2004). Extracting the green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical review E*, *69*(4), 046610.

Stehly, L., Cupillard, P., & Romanowicz, B. (2011). Towards improving ambient noise tomography using curvelet denoising filters and sem simulations of seismic ambient noise simultaneously. *Comptes Rendus Geoscience*, *343*(8-9), 591–599.

Tromp, J., Luo, Y., Hanasoge, S., & Peter, D. (2010). Noise cross-correlation sensitivity kernels. *Geophysical Journal International*, *183*(2), 791–819.

Van Vliet, L. J., Young, I. T., & Verbeek, P. W. (1998). Recursive gaussian derivative filters. In *Proceedings. fourteenth international conference on pattern recognition (cat. no. 98ex170)* (Vol. 1, pp. 509–514).

Wang, K., Yang, Y., Basini, P., Tong, P., Tape, C., & Liu, Q. (2018). Refined crustal and uppermost mantle structure of southern california by ambient noise adjoint tomography. *Geophysical Journal International*, *215*(2), 844–863.

Wang, Q.-Y., & Yao, H. (2020). Monitoring of velocity changes based on seismic ambient noise: A brief review and perspective. *Earth and Planetary Physics*, *4*(5), 532–542.

Wang, X., Zhan, Z., Zhong, M., Persaud, P., & Clayton, R. W. (2021). Urban basin structure imaging based on dense arrays and bayesian array-based coherent receiver functions. *Journal of Geophysical Research: Solid Earth*, *126*(9), e2021JB022279.

Wapenaar, K. (2004). Retrieving the elastodynamic green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical review letters*, *93*(25), 254301.

Weaver, R. L., & Lobkis, O. I. (2001). Ultrasonics without a source: Thermal fluctuation correlations at mhz frequencies. *Physical Review Letters*, *87*(13), 134301.

Weaver, R. L., & Yoritomo, J. Y. (2018). Temporally weighting a time varying noise field to improve green function retrieval. *The Journal of the Acoustical Society of America*, *143*(6), 3706–3719.

Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, *20*(11), 5556–5564.

Wolin, E., van der Lee, S., Bollmann, T. A., Wiens, D. A., Revenaugh, J., Darbyshire, F. A., ... Wysession, M. E. (2015). Seasonal and diurnal variations in long-period noise at spree stations: The influence of soil characteristics on shallow stations' performance. *Bulletin of the Seismological Society of America*, *105*(5), 2433–2452.

Woollard, G. (1965). The bouguer gravity anomaly map of the united states. *Eos, Transactions American Geophysical Union*, *46*(1), 197–202.

Wu, S.-M., Huang, H.-H., Lin, F.-C., Farrell, J., & Schmandt, B. (2023). Extreme seismic anisotropy indicates shallow accumulation of magmatic sills beneath yellowstone caldera. *Earth and Planetary Science Letters*, *616*, 118244.

Xie, J., Yang, Y., & Luo, Y. (2020). Improving cross-correlations of ambient noise using an rms-ratio selection stacking method. *Geophysical Journal International*, *222*(2), 989–1002.

Yang, X., Bryan, J., Okubo, K., Jiang, C., Clements, T., & Denolle, M. A. (2023). Optimal stacking of noise cross-correlation functions. *Geophysical Journal International*, *232*(3), 1600–1618.

Yang, Y., Ritzwoller, M. H., Levshin, A. L., & Shapiro, N. M. (2007). Ambient noise rayleigh wave tomography across europe. *Geophysical Journal International*, *168*(1), 259–274.

Yao, H., van Der Hilst, R. D., & De Hoop, M. V. (2006). Surface-wave array tomography in se tibet from ambient seismic noise and two-station analysis—i. phase velocity maps. *Geophysical Journal International*, *166*(2), 732–744.

Yao, H., Van Der Hilst, R. D., & Montagner, J.-P. (2010). Heterogeneity and anisotropy of the lithosphere of se tibet from surface wave array tomography. *Journal of Geophysical Research: Solid Earth*, *115*(B12).

Zhan, Z., Ni, S., Helmberger, D. V., & Clayton, R. W. (2010). Retrieval of moho-reflected shear wave arrivals from ambient seismic noise. *Geophysical Journal International*, *182*(1), 408–420.

Zhang, Z., Nakata, N., Karplus, M., Kaip, G., & Yi, J. (2022). Shallow ice-sheet composite structure revealed by seismic imaging near the west antarctic ice

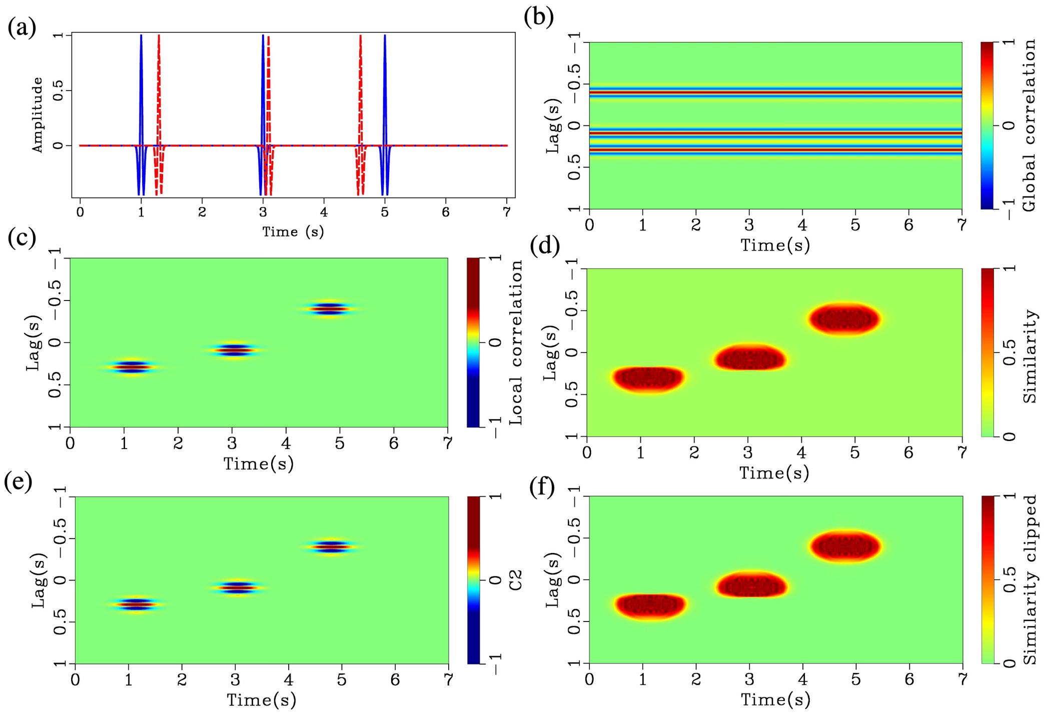


Figure 1. Comparison of local and global cross-correlations. Panel(a) shows two signals with non-stationary time shifts. Three events with time shifts of -0.4 s, 0.1 s and 0.3 s are used. Panels (b) and (c) present global and local cross-correlations, respectively. Here, the 1D global cross-correlations are extended along the time dimension for better comparisons with local cross-correlations. σ is set to 0.2 s so that it is small enough to capture the non-stationary property of these two signals. Panels (d) and (f) show the similarity ($\mathbf{s}$ in Equation 16) between local and global cross-correlations before and after applying a threshold ($\alpha(t, \tau)$ in Equation 19), respectively. Panel e shows the prestack local cross-correlations weighted by local similarity (c_2 in Equation 18)

.

sheet (wais) divide camp. *Journal of Geophysical Research: Earth Surface*, e2022JF006777.

Zhu, H. (2018). Crustal wave speed structure of north texas and oklahoma based on ambient noise cross-correlation functions and adjoint tomography. *Geophysical Journal International*, *214*(1), 716–730.

