Geophysical Journal International

Geophys. J. Int. (2016) **205,** 301–313
GJI Seismology

doi: 10.1093/gji/ggw033

Seismic interferometry of railroad induced ground motions: body and surface wave imaging

Diego A. Quiros, Larry D. Brown and Doyeon Kim

Department of Earth and Atmospheric Sciences, Cornell University, 112 *Hollister Drive, Ithaca, NY, USA. E-mail: daq7@cornell.edu*

Accepted 2016 January 19. Received 2016 January 5; in original form 2015 September 1

SUMMARY

Seismic interferometry applied to 120 hr of railroad traffic recorded by an array of vertical component seismographs along a railway within the Rio Grande rift has recovered surface and body waves characteristic of the geology beneath the railway. Linear and hyperbolic arrivals are retrieved that agree with surface (Rayleigh), direct and reflected *P* waves observed by nearby conventional seismic surveys. Train-generated Rayleigh waves span a range of frequencies significantly higher than those recovered from typical ambient noise interferometry studies. Direct *P*-wave arrivals have apparent velocities appropriate for the shallow geology of the survey area. Significant reflected *P*-wave energy is also present at relatively large offsets. A common midpoint stack produces a reflection image consistent with nearby conventional reflection data. We suggest that for sources at the free surface (e.g. trains) increasing the aperture of the array to record wide angle reflections, in addition to longer recording intervals, might allow the recovery of deeper geological structure from railroad traffic. Frequency–wavenumber analyses of these recordings indicate that the train source is symmetrical (i.e. approaching and receding) and that deeper refracted energy is present although not evident in the time-offset domain. These results confirm that train-generated vibrations represent a practical source of high-resolution subsurface information, with particular relevance to geotechnical and environmental applications.

Key words: Interferometry; Body waves; Surface waves; Wave propagation.

1 INTRODUCTION

Claerbout (1968) showed that the autocorrelation of the transmission response of a horizontally layered medium results in its reflection response. He later suggested that by cross-correlating the records from two locations on the surface, one can retrieve the response that would be recorded at one of these locations as if there were a source at the other. Although early seismological attempts to confirm Claerbout's conjecture were inconclusive (Baskir & Weller 1975; Cole 1995), independent work in helioseismology (Duvall *et al*. 1993) and ultrasonics (Lobkis & Weaver 2001; Weaver & Lobkis 2001) demonstrated its validity. The work of Lobkis & Weaver (2001) on diffuse ultrasonic fields, for example, established from both theoretical and experimental approaches that the Green's function (i.e. the response to an impulsive source) between two receivers could be recovered by temporal cross-correlation of those records. This concept, now known as seismic interferometry, has been developed by a number of authors, including Schuster (2001), Snieder *et al*. (2002), Wapenaar (2003, 2004), Bakulin & Calvert (2004), Schuster *et al*. (2004), Snieder (2004) and Wapenaar & Fokkema (2006).

Many seismic interferometry studies to date have concentrated on the ambient seismic field generated by natural processes. This natural field has been shown to be generated by different mechanisms at different frequency bands. For example, one widely used source of natural energy are microseisms (surface waves) observed between 0.04 and 0.08 Hz, which have been linked to ocean swells at coastal regions (Haubrich *et al*. 1963). The random pressure fluctuations generated by ocean waves have been shown sufficient to generate microseisms of considerable amplitude (Hasselmann 1963). Longuet-Higgins (1950) established that a secondary type of microseismic energy, with dominant peak frequencies between 0.1 and 0.16 Hz, results from the nonlinear interaction of pairs of ocean swell trains. The low-frequency (0.001–0.01 Hz) part of the Earth's background free oscillations, the 'hum', is the least understood. Some have attributed this phenomenon to atmospheric turbulence (Nishida *et al*. 2000; Tanimoto 2001), though recent work indicates that the 'hum' results from the mechanisms responsible for the primary and/or secondary microseism (Webb 2007, 2008; Traer & Gerstoft 2014; Ardhuin *et al*. 2015).

The ambient seismic field from natural sources is well known to be dominated by fundamental mode surface waves below frequencies of 0.1 Hz (Haubrich *et al*. 1963; Toksöz & Lacoss 1968; Ekström 2001). As a result ambient noise tomography using surface waves has developed into a robust imaging technique, allowing seismologists to image the Earth at local (Brenguier *et al*. 2007; Lin *et al*. 2013a), regional (Shapiro *et al*. 2005; Kang & Shin 2006; Yao *et al*. 2006; Yang *et al*. 2007) and global scales (Nishida *et al*.

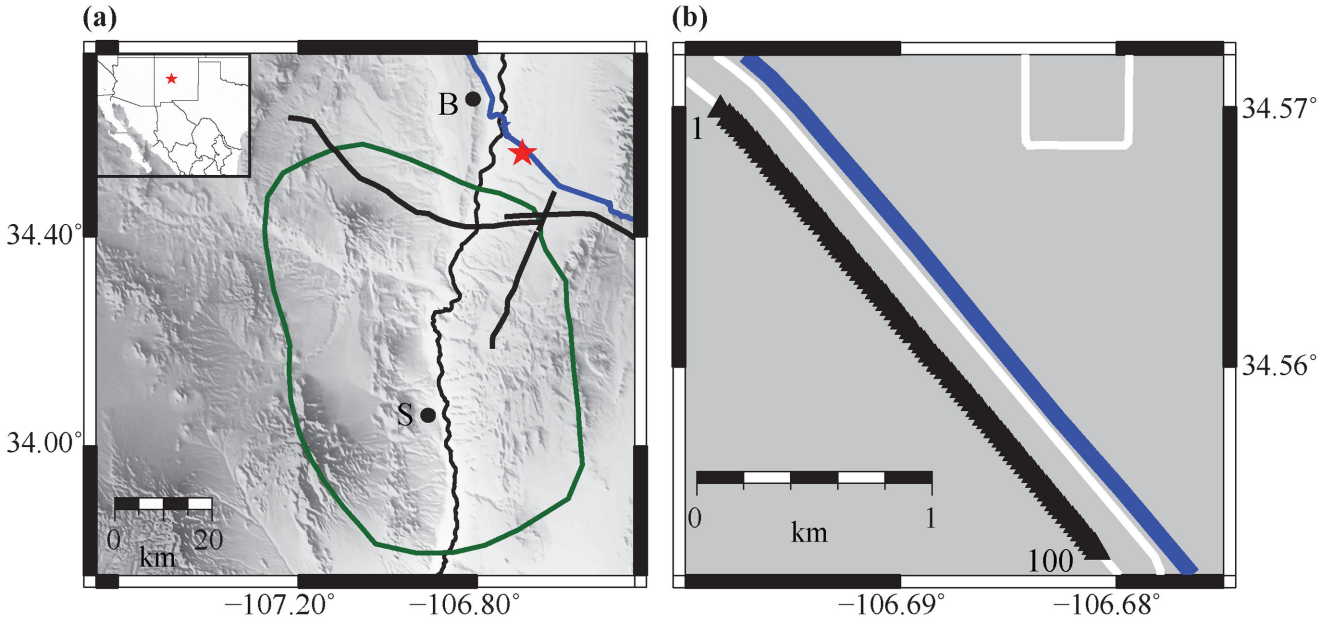

Figure 1. (a) Survey location (red star) relative to the towns of Belen (B) and Socorro (S). Also shown are the Rio Grande river (meandering north–south black line), the outline (green line) of the Socorro Magma Body (Balch *et al*. 1997), and the COCORP deep seismic profiles (bold black lines) (Brown *et al*. 1979). Inset map shows survey location (red star) within New Mexico. (b) Distribution of seismographs (black triangles) paralleling the BNSF railway (blue line) with local gravel roads shown in white.

2009). At frequencies above 0.1 Hz the ambient seismic field is a complicated mixture of fundamental mode surface waves, higher mode surface waves and body waves (Bonnefoy-Claudet *et al*. 2006; Koper *et al*. 2010). The extraction of body waves from the ambient seismic field generated by natural processes has proven to be more difficult than for surface waves, although various degrees of success have been reported (e.g. Roux *et al*. 2005; Draganov *et al*. 2007, 2009; Zhan *et al*. 2010; Poli *et al*. 2011; Ruigrok *et al*. 2011; Ryberg 2011; Lin *et al*. 2013b).

The anthropogenic contribution to the ambient seismic field has been addressed by only a few interferometry studies. Thus far, Matsuoka *et al*. (2006) and Shiraishi *et al*. (2006), to our knowledge, are the first attempts to image geological structure using seismic interferometry with vehicles as sources. Halliday *et al*. (2008) also recovered surface waves from vehicles. Miyazawa *et al*. (2008) observed vertical *P*- and *S*-wave propagation in a monitoring well from vibrations caused by mechanical equipment operating around the top of the well. Nakata *et al*. (2011) retrieved shear wave (i.e. refracted and reflected) and Love-wave energy from recordings of railways and highways. Behm & Snieder (2013) and Behm *et al*. (2014) extracted surface waves from highway records. Imaging with drilling noise in oil exploration is perhaps the most mature example of the use of anthropogenic 'noise' in this context (Rector & Marion 1991; Poletto & Miranda 2004).

The work presented here reports our attempt to use the ambient seismic field generated by trains to produce virtual records of surface and body waves for imaging the subsurface.

2 THE BELEN EXPERIMENT

In 2014 May, 100 vertical component geophones (Geospace GS-11D 4.5 Hz) coupled with Ref Tek 125A recorders (Texans) were deployed parallel to a busy section of the Burlington Northern Santa Fe (BNSF) railway near Belen, New Mexico (Fig. 1). The array used a spacing of approximately 25 m between seismographs for a total array length of 2475 m. The Texans acquired data continuously for 120 hr between 2014 May 29 and June 3. We had excellent data recovery, with only one station failing completely. A sampling rate of 0.004 s (250 Hz) was used, corresponding to a Nyquist frequency of 125 Hz.

The site was chosen for several reasons: (1) it has one of the largest volumes of train traffic in the southern transcontinental corridor, approximately 90 trains per day with some being 3 km long (Frailey 2007), (2) it is in close proximity to the IRIS-PASSCAL instrument centre which provided the instrumentation for the deployment and (3) it is located within the Rio Grande rift, an area which has been subject to many previous geophysical studies (Sanford & Long 1965; Schmucker 1970; Decker & Smithson 1975; Ramberg & Smithson 1975), including the COCORP deep seismic reflection surveys in 1976 and 1977 (Brown *et al*. 1979).

