Epicentral Location Based on Rayleigh Wave Empirical Green's Functions from Ambient Seismic Noise

M.P. Barmin, A.L. Levshin, Y. Yang, and M.H. Ritzwoller
Center for Imaging the Earth's Interior
Department of Physics
University of Colorado at Boulder, Boulder, CO 80309-0390

Abstract

A new method to locate the epicenter of regional seismic events is developed with strengths and limitations complementary to existing location methods. This new technique is based on applying Empirical Green's Functions (EGFs) for Rayleigh waves between 7 and 15 sec period that are determined by cross-correlation of ambient noise time-series recorded at pairs of seismic receivers. The important advantage of this method, in comparison with standard procedures based on use of body wave travel times, is that it does not rely on an earth model. Rather it is based on interpolating the EGFs to arbitrary hypothetical event locations. The method is tested by locating well known "Ground Truth" crustal events in the western US (magnitude 4 earthquakes in California, magnitude 3 earthquakes in Utah, and a mine collapse) as well as locating seismic stations as virtual events. In these applications, location errors average less than 1 km, but are expected to vary with event mechanism and depth. Numerical simulations show that the method optimizes when source depth is less than 1 km or more than 5 km, and when the source mechanism is nearly purely strike-slip, thrust, or normal.

1. Introduction

In principle, seismic surface waves possess valuable information about seismic source epicentral location and depth. Surface waves travel slower than body waves, they are excited by even relatively small crustal events, and the frequency dependence of their amplitudes depends strongly on event depth. Nevertheless, with a few exceptions (e.g., Ammon et al., 2001; Ekström, 2006), surface waves are not commonly employed to constrain epicentral location for small or moderate sized earthquakes, although long period surface waves are used on a regular basis to determine the centroid of large earthquakes (Dziewonski et al., 1981). This is not because standard procedures for determining the location of individual seismic events are perfect. Standard methods rely predominantly on body wave travel times and most depend on the ability to predict travel times through a reference earth model. Such models are often 1D and in some cases may not be tuned regionally. One-dimensional models such as IASPEI91 (Kennett and Engdahl, 1991), for example, are commonly used even for locations based exclusively on regional phases. One-dimensional models, however, do not capture the effects of 3D structural variations on travel times, which for regional phases such as Pg and Pn can be very strong. The inability to model 3D travel time effects produces location bias which, in the context of body wave location methods, can only be mitigated by using an accurate underlying 3D model (e.g., Shapiro and Ritzwoller, 2002), by applying empirical path corrections (e.g., Ritzwoller et al., 2003; Yang et al., 2004), by utilizing multiple event methods (e.g., Richards et al., 2006), or in the presence of exceptional azimuthal coverage (which commonly only occurs for large events). Irrespective of the shortcomings of traditional event location methods, the effect of 3D structural variations on surface waves is even more profound than that on body waves. This has led seismologists largely to ignore the potential for surface waves to improve the epicentral and hypocentral location of small to intermediate magnitude seismic events.

In this paper, we develop a new method and present pilot tests to locate regional seismic events based on surface wave information. This technique is based on applying Empirical Green's Functions (EGFs) for Rayleigh waves between 7 and 15 sec period, which are produced from ambient noise cross-correlation analysis (e.g., Shapiro & Campillo, 2004; Snieder, 2004; Shapiro et al., 2005; Bensen et al., 2007). In this period band, ambient noise is strong and shallow crustal earthquakes are energetic. Elastic EGFs are determined by cross-correlating ambient noise time-series recorded at pairs of stations. The important advantage of the application of ambient noise EGFs is that they do not depend on an underlying velocity reference model for detection and location. Rather, the empirical Green's functions contain full structural information without reference to a 3D model. Location errors are, therefore, expected to be largely unbiased by structural effects. In this paper, by

"location" we refer to epicentral location. Hypocentral location is the planned subject of a future contribution.

The idea of the new location method is straightforward, but it rests on the existence of at least a temporary dense local array, termed the *base stations*, that are deployed in the "region of interest" where interesting seismic events may occur. Second, it is also assumed that there is a more distant long duration (but potentially sparse) regional network of stations termed the *remote stations*. This lay-out is depicted schematically in Figure 1. Using what are now well established methods (e.g, Bensen et al., 2007), the EGFs from every base station to every remote station are computed. For this paper, we use only the EGFs from the vertical component of each station-pair, which are dominated by Rayleigh waves. The set of EGFs observed at the base stations in the neighborhood of a seismic event to a given remote station are used to produce an interpolated "Composite" EGF (CEGF) to the remote station for any hypothetical event location enclosed by the base stations. These Composite EGFs then can be used to locate events within the region of interest using waveform correlation with event (e.g., earthquake) records observed at the remote stations. In this paper, we use a simple grid search method for location, but the CEGFs could be used to compute partial derivatives with respect to the event coordinates, if desired. The key to the location method is the transformation of the observed EGFs to a hypothetical event location, which is described in detail in Section 3. The transformation improves the signal/noise ratio and minimizes the effect of waveform differences between event records and CEGFs.