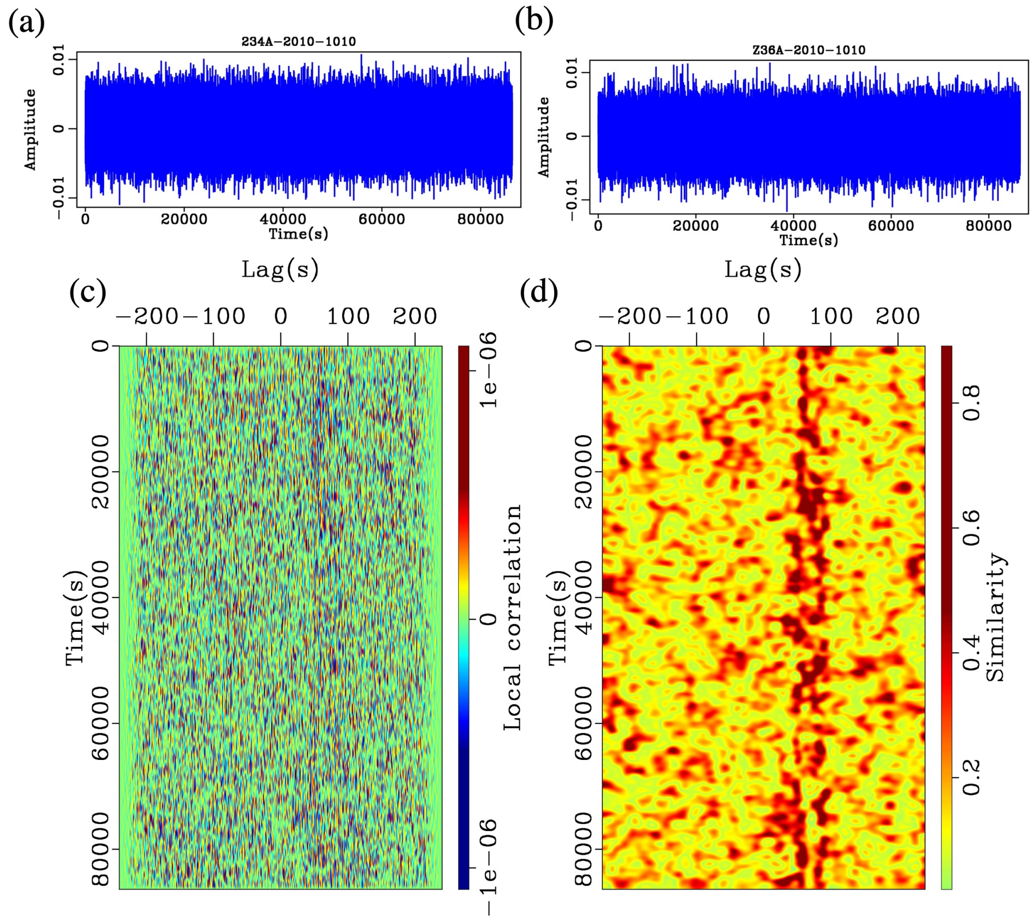


Figure 2. Local cross-correlation and local similarity for one-day CCFs between stations TA.234A and TA.Z36A on October 10, 2010. Panels (a) and (b) show one-day records for stations TA.234A and TA.Z36A, respectively. Panel (c) presents local cross-correlation between these two signals. Panel (d) shows the local similarity between panel (c) and the linearly stacked signal shown in Figure 3a. The causal Rayleigh waves between 60-80 s can be clearly observed in panel (d).

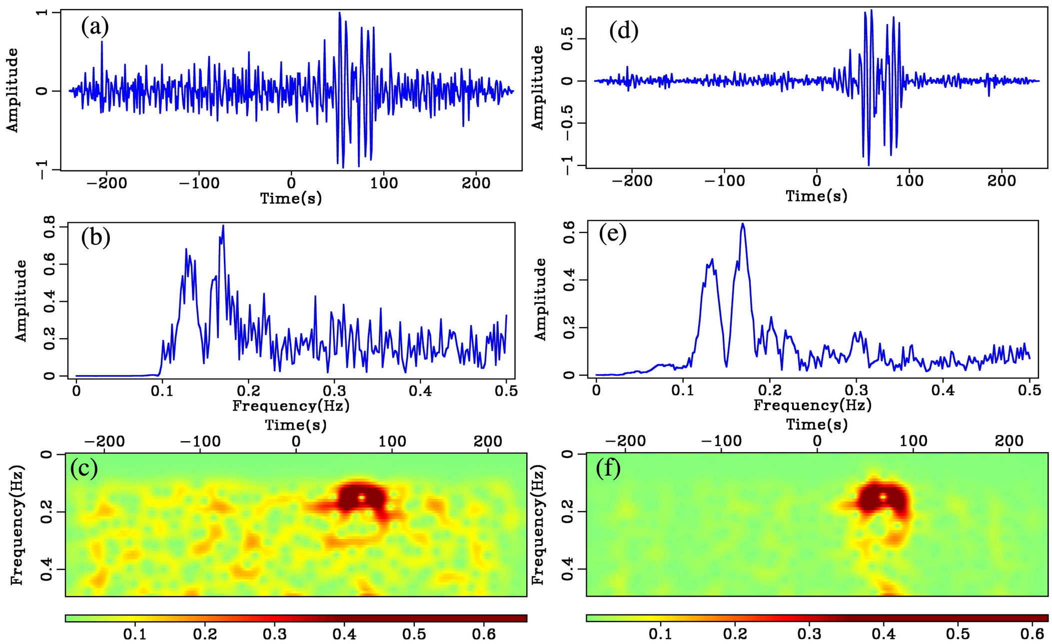


Figure 3. Comparisons of one-day CCFs from simple stacking and stacking with local attributes. Panel (a) shows the CCF with simple linear stacking (Equation 1). Panels (b) and (c) are the amplitude and time-frequency spectra of panel (a), respectively. Panel (d) is the CCF stacked with local similarity. Panels (e) and (f) are the amplitude and time-frequency spectra of panel (d).

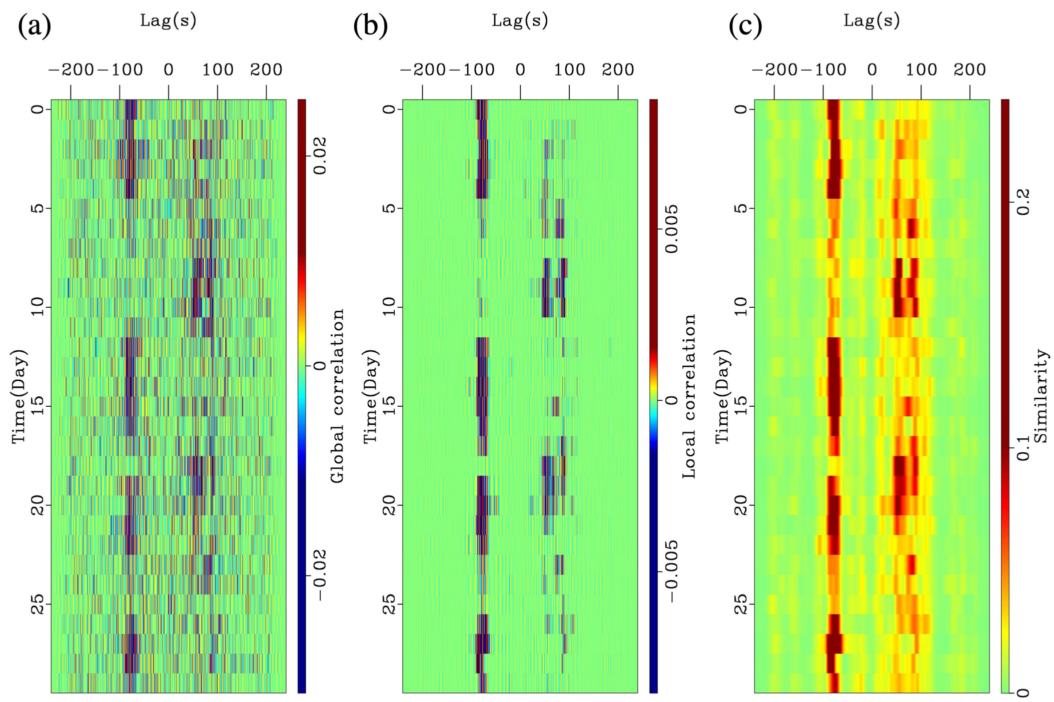


Figure 4. Comparisons of daily CCFs within one month (October 2010) for simple stacking (a) and stacking with local attribute (b). Panel (c) shows the local similarity between panel (b) and the linear stacked result shown in Figure 5a.

