Green's functions extraction and surface-wave tomography from microseisms in southern California

Peter Gerstoft[1], Karim G. Sabra[1], Philippe Roux[1], W. A. Kuperman[1], and Michael C. Fehler[2]

ABSTRACT

We use crosscorrelations of seismic noise data from 151 stations in southern California to extract the group velocities of surface waves between the station pairs for the purpose of determining the surface-wave velocity structure. We developed an automated procedure for estimating the Green's functions and subsequent tomographic inversion from the 11,325 station pairs based on the characteristics of the noise field. We eliminate specific events by a procedure that does not introduce any spurious spectral distortion in the band of interest, 0.05–0.2 Hz. Further, we only used the emerging arrival structure above a threshold signal-to-noise ratio. The result is that mostly station pairs with their axes oriented toward the sea are used, consistent with the noise having a microseism origin. Finally, it is the time derivative of the correlation function that is actually related to the Green's function. The emergence of the time-domain Green's function is proportional to the square root of the recording time and inversely proportional to the square root of the distance between stations. The tomographic inversion yields a surface-wave velocity map that compares favorably with more conventional and elaborate experimental procedures.

INTRODUCTION

By crosscorrelating noise recordings at two seismic stations over a sufficiently long time, it is possible to extract the time-domain Green's function (TDGF) between the two stations. The TDGF is derived from the part of the signal that remains coherent between the two recordings. Thus, it is necessary that a small fraction of the coherent energy remain in both recordings in the frequency band of interest. The noise crosscorrelation function (NCF) is defined as the crosscorrelation of the noise-only time series, i.e., major earth-

quakes have been removed. Experimental and theoretical studies have shown that the arrival-time structure of the TDGF can be estimated from the time derivative of the NCF in various environments and frequency ranges of interest: helioseismology (Rickett and Claerbout, 1999), ultrasonics (Weaver and Lobkis, 2001, 2003, 2004; Larose et al., 2004; Malcolm et al., 2004), ocean acoustics (Roux et al., 2004, 2005b; Sabra et al., 2005c), and seismology (Shapiro and Campillo, 2004; Snieder, 2004; Wapenaar, 2004; Roux et al., 2005a; Sabra et al., 2005a, 2005b; Shapiro et al., 2005).

The seismic noise field might not be isotropic when dominated by directional noise sources, such as ocean microseisms (Bromirsky and Duennebier, 2002). In such cases, the station-pair orientation will matter. To a first-order approximation for a weakly scattering medium, the main contribution to the NCF comes from noise sources close to the station-pair axis (Roux et al., 2004; Snieder, 2004; Sabra et al., 2005a). Noise sources away from the station-pair axis typically have a weaker contribution to the NCF. Thus, for a weakly scattering medium and an anisotropic noise-source distribution, coherent waveforms corresponding to the TDGF arrival structure do not merge well for station pairs perpendicular to the main propagation direction of the noise sources.

As suggested by Shapiro et al. (2005) and Sabra et al. (2005b), these extracted TDGFs can then be used for tomographic imaging purposes. This possibility is very attractive, as the existing worldwide networks of broadband seismic stations could form a potentially much denser set of observations than if only earthquakes were used. Such an approach could also easily be used for monitoring change in the medium between the sensors.

We estimate the TDGF and associated surface-velocity maps by using continuous-noise data collected on 151 broadband stations of the southern California network sampled at 1 Hz for the month of July 2004. Our basic approach is the same as used in Sabra et al. (2005a, 2005b), but we provide much more detail in the analysis of the TDGF extraction and in the tomographic approach. We extract the three-component Green's function, and then we discuss signal-to-noise ratio (SNR) convergence (which is important for the emer-

gence of the TDGF) and spatial dependence of the TDGF convergence. In addition, we use a slightly different data set with longer observation time (30 versus 18 days) and lower sampling rate (1 Hz versus 20 Hz) and slightly different stations (151 versus 148). Shapiro et al. (2005) have also implemented a tomographic inversion of the southern California data set by using TDGF estimated from noise. Our approach differs on several points from theirs: (1) We use the time derivative of NCF (and not the NCF directly) as an estimate of the TDGF. (2) The influence of large events is reduced by clipping of the amplitudes of the noise time series (not by reducing the noise time series to one-bit time series). (3) A denser seismic array of 151 stations ($151 \times 150/2 = 11,325$ potential paths) instead of 62 ($62 \times 61/2 = 1891$ potential paths) is used, and a smaller covered area results in a higher path coverage and thus a higher tomographic resolution across a finer grid.

NOISE CROSSCORRELATIONS

For a spatially uniform broadband noise distribution in a homogeneous medium with sound speed c, it can be argued (Snieder, 2004; Roux et al., 2005a) that the NCF denoted by C of the traces at two receivers separated by distance L is zero for time $|t| > L/c$ (no wave can travel slower than c), noncontinuous at $t = \pm L/c$, and continuous for $|t| < L/c$ (waves originating from broadside will appear to travel faster than c). Thus, the derivative of the NCF is

$$\frac{dC}{dt} \approx -\delta(t - L/c) + \delta(t + L/c). \quad (1)$$

This relationship is indicated in Figure 1a for the free-space case where the TDGF is a delta function at $t = L/c$. Note that the NCF does not resemble the TDGF, but the derivative of the NCF is proportional to the free-space TDGF and the time-reversed TDGF. However, we usually work with band-limited signals, and then both the NCF and its derivative will have a waveformlike structure, as indicated for the two filtered traces.
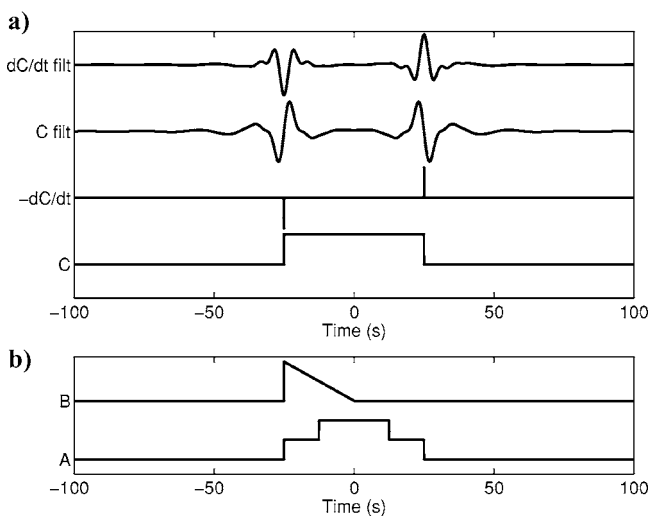

Figure 1. Examples of noise crosscorrelation. (a) Bottom to top: a simple NCF and its derivative, and the filtered (0.05–0.2 Hz) NCF and its derivative. (b) Two possible noise crosscorrelation functions for a single-sided noise distribution (trace B) and a noise crosscorrelation function from two dominating noise-propagation scenarios.

