Noise characterization and ambient noise cross-correlations at Long Beach

Jason P. Chang, Sjoerd de Ridder, and Biondo Biondi

ABSTRACT

The dense seismic array in Long Beach, California is located in an urban environment along the Pacific Ocean. There are a variety of noise sources influencing the ambient seismic noise field, both natural and anthropogenic in origin. To understand the temporal and spatial influences of these sources, we calculate power spectral densities (PSD) and apply beamforming to ambient seismic noise data. From spatial distribution maps of noise PSD, we find that energy from the Pacific Ocean dominates the noise field at frequencies below 2 Hz, while energy from local roads and Interstate 405 dominate the noise field at frequencies above 2 Hz. From spectrograms, we observe diurnal fluctuations in energy that are in accord with expected patterns in human activity. From beamforming, we find that directed, low-frequency energy from the Pacific Ocean is prevalent throughout the array, while the directivity of high-frequency energy varies throughout the array. Near Interstate 405, noise energy is clearly directed outwards from the freeway, but at a distance from the freeway, noise energy arrives more evenly across azimuths. Based on these observations, we expect the noise source distribution to be generally more homogeneous at higher frequencies than at lower frequencies. Ambient noise cross-correlation results at frequencies spanning 0.5-2 Hz and 2-4 Hz reinforce the validity of this expectation. A more exciting observation is the emergence of a P-wave at the higher-frequency range in our virtual super-source gather. This is a promising first step toward potentially retrieving other body-wave arrivals.

INTRODUCTION

The Long Beach oil field is a producing oil field located beneath the city of Long Beach. Due to its urban location, traditional techniques for collecting data for seismic imaging and velocity analysis are disruptive and difficult to perform. One alternative is to use passive seismic interferometry, or ambient noise cross-correlation, for this type of subsurface analysis. By cross-correlating ambient seismic noise recorded at two simultaneously recording receivers, the Green's function between the two receiver locations can be estimated (Bensen et al., 2007; Wapenaar et al., 2010). Because surface waves are the strongest events in ambient noise, the surface wave part of the Green's function has been widely used for tomographic imaging at both the regional

(Shapiro et al., 2005; Sabra et al., 2005) and continental scales (Yang et al., 2008; Bensen et al., 2008). The body wave part of the Green's function is beginning to gain attention, with Roux et al. (2005) retrieving direct P-waves at the regional scale, Draganov et al. (2007, 2009) retrieving P-wave reflections at the exploration scale, and Zhan et al. (2010) and Poli et al. (2012) retrieving P- and S-wave reflections at the continental scale.

In January 2012, Nodal Seismic, LLC, deployed Phase II of its dense seismic array in Long Beach that is well-suited for testing the effectiveness of using ambient seismic noise for resolving subsurface structure at the exploration scale (see Figure 1). The array consists of 2400 vertical-component geophones with an average spacing of 100 m. The geophones are continuously recording (24 hrs/day) for over three months. Using data from Phase I, Lin et al. (2012) have shown that Green's functions can be successfully estimated at the neighboring Long Beach array using seismic interferometry.

Under ideal conditions, noise sources would be evenly distributed throughout Long Beach, thereby allowing seismic interferometry to better estimate Green's functions (Snieder, 2004). Due to the presence of obvious sources of seismic noise, this is unlikely to be the case at this site. Our goal is thus to characterize the ambient seismic noise field at Long Beach. We want to identify the various noise sources and their influence on the ambient seismic field in both time and space. To achieve this, we first compute power spectral densities (PSD) of recorded ambient seismic noise. We plot spatial distribution maps to locate potential sources of noise, and then we generate spectrograms for those energetic areas to observe how they vary over time. Next, we apply beamforming to the ambient seismic noise data to determine the slowness and azimuthal distribution of noise. Finally, we relate our beamforming results to estimated Green's functions from seismic interferometry at frequencies between 0.5-2 Hz and 2-4 Hz.

NOISE CHARACTERIZATION

PSD Spatial Distribution Maps

Generating and examining noise PSD, or power spectrums, is the standard method for quantifying seismic background noise (McNamara and Boaz, 2005). PSD reveal how power in a signal is distributed over a range of frequencies. Before calculating the PSD at each receiver, we need to prepare the seismic recordings. We first break our time series into 5-minute patches with 50% overlap. Next we subtract the mean from each time segment to remove any systematic offset or zero-frequency component. We then apply a smooth Hanning taper to each time segment to reduce the artifacts caused by Gibbs phenomenon when performing a Fourier transform on a truncated time series. To compute the PSD for a time segment, we take the one-dimensional, discrete Fourier transform and then compute the square of the complex amplitude.

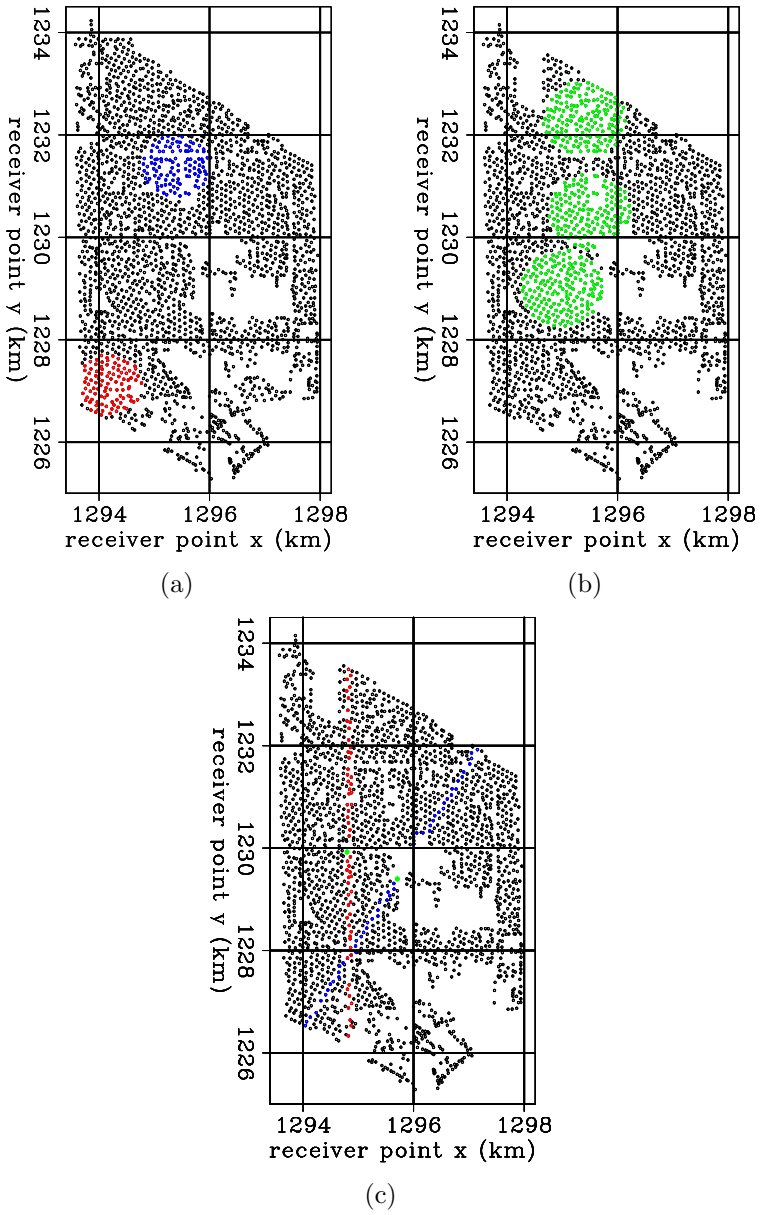
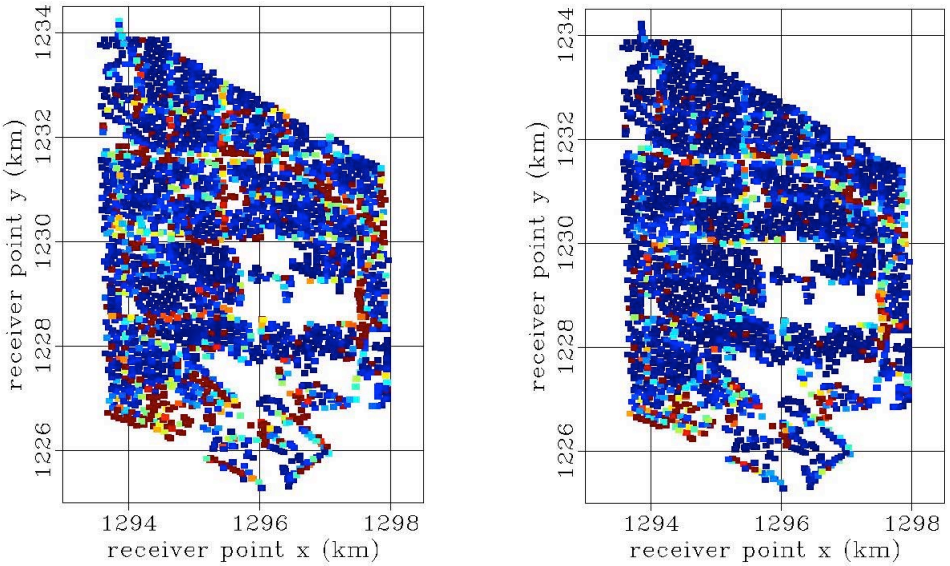