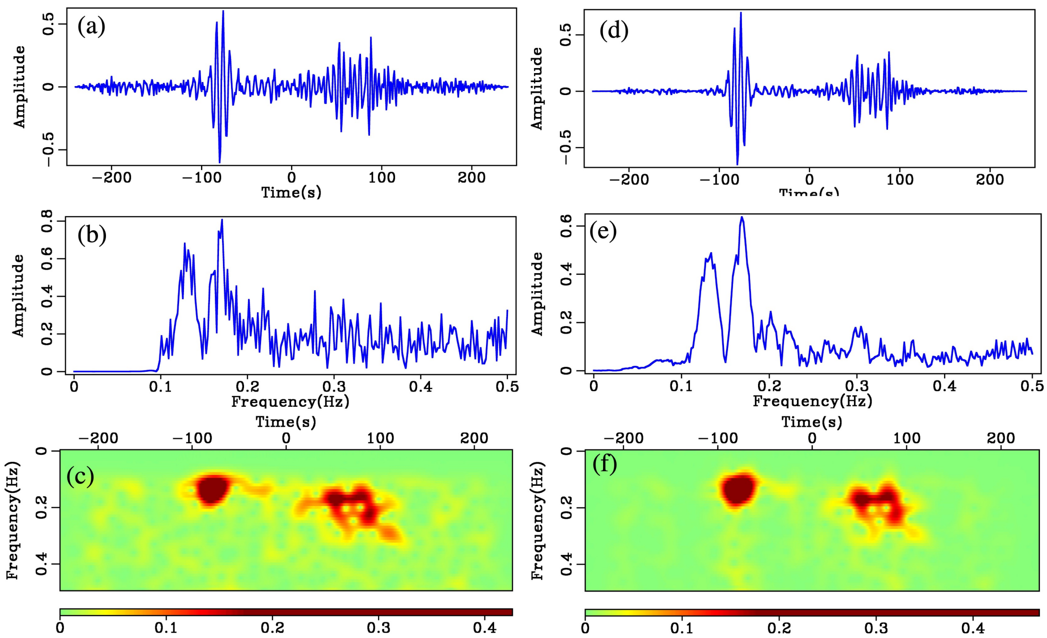


Figure 5. The same setting as Figure 3 except for monthly stacked results on October 2010.

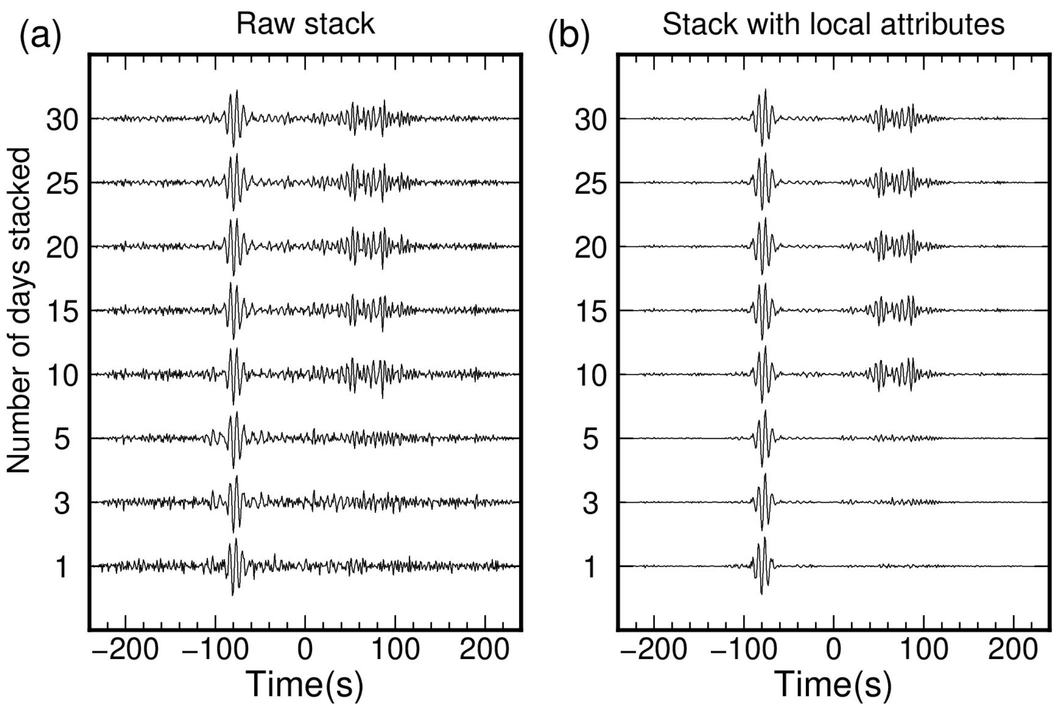


Figure 6. Comparisons of convergence for conventional (a) and the proposed procedure (b). The causal and acasual signals could be clearly identified around ±80 s with an apparent group velocity of 2.7 km/s. After 10 days of stacking, we observe a causal coherent signal appear around 50 s with an apparent velocity of 4.3 km/s, possibly representing body waves.

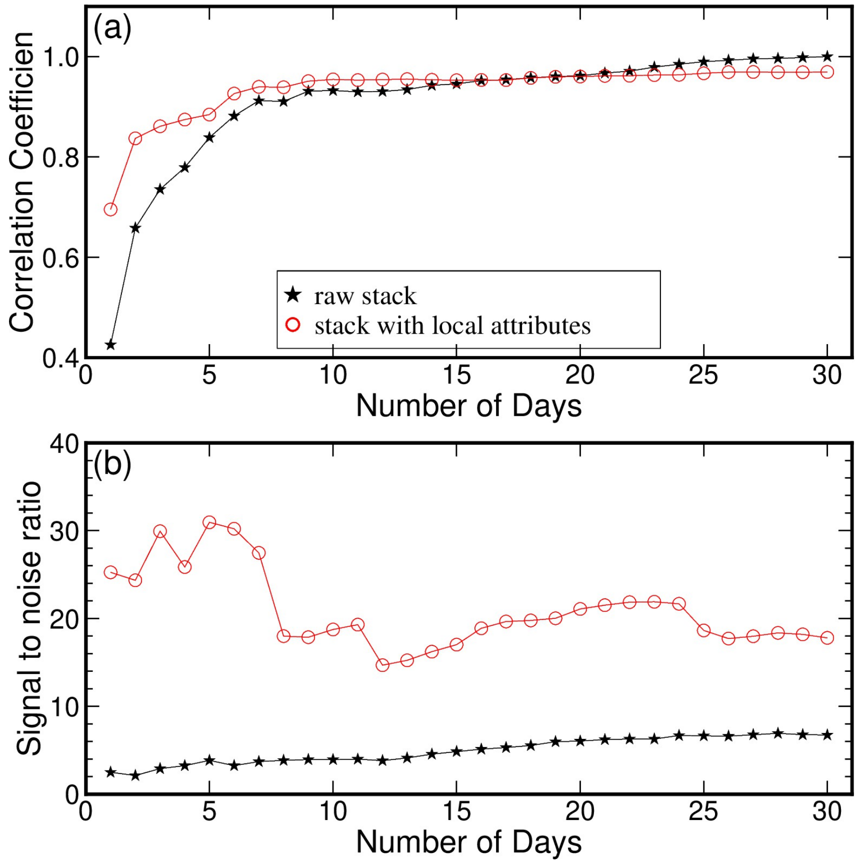


Figure 7. Improvements of correlation coefficients (a) and SNRs (b) for one-month stacking results. Black stars and open red circles are the results of the conventional stacking procedure and the proposed approach. The reference trace for calculating the correlation coefficients is the 30-day raw stack.