The NCF depends on the noise distribution, and all theory developed to date assumes a uniform spatial distribution (e.g., Weaver and Lobkis, 2001, 2004; Snieder, 2004; Roux et al., 2005a). However, it can be argued that a noise field dominated by a one-sided noise distribution centered in a direction toward the axis pair would quantitatively have a shape as suggested in trace A in Figure 1b. This shape might represent noise from microseisms from the direction of the ocean. Another important case is a noise field that has two uncorrelated propagating wave components, as is illustrated by trace B (Figure 1b), which might represent noise signals propagating as surface and body waves.

The crosscorrelation C_{ij} is computed from the observed displacement fields $v_i(\mathbf{r}_1,t)$ at location $\mathbf{r}_1$ in direction i and $v_j(\mathbf{r}_2,t)$ at location $\mathbf{r}_2$ in direction j by integrating the time t over the whole observation period T,

$$C_{ij}(1,2,t) = \int_0^T v_i(\mathbf{r}_1,\tau)v_j(\mathbf{r}_2,t + \tau)d\tau. \quad (2)$$

In practice, it is faster to compute the crosscorrelation by sectioning the data and computing the crosscorrelation for each time period. We section the data for each day and then average over all observed days.

On the basis of analytic derivations for specific propagation models (Weaver and Lobkis, 2001, 2004; Snieder, 2004; Roux et al., 2005a), we construct the following relationship for the time derivative of the noise crosscorrelation $C_{ij}(\mathbf{r}_1,\mathbf{r}_2,t)$ between two seismic stations, station 1 (located at $\mathbf{r}_1$ recording component i) and station 2 (located at $\mathbf{r}_2$ recording component j) for the TDGF $G_{ij}(\mathbf{r}_1;\mathbf{r}_2,t)$:

$$\frac{dC_{ij}(\mathbf{r}_1,\mathbf{r}_2,t)}{dt} \approx -G_{ij}(\mathbf{r}_1;\mathbf{r}_2,t) + G_{ij}(\mathbf{r}_1;\mathbf{r}_2,-t). \quad (3)$$

The TDGF $G_{ij}(\mathbf{r}_1;\mathbf{r}_2,t)$ relates a unit force in direction i at $\mathbf{r}_1$ to the displacement response in direction j at receiver $\mathbf{r}_2$.

In equation 3, the terms on the right-hand side are as follows: (1) the TDGF that comes from noise events that propagate from station 1 to 2 and yields a positive correlation time delay t and (2) the time-reversed TDGF that comes from noise events that propagate from station 2 to 1 and yields a negative correlation time delay $-t$. Thus, for a uniform noise-source distribution surrounding the two stations, the derivative of the NCF will be a symmetric function with respect to the arrival time because seismic noise sources are distributed on both sides of the station pair. However, in the case of a predominant directional noise source (e.g., noise originating from ocean microseisms; see next section) the NCF and its time derivative likely will be one-sided.

Several researchers have used a form similar to equation 3, but instead of the derivative, they just used the crosscorrelation (e.g., Campillo and Paul, 2003; Shapiro and Campillo, 2004; Weaver and Lobkis, 2003). The noise recorded from broadband seismic stations is typically a process with a relatively small bandwidth. Hence the time series of the NCF and its time derivative will have similar arrival structure but with a phase shift ($\pi/2$ for narrow-band signals). For a narrow-band signal, the variation of the envelope of the signal and of the time-derivative signal is similar. For the traveltime estimation, the envelope of the NCF is used.

PROCESSING

For demonstrating the approach, data from the month of July 2004 for all seismic stations available from the Southern California Data Center were used, as indicated in Figure 2. High-frequency effects were investigated by sampling the broadband data at 20 Hz (Figures 3 and 4). All other data processing (Figures 5–16) is based on data that are sampled at 1 Hz. The network contains 151 three-component stations; crosscorrelating the observed time series for an entire month between all the stations gives potentially $(151 \times 150/2) \times 3 \times 3 = 11{,}325 \times 9 = 101\,925$ seismic traces. Thus, it is important that all the processing is automatic and efficient.

Crosscorrelating one station pair

In order to estimate the NCF, the noise traces are first band-pass filtered at the frequency band 0.05–0.2 Hz, roughly corresponding to the main band of the microseisms (Bromirsky and Duennebier, 2002). Next, we note that this crosscorrelation technique works best when the noise distribution is uniform in space and time (Larose et al., 2004; Snieder, 2004; Shapiro and Campillo, 2004; Roux et al., 2004). Hence, the effects of large seismic events should be minimized because they would otherwise dominate the arrival-time structure of the NCF. Clipping all signals above a certain threshold reduces these events. The clipping threshold is determined as the minimum of the standard deviations measured over each day, because of the expectation that few events appear in that time interval, and the threshold is only determined by the noise. In this way, the effect of large events is reduced, but the high-frequency content of the seismic noise spectrum is less distorted. An example of the processing is shown in Figure 3a for station-pair JVA-SDG. An arrival is clearly seen centered at $t = 60$ s.

A straightforward way to avoid the influence of large events is to disregard the amplitude completely by keeping only the sign of the signals (Campillo and Paul, 2003; Shapiro and Campillo, 2004). An example of this single-bit truncation, carried out for station-pair

JVA-SDG in Figure 3b, gives identical results to the threshold clipping just described. In multiple scattering experiments for ultrasonics, it has been reported that single-bit truncation gives a faster convergence because the one-bit correlation tends to give more weight to the longest and most diffracted paths (Larose et al., 2004). In our processing, faster convergence was not observed for the single-bit crosscorrelation, likely because we only had one dominant arrival. However, as indicated in Figure 4, the rougher clipping (solid line) artificially modifies the low-amplitude parts of the seismic noise spectrum. This result could become important when processing higher frequencies where the signal is weak. The strong peak in ener-