2.1 Inspection of recordings

The continuous data set of 120 hr was divided into hour-long sections and filtered to remove the non-zero mean value (DC component) of the waveforms inherent in the Texan recorders. While inspecting the raw records we recognized two major types of signals: trains and road vehicles (cars and trucks). Trains were identified based on the temporal length of the vibrations at each station and the velocity of the signal as it traversed the array, while vehicle signatures on the two main gravel roads parallel to the railroad are identified from field notes, relative amplitudes and temporal length of the vibrations. A third type of signal, bursts of 'noise' of much lower relative amplitude to trains and vehicles, we associate with wind gusts. The records contained 231 clear train signals traveling westbound and 199 clear train signals traveling eastbound, along with 59 vehicles and 94 bursts of noise. These represent a total of 583 specific 'events' averaging 116.6 per day. The time interval between trains ranged from a few seconds for trains traveling in

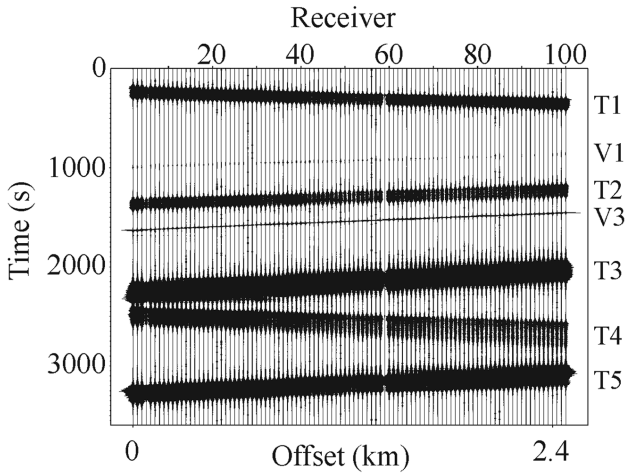

Figure 2. Raw seismograms of the 17th hour of recording. Trains (T) and vehicles (V) are visible on the record. Events that dip down to the right are moving towards the southeast (e.g. T1), events that dip down to the left are moving towards the northwest (e.g. T5). Vehicle 1 is moving on the gravel road closer to the railway, vehicle 2 is moving along the gravel road next to the array (see Fig. 1).

opposite directions (the railroad consisted of two parallel tracks) to over an hour between trains. Fig. 2 shows an example of an hour-long record that contains signals that correspond to trains and other vehicles.

2.2 Spectral character of recordings

The train signals T1, T2 and T5 in Fig. 2 were investigated using a master trace (top row Fig. 3) derived from summing all traces after aligning the train signals using a time-shift (i.e. linear moveout correction) obtained from the average velocity of the train. Spectrograms and amplitude spectra for these master traces are shown in Fig. 3. The spectra show substantial energy from about 4.5 Hz (the geophone natural frequency) up to about 50 Hz, with discernible energy as high as 110 Hz. Peak amplitudes occur between 5 and 20 Hz. Note that a Doppler effect is evident in the spectrograms, with higher frequency content present at earlier times (i.e. approaching train) and the absence of those higher frequencies at later times (i.e. receding train).

3 SEISMIC INTERFEROMETRY

Initial interferometric processing consisted of treating all hour-long records equally, without making a distinction among different signals (i.e. trains, cars, wind noise, quiet periods). Several frequency bandpass filters and amplitude-normalization routines were tested for pre-processing of the records (details below). After pre-processing, the positive lags of the cross-correlation are computed in the time domain for each hour-long record. The procedure selects a station to cross-correlate against the remaining stations, resulting in 120 cross-correlation records for the selected station. These cross-correlation panels are then normalized by the root mean square

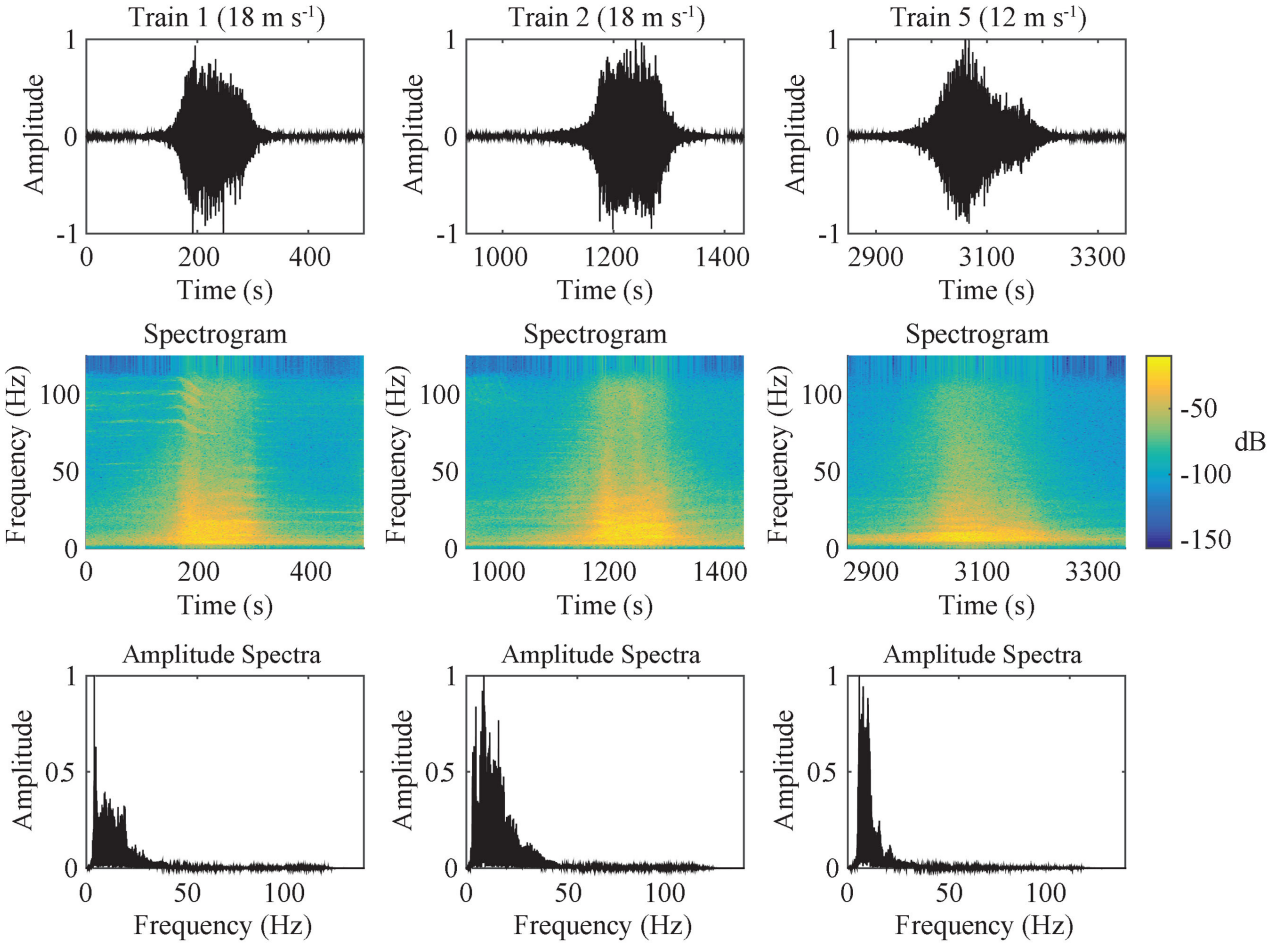

Figure 3. Fourier analysis of the train records shown in Fig. 2. Top row: master trace for train signals T1, T2 and T5, obtained by time-shifting the signals using the average velocity of the train and summing all traces. Middle row: spectrogram of master trace. Bottom row: amplitude spectra of master trace.

