JOURNAL OF GEOPHYSICAL RESEARCH: SOLID EARTH, VOL. 118, 4345–4360, doi:10.1002/jgrb.50339, 2013

Seismic exploration-scale velocities and structure from ambient seismic noise (> 1 Hz)

Deyan Draganov,[1] Xander Campman,[2] Jan Thorbecke,[1] Arie Verdel,[3] and Kees Wapenaar[1]

Received 16 May 2013; revised 27 July 2013; accepted 5 August 2013; published 30 August 2013.

[1] The successful surface waves retrieval in solid-Earth seismology using long-time correlations and subsequent tomographic images of the crust have sparked interest in extraction of subsurface information from noise in the exploration seismology. Subsurface information in exploration seismology is usually derived from body-wave reflections > 1 Hz, which is challenging for utilization of ambient noise. We use 11 h of noise recorded in the Sirte basin, Libya. First, we study the characteristics of the noise. We show that the bulk of the noise is composed of surface waves at frequencies below 6 Hz. Some noise panels contain nearly vertically traveling events. We further characterize these events using a beamforming algorithm. From the beamforming, we conclude that these events represent body-wave arrivals with a fairly rich azimuthal distribution. Having body-wave arrivals in the noise is a prerequisite for body-wave reflections retrieval. We crosscorrelate and sum the recorded ambient-noise panels to retrieve common-source gathers, following two approaches—using all the noise and using only noise panels containing body-wave arrivals likely to contribute to the reflections retrieval. Comparing the retrieved gathers with active seismic data, we show that the two-way traveltimes at short offsets of several retrieved events coincide with those of reflections in the active data and thus correspond to apexes of reflections. We then compare retrieved stacked sections of the subsurface from both approaches with the active-data stacked section and show that the reflectors are consistent along a line. The results from the second approach exhibit the reflectors better.

Citation: Draganov, D., X. Campman, J. Thorbecke, A. Verdel, and K. Wapenaar (2013), Seismic exploration-scale velocities and structure from ambient seismic noise (> 1 Hz), *J. Geophys. Res. Solid Earth*, *118*, 4345–4360, doi:10.1002/jgrb.50339.

1. Introduction

[2] Ambient-noise seismic interferometry (ANSI) is a recently developed method that allows extraction (retrieval) of the Green's function between two receivers from seismic noise, as if one of them were a virtual source. *Wapenaar and Fokkema* [2006] and *van Manen et al.* [2006] show that the exact Green's function can be obtained from a sum of crosscorrelations of the traces at the receivers due to individual monopole and dipole sources, provided that the receivers are completely surrounded by a dense, homogeneous source distribution. Hence, the successful extraction of the relevant part of the Green's function from ambient noise relies entirely on the characteristics and distribution of the noise sources. Unfortunately, in practice, the strong assumptions on the noise sources are not easily met.

[3] Till now, ANSI is mainly applied to noise recordings in the frequency band of the secondary microseism approximately between 0.1 and 0.4 Hz. [e.g., *Shapiro and Campillo*, 2004; *Sabra et al.*, 2005; *Yang et al.*, 2007]. One reason is that the secondary microseism usually dominates noise recordings. The abundant availability of data from regional and local seismic networks and the increased resolution of tomographic images of the crust afforded by surface wave ANSI has produced a spectacular uptake of this method in the solid-Earth seismology community [e.g., *Shapiro and Campillo*, 2004; *Yao et al.*, 2006, *Harmon et al.*, 2008; *Lin et al.*, 2008; *Yang et al.*, 2007].

[4] In the exploration community, *Claerbout* [1968] proposed to use autocorrelation [and later crosscorrelation, see *Rickett and Claerbout*, 1996] with noise for retrieving the reflection response. Claerbout's idea has been used in global seismology [see, e.g, *Scherbaum*, 1987; *Kawase et al.*, 2011], but only a few studies have focused on ANSI for seismic exploration [e.g., *Baskir and Weller*, 1975; *Draganov et al.*, 2007, 2009; *Nakata et al.*, 2011; *Xu et al.*,

[1]Department of Geoscience and Engineering, Delft University of Technology, Delft, Netherlands.
[2]Shell Global Solutions International B.V., Rijswijk, Netherlands.
[3]TNO Geo Energy, Utrecht, Netherlands.

Corresponding author: D. Draganov, Department of Geoscience and Engineering, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, Netherlands. (d.s.draganov@tudelft.nl)

©2013. American Geophysical Union. All Rights Reserved.
2169-9313/13/10.1002/jgrb.50339

2012]. A reason for the limited application of ANSI in seismic exploration is the lack of data. Another reason is that to retrieve Green's functions relevant for the traditional exploration purposes, the noise must include body-wave energy and have a bandwidth relevant for exploration (> 1 Hz). Probably the most significant problem in the application of ANSI to the exploration scale is the lack of understanding of the properties of the body-wave noise in this frequency range.

[5] Body-wave phases have been detected in the noise also at frequencies > 1 Hz [e.g., *Toksöz and Lacoss*, 1968; *Zhang et al.*, 2009; *Koper et al.*, 2010]. These high-frequency body waves originate from local or regional sources and may propagate as crustal *P* wave phases such as P_n, *PmP*, or P_g (see *Kulhánek* [2002] for a description of these phases and their velocities). Of course, *S* wave phases can also be recorded, but since in exploration seismics mainly vertical-particle-velocity geophones are used and small slownesses are of interest, the *S* phases are not likely to interfere.

[6] As can be expected in practical applications, the retrieval of body waves has proven more difficult than the retrieval of surface waves. The amplitude decay with distance of surface waves is smaller than that of body waves, and so in the recorded noise, the surface waves normally completely dominate. As stated above, there is evidence that in certain frequency bands, body waves are abundant or even dominant. Correlating recorded noise in such frequency bands allows retrieval of body waves. For example, *Roux et al.* [2005] and *Gerstoft et al.* [2008] retrieved diving *P* waves in the band of the double-frequency microseismic peak. More recently, *Ruigrok et al.* [2011] correlated ambient seismic noise recorded in Egypt to retrieve (Moho) reflected waves for frequencies between 0.03 and 1 Hz. *Zhan et al.* [2010] retrieved Moho-reflected *S* wave phases (SmS) between 0.1 and 1 Hz using stations near the critical distance for the retrieved phase, because at such distances, the SmS phase becomes comparable in amplitude to the amplitudes of surface waves.

[7] At frequencies > 1 Hz, surface waves due to surface processes become dominant. To retrieve body waves in this case, one would need to suppress the surface waves. Limiting the presence of anthropogenic surface waves by using ambient noise recorded during the night hours in a desert area, *Draganov et al.* [2007] retrieved reflection arrivals between 2 and 10 Hz. *Draganov et al.* [2009] showed preliminary results of retrieval of reflection arrivals between 6 and 24 Hz by advantageously using patterns of geophones that suppressed the dominant surface waves above 5 Hz in the field. However, body-wave extraction at frequencies > 1 Hz remains a challenge. There is also the question what mechanisms contribute direct and reflected body waves to the ambient noise. The local and regional phases, such as the diving *P* wave (P_g) and *P* wave reflected from the Moho (*PmP*), could contribute to retrieval of reflected *P* waves. Another contribution could come from reflected waves due to sources at or very close to the surface. Seismic exploration geometries generally have relatively short source-receiver offsets (< 10 km) to maximize the amount of reflection energy recorded. Hence, surface noise sources should be located at relatively short offsets from the receivers in order to extract reflection data from depths shallower than, say, 5 km which is approximately the range of interest for hydrocarbon exploration.

[8] In this paper, we present an overview of our efforts to apply ANSI on the scale of exploration seismics. We start with a short review of the theory of ANSI. We then describe the survey area and the recording parameters of the survey geometry. After that, we present a study of the characteristics of the noise to assess its usefulness for body-wave ANSI. Then, we discuss the processing involved in extracting the reflected-wave part of the Green's function. To obtain the body-wave responses, we follow two approaches: (1) following *Draganov et al.* [2009], we correlate and sum all recorded noise and determine whether we observe body-wave reflections in the results, and (2) we apply ANSI only to noise panels containing identified body-wave arrivals [*Draganov*, 2007] after characterizing the slowness of the arrivals, as such knowledge might enhance the signal-to-noise ratio of the retrieved reflections (see *Ruigrok et al.* [2011] for a more extensive discussion of this procedure). We judge the success of the two approaches by a characteristic of the reflections—hyperbolic traveltime difference between two offsets, and by direct comparison with active shot records where available. In the end, we describe the utilization of the retrieved results from ANSI from both approaches for extraction of velocity information of the subsurface and for obtaining an image of the subsurface structure.

2. Method

[9] Retrieval of the Green's function between two points from ambient seismic noise is achieved using the relation [*Wapenaar*, 2004]

$$\left\{G_{p,q}(\mathbf{x}_\mathcal{A}, \mathbf{x}_\mathcal{B}, t) + G_{p,q}(\mathbf{x}_\mathcal{A}, \mathbf{x}_\mathcal{B}, -t)\right\} * \mathcal{F}(t) \approx \sum_i^N v_p^i(\mathbf{x}_\mathcal{A}, -t) * v_q^i(\mathbf{x}_\mathcal{B}, t). \quad (1)$$

The right-hand side of the relation denotes the convolution (the asterisk sign) of time-advanced and time-retarded observations, that is, it represents crosscorrelation. The observations denote particle velocities $v_p^i(\mathbf{x}_\mathcal{A}, t)$ and $v_q^i(\mathbf{x}_\mathcal{B}, t)$ at points $\mathcal{A}$ and $\mathcal{B}$ ($\mathbf{x}$ stands for the coordinate vector), measured in the x_p and x_q directions, respectively, while the superscript i denotes the number of the consecutive observation (noise panel). In the left-hand side, we have Green's function, in terms of particle velocity, measured in the x_p direction at point $\mathcal{A}$ due to an impulsive force in the x_q direction at point $\mathcal{B}$. Relation (1) states that the superposition of the Green's function and its time-reversed version between two points can be retrieved from the crosscorrelation and summation of N noise panels. Note that the so-obtained Green's functions are filtered by the autocorrelation $\mathcal{F}(t)$ of the source function. The relation between the two sides is approximate as several assumptions have been made in the derivation: the measurements are in the far field of the sources; the medium parameters around the noise-source boundary are smoothly varying. When these assumptions are not met, the retrieved Green's functions will have amplitude errors [*Wapenaar and Fokkema*, 2006]. In the original formulation, an ensemble average is taken of the crosscorrelations in the right-hand side of equation (1). In practice,