Figure 3. NCF for station pairs JVA-SDG for each of 30 days during July 2004 after application of band-pass filter at 0.025–1 Hz. (a) Crosscorrelation based on the truncated noise trace for each station. (b) Crosscorrelation based on the single-bit noise trace. The top trace in each box shows mean traces for the 30-day period.
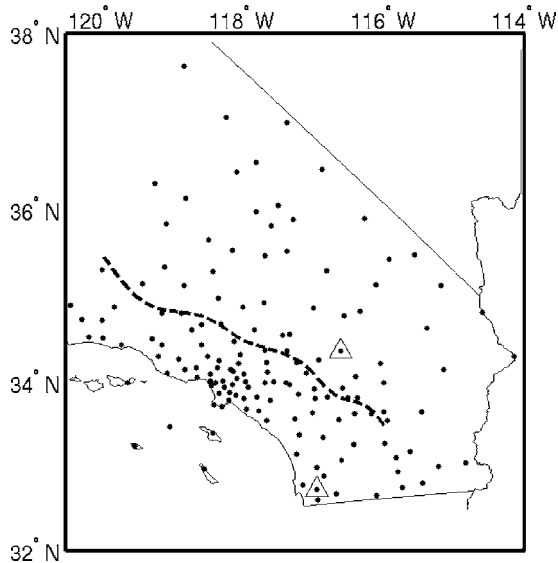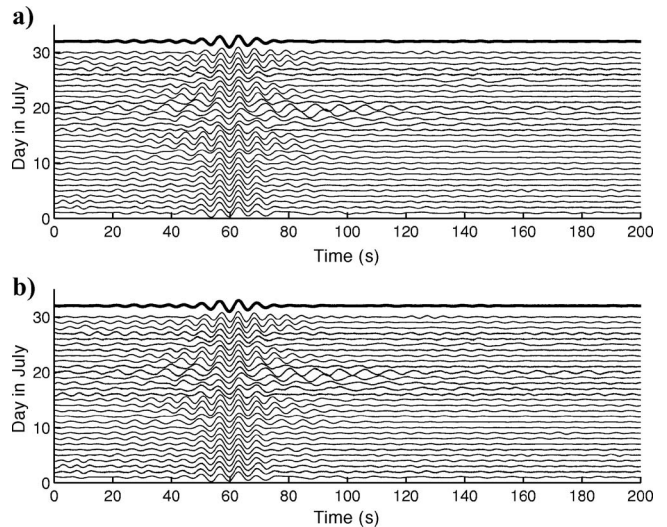

Figure 2. Map of the 151 online stations in the Southern California Seismic Network. Stations JVA (on the north) and SDG (on the south) are marked with triangles. The dashed line indicates the San Andreas Fault.
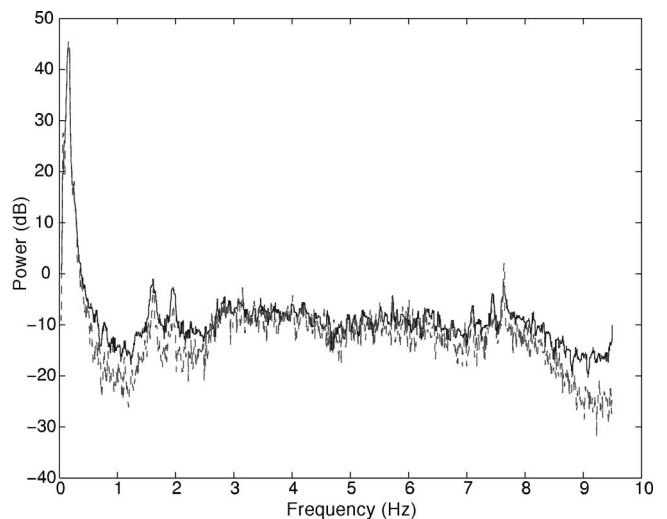

Figure 4. Frequency-domain response of the 30-day NCF in Figure 3 using full-noise trace (dashed) for each station and single-bit trace (solid).

gy at 0.1–0.2 Hz is very clear (50 dB above average noise level) and is consistent with the peak in microseisms (Aki and Richards, 1980; Bromirsky and Duennebier, 2002).

To mitigate the frequency dependence of the noise spectrum, equalization in the frequency interval 0.05–0.2 Hz is applied in the remainder of the paper. The purpose of this processing is to homogenize the noise signal over the frequency interval for which seismic noise is still present. Thus, the results are less sensitive to spectral characteristics of the noise. The crosscorrelation is usually done by using two forward (one for each component) and one inverse Fourier transform. When combining all station pairs to form the crosscorrelation, it is computationally efficient to first transform all noise re-

cordings to the frequency domain, whereby to a first-order approximation, the computational effort of the crosscorrelation is reduced to just one inverse Fourier transform.

Extracting record sections

The just-described crosscorrelation of the vertical noise components was applied to all $151 \times 150/2 = 11,325$ station pairs in the Southern California Seismic Network by using a 0.05–0.2 Hz passband and averaging over all days in July. In all cases, we choose the first station to be west of the second station, in order to remove orientation ambiguity. Figure 5 shows record sections from this process-
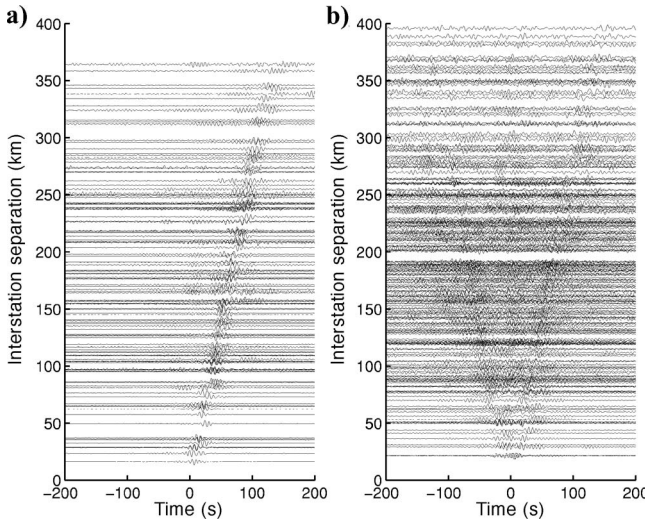

Figure 5. Time series of the derivative of the NCF for all station paths with a distance of < 400 km. Station pairs having an azimuth (a) between 18° and 20° and (b) between 130° and 132° for vertical components band-pass filtered at 0.1–0.2 Hz. Each trace has been normalized to the same amplitude.