Figure 1: Maps of the dense Long Beach array overlain with station used for (a) spectrograms, (b) beamforming analysis, and (c) common virtual source gathers. In (c), solid green circles mark the locations of the virtual sources. Coordinates are NAD27, CA State Plane, Zone 7, kilometers. [**ER**]

To identify potential sources of noise at Long Beach, we generate spatial distribution maps of the noise PSD by plotting PSD at their receiver locations. We average the PSD derived from twelve consecutive overlapping 5-minute time segments to obtain a single PSD for a 32.5-minute time window. By averaging over multiple overlapping time segments, we decrease the variance in our PSD estimates and minimize the influence of spurious events. We show results from two different time windows to get a sense of how the spatial noise distribution differs during peak human activity (5pm) and during calmer periods of human activity (midnight).

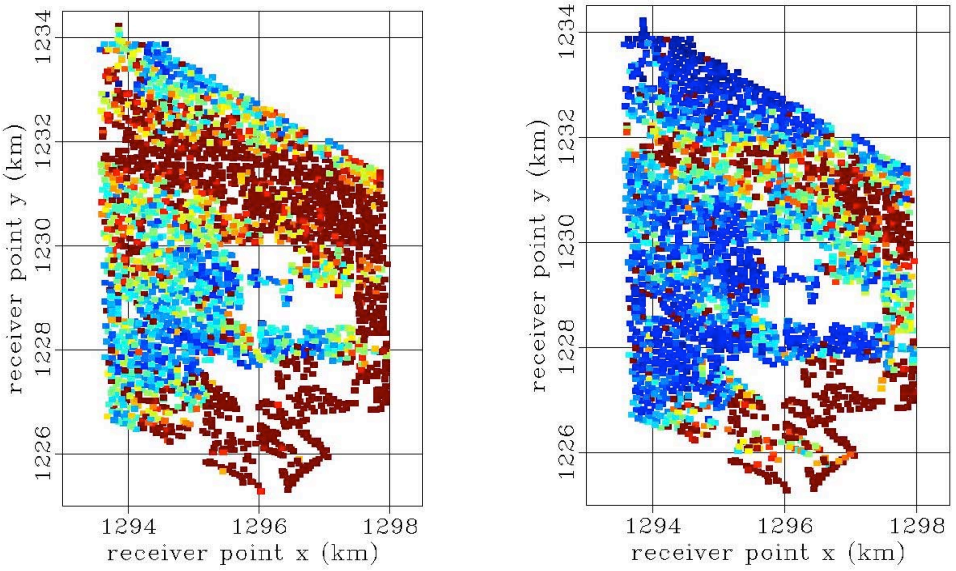
Figures 2, 3, and 4 display the spatial distribution maps of noise PSD for frequencies of 15 Hz, 4 Hz, and 1 Hz, respectively. Figures 2-4(a) are averaged PSD results from a 32.5-minute time window at 5pm, and Figures 2-4(b) are averaged PSD results from a 32.5-minute time window at midnight. Warmer colors (red) indicate higher energy at the given frequency at a particular station. For a given pair of maps at the same frequency, the amplitude scale is the same so that a side by side comparison of relative energy can be made. The amplitude scale of figures from frequency to frequency is not the same, however. We received the data with a 3 Hz low-cut filter, so PSD amplitudes at 1 Hz are much lower than at 15 Hz.



(a) (b)

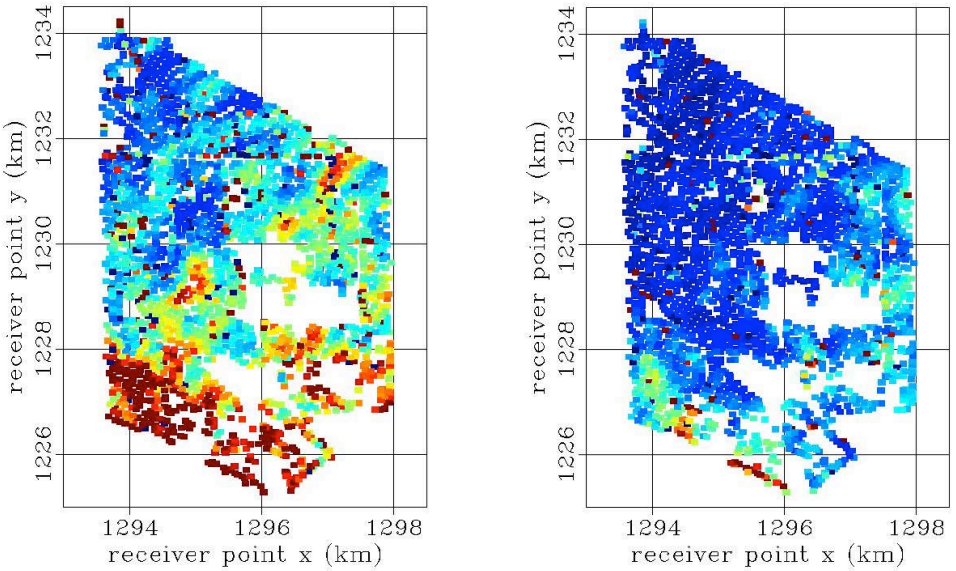
Figure 2: Spatial distribution maps of noise PSD at 15 Hz. (a) at 5pm. (b) at midnight. Warmer color (red) points indicate stations that are recording higher energy at this frequency. The Pacific Ocean is immediately south of the array, and Interstate 405 runs east-west through the north of the array.. [**CR**]

The PSD spatial distribution maps reveal regions where high amounts of seismic energy are being recorded and potentially produced. From Figure 2, we see high amounts of energy at 15 Hz localized at receivers along Interstate 405 (which runs



(a) (b)

Figure 3: Spatial distribution maps of noise PSD at 4 Hz. (a) at 5pm. (b) at midnight. Warmer color (red) points represent stations that are recording higher energy at this frequency. . [**CR**]



(a) (b)

Figure 4: Spatial distribution maps of noise PSD at 1 Hz. (a) at 5pm. (b) at midnight. Warmer color (red) points represent stations that are recording higher energy at this frequency. . [**CR**]

roughly east-west through the northern part of the array) and local roads. This suggests that the seismic noise field is dominated by traffic noise at higher frequencies. We observe a similar trend in Figure 3, where the dominant energy source at 4 Hz is clearly Interstate 405. The difference in spatial extent of freeway energy at 4 Hz and 15 Hz can be attributed to attenuation. The freeway energy at 15 Hz is localized because higher frequencies attenuate more rapidly with distance, whereas the freeway energy at 4 Hz is more spread out because lower frequencies attenuate more gradually with distance.