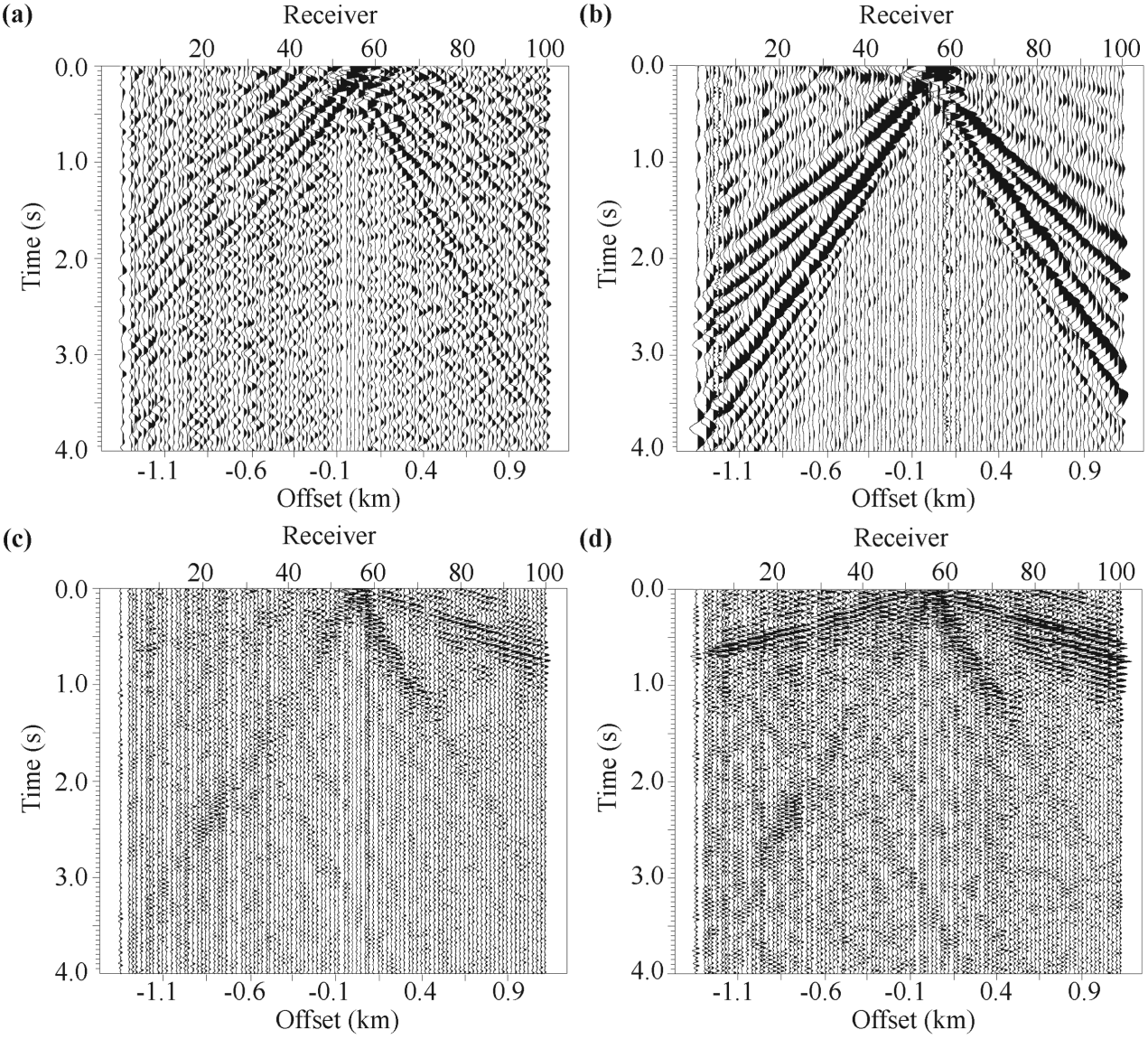

Figure 4. Virtual-shot gathers generated by cross-correlation, rms scaling to equalize the amplitude among all panels and stacking. Various pre-processing steps were applied. Only positive correlation lags are shown. Different panels illustrate the results of pre-processing with (a) rms amplitude scaling with a window of 1 hr, then rms panel-to-panel normalization (b) one-bit amplitude normalization, (c) rms amplitude scaling with a window of 50 s, then rms panel-to-panel normalization and bandpass filtering (16–100 Hz), (d) bandpass filtering (16–100 Hz) and one-bit amplitude normalization. The virtual source in each case is at receiver 55.

(rms) value of the amplitude of the panel, to equalize the amplitude from panel to panel, and then summed (stacked). This is then repeated for all stations, resulting in the computation of the negative lags of the cross-correlation (note that computing the positive lags of the cross-correlation of station A with station B and of station B with station A results in the complete cross-correlation function). The positive lags of the correlation are what we here refer to as a virtual-shot gather, while the negative lags are what we here refer to as a virtual-receiver gather. The sum of both the positive and negative lags for each station results in what we here refer to as a virtual-source gather.

Fig. 4 shows several versions of a typical virtual-shot gather (positive lags only) corresponding to the selection of receiver 55 as the source. Each panel varies in terms of the pre-processing steps applied. Fig. 4(a) results from applying an rms normalization with a 1 hr window to each trace and an rms panel-to-panel normalization prior to the correlation step. The result shows relatively weak energy propagating linearly with apparent velocities between 300 and 800 m s^{-1}, velocities that suggest surface wave energy. One-bit amplitude normalization (e.g. sign-bit; Bond & Cahn 1958) applied to the raw recordings prior to correlation results in the recovery of much stronger linear, and quasi-linear arrivals (Fig. 4b). Again, based on their apparent velocities, we interpret these as Rayleigh waves. Moreover, as a range of slopes (apparent velocity) is evident for this energy, we infer that it is dispersive. As we will show later in this paper, these arrivals are a robust basis for velocity inversion using surface wave techniques. Because normalizing with respect to the entire trace (1 hr window) runs the risk of anomalous scaling by strong amplitude signals (e.g. noise bursts or spikes), various moving windows were also tested to normalize the amplitude prior to cross-correlation. Those tests mimic the results obtained in Fig. 4(a) for very large windows (30 min),

and those obtained in Fig. 4(b) for very small windows (20 s). In order to suppress the surface wave contributions and enhance body wave recovery, an rms amplitude scaling using a window of 12 500 samples (i.e. 50 s windows) and a 16–100 Hz bandpass filter were applied prior to computing the correlations. The resulting virtual-shot gather (Fig. 4c) clearly exhibits linear arrivals with much higher apparent velocities for at least the positive offsets. Application of a 16–100 Hz filter followed by one-bit normalization prior to correlation (Fig. 4d) recovered the same linear events observed in Fig. 4(c), propagating in the positive and negative offset directions with much stronger amplitudes. These events have much higher apparent velocities than those of the surface waves observed in Figs 4(a) and (b), and as argued below, represent *P* waves. Clearly one-bit amplitude normalization is an effective tool for enhancing virtual energy associated with both surface and body waves when applied in conjunction with the appropriate bandpass filter. As noted by Cupillard *et al*. (2011), this method retrieves any coherent energy present between two receivers, in the case of Fig. 4(b) the most abundant energy corresponds to surface waves, while in Fig. 4(d), one-bit pre-processing is retrieving coherent energy present at higher frequencies.

Inspection of the positive and negative lags of the cross-correlations shows that for the retrieved surface (Fig. 5a) and body (Fig. 5b) waves the illumination can be considered close to symmetric, implying that railroad vibrations illuminate the array in the same manner for either travel direction of the trains. This is

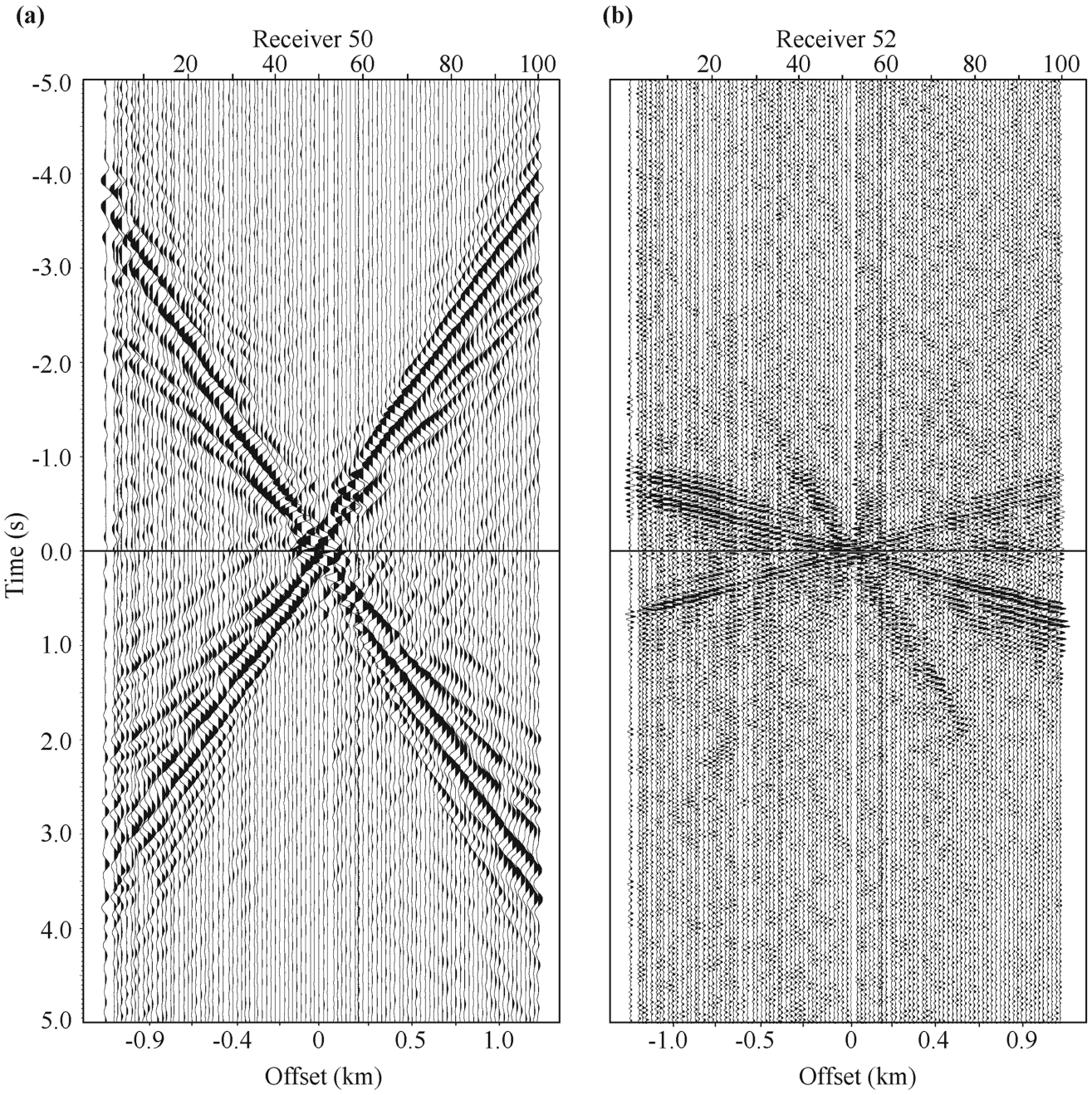

Figure 5. Positive and negative lags of the cross-correlation for (a) the source at station 50, with one-bit normalization pre-processing, and for (b) the source at station 52, with bandpass (16–100 Hz) and one-bit normalization pre-processing.