Figure 7. SNR versus range for each seismic trace in Figure 5 for station pairs having a bearing (a) between 18° and 20° and (b) between 130° and 132°. The solid line shows a $1/\sqrt{R}$ decay.
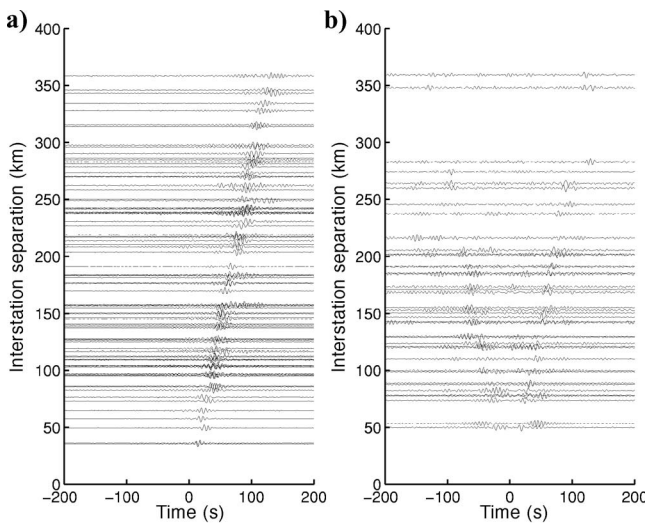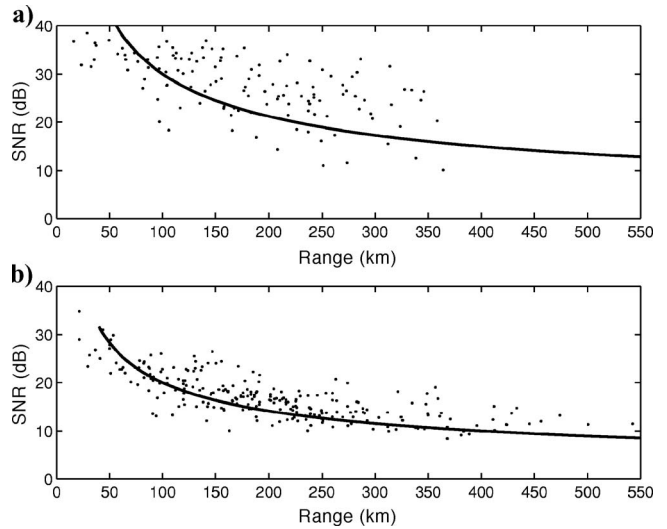

Figure 6. Time series of the derivative of the NCF for all station paths used in the tomography with a distance of < 400 km. Station pairs having an azimuth (a) between 18° and 20° and (b) between 130° and 132° for vertical components band-pass filtered at 0.05–0.2 Hz. Each trace has been normalized to the same amplitude.
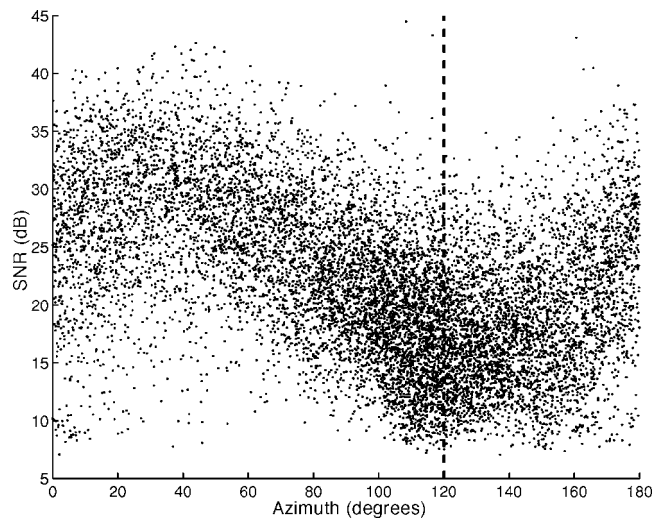

Figure 8. Distribution of SNR of the NCF versus the azimuth of all station pairs. Because pairs were ordered with first station to west of second station, all azimuths are between 0° and 180°. The coastline orientation has a bearing of approximately 120° (vertical dashed). The overall SNR of the NCF for station pairs oriented perpendicular to the coast (bearing angle about 10°–50°) is higher than for station pairs parallel to the coast.

ing based on the NCF derivative. The envelope is computed by using a Hilbert transform. For the station pairs oriented parallel to the main noise-propagation direction (perpendicular to the coast) (Figure 5a), the recovered signals are propagating wavetrains with an average group velocity of 2.8 km/s. Most of the noise originates from microseisms from the direction of the ocean (between southwest and west); therefore, the time series for station pairs parallel to this noise-propagation direction (Figure 5a) emphasize the forward-propagating Green's function. For the station pairs perpendicular to the main noise-propagation direction (Figure 5b), two wavetrains can be seen propagating in both the forward and backward directions with an average group velocity of 2.8 km/s. Clearly, some of these traces are very noisy, and for the tomography, we use empirical rules (see the Tomography section) to remove the traces of lower quality. This procedure produces a much clearer record section, as shown in Figure 6.

Figure 11. Average velocity for all station pairs with an SNR > 25 dB, (a) as a function of angle and (b) as function of range.
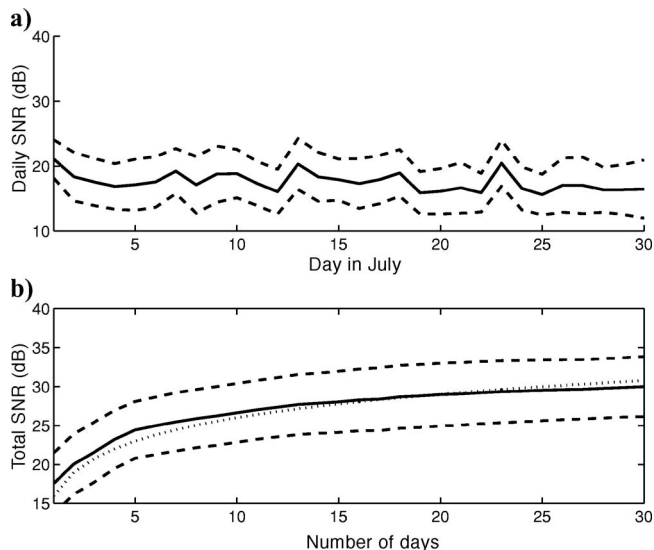

Figure 9. (a) Mean and standard deviation of SNR for each day for stations with a range separation of 200–220 km and an azimuth between 0° and 20°. (b) Mean and standard deviation of SNR computed by summing all combinations of traces for a number of days; a square-root dependence is shown as a dotted line.
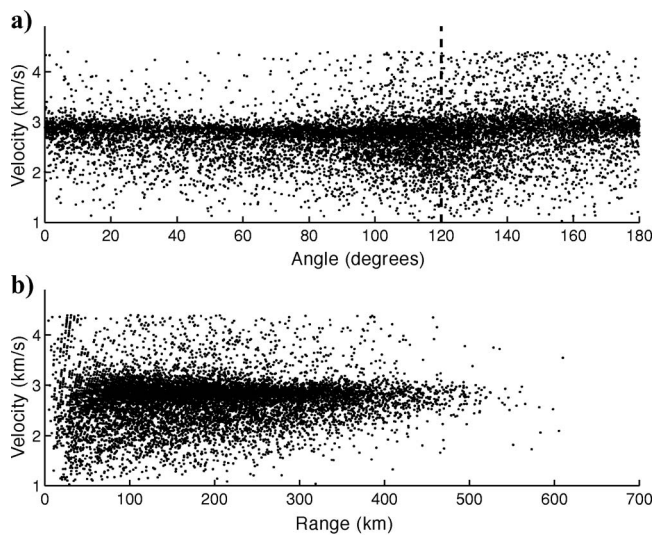

Figure 10. Average velocity for all station pairs with an SNR > 10 dB, (a) as a function of angle and (b) as function of range.