At 4 Hz frequency, we also see the appearance of high-energy noise in the southeast region of the array. The border of this region is fairly consistent over the two time windows, which suggests that we could be looking at a geological boundary rather than a source of noise. According to the California Department of Conservation (Greenwoood, 1998), parts of this region, and specifically margins of Alamitos Bay, are built on artificial fill that consists of soft and silty sand. Soft material can amplify seismic amplitudes (Delgado et al., 2000), so we could be seeing artificial fill or some other type of geology amplifying seismic signal at around 4 Hz in this region. A tomography study would address this conjecture, and it is the next objective of our research. At 1 Hz (Figure 4), we see energy primarily in the southern part of the array along the coastline, which suggests that energy from the Pacific Ocean dominates at these low frequencies. In general, there is more seismic energy at 5pm than at midnight for all frequencies.

Spectrograms

To better resolve how the noise field at Long Beach changes over time, we generate spectrograms (plots of frequency versus time) for receivers near the coastline (red circle in Figure 1(a)) and receivers overlapping Interstate 405 (blue circle in Figure 1(a)). We first discard recordings in those regions that have maximum PSD amplitudes above a certain threshold, since they will bias the spatial average. We then average the PSD from the remaining receivers in those regions for each 5-minute time segment and plot those PSD side by side to obtain a map showing how frequency content changes with time.

Figure 5 displays spectrograms spanning over two days in February for frequencies up to 4 Hz. Figure 5(a) represents the freeway region and Figure 5(b) represents the ocean region. The color scale represents pseudo-amplitudes after taking the logarithm of the PSD amplitudes. A better comparison of logarithmic PSD amplitudes can be found in Figure 6(a) and Figure 6(b), where we compute the average PSD over one-hour time windows at noon and midnight of the 17th, respectively, for the freeway (blue curve) and ocean (red curve) regions.

From the noise PSD spatial distribution maps, it is clear that the two primary sources of energy at Long Beach are the Pacific Ocean and Interstate 405. We also observe that energy in the ambient noise field is generally weaker at midnight than at

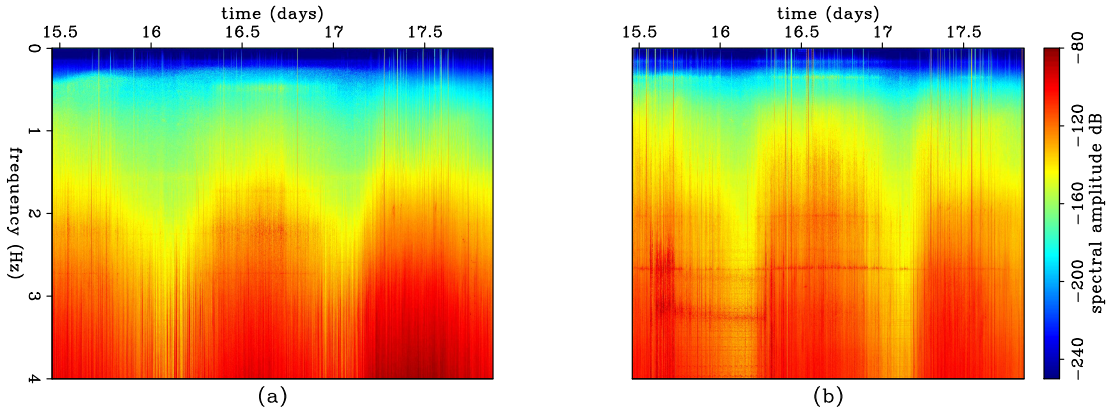


Figure 5: Spectrogram spanning over two days in February for frequencies up to 4 Hz. (a) freeway. (b) ocean. Color scale corresponds to the logarithm of the PSD amplitudes, with warmer colors indicating higher energy and cooler colors indicating lower energy. [**CR**]

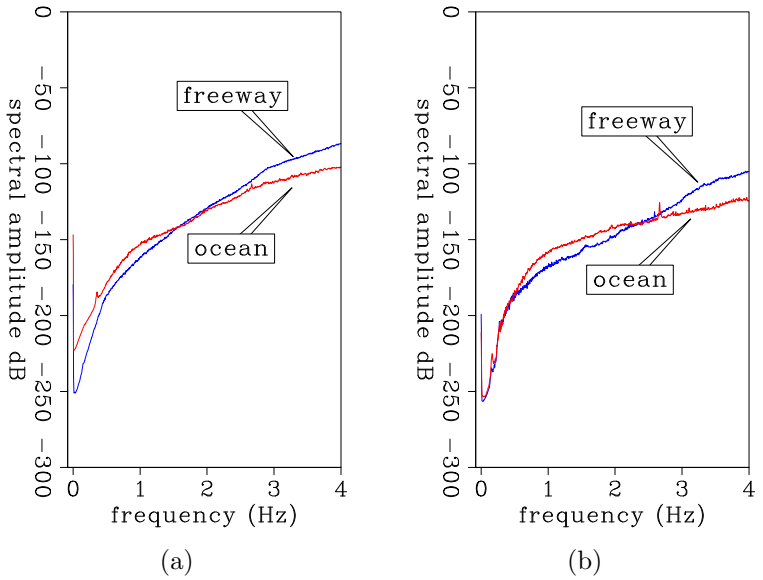


Figure 6: Average PSD over one-hour time windows at (a) noon, and at (b) midnight for the freeway (blue) and ocean (red) regions. [**ER**]

5pm, which suggests that anthropogenic noise sources are highly influential at Long Beach. Spectrograms in Figure 5 and the averaged PSD in Figure 6 confirm both these observations. The averaged PSD show that energy is always greatest near the freeway at frequencies above 2 Hz, hence identifying Interstate 405 as a major source of seismic noise, and greatest at the ocean receivers at frequencies below 2 Hz, hence identifying the Pacific Ocean as another major source of seismic noise. Diurnal fluctuations are clearly seen in the spectrograms, with peaks corresponding to daytime periods when human activity is most active and valleys corresponding to early morning periods when human activity is more subdued. Along with the differences in the averaged PSD at noon and midnight, it is apparent that human activity is primarily responsible for the diurnal variations seen in the noise field.

Beamforming

Having identified two major sources of noise at Long Beach (the Pacific Ocean and traffic noise), we now want to determine the extent of their influence over the ambient noise field. We do this by applying beamforming to the ambient seismic noise data, which provides the slowness and azimuth at which seismic energy arrives (Rost and Thomas, 2002). We first select two hours of ambient seismic noise recorded at groups of stations throughout the array. A long time window is chosen because our ambient noise cross-correlations are averaged over long periods of time. We whiten the spectrum of each recording by dividing the complex spectrum by a smoothed version of its amplitude spectrum. This step maintains the crucial phase information while removing the bias in the recordings toward higher frequencies (due to the 3 Hz low-cut filter) and mitigating the influence of anomalous recordings on the beamforming output. We then apply a series of overlapping 0.5 Hz bandwidth filters centered at intervals 0.25 Hz apart, beginning at 0.5 Hz and ending at 3.75 Hz. By applying beamforming to each narrow frequency range, we can better resolve how slowness and azimuthal distribution of noise vary with frequency. Next we transform the data to the τ-p domain. This can be thought of as a series of slant stacks along different azimuths. If there is any coherent energy traveling across an array at a certain slowness and azimuth, it will stack constructively and appear as a strong event in the τ-p domain. We then take the absolute value of the transformed data and average over τ. This produces a map depicting the slowness and azimuthal distribution of noise for those set of stations, representative of the time window.

