

Ground motion prediction of realistic earthquake sources using the ambient seismic field

M. A. Denolle,¹ E. M. Dunham,¹ G. A. Prieto,² and G. C. Beroza¹

Received 4 July 2012; revised 12 November 2012; accepted 11 December 2012.

[1] Predicting accurate ground motion is critical for earthquake hazard analysis, particularly in situations where sedimentary basins trap and amplify seismic waves. We exploit the information carried by the ambient seismic field to extract surface-wave Green's functions between seismic stations and to predict long-period ground motion from earthquakes. To do so, we modify the surface impulse response to correct for the source depth and for the double-couple focal mechanism. These corrections are derived under the assumption that material properties in the immediate vicinity of the source depend only on depth. Using this local 1-D assumption, we solve the surface-wave eigenproblem and compute the fundamental-mode displacement eigenfunctions to express the surface-wave excitation at the source. We validate this technique, which we call the virtual earthquake approach, by comparing computed seismograms with earthquake waveforms from four moderate earthquakes that occur near broadband stations in southern California. The depth and mechanism corrections show clear improvements of the predicted ground motion relative to the surface impulse response.

Citation: Denolle, M. A., E. M. Dunham, G. A. Prieto, and G. C. Beroza (2013), Ground motion prediction of realistic earthquake sources using the ambient seismic field, *J. Geophys. Res. Solid Earth*, 118, doi:10.1029/2012JB009603.

1. Introduction

[2] Ground motion prediction is a central component of seismic hazard analysis. Until recently, it has been based primarily on ground motion prediction equations, which are regressions of observed ground motion intensity measurements [Abrahamson and Shedlock, 1997; Toro *et al.*, 1997] against source, path, and site descriptions. Those empirical equations suffer from a shortage of data for large seismic events at short distances and only account for wave propagation effects, such as amplification in sedimentary basins, approximately, if at all. Simulations of earthquake rupture and the resulting ground motion have the potential to overcome the lack of data and to properly model wave propagation in basins, so seismologists are increasingly turning to such physics-based methods [Olsen *et al.*, 2006, 2009] for ground motion prediction. If such simulations are to be trusted, however, their accuracy must be established [Hartzell *et al.*, 2011]. Three main sources of uncertainty need to be addressed: the earthquake rupture process, wave propagation in a complex three-dimensional Earth,

and nonlinear site effects. We focus this study on the linear wave propagation, or path effects, and only consider moderate-sized earthquakes, for which a simple source description suffices.

[3] New opportunities in seismology have emerged over the past decade through analysis of the ambient seismic field. This so-called “seismic noise” includes a coherent part that can be extracted through careful time-series analysis. The coherent signal between two stations directly relates to the Green's function, or impulse response function (IRF), from one station to another. The exact IRF between a pair of stations is recoverable by cross-correlating simultaneous time series and stacking (averaging) over time. The ambient noise IRFs (ANIRFs) are dominated by surface-wave fundamental modes at the microseismic period band of 5–20 s [Tanimoto and Alvizuri, 2006; Koper *et al.*, 2011], although several studies [Brooks *et al.*, 2009; Nishida *et al.*, 2008] found higher modes at shorter periods.

[4] To date, ambient noise Green's functions have been used primarily to estimate the velocity structure from Rayleigh-wave [Shapiro and Campillo, 2004; Shapiro *et al.*, 2005; Sabra *et al.*, 2005; Bensen *et al.*, 2007] and Love-wave [Lin *et al.*, 2008] arrival time measurements. For these applications, it is sufficient to extract arrival time information, and the amplitude of the signal need not be preserved. To estimate ground motion, however, we need to capture both elastic (focusing and de-focusing) and inelastic (attenuation) propagation effects that affect wave amplitudes. Validity and accuracy of the amplitudes and dispersion measurements for non-uniform source

¹Department of Geophysics, Stanford University, Stanford, California, USA.

²Departamento de Física, Universidad de los Andes, Bogotá, Colombia.

Corresponding author: M. A. Denolle, Department of Geophysics, Stanford University, 397 Panama Mall, Stanford, CA 94305, USA. (mdenolle@stanford.edu)

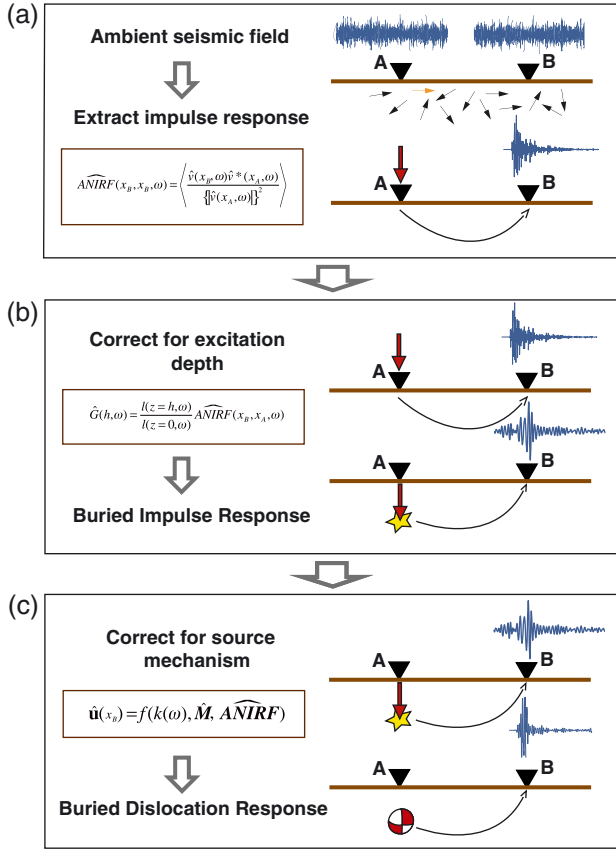


Figure 1. Schema to predict surface-wave response to buried point dislocations using the ambient seismic field. (a) We compute the impulse responses $\widehat{ANIRF}(x_B, x_A, \omega)$ from the ambient noise displacement records at the virtual source A, $\hat{v}(x_A, \omega)$, and receiver B, $\hat{v}(x_B, \omega)$. (b) We solve the surface-wave eigenproblem and use the displacement eigenfunctions $l(z)$ to predict the response to a buried point force (symbolized by red arrows) $\hat{G}(h)$. (c) We use those eigenfunctions and moment tensor (\hat{M}) to predict ground motion in B, $\hat{u}(x_B)$.

distribution are highly debated. Extensive analytical work [Cupillard *et al.*, 2011; Tsai, 2011] has attempted to express the empirical cross-correlation technique to show the effect of the noise source distribution on the amplitude measurements. We address numerically this concern in a recent study (Lawrence, J. F., *et al.*, A numeric evaluation of attenuation from ambient noise correlation functions, submitted to the *Journal of Geophysical Research*, 2013), in which we show that we retrieve accurate amplitude and velocity measurements with the coherence technique for diverse noise source distribution scenarios. Although the noise sources consistently originate from the Pacific coast [Stehly *et al.*, 2006] in the period band of 5–10 s, Prieto and Beroza [2008] use deconvolution of the ambient seismic field without pre-processing and show that the relative amplitude of the ANIRF is preserved and exhibits similar propagation effects to those observed in records of a moderate earthquake. Prieto *et al.* [2009] and Lawrence and Prieto [2011] further show that surface-wave attenuation measured from the raw ambient seismic

field using the SPatial AutoCorrelation (SPAC) method is consistent with previous studies that use surface waves from earthquakes.

[5] Figure 1 shows schematically the steps we undertake to obtain a realistic displacement response to a buried double-couple source. The ANIRF is the surface-wave response of the Earth for a virtual source (a point force) at the surface (station-source) and recorded at the surface (station-receiver). Because the point force and recorded displacement are both three-component vectors, the ANIRF is a rank two tensor with nine components. We compute the nine components of the ANIR tensor following Prieto and Beroza [2008] (Figure 1(a)) and show in the first section that we retrieve reliable propagation information with the ambient noise surface-impulse response. The source depth strongly affects the fundamental mode excitation at short periods, especially for complex velocity structures, such as in southern California, and we must account for that. However, the impulse responses obtained from the ambient noise, which account for the complex 3-D wave propagation, restrict the surface-wave excitation to be at the surface. If we describe the medium surrounding the source to be locally 1-D, we can express the excitation as solution of the surface-wave eigenproblem. In the second section (Figure 1(b)), we use the Generalized Eigenproblem Spectral Collocation (GESC) method [Denolle *et al.*, 2012] to solve the surface-wave eigenproblem by assuming locally a 1-D medium at the virtual source. We use the displacement eigenfunctions to correct the ANIRF and simulate the response of the Earth to a buried point force. In the following section (Figure 1 (c)), we account for radiation pattern effects due to a double couple at depth.

[6] We compare the predicted surface-wave ground motion from the depth- and source-corrected ANIRFs, referred to as “virtual earthquakes,” with records from moderate earthquakes, which we treat as point sources at the wavelengths/periods of interest. We choose four events that occurred close to permanent broadband seismic stations in southern California to validate the approach: the 2009 M 4.5 San Bernardino, 2008 M 5.1 Hector Road, 2008 M 5.4 Chino Hills, and 2011 M 4.2 San Fernando earthquakes. By validating the virtual earthquakes against data, we demonstrate that it should be possible to simulate long-period ground motion from larger earthquakes with this approximation.

2. Impulse Response Functions

[7] Extracting the Green’s function from the ambient seismic field is now common practice in seismology.

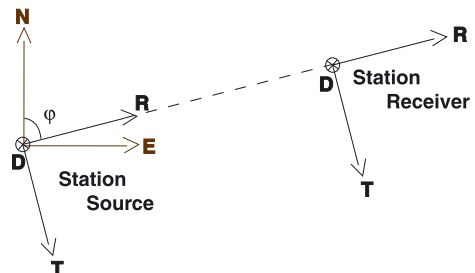


Figure 2. Coordinate system.

Table 1. Earthquake Moment Magnitudes, Dates, and Hypocenters From www.data.sceec.org [Hauksson et al., 2012]; Virtual Source (Seismic Station Closest to Epicenter) Locations; and Range Between Epicenter and Virtual Source

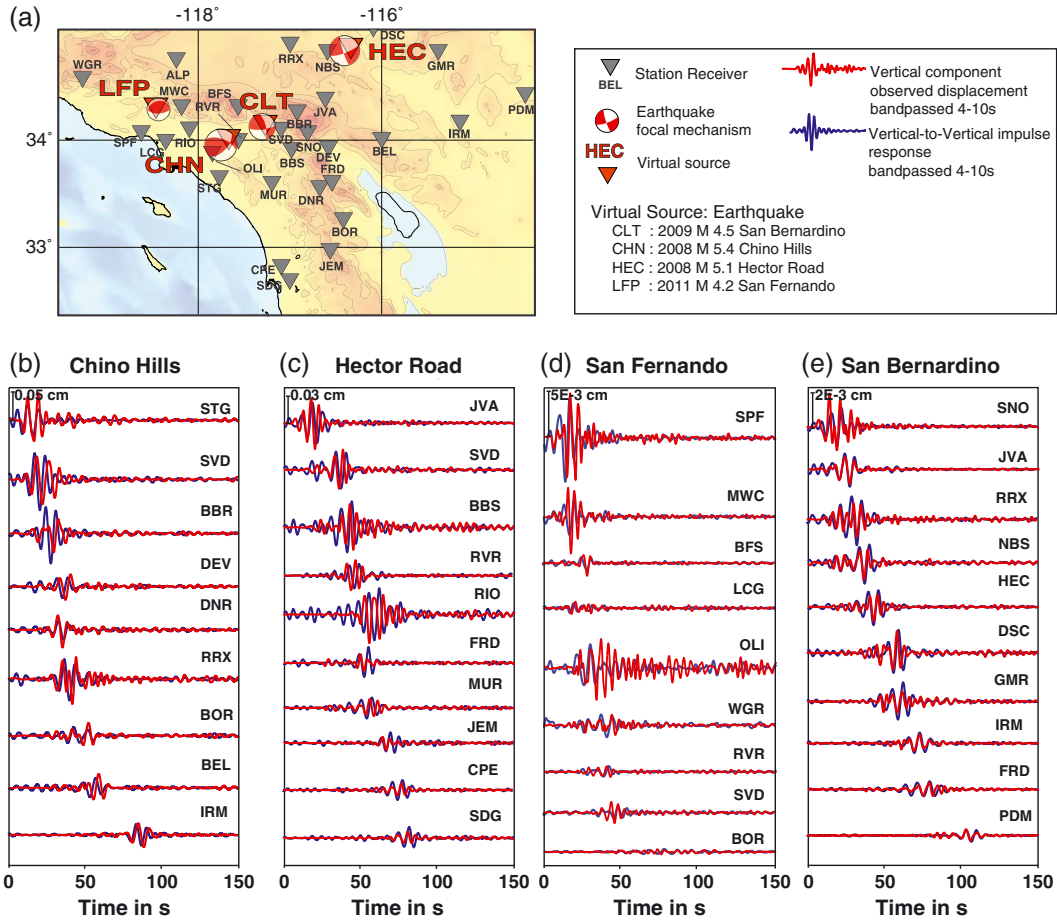
Location	M_w	Date	Latitude (deg)	Longitude (deg)	Depth (km)	Virtual source	Latitude (deg)	Longitude (deg)	Range (km)
Hector Road	5.06	06 Dec 2008	34.813	-116.419	5	HEC	34.829	-116.335	7.89
Chino Hills	5.39	29 Jul 2008	33.953	-117.761	15	CHN	33.998	-117.680	9.03
San Bernardino	4.45	09 Jan 2009	34.107	-117.304	13.8	CLT	34.093	-117.316	1.98
San Fernando	4.24	01 Aug 2011	34.339	-118.475	7.3	LFP	34.305	-118.488	3.95