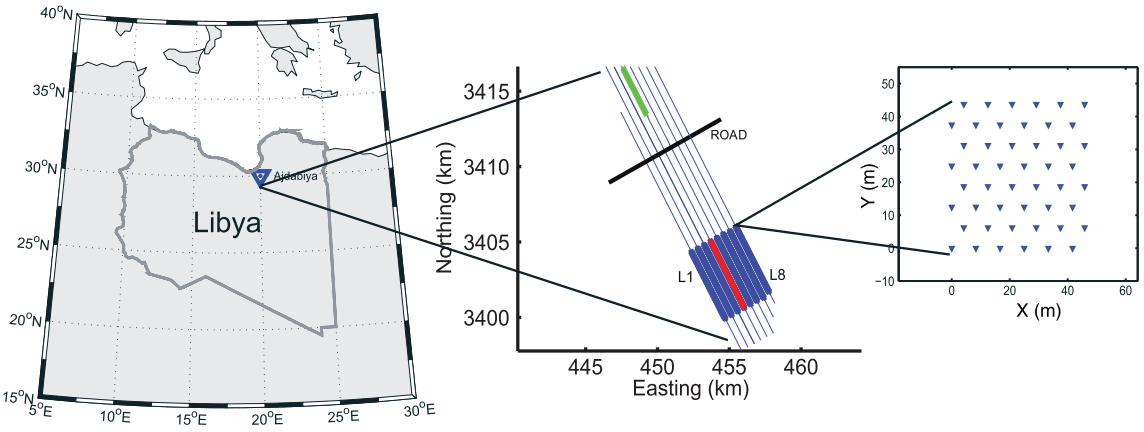


Figure 1. (left) Approximate location of the survey area near the town of Ajdabiya, Libya. (middle) Recording geometry with each of the eight lines consisting of around 400 receiver stations. The fat black line indicates a road. The colored parts of the lines indicate the receiver stations used in a line analysis of the noise (green and red with 70 and 100 stations each, respectively) and in an areal analysis of the noise (blue with 100 stations per line). (right) Each receiver station is represented by a pattern of 48 geophones.

the averaging is performed over different time intervals and the ensemble average is replaced by summation.

[10] The theory requires that the noise sources, primary or secondary (for example, due to scattering), illuminate the observation points with equal strength from all directions, meaning that the noise field should be equipartitioned. Furthermore, it also requires sufficient density of uncorrelated noise sources (at least two sources per wavelength for a general inhomogeneous medium). When the two mentioned conditions are not fulfilled, phase errors [*Froment et al.*, 2010] and nonphysical arrivals [*Snieder et al.*, 2006] will be introduced in the retrieved results.

[11] In field applications for retrieval of body-wave reflections, these requirements on the noise sources are almost never met. The noise sources that would contribute to reflection retrieval will be illuminating the recording stations from limited directions and will be irregularly distributed in space and time. On the other hand, the sources of surface wave noise will be more frequently active and will more easily result in an equipartitioned field. This means that correlating and summing all recorded noise panels will most likely retrieve the surface waves as dominant arrivals, while retrieved reflections will be drowned under the retrieved surface waves. For this reason, we also apply ANSI relation (1) only to noise panels that contain body-wave arrivals. *Roux et al*. [2005], for example, shows that the main contribution to the extracted Green's function comes from points that lie in a stationary-phase region for the integrand of the correlation summation. In physical terms, it means that the noise source should lie on (or close to) a line connecting the receivers (for direct waves) or in a region where it produces a specular reflection between the two relevant receivers (for reflections). Thus, using for ANSI only noise panels that have body-wave arrivals would lower the chances of capturing sources that contribute to retrieval of surface wave, while the main noise-source contributions for reflection retrieval would still be kept. The resulting retrieved surface waves would be suppressed and the retrieved reflections would be relatively stronger. Furthermore, using only selected parts of the noise would allow separate handling after crosscorrelation and before summation to those of the correlated noise panels. Such processing could include removal of the autocorrelation of the source function and summation of the retrieved positive and negative parts of the correlation result per noise panel. *Ruigrok et al.* [2011] apply such an approach for retrieval of *P* wave reflections in the frequency band between 0.4 and 1 Hz.

3. Survey Area and Recording Parameters

3.1. Survey Area

[12] In September 2007, Shell recorded approximately 11 h of ambient seismic noise during one night and during a 3-D seismic-exploration campaign in the Eastern part of the Sirte Basin near the town of Ajdabiya in Libya (Figure 1).

[13] Libya is generally not considered an active seismic area although several large earthquakes have occurred in the past. The most active part is the eastern flank of the Hun graben in the western part of the Sirte Basin. The eastern part of the basin is quieter with no activity in the basin itself [*Suleiman and Doser*, 1995]. The Cyrenaica platform, to the NE of the survey area, is also an area of recent seismicity. Not much is known (at least to the authors) about the specific characteristics of the ambient seismic noise in Libya, since only few seismometers have been deployed in the past.

[14] The proximity of Libya to the Mediterranean sea suggests that noise induced by storms on the Mediterranean may also contribute to the ambient seismic-noise environment (see *Ruigrok et al.* [2011] for example). Such noise sources could generate low-frequency crustal phases such as Lg (note that we are interested in higher frequencies), but also body-wave phases with higher frequencies [*Zhang et al*., 2009; *Koper et al.*, 2010].

3.2. Recording Equipment and Geometry

[15] The recording geometry is shown in the middle and right pictures in Figure 1. A fairly busy blacktop road (the black line) cuts almost perpendicularly through the receiver lines. The receiver spread consists of eight parallel lines with 500 m spacing between them. Each line consists

of approximately 400 receiver stations, the number varied slightly per line, with spacing Δx = 50 m. Each receiver station consists of 48 geophones in a staggered pattern of 50 × 50 m^2 with inline geophone spacing of dx = 6.25 m and crossline spacing of dy = 8.3 m. The signals from the 48 geophones are summed in the field to form a single output at the location of the receiver station. Hence, the pattern acts as a resampling operator from the pattern interval dx to the station interval Δx. In current seismic-exploration practice, these patterns are also designed to act as anti-alias filters for both the noise and the signal [*Vermeer*, 1990].

[16] The sensors used in this survey are standard 10 Hz vertical-component geophones that are sensitive to particle velocity. Data were acquired on a Sercel 428 recorder with data being recorded continuously and written in 47 s records (the maximum length allowed by the Society of Exploration Geophysicists Y format of the files with 4 ms sampling).

[17] The equipment used in this experiment does not allow a full characterization of the ambient noise below a few Hz. Because of this, we have not attempted to quantify the noise in terms of absolute velocities to compare them with the standard noise models [*Peterson*, 1993]. The response of a standard 10 Hz geophone is usually not adequate to record very low-amplitude, low-frequency signals because of its relatively low sensitivity and the 12 dB/Oct decay below 10 Hz.

[18] Nevertheless, the output from the receiver stations show a strong response down to about 1 Hz. This could be attributed to the fact that the signals from the 48 geophones from the pattern in each station are summed, thereby boosting coherent parts of the wavefield with low wave numbers while suppressing incoherent parts between the 48 sensors. Since these coherent parts consist mainly of the low-frequency part of the strong surface wave noise from the road, this noise dominates our records, notwithstanding the low-frequency characteristics of the 10 Hz geophones.

4. Noise Analysis

4.1. Geophone Pattern, Line Array, and Areal Array

[19] To characterize the noise, we use power spectral density (PSD) and frequency–wave number (FK) power spectral density analysis. The latter provides insight in the propagation velocities of the noise, the propagation directions, and the phase composition of the noise wavefield. Noise properties and sources generally vary in and between different frequency bands, so we cannot expect to determine all propagation modes in all slowness ranges for all frequencies of interest. As discussed, a frequency of about 1 Hz is the low limit determined by the geophone response. However, the recording geometry does offer some flexibility.

[20] On the smallest scale, the geophone patterns actually constitute arrays and we take advantage of their anti-alias properties. Figure 2 shows the inline pattern-slowness response as a function of frequency. The horizontal slowness p is defined as $p = \frac{\sin\gamma}{c}$, where γ is the inclination of the event with respect to the vertical and c is the propagation velocity. We observe that at small slownesses, all frequencies are passed, but that at higher slownesses, only low frequencies are passed such as to prevent aliasing. Effectively, this means that high-frequency, low-velocity surface waves are suppressed by the patterns. The fact that low-frequency surface waves are passed by the array did not cause any problems in the active-source survey, as the used active sources (seismic vibrators) only excited energy above 5 Hz.

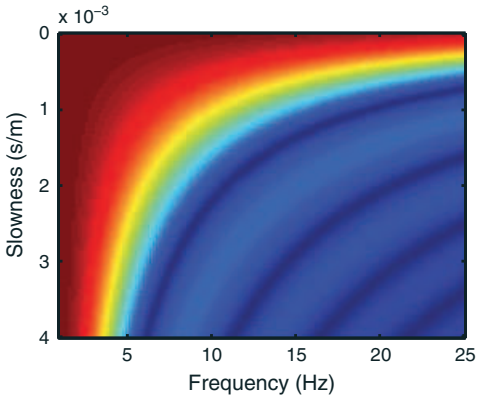


Figure 2. The slowness response of the geophone patterns in the direction parallel to the receiver lines.

[21] For an intermediate-scale analysis, we use line L4 of the survey (see middle picture in Figure 1) also as an array with a pattern sampling of Δx = 50. Because the wavefield propagates in two dimensions along the surface, this array could cause ambiguities in the analysis. However, careful interpretation still yields useful information.

[22] On the largest scale, we consider the eight parallel lines as one areal array with an inline spacing of Δx = 50 m and a crossline spacing of Δy = 500 m. While the geophone patterns ensure alias-free (surface) waves in the inline direction, the crossline sampling interval would result in strong aliasing in the FK response of the low-frequency surface waves that are passed by the geophone patterns. This can be overcome by limiting the slowness and/or frequency range in determining the FK spectrum. We use the areal array for FK analysis using a frequency-domain beamforming (FDB) algorithm [*Lacoss et al.*, 1969] when analyzing possible body-wave energy in the recorded noise.

[23] We limit the inline aperture of both the line and the areal arrays. Array analysis does assume some degree of space stationarity of the noise [*Capon*, 1969; *Aki and Richards*, 1980]. From the PSD analysis, it will become clear that this assumption is met approximately only if the area around the road is excluded from the analysis. An additional advantage of excluding that area is that we can consider the noise emitted by cars on the road as plane waves.