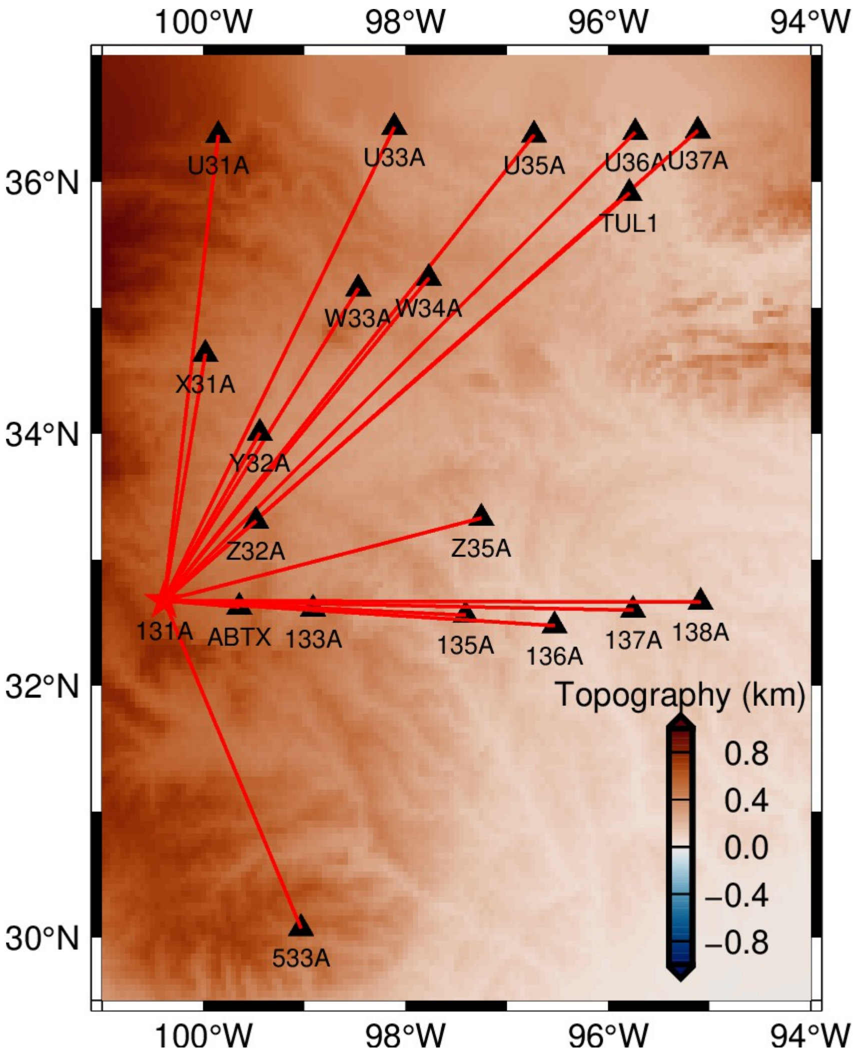


Figure 8. Distributions of the USArray stations in north Texas and Oklahoma used for comparisons in Figure 9.

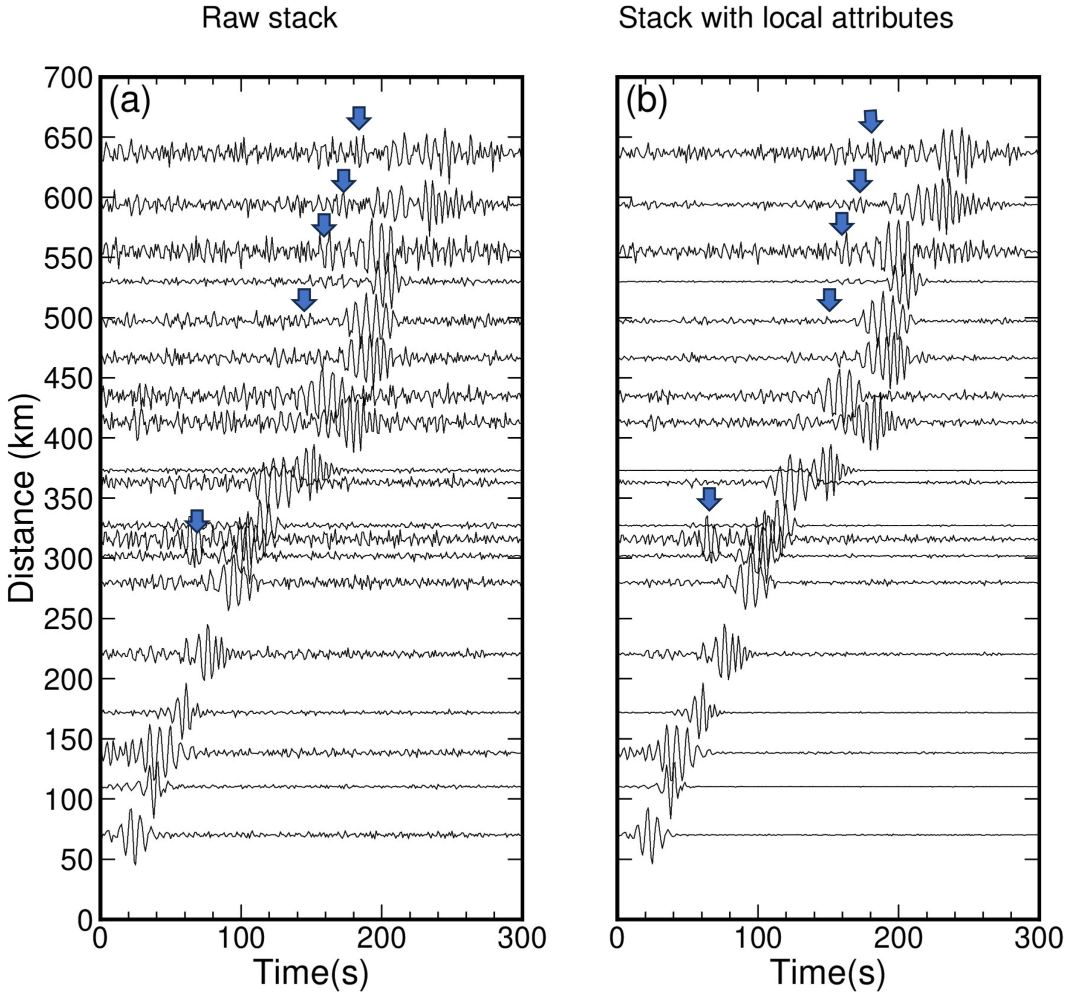


Figure 9. Comparisons of one-month stacked CCFs averaged from negative and positive lags from conventional approach (a) and the proposed procedure (b) for station pairs shown in Figure 8.