The cross-correlation function of simultaneously recorded ambient noise records is, under the assumption of homogeneous distribution of noise sources, proportional to the Green's function [Lobkis and Weaver, 2001; Weaver and Lobkis, 2006; Snieder, 2004; Sánchez-Sesma and Campillo, 2006]. To improve the distribution of noise sources, we estimate the impulse response from multiple time windows of the ambient seismic field and then stack the result from many windows that span a long period of time. There are numerous processing techniques to minimize the contribution from unwanted sources and to improve the convergence to the Green's function: e.g., sign-bit normalization [Campillo and Paul, 2003; Shapiro and Campillo, 2004] or pre-whitening

[Bensen et al., 2007]. These methods have shown great efficiency in providing stable results in arrival time and dispersion measurements, but they have the disadvantage of diminishing relative amplitude information.

[8] We use the raw ambient seismic field to compute the IRFs for 1 year of continuous data (during 2010–2011) from the Southern California Seismic Network and ANZA seismic network. We select 1 h long time series, and discard the ones with spikes larger than 10 times the standard deviation of the window, and compute the frequency domain IRF [Prieto et al., 2009].

[9] Stehly et al. [2006] showed that noise sources in southern California vary annually but mainly originate from


Figure 3. DD impulse responses compared with earthquake observations. We show in (a) the earthquake locations and mechanisms, the virtual sources, and the receiver locations. We show in (b)–(e) the vertical-to-vertical ANIRFs (in blue) against the displacements earthquake waveforms (in red), band-passed 4–10 s, for Chino Hills, Hector Road, San Fernando, and San Bernardino earthquakes, respectively.

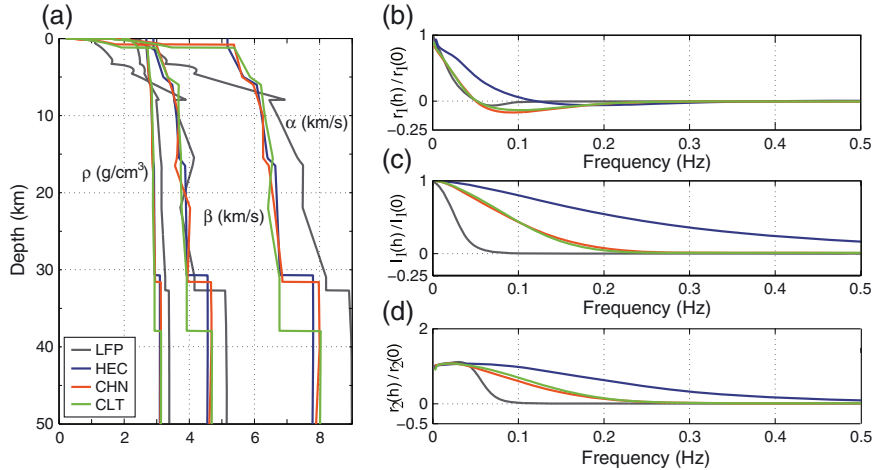


Figure 4. (a) Velocity and density profiles under seismic stations LFP, HEC, CLT, and CHN. (b)–(d) Ratio of the radial (Rayleigh), transverse (Love), and vertical (Rayleigh) displacement eigenfunctions taken at the source depth and the surface for the four respective seismic stations located closest to the earthquakes of interest.

the oceans. We use the technique of *Seats et al.* [2011] to improve the apparent azimuthal distribution of the noise sources, which accelerates convergence to the Green’s function. We divide the day-long records into 70 time windows of 30 min duration, overlapped by 20 min.

[10] In the context of predicting ground motion, we define the ANIR tensor to be proportional to the surface-wave Green tensor $\mathbf{G}(\mathbf{x}, \mathbf{x}', t)$, up to a normalization factor that is common to all receivers. The Green tensor component $G_{ij}(\mathbf{x}, \mathbf{x}', t)$ is the i th component of displacement at receiver position \mathbf{x} and time t by a unit impulse in the j th direction applied at source position \mathbf{x}' and time $t = 0$. The Fourier transform of $\mathbf{G}(\mathbf{x}, \mathbf{x}', t)$ is $\hat{\mathbf{G}}(\mathbf{x}, \mathbf{x}', \omega)$; a similar notation applies to other fields throughout this manuscript. For each station pair (A,B), we compute the ANIR tensor component averaged over many time windows:

$$\widehat{ANIR}_{ij}(\mathbf{x}_B, \mathbf{x}_A, \omega) = \left\langle \frac{\hat{v}_i(\mathbf{x}_B, \omega) \hat{v}_j^*(\mathbf{x}_A, \omega)}{\{\hat{v}_j(\mathbf{x}_A, \omega)\}^2} \right\rangle, \quad (1)$$

where A is the virtual source (seismic station), B is the receiver (seismic station), and $\hat{v}_i(\mathbf{x}_A, \omega)$ and $\hat{v}_j(\mathbf{x}_B, \omega)$ are their respective noise displacement spectra. The operator $\langle \rangle$ denotes stacking over time windows, and $\{ \}$ denotes smoothing over the virtual source spectrum (10-points running average) to ensure stability in the deconvolution. Along with the smoothing operation, we apply a water level if necessary to avoid singularities in the deconvolution.

[11] To extract both Rayleigh and Love waves, we rotate the tensor from the coordinate system North-East-Down (NED) to Radial-Transverse-Down (RTD). The vertical component D is positive downward, and the horizontal rotation is shown for the radial and transverse components, respectively, R and T in Figure 2.

[12] We compute the nine components of the Green’s tensor, given the three channels at each station. We make the assumption that the Rayleigh waves are fully described in the radial and vertical planes (on the tensor components RR, RD, DR, and DD) and that the Love waves are on the

transverse components only (TT). This is an approximation since we ignore off-great-circle propagation and any Love-to-Rayleigh wave conversion (and vice versa), surface wave-to-body wave conversion from complex 3-D structure [Gregersen, 1978; Yoshida, 2003; Langston et al., 2009], or anisotropy [Yao et al., 2011]. These effects may be present in the Green’s function, but we do not correct for them in the excitation nor in the recording at the specific locations. In southern California, we see some leakage of energy on the cross-terms DT, TD, RT, and TR in the period band 4–10 s. Although this is not the focus of our study, it is a useful observation that could be used to constrain crustal structure.

[13] Once we account for the common proportionality factor between the Green’s function and the ANIRF, we use the superscript *AN* to refer to the ANIR tensor \mathbf{G}^{AN} and we assume that $\hat{G}_{TD}^{AN} = \hat{G}_{TR}^{AN} = \hat{G}_{DT}^{AN} = \hat{G}_{RT}^{AN} = 0$ at all frequencies. Note that \mathbf{G}^{AN} refers specifically to the Green tensor between two surface locations. We exploit the causal and anti-causal symmetry of the Green’s function [Snieder, 2004; Bensen et al., 2007] by averaging the causal and anti-causal time series.

[14] To verify that we retrieve correct path effects from the ANIR tensor, we compute the impulse responses from the stations closest to the four epicenters, or virtual sources: CHN for Chino Hills, HEC for Hector Road, LFP for San Fernando, and CLT for San Bernardino (Table 1). First, we compare the vertical-to-vertical impulse responses with the vertical displacement earthquake records, band-passed 4–10 s. We calibrate the ANIRFs to the observed displacement amplitudes using a normalization factor (peak displacement amplitude) common to all stations pairs that accounts for the strength of the coherent signal traveling from the virtual source.

[15] We correct the ANIRF for the distance between the station source and the estimated epicenters, listed in Table 1, by using the phase velocity dispersion curve computed from GESC at the virtual source location.

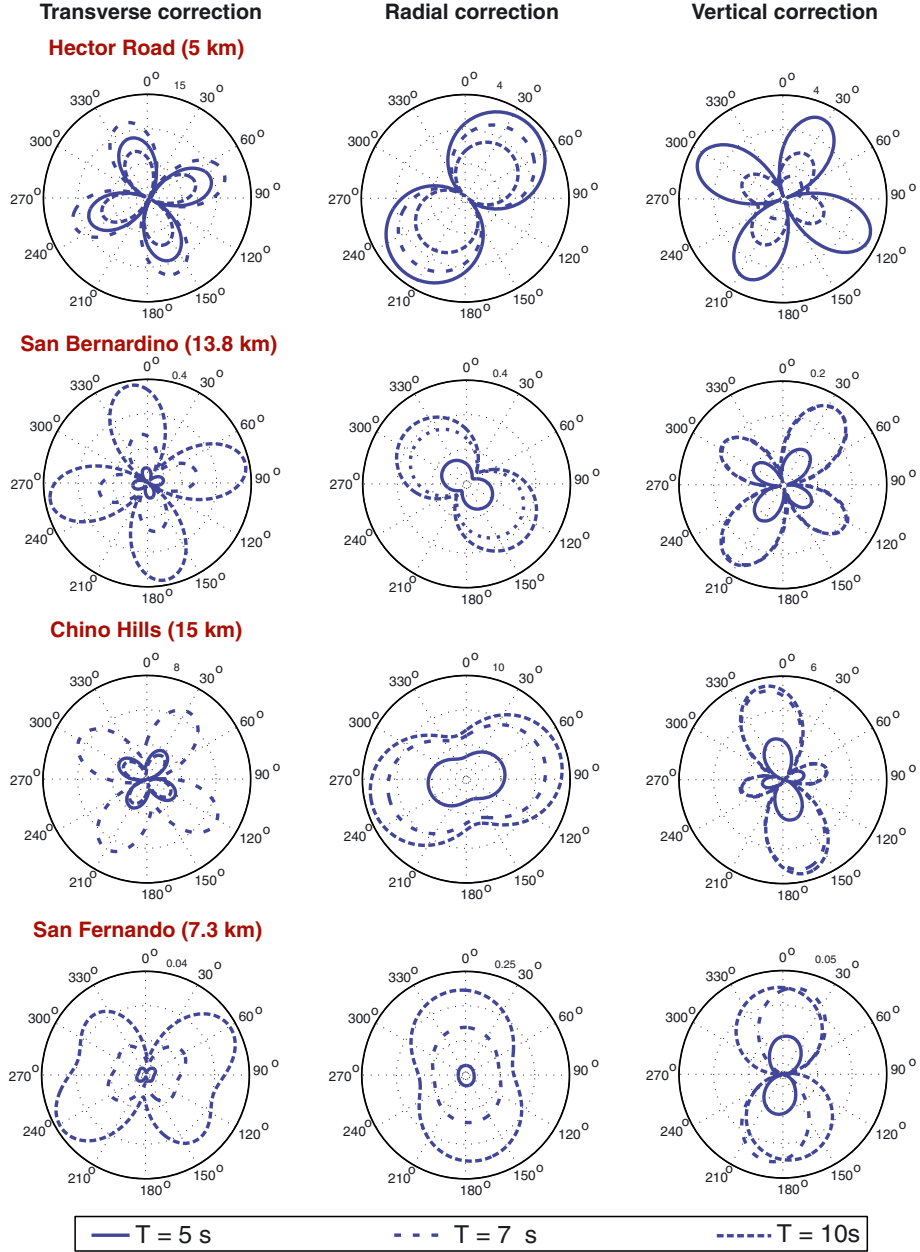


Figure 5. Correction terms to convert surface-impulse responses to buried double-couple radiation (displacements). The absolute values of the three factors of conversion described in (13) and (15) are shown in polar plots and their azimuth and period dependence using the four earthquake moment tensors listed in Table 2. For each polar representation, the maximum amplitude is shown at azimuth 15° . We impose a flat response of the moment-rate function at those periods.