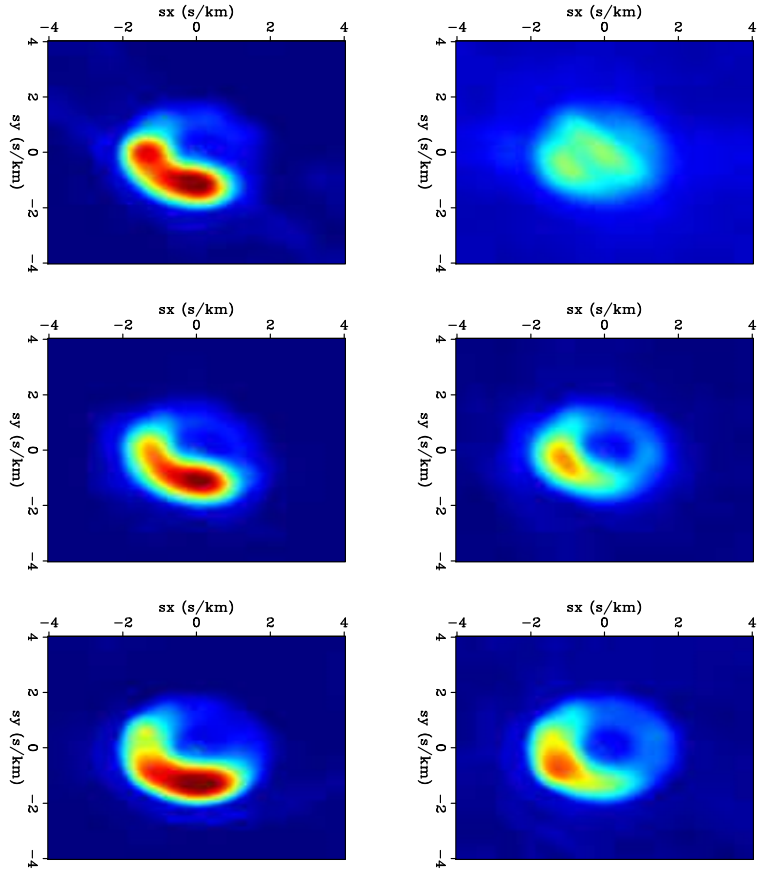
Figure 7 and Figure 8 display beamforming results for frequency bands 0.5-1.0 Hz and 3.25-3.75 Hz, respectively. For both figures, columns from left to right represent results from January 18 at 2 am, 10 am, and 6 pm, while rows from top to bottom represent results from clusters of stations located just north of Interstate 405, just south of Interstate 405, and near the center of the array (for station clusters see Figure 1(b)). The azimuth of a high-beampower anomaly represents the azimuth from which noise is coming (eg. an anomaly on the right represents noise coming from the east) while the radial distance of the anomaly represents the slowness at

which that noise is traveling across the array. The amplitude of the beampower is scaled the same for panels in a given row. Warmer colors represent higher beampower.

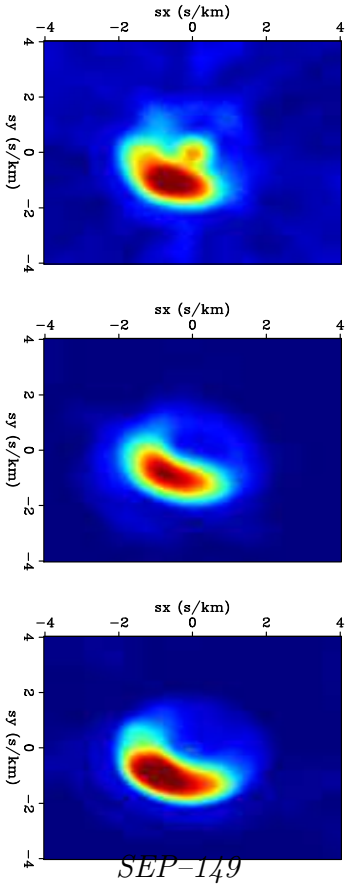
At the frequency band 0.5-1 Hz, it is clear from Figure 7 that noise is arriving from the southwest at all locations and times. This is yet another indication .0that the Pacific Ocean is the dominant source of energy at low frequencies, and that its influence extends throughout the entire array. For this particular frequency band, noise from the ocean arrives generally at a slowness of .001 s/m, or a velocity of 1 km/s. Based on the frequency and velocity, this is likely Rayleigh wave energy.

At the frequency band 3.25-3.75 Hz, there is more variation in the ambient seismic noise field from region to region, and in some cases variation from time to time. Looking at different regions of the array (different rows), we see clear differences in the azimuths at which energy is distributed. Stations just north of the freeway (top row) reveal that noise is primarily approaching from the south, while stations just south of the freeway (middle row) reveal that noise is primarily approaching from the north. These observations indicate that Interstate 405 is the dominant source of noise energy for regions near the freeway. Strong arrivals at zero slowness during the day indicate a signal that arrives at those stations at the same time and could be caused by noise from airplanes flying over the stations after taking off or before landing at the nearby Long Beach airport. Airplane noise might also explain the generally high beampower in these panels. Noise energy is better equipartitioned near the center of the array (bottom row), which suggests that noise from nearby local roads could be the dominant source of energy away from the freeway. In all these regions, there appears to be noise consistently arriving at approximately 333 m/s and 500 m/s. Because strong energy exists at similar azimuths for both arrivals, we are likely observing the slower fundamental Rayleigh wave mode and a faster higher-order Rayleigh wave mode rather than arrivals from two different sources. At these frequencies, nearly all panels exhibit crossing anomalies; these are spatial aliasing artifacts. The two surface wave modes are affected by it, but it appears that any potential higher-velocity body wave will not have crossing events and hence will not be aliased. Looking at different times (different columns), we see that the relative effect of the freeway on nearby stations is consistent over time. At stations away from the freeway, noise is slightly better equipartitioned at 10 am and 6 pm than at 2 am, suggesting that the noise source distribution is more even at peak periods of human activity.

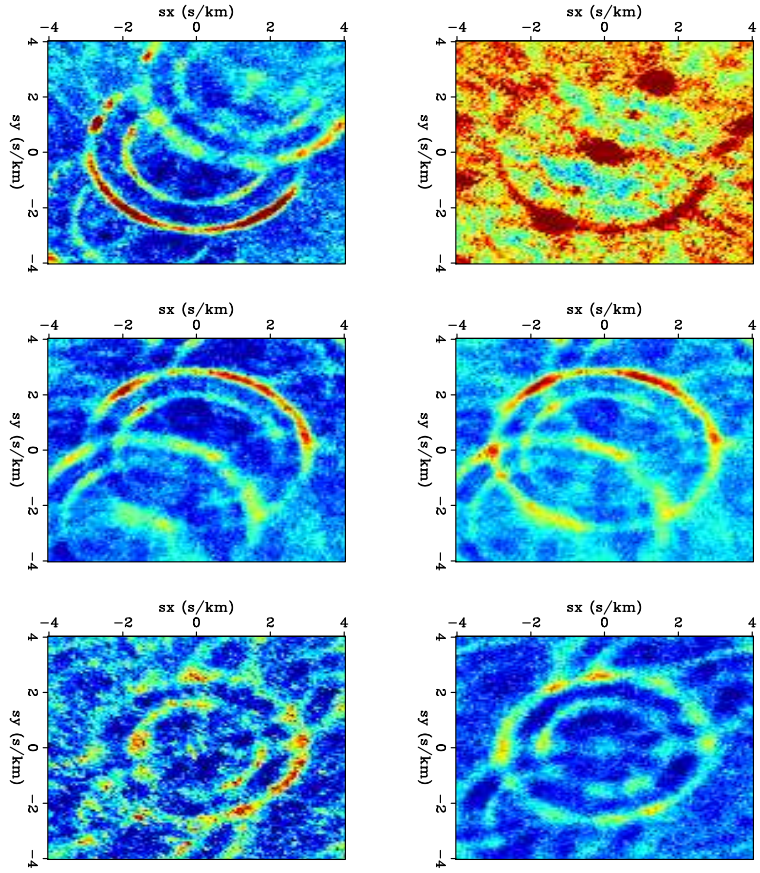
From looking at PSD spatial distribution maps, we determined that the major sources of noise energy at Long Beach were the Pacific Ocean at frequencies below 2 Hz, and traffic noise, particularly Interstate 405, at frequencies above 2 Hz. Spectrograms revealed that the ambient seismic noise field goes through diurnal fluctuations in energy that can be attributed to diurnal patterns in human activity. Beamforming analysis showed that noise is consistently arriving from the southwest at low frequencies, confirming that the Pacific Ocean is a dominant noise source. This analysis also showed that high-frequency noise energy near the freeway is dominated by Interstate 405. However, further away from the freeway, the noise arrived at a wider range of azimuths and could potentially be attributed to energy generated from traffic on local



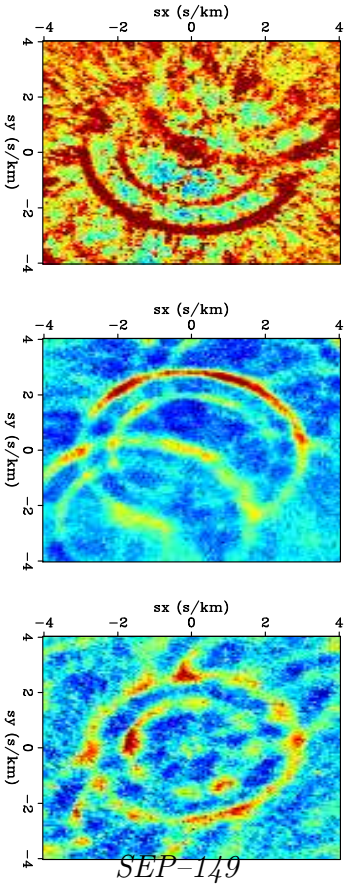
(a) (b)



(c)



(a) (b)



(c)

roads.