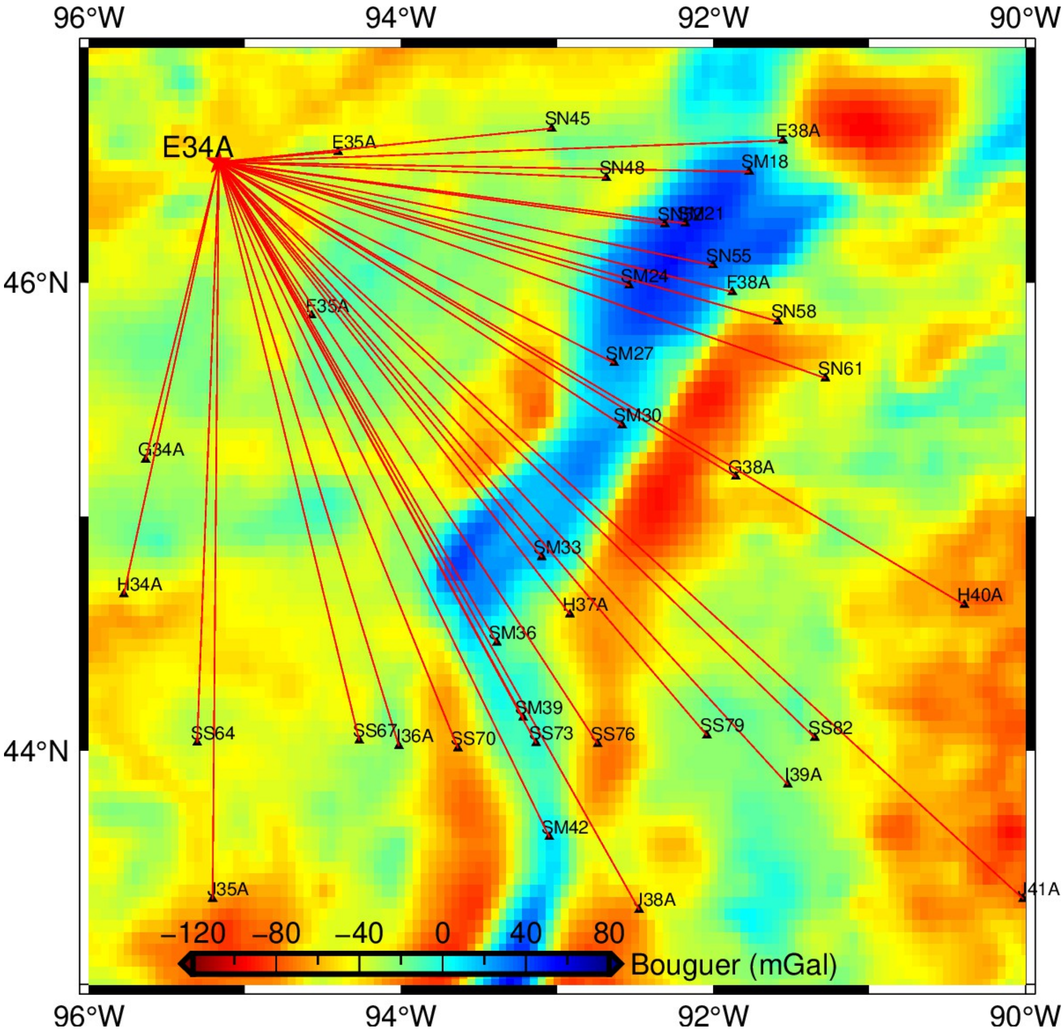


Figure 10. 35 seismic stations used for comparisons shown in Figures 11-13. The SM, SN and SS stations are parts of the Superior Province Rifting Earthscope Experiment (Wolin et al., 2015). Other stations come from the USArray Transportable Array. The background color is the Bouguer gravity anomaly (Woollard, 1965), where the linear feature with positive (blue) values highlights the extension of the Mid-continent Rift.

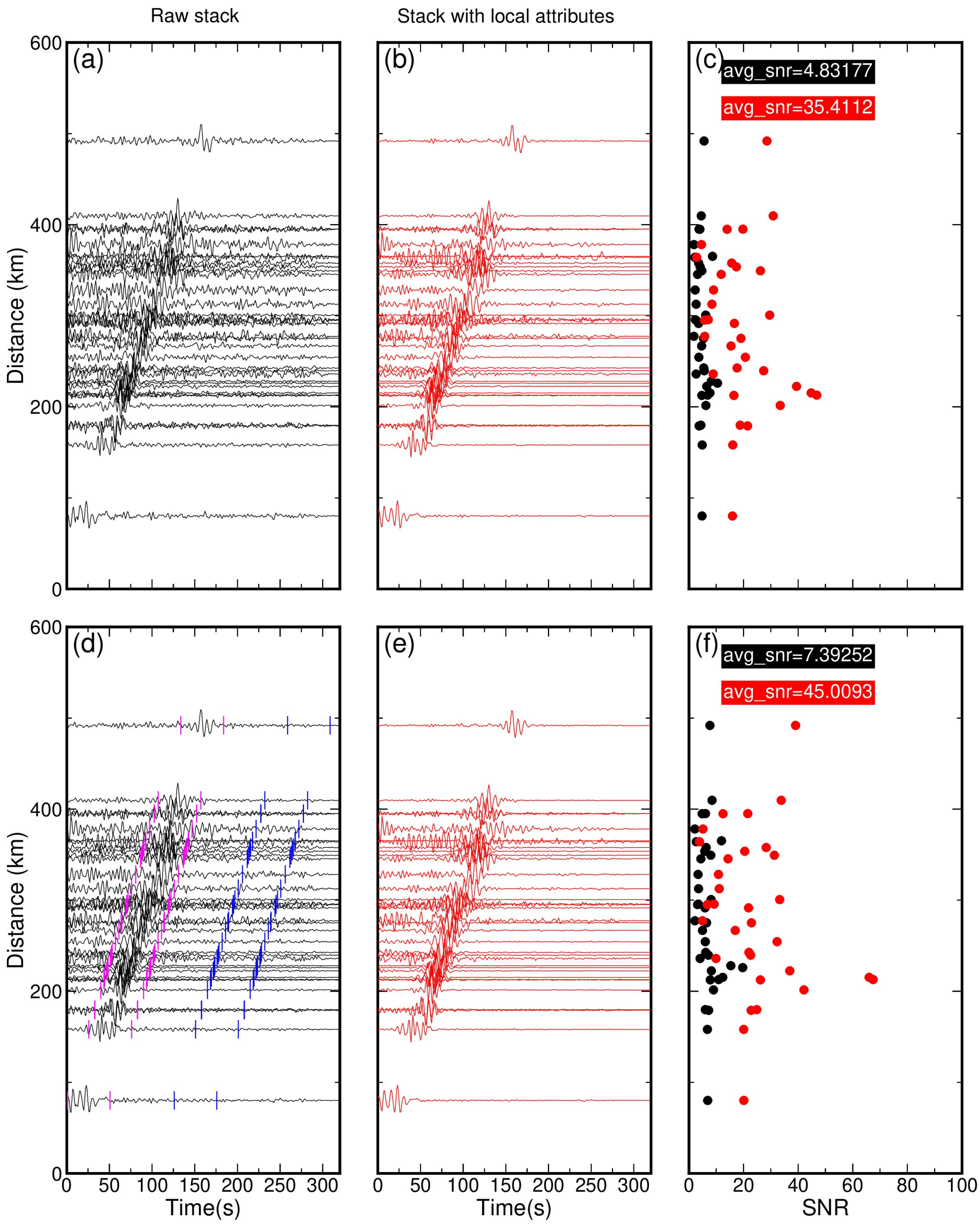


Figure 11. Comparisons of one-month stacked CCFs averaged from negative and positive lags (without bandpass filter) from conventional approach (a) and proposed procedure (b). The improvement of averaged SNR for each trace is shown in panel (c) with black (conventional) and red (proposed) dots. Panels d-f are the same as panels a-c except for two-month stacking. The short magenta and blue solid lines in panel (d) represent the signal and noise windows used to calculate the SNRs shown in panels c and f.