[16] In Figure 3, we show the similarity between the observed records and the vertical-to-vertical component of the ANIR tensor only. The horizontal components reveal similar results. By rotating the Green tensor into the RTD coordinate system, we align the orientation of the coordinate system into the maximum of the single-force radiation lobes. We therefore expect some misfit between the earthquake records that include the dislocation source mechanism radiation and the impulse responses. However, in the 4–10 s period band, we see a reasonable match of the amplitudes between both waveforms for the specific examples shown in Figure 3. This implies that the impulse responses

capture, to first order, the path effects from the virtual source as found by *Prieto and Beroza* [2008]. Figures 3(b)–3(e) illustrate the power of using the ANIRFs as a tool for ground motion prediction. The geometrical decay clearly dominates the surface-wave amplitudes; however, for receivers located in the Los Angeles Basin, the ANIRFs also capture the observed local surface-wave amplification (stations RIO in Figure 3(d) and OLI in Figure 3(e)) and extended duration.

[17] Surface-wave excitation depends on the complexity of the velocity structure and is strongly frequency dependent. A source at the surface will more efficiently excite

Table 2. Virtual Source Parameters: Seismic Moment, Deviatoric Solution of the Moment Tensor (From SCSN, Normalized to 10^{16} Nm With the Convention of z Positive Downward), and Estimated Corner Frequency f_c

Virtual Source	Mo (Nm)	Mxx	Mxy	Mxz	Myy	Myz	Mzz	f_c (Hz)
HEC	4.9×10^{16}	-2.749	-3.734	-0.959	3.052	-0.902	-0.304	0.58
CHN	1.53×10^{17}	-14.17	4.9	-1.9	5.85	-7.39	8.32	0.39
CLT	5.96×10^{15}	-0.2507	-0.4974	0.1639	0.1259	-0.1828	0.1248	1.17
LFP	1.74×10^{15}	-0.14487	-0.02634	0.09129	-0.0004	-0.0116	0.14574	1.71

short-period waves than a buried source. Therefore, for a better representation of earthquake ground motion, we have to correct the ANIRF for the effect of depth on the surface-wave excitation.

3. Correction for Source Depth

[18] In this section, we focus on the fundamental surface-wave modes and assume that the medium surrounding the source can be approximated by a 1-D vertical structure. We express the surface-wave part of Green tensor in the RTD system as the contribution of both Rayleigh and Love waves:

$$\mathbf{G} \approx \mathbf{G}^L + \mathbf{G}^R = \begin{pmatrix} G_{RR} & 0 & G_{RD} \\ 0 & G_{TT} & 0 \\ G_{DR} & 0 & G_{DD} \end{pmatrix}. \quad (2)$$

[19] We use the convention of *Aki and Richards* [2002] and derive the source-depth dependence of the Green tensor in cylindrical coordinates. For a surface receiver located at \mathbf{x} and a source located at \mathbf{x}' , we define horizontal distance between the source and receiver r , azimuthal angle ϕ , and source depth h . To simplify the notation, we suppress explicit ω -dependence while retaining the source-depth dependence h . The Love-wave Green tensor is naturally expressed in the frequency domain as

$$\hat{\mathbf{G}}^L(h) = \frac{1}{8c_L U_L I_1} \sqrt{\frac{2}{\pi k_L r}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & l_1(0)l_1(h) & 0 \\ 0 & 0 & 0 \end{pmatrix} \exp(i(k_L r + \pi/4)), \quad (3)$$

where $l_1(z)$ is the Love-wave displacement eigenfunction at depth z , c_L the phase velocity, U_L the group velocity, k_L the

wave number, and $I_1 = 1/2 \int_0^\infty \rho(z) l_1^2(z) dz$ the first energy integral. The Rayleigh-wave Green tensor is

$$\hat{\mathbf{G}}^R(h) = \frac{1}{8c_R U_R I_1} \sqrt{\frac{2}{\pi k_R r}} \times \begin{pmatrix} r_1(0)r_1(h) & 0 & -ir_1(0)r_2(h) \\ 0 & 0 & 0 \\ ir_2(0)r_1(h) & 0 & r_2(0)r_2(h) \end{pmatrix} \exp(i(k_R r + \pi/4)), \quad (4)$$

with similar notation as in the Love case; the Rayleigh-wave group velocity, phase velocity, and wave numbers have the subscript R and $I_1 = 1/2 \int_0^\infty \rho(z) (r_1^2(z) + r_2^2(z)) dz$. The horizontal and vertical displacement eigenfunctions, respectively, $r_1(z)$ and $r_2(z)$, are frequency and depth dependent.

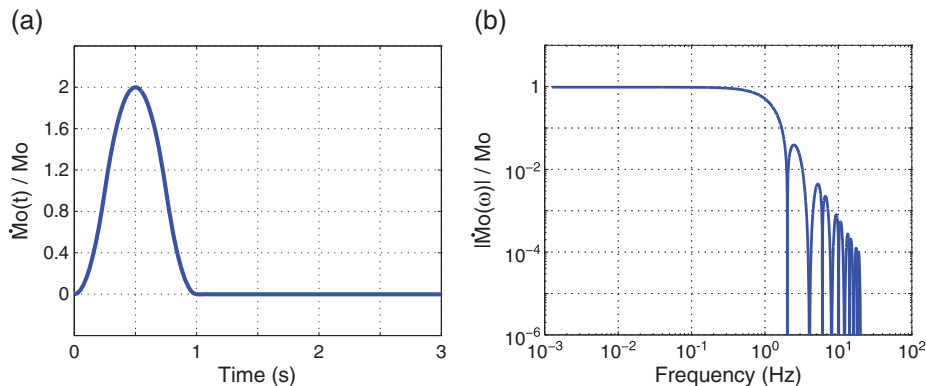
[20] We note from (3) that

$$\hat{G}_{TT}(h) = \frac{l_1(h)}{l_1(0)} \hat{G}_{TT}(0). \quad (5)$$

[21] For Love waves, (5) shows a linear relationship between the Green tensor component for a source at depth and the Green tensor component for a source at the surface. The correction factor is the ratio of the displacement eigenfunctions taken at the source depth and surface. The ambient noise Green tensor \mathbf{G}^{AN} contains the information on the surface-wave 3-D propagation in the true complex crustal structure. We replace $G_{TT}(0)$ with G_{TT}^{AN} to obtain

$$\hat{G}_{TT}(h) \approx \frac{l_1(h)}{l_1(0)} \hat{G}_{TT}^{AN}. \quad (6)$$

[22] This relation is always stable, in the sense that division by zero or nearly zero is avoided, because the


Figure 6. Normalized moment-rate function (a) and spectrum (b) for a corner frequency $f_c = 0.5$ Hz.

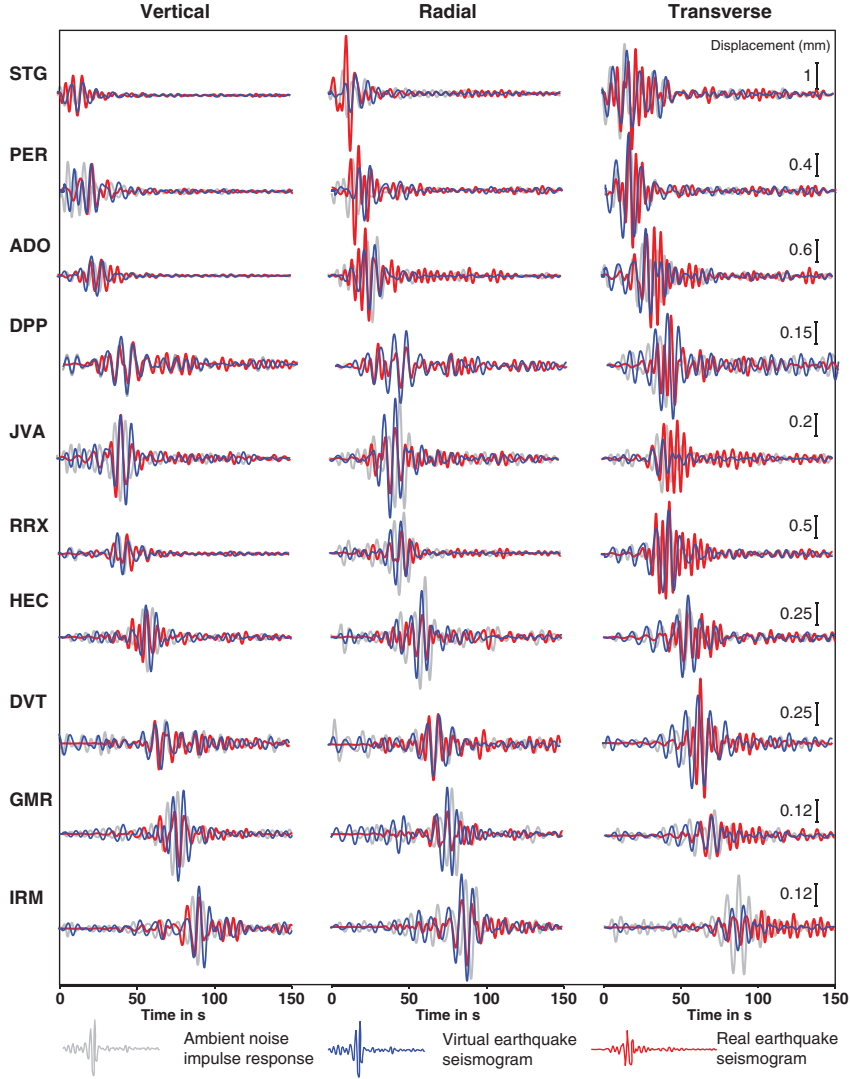


Figure 7. Waveform comparison, band-passed 4–10 s, for the 2008 M 5.4 Chino Hills earthquake for all three components: the initial diagonal terms of the ANIRF tensor (DD, RR, and TT in gray), the earthquake records (D, R, and T in red), and the VEA waveforms (D, R, and T in blue). We show the receiver locations in the black upside-down triangles in all panels of Figure 11.

Love-wave displacement eigenfunctions are nonzero at the surface.