AMBIENT NOISE CROSS-CORRELATIONS

Method

Having characterized the noise field at Long Beach, we now perform ambient noise cross-correlations to estimate the Green's function at low frequencies (0.5-2 Hz) where noise from the Pacific Ocean is dominant, and at high frequencies (2-4 Hz) where traffic noise is dominant. This estimate is typically made by cross-correlating recordings of ambient seismic noise from two simultaneously recording receivers over a long period of time. This process recovers the Green's function and its time-reversed version between the two receivers, convolved with the autocorrelation of a source function such as noise (Wapenaar et al., 2010). Here, we implement a processing method adapted from Bensen et al. (2007). We first break up our time series into tapered, non-overlapping two-hour time windows and then bandpass accordingly. We then whiten the input traces prior to cross-correlating. In the frequency domain, the cross-correlation of pre-whitened traces can be expressed as

$$[G(x_B, x_A, \omega) + G^*(x_B, x_A, \omega)] = \left\langle \left(\frac{U(x_B, \omega)}{\{|U(x_B, \omega)|\}} \right) \left(\frac{U^*(x_A, \omega)}{\{|U(x_A, \omega)|\}} \right) \right\rangle \quad , \qquad (1)$$

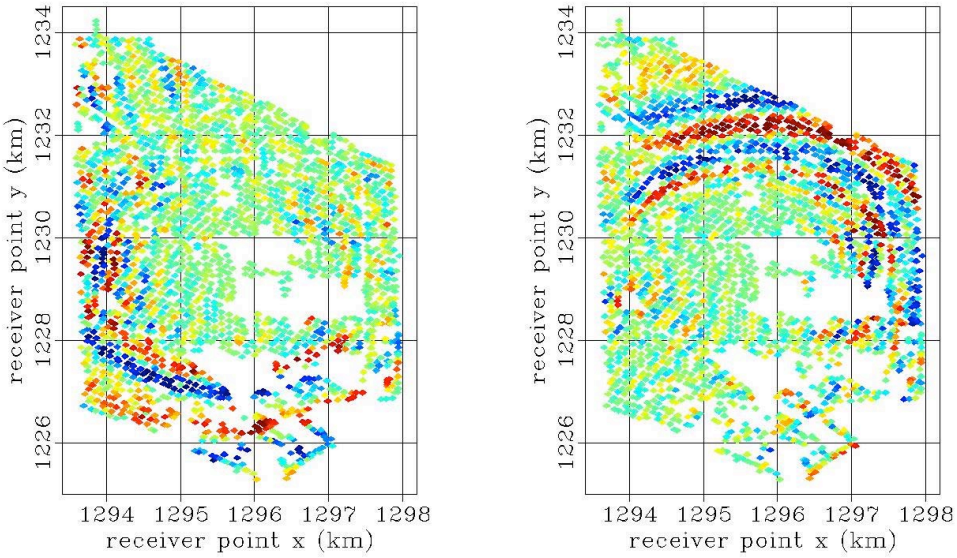
where G is the Green's function between two receiver locations (x_A, x_B), $U(x, \omega)$ is the spectrum of the wavefield at a given receiver location x, $*$ represents the complex conjugate, $\langle \cdot \rangle$ represents the time-averaged ensemble, $|\cdot|$ represents the real absolute value of the spectrum, and $\{\cdot\}$ represents a 0.003 Hz running window average used for whitening the signal.

By dividing the standard cross-correlation procedure by smoothed amplitude spectrums at the two receivers, we are deconvolving an estimate of the noise source signal from the correlations in the time domain to better estimate the Green's function (see Appendix). Pre-whitening limits the amplitude variation from trace to trace while maintaining the phase information and increasing the bandwidth of the estimated Green's function (Bensen et al., 2007). This step is particularly important at the frequencies we are looking at (under 4 Hz) because it removes the effect of the 3 Hz low-cut filter applied to the data. Although true amplitude information is lost, it is not a major problem at Long Beach because amplitudes from trace to trace can vary significantly and because the data is very noisy.

After whitening the input traces and cross-correlating each station pair, we normalize each result by the maximum amplitude. By doing this, we are further limiting the effect of spurious events and instrument spikes in our correlations. For the lower frequency range, we average correlations from 100 two-hour time windows, resulting in an average over 8 days. For the higher frequency range, we average correlations from 240 two-hour time windows, resulting in an average over nearly 3 weeks.

Results

We first show correlations in map view for a virtual source near the center of the array. Figure 9 and Figure 10 display snap shots for the lower frequency range and the higher frequency range, respectively, at -5 s and 5 s time lags. For frequencies below 2 Hz, the estimated Green's function shows significant directivity, as it is strongest in the southwest portion of the array at acausal time lags (when energy is moving toward the receiver) and strongest in the northeast portion of the array at causal time lags (when energy is moving away from the receiver). This suggests that the Pacific Ocean, which is south of the array, is a significant source of directed, low-frequency energy. For frequencies above 2 Hz, the estimated Green's function is more symmetric, suggesting that the noise source is more homogeneously distributed at these frequencies.



(a) (b)

Figure 9: Snapshots from a virtual source in the center of the array at (a) -5 s lag and (b) 5 s lag for frequencies between 0.5 Hz and 2 Hz. Note the directivity of the Green's function. [**CR**]

These correlation results are in line with our observations from the PSD and beamforming analysis. Noise characterization revealed that energy at low frequencies is dominantly generated by noise along the coastline in the southern part of the array, thereby explaining the inhomogeneous noise source distribution that is leading to strong asymmetry in the estimated Green's function. As we look at increasingly higher frequencies starting from 2 Hz, the noise source distribution becomes more homogeneous as the noise field becomes more dominated by anthropogenic sources (particularly traffic noise), thus leading to more symmetric Green's functions. The influence of Interstate 405 can be seen in the acausal snapshot, where the wavefront

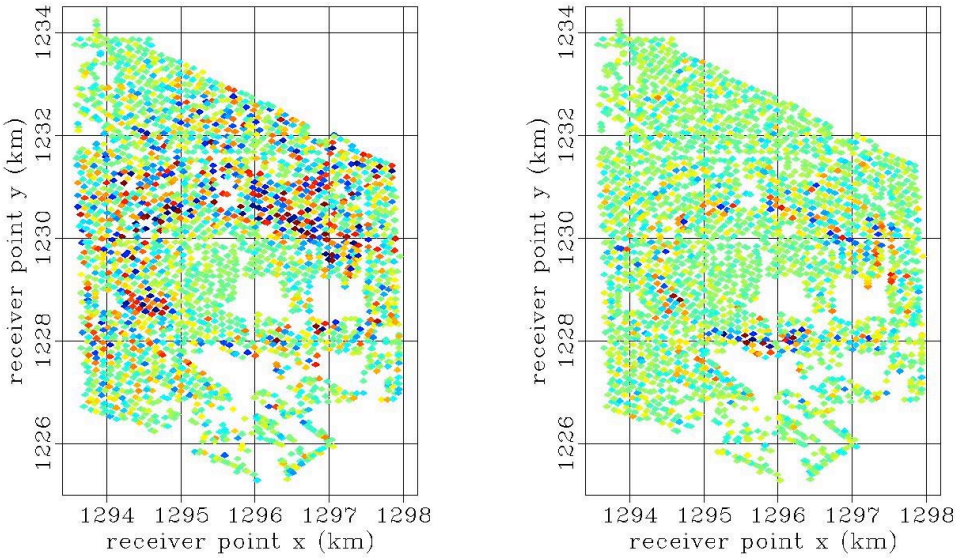


Figure 10: Snapshots from a virtual source in the center of the array at (a) -5 s lag and (b) 5 s lag for frequencies between 2 Hz and 4 Hz. Note the improved symmetry of the Green's function. [**CR**]

in the north (which is moving south from the freeway toward the virtual source) is relatively strong compared to the rest of the wavefront. Furthermore, we can see the dispersive nature of Rayleigh waves in vertically inhomogeneous media from these snapshots. At the time lags shown here, the moveout from the virtual source is greater at lower frequencies than at higher frequencies because lower frequencies sense deeper and are thus expected to travel faster.