Figure 12. Time derivative of the NCF for all station pairs used in the tomography with a distance of <400 km. Station pairs having a bearing (a and c) between 18° and 20° and (b and d) between 130° and 132° for radial (a and b) and transverse (c and d) components band-pass filtered at 0.05–0.2 Hz. Each trace has been normalized to the same amplitude.

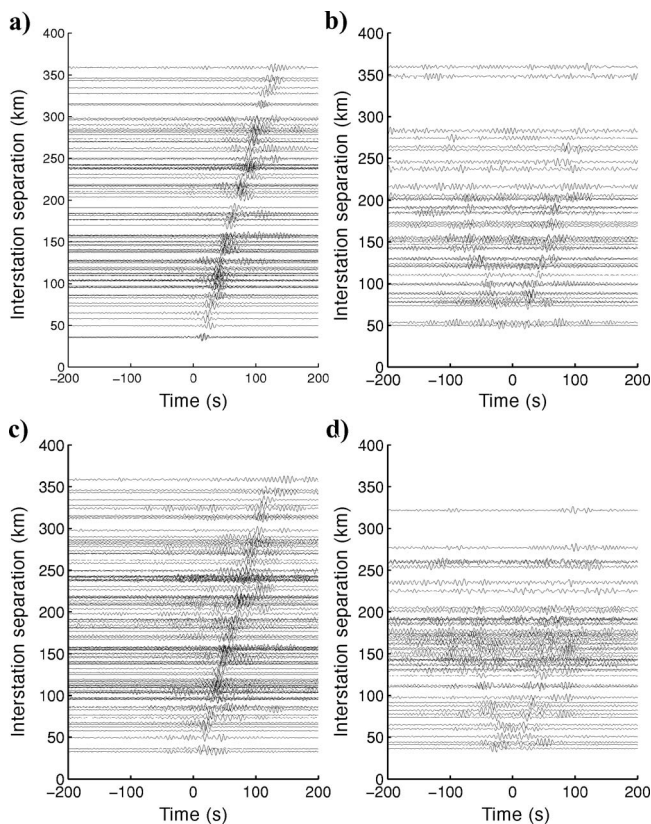Many of the time series perpendicular to the main noise-propagation direction are two-sided. Scattering from heterogeneities (Hennino et al., 2001) and geometric effects [e.g., such as reflections from the edge of a basin (Aki and Richards, 1980)] randomize the noise field and thus partially redistribute the ocean microseism arrival direction more uniformly. Some station pairs oriented parallel to the coast have a clear waveform, which may be due to scattered ocean noise propagating along their station-pair axes. However, the NCF for station pairs oriented perpendicular to the coast emerge overall more reliably and have sharper waveforms than the NCF for station pairs oriented parallel to the coast.

We estimated the group velocity of Rayleigh waves within the region containing the stations used in calculating the NCF record section by extracting an average layered P- and S-wave velocity model for the region from the 3D model of Kohler et al. (2003). Group velocity as a function of period was calculated for the model structure.
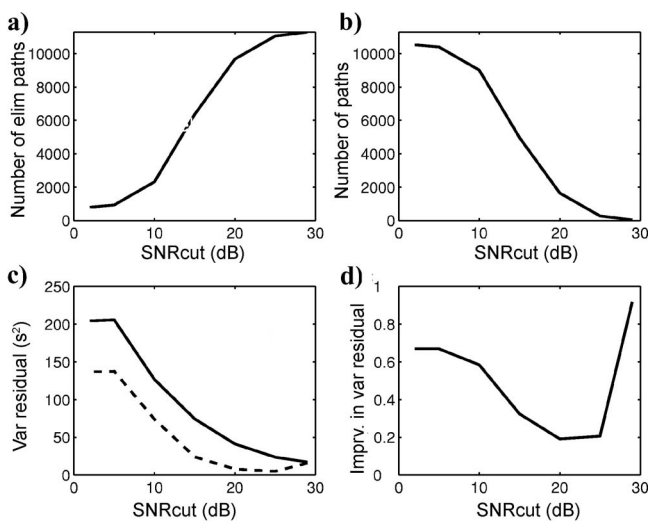

Figure 13. The tomography dependence of the cutoff SNR. (a) Number of eliminated paths. (b) Number of remaining paths. (c) A priori (solid) and a posteriori (dashed) variances of the residuals. (d) Improvement in variance of residuals.
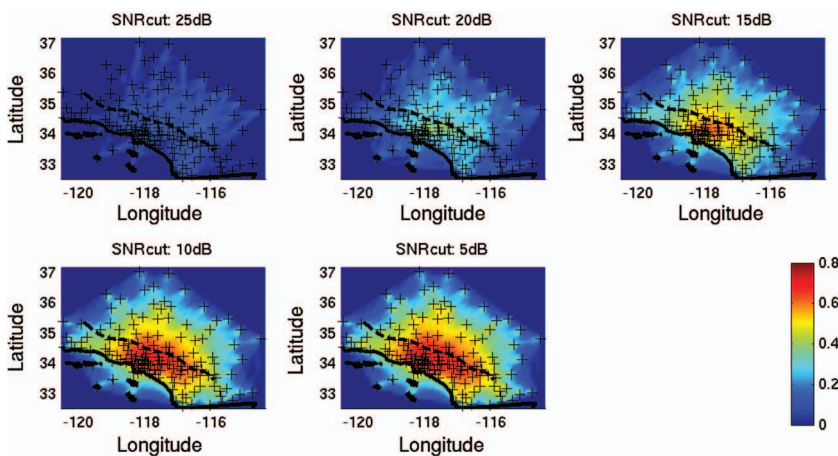
The estimated group velocity ranges from 2.8 km/s at a 5-s period to 3.0 km/s at a 10-s period, which agrees well with the estimated group velocity of 2.8 km/s at a period of approximately 6 s from our NCF record section.