[23] For the Rayleigh-wave components of the Green tensor, we have analogous expressions:

$$\hat{G}_{RR}(h) \approx \frac{r_1(h)}{r_1(0)} \hat{G}_{RR}^{AN} \quad \text{and} \quad \hat{G}_{RD}(h) \approx \frac{r_2(h)}{r_2(0)} \hat{G}_{RD}^{AN}, \quad (7)$$

$$\hat{G}_{DR}(h) \approx \frac{r_1(h)}{r_1(0)} \hat{G}_{DR}^{AN} \quad \text{and} \quad \hat{G}_{DD}(h) \approx \frac{r_2(h)}{r_2(0)} \hat{G}_{DD}^{AN}. \quad (8)$$

[24] These relationships are stable in most cases for similar reasons. *Tanimoto and Rivera* [2005] and *Denolle et al.* [2012], however, highlight changes from the usual retrograde to prograde Rayleigh-wave particle motion at the free surface for certain velocity models and frequencies. In such cases, the $r_2(0)$ can be small (or even zero) and this correction becomes ill-conditioned. We use the Southern California Earthquake Center Community Velocity Model Version 4.0 (CVM4.0) [*Magistrale et al.*, 2000] for

which we do not see such changes in particle motion at the frequencies of interest.

[25] We extract from CVM4.0 the velocity profiles at the four seismic stations closest to the epicenters and represent them in Figure 4(a). Table 1 contains the locations of the earthquake hypocenters [*Hauksson et al.*, 2012] with their respective virtual source (seismic station) name and location. LFP, CHN, and CLT are located in sedimentary basins, which involve strong velocity gradients at shallow depth. HEC is located in the Mojave Desert and presents characteristics of the shallow crust that are closer to bedrock. The Hector Road event occurred shallow (~ 5 km) on a strike-slip fault that accommodates part of the distributed right-lateral motion of the Eastern California Shear Zone [*Savage et al.*, 2001]. The Chino Hills earthquake occurred on the blind thrust underlying the sedimentary basin [*Hauksson et al.*, 2008] at greater depth (~ 15 km), whereas the San Bernardino earthquake, with similar depth, is near the deepest part of the San Andreas Fault system. The 2011

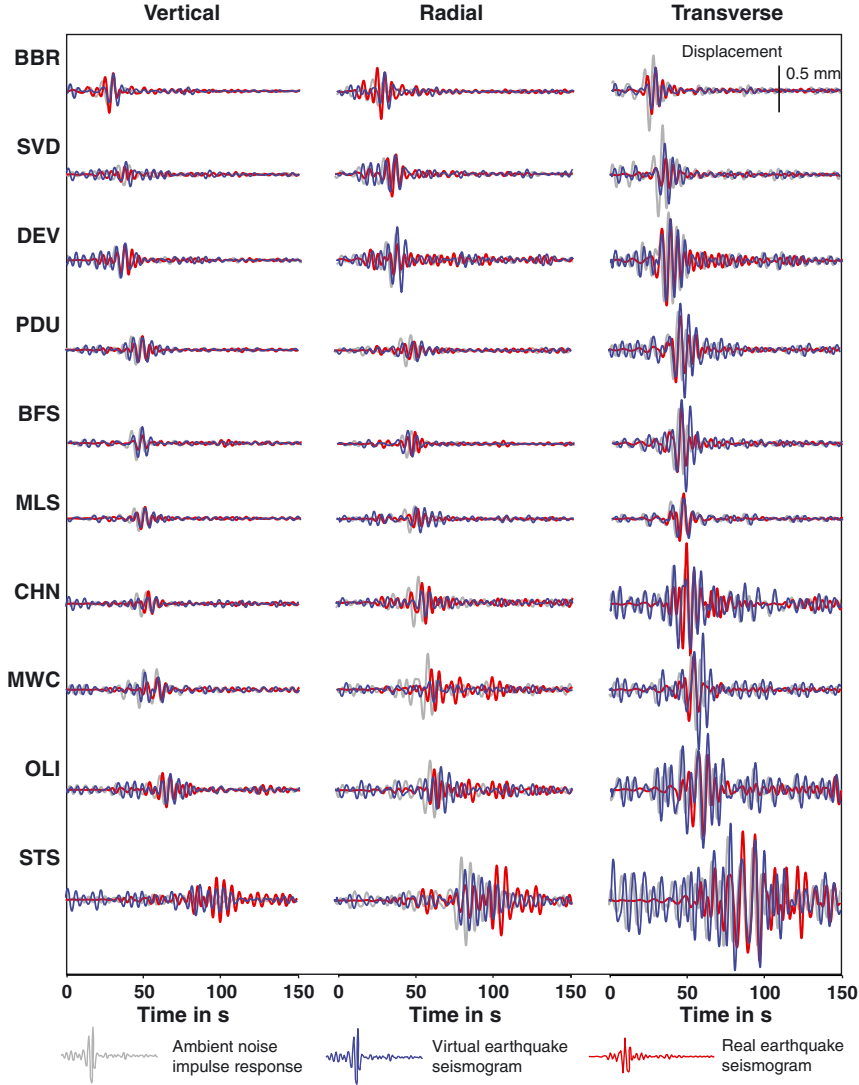


Figure 8. Same as Figure 7 for the 2008 M 5.1 Hector Road earthquake.

San Fernando earthquake occurred at ~ 7.3 km near the location of the 1994 M 6.7 Northridge earthquake.

[26] We represent in Figure 4 the source-depth correction to the Green tensor in the 1-D approximation. The dominant pattern of this correction (right panel of Figure 4) is the filtering of the high frequencies for sources at depth. For the radial components, the second feature to note is the change of sign in the spectral correction related to the frequency-dependent zero crossing of the radial eigenfunction at depth.

4. Radiation Pattern Correction

[27] We write the surface-wave displacements, \hat{u}_i , at \mathbf{x} generated by a point source described by the moment tensor $\hat{\mathbf{M}}$ at the source located at \mathbf{x}' as

$$\hat{u}_i(\mathbf{x}) = \hat{M}_{pq} \frac{\partial}{\partial x'_q} \hat{G}_{ip}(\mathbf{x}, \mathbf{x}'). \quad (9)$$

[28] Assuming that $\hat{\mathbf{G}}$ can be expressed in the form given in equations (2)–(4), we follow *Aki and Richards*

[2002], assuming that the largest contributions are from depth derivatives of the eigenfunctions and the horizontal derivatives of $\exp(ikr)$. We rotate the coordinate system from NED to RTD and, under the approximation described earlier, simplify the horizontal partial derivatives to

$$\frac{\partial \hat{G}}{\partial R} = -ik\hat{G} \quad \text{and} \quad \frac{\partial \hat{G}}{\partial T} = 0. \quad (10)$$

[29] For Love waves, we expand (9) using the depth-corrected Green tensor $\hat{\mathbf{G}}(h)$ and approximate

$$\hat{u}_T = \hat{M}_{TD} \frac{\partial \hat{G}_{TT}}{\partial Z} \Big|_{Z=h} - ik_L \hat{M}_{TR} \hat{G}_{TT}(h). \quad (11)$$

[30] From (3), we see that

$$\frac{\partial \hat{G}_{TT}}{\partial Z} \Big|_{Z=h} = \frac{l'_1(h)}{l_1(h)} \hat{G}_{TT}(h) = \frac{l'_1(h)}{l_1(0)} \hat{G}_{TT}(0). \quad (12)$$

[31] And approximating $\hat{G}_{TT}(0) \approx \hat{G}_{TT}^{AN}$,

$$\hat{u}_T \approx \frac{1}{l_1(0)} \left[-ik_L \hat{M}_{TR} l_1(h) + \hat{M}_{TD} l'_1(h) \right] \hat{G}_{TT}^{AN}. \quad (13)$$

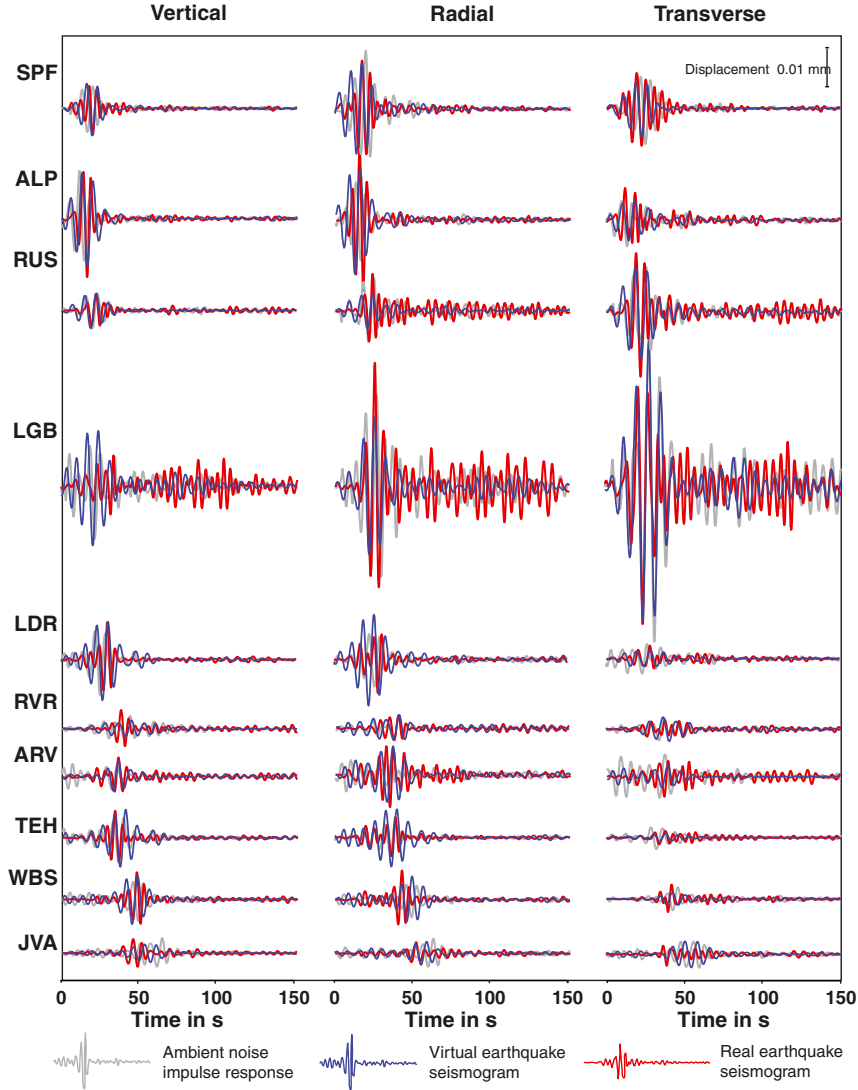


Figure 9. Same as Figure 7 for the 2011 M 4.2 San Fernando earthquake.

[32] In a similar manner, one can show that for Rayleigh waves,

$$\hat{u}_D \approx \frac{1}{r_1(0)} \left[-ik_R \hat{M}_{RR} r_1'(h) + \hat{M}_{RD} r_1'(h) \right] \hat{G}_{ZD}^{AN} \quad (14)$$

$$+ \frac{1}{r_2(0)} \left[-ik_R \hat{M}_{DR} r_2'(h) + \hat{M}_{DD} r_2'(h) \right] \hat{G}_{DD}^{AN},$$

$$\hat{u}_R \approx \frac{1}{r_2(0)} \left[-ik_R \hat{M}_{DR} r_2'(h) + \hat{M}_{DD} r_2'(h) \right] \hat{G}_{RD}^{AN} \quad (15)$$

$$+ \frac{1}{r_1(0)} \left[-ik_R \hat{M}_{RR} r_1'(h) + \hat{M}_{RD} r_1'(h) \right] \hat{G}_{RR}^{AN}.$$

[33] For Rayleigh waves, the conditioning of the correction strongly depends on the particle motion at the surface. The vertical displacement eigenfunction $r_2(0)$ becomes zero when particle motion changes from retrograde to prograde (and vice versa). Those changes introduce singularities in the conversion that we do not encounter for the frequencies of interest when using CVM4.0.

[34] We see that (13)–(15) relate the ANIR tensor components to the earthquake displacements with three main factors depending on the vertical, radial or transverse components. We represent their absolute values at given periods (5, 7, and 10 s) in Figure 5 and their variations with respect to azimuth and frequency of the source–depth and double-couple correction with a flat moment–rate function spectrum. The effect of the source depth is clearly expressed because buried sources excite short-period surface waves less efficiently than shallower sources. The other main feature illustrated in Figure 5 is the presence of four lobes at the transverse and vertical components for the pure strike-slip events (Hector Road and San Bernardino).