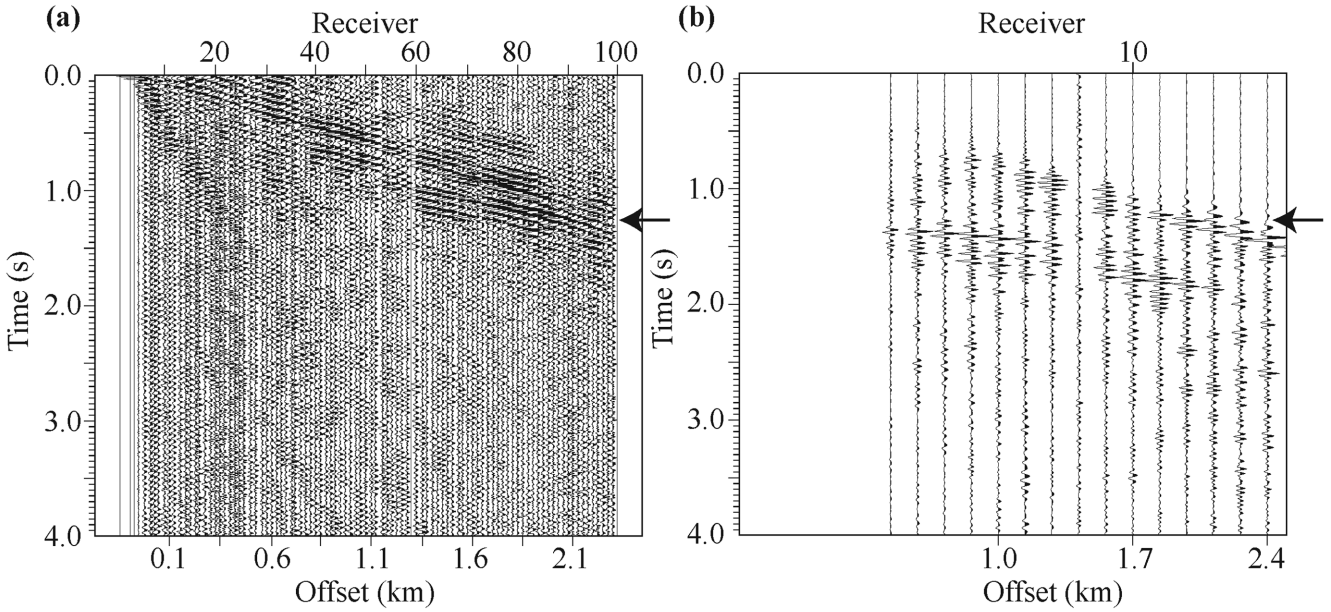

Figure 6. (a) Virtual-shot gather (positive lags) at receiver 5. (b) COCORP line 1-A real shot gather (vibration point 105) presumed to sample similar geology to that beneath the Belen array. Only 15 traces of the COCORP data are shown to match the aperture of the Belen railroad array (*ca*. 2.4 km). The apparent velocity of the first arrivals (arrows) in both cases is ∼2100 m s^{-1}. The curved second arrival on the COCORP data is a reflection at ∼1200 m depth. Hyperbolic events at large offsets on the virtual-shot gathers (e.g. Fig. 6a) are here interpreted as reflected *P* waves as well (see Fig. 7).

to be expected since each train 'broadcasts' in both directions to the array as it approaches and recedes (see also f–k analysis below).

Comparison of our virtual-shot gathers (e.g. Fig. 4d) with physical shot gathers from the COCORP survey collected within similar Tertiary-Quaternary sedimentary rocks of the Rio Grande valley shows that the high velocities (i.e. ∼2100 m s^{-1}) associated with the linear first arrivals recovered from interferometry correspond closely with the velocity of the real *P* waves traveling in the near-surface sedimentary rocks (Fig. 6).

In addition to what we interpret as clear direct *P* waves on the virtual gathers, there are later arrivals that exhibit curvature suggestive of the hyperbolic moveout expected of reflections. The hyperbolic events, most obvious in virtual gathers at larger offsets, correspond to a reflector at approximately 0.72 s vertical two-way time (twt) or ∼750 m below the surface, and a normal moveout (NMO) velocity of approximately 2075 m s^{-1} (Fig. 7). The reflected arrival is not fully aligned at large offsets by the standard NMO correction (e.g. Fig. 7c). This may be due to the presence of a shallow weathering layer and/or by use of an NMO correction where the fourth and higher order terms are neglected (Yilmaz 2001). Frequency distortion (NMO stretching) is common at large offsets and shallow arrival times (Yilmaz 2001). Usually data that suffer NMO stretching are muted prior to processing. However, since the energy of interest lies completely within the 'stretch' regime, an alternative approach is taken here. Fig. 7(e) shows the virtual-shot gather in Fig. 6(a) after applying a constant normal moveout correction (Shatilo & Aminzadeh 2000), which assumes that the seismic wavelet is primarily from a single interface at depth. The traveltimes of this particular reflected phase agree with the two-way times (0.6–0.9 s) observed by Russell & Snelson (1994) for the contact between upper Tertiary sedimentary rocks (Santa Fe group) and Palaeozoic sedimentary rocks 10 km south of our array.

4 FREQUENCY–WAVENUMBER ANALYSIS OF TRAIN SOURCES

While body and surface waves can be retrieved with relative simple pre-processing, up to this point it is unclear which components of the recordings (e.g. background 'noise', approaching versus receding train, etc.) contribute to the emergence of the different seismic phases observed. To address this issue we carried systematic frequency–wavenumber (f–k) analyses of the raw recordings using short time windows (i.e. 100 s) on several records. Fig. 8 shows three-hour long records of raw data (filtered to remove non-zero mean value and amplitude normalized) that are adjacent in time (hour 12, 13 and 14). Hour 12 contains four train signals (the signal at the beginning of the record corresponds to a train close to the one hour mark on the record for the preceding hour 11). Three windows of data were extracted from hour 12 to represent the approaching, passing and receding part of the fourth train signal. Hour 13 is the only 60 min record in the entire data set that does not appear to contain any train signals or vehicles. The first window in this record was chosen to start after an hour had passed since the last train went by the array. Hour 14 contains three train signals. The first window extracted from this record represents a 'quiet' period; the second and third windows correspond to the approaching and receding parts of the train signal in the record. Each window (red rectangle) in Fig. 8 was transformed into the f–k domain. The convention used in the f–k plots shown here (Fig. 9) is that events that dip down to the right (i.e. direction of travel is northwest to southeast in Fig. 1b) are assigned negative wavenumbers, while events that dip up to the right (i.e. direction of travel is southeast to northwest in Fig. 1b) are assigned positive wavenumbers.

The f–k transforms in Figs 9(a)–(c) correspond to the three windows selected in hour 12 (Fig. 8a), from earliest to last. The approaching train signal in Fig. 8(a) (earliest window) has a negative wavenumber according to the convention used here, corresponding to travel direction from northwest to southeast. Fig. 9(a) shows

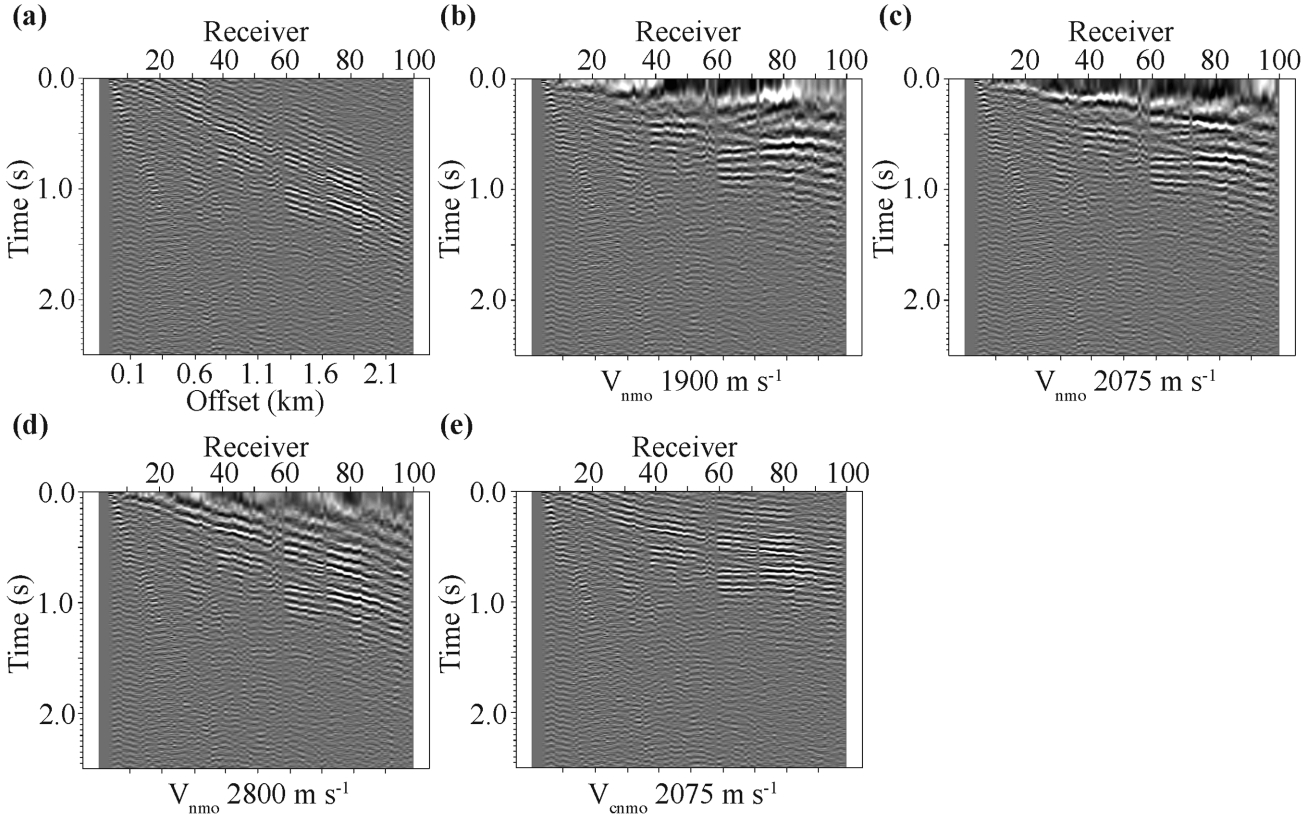

Figure 7. Virtual-shot gather (positive lags) for receiver 5 with (a) no NMO correction, (b) $V_{\rm nmo} = 1900$ m s^{-1}, (c) $V_{\rm nmo} = 2075$ m s^{-1} and (d) $V_{\rm nmo} = 2800$ m s^{-1}. Because of the severe NMO stretch at large offsets for shallow events we have used the constant normal moveout (Shatilo & Aminzadeh 2000) correction for $V = 2075$ m s^{-1} in (e).

strong-amplitude signals with negative wavenumbers. The linear energy bands in the f–k plots represent surface and body waves, with the latter corresponding to the steeper slopes (high apparent velocities). Two distinct bands of 'slow' energy are apparent in Fig. 9(a), which we attribute to the fundamental mode ($V \sim 400$ m s^{-1}), spatially aliased at about 8 Hz, and a higher mode ($V \sim 600$ m s^{-1}), spatially aliased at about 11 Hz, of surface wave propagation.

Fig. 9(b) suggests that most of the energy arrives with infinite (e.g. horizontal) apparent velocity, consistent with the train being largely parallel to and spanning the array. In this case, the seismic energy is largely arriving broadside to the array.