Extracting traveltime and SNR

An important issue is the rate at which the waveforms emerge as a function of averaging time and separation distance (range). We define the SNR as the ratio (in decibels) of the peak in a 50-s time window centered around the main arrival r/c ($c = 2.8$ km/s, and r is station-pair distance) and the standard deviation for a noise-only time window, -400 to -350 s. For the two record sections in Figure 5 we see that the corresponding SNR in Figure 7 is much higher for the station pairs oriented perpendicular to the coast (Figure 7a) than for those oriented parallel to the coast (Figure 7b); thus, an azimuthal dependence clearly exists. This dependence is emphasized in Figure 8, which depicts the SNR for all station pairs as a function of azimuth. For the range dependence of SNR, intuitively and to a first-order approximation, we expect the SNR to be proportional to $1/\sqrt{R}$ (solid line in Figure 7) because of the geometrical spreading and because the signal is dominated by Rayleigh waves. This proportionality seems to hold reasonably well. The SNR in the direction perpendicular to the coast has a larger spread, possibly because more spread in the noise level takes place in this direction.

In Figure 9, we focus on stations within a specific range in the far-field, 200–220 km. The SNR for each day (Figure 9a) shows moderate variation as the noise and seismicity vary. We sum the crosscorrelation time series for each day to obtain the time series for N days and compute the SNR. Each day is treated as an independent observation; therefore, for each N the sum is computed for all combinations of all days. The accumulated SNR based on the summed traces (Figure 9b) shows clearly that the SNR increase is proportional to the square root of recording time, in agreement with Roux et al. (2004) and Snieder (2004).

For each waveform, the traveltime is estimated as the maximum of the envelope, resulting in $151 \times 150/2 = 11{,}325$ estimates of traveltime. By using these traveltimes and the distance between the stations, the average group velocity for each station pair is computed. The extracted velocities are summarized in Figure 10 for SNR > 10 dB and in Figure 11 for SNR > 25 dB. Despite the variation in geology, the propagation velocity between the stations is commonly approximately 2.8 km/s. For the high SNR ratio (Figure 10b), it is seen that there is a major variation in velocity only at short ranges. This finding could be because of (1) traveltime being more difficult to pick at short ranges, as many traces have some incoherent noise at $t = 0$, or (2) real velocity fluctuations in the medium that produce fluctuations in the traveltime due to less path averaging. Further, the low-frequency waveforms are not as developed for short ranges.

Most of the noise comes from the direction of the ocean; therefore, the velocities might be biased if, by using a plane-wave analogy, the station pairs are not aligned with this propagation direction. Figure 11 shows the estimated velocity versus azimuth (Figure 11a) and range (Figure 11b) for all paths having an SNR > 25 dB. If only

Figure 14. Diagonal of the resolution matrix for the inversion model keeping only station pairs having an SNR $> S_{cut}$. Crosses indicate the location of the 151 seismic stations of the Southern California Seismic Network; scale has no units.

paths with an SNR > 25 dB (Figure 11) are selected, the paths parallel to the coast are nearly eliminated as these have small SNR values, in accordance with the results shown in Figure 8. On the basis of the information available, it is not possible to determine whether the estimated velocities for station pairs parallel to the coast are biased, or whether the velocities are more difficult to determine owing to the lower SNR levels. These traveltimes could influence the quality of our tomography.

Other Green's function components

Because we can crosscorrelate all components of motion, it is also possible to estimate all nine elements of the Green's function between station pairs. We begin by projecting the horizontal north and east components of motion for each station into radial and transverse components of motion. The vertical component is obtained directly from the seismometers. We have indicated (Sabra et al., 2005b, Figure 6) that the arrival and polarization plots show strong correlations between the vertical and radial components of motion, indicating that Rayleigh wave propagation is dominant for these, in agreement with Campillo and Paul (2003). The transverse part seems slightly slower and decoupled from the other two directions, and this circumstance could indicate a Love wave.

The elements containing the radial and horizontal components of motion are more noisy than those containing only vertical components of motion, as shown by comparing Figure 6 for the vertical-vertical correlation with the radial-radial correlation in Figure 12a and b and transverse-transverse correlation, Figure 12c and d. The lower SNR ratio is especially evident for the station pairs parallel to the coast (Figures 6b, 12b, and 12d).

TOMOGRAPHY USING VERTICAL COMPONENT OF THE GREEN'S FUNCTION

As shown in Shapiro et al. (2005) and Sabra et al. (2005b), it is possible to use the traveltimes for a detailed Rayleigh-wave-based tomographic inversion. We investigate this possibility in more detail. The southern California region is divided into 13×16 km cells of constant group velocity for the tomographic inversion. The propagation paths are assumed to be straight rays. A classic tomographic method is used to construct the group-velocity map (Tarantola, 1987; Menke, 1989; Rodgers, 2000). If $\mathbf{T}$ is the column vector of the arrival times estimated from the time derivative of the NCF waveforms, $\mathbf{S}$ is the column vector of the slowness of each grid cell (arranged in a lexicographic order), and if a simple linear model is assumed, the inversion problem reduces to

$$\mathbf{T} = \mathbf{KS}, \tag{4}$$

where $\mathbf{K}$ is the forward-mapping matrix (or kernel) indicating for each particular straight ray its path length across each cell. The group-velocity value is assigned to the full rectangular cell.

An averaged arrival-time measurement error of $\sigma_T = 2$ s is assumed (and thus a simple diagonal measurement-error covariance matrix), in agreement with the frequency-bandwidth resolution. The

elements of the a priori error covariance matrix of the cell slownesses $\boldsymbol{\Sigma}_s$ are

$$\boldsymbol{\Sigma}_s(i,j) = \sigma_s^2 \exp(- D_{i,j}/L), \tag{5}$$

where $\sigma_s = \sigma_c/c_0^2$, $c_0 = 2.8$ km/s is average regional group velocity, and σ_c is set to 0.15 km/s on the basis of group-velocity measure-
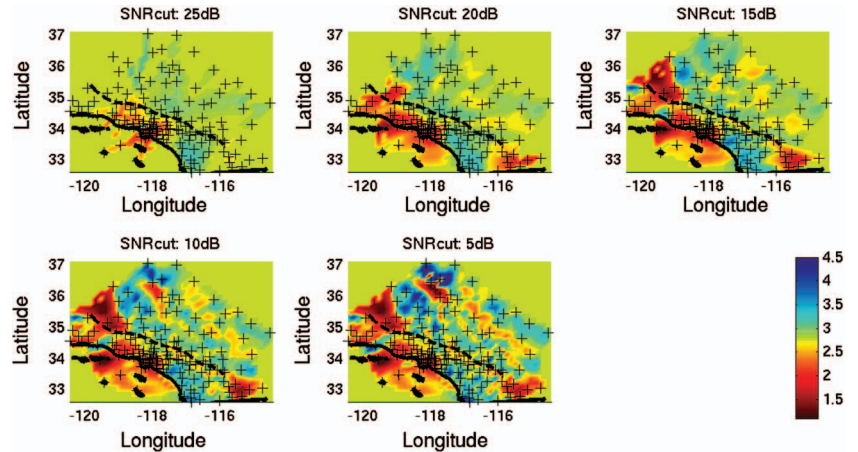