[35] As mentioned earlier, the 2008 M 5.1 Hector Road earthquake occurred in the Eastern California Shear Zone and the moment tensor solution exhibits almost pure strike-slip (Figure 3(a)) motion. The 2009 M 5.4 Chino Hills and 2009 M 4.5 San Bernardino earthquake mechanisms feature oblique strike-slip motion, as shown in (Figure 3(a)). The 2011 M 4.2 San Fernando earthquake mechanism is almost almost pure reverse faulting.

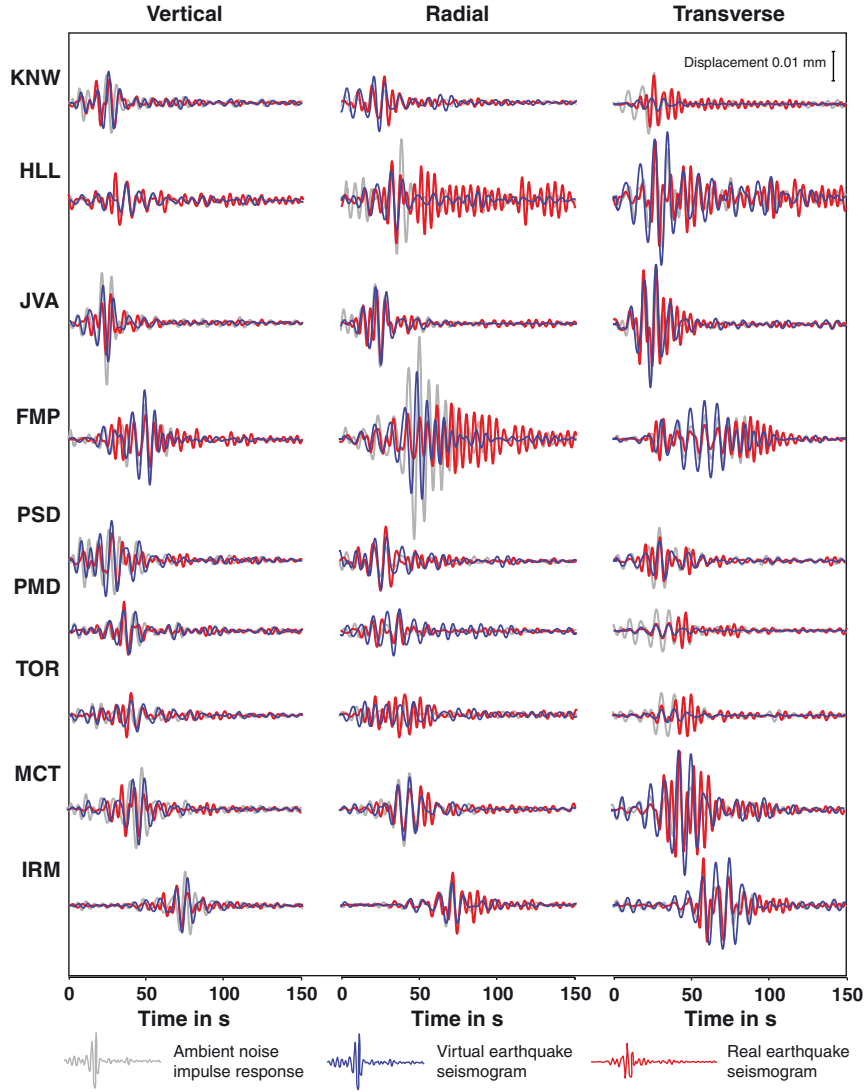


Figure 10. Same as Figure 7 for the 2011 M 4.5 San Bernardino earthquake.

5. Source Duration

[36] In this study, we choose earthquakes of moderate magnitude so that finite-fault effects are minor at the periods of interest. There is a trade-off, however, because we require adequate signal-to-noise ratio at longer periods. Earthquakes in the range of M 4.5–5.5 provide a good balance between these two considerations. We account for the finite duration of the event with an assumed pulse width for the moment-rate function. The duration T of the pulse is controlled by the corner frequency such that $T = 1/2f_c$. We estimate the corner frequency of the events based on *Hanks and Thatcher* [1972] with an assumed stress drop $\Delta\sigma = 3$ MPa, the observed seismic moment, M_0 , and shear velocity $\beta = 3$ km/s:

$$f_c = 0.491\beta \left(\frac{\Delta\sigma}{M_0} \right)^{1/3}. \quad (16)$$

[37] We list each earthquake corner frequency in Table 2. In the far-field approximation, the displacement field is proportional to the moment-rate function. We use a

parabolic moment-rate function [*Herrmann, 1978*] with the Fourier spectrum

$$\hat{S}(\omega) = \exp(-i\omega T/2) \frac{4 \sin^2(\omega T/8) \sin(\omega T/4)}{(\omega T/4)^3}, \quad (17)$$

where the moment-rate function $\dot{M}_0(t) = M_0 S(t)$ is shown in Figure 6.

6. Validating the Virtual Earthquake Approach With Seismic Observations

[38] We have shown how to incorporate the effects of more realistic source parameters in the ANIRFs to produce far-field surface-wave seismograms for a buried double-couple source that can be directly compared with earthquake observations. Once we account for the source depth, double-couple mechanism, and finite duration of the moment-rate pulse, we have constructed the virtual earthquake. To validate what we refer to as the Virtual Earthquake Approach (VEA), we compare the virtual earthquake seismograms with the earthquake records, between 4 and 10 s period.

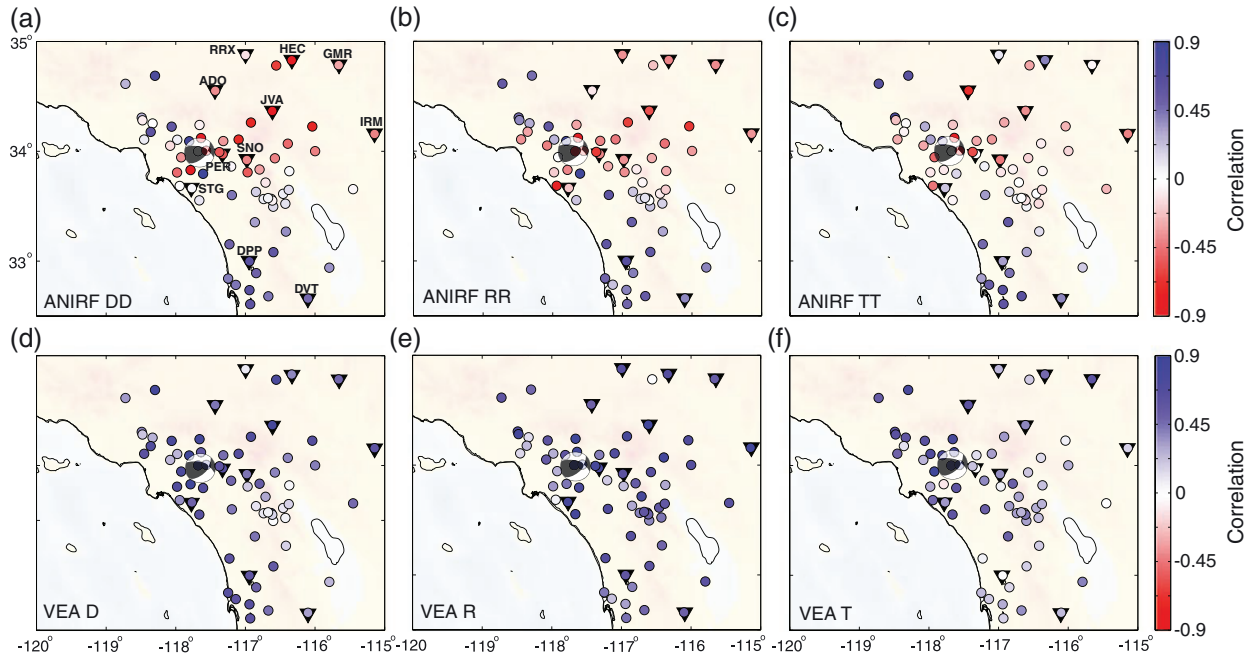


Figure 11. Maps representing the spatial distribution of the normalized correlation coefficients (CCs) between the virtual earthquake and real earthquake waveforms, at each receiver, for the 2008 M 5.4 Chino Hills earthquake. The color scale shows the CC values ranging between -0.9 (red) and 0.9 (blue). The top panels show the correlation between the initial ANIRFs and observed displacements at all three components: DD (a), RR (b), and TT (c). The black upside-down triangles show the locations of the receivers used in Figure 7. The bottom panels show the correlation between the VEA and the earthquake waveforms at all three components D (d), R (e), and T (f).

[39] The epicenters are not exactly collocated with the station source (Table 1); the distance between these and the seismic stations varies between 3 and 8 km. We combine the estimated surface-wave dispersion information, taken at the virtual source, and the difference between epicenter-receiver and virtual source-receiver locations to account for

the expected time shift. For the earthquakes considered, this correction is as large as 3 s, which is significant over the period band of interest.

[40] In making this comparison, we need to account for several sources of uncertainty in our system. First, we approximate the moment-rate pulse width based on an

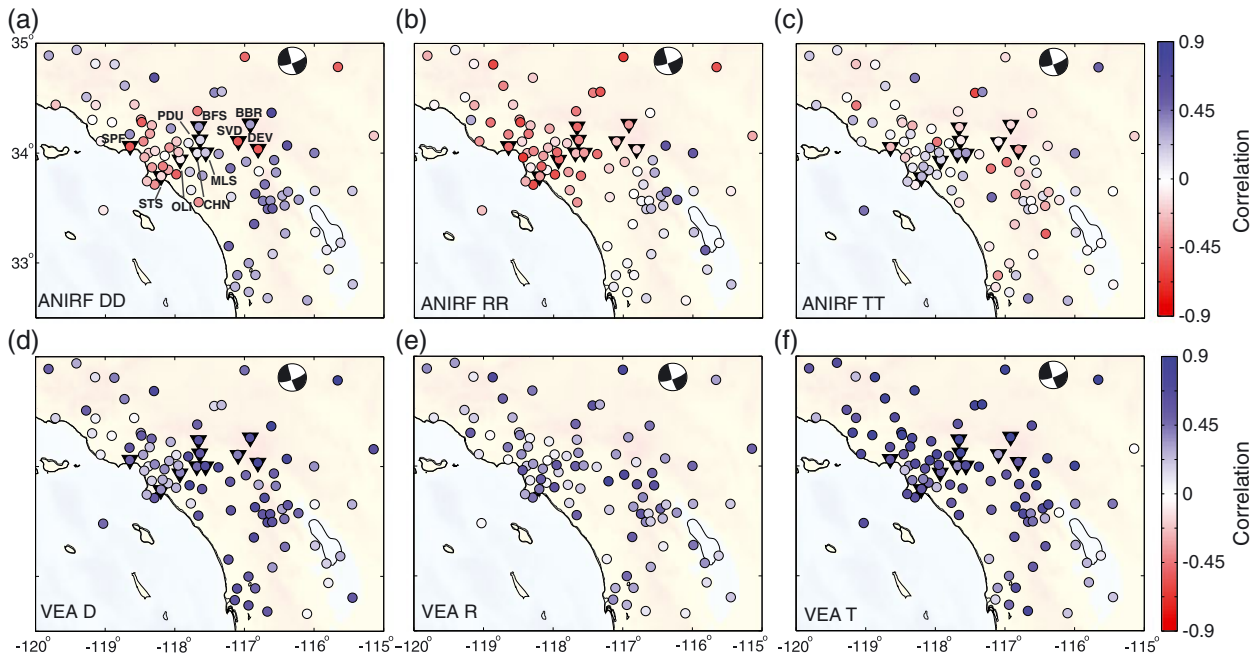


Figure 12. Same as Figure 11 for the 2008 M 5.1 Hector Road earthquake.