Fig. 9(c) shows the train signal as it is receding. As expected, the energy propagates in the direction opposite to that when the train was approaching (e.g. Fig. 9a). Similarly to Fig. 9(a), Fig. 9(c) clearly shows events whose slopes suggest both body and surface waves. Here, however, at least three distinct bands of surface wave energy are evident, with apparent velocities of 400, 600 and 750 m s^{-1}. We again interpret these as fundamental and higher mode surface waves. In Figs 9(a)–(c) there are also horizontal bands of energy (e.g. at 12, 18, 24 and 30 Hz). Such monotonic noise can be a serious contaminant in cross-correlations, especially if this noise is stationary in time and space, effectively 'drowning out' recovery of propagating seismic energy (Draganov *et al*. 2009). This does not seem to be a problem with this data set, possibly due to these energy bands being unstable in time.

Figs 9(d)–(f) correspond to the time windows in Fig. 8(b) where much less, if any, train energy should be present. These plots therefore characterize the 'background' seismic energy upon which the train energy is presumably superimposed. All three of the background f–k plots show surface wave energy propagating somewhat more symmetrically than when the train energy is dominant. Slopes corresponding to body wave energy are also present, albeit at somewhat weaker levels and dominantly traveling from northwest to southeast. We speculate that the surface wave energy may be due to cultural noise from the town of Belen (Fig. 1), or perhaps highway traffic noise from the several highways which pass near Belen with a north–south orientation. Whereas the body wave energy is perhaps excited by a distant train traveling from northwest to southeast, towards the railroad station in the town of Belen.

Figs 9(g)–(i) correspond to the three windows in hour 14 (Fig. 8c). The f–k plot in Fig. 9(g) examines a 'quiet' period within hour 14. Although hints of surface wave energy are present but no body wave energy is visible. Note that background amplitudes for this window are even lower than the ones in record 13. Figs 9(h) and (i) represent the approaching and receding train signal, respectively. The reversals in the slopes of body and surface wave energy correspond to the signal directivity.

The direct wave that is so clear in the virtual-shot gathers (e.g. Figs 4c and d) is also prominent in all of the f–k spectra with a velocity ($V \sim 2.1$ km s^{-1}) that matches that derived from the slope of the virtual traveltime curves. Like the surface waves, it appears strongest in the direction toward the approaching or receding train source. However, we also note that in some of the spectra there appears to be two linear arrivals with body wave type velocities. These are most evident in Figs 9(c) and (h), with apparent velocities of approximately 4 km s^{-1}. We interpret these (e.g. $P_{\rm r}$ in Fig. 9h) as corresponding to critically refracted, or head wave, energy from

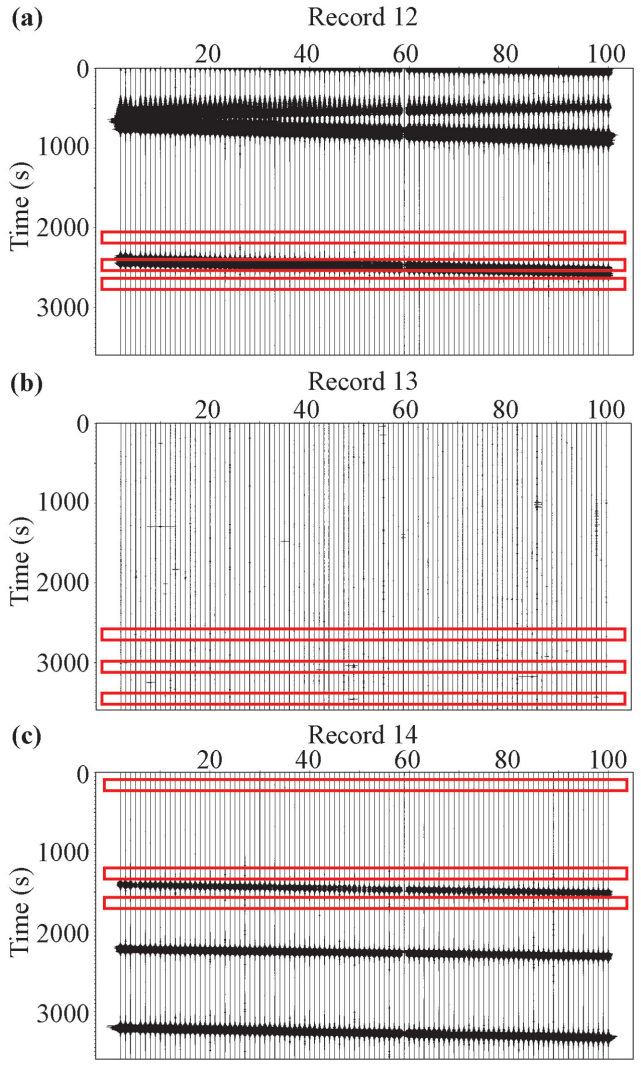

Figure 8. Three-hour-long records normalized by the max amplitude in record 12, each with three 100 s windows (red rectangles). The records are the (a) 12th, (b) 13th and (c) 14th hour of recording. Record 13 is the only hour in the entire data set that did not appear to contain any train or vehicle signals.

deeper subsurface discontinuities. Corresponding linear arrivals are not evident in any of the virtual gathers computed by the various pre-processing sequences (e.g. Fig. 4), implying that such energy is fairly weak compared to the direct *P* energy, which is often the case for head waves (e.g. Grant & West 1965). Mikesell *et al*. (2009) describe the interferometric treatment of these refracted arrivals, and introduce the term *virtual refraction* to refer to the spurious energy observed in virtual gathers due to these phases. Interestingly, these faster linear arrivals are also evident in the f–k spectra for the background noise (e.g. Figs 9e and f) as well as those for the train sources.

The f–k analyses in Fig. 9 indicate that the most useful energy from the train sources occur when the trains are either approaching or leaving the array. When the train is 'on the array', energy arriving 'broadside' is clearly swamping any useful surface or body wave information. However, this energy could be useful for instruments placed perpendicular to our array.

While applying the same f–k analysis to signals such as vehicles and sources other than trains (i.e. wind gusts) surface waves are clearly observed but body waves do not appear to be present (Fig. 10). This suggests that these sources are not as efficient as trains, at least in generating body wave energy.

4.1 Seismic interferometry guided by f–k analysis

To determine if body waves might be more effectively extracted if the 'broadside' train, vehicle and wind energy were removed from the correlations, the results shown in Fig. 11 were computed. Fig. 11(a) shows a brute correlation using all the energy (as in Fig. 4d). Fig. 11(b) shows the result of pre-processing with a 50 s rms window for amplitude normalization, panel-to-panel equalization, a bandpass filter (16–100) and muting (i.e. deleting) of signals from vehicles, wind, and broadside train energy. Similarly, Fig. 11(c) shows the result of the same pre-processing with the addition of a 20 Hz notch filter. The notch filter was chosen to eliminate a narrow peak present in the amplitude spectrum of Figs 11(a) and (b). Fig. 11(d) shows the result of pre-processing with a 50 s rms window for amplitude normalization, panel-to-panel equalization, an f–k filter and the same muting used in Figs 11(b) and (c). In spite of our expectation that this muting would be helpful, the resulting virtual-shot gathers (Figs 11b–d) show no significant improvement in terms of recovery of body wave energy. As mentioned earlier, we also tried applying a notch filter prior to correlation and stacking to remove a 20 Hz monotonic energy present in Figs 11(b) and (c). Monotonic 'noise' can frequently obscure propagating seismic energy. However, the removal of this particular frequency does not result in any improvement in the virtual shot gather (Fig. 11c). In general, we see little evidence of the influence of monotonic noise in our correlations, perhaps because the frequencies of such noise are unstable in time (e.g. Fig. 9). In addition, f–k filtering was tested as a means to remove surface wave energy prior to computing the correlations without sacrificing any low-frequency body wave energy that might be present. However, the result (Fig. 11d) fails to show any significant new body wave arrivals.

5 REFLECTION IMAGING

The obvious next step in using the virtual-source gathers (i.e. sum of positive and negative correlation lags) is to apply common midpoint (CMP) processing to generate a stacked seismic section. The processing sequence included linear moveout velocity picks to guide automated first arrival picks, editing of first arrival picks, mute function for first arrivals, elevation statics, amplitude scaling, surface consistent deconvolution, constant NMO correction (Shatilo & Aminzadeh 2000), CMP stacking and amplitude scaling.

The resulting stacked section (Fig. 12) shows coherent events, but only down to 0.8 s twt. The most prominent event is the reflector previously identified in Fig. 7 at about 0.72 s twt. We associate this reflector with the contact between Tertiary and Palaeozoic sedimentary rocks based on the interpretation of conventional seismic data nearby (e.g. Russell & Snelson 1994). Although shallow reflectivity is apparently recovered, it is far from strong and deeper events are not observed. It is also apparent from inspection of the virtual-shot gathers that the energy that is contributing to the reflection stack is coming at 'wide' versus 'near-vertical' angles. Given the limited aperture of our spread (*ca*. 2.5 km), only the shallow reflectors correspond to illumination at relatively wide angles.