Figure 15. Surface group-velocity (scale in km/s) maps corresponding to the maximum a posteriori solution for the tomographic inversion as a function of S_{cut}.
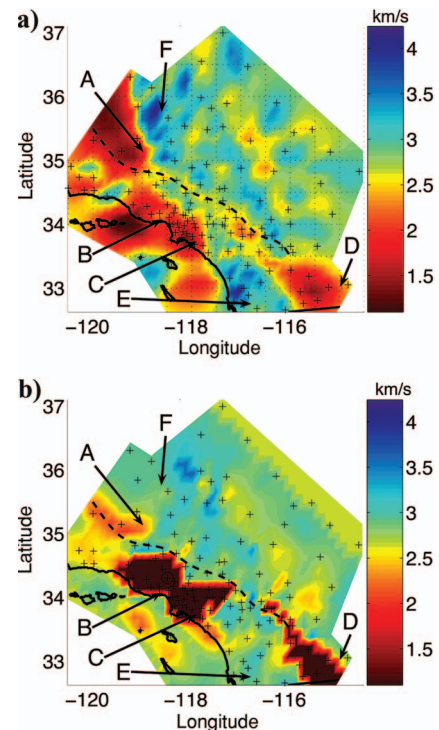
Figure 16. Surface group-velocity maps. (a) The maximum a posteriori solution for the tomographic-inversion scheme. Low-resolution areas (determined from Figure 13) are masked. (b) Estimated map of the Rayleigh wave group velocity at 7.5 s, constructed from dispersion analysis (Herrmann, 2002) based on a 3D velocity model for the southern California region (Kohler et al., 2003). A = San Joaquin valley; B = Ventura; C = Los Angeles; D = Salton Sea Trough; E = Peninsular Ranges; F = Sierra Nevada. The dashed line indicates the San Andreas Fault. From Sabra et al. (2005c).

ments from 1.1 to 4.5 km/s. $D_{i,j}$ is the distance (in kilometers) between the centers of the ith and jth grid cells, and L is a smoothness scale set to 30 km so that the smoothing extends over two grid cells (Tarantola, 1987; Rodgers, 2000). The resolution matrix (or averaging kernel matrix) is then

$$\mathbf{R} = (\mathbf{K}^t \boldsymbol{\Sigma}_T^{-1} \mathbf{K} + \boldsymbol{\Sigma}_S^{-1})^{-1} \mathbf{K}^t \boldsymbol{\Sigma}_T^{-1} \mathbf{K}, \qquad (6)$$

where the superscript t indicates a transpose matrix. We use a homogeneous starting model based on the a priori average regional group velocity $c_0 = 2.8$ km/s. Hence a priori slowness vector $\mathbf{S}_0$ was set to have constant elements $S_0(i,j) = 1/c_0$. The inversion is then performed on the arrival-time residuals. The maximum a posteriori solution $\mathbf{S}_m$ for the linear problem defined by equation 4 is then

$$\mathbf{S}_m = \mathbf{S}_0 + (\mathbf{K}^t \boldsymbol{\Sigma}_T^{-1} \mathbf{K} + \boldsymbol{\Sigma}_S^{-1})^{-1} \mathbf{K}^t \boldsymbol{\Sigma}_T^{-1} (\mathbf{T} - \mathbf{KS}_0). \quad (7)$$

The a posteriori covariance matrix map $\boldsymbol{\Sigma}$ associated with $\boldsymbol{\Sigma}_m$ (i.e., the a posteriori error) is directly related to the resolution matrix $\mathbf{R}$:

$$\boldsymbol{\Sigma}_m = (\mathbf{I} - \mathbf{R}) \boldsymbol{\Sigma}_S. \qquad (8)$$

The paths used in the inversion are selected automatically, by using the following four criteria:

1) The range has to be larger than 30 km. This size removes arrivals in which the surface wave has not completely developed.
2) On the basis of a priori estimates of the variations of the group velocity in the regions, the traveltimes were picked only when the corresponding straight-path velocity was in the interval 1.1–4.5 km/s.
3) The SNR has to be larger than a certain threshold, designated the S_{cut}.
4) Some waveforms appear noisy from $t = 0$ until the surface-wave arrival, and this fact makes the automatic arrival-time picking difficult. Such data are removed by requiring that the ratio of the main arrival to the average signal around $t = 0$ be larger than S_{zero}. We have found that $S_{zero} = S_{cut} - 3$ works well. Much energy around $t = 0$ is due to correlation of nonpropagating energy. Arrivals appearing before the surface-wave arrival can be due to other wave types (e.g., body waves at short separation distances). Spurious early arrivals can also appear if the noise distribution is dominated by a single wave propagating at an angle to the station pair.

The most important of these criteria is the SNR criteria (item 3). The number of eliminated (Figure 13a) and used (Figure 13b) station paths in the inversion versus the SNR cutoff ratio changes with increasing S_{cut}. The less reliable station paths tend to have errors so large that the inversion cannot improve the traveltimes mismatch, and as these are removed, the posterior fit becomes better, as indicated in Figure 13c and d, which show the variance of the traveltime error. Contrary to Shapiro et al. (2005), we do not remove large outliers. Retaining them is simpler, and the outliers can potentially contain useful information, but our variance reduction will not be as large. When station pairs are removed, the resolution also becomes smaller, as indicated in Figure 14 for several values of S_{cut}. The best resolution is naturally obtained when all station pairs are used in the inversion.

A more direct approach is to just perform the inversion for each value of S_{cut} (Figure 15). For high values of S_{cut}, only a few paths exist, and not much detail can be seen. Small S_{cut} values yield just too

much detail. We judge that a value of S_{cut} of 10–15 seems appropriate.