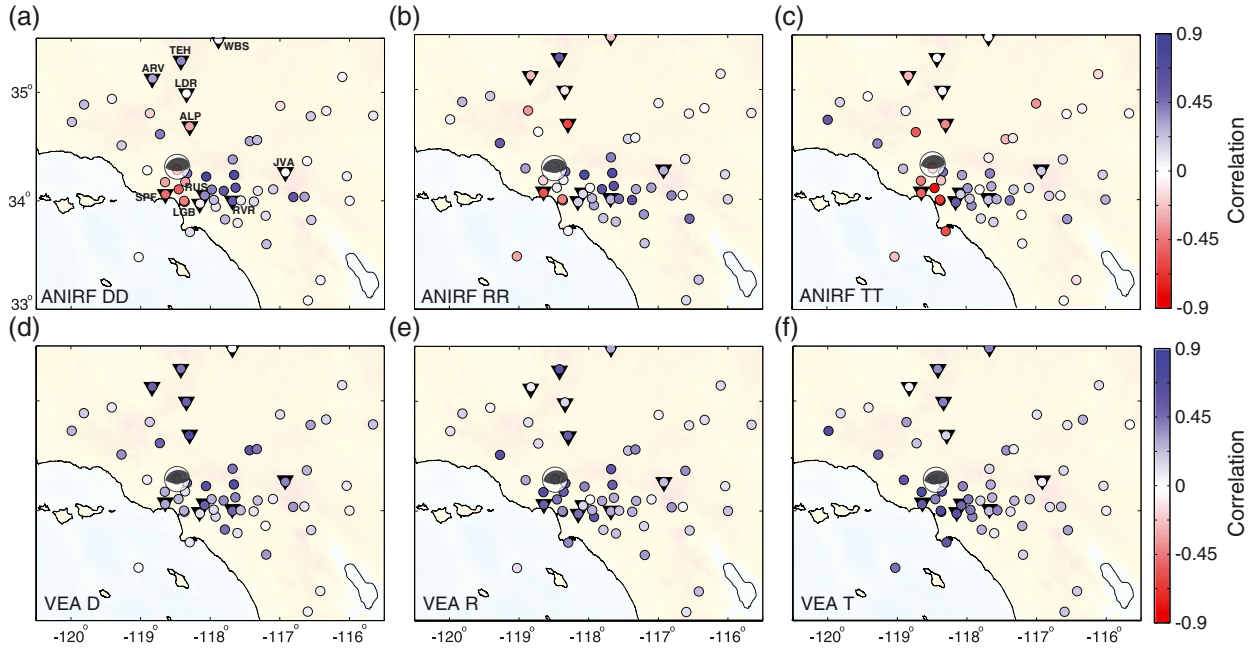


Figure 13. Same as Figure 11 for the 2011 M 4.2 San Fernando earthquake.

assumed corner frequency, which we calculate for constant stress drop of 3 MPa despite its expected strong variability [Baltay *et al.*, 2011]. Depending on the seismic moment and stress drop, the time delay due to the finite width of the pulse ranges between 0.3 and 1 s. This particular source of uncertainty will be consistent over the entire seismic network. Second, we use the surface-wave eigenfunctions for a 1-D velocity profile at the station source. By comparing velocity profiles from CVM4.0 at the epicenter and station-source locations, we find variations in phase velocity that lead to

variations in the phase shift of at most 2–3 s for stronger variations in the velocity structure. Finally, we allow for arrival time uncertainty of 0.2 s due to the SCSN hypocenter location uncertainties. We combine the possible effects of all these uncertainties by allowing a conservative free phase shift of 1.5 s to maximize the correlation between the virtual and real earthquake waveforms. To isolate the changes in the waveforms only due to the VEA, we also correct the initial IRFs in the same way by allowing the station source and the epicenter the same 1.5 s shift.

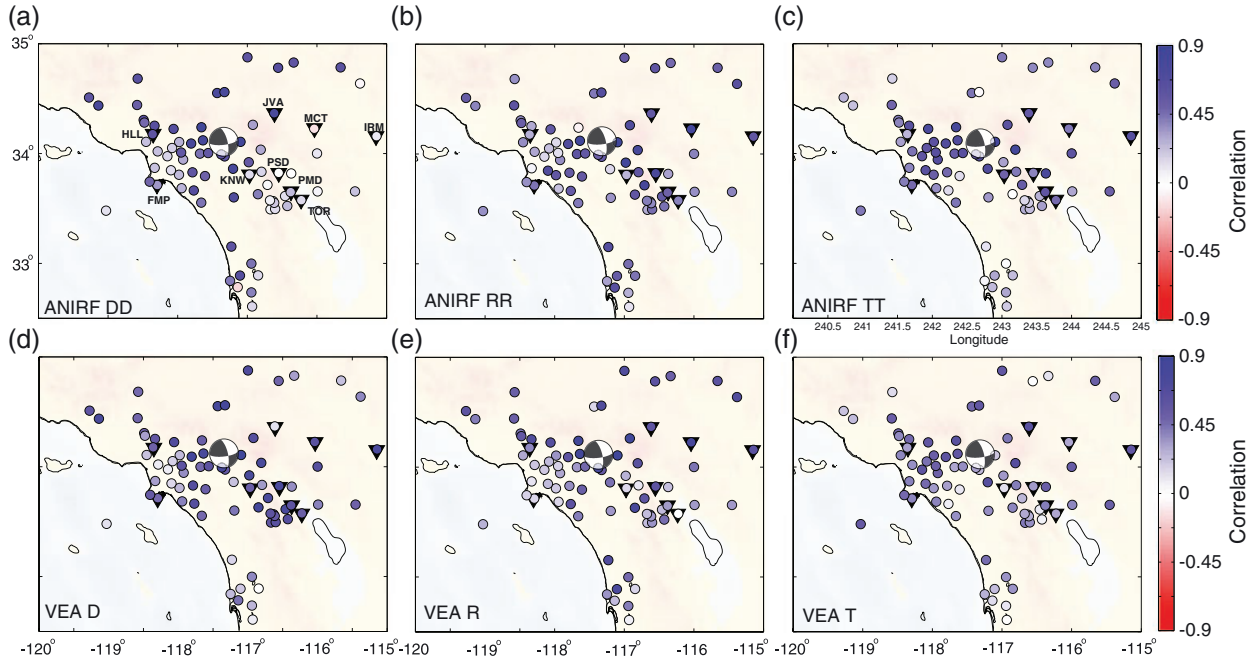


Figure 14. Same as Figure 11 for the 2011 M 4.5 San Bernardino earthquake.

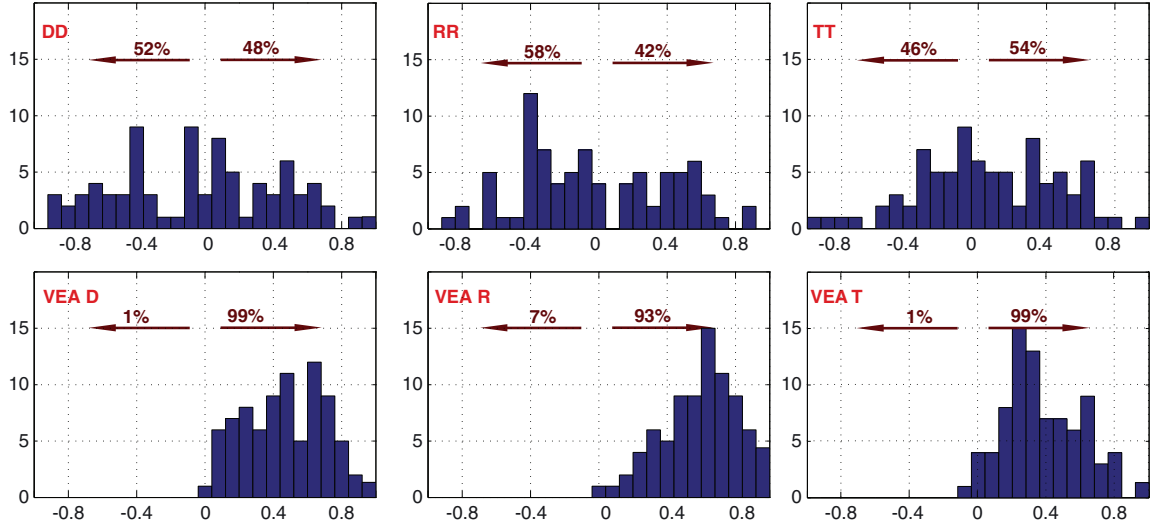


Figure 15. Histograms of the correlation coefficients for 2008 M 5.4 Chino Hills earthquake at all three components (from left to right: R, T, and D): the CC values for the initial ANIRFs diagonal components on the top panels and the CC values for the VEA waveforms in the middle panels. The bottom panels show consistently better fits to the earthquake data than the upper panels.

[41] For each earthquake and each component, we calibrate the ANIR tensor and VEA waveform amplitudes with the earthquake records by taking the peak displacement amplitudes (between 4 and 10 s), averaged over all the station pairs. To first order, this normalization accounts for the difference in strength of coherence between the virtual source and receivers.

[42] We show in Figures 7–10 comparisons between the earthquake displacement waveforms, the initial impulse responses (diagonal terms of the Green tensor), and the displacements calculated with the VEA, all band-passed from 4–10 s. The VEA waveforms show a much better fit to the earthquake records for all three components for most of the stations. The initial IRFs show strong similarity with the waveforms, both in phase and amplitude,

for the San Bernardino and San Fernando earthquakes. This occurs because the diagonal terms significantly dominate the response to the buried double couple for these two examples. We find greater improvement for the Hector Road and Chino Hills events, where the diagonal terms do not dominate.

[43] To evaluate the waveform fit, we calculate the normalized correlation coefficient (CC) for each component and at each station, between the ambient-noise-derived responses \mathbf{u} and the earthquake records \mathbf{v} . We compute this coefficient on a variable time window that contains most of the surface-wave energy. We calculate the cumulative energy (in the root mean square sense) of the waveform and select the time window that contains between 1% of the energy and 90%. We index those,

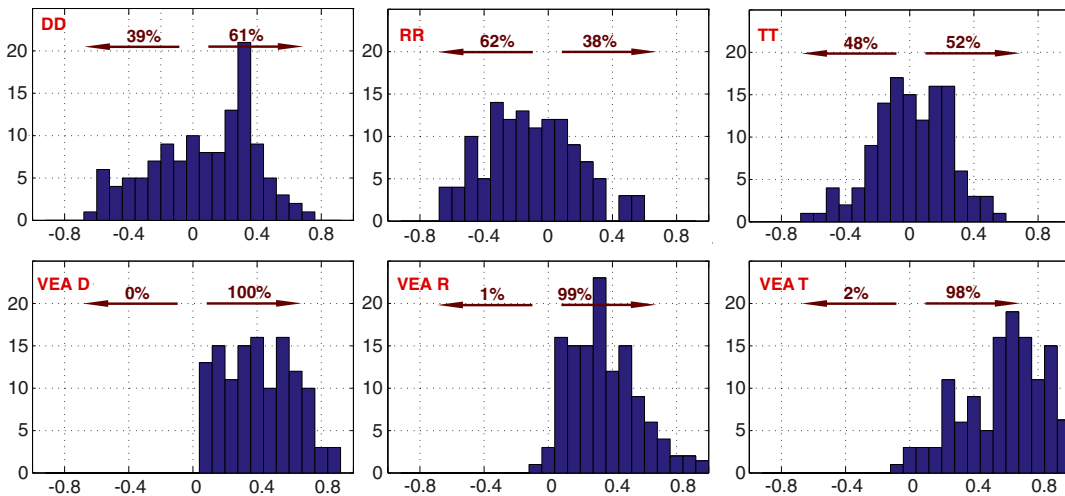


Figure 16. Same as Figure 15 for the 2008 M 5.1 Hector Road earthquake.