4.2. Power Spectral Densities

[24] Noise at higher frequencies is usually attributed to human activities, local meteorological conditions, such as wind or rain, but also to natural noise like geothermal sources [*Bnefoy-Claudet et al.*, 2006; *Asten and Henstridge*, 1984; *Okada*, 2003; *Wilson et al.*, 2002]. Because the majority of the noise sources are at the Earth's surface, they generate predominantly surface waves. In our case, the road is a cause of strong surface wave noise observed in the records. Figure 3a shows the PSD estimates for all traces along line L4, which is in the middle of the survey area. These spectra are calculated using Welch's method

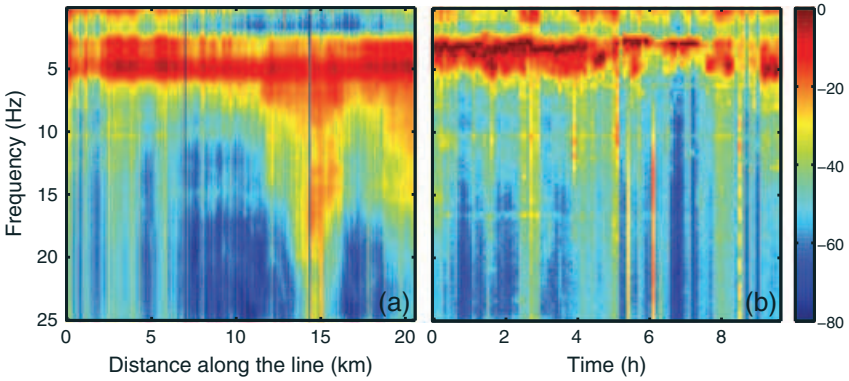


Figure 3. (a) Estimate of the power spectral density of approximately 2 h of noise at every receiver station along line L4, see Figure 1 (middle). (b) Power spectral density as a function of time for station 21 on line L4 (at 1 km from the line's beginning).

[e.g., *Kay*, 1969]. We use approximately 2 h of noise panels with 50% overlap. Most of the noise is concentrated below about 6 Hz. The road can be identified by the increase in amplitude of the noise, which is evident especially at the frequencies up to about 20 Hz around 14 km along the line. The power at frequencies between 20 and 5 Hz then quickly drops away from the road due to the higher attenuation at these frequencies.

[25] Figure 3b shows the PSD as a function of time for station 21 along line L4 (at about 1 km from the line's beginning). We calculate PSDs from each noise panel without overlap. Then they are averaged in a window of 6 min. We do this for almost 10 h of recording. We observe that the noise below 6 Hz is continuously generated, but the power decreases after about 5 h into the recording and increases again after 9 h. The quieter interval corresponds to the nighttime between 23:00 and 03:00 local time.

4.3. Body-Wave Noise

[26] To retrieve body-wave reflections from ANSI, body-wave arrivals should be present in the recorded noise. Because of this, we examine the noise for body-wave arrivals and then analyze whether utilization of only selected noise panels with body waves in them could contribute to the retrieval of the reflected-wave part of the Green's function. Part of the noise analysis consists of visual inspection of the nearly 900 noise panels. Figure 4a shows a part of a noise panel as recorded by the receiver stations along the red-marked part of line L4 in Figure 1. In it, we can see the dominant surface waves originating from the road. Such arrivals are representative of the continuous noise. Figure 4b shows the same part of the noise panel after application of a low-cut filter with a cutoff at 6 Hz to suppress the main surface wave noise. Apart from some remaining (aliased)

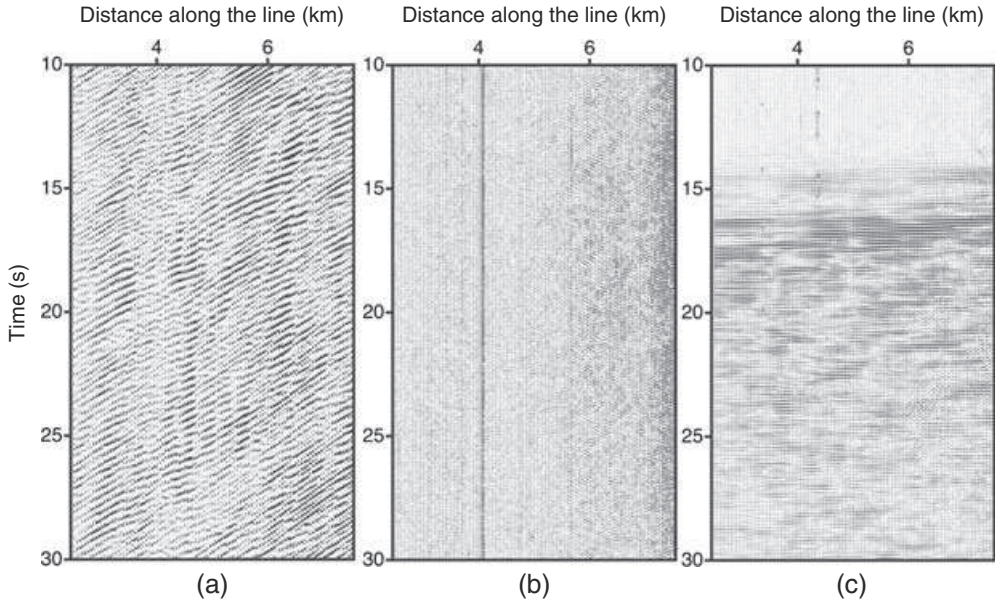


Figure 4. (a) Part of a noise panel without filtering. (b) Part of a filtered noise panel without any noticeable events. We call this continuous noise. (c) Part of a filtered noise panel showing coherently aligned, almost vertically traveling energy. We call this "event." Events occur intermittently. Sometimes they last for several minutes.

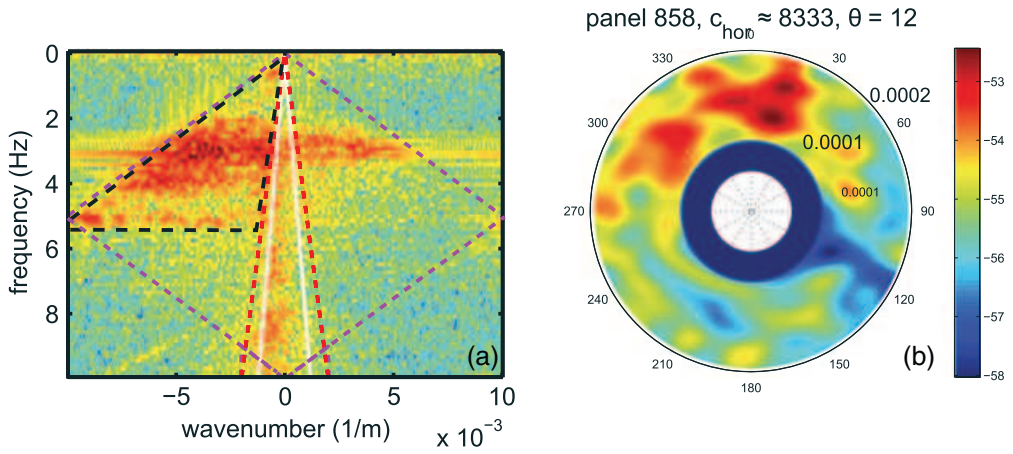


Figure 5. (a) Frequency–wave number power spectrum between 0 and 10 Hz of panel 858 (a part of which is shown in Figure 4c) along the red-marked part of line L4. The dashed magenta line corresponds to the lowest surface wave velocity (approximately 500 m/s). The dashed black line corresponds to the lowest velocity (highest ray-parameter) used for beamforming. The white lines correspond to the horizontal velocity of the event picked from the beam in Figure 5b. The dashed red lines delimit the velocities for retrieval of reflections. (b) Wave number spectrum (summed over frequency bins from 5 to 10 Hz) determined using panel 858 (from Figure 4c). The horizontal velocity corresponding to the dominant slowness is 8333 m/s and the estimated azimuth is 12°.

surface wave noise (the inclined events), there does not appear to be any other coherently aligned energy. To extract reflections, we need (nearly) vertically traveling events and we identify them using visual inspection. Figure 4c shows part of a band-pass filtered noise panel (panel 858) with such an event. To further characterize these events, we estimate their slowness spectra using the FDB algorithm.

[27] For parts of the wavefield characterized by small slownesses, such as body waves, the areal array produces a reliable wave number spectrum. In order not to be biased too much by the array geometry, in the following, we only consider the areal array indicated by the fat blue lines in Figure 1 after downsampling it in the inline direction to $\Delta x = 200$ m.

[28] The FDB output is a function of frequency and two horizontal wave number components (k_x and k_y). We display it for a sum of frequency bins as a wave number spectrum parameterized with the horizontal slowness or ray parameter p and a back azimuth φ. Since the wave number vector $\mathbf{k} = \frac{\omega}{c_{\text{hor}}}(\cos\varphi, \sin\varphi)$, where ω is the radial frequency and c_{hor} is the horizontal velocity, the horizontal slowness is proportional to the magnitude of $\mathbf{k}$, while the back azimuth determines the direction of $\mathbf{k}$ [see also *Rost and Thomas*, 2002]. The horizontal velocity c_{hor} is defined as $c_{\text{hor}} = \frac{c}{\sin\gamma}$.

[29] In order to gain the most complete understanding of the noise wavefield, we first show in Figure 5a the FK spectrum of noise panel 858 determined along line L4. It shows frequency as a function of the apparent wave number $k_{\text{app}} = \frac{\omega}{c_{\text{app}}}$, where $c_{\text{app}} = \frac{c_{\text{hor}}}{\sin\varphi} = \frac{c}{\sin\varphi\sin\gamma}$. The dashed magenta line illustrates the lowest (surface) wave velocity of approximately 500 m/s found in the noise. If no geophone patterns are used as anti-alias filters, the surface waves would have been aliased and would have followed the dashed magenta lines between 5 and 10 Hz as well. The black dashed triangle indicates the area where we would expect the surface wave energy from the panel. Since the surface wave noise is generated by cars along the road, the surface wave energy in this panel could have been generated by several cars at various distances away from the line. In such a case, the FK spectrum of the surface waves will appear to be spread between the dashed magenta line and $k = 0$, as a car at each distance would generate a surface wave propagating with a different apparent velocity (or wave number). This appears to have been the case when this panel was recorded.

[30] The area between the dashed red lines indicates the range of velocities we are interested in—horizontal velocities > 5 km/s ($|p| < 0.2$ s/km) corresponding to P wave components in the noise. Surface waves could also appear closer to the $k = 0$ axes (between the dashed red lines) if they were propagating in the direction perpendicular to the line of receivers. Hence, on an FK spectrum, these could easily be misinterpreted for body waves. This can be resolved by forming a slowness spectrum using the areal array. To reduce computational time, we restrict our frequency range between 5 and 10 Hz. Then, to reduce the potential of aliased surface wave remnants and aliasing of other wave types, we limit the beamforming analysis to velocities falling inside the area delimited by the dashed red lines in Figure 5a. The slowness range limits are such that aliased surface waves will only interfere between 9 and 10 Hz, as shown by the dashed magenta lines crossing the red dashed lines in Figure 5a. If the aliasing is not too strong, this should not be a problem, because the pattern response in Figure 2 shows that the surface waves are sufficiently attenuated at 9–10 Hz by the anti-alias filter. The wavenumber response of the array (along the receiver line) is a sinc function with its first zero at twice the Nyquist wavenumber for sampling at the station spacing. Hence, in the absence of array perturbations aliased noise is completely rejected at $k = 0$. If aliased noise is strong, then the attenuation achieved by the geophone pattern may not be sufficiently strong to prevent the retrieval at small wavenumbers.