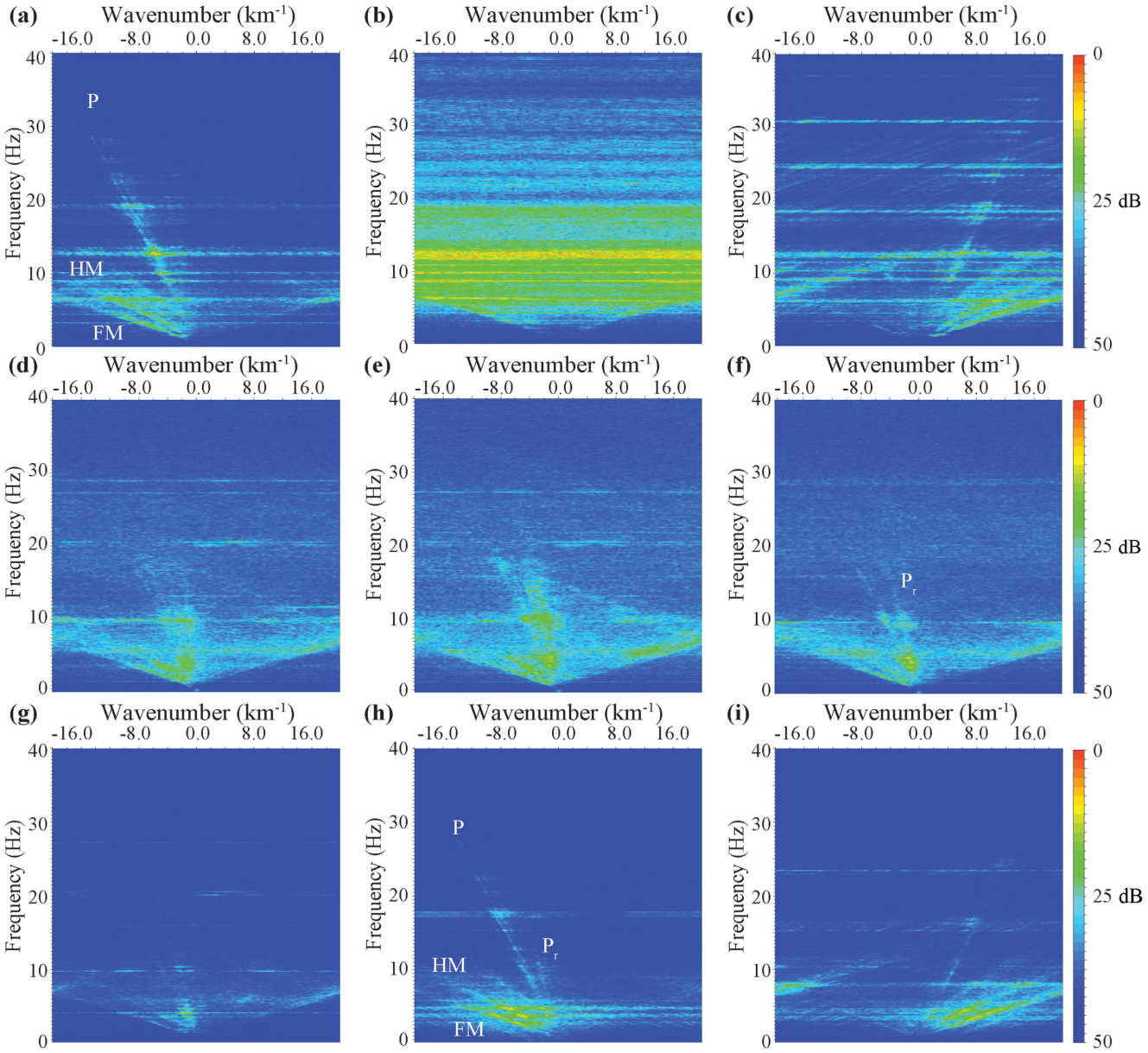

Figure 9. Frequency–wavenumber analysis of Fig. 8. f–k analysis of window (a) 2150–2250 s, (b) 2430–2530 s, (c) 2620–2720 s, in record 12. f–k analysis of window (d) 2600–2700 s, (e) 3000–3100 s, (f) 3400–3500 s, in record 13. f–k analysis of window (g) 100–200 s, (h) 1200–1300 s, (i) 1550–1650 s, in record 14. P, FM and HM mark the direct *P*-wave, and the fundamental and higher mode surface waves, respectively. P_r marks linear arrivals which we interpret as energy (head waves) refracted from deeper interfaces. See the text for discussion.

Draganov *et al*. (2006) and Forghani & Snieder (2010) make clear the importance of the source distribution in seismic interferometry. Ideally for near-vertical reflection imaging using interferometry, sources should be distributed beneath the recording array (Draganov *et al*. 2006; Ruigrok *et al*. 2010). In contrast, the train source is restricted to the surface with most of its energy propagating horizontally across our array.

The studies of Zhan *et al*. (2010) and Poli *et al*. (2011) which report deep reflections (e.g. Moho) suggest that relevant illumination occurs near-horizontally, with the retrieved reflections observed at wide angles. Zhan *et al*. (2010) reason that sources must be distributed within the crust (at the free surface and at depth), but that the crust behaves as a waveguide for near-horizontally traveling body wave energy (i.e. SmS^n).

The lack of deeper reflections recovered by our analysis could also be due in part to the very short length of our recording period, 5 d. Zhan *et al*. (2010) and Poli *et al*. (2011) reported their wide-angle Moho reflections based on 6 months and 1 yr of data, respectively. Although Ruigrok *et al*. (2011) report Moho reflections from approximately 40 hr of data collected over the Northeast Abu Gharadig Basin in Egypt, they attribute their success to having near-surface noise sources sufficiently far away that the relevant energy is actually arriving near vertically from beneath the array due to the velocity gradient in the crust and mantle.

Based on these observations as well as our own analysis, we suggest that deeper reflection imaging with train sources will require a much larger virtual source to receiver offsets and longer recording time.

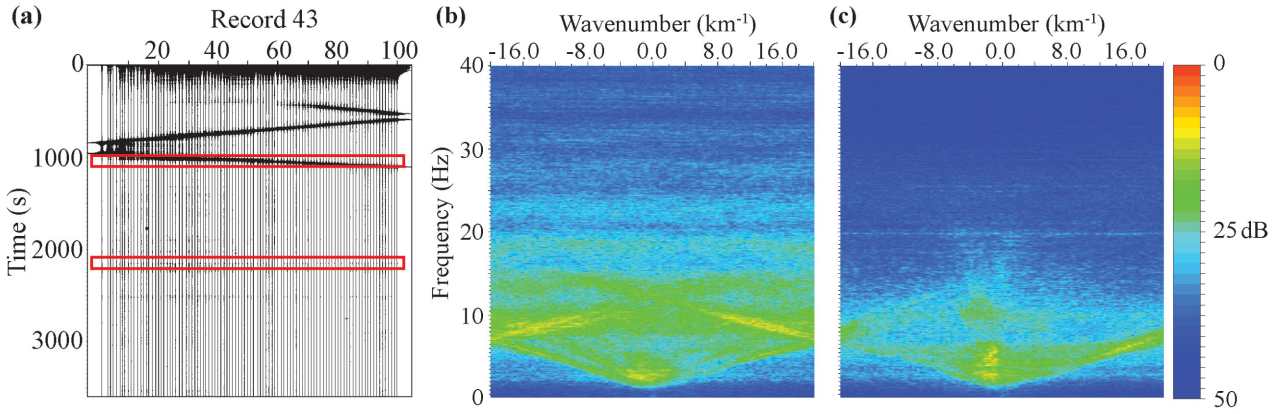

Figure 10. Frequency–wavenumber analysis of strong sources other than trains. (a) Seismograms of 43rd hour showing part of a train signal at the start of the record, followed by three car signals and what is interpreted as wind gusts. f–k analysis of window (b) 990–1090 s containing a car signal, and of window (c) 2100–2200 s containing a wind gust.

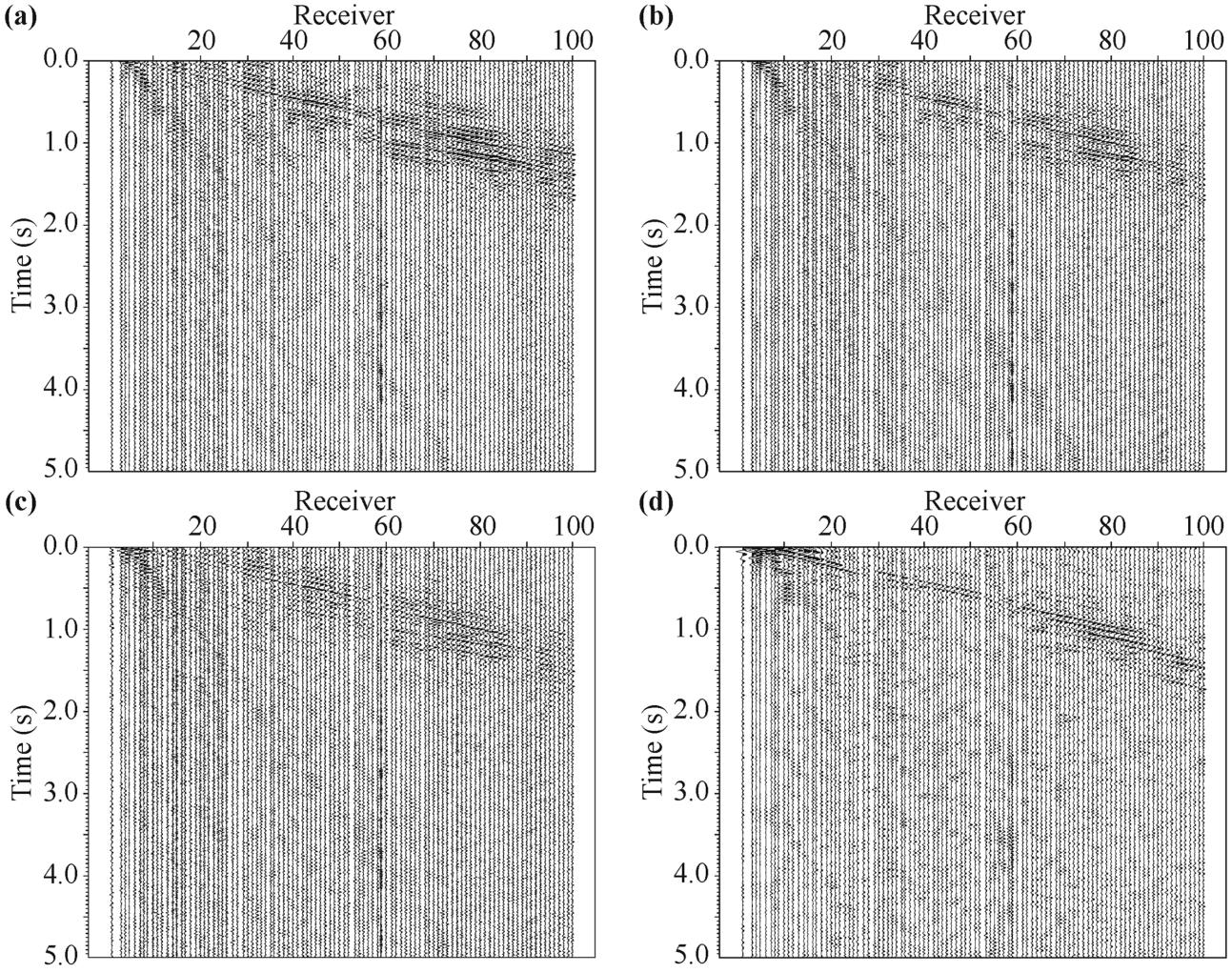

Figure 11. Comparison of several pre-processing routines guided by the f–k analysis. (a) Brute correlation using the same parameters as in Fig. 4(d). (b) rms amplitude scaling with a window of 50 s, then rms panel-to-panel normalization, followed by a bandpass filtering (16–100 Hz) and muting of vehicles, wind gusts and broadside train energy prior to correlation. (c) Same pre-processing as in (b) with an added notch filter to remove a monotonic 20 Hz signal (see the text for discussion). (d) rms amplitude scaling with a window of 50 s, then rms panel-to-panel normalization, followed by an f–k filter to remove surface wave energy before correlation.