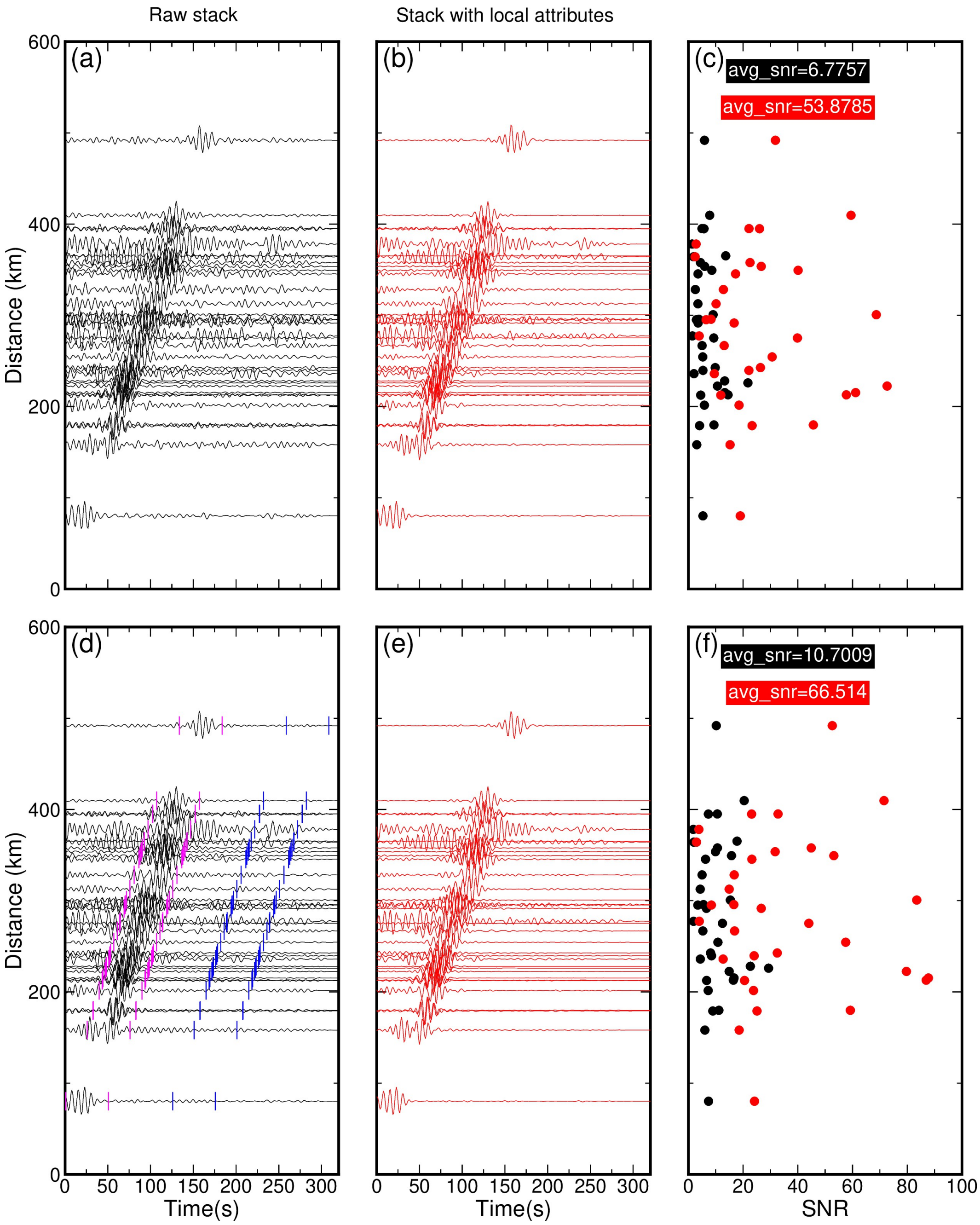


Figure 12. The same setting as Figure 11, but for the stacked data bandpass filtered between 6-15 s.

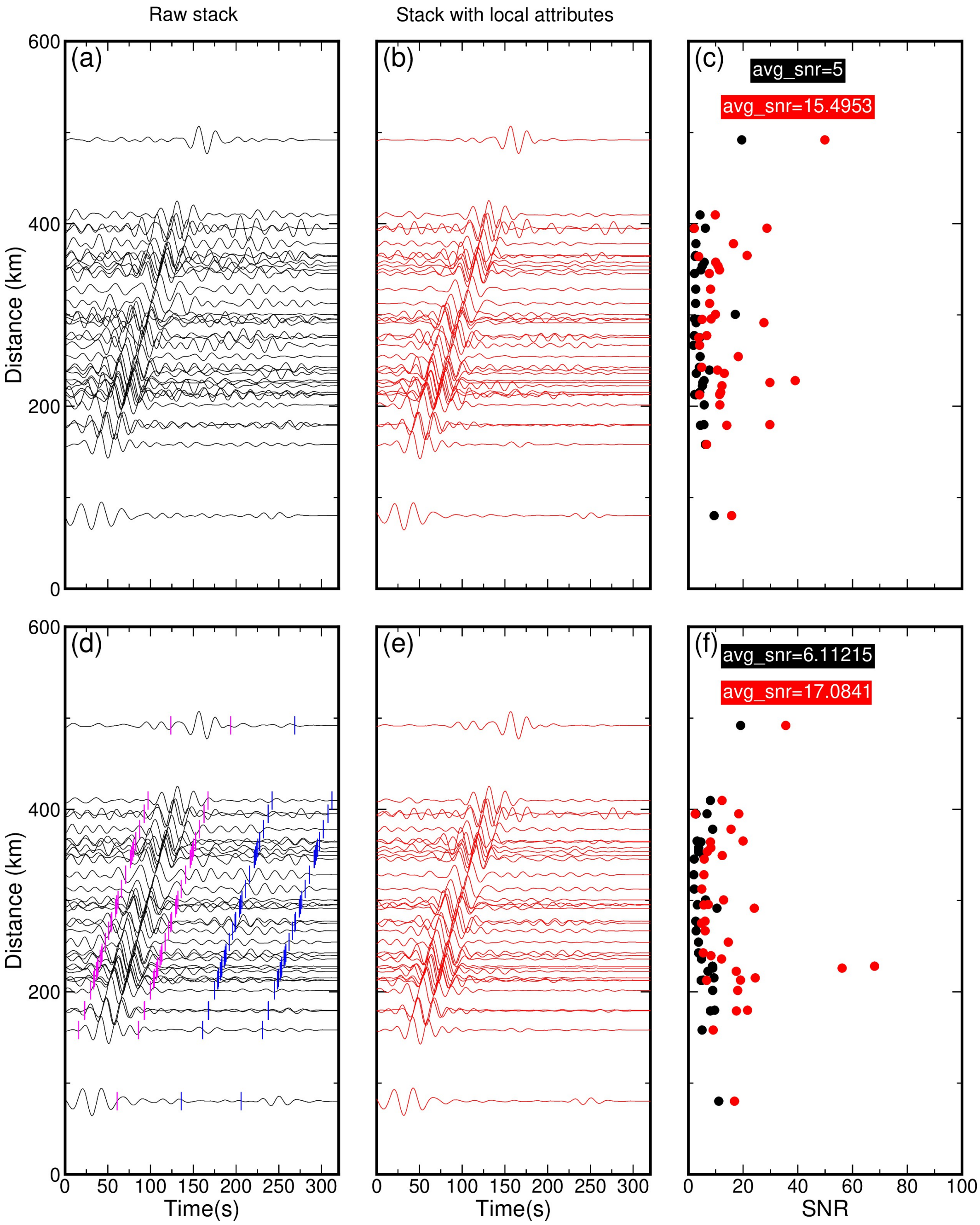


Figure 13. Same as Figure 11 but from the stacked data bandpass filtered between 15-30 s.