To better measure these velocities, we display virtual common source gathers for lines of receivers shown in Figure 1(c). The lower-frequency source gather in Figure 11(a) is generated along a seismic line roughly perpendicular to the coast, while the higher-frequency source gather in Figure 11(b) is generated along a seismic line running north-south and perpendicular through a section of Interstate 405. Sources in stationary phase locations, such as locations at the ends of a seismic line, have the most influence over the estimates of Green's functions (Snieder, 2004). Therefore, the seismic lines are chosen so that the dominant noise sources (the Pacific Ocean and Interstate 405) are located in these stationary phase locations. The lower-frequency gather has the same virtual source as in the snapshots, while the higher-frequency gather has a virtual source in a different location in order to limit the gaps in the north-south seismic line. Negative offsets are stations to the south or southwest of the virtual source, and positive offsets are stations to the north or northeast of the virtual source. Acausal time lags represent energy moving toward the virtual source, while causal time lags represent energy moving away from the virtual source.

For the lower-frequency source gather, we see propagation of the fundamental Rayleigh wave from southwest to northeast at a velocity of approximately 630 m/s. The directivity of the wave is apparent from the asymmetry about zero-time lag. For the higher-frequency source gather, we see that the dominant Rayleigh wave is propagating from north to south. This wave energy begins approximately 1.5 km north of the virtual source, which is the location of Interstate 405. The dominant Rayleigh wave energy is generated by the freeway, and this is further evidenced by the correlating energy traveling northward at offsets greater than 1.5 km. At receivers south of the virtual source, the energy from the freeway has attenuated, allowing weaker Rayleigh wave energy traveling north from the Pacific Ocean toward the virtual source to become visible. Both these fundamental Rayleigh waves are traveling at approximately 250 m/s, which is slower than the fundamental Rayleigh wave velocity at lower frequencies, as we expect. Focusing further on stations south of the virtual source, we see hints of a higher-order Rayleigh wave mode traveling away from the virtual source at a velocity of approximately 500 m/s. The directionality of this wave mode suggests that it is being generated by Interstate 405.

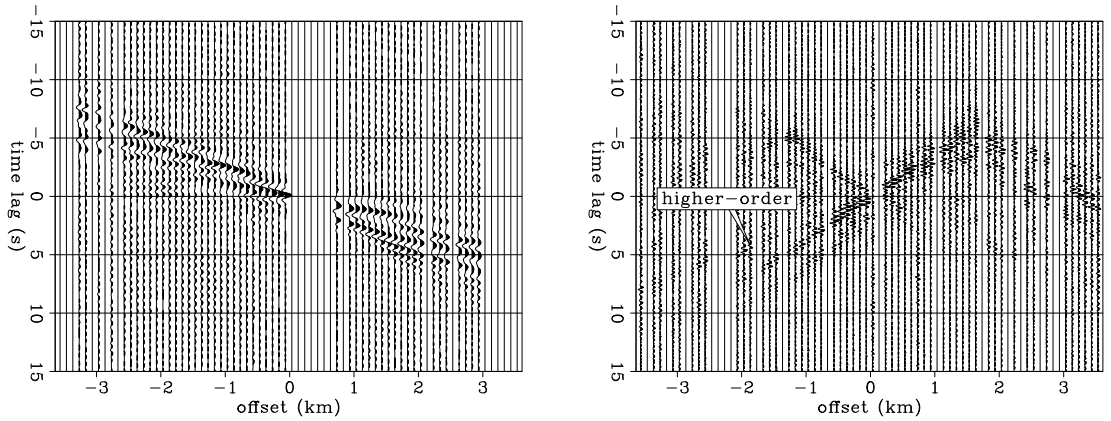
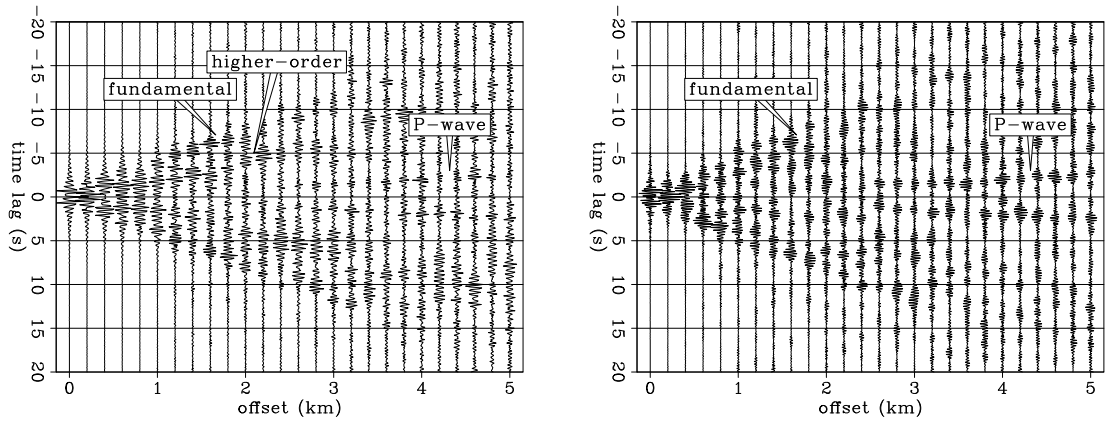


Figure 11: (a) Common source gather for frequencies between 0.5 Hz and 2 Hz. (b) Common source gather for frequencies between 2 Hz and 4 Hz. [**CR**]

To enhance the the higher-order Rayleigh wave mode, we construct a super source-gather, as seen in Figure 12(a). This is done by creating virtual source gathers from 100 randomly distributed source locations for each of the 240 two-hour time windows, sorting each resulting correlation by radial offset into 200 m bins, and then stacking. This amounted to stacking over 50 million correlations. The fundamental and higher-order Rayleigh wave modes are more distinct now, with velocities of approximately 250 m/s and 500 m/s, respectively. More importantly, we see the emergence of a refracted P-wave traveling at approximately 2 km/s, as also observed in Lin et al. (2012). We display the same super source-gather for a frequency range of 3-4 Hz to bring out this arrival in Figure 12(b). One may speculate that the P-wave energy is generated from converted surface wave energy due to heterogeneities in Earth's crust, but since these P-waves are more distinct at frequencies higher than 3 Hz, there

is a chance they could be generated by anthropogenic sources such as traffic noise. For this P-wave arrival to be useful, though, we must extract it from correlations between a single station pair. This is one of our research focuses for the coming year. Regardless, the presence of a refracted P-wave from ambient noise cross-correlations at Long Beach is a first step toward potentially retrieving other body-wave arrivals.