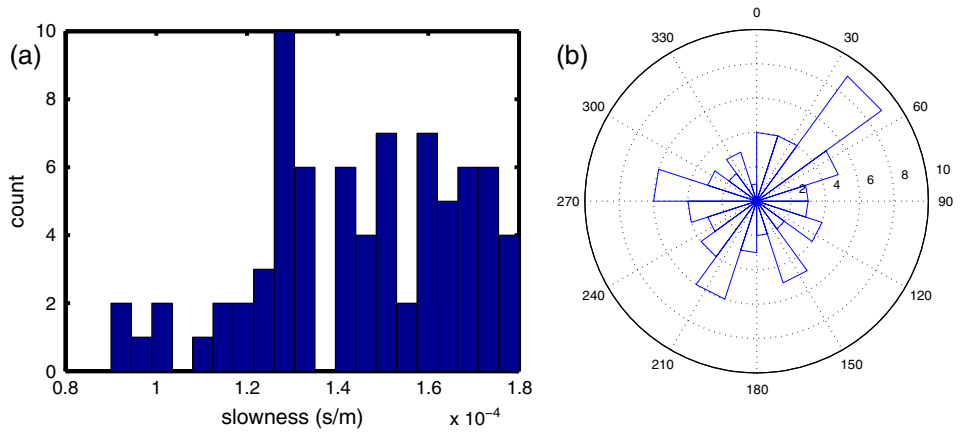


Figure 6. (a) Distribution of slownesses estimated from beamforming of records containing body-wave noise. (b) Rose diagram with the azimuths corresponding to the slownesses in Figure 6a.

[31] Figure 5b shows the FDB output for noise panel 858. We determine the individual FK spectra for small frequency bins and sum the individual spectra to produce the output slowness spectrum. The largest red blob corresponds to an event with a horizontal velocity of 8333 m/s. This is a body wave which, depending on the distance to its source, could contribute to the retrieval of reflections from the noise. The white lines in Figure 5a indicate this velocity in the FK spectrum obtained from line L4.

[32] In this way, we determine FK spectra and FDB output for all records that are visually inspected and determined to contain at least some nearly vertically traveling waves. Figure 6a shows a histogram of all picked slownesses $|p| <$ 0.2 s/km (about 80 records). We only pick the dominant slowness from each slowness spectrum of a particular panel. However, some spectra, such as the one from panel 858 shown in Figure 5b contain two (or more) clear peaks. These secondary peaks may be due to reflections, different phases from the same sources or wavefields from another source being captured at the same time as the dominant wavefield. Hence, there could be more events with different slownesses present in the data. From Figure 6a, we observe that there is a spread in the slownesses, with one slowness being dominant. This slowness corresponds for the larger part to a train of fairly strong events coming from the same direction in the time span of a few consecutive records. We also note that the azimuthal distribution of the noise is quite diverse, see Figure 6b. Together, these observations suggest that we should be able to retrieve body-wave reflections by applying ANSI to the recorded noise. They also suggest that we should be able to retrieve the reflections even only from correlating and summing the noise panels with the identified events in them albeit with a limited move-out because of the limited slowness range (which is effectively the noise-source aperture).

[33] We note that it is as yet unclear what and where are the sources for the body waves seen in the noise. Considering the short-period contents and the fact that velocities of most of the events lie between approximately 5.5 and 8 km/s, the identified events might correspond to crustal phases. A more comprehensive evaluation of the measured body waves would have been possible with 3-component short-period seismometers. With the current equipment, it is not possible to identify any clear phases in the events.

[34] Having obtained an understanding of the noise environment and of the effective illumination of the array, we proceed to apply ANSI to the data.

5. Application of ANSI for the Extraction of Body-Wave Reflections

5.1. Noise Preprocessing for Body-Wave Retrieval

[35] To apply relation (1) to the recorded ambient noise, we follow the procedure outlined in Figure 7. We begin the processing by applying a band-pass filter between 6 and 24 Hz to the recorded noise. The lower limit of this filter is aimed at suppression of the present surface wave energy and is dictated by the dominance of the surface waves that were not suppressed by the geophone patterns; it is also dictated by the desire to preserve body-wave arrivals.

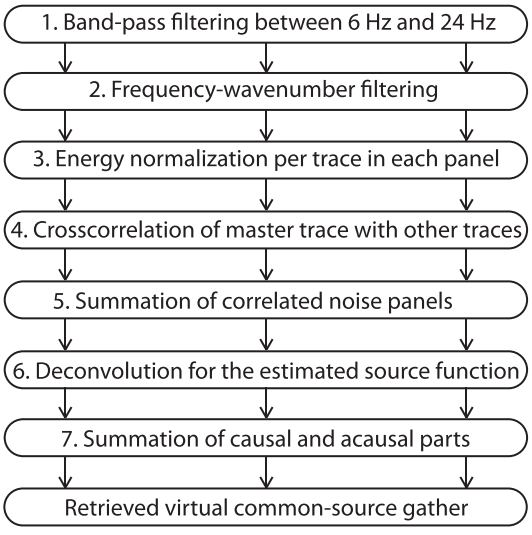


Figure 7. Processing flow for the application of seismic interferometry by crosscorrelation to the recorded ambient seismic noise.

[36] The upper limit of the filter is dictated by the lack of useful energy above 24 Hz and the presence of high-amplitude peaks in the spectrum. The latter may have been caused by electromagnetic noise due to the power lines that were feeding the equipment, and thus, these peaks do not carry information of the subsurface. In the beamforming analysis, we placed an upper limit of 10 Hz to avoid misinterpretation of aliased waves, where the aliasing depends on the station spacing. The correlation in the ANSI is a point-to-point process and does not suffer from distance limitations (even though correlation of aliased arrivals will result in retrieval of aliased arrivals).

[37] As a second step in the processing flow, we apply an FK fan filter for waves with apparent velocities between 230 and 1500 m/s. Such a filter has the potential of suppressing remaining surface wave energy (higher modes) and guided wave types that have the same frequency content as the body-wave noise; it could also suppress possible retrieval of direct body waves. From the active data, we know that the surface wave velocities are around 500 m/s, but we take wider limits to suppress also surface wave energy coming from the crossline directions.

[38] Relation (1) is derived assuming uncorrelated noise sources that act simultaneously and have equal amplitudes. From the noise analysis, we know that most of the body-wave noise is recorded at different time intervals. For this reason, the longer the recording, the bigger the chance that body-wave noise from more sources would be captured. On the other hand, the noise sources might be emitting energy with different strength at different times. Because of this, the noise in some panels might happen to be magnitudes stronger than in others. In the summation process after crosscorrelation, such panels would give by far a dominant contribution, and this would effectively mean using a much lower number of summed correlated noise panels. To prevent this, as a third step in our processing, we apply per noise panel normalization to each of the traces by dividing over the trace's energy (energy normalization). This normalization in the time domain is comparable to the running-absolute-mean normalization in *Bensen et al.* [2007] when the running window is taken equal to the length of one noise panel. The difference is that we use the root mean square of the energy of each trace to normalize the amplitudes in this trace. In this way, the relative amplitudes in each trace are kept the same, while among traces, the amplitudes are equalized.

5.2. ANSI With All the Recorded Noise

[39] After the above three steps, we are ready to apply relation (1). We choose a receiver at which we want to create a virtual source; we call the trace recorded by this receiver "master." We take a noise panel, extract the master trace, and correlate the latter with all the traces in the same noise panel (the fourth step in Figure 7). This is repeated for all noise panels. The so-obtained correlated noise panels are summed together per trace location (the fifth step in Figure 7). If the noise were caused by white-noise sources, then we would retrieve the Green's function and its time-reversed variant (or parts of them). In practice, the sources are always band limited, so the retrieved result will be filtered by (convolved with) the autocorrelation of the source function ($\mathcal{F}(t)$ in relation (1)). To compensate for this, *Bensen et al.* [2007] propose to apply spectral whitening before correlation. We compensate for $\mathcal{F}(t)$ after correlation—we deconvolve the retrieved result by an estimate of the autocorrelation of the source function (step six in Figure 7). The estimate is obtained by extracting a short window around $t = 0$ s from the autocorrelated master trace. Then we use the spectrum of this time window to divide by the spectrum of the result from the fifth step. Note that the processing we propose of energy normalization before correlation and wavelet deconvolution for the source function after correlation is akin to the utilization of the coherence function [see, e.g., *Prieto et al.*, 2009; *Nakata et al.*, 2011].

[40] If the ambient-noise sources illuminate the receiver stations from all directions with comparable strength (homogeneous illumination), then the stationary-phase regions of all events would be sampled, and the Green's function and its time-reversed variant would be retrieved equally well. From the noise analysis in section 4.3, we know that the receivers are preferentially illuminated by body-wave noise from several directions. This means that some parts of the Green's function might be retrieved only at positive times, while others only at negative times. In such cases, more complete results are obtained when the causal and time-reversed acausal correlation results are summed together and so we do that (step seven in Figure 7).

[41] Figures 8a, 8d, and 8f show examples of retrieved common-source gathers along line L4 for virtual sources at 1, 2.5, and 13 km, respectively. In common-source gathers from active surveys, reflection arrivals from subsurface layers are characterized by hyperbolic moveout. (For one horizontal layer in the subsurface, the relation between the traveltime t a reflected wave will take to propagate between a source and a receiver, the offset y between the source and the receiver, and the velocity c down to the layer is $t^2 = {t_0}^2 + \frac{y^2}{c^2}$, where t_0 is the traveltime at zero offset. The traveltime difference between two offsets is called moveout.) In the gathers in Figures 8a and 8d, we can observe several coherent retrieved events, highlighted in transparent red, that exhibit approximately hyperbolic moveout, so these could be potentially retrieved reflections. In the gather in Figure 8f, we cannot distinguish coherent approximately hyperbolic events. Nevertheless, we can observe a few inclined linear events, highlighted in transparent yellow, that might be retrieved body-wave arrivals (tails of reflection hyperbolas or refracted waves).