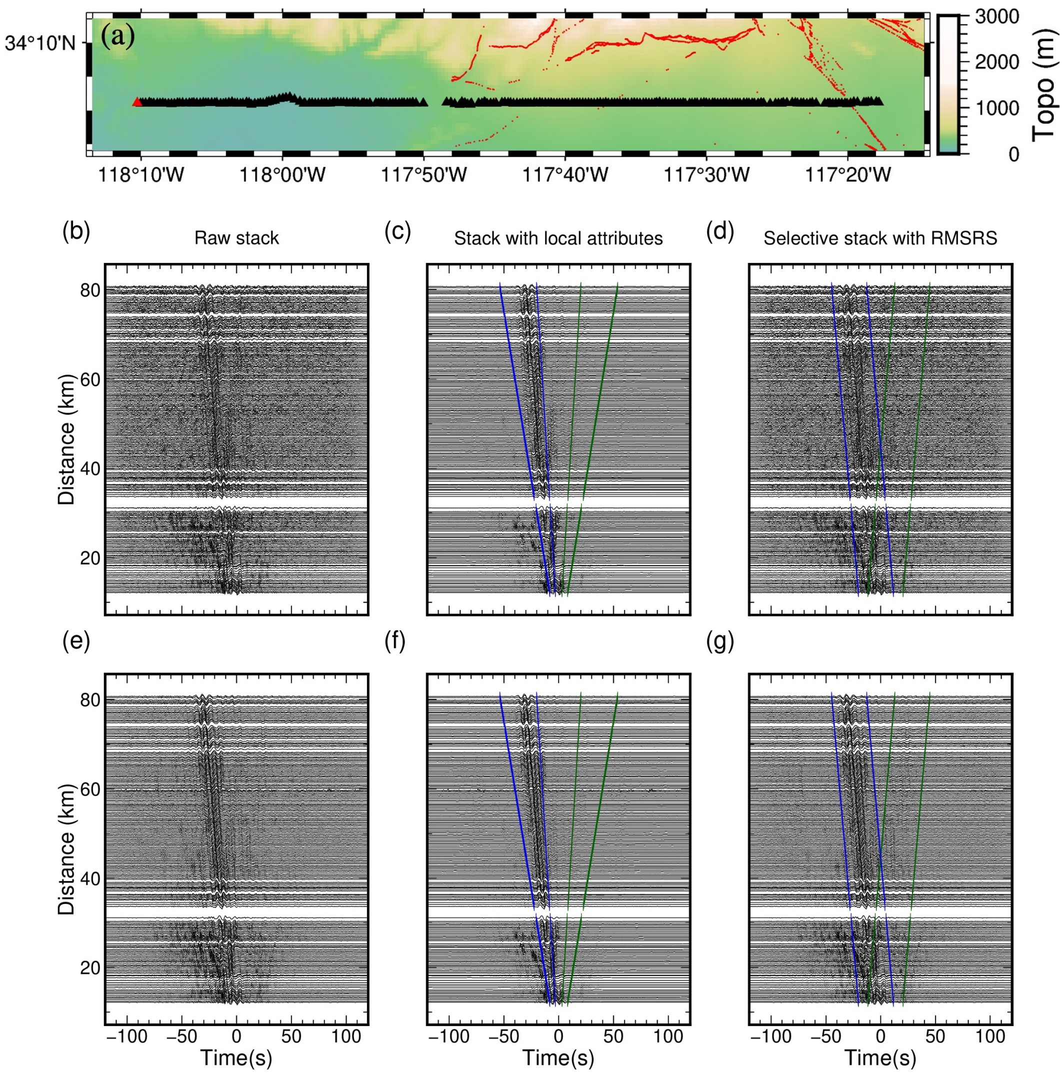


Figure 14. (a) Station distribution of one dense array deployed in the San Bernardino basin. The red triangle denotes the master station while the other black ones are stations used to calculate the CCFs. The red dots represent faults from the U.S. Geological Surveys. Shot gathers of 5 days (b) linear, (e) proposed local and (d) root-mean-square ratio based selective (RMSRS) (Xie et al., 2020) stacking. The blue and dark green lines in panels c and e are acausal and causal arrival times with a group velocity of 1.5 km/s and 3.5 km/s, respectively. They are used to highlight the potential ranges of Rayleigh wave arrivals. The blue and dark green lines in panels d and g are used to define the signal windows for root-mean-square ratio calculation. Panels (e-g) display the corresponding 30-day stacking results.

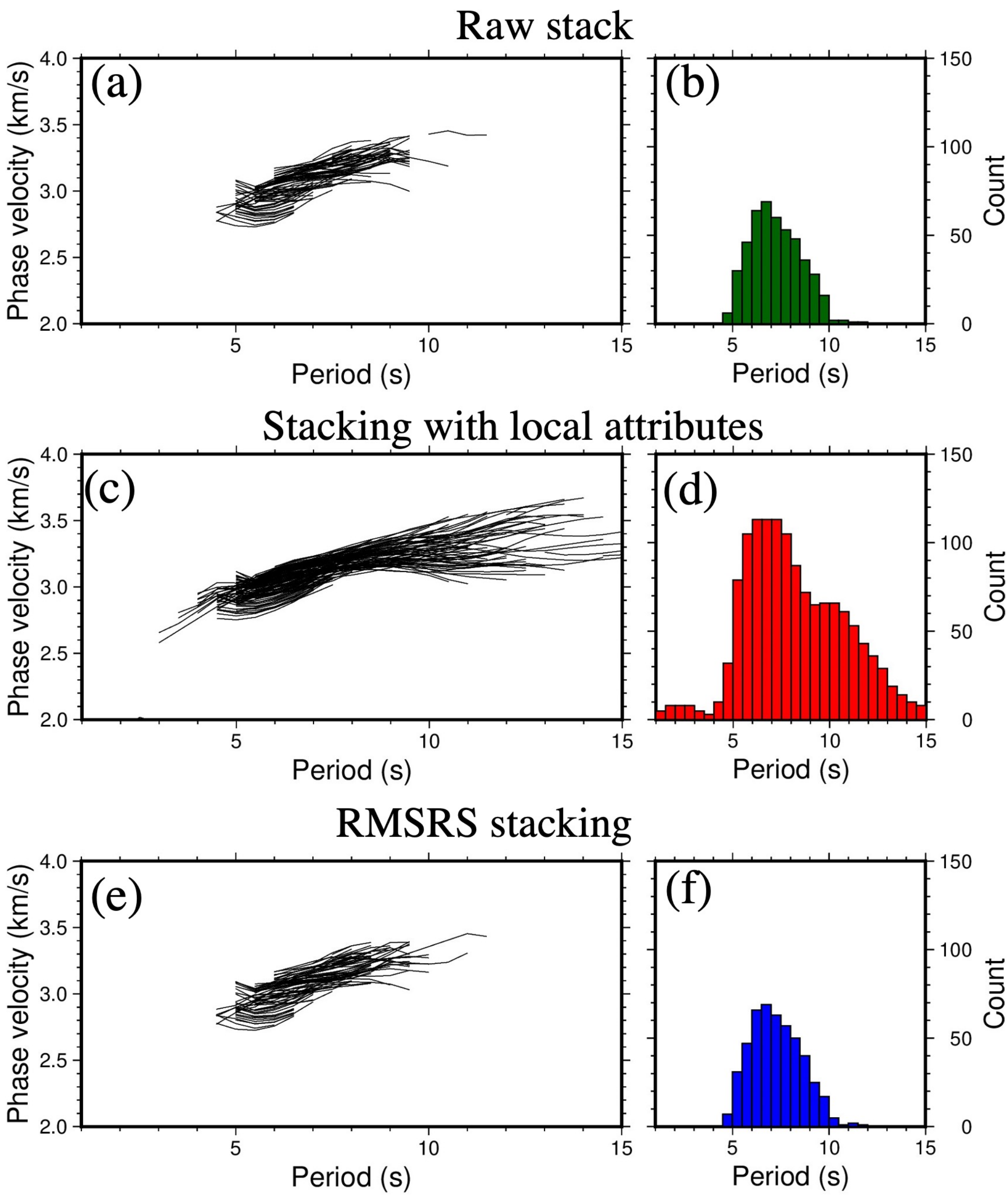


Figure 15. Comparison of selected phase dispersion curves from 5-day raw (a), proposed (b) and RMSRS (c) stacking CCFs displayed in Figures 14b-d. Panels (b), (d) and (f) show the number of measurements for every 0.5 s from 1.0 s to 15.0 s.