There are two principal postive aspects of the method worth discussing at the outset. First, as mentioned above, the EGFs are empirical and, therefore, contain the effect of 3D structure. Second, the base stations do not have to be installed when the event of interest occurs. The base and remote stations must be emplaced simultaneously for some period of time in order to produce the set of EGFs between the base and remote stations, but the base stations need not be in place during the event. A temporary deployment of base stations will suffice to produce the EGFs for the grid of points within the footprint of the base stations to each of the remote stations. Thus, when an interesting event occurs, observations of the event are only needed at the remote stations, which are assumed to be "permanent". The longer the base stations can be operated the better, because longer time-series will produce higher signal-to-noise ratio (SNR) EGFs, although quite short deployments can produce surprisingly good EGFs at periods below about 15 sec (Bensen et al., 2007). The base stations can be installed even after an event of interest has occurred.

Section 2 describes the basic idea of the method and provides some explanatory illustrations for why the method succeeds. The formal description of the technique is given in Section 3. Examples of the application of the method are presented as tests in the following

sections. These include locating stations as virtual events as well as locating the Crandall Canyon Mine collapse in Utah on August 6, 2007, fourteen (magnitude 4 and 5) GT events in California, and four smaller (magnitude 3) events in Utah. Results are summarized in Tables 1-3. In Section 5 we discuss several factors that affect location accuracy, most notably source mechanism and depth.

2. Epicentral Location Method: The Idea

Consider, for example, locating an earthquake, such as the $m_b = 4.5$ event that occurred north of the San Francisco Bay area on May 12, 2006, Figure 2d. The idea of the method is to compare the event (earthquake, explosion, mine collapse, landslide, etc) records observed at the remote stations to the ambient noise Rayleigh wave Empirical Green's functions (EGFs) between the base stations and the remote stations. To do so, we systematically move the hypothesized location of the event until the match between the EGFs and the earthquake records is optimized. There are two principal wrinkles to the method as it currently exists. First, when we stack the individual EGFs over all of the base stations to produce a Composite EGF (CEGF) for each hypothesized event location, we currently ignore phase information by summarizing both the CEGFs and the event records with envelope functions. Second, the transformation or "interpolation" of the individual EGFs to form a CEGF is performed in the frequency-time domain. These details are discussed in the following section. The basic idea of the use of Rayleigh wave CEGFs determined from ambient noise to locate regional events is illustrated in Figures 2-4.

An example vertical component earthquake (event) record observed at station V04C in southern California is shown in Figure 2a together with the vertical-vertical EGF from a station near the epicenter (GASB) with station V04C in Figure 2b. Station GASB is about 100 km from the earthquake epicenter. Both time-series are band-pass filtered between 7 and 15 sec period and are dominated by Rayleigh waves. The earthquake and the EGF Rayleigh waveforms are similar, but are off-set in time because the epicentral distance for V04C is smaller than the inter-station distance V04C-GASB. The EGF can be shifted in time and deformed to match the epicentral distance, however, as seen in Figure 2c. The process of time shifting and deformation is described in detail in section 3.

The phase content of the earthquake and EGF records differ appreciably, however, because earthquakes impart an initial phase to surface waves that depends on the hypocentral depth and the source mechanism. For this reason, at this stage we ignore phase information and summarize both the EGFs and the earthquake records with their envelope functions, as seen in Figure 2a-c as dashed lines. There are, however, several base stations from the

EarthScope/USArray Transportable Array and the Berkeley Seismic Network near to this earthquake. Figure 3 shows the envelope functions of the EGFs for nine stations near to the earthquake to the remote station V04C compared with the envelope function of the earthquake record. Each of these EGFs were similarly transformed to the earthquake epicentral distance and the amplitude content of the EGFs and the earthquake are seen to align approximately, but there is some variability between them. To reduce the effect of the variability between the individual EGFs, we stack them to produce a Composite EGF (CEGF) for this particular source location and remote station. Typically, the CEGF agrees with the earthquake record better than the individual EGFs from the individual base stations.