[42] To be able to interpret what the retrieved events are, we compare the retrieved virtual common-source gathers with common-source gathers recorded using active sources at the positions of the virtual sources. For these purposes, we have at our disposal two gathers from places where the zigzag active-source acquisition geometry turned at line L4—for a source at 1 km and for a source at 13 km, see Figures 8b and 8g, respectively. For these measurements, the active vibrator sources were placed 12.5 m away in the crossline direction from the positions of the virtual sources. It is possible to compare the active and the retrieved gathers as the wavelets of both records are zero-phase. The active wavelet is obtained from a correlation of the raw recorded data with the vibrator sweep; the ANSI wavelet comes from a sum of correlations. To facilitate the comparison, we apply an extra FK filter to the retrieved common-source gathers to suppress events with apparent velocities lower than 2700 m/s. In the active data, the direct body-wave

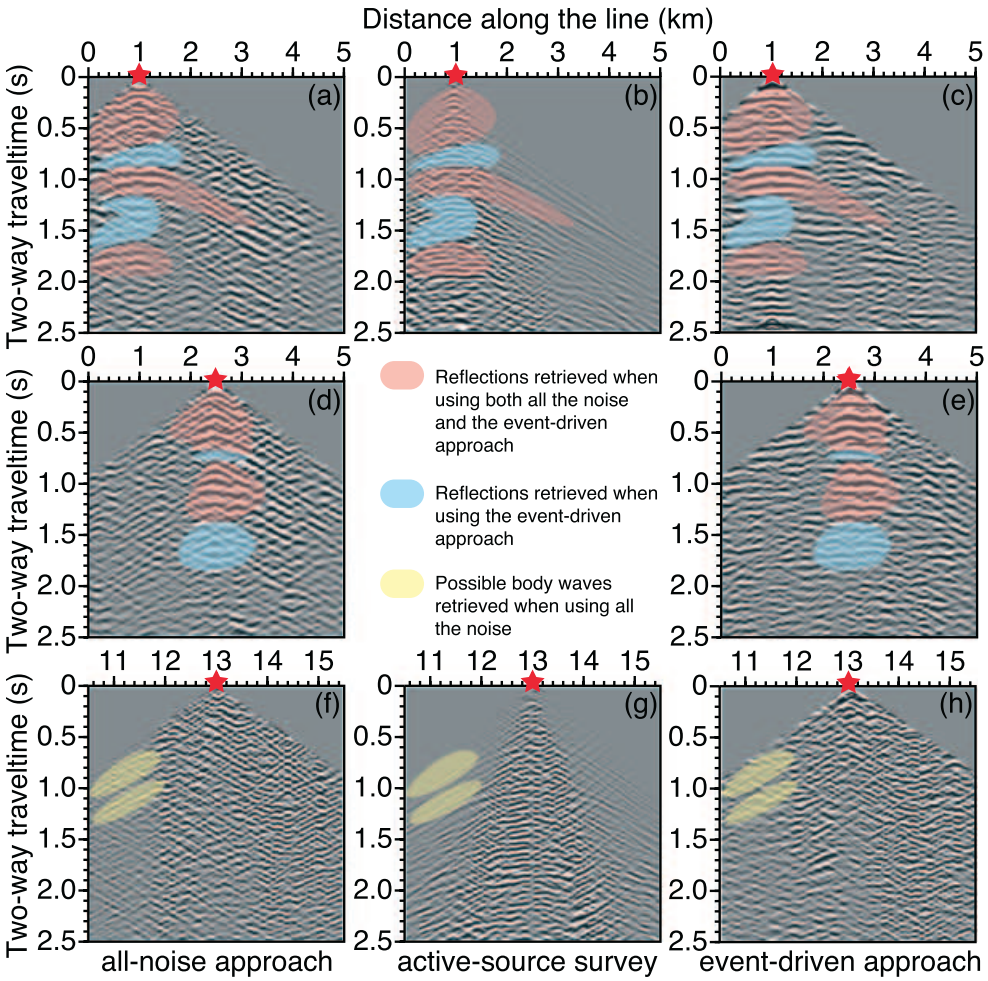


Figure 8. Common-source gathers along line L4: (a) for a virtual source at 1 km using all the noise; (b) for an active source at 1 km; (c) for a virtual source at 1 km using events only; (d) for a virtual source at 2.5 km using all the noise; (e) for a virtual source at 2.5 km using events only; (f) for a virtual source at 13 km using all the noise; (g) for an active source at 13 km; and (h) for a virtual source at 13 km using events only. The transparent colored areas highlight reflection arrivals.

arrival at near offsets has a velocity of 2000 m/s. By again taking a filter that rejects slightly higher velocities, we suppress most such arrivals. For comparison purposes, we apply the same frequency and FK filters to the two active-source gathers as well. Because the FK filter becomes very short at low frequencies, it will create in both the active and passive data linear artifacts that would appear to propagate in both directions. Inspection of the events highlighted in red in Figures 8a and 8b reveals that the two-way traveltimes of the retrieved approximately hyperbolic events coincide with those of the reflection hyperbolas in the active data, especially for short offsets. We therefore conclude that at short offsets, these retrieved events correspond to the apexes of the reflection hyperbolas in the active data. We also note that there are events in either data set that do not have similar two-way traveltimes.

[43] Even though the same filtering is applied to both data sets, we can see from the comparison that the frequency character and the hyperbolicity of the highlighted events is different. This can also be observed in Figures 9a and 9b, which show the FK spectra of the common-source gathers from Figures 8a and 8b, respectively—the main energy in the spectrum of the retrieved gather in Figure 9a is concentrated up to 15 Hz, while the active-data spectrum in Figure 9b is flat inside the complete frequency band. Taking a closer look at the present wave numbers in Figures 9a and 9b reveals that their spread in the retrieved gather is not as flat as in the active gather, with the smaller wave numbers being more energetic, which again shows that the retrieved events would have a different moveout at longer offsets than the reflection hyperbolas in the active data.

[44] Comparison of the result in Figures 8f with the active data in Figure 8g shows that even though the highlighted events in the retrieved and the active panel appear similar, they have different apparent propagation velocities—the retrieved events are faster. We cannot interpret these retrieved events unambiguously: they might be retrieved tails of reflection arrivals that are not visible on the active panel, because they overlap with the stronger-in-amplitude

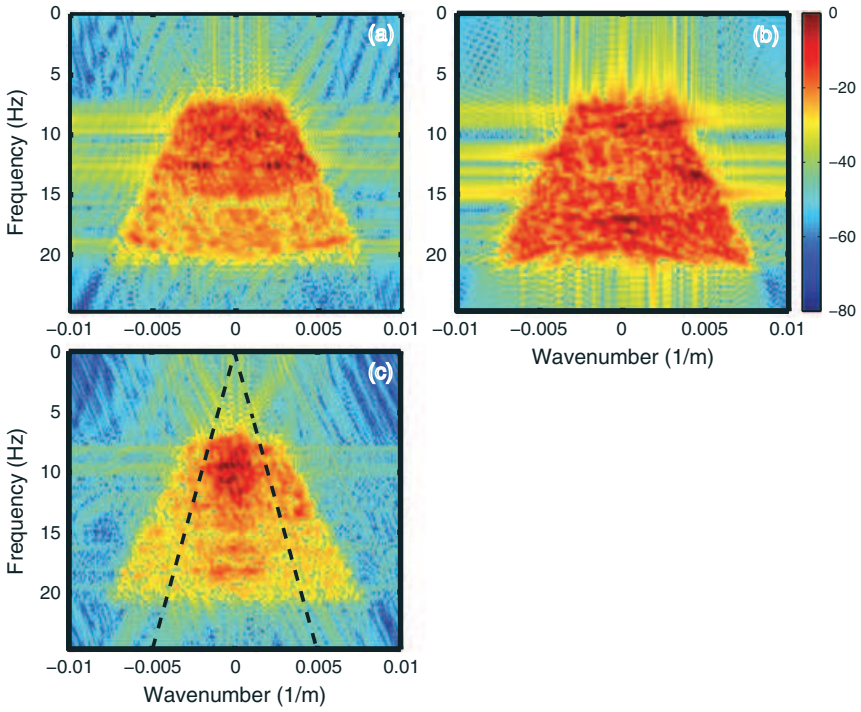


Figure 9. Frequency–wave number amplitude spectra of the common-source gathers from (a) Figure 8a, (b) Figure 8b, and (c) Figure 8c. The dashed lines correspond to the lowest velocity as used in the beamforming procedure for the identified events.

refracted waves; they might also be reflected or refracted arrivals retrieved with a wrong moveout velocity due to the stationary-phase regions not being captured.

[45] Even though we do not have at our disposal active measurements for a source at 2.5 km, we can use for qualitative comparison the active common-source gather at 1 km, because from the active data, we know that the subsurface is close to horizontally layered for such offsets. Comparing the retrieved events highlighted in red in Figures 8d with the reflection hyperbolas in Figure 8b, we can see that, for example, the events with apexes at 0.2, 0.4, and 1.0 s for a virtual source at 2.5 km correspond, for short offsets, to hyperbolic events in the active data at these times. Based on this consistency, we conclude that the red-highlighted events in Figure 8d are also actual reflection arrivals.

[46] Comparing the quality of the retrieved reflections in Figures 8a, 8d, and 8f to the active data, we see clear differences as well. The first thing that is seen is that the data retrieved from ANSI contains many events that do not correspond to physical arrivals in the active data. As mentioned in section 2, inhomogeneous illumination of the passive array from the noise sources would give rise to spurious events. As shown in Figure 6, the illumination of our array is indeed inhomogeneous: the ambient noise contains body-wave arrivals, but with limited and irregular slowness distribution.

[47] Another difference that we can observe is that the closer the virtual source position is to the traffic road (around kilometer 14), the less observable the retrieved reflections. We can see that in the proximity of the road, the lateral coherency of the apexes and the tails of the retrieved reflections disappears. As the spectrogram in Figure 3a shows, for stations close to the road surface wave energy is present up to 20 Hz. Because this energy is severely aliased, our filtering steps 1 and 2 (Figure 7) do not suppress it. This means that even though we did our best to suppress the remnants of the surface waves, they are still present and dominant at recording stations closer to the road. In the retrieval process, these remnants will overwhelm the (weaker) body-wave noise.

5.3. ANSI With Selected Noise Panels

[48] We now follow the second approach and apply ANSI only to the approximately 80 noise panels containing body-wave arrivals as identified using the beamforming. This has the advantage that the wavefield components found in these panels are expected to contribute to retrieval of reflections. Another significant advantage of this approach is that it minimizes the absolute amount of interference from the attenuated (but not fully removed) aliased surface waves and other arrivals with higher wave numbers. However, a potential drawback is that we will not use noise panels containing weak body-wave arrivals, which could still contribute to the reflections we aim to retrieve, but are not selected for correlation because they are dominated by surface waves.