(a) (b)

Figure 12: (a) super source-gather for frequencies between 2 Hz and 4 Hz created from stacking over 50 million correlations. Note the emergence of a P-wave traveling at approximately 2 km/s. (b) Same super source-gather as in (a), but for frequencies between 3 Hz and 4 Hz. Note that the P-wave arrival is slightly clearer now. [**CR**]

CONCLUSIONS AND FUTURE WORK

The distribution of noise sources plays a significant role in the quality of the Green's functions retrieved from ambient noise cross-correlations. Thus, the primary goal of our paper is to characterize the noise at Long Beach. By looking at PSD spatial distribution maps, we found that the Pacific Ocean dominates the noise field at frequencies below 2 Hz, and that human activity dominates the noise field at frequencies above 2 Hz. By looking at spectrograms, we found that noise energy goes through daily fluctuations for frequencies above 2 Hz. These fluctuations fall in line with periods of human activity. Beamforming analysis revealed that the influence of the Pacific Ocean noise is consistent throughout the array, while the influence of Interstate 405 has a limited extent. Away from the freeway, noise arrives at a wider range of azimuths at higher frequencies. From these results, we expect the noise field to be more homogeneous at higher frequencies than at lower frequencies. Ambient noise cross-correlation results from 0.5-2 Hz and 2-4 Hz confirm these observations, with estimated Green's functions showing strong asymmetry at lower frequencies and better symmetry at higher frequencies. We also created virtual source gathers and found the fundamental Rayleigh wave to be dispersive by having velocities of approximately 630 m/s at the lower-frequency range and approximately 250 m/s at the higher-frequency range. We also observed a weak high-order Rayleigh wave mode traveling at 500 m/s at the

higher-frequency range. Lastly, we created a super source-gather, which revealed a refracted P-wave traveling at 2 km/s.

Our future work will focus on exploiting the station density and time duration of the Long Beach dataset for surface wave tomographic studies, with an eye toward time-lapse monitoring. The parameters of the array allow for the retrieval of high-frequency (¡ 2 Hz) Green's functions from cross-correlations, which will provide better subsurface resolution at shallow depths. This level of resolution is particularly useful for earthquake hazard analysis, which is of importance due to the presence of the Newport-Inglewood fault running through the region. Furthermore, we want to improve the refracted P-wave arrival. We obtained that arrival by stacking over many virtual sources, so we would like to retrieve the P-wave for a single virtual source. A potential strategy is to employ a more selective stacking procedure. Also, because the P-wave arrival was enhanced at frequencies above 3 Hz, we want to determine if it was generated by the strongest anthropogenic source in Long Beach: Interstate 405. If we can show that the freeway generated those refracted P-waves, then we can use it as a passive, yet very active, seismic source.

ACKNOWLEDGMENTS

We gratefully acknowledge Signal Hill Petroleum, Inc. for access to this dataset and permission to publish. Thanks to Dan Hollis for his enthusiasm and cooperation throughout this research. Thanks to Stew Levin and Bob Clapp for their help converting the data.

APPENDIX

Estimates of Green's functions from ambient seismic noise are typically obtained by cross-correlating traces from two simultaneously recording receivers. This process recovers the Green's function and its time-reversed version between the two receivers, convolved with the autocorrelation of a source function such as noise (Wapenaar et al., 2010). In the frequency domain, this can be expressed as

$$[G(x_B, x_A, \omega) + G^*(x_B, x_A, \omega)]|S_N(\omega)|^2 = \langle U(x_B, \omega)U^*(x_A, \omega)\rangle \quad , \tag{A-1}$$

where G is the Green's function between two receiver locations (x_A, x_B), $|S_N(\omega)|^2$ is the power spectrum of the noise, $U(x, \omega)$ is the spectrum of the wavefield at a given receiver location, $*$ represents the complex conjugate, and $\langle\cdot\rangle$ represents the time averaged ensemble. This equation suggests that to accurately estimate Green's functions from cross-correlations of ambient noise, the power spectrum of the noise must be known. However, the power spectrum of noise is typically not known. Furthermore, PSD analysis at Long Beach revealed that major sources of noise, particularly of anthropogenic origin, have highly-varying power spectrums in time and space.

To better estimate the Green's function between two receivers, we whiten the input traces prior to cross-correlation. In the frequency domain, this pre-whitened cross-correlation can be expressed as

$$[G(x_B, x_A, \omega) + G^*(x_B, x_A, \omega)] = \left\langle \left(\frac{U(x_B, \omega)}{\{|U(x_B, \omega)|\}} \right) \left(\frac{U^*(x_A, \omega)}{\{|U(x_A, \omega)|\}} \right) \right\rangle \quad , \quad \text{(A-2)}$$

where $|\cdot|$ indicates the real absolute value of the spectrum and $\{\cdot\}$ indicates a 0.003 Hz running window for whitening the signal. In words, we divide the spectrum of each input trace by its smoothed amplitude spectrum prior to computing the cross-correlation. The denominator of the right side of Equation A-2 can be viewed as an estimate of the unknown power spectrum of noise in Equation A-1:

$$\{|U(x_B, \omega)|\}\{|U(x_A, \omega)|\} \approx |S_N(\omega)|^2 \quad . \quad \text{(A-3)}$$

By dividing the standard cross-correlation procedure by an estimate of the power spectrum of noise in the frequency domain, we are trying to deconvolve the noise source signal from the correlations in the time domain to better estimate the Green's function. Although this process compromises the amplitude information of the correlation, it does not effect the phase information. We then normalize all the correlations by their respective maximum amplitude to limit the effect of anomalous events and instrument spikes on the correlation results.

To demonstrate the effectiveness of pre-whitening the correlations, we apply multiple processing techniques to the ambient seismic noise to create similar virtual source gathers as seen in the main body. Figure A-1 shows gathers at low frequencies (0.5 Hz to 2 Hz), while Figure A-2 shows gathers at high frequencies (2 Hz to 4 Hz). In both figures, (a) represents the standard correlation procedure, (b) represents the standard correlation procedure with normalization, (c) represents the correlation procedure with pre-whitening, and (d) represents the correlation procedure with pre-whitening and normalization (which is what is displayed in the main body). Again, for the lower frequency range we average correlations from 100 two-hour time windows, and for the higher frequency range we average correlations from 240 two-hour time windows.

At both frequency ranges, we see that the standard correlation procedure allows anomalous events to overwhelm the correlation results, rendering the results unusable. Normalization of the correlations prior to stacking greatly improves the results, allowing the moveouts to be clearly observed. However, there is clearly a bias toward the higher frequencies in both cases. Pre-whitening broadens the spectrums of the correlations and balances the amplitudes, thus improving the quality of the correlations. Although further amplitude normalization of each correlation prior to stacking does not greatly enhance the signal to noise ratio of the resulting gather at lower frequencies, it does slightly improve the signal to noise ratio for the resulting gather at higher frequencies.