To further validate our results we show the group-velocity map of Sabra et al. (2005b) (Figure 16a). This map is based on similar processing as used here, but the stations are slightly different, data are sampled at 20 Hz, and data are averaged over 18 days. The Sabra et al. (2005b) processing gave results similar to those shown in Figure 15 and produced a residual variance reduction of 50% relative to residuals for the homogeneous model. The group velocity obtained from the surface-wave inversion map (Figure 16a) correlates well with several geologic features of southern California (Jennings, 1977; Shapiro et al., 2005). Regions of slow surface-wave velocity (labels A–D) correspond to sedimentary basins. Fast group velocities occur in regions with mountain ranges (the Peninsular Ranges and Sierra Nevada; labels E and F, respectively). The main imaged geologic structure in Figure 15a is consistent with the results of Shapiro et al. (2005). However, a greater velocity contrast is observed (1) along the San Andreas fault and (2) on the west side of the Great Basin and the Mojave Desert, along a low-velocity zone extending approximately from 36°N, 117°W to 33.5°N, 116°W.

This tomographic map also agrees quantitatively with an estimated group-velocity map computed from analysis of dispersion curves at 7.5 s (Herrmann, 2002) based on a previous 3D model for P- and S-wave variations in the southern California region (Kohler et al., 2003) (Figure 16b). The average background group velocity and main geologic units in Figure 16 agree with the obtained tomographic maps, but overall the image resolution is lower. The very low group velocity for the sedimentary basins (A, B, and D) predicted by the 3D model (~ 0.5 km/s) differs from the obtained tomographic maps.

CONCLUSION

Estimates of the surface-wave components of the time-domain Green's function (TDGF) can be extracted from the time derivative of the noise crosscorrelation between two stations. These waveforms can be used as a basis for constructing geophysical models. The TDGF was estimated from crosscorrelations of continuous-noise recordings dominated by ocean microseisms in the frequency band 0.05–0.2 Hz. The observed emergence of the coherent waveforms confirms that they are proportional to the square root of the recording time and inversely proportional to the square root of the range between the stations.

An automatic procedure for estimating the Green's functions and the subsequent tomographic inversions was applied to 151 seismic stations (corresponding to 11,325 station pairs) in southern California. The traveltimes were estimated as the maximum of the envelope of the time derivative of the noise crosscorrelation, and only arrivals above a threshold signal-to-noise ratio were used. On the basis of the extracted traveltimes, a high-resolution tomographic map showing surface-wave group velocity in southern California was constructed. This map of surface-wave group velocity showed good agreement with maps obtained by more conventional and elaborate procedures. This success indicates that noise-only data could be used for extracting information about the medium it has propagated in.

ACKNOWLEDGMENTS

(Cooperative Agreement on Research and Education) program. Data were downloaded from the Southern California Earthquake Data Center (http://www.data.scec.org).

REFERENCES

Aki, K., and P. G. Richards, 1980, Qualitative seismology: W. H. Freeman and Company.

Bromirsky, P. D., and F. K. Duennebier, 2002, The near-coastal microseism spectrum: Spatial and temporal wave climate relationships: Journal of Geophysical Research, **107**, 2166, doi: 10.1029/2001JB000265.

Campillo, M., and A. Paul, 2003, Long-range correlations in the diffuse seismic coda: Science **229**, 547–549.

Hennino, R., N. Trégourès, N. M. Shapiro, L. Margerin, M. Campillo, B. A. van Tiggelen, and R. L. Weaver, 2001, Observation of equipartition of seismic waves: Physical Review Letters, **86**, 3447–3450.

Herrmann, R., 2002, Computer programs in seismology, version 3.15: Saint Louis University.

Jennings, C. W., 1977, Geologic map of California: California Division of Mines and Geology, map no. 2.

Kohler, M. D., H. Magistrale, and R. W. Clayton, 2003, Mantle heterogeneities and the SCEC reference three-dimensional seismic velocity model, version 3: Seismological Society of America Bulletin **93**, 757–774.

Larose, E., A. Derode, M. Campillo, and M. Fink, 2004, Imaging from one-bit correlations of wideband diffuse wavefields: Journal of Applied Physics, **95**, 8393–8399.

Malcolm, A. E., J. A. Scales, and B. A. van der Tiggelen, 2004, Retrieving the Green function from diffuse, equipartitioned waves: Physical Review E, **70**, 015601(R).

Menke, W., 1989, Geophysical data analysis: Discrete inverse theory: Academic Press, Inc.

Rickett, J., and J. Claerbout, 1999, Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring: The Leading Edge, **18**, 957–960.

Rodgers, C. D., 2000, Inverse methods for atmospheric sounding theory and practice: World Scientific Publishing Company.

Roux, P., W. A. Kuperman, and the NPAL Group, 2004, Extracting coherent wavefronts from acoustic ambient noise in the ocean: Journal of the Acoustical Society of America, **116**, 1995–2003.

Roux, P., K. G. Sabra, P. Gerstoft, W. A. Kuperman, and M. C. Fehler, 2005a, P-waves from cross correlation of seismic noise: Geophysical Research Letters, **32**, L19303, doi: 10.1029/2005GL023803.

Roux, P., K. G. Sabra, W. A. Kuperman, and A. Roux, 2005b, Ambient noise cross-correlation in free space: Theoretical approach: Journal of the Acoustical Society of America, **117**, 79–84.

Sabra, K. G., P. Gerstoft, P. Roux, W. A. Kuperman, and M. C. Fehler, 2005a, Extracting time-domain Green's function estimates from ambient seismic noise: Geophysical Research Letters, **32**, L03310, doi: 10.1029/2004 GL021862.

——, 2005b, Surface wave tomography from microseisms in southern California: Geophysical Research Letters, **32**, L14311, doi: 10.1029/2005 GL023155.

Sabra, K. G., P. Roux, and W. A. Kuperman, 2005c, Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide: Journal of the Acoustical Society of America, **117**, 164–174.

Shapiro, N. M., and M. Campillo, 2004, Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise: Geophysical Research Letters, **31**, L07614, doi: 10.1029/2004GL019491.

Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller, 2005, High-resolution surface-wave tomography from ambient seismic noise: Science **307**, 1615–1617.

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

Tarantola, A., 1987, Inverse problem theory: Elsevier Science Publishing Company, Inc.

Wapenaar, K., 2004, Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation: Physical Review Letters, **93**, 254301.

Weaver, R. L., and O. I. Lobkis, 2001, Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies: Physical Review Letters, **87**, 134301.

——, 2003, Elastic wave thermal fluctuations, ultrasonic waveforms by correlation of thermal photons: Journal of the Acoustical Society of America, **113**, 2611–2621.

——, 2004, Diffuse fields in open systems and the emergence of the Green's function: Journal of the Acoustical Society of America, **116**, 2731–2734.