[49] We process the approximately 80 selected event noise following the procedure from Figure 7 with the change that steps 5 is now performed after step 7. Knowing the illumination characteristics of the noise allows application of steps 6 and 7 to the individual correlated noise panels. The deconvolution step 6 can be applied by assuming that each event noise panel contains arrivals from only one source of body waves. Deconvolving per noise panel could potentially increase the overall resolution of the retrieved results, as there would be no need to assume that the power spectra of

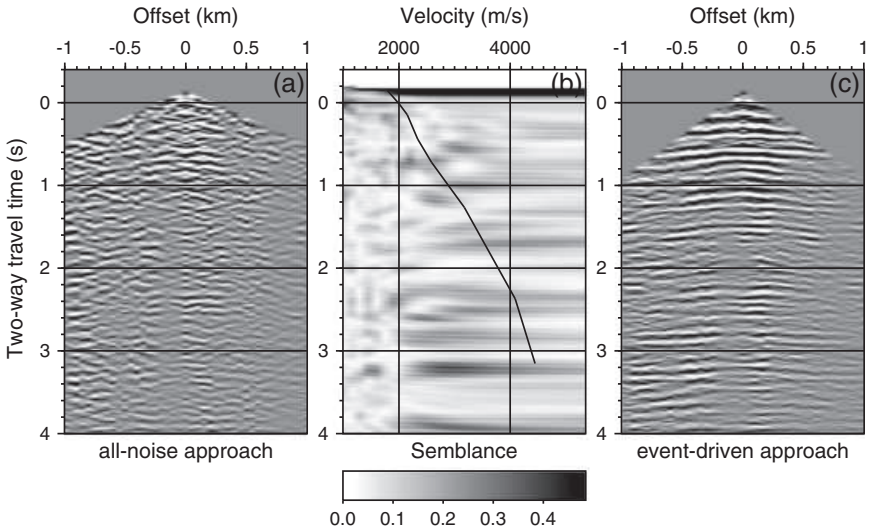


Figure 10. (a) Common midpoint (CMP) gather for a midpoint at 1 km obtained when all noise panels are used. (b) Stacking-velocity plot and picked stacking velocities (black line) calculated for the CMP gather in Figure 10a. (c) As in Figure 10a, but using the event-driven approach.

the noise sources in the different noise panels are the same. On the other hand, by applying the deconvolution for the autocorrelation of the source function after the summation over the correlated noise panels (when all noise panels are used), we do not have to assume that we have only one source per noise panel.

[50] In the consecutive step 7, by summing the causal and acausal parts of each correlated noise panel, we can take into account the specific illumination of the passive array for the specific noise-panel time period. As can be seen in Figure 6a, the identified events arrive at the recording lines with horizontal slownesses between 0.125 and 0.18 s/km; these events are recorded as approximately plane waves with small dips. We know from the active data that the subsurface is close to horizontally layered. In such a case, cross correlation of an incoming train of dipping plane waves would also result in a retrieved train of plane waves with the same dip in both causal and acausal times. To minimize the introduction of correlation artifacts during the implementation of step 7, instead of summing the complete retrieved causal and time-reversed acausal parts, we use only specific parts from both. For an apparent dip toward the smaller x values along the lines, we sum the causal part of the correlation result for traces that lie to the left of the master trace and the time-reversed acausal part of the correlation result for traces to the right of the master trace. For an apparent dip toward the bigger x values along the lines, we sum the opposite parts. For details of this technique, see *Ruigrok et al.* [2010].

[51] Figures 8c, 8e, and 8h show the common-source gathers for virtual source positions at 1, 2.5, and 13 km, respectively, retrieved using the event-driven approach. Comparing these results with the results retrieved using all the noise (see Figures 8a, 8d, and 8f, respectively) we see that in the red-highlighted areas, we have retrieved arrivals at the same two-way traveltimes, but flatter than the arrivals in the results from all the noise. This is explained with the fact that in the event-driven approach, we have used noise panels containing body-wave noise with very low horizontal slownesses. The retrieved arrivals from the event-driven approach are clearer, as the signal-to-noise ratio is higher and the events are more continuous. Comparing the blue-highlighted areas in the retrieved and the active panels, we can identify retrieved reflections in the event-driven results that are impossible to distinguish in the results from all the noise.

[52] From the comparison of the active and the retrieved data in Figure 8, we can also see that using the event-driven approach, which means limiting the amount of the recorded ambient noise, results in our case in limiting the illumination aperture of the noise sources: we could retrieve the apexes of the reflection arrivals and the short offsets (compare, for example, the arrival times at distance 0 km for the event with apex at 0.5 s in Figures 8a, 8b, and 8c). This can also be seen in the frequency–wave number spectra of the common-source gathers from Figures 8b and 8c, as shown in Figures 9b and 8c. We can observe that, as a result from the event-driven approach, the main energy is retrieved inside the triangle bound by velocity of 5000 m/s (the dashed black lines).

6. Extraction of Velocity and Structural Information From Retrieved Virtual Gathers

[53] After reflection information of the subsurface has been obtained from the application of ANSI, one could apply to the retrieved results processing algorithms from the seismic-exploration industry used for the extraction of subsurface velocity and structural information. In the previous section, we showed that we have retrieved reflection arrivals from ANSI at several places along Line L4. Obtaining a stacked section of the subsurface would also show whether reflections are retrieved at more places along the line. (We compare stacked sections along line L4, as along this line we have velocity and structural information available for comparison from the active data.) To extract velocity and structural information, we follow a standard processing scheme [*Claerbout*, 1985; *Yilmaz*, 1999] and compare the results with information extracted from the active

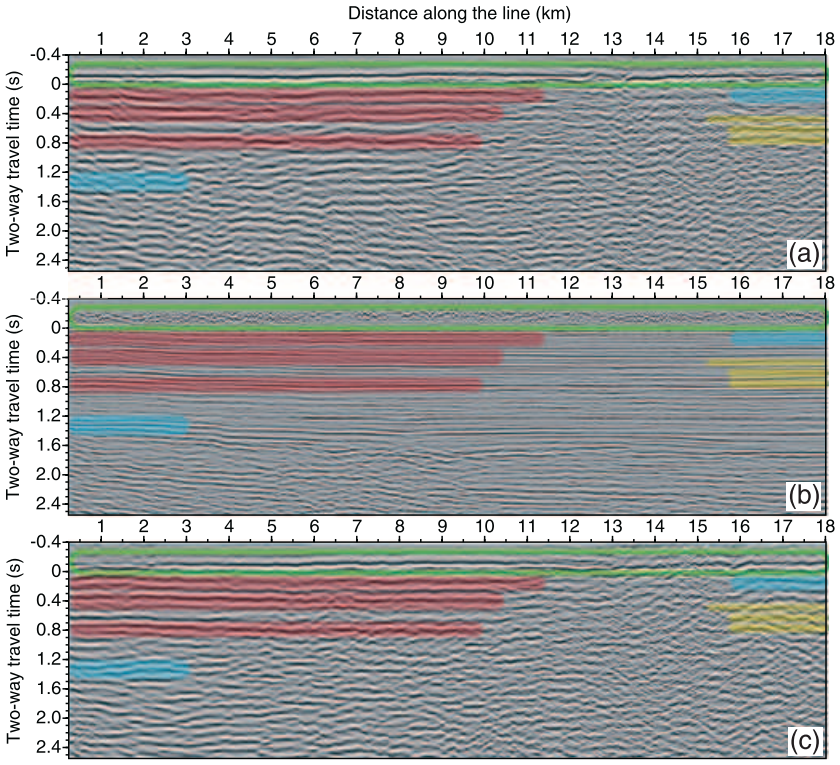


Figure 11. Time-migrated stacked section under L4 obtained (a) from ANSI applied to all the recorded noise, (b) from the active data, and (c) from event-driven ANSI using the final stacking velocities. The green curve encloses the Earth's surface. The transparent colored areas indicate zones with imaged reflectors that are present in the active section and the retrieved sections. When comparing with the active data, the red color indicates interpretable reflectors that are retrieved using both all the noise and the event-driven approach, the blue color indicates retrieved reflectors interpretable only in the section from the event-driven approach, and the yellow color indicates retrieved reflectors interpretable only in the section from all the noise.

data. The processing scheme consists of resorting to common midpoint (CMP) gathers, statics correction, interactive velocity analysis, normal moveout correction, stacking, and phase-shift time migration.

[54] We start the processing by resorting the retrieved common-source gathers to CMP gathers. In field measurements, the first seismic arrival to be detected at shorter offsets is the direct *P* wave arrival, while at longer offsets, refracted waves arrive first. This means that all events in the retrieved CMP gathers appearing earlier than these arrivals are not physical. For this reason, we mute all correlation artifacts earlier than the first arrival detected in the active data at the longest offset and whose apparent propagation velocity is 2700 m/s. Then, we apply so-called statics correction, which compensates for the effect of irregular topography, differences in the elevation of sources and geophones relative to a predefined datum, and low-velocity near-subsurface layers. The corrections applied here result in the retrieved CMP gathers being redatumed from the surface to a mean-sea level using a constant replacement velocity of 800 m/s for the near subsurface. Figure 10a shows the results from the described three steps for a CMP at 1 km obtained when using all-noise ANSI approach.

[55] Next, we use common-velocity stacks in an interactive fashion to pick stacking velocities along the retrieved reflections, i.e., along approximately hyperbolic arrivals. The picking is performed by fitting hyperbolic curves through the retrieved CMP gathers [*Taner and Koehler*, 1969]. In a CMP gather, the hyperbolic curves have their apexes at the zero-offset trace and have their curvature controlled by the stacking velocity. The analysis is performed by plotting the ratio of the square of summed amplitudes of the time samples fitted by calculated hyperbolic curves and of the summed squared amplitudes of the same time samples. Both the nominator and the denominator are smoothed by performing extra summation over time in a short time window centered at the time sample for which the velocity stacks are computed. In the resulting panel, we look for the maxima in the velocity range. Figure 10b shows the stacking-velocity plot for the CMP gather in Figure 10a. The picking should be done at the apexes of reflection events and thus is done in combination with the corresponding CMP gather, which shows where the reflection apexes are.

[56] From the beamforming analysis, we know that the recorded noise provides limited illumination angles of the passive array. This means that we should not expect the picked velocities from the retrieved data to be accurate, as the accuracy would depend a lot on our ability to reconstruct the asymptotes of the reflection-hyperbola, which in turn is defined by the illumination characteristics of the noise and our selected slownesses. The black curve in the stacking-

velocity plot in Figure 10b shows the final stacking velocities that we have picked. As we mentioned above, one is looking for the maxima in the velocity range in the resulting velocity plot. It can be seen that the lines do not strictly follow the maxima. There are two reasons for this. The first one is our desire to suppress free-surface multiples during the following steps of normal-moveout correction and stacking. For this purpose, the stacking velocities should generally increase with increasing two-wave traveltime, i.e, the black curve in Figure 10b should be a monotonically increasing function of time. The presence of free-surface multiples will cause local maxima in the stacking-velocity plots at velocities equal to already picked values at earlier two-way traveltimes, thus disturbing the monotonic character of the curve in the stacking-velocity plots.

[57] The second reason is the presence of correlation artifacts like the events visible along the time axis at 0.6 and 1.6 s for negative offsets around 1 km in Figure 10a. Such events do not form parts of hyperbolic arrivals. These artifacts, when stacked along a hyperbola in the CMP gathers, will cause ambiguity in the stacking-velocity plots by contributing to local maxima. When the signal-to-noise ratio of the retrieved reflections is low, like in the case of our all-noise retrieved results (Figure 10a), these artifacts will give a significant contribution to the stacking velocities and bias the velocity picking. In our case, velocities picked in this way underestimate the velocities estimated using the active data by up to 1000 m/s.