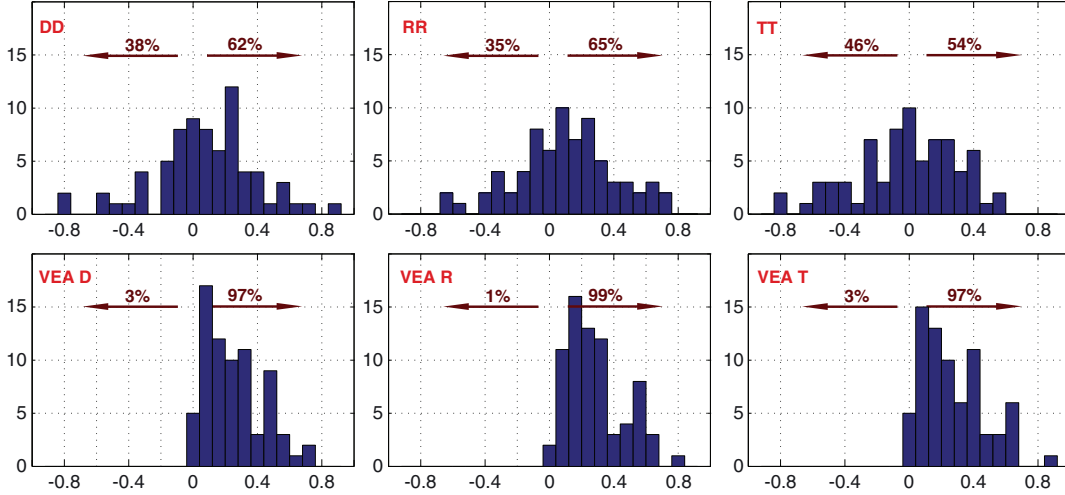


Figure 17. Same as Figure 15 for the 2011 M 4.2 San Fernando earthquake.

respectively, N_1 and N_{90} . The correlation coefficient at each station is then

$$CC = \frac{\sum_{i=N_1}^{N_{90}} u_i v_i}{\sqrt{\sum_{i=N_1}^{N_{90}} u_i^2 \sum_{i=N_1}^{N_{90}} v_i^2}}. \quad (18)$$

[44] For each earthquake (Figures 11–14), we show a map view of the CCs across southern California. For the smaller earthquakes (San Bernardino in Figure 13 and San Fernando in Figure 10), the correlation is initially high on all three components, as explained earlier, validating again the use of the ambient seismic field for ground motion prediction. For these earthquakes, the conversion between surface-impulse response to buried dislocation preserves the goodness of fit between the observed data and the ANIRFs. For Hector Road (Figure 11) and Chino Hills (Figure 12) events, there

is no obvious correlation between the diagonal terms of the ANIR tensor and the observed waveforms. The VEA clearly improves the accuracy of the predicted ground motion relative to the ambient noise surface-impulse responses. Apart from isolated cases, all the correlations between the new waveforms and the earthquake records are positive at most stations and components.

[45] We can represent the CCs distribution in another, more quantitative way. In Figures 15–18, we show the distribution of the CC values. The top panels describe the ranges of CCs for all four earthquakes between the diagonal terms of the Green tensor and the observed waveforms. We find positive correlations for the San Fernando and San Bernardino events and the missing radiation pattern for the Hector Road and Chino Hills events. The bottom panels show the results of applying our technique to the ANIRFs, and the overall improvement. We confirm that the VEA waveforms show overall a good match in the phase of the observed records.

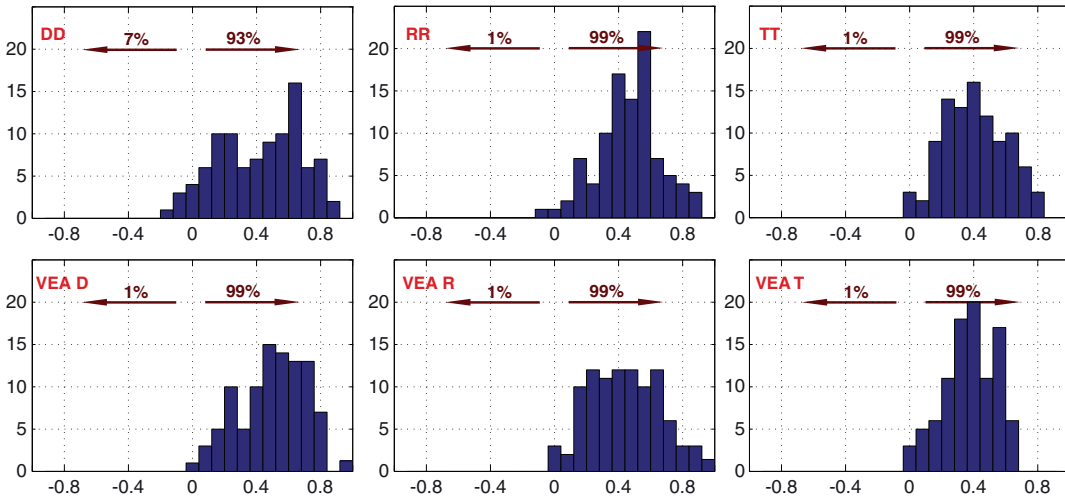


Figure 18. Same as Figure 15 for the 2011 M 4.5 San Bernardino earthquake.

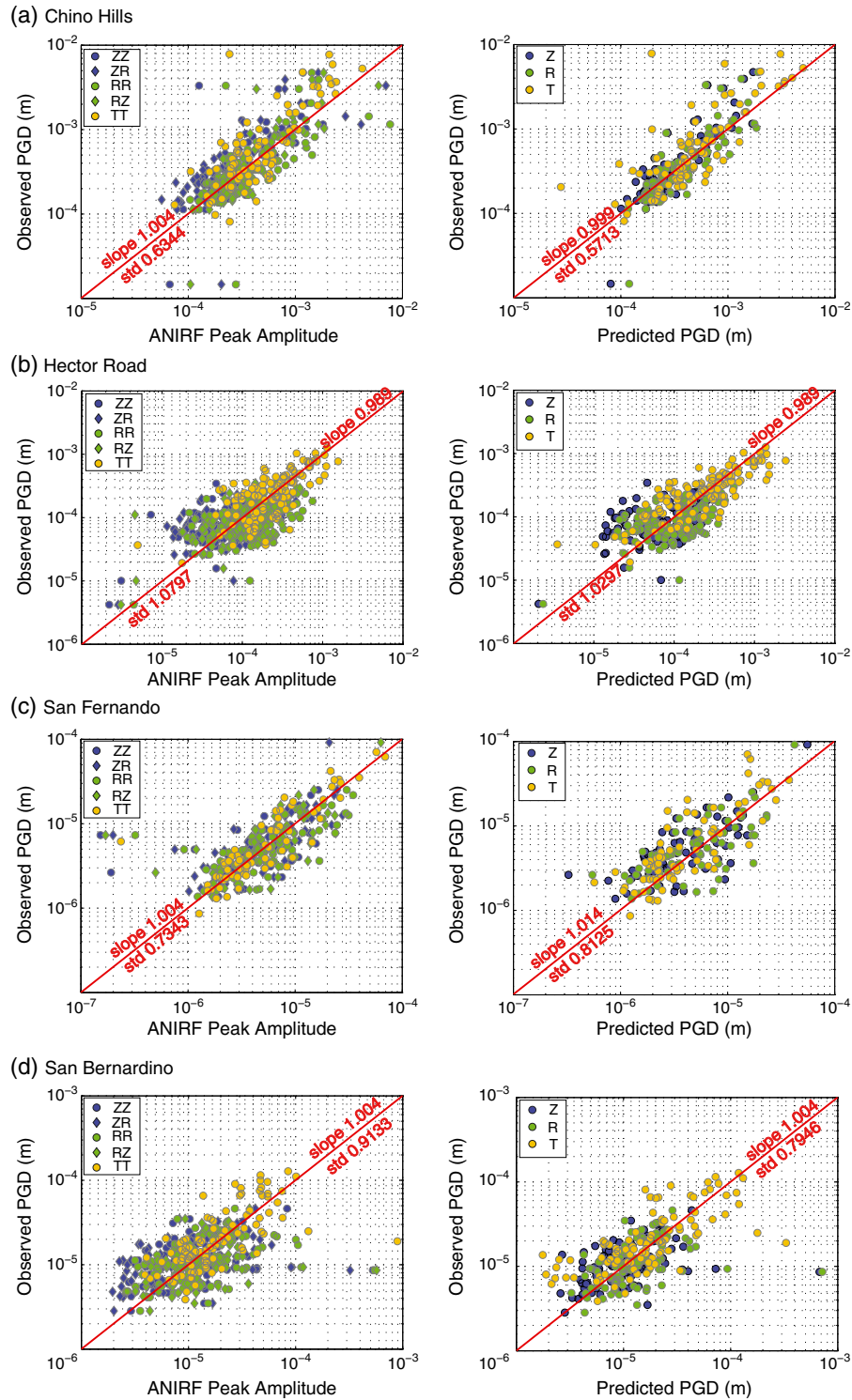


Figure 19. Observed and predicted peak amplitudes, filtered 4–10 s, at all stations for all four earthquakes: (a) Chino Hills, (b) Hector Road, (c) San Bernardino, and (d) San Fernando. We compare the vertical (blue), radial (green), and tangential (yellow) observed peak ground displacements PGD, with the ANIRF amplitudes (left panels) and with the VEA waveforms (right panels). The red line in each panel represents the the L1-linear regression, with each slope and standard deviation to ideal fit.

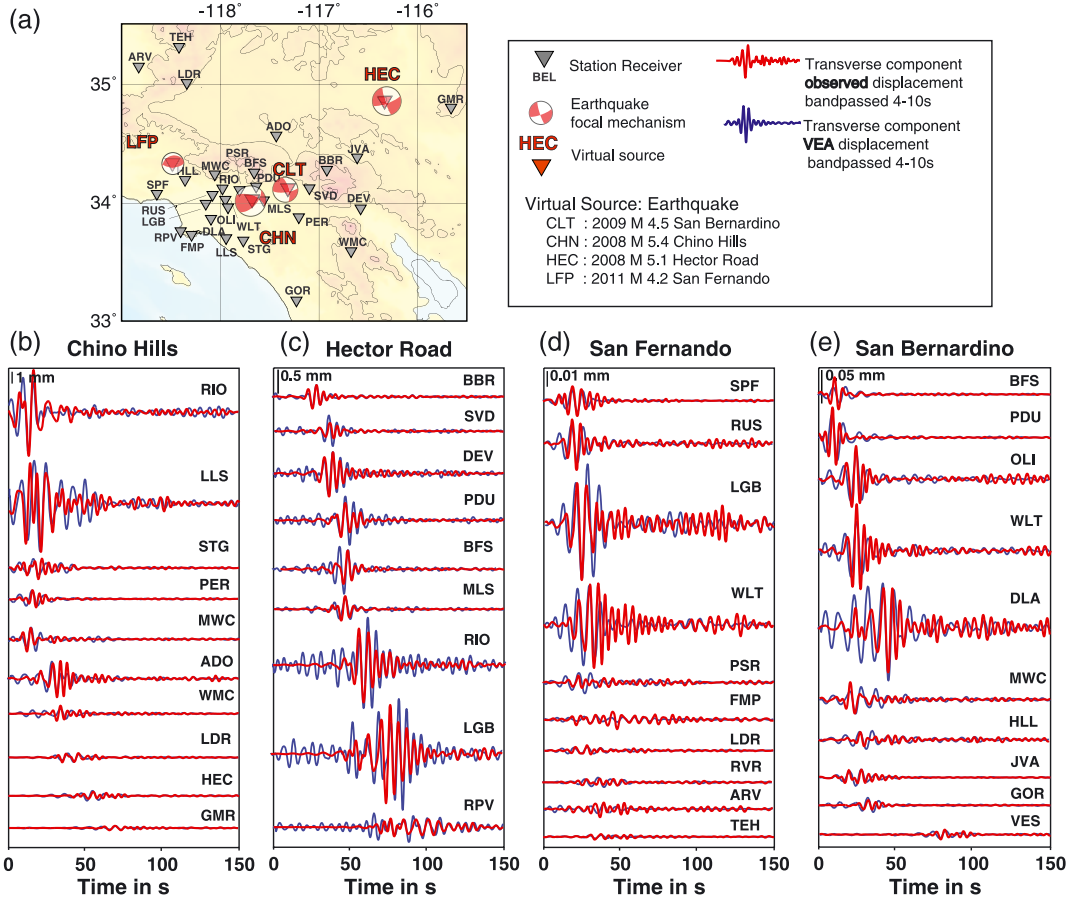


Figure 20. Same as Figure 3 but for depth and mechanism corrected response.