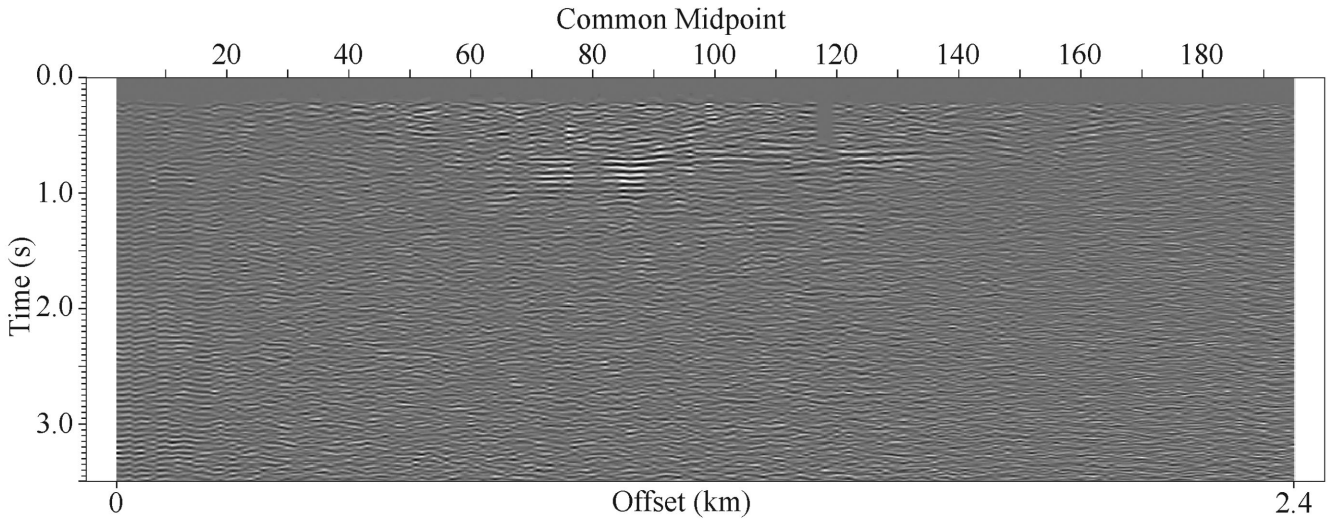

Figure 12. Common midpoint stack of virtual CMP gathers. The coherent event at 0.72 s twt (*ca*. 750 m below the surface) is interpreted as the base of the Tertiary sedimentary section.

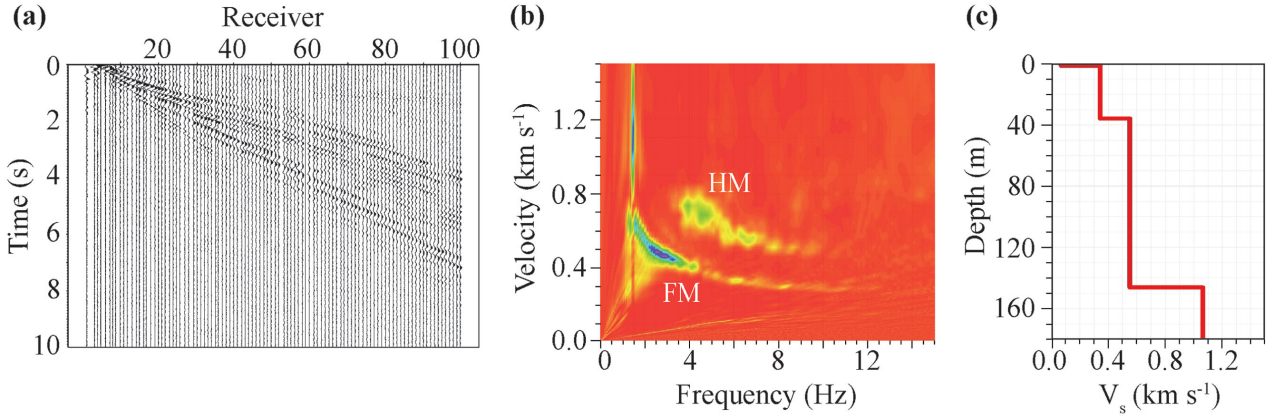

Figure 13. (a) Virtual-shot gather with station 5 as the source. (b) Dispersion curve for the virtual-shot gather at receiver 5, with the fundamental mode (FM) and a higher mode (HM) labeled. (c) 1-D shear wave velocity model obtained from the inversion in the frequency range 0–8 Hz of the dispersion curve from the fundamental mode.

6 SURFACE WAVES

It is abundantly clear from Figs 4 and 9 that trains are an exceptionally good source of surface wave energy, with substantial potential for mapping subsurface structure via inversion of their dispersion curves. Moreover, the surface wave energy provided by trains is richer in high frequencies (1–14 Hz) than those from microseism-based interferometry at both crustal (Bensen *et al*. 2007) and shallow (Lin *et al*. 2013a) scale frequencies (i.e. crustal 0.05–2 Hz, shallow 0.5–4 Hz).

Here, by way of demonstration, we computed a simple 1-D shear wave velocity model by inversion of the dispersion curve for a particular virtual-shot gather. Fig. 13(a) shows the virtual-shot gather for station 5 generated via one-bit normalization pre-processing, together with its corresponding dispersion curve (Fig. 13b). The dispersion curve clearly shows both fundamental and higher mode Rayleigh wave energy. Inversion of the fundamental mode (Wathelet *et al*. 2004; Wathelet 2008) in the frequency range 0–8 Hz (unaliased surface waves) resulted in the 1-D shear wave velocity model shown in Fig. 13(c). We did not attempt inversion of the higher mode energy as it may have been inaccurately recovered (Halliday & Curtis 2008; Kimman & Trampert 2010). Note that direct correlation of the surface wave velocity structure with the body wave reflectivity is of limited valued, since the surface waves probe only the very shallowest portion (e.g. less than 200 m, or 0.2 s twt) of the subsurface, a portion that is non-reflective in the seismic section in Fig. 12. However, the *P*-wave velocity of ca 2.0 km s^{-1} derived from the virtual shot gather is consistent with the *S*-wave velocities indicated by the surface wave inversion (*ca*. 1.0 km s^{-1}).

Our surface wave results are comparable to those reported by Nakata *et al*. (2011) also using highway and train recordings. These suggest that train energy represents a promising source for mapping shallow shear wave velocity variations at shallow depths, with possible application to geotechnical issues, ranging from hydrology to seismic site response.

7 CONCLUSIONS

Virtual gathers obtained from cross-correlation of railroad traffic exhibit robust dispersive surface wave (Rayleigh) energy at frequencies higher than those obtained from conventional interferometry

using microseism energy. As such, data generated by train traffic should prove valuable for a range of near-surface applications of surface wave tomography. Potentially useful body wave (*P*) energy can also be recovered for refraction/reflection imaging of the subsurface. Although the reflection image produced in this case only reaches to 1 km depth, in principle greater depths and fidelity should be attained by simply recording/stacking for longer periods of time and/or using more favourable recording geometries (e.g. larger apertures).

The results presented here and from previous studies suggest that careful consideration should be given to using the 'noise' from railroads as an effective source for surface wave and reflection imaging of at least the shallow subsurface over large areas at relatively low cost.

ACKNOWLEDGEMENTS

The authors wish to thank Bridget O'Neill, Noël Barstow, Steve Azevedo, Patricia Griego, George Slad, Pnina Miller and Bruce Beaudoin from IRIS-PASSCAL, and Steve Bernsen, Dave Parmalee, Dylan Rose-Coss, Kyle Murray and Susan Bilek from New Mexico Tech, for their collaboration with the experiment, Schlumberger for providing Vista 3-D seismic analysis software. We are very grateful to one anonymous reviewer, Trond Ryberg and Deyan Draganov whose comments substantially improved the manuscript. The data can be accessed by request through the IRIS DMC repository. The data set BelenRT has the assembled data set ID 14–024. Seismic instruments were provided by the IRIS-PASSCAL instrument centre. Generic Mapping Tools (Wessel *et al*. 2013) and Vista 3-D were used to create the figures presented.

REFERENCES

Ardhuin, F., Gualtieri, L. & Stutzmann, E., 2015. How ocean waves rock the Earth: two mechanisms explain microseisms with periods 3 to 300 s, *Geophys. Res. Lett.,* **42,** 765–772.

Bakulin, A. & Calvert, R., 2004. Virtual source: new method for imaging and 4D below complex overburden, in *74th Annual International Meeting,* SEG, Expanded Abstracts, 2477–2480.

Balch, R.S., Hartse, H.E., Sanford, A.R. & Lin, K., 1997. A new map of the geographic extent of the Socorro mid-crustal magma body, *Bull. seism. Soc. Am.,* **87,** 174–182.

Baskir, E. & Weller, C.E., 1975. Sourceless reflection seismic exploration, *Geophysics,* **40,** 158–159.

Behm, M., Leahy, G.M. & Snieder, R., 2014. Retrieval of local surface wave velocities from traffic noise—an example from the La Barge basin (Wyoming), *Geophys. Prospect.,* **62,** 223–243.

Behm, M. & Snieder, R., 2013. Love waves from local traffic noise interferometry, *Leading Edge,* **32,** 628–632.

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

Bond, F.E. & Cahn, C.R., 1958. On the sampling the zeros of bandwidth limited signals, *IRE Trans. Inform. Theory,* **4,** 110–113.