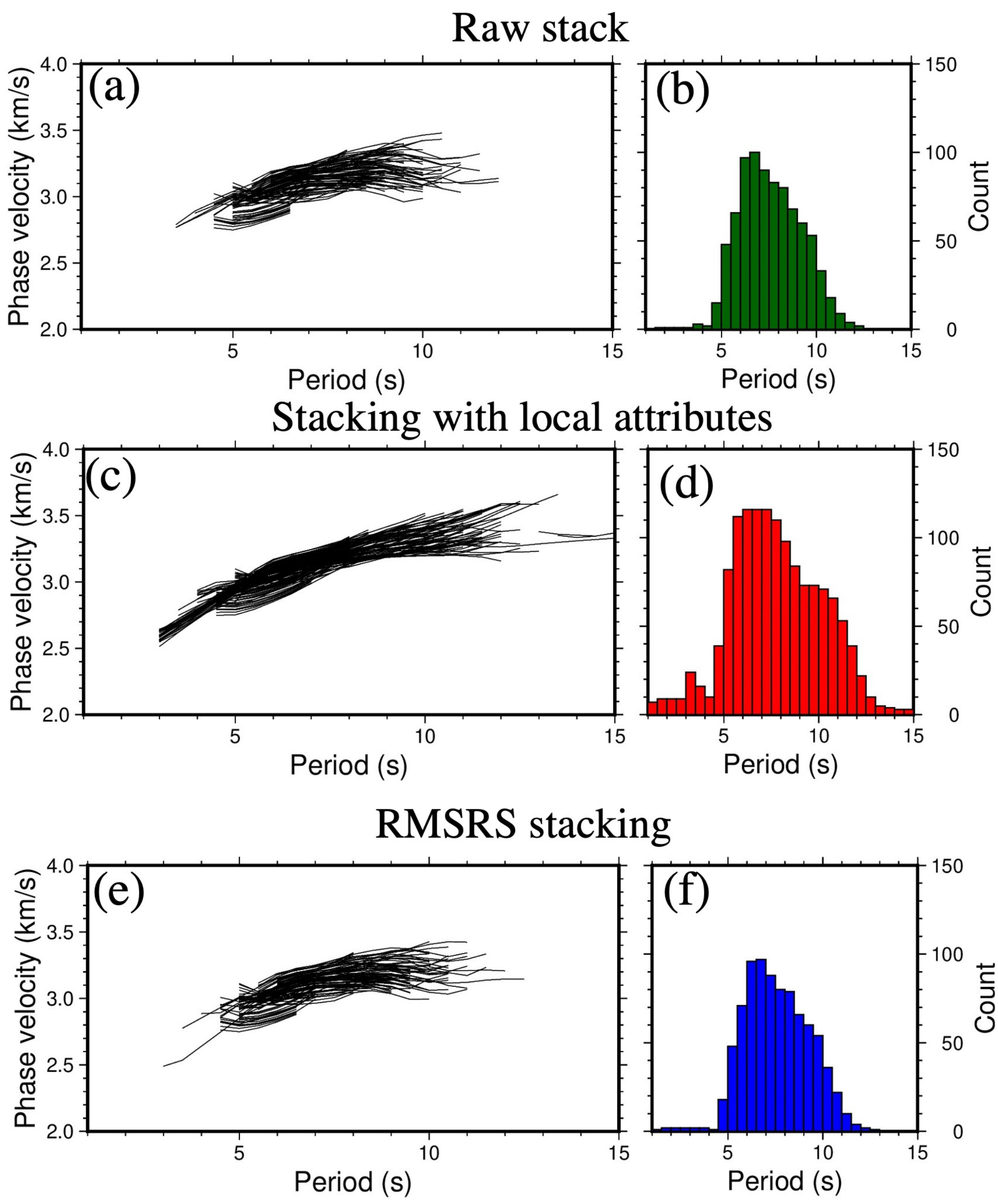


Figure 16. Similar to Figure 15 but from 30-day stacking CCFs displayed in Figures 14e-g

.

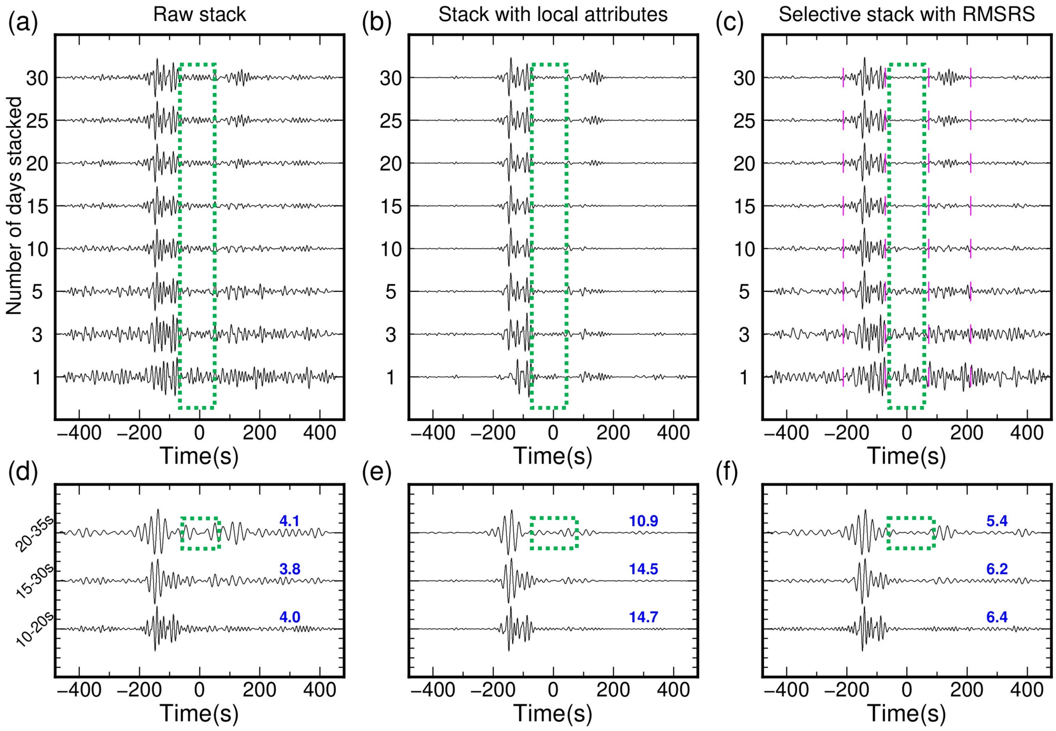


Figure 17. Comparisons of convergence for conventional (a), the proposed procedure (b) and root-mean-square-ratio selective (RMSRS) (Xie et al., 2020) stacking of the TA.A12A-TA.A18A station pair. Panels (d-e) compare 15-day stacked CCFs of the three stacking methods at 10-20, 15-30 and 20-35 s period bands. The blue numbers are the SNR for the corresponding trace. To calculate the SNR, we choose 57-221 s as the signal window and 221-357 s as the noise window. The magenta short lines (57-221 s) are used to define the signal windows for RMSRS calculation. The green dashed rectangles are used to highlight the improvement of the RMSS stacking method. Both RMSRS stacking and our proposed approaches help us to improve the SNR at different period bands.