[46] For a better understanding of the accuracy of the VEA-predicted amplitudes, we show in Figure 19 peak amplitudes for the ANIRF, the virtual and real earthquake waveforms for the four events. There is a good match between the observed and predicted amplitudes, both for the impulse responses and the VEA waveforms. We estimate the best fitting linear trend, using L_1 norm minimization between the predicted and observed peak amplitudes. The slopes, indicated in Figure 19, are close to one for all cases. Despite the scatter in the amplitude, the standard deviation around the ideal fit is reduced with the Virtual Earthquake Approach. This clearly shows, again, that the VEA predicts, along with the correct phase, reliable amplitudes.

[47] There are several explanations for the isolated cases where we do not see a good match between observed and predicted displacements. Even though we retrieve the 3-D path effects from the ambient noise Green's functions, the 1-D approximation at the source ignores potential coupling between Love and Rayleigh waves that would occur locally for highly heterogeneous media. Moreover, the accuracy of the velocity profiles extracted at the virtual source locations is somewhat uncertain. Local noise may affect certain components of the ANIR tensor and hence the overall accuracy of the resulting VEA displacements. Finally, we used a far-field approximation for the expression of the surface-wave displacement, and this method requires modification for receivers within a wavelength of the epicenters.

In parallel to Figure 3, the summary Figure 20 demonstrates the potential of the VEA for ground motion prediction.

7. Conclusions

[48] We have shown that the virtual earthquake approach can successfully predict ground motion for moderate magnitude earthquakes. We first extracted the surface-wave impulse response, or Green tensor, using the ambient seismic field for each station pair, and then we corrected the impulse response for excitation depth and the double-couple response for a single forces to the response due to a dislocation. These conversions require accurate estimation of the surface-wave excitation at the earthquake source for which we used a new surface-wave eigenproblem solver (GESC) [Denolle *et al.*, 2012] that allowed us to calculate the displacement eigenfunctions and account for the locally complex vertical structure at the virtual source location.

[49] We validate this technique by reproducing far-field terms of the seismograms for the period band of 4–10 s. Given the approximations and sources of uncertainty, we consider the validation successful, in that the constructed seismograms match real earthquake records, in both phase and amplitude. For the four moderate earthquakes in southern California with diverse focal mechanisms, the virtual earthquake approach provides reliable prediction of the ground motion over this period band.

[50] Damaging earthquakes that are of most interest ($M > 6$) cannot be approximated as point sources at these periods. To predict ground motion for large seismic events, we need to consider the spatial variability of the Green's functions along extended ruptures. A deployment of seismometers along a fault of concern to record the ambient seismic field would provide the Green's functions needed to predict more complex displacement fields from scenario earthquakes.

[51] **Acknowledgments.** We thank Jesse Lawrence for his helpful discussions. This work was supported by NSF grant EAR-0943885 and by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR-0529922 and USGS Cooperative Agreement 07HQAG0008. The SCEC contribution number for this paper is 1642.

References

- Abrahamson, N. A., and K. M. Shedlock (1997), Overview, *Seismol. Res. Lett.*, *68*, 9–23, doi:10.1785/gssrl.68.1.9.
- Aki, K., and P. G. Richards (2002), *Quantitative Seismology*, Second Edition, University Science Books, Sausalito, Calif.
- Baltay, A. A., S. Ide, G. Prieto, and G. C. Beroza (2011), Variability in earthquakes stress drop and apparent stress, *Geophys. Res. Lett.*, *38*, L06,303, doi:10.1029/2011GL046698.
- Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007), Processing seismic ambient noise data to obtain reliable broad-band surface waves dispersion measurements, *Geophys. J. Int.*, *169*, 1239–1260, doi:10.1111/j.1365-246X.2007.03374.X.
- Brooks, L. A., J. Townend, P. Gerstoft, S. Bannister, and L. Carter (2009), Fundamental and higher-mode Rayleigh wave characteristics of ambient seismic noise in New Zealand, *Geophys. Res. Lett.*, *36*, L23,303, doi:10.1029/2009GL040434.
- Campillo, M., and A. Paul (2003), Long-range correlations in the diffuse seismic coda, *Science*, *299*, 547, doi:10.1126/science.1078551.
- Cupillard, P., L. Stehly, and B. Romanowicz (2011), The one-bit noise correlation: A theory based on the concepts of coherent and incoherent noise, *Geophys. J. Int.*, *184*, 1397–1414.
- Denolle, M. A., E. M. Dunham, and G. C. Beroza (2012), Solving the surface-wave eigenproblem with Chebyshev spectral collocation, *Bull. Seismol. Soc. Am.*, *102*, 1214–1223, doi:10.1785/0120110183.
- Gregersen, S. (1978), Possible mode conversion between Love and Rayleigh waves at a continental margin, *Geophys. J. Roy. Astr. Soc.*, *54*, 121–127, doi:10.1111/j.1365-246X.1978.tb06759.X.
- Hanks, T., and W. Thatcher (1972), A graphical representation of seismic source parameters, *J. Geophys. Res.*, *77*(23), 4393–4405, doi:10.1029/JB077i023p04393.
- Hartzell, S., A. Frankel, P. Liu, Y. Zeng, and S. Rahaman (2011), Model and parametric uncertainty in source-based kinematic models of earthquake ground motion, *Bull. Seismol. Soc. Am.*, *101*(5), 2431–2452, doi:10.1785/0120110028.
- Hauksson, E., et al. (2008), Preliminary report on the 29 July 2008 Mw 5.4 Chino Hills, Eastern Los Angeles Basin, California, earthquake sequence, *Seismol. Res. Lett.*, *79*(6), 855–866, doi:10.1785/gssrl.79.6.855.
- Hauksson, E., W. Yang, and P. M. Shearer (2012), Waveform relocated earthquake catalog for Southern California (1981 to 2011), *Submitted to: Bull. Seismol. Soc. Am.*, *102*(5), 2239–2244, doi:10.1785/0120120010.
- Herrmann, R. B., (1978), Computer programs in earthquake seismology, Volume 2: General programs, *Tech. Rep.*, Saint Louis University.
- Koper, K. D., K. Seats, and H. Benz (2011), On the composition of Earth's short-period seismic noise field, *Bull. Seismol. Soc. Am.*, *100*(2), 606–617, doi:10.1785/0120090120.
- Langston, C. A., S.-C. C. Chiu, Z. Lawrence, P. Bodin, and S. Horton (2009), Array observations of microseismic noise and the nature of H/V in the Mississippi Embayment, *Bull. Seismol. Soc. Am.*, *99*(5), 2893–2911, doi:10.1785/0120080189.
- Lawrence, J. F., and G. A. Prieto (2011), Attenuation tomography of the western United States from ambient seismic noise, *J. Geophys. Res.*, *116*, B06,302, doi:10.1029/2010JB007836.
- Lin, F.-C., M. P. Moschetti, and M. Ritzwoller (2008), Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.*, *2007*, doi:10.1111/j.1365-246X.2008.03720.X.
- Lobkis, O. I., and R. L. Weaver (2001), On the emergence of the Green's function in the correlations of a diffuse field, *J. Acoust. Soc. of Am.*, *110*, 3001–3017, doi:10.1121/1.1417528.
- Magistrale, H., S. Day, R. W. Clayton, and R. Graves (2000), The SCEC Southern California reference three-dimensional seismic velocity model version 2, *Bull. Seismol. Soc. Am.*, *90*, 6B, S65–S76.
- Nishida, K., H. Kawakatsu, and K. Obara (2008), Three-dimensional crustal S wave velocity structure in Japan using microseismic data recorded by Hi-net tiltmeters, *J. Geophys. Res.*, *113*, B10,302, doi:10.1029/2007JB005395.
- Olsen, K., S. M. Day, J. Minster, Y. Cui, A. Chourasia, M. Faerman, R. Moore, P. Maechling, and T. Jordan (2006), Strong shaking in Los Angeles expected from southern San Andreas earthquakes, *Geophys. Res. Lett.*, *33*, L07,305, doi:10.1029/2005GL025472.
- Olsen, K. B., et al. (2009), ShakeOut-D: Ground motion estimates using an ensemble of large earthquakes on the San Andreas fault with spontaneous rupture propagation, *Geophys. Res. Lett.*, *36*, L04,303, doi:10.1029/2008GL036832.
- Prieto, G. A., and G. C. Beroza (2008), Earthquake ground motion prediction using the ambient seismic field, *Geophys. Res. Lett.*, *35*, L14304, doi:10.1029/2008GL034428.
- Prieto, G. A., J. F. Lawrence, and G. C. Beroza (2009), An elastic Earth structure from the coherency of the ambient seismic field, *J. Geophys. Res.*, *114*, B07303, doi:10.1029/2008JB006067.
- Sabra, K. G., P. Gerstoft, P. Roux, W. A. Kuperman, and M. C. Fehler (2005), Surface wave tomography from microseisms in Southern California, *Geophys. Res. Lett.*, *32*, L14,311, doi:10.1029/2005GL023155.
- Sánchez-Sesma, F. J., and M. Campillo (2006), Retrieval of the Green's function from cross correlation: The canonical elastic problems, *Bull. Seismol. Soc. Am.*, *96*, 1182–1191, doi:10.1785/0120050181.
- Savage, J. C., W. Gan, and J. L. Svarc (2001), Strain accumulation and rotation in the Eastern California shear zone, *J. Geophys. Res.*, *106*(B10), 21,995–22,007, doi:10.1029/2000JB000127.
- Seats, K. J., J. F. Lawrence, and G. A. Prieto (2011), Improved ambient noise correlation functions using Welch's methods, *Geophys. J. Int.*, *188*, 513–523, doi:10.1111/j.1365-246X.2011.05263.X.
- Shapiro, N. M., and M. Campillo (2004), Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, *31*, L07,614, 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, doi:10.1126/science.1108339.
- Snieder, R. (2004), Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase, *Phys. Rev. E*, *69*(4), 046,610–1–7, doi:10.1103/PhysRevE.69.046610.
- Stehly, L., M. Campillo, and N. M. Shapiro (2006), A study of the seismic noise from its long-range correlation properties, *J. Geophys. Res.*, *111*, B10,306.
- Tanimoto, T., and C. Alvizuri (2006), Inversion of the HZ ratio of microseisms for S-wave velocity in the crust, *Geophys. J. Int.*, *165*, 323–335, doi:10.1111/j.1365-246X.2006.02905.X.
- Tanimoto, T., and L. Rivera (2005), Prograde Rayleigh wave particle motion, *Geophys. J. Int.*, *162*, 399–405, doi:10.1111/j.1365-246X.2005.02481.X.
- Toro, G. R., N. A. Abrahamson, and J. F. Schneider (1997), Model of strong ground motion from earthquakes in Central and Eastern North America: Best estimates and uncertainties, *Seismol. Res. Lett.*, *68*, 41–57, doi:10.1785/gssrl.68.1.41.
- Tsai, V. C. (2011), Understanding the amplitudes of noise correlation meas, *J. Geophys. Res.*, *116*, B09,311.
- Weaver, R. L., and O. I. Lobkis (2006), Diffuse fields in ultrasonics and seismology, *Geophysics*, *71*(4), S15–S19, doi:10.1190/1.2212247.
- Yao, H., R. D. van der Hilst, and J.-P. Montagner (2011), Heterogeneity and anisotropy of the lithosphere of SE Tibet surface-wave array tomography, *J. Geophys. Res.*, *115*, B12,307, 24 PP., doi:10.1029/2009JB007142.
- Yoshida, M. (2003), Rayleigh to Love wave conversion in a mountain root structure, *Bull. Earthq. Res. I. Tokyo*, *78*(1), 1–18.