[58] In our case, the retrieved reflections from the event-driven approach are quite flat and thus cannot be used alone in velocity analysis. On the other hand, the signal-to-noise ratio at their apexes in the CMP gathers is much higher (see Figure 10c). Because of this, we use the event-driven CMP gathers together with the stacking velocity plots, obtained from the all-noise CMP gathers, to guide us where to pick the stacking velocities (the black line in Figure 10b). The so-picked final stacking velocities follow the trend of the active-data velocities, but generally underestimate them by up to 600 m/s. Compared to the underestimation when using only the all-noise CMP panels, we managed to reduce the maximum relative error from 41% to 21%. The fact that the velocities picked from the retrieved data differ from the active-data velocities shows again that the curvatures of the events in both data sets are different, as expected from the illumination analysis. We can conclude that the 11 h of recorded noise do not permit accurate velocity information extraction. Depending on the noise environment, longer noise recordings might result in more events being recorded, better illumination of the passive array and thus better retrieval of the hyperbolic curvature of the retrieved reflections.

[59] Having extracted stacking velocities from the retrieved data, we can proceed to obtain an image of the subsurface structures. We use the extracted final stacking velocities in a so-called normal-moveout correction, which flattens the reflection arrivals in the CMP gathers by correcting the reflections for their hyperbolic moveout. The traces in a flattened CMP gather are then stacked together, and the obtained stacked trace is assigned to the position of the midpoint of this gather. In ideal circumstances, without normal-moveout stretch and/or velocity misestimates, the result of the stacking process is a zero-offset trace that would have been recorded with coinciding source and receiver. The stacking process improves the signal-to-noise ratio of the reflections in the stacked trace, compared to the signal-to-noise ratio of the individual traces, by $\sqrt{M}$, where M is the number of stacked traces. This is true only for uncorrelated noise. In our case, the stacking process may not be that efficient as the traces to be stacked are themselves the result of a correlation process and much of the noise that did not contribute to the retrieved reflections is not uncorrelated (such as the remaining aliased surface waves).

[60] We use the final picked velocities to stack CMP gathers from both the all-noise and the event-driven results. As the retrieved reflections in the event-driven CMP gathers with very high apparent velocities, normal-moveout correction with the picked velocities would have the effect of curving the arrivals toward earlier time (overcorrection). During the stacking process, this would lead to a reduced number of stacked individual traces contributing to a reflection, but also to an increased noise level between the reflections. To reduce the influence of the latter, we limit the stacking offsets in the event-driven CMP gathers to 2000 m (in the all-noise CMP gathers, the stacking offsets are 8000 m). In general, using inaccurate stacking velocities would result in lower signal-to-noise ratio of the stacked trace.

[61] After the stacking process, we apply a phase-shift time migration, which maps possibly inclined subsurface reflectors to their real position in time [*Yilmaz*, 1999]. We perform the migration using velocities from the active-source data to be able to obtain more accurate relocation of the reflectors. Figures 11a, 11b, and 11c show the results from the migration along line L4 when using CMP gathers obtained from the all-noise ANSI, from the active data, and from the event-driven ANSI, respectively. The green curves enclose the Earth's surface. Comparing the retrieved image in Figure 11a with the image in Figure 11b, we can see that the retrieved events at 0.2, 0.3, 0.4, and 0.9 s, highlighted with transparent red and yellow, are imaged quite consistently over some distance along line L4. This shows that the retrieved arrivals in Figure 8, which we interpreted as retrieved reflections for virtual sources at 1, 2.5, and 13 km, are retrieved also at other virtual source locations along the line. As these imaged events are obtained from retrieved reflections and because they coincide very well in traveltime sense with imaged reflectors in the active data, we conclude that they are imaged reflectors. Note that the red-highlighted reflectors in Figure 11a are imaged with lower frequency and appear to represent several thin layers in the active-data image. Using incorrect stacking velocities results in smearing of the reflections during the stacking process. This might also have contributed to the obtained lower frequencies of the retrieved images. On the other hand, the yellow-highlighted reflectors in Figure 11a, visible in the part of the section with x values higher than the x value of the blacktop road, appear to contain higher frequencies, and coincide with separate reflectors in the image from the active data in Figure 11b. Such retrieved events containing higher frequencies can be observed only relatively close to the place where the road intersects the line, and even though they are also present in the part of the section with x values lower than the x value of the road, there they do not form laterally

continuous features. Having higher-frequencies reflections interpretable only close to the road might be an indication that body-wave noise with higher frequencies is generated by the traffic on the road.

[62] Comparing the retrieved image in Figure 11c with the active-data image in Figure 11b, we see that the event-driven approach has also imaged subsurface reflectors. The red-highlighted reflectors are imaged also in this result, but compared to the ones in Figure 11a, they are laterally more continuous, as expected from the active data (Figure 11b). Comparing the yellow-highlighted zones in the three images, in accordance with the observations from the comparison of the common-source gathers in Figures 8f–8h, we see that the event-driven approach has done a poorer job in these zones. Comparing the active section in Figure 11b with the event-driven section, we can see that in the latter, we can easily interpret the blue-highlighted events as retrieved reflectors as well. In Figure 11c, we can interpret part of a deeper reflector at around 1.3 s, which is not possible in Figure 11a. The event-driven approach also adds to the retrieved subsurface information a piece of the shallowest reflector in the direction of increasing x values from the blacktop road and thus improves this reflector's continuity. Interpretation of the blue-highlighted events is not possible in the all-noise image, as at these times, the events are laterally incoherent or too weak.

7. Discussion

[63] We show that reflection characteristics of the subsurface can be obtained using ANSI. Our results exhibit lower frequencies when compared to the active data. The frequency content of the retrieved data depends on the nature of the recorded body-wave noise (source mechanisms and related signal and radiation characteristics), but also on the locations of the noise sources. The active sources were tuned to induce energy in the ground that has a flat spectrum between 6 and 80 Hz. As the active sources are situated relatively close to the receivers, the recorded reflected energy would also have a comparable frequency spectrum. On the other hand, the reflections retrieved with ANSI would inherit the frequency characteristics of the body-wave events that are present in the recorded ambient noise. The body-wave noise might be coming from sources situated at large distances from the recording lines, and due to the intrinsic absorption in the subsurface, the higher frequencies in the recorded signals would be attenuated. In addition, the noise sources, like local earthquakes for example, might emit only lower frequencies. As for the moment we cannot conclude what noise sources caused the body-wave noise characterized by the beamforming, we cannot say if also higher-frequency reflections could be obtained from ANSI from noise other than the blacktop-road noise. Even if only reflections containing lower frequencies are possible to retrieve, they could still be used to our advantage by combining them with the active data to obtain data with a broader spectrum. It should be investigated how this could be done.

[64] The behavior and strength of seismic noise depends on several factors, such as the depth of the receivers, atmospheric and weather conditions, geographical location, local geology, tectonic activity, and proximity to anthropogenic noise sources. This makes it difficult to predict the outcome of an experiment with limited duration. Nevertheless, our results show that ANSI has a potential for frontier exploration of large areas after establishing the confidence level in the data using evaluation of the illumination characteristics of the noise. Longer-term continuous passive monitoring experiments would be a way to further understand and develop ANSI as a tool for subsurface characterization for exploration purposes. Depending on the factors described above, the results from continuous data, acquired during an extended period of time, could be improved compared to the results of the current study, because longer monitoring could increase the chance of accumulating noise sources with a greater diversity of ray parameters, implying more comprehensive illumination of the array.

[65] In exploration seismology, velocity information is traditionally estimated from the typical hyperbolic character of reflections. The continuity, size, and curvature of the reflector is mainly determined by the effective illumination of the incident noise wavefield. For reliable velocity information, it is therefore important to have a comprehensive illumination by noise sources. When this is not the case, we should gain sufficient knowledge of the noise-source characteristics (the effective illumination) for interpretation of the reliability of the velocity information. For seismic exploration with ANSI, these noise characteristics will vary with geographical location, atmospheric conditions, geology, tectonic setting, and proximity to anthropogenic noise sources. The velocity analysis is sensitive to the tails of the hyperbolas, which in our case are not well retrieved. Nevertheless, we show how common-source gathers retrieved using the all-noise approach and retrieved using the event-driven approach could be used in a complementary fashion to apply standard exploration-seismic processing for obtaining structural information of the subsurface. In practice, one can first apply the all-noise ANSI aiming to extract preliminary information of the subsurface. This information can then be used to decide to search for the body-wave parts of the noise, which can in turn be used in the described event-driven approach. If we assume that the noise is incident as plane-wave events, then relation (1) can be re-parametrized as a slowness integral or sum over slownesses [see *Ruigrok et al.*, 2010, 2011]. By already having some idea of the subsurface image from the all-noise approach and by characterizing the noise events in terms of their slowness, for example, using beamforming, one could select those events that lie in the stationary-phase regions of the integral in relation (1). Furthermore, if the distribution of noise sources is limited or too sparse, the correlated events in the integral would not interfere destructively properly thereby producing artifacts. By using only noise panels whose illumination characteristics fall inside the stationary-phase regions of the integral, such artifacts would be suppressed, and consequently, one would synthesize a kinematically correct estimate of the partial Green's function at a faster convergence rate and with a higher signal-to-noise ratio than when all available noise records are used. A similar process of limiting the slowness range and stacking only over the Fresnel zone has been used in seismic migration to reduce imaging artifacts when data are sparsely sampled and/or have limited aperture [e.g., *Buske et al.*, 2009].

[66] Extraction of subsurface structural information from ANSI could be used as a monitoring tool of subsurface

processes, for example, during exploitation of petroleum fields or CO_2 sequestration. In cases like the two example, velocity information from active seismic surveys would be available and could be used for the migration purposes to eliminate smearing of the imaged reflectors due to incorrectly picked velocities. The active seismic surveys are expensive and thus performed infrequently. ANSI could be used to monitor changes in the subsurface between the active surveys on a more frequent basis.

8. Conclusions

[67] We showed how velocity and structural information of the subsurface at the exploration scale could be extracted from ambient-noise seismic interferometry. For this, we first investigated some properties of ambient noise recorded for 11 h at one night during an active-source 3-D exploration survey in the eastern part of the Sirte basin in Libya. We showed that the bulk of the energy in the noise is composed of surface waves at frequencies below 6 Hz. These surface waves are generated by cars along a blacktop road that intersects the survey area. We visually inspected the noise and determined that it contains nearly vertically traveling events. We further characterized these events using a frequency-domain beamforming algorithm. The beamforming output for the most of these noise parts showed a dominant ray-parameter approximately between 0.125 and 0.18 s/km. Considering this ray-parameter distribution, these events probably correspond to crustal *P* wave phases. In addition, these body waves had a fairly rich azimuthal distribution. So far, it remains unclear what is the origin of the body-wave phases seen in the noise panels. Recording body-wave noise is a prerequisite for retrieving body-wave reflections.