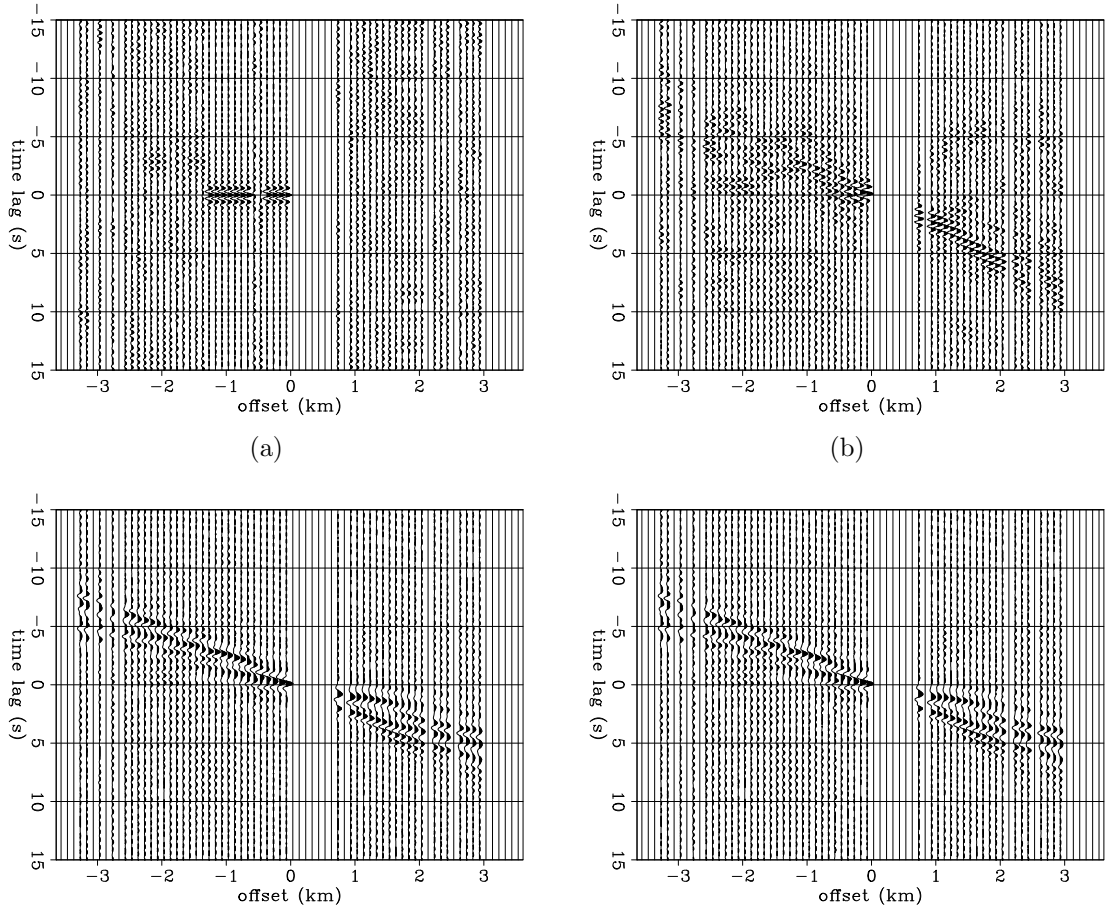


Figure A-1: Virtual source gathers at the frequency range 0.5 Hz to 2 Hz from various processing procedures. (a) standard correlation procedure, (b) standard correlation procedure with normalization, (c) correlation procedure with pre-whitening, and (d) correlation procedure with pre-whitening and normalization. [**CR**]

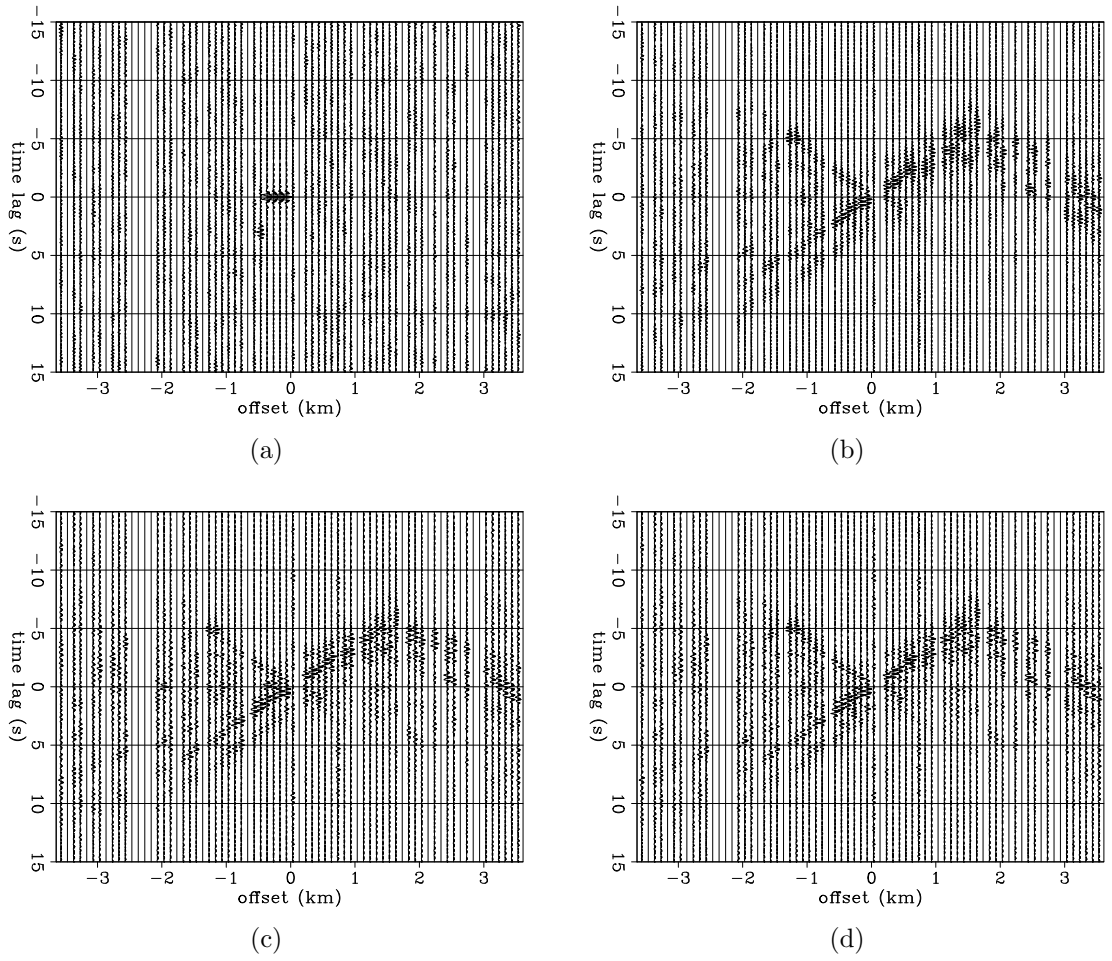


Figure A-2: Virtual source gathers at the frequency range 2 Hz to 4 Hz from various processing procedures.(a) standard correlation procedure, (b) standard correlation procedure with normalization, (c) correlation procedure with pre-whitening, and (d) correlation procedure with pre-whitening and normalization. [**CR**]

REFERENCES

Bensen, G., M. Ritzwoller, M. Barmin, A. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang, 2007, Processing seismic ambient noise data to obtain reliable broadband surface wave dispersion measurements: Geophysical Journal International, **169**, 1239–1260.

Bensen, G., M. Ritzwoller, and N. Shapiro, 2008, Broadband ambient noise surface wave tomography across the United States: Journal of Geophysical Research, **113**, 1–21.

Delgado, J., C. Lopez Casado, A. Estevez, J. Giner, A. Cuenca, and S. Molina, 2000, Mapping soft soils in the segura river valley (se spain): a case study of microtremors as an exploration tool: Journal of Applied Geophysics, **45**, 19–32.

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

Draganov, D., K. Wapenaar, W. Mulder, J. Singer, and A. Verdel, 2007, Retrieval of reflections from seismic background-noise measurements: Geophysical Research Letters, **34**, 1–4.

Greenwoood, R., 1998, Liquefaction zones in the Long Beach 7.5-minute quadrangle, Los Angeles County, California: California Department of Conservation, Division of Mines and Geology, Seismic Hazard Zone Report, **28**, 3–20.

Lin, F.-C., D. Li, R. Clayton, and D. Hollis, 2012, Interferometry with a dense 3D dataset: SEG Technical Program Expanded Abstracts, **1**, 1–6.

McNamara, D. and R. Boaz, 2005, Seismic noise analysis system using power spectral density probability density functions: A stand-alone software package: USGS Open-File Report, 1–14.

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

Rost, S. and C. Thomas, 2002, Array seismology: Methods and applications: Reviews of Geophysics, **40**, 2–1–2–27.

Roux, P., K. Sabra, P. Gerstoft, W. Kuperman, and M. Fehler, 2005, P-waves from cross-correlation of seismic noise: Geophysical Research Letters, **32**, 1–4.

Sabra, K., P. Gerstoft, P. Roux, W. Kuperman, and M. Fehler, 2005, Surface wave tomography from microseisms in Southern California: Geophysical Research Letters, **32**, 1–4.

Shapiro, N., M. Campillo, S. L., and M. Ritzwoller, 2005, High-resolution surface-wave tomography from ambient seismic noise: Science, **307**, 1615–1618.

Snieder, R., 2004, Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase: Physical Review E, **69**, 046610–1–046610–8.

Wapenaar, K., D. Draganov, R. Snieder, X. Campman, and A. Verdel, 2010, Tutorial on seismic interferometry: Part 1 - Basic principles and applications: Geophysics, **75**, 75A195–75A209.

Yang, Y., M. Ritzwoller, F. Lin, M. Moschetti, and N. Shapiro, 2008, Structure of the crust in uppermost mantle beneath the western United States revealed by amibent

noise and earthquake tomography: Journal of Geophysical Research, **113**, 1–9.
Zhan, Z., S. Ni, D. Helmberger, and R. Clayton, 2010, Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise: Geophysical Journal International, **182**, 408–420.