Bonnefoy-Claudet, S., Cotton, F. & Bard, P.-Y., 2006. The nature of noise wavefield and its applications for site effects studies: a literature review, *Earth-Sci. Rev.,* **79,** 205–227.

Brenguier, F., Shapiro, N.M., Campillo, M., Nercessian, A. & Ferrazzini, V., 2007. 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations, *Geophys. Res. Lett.,* **34,** L02305, doi:10.1029/2006GL028586.

Brown, L.D., Krumhansl, P.A., Chapin, C.E., Sanford, A.R., Cook, F.A., Kaufman, S., Oliver, J.E. & Schilt, F.S., 1979. COCORP seismic reflection studies of the Rio Grande rift, in *Rio Grande Rift: Tectonics Magmatism*, pp. 169–184, ed. Rieker, R.E., American Geophysical Union.

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

Cole, S., 1995. Passive seismic and drill-bit experiments using 2-D arrays, *PhD thesis,* Stanford University, Palo Alto, USA.

Cupillard, P., Stehly, L. & Romanowicz, B., 2011. The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise, *Geophys. J. Int.,* **184,** 1397–1414.

Decker, E.R. & Smithson, S.B., 1975. Heat flow and gravity interpretation across the Rio Grande rift in southern New Mexico and west Texas, *J. geophys. Res.,* **80,** 2542–2552.

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

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

Draganov, D., Wapenaar, K. & Thorbecke, J., 2006. Seismic interferometry: reconstructing the earth's reflection response, *Geophysics,* **71**(4), SI61–SI170.

Duvall, T.L., Jeffferies, S.M., Harvey, J.W. & Pomerantz, M.A., 1993. Time–distance helioseismology, *Nature,* **362,** 430–432.

Ekström, G., 2001. Time domain analysis of Earth's long-period background seismic radiation, *J. geophys. Res.,* **106,** 26 483–26 493.

Forghani, F. & Snieder, R., 2010. Underestimation of body waves and feasibility of surface-wave reconstruction by seismic interferometry, *Leading Edge,* **29,** 790–794.

Frailey, F.W., 2007. Birthplace of the Transcon, *Trains Magazine,* April Issue, Kalmbach Publishing Co.

Grant, F.S. & West, G.F., 1965. *Interpretation Theory in Applied Geophysics,* McGraw-Hill.

Halliday, D. & Curtis, A., 2008. Seismic interferometry, surface waves and source distribution, *Geophys. J. Int.,* **175,** 1067–1087.

Halliday, D., Curtis, A. & Kragh, E., 2008. Seismic surface waves in a suburban environment: active and passive interferometric methods, *Leading Edge,* **27,** 210–218.

Hasselmann, K., 1963. A statistical analysis of the generation of microseisms, *Rev. Geophys.,* **1,** 177–210.

Haubrich, R.A., Munk, W.H. & Snodgrass, F.E., 1963. Comparative spectra of microseisms and swell, *Bull. seism. Soc. Am.,* **53,** 27–37.

Kang, T.-S. & Shin, J.S., 2006. Surface-wave tomography from ambient seismic noise of accelerograph networks in southern Korea, *Geophys. Res. Lett.,* **33,** L17303, doi:10.1029/2006GL027044.

Kimman, W.P. & Trampert, J., 2010. Approximations in seismic interferometry and their effects on surface waves, *Geophys. J. Int.,* **182,** 461–476.

Koper, K.D., Seats, K. & Benz, H., 2010. On the composition of Earth's short-period seismic noise field, *Bull. seism. Soc. Am.,* **100,** 606–617.

Lin, F.-C., Li, D., Clayton, R.W. & Hollis, D., 2013a. High-resolution 3D shallow crustal structure in Long Beach, California: application of ambient noise tomography on a dense seismic array, *Geophysics,* **78,** Q45–Q56.

Lin, F.-C., Tsai, V.C., Schmandt, B., Duputel, Z. & Zhan, Z., 2013b. Extracting seismic core phases with array interferometry, *Geophys. Res. Lett.,* **40,** 1049–1053.

Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *J. acoust. Soc. Am.,* **110,** 3011–3017.

Longuet-Higgins, M.S., 1950. A theory of the origin of microseisms, *Phil. Trans. R. Soc. Lond. Math. Phys. Eng. Sci.,* **243,** 1–35.

Matsuoka, T., Shiraishi, K., Onishi, K. & Aizawa, T., 2006. Application of seismic interferometry to subsurface imaging (1), in *Proceedings of 10th International Symposium on RAEG,* Kyoto University Geophysical Society, pp. 35–38.

Mikesell, D., van Wijk, K., Calvert, A. & Haney, M., 2009. The virtual refraction: useful spurious energy in seismic interferometry, *Geophysics,* **74,** A13–A17.

Miyazawa, M., Snieder, R. & Venkataraman, A., 2008. Application of seismic interferometry to extract *P*- and *S*-wave propagation and observation of shear-wave splitting from noise data at Cold Lake, Alberta, Canada, *Geophysics,* **73,** D35–D40.

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

Nishida, K., Kobayashi, N. & Fukao, Y., 2000. Resonant oscillations between the solid earth and the atmosphere, *Science,* **287,** 2244–2246.

Nishida, K., Montagner, J.-P. & Kawakatsu, H., 2009. Global surface wave tomography using seismic hum, *Science,* **326,** 112–112.

Poletto, F.B. & Miranda, F., 2004. *Seismic While Drilling: Fundamentals of Drill-Bit Seismic Exploration,* Vol. 35, Elsevier.

Poli, P., Pedersen, H.A. & Campillo, M., 2011. Emergence of body waves from cross-correlation of short period seismic noise, *Geophys. J. Int.,* **188,** 549–558.

Ramberg, I.B. & Smithson, S.B., 1975. Gridded fault patterns in a late Cenozoic and a Paleozoic continental rift, *Geology,* **3,** 201–205.

Rector, J.W., III & Marion, B.P., 1991. The use of drill-bit energy as a downhole seismic source, *Geophysics,* **56,** 628–634.

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

Ruigrok, E., Campman, X., Draganov, D. & Wapenaar, K., 2010. High-resolution lithospheric imaging with seismic interferometry, *Geophys. J. Int.,* **183,** 339–357.

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

Russell, L.R. & Snelson, S., 1994. Structure and tectonics of the Albuquerque Basin segment of the Rio Grande rift: insights from reflection seismic data, *Geol. Soc. Am. Spec. Paper,* **291,** 83–112.

Ryberg, T., 2011. Body wave observations from cross-correlations of ambient seismic noise: a case study from the Karoo, RSA, *Geophys. Res. Lett.,* **38,** L13311.

Sanford, A.R. & Long, L.T., 1965. Microearthquake crustal reflections, Socorro, New Mexico, *Bull. seism. Soc. Am.,* **55,** 579–586.

Schmucker, U., 1970. Anomalies of geomagnetic variations in the southwestern United States, *Bulletin of the Scripps Institution of Oceanography,* University of California Press, 13 pp.

Schuster, G., 2001. Theory of daylight/interferometric imaging: tutorial, in *63rd Conference & Technical Exhibition,* European Association of Geoscientists and Engineers, Extended Abstracts, Session A32.

Schuster, G.T., Yu, J., Sheng, J. & Rickett, J., 2004. Interferometric/daylight seismic imaging. *Geophys. J. Int.,* **157,** 838–852.

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

Shatilo, A. & Aminzadeh, F., 2000. Constant normal-moveout (CNMO) correction: a technique and test results, *Geophys. Prospect.,* **48,** 473–488.

Shiraishi, K., Tanaka, M., Onishi, K., Matsuoka, T. & Yamaguchi, S., 2006. Application of seismic interferometry to subsurface imaging (2), in *Proceedings of 10th International Symposium on RAEG,*, Kyoto University Geophysical Society, doi:10.3997/2352-8265.20140076.

Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E.,* **69,** 46610, doi:10.1103/PhysRevE.69.046610.

Snieder, R., Grêt, A., Douma, H. & Scales, J., 2002. Coda wave interferometry for estimating nonlinear behavior in seismic velocity, *Science,* **295,** 2253–2255.

Tanimoto, T., 2001. Continuous free oscillations: atmosphere-solid earth coupling, *Annu. Rev. Earth Planet. Sci.,* **29,** 563–584.

Toksöz, M.N. & Lacoss, R.T., 1968. Microseisms: mode structure and sources, *Science,* **159,** 872–873.

Traer, J. & Gerstoft, P., 2014. A unified theory of microseisms and hum, *J. geophys. Res.,* **119,** 3317–3339.

Wapenaar, K., 2003. Synthesis of an inhomogeneous medium from its acoustic transmission response, *Geophysics,* **68,** 1756–1759.

Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett.,* **93,** 254301, doi:10.1103/PhysRevLett.93.254301.

Wapenaar, K. & Fokkema, J., 2006. Green's function representations for seismic interferometry, *Geophysics,* **71,** SI33–SI46.

Wathelet, M., 2008. An improved neighborhood algorithm: parameter conditions and dynamic scaling, *Geophys. Res. Lett.,* **35,** L09301, doi:10.1029/2008GL033256.

Wathelet, M., Jongmans, D. & Ohrnberger, M., 2004. Surface wave inversion using a direct search algorithm and its application to ambient vibration measurements, *Surf. Geophys.,* **2,** 211–221.

Weaver, R.L. & Lobkis, O.I., 2001. Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies, *Phys. Rev. Lett.,* **87,** 134301.

Webb, S.C., 2007. The Earth's "hum" is driven by ocean waves over the continental shelves, *Nature,* **445,** 754–756.

Webb, S.C., 2008. The Earth's hum: the excitation of Earth normal modes by ocean waves, *Geophys. J. Int.,* **174,** 542–566.

Wessel, P.W., Smith, R., Scharroo, J.L. & Wobbe, F., 2013. Generic mapping tools: improved version released, *EOS, Trans. Am. geophys. Un*., **94,** 409–410.

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

Yao, H., van Der Hilst, R.D. & Maarten, V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—I. Phase velocity maps, *Geophys. J. Int.,* **166,** 732–744.

Yilmaz, Ö., 2001. *Seismic Data Analysis,* Vols 1 and 2, pp. 1–2028, Society of Exploration Geophysicists.

Zhan, Z., Ni, S., Helmberger, D.V. & Clayton, R.W., 2010. Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise, *Geophys. J. Int.,* **182,** 408–420.