[68] We then applied seismic interferometry to the recorded noise using two approaches. In the first one, we correlated and summed all the recorded noise. Using the fact that retrieval of reflections would result from correlation of body waves noise, in the second approach, we correlated and summed only the parts of the noise that contained body-wave arrivals (events) as identified by the beamforming. Both approaches result in obtaining common-source gathers. Comparison of the retrieved results to common-source gathers from active exploration sources showed that we retrieved several shallow reflection events at 0.2, 0.3, 0.4, and 0.9 s that also appear in the active data. The retrieved results from the event-driven (second) approach exhibited a much better signal-to-noise ratio when compared to the results from the all-noise (first) approach. On the other hand, the event-driven approach limited the illumination angles of the array from the noise sources, thus resulting in retrieved reflection arrivals with faster move out than the retrieved arrivals from the all-noise approach. To retrieve reflection arrivals, we needed to suppress the present surface waves, which were drowning out the useful body-wave noise. The suppression was achieved in the field using groups of receivers, followed by frequency filtering before correlation. We then applied frequency–wave number filtering before correlation to suppress other events.

[69] After retrieval of reflection gathers, we proceeded to apply an interactive velocity analysis to extract stacking velocities. The applied velocity analysis on the retrieved data did not result in accurate extracted velocity information when compared to the velocities from the active data. This might be due to insufficient temporal length of the recorded noise. Nevertheless, we showed that analysis of the reflection arrivals obtained using the event-driven approach can be useful during the velocity analysis of reflections retrieved using all the noise, in this way helping to extract velocities of the subsurface closer to the active-data velocities. As a final step, we obtained an image of the subsurface reflectors after stacking the retrieved gathers using the picked stacking velocities. Comparison with a stacked image from the active data along one of the lines showed that the retrieved shallow reflection events are consistent along more than half of the line.

[70] **Acknowledgments.** We thank SEPLG (Erik Kleiss, Rian de Jong, and Alan Smith) and the Libyan National Oil Corporation for releasing the data and for permission to publish the results. We thank Wim Mulder (Shell) for his review of and constructive comments on an earlier manuscript and Elmer Ruigrok (TUDelft), Kees Hornman, and Boris Kuvshinov (Shell) for fruitful discussions and suggestions for improvement of the manuscript. This work was carried out while D.D. was supported by the Technology Foundation STW, applied science division of NWO (project VENI.08115).

References

Aki, K., and P. G. Richards (1980), *Quantitative Seismology: Theory and Methods,* vols. I, II, W.H. Freeman and Co, San Francisco, Calif.

Asten, M. W., and J. D. Henstridge (1984), Array estimators and the use of microseisms for reconnaissance of sedimentary basins, *Geophysics*, *49* (11), 1828 –1837.

Baskir, E., and C. Weller (1975), Sourceless reflection seismic exploration, *Geophysics*, *40*, 158.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007), Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.*, *169*, 1239–1260.

Bnefoy-Claudet, S., F. Cotton, P.-Y. Bard, and The nature of noise wave-field and its application for site effects studies: A literature review (2006), *Earth Sci. Rev.*, *79*, 205-227, doi:10.1016/j.earscirev.2006.07.004.

Buske, S., S. Guthjar, and C. Sick (2009), Fresnel volume migration of single-component seismic data, *Geophysics*, *74*(6), WCA47–WCA55.

Capon, J. (1969), High-resolution frequency-wavenumber spectrum analysis, *Proceed. IEEE*, *57*(8), 1408–1418.

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

Claerbout, J. F. (1985), *Imaging the Earth's Interior*, Blackwell Science, Oxford, U. K.

Draganov, D. (2007), Seismic and electromagnetic interferometry: Retrieval of the Earth's reflection response using crosscorrelation, PhD thesis, Delft University of Technology.

Draganov, D., K. Wapenaar, W. Mulder, J. Singer, and A. Verdel (2007), Retrieval of reflections from seismic background-noise measurements, *Geophys. Res. Lett.*, *34*, L04305, doi:10.1029/2006GL028735.

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

Froment, B., M. Campillo, P. Roux, P. Gouédard, A. Verdel, and R. Weaver (2010), Estimation of the effect of nonisotropically distributed energy on the apparent arrival time in correlations, *Geophysics*, *75*, SA85–SA93, doi:10.1190/1.3483102.

Gerstoft, P., P. M. Shearer, N. Harmon, and J. Zhang (2008), *Geophys. Res. Lett.*, *35*, L23306, doi:10.1029/2008GL036111.

Harmon, N., P. Gerstoft, C. A. Rychert, G. A. Abers, M. Salas de la Cruz, and K. M. Fisher (2008), Phase velocities from seismic noise using beamforming and crosscorrelation in Costa Rica and Nicaragua, *Geophys. Res. Lett.*, *35*, L19303, doi:10.1029/2008GL035387.

Kawase, H., F. J. Sánchez-Sesma, and S. Matsushima (2011), The optimal use of horizontal-to-vertical spectral ratios of earthquake motions for velocity inversions based on diffuse-field theory for plane waves, *Bull. Seismol. Soc. Am.*, *101*, 2001–2012, doi:10.1785/0120100263.

Kay, S. M. (1969), *Modern Spectral Estimation,* Prentice Hall, Englewood Cliffs, N. J.

Koper, K. D., K. Seats, and H. Benz (2010), On the composition of the Earth's short-period seismic noise field, *Bull. Seismol. Soc. of Am.*, *100*, 606–617, doi:10.1785/0120090120.

Kulhánek, O. (2002), The structure and interpretation of seismograms, in *International Handbook of Earthquake and Engineering Seismology*, edited by W. H. K. Lee et al., Part A, pp. 333–348, Academic Press, Amsterdam, The Netherlands.

Lacoss, R. T., E. J. Kelly, and M. N. Toksöz (1969), Estimation of seismic noise structure using arrays, *Geophysics*, *34*(1), 21–38.

Lin, F. C., M. P. Moschetti, and M. H. Ritzwoller (2008), Surface-wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.*, *173*, 281–298, doi:10.1111/j1365-246X.2008.03720.x.

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

Okada, H. (2003), The microtremor survey method, Society of Exploration Geophysicists, *Geophys. Monogr. Ser.*, doi:10.1190/1.9781560801740.

Peterson, J., (1993), Observation and modeling of seismic background noise, *USGS Open-File Report 93–322.*

Prieto, G. A., J. F. Lawrence, and G. C. Beroza (2009), Anelastic Earth structure from the coherency of the ambient seismic field, *J. Geophys. Res.*, *114*(B07303), doi:10.1029/2008JB006067.

Rickett, J., and J. F. Claerbout (1996), Passive seismic imaging applied to synthetic data, *Stanford Exploration Project Report 92, 87–94.*

Roux, P., K. G. Sabra, W. A. Kuperman, and A. Roux (2005), Ambient noise cross correlation in free space: Theoretical approach, *J. Acoust. Soc. Am.*, *117*(1), 79–84.

Roux, P., K. G. Sabra, P. Gerstoft, W. A. Kuperman, and M. C. Fehler (2005), P-waves from cross-correlation of seismic noise, *Geophys. Res. Lett.*, *32*, L19303, doi:10.1029/2005GL023803.

Rost, S., and C. Thomas (2002), Array seismology: Methods and applications, *Rev. Geophys.*, *40*(3), 2-1–2-27, doi:10.1029/2000RG0001002002.

Ruigrok, E., X. Campman, D. Draganov, and K. Wapenaar (2010), High-resolution lithospheric imaging with seismic interferometry, *Geophys. J. Int.*, *183*, 339–357, doi:10.1111/j.1365-246X.2010.04724.x.

Ruigrok, E., X. Campman, and K. Wapenaar (2011), Extraction of P-wave reflections from microseisms, *C. R. Geosci.*, *343*, 512–525.

Sabra, K. G., P. Gerstoft, P. Roux, and W. A. Kuperman (2005), Surface wave tomography from microseisms in Southern California, *Geophys. Res. Lett.*, *32*, L14311, doi:10.1029/2005GL023155.

Scherbaum, F. (1987), Seismic imaging of the site response using microearthquake recordings. Part I. Method, *Bull. Seismol. Soc. Am.*, *77*, 1905–1923.

Shapiro, N. M., and M. Campillo (2004), Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, *31*, L07614, doi:10.1029/2004GL019491.

Snieder, R., K. Wapenaar, and Larner K. (2006), Spurious multiples in seismic interferometry of primaries, *Geophysics*, *71*, SI111–SI124.

Suleiman, A. S., and D. I. Doser (1995), The seismicity, seismotectonics and earthquake hazards of Libya, with a detailed analysis of the 1935 April 19, M=7.1 earthquake sequence, *Geophys. J. Int.*, *120*, 312–332.

Taner, M. T., and F. Koehler (1969), Velocity spectra—Digital computer derivation applications of velocity functions, *Geophysics*, *34*, 859–881, doi:10.1190/1.1440058.

Toksöz, M. N., and R. T. Lacoss (1968), Microseism: Mode-structure and sources, *Science*, *159*(6), 872–873.

van Manen, D.-J., A. Curtis, and J. O. A. Robertsson (2006), Interferometric modeling of wave propagation in inhomogeneous elastic media using time reversal and reciprocity, *Geophysics*, *71*, SI47–SI60.

Vermeer, G. J. O. (1990), *Seismic Wavefield Sampling*, Society of Exploration Geophysicists, Tulsa, Okla.

Wapenaar, K. (2004), Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross-correlation, *Phys. Rev. Lett.*, *93*, 254301.

Wapenaar, K., and J. T. Fokkema (2006), Green's function representations for seismic interferometry, *Geophysics*, *71*, SI33–SI46.

Wilson, D., J. Leon, R. Aster, J. Ni, J. Schlue, S. Grand, S. Semken, S. Baldridge, and W. Gao (2002), Broadband seismic background noise at temporary seismic stations observed on a regional scale in the southwestern United States, *Bull. Seismol. Soc. Am.*, *92*(8), 3335–3341.

Xu, Z., C. Juhlin, O. Gudmunsson, F. Zhang, C. Yang, A. Kashubin, and S. L uth (2012), Reconstruction of subsurface structure from ambient seismic noise: An example from Ketzin, Germany, *Geophys. J. Int.*, *189*, 1085–1102.

Yang, Y., M. H. Ritzwoller, A. L. Levshin, and N. M. Shapiro (2007), Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.*, *168*, 259–274.

Yao, H. J., R. D. van der Hilst, and M. V. de Hoop (2006), Surface-wave array tomography in SE Tibet from ambient-seismic noise and two-station analysis, *Geophys. J. Int.*, *166*, 732–744.

Yilmaz, O. (1999), *Seismic Data Processing* 9th ed., Society of Exploration Geophysicists, Tulsa, Okla.

Zhan, Z., N. Sidao, D. V. Helmberger, and R. W. Clayton (2010), Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise, *Geophys. J. Int.*, *182*, 408–420, doi:10.1111/j.1365-246X.2010.04625.x.

Zhang, J., P. Gerstoft, and P. Shearer (2009), High-frequency P-wave seismic noise driven by ocean winds, *Geophys. Res. Lett.*, *36*, L09302, doi:10.1029/2009GL037761l.