The location method is based on comparing the envelope of the Composite EGFs for each hypothetical event location to the envelope of the earthquake record observed at all remote stations. An example of such a comparison is presented in the record section shown in Figure 4. Inspection of the fit between the Composite EGFs and the earthquake envelope functions reveals that 3-D structure similarly affects the CEGFs and the earthquake records. For example, consider the arrivals at stations HUMO, Q08A, and PHL, which are within several kilometers in epicentral distance from each other. The Rayleigh wave arrives much earlier at Q08A than at the other stations because of fast propagation through the Sierra Nevada. In fact, the fast arrivals in the record section all propagate in part through the Sierra Nevada. In addition, some of the envelope functions are significantly broader than others. Stations ELFS and T06C present examples for wave propagation across and oblique to the Great Valley sediments of central California which is characterized by very steep velocity dispersion between 7 and 15 sec period.

3. Epicentral Location Method: The Technique

3a. Notation, observational setting, and input data. The epicentral location of a seismic point source is specified by the time-space location vector $\mathbf{z}_0 = (t_0,\ \mathbf{x}_0)$, whose components are source time, latitude, and longitude. In this paper we ignore source depth, and focus on determining epicentral location. Let $\rho(\mathbf{x}, \mathbf{y})$ be the great-circle distance between the surface points $\mathbf{x}$ and $\mathbf{y}$. Consider the set of observation points (seismic stations) deployed in the circular vicinity of radius ρ_1 centered on a hypothetical source region. This set of stations may be split into two subsets: base stations $\mathbf{x}_i^B,\ i = 1, ..., n$, and remote stations $\mathbf{x}_j^R,\ j = 1, ..., m$. The base stations are deployed in the vicinity of the source location (local area) $\mathbf{x}_0$, $\rho(\mathbf{x}_0, \mathbf{x}_i^B) \leq \rho_2$, where we define ρ_2 to be not greater than 100-150 km. The radius ρ_1 of the entire set of receiving stations is taken here to be between 400 and 500 km. The remote stations are deployed in the annulus formed by radii ρ_1 and ρ_2, $\rho_2 < \rho(\mathbf{x}_0, \mathbf{x}_j^R) \leq \rho_1$.

This lay-out is depicted schematically in Figure 1.

To locate a particular event we require two data sets: vertical component event records for all remote stations and vertical-vertical component EGFs determined from ambient noise. (In principle, horizontal components also could be used for both Rayleigh and Love waves.) The event records observed at the remote stations, $w_j(t, \mathbf{x}_j^R)$, $j = 1, ..., m$, or simply $w_j(t)$, must include the Rayleigh waves and must be a continuous segment of data starting from some fixed time t_d. The EGFs, $r_{ij}(t, \mathbf{x}_i^B, \mathbf{x}_j^R)$, or simply $r_{ij}(t)$, must be present for some subset of station indices i and j, where $i = 1, ..., n$ refers to the base stations and $j = 1, ..., m$ to the remote stations. We use the symmetric representation of the EGFs, which is the half-sum of the positive and negative correlation lags. Thus, each EGF, $r_{ij}(t)$, starts from zero time. For simplicity, we will assume that EGFs exist for all combinations of the indices i and j, but this will generally not be the case in practice.

3.b Equalization of spectral amplitudes. As pointed out by Shapiro et al. (2004) and Bensen et al. (2007) as well as others, the procedure used to find the EGFs between a pair of stations does not preserve the amplitude spectrum of these functions. To overcome this problem, we equalize the amplitude spectra of the event record and the EGF without changing its phase spectrum. This could be done, of course, in a variety of ways; for example,

(1) by replacing the amplitude spectrum of the event with the amplitude spectrum of the EGF,

(2) by replacing the EGF amplitude spectrum with the event spectrum, or

(3) by replacing both the EGF and event amplitude spectra with some average of the two, for example, the square root of the product of the amplitude spectra of the event and the EGF.

The first and the third equalization techniques provide quite similar results. The second technique may amplify noise in the case of weak events. For this reason, we have chosen procedure (1).

After amplitude equalization the event and CEGF signals typically become quite similar, as can be seen in Figure 5c and Figure 5d. To select the surface waves from both signals, cosine tapering in time is applied to the beginning and ending parts of both signals.

3.c Frequency-time transformation. We transform the event record $w_j(t)$ into the frequency-time field $W_j(t, \omega_l)$, and the EGF $r_{ij}(t)$ into the frequency-time field $R_{ij}(t, \omega_l)$ by passing both equalized signals through a set of $l = 1, 2, ..., L$ narrow-band frequency domain

filters. For all event location examples in this paper $L = 32$, with filters within the bandpass period range from 7 to 15 sec.

Both frequency-time fields are formed by a set of envelopes obtained using narrow-band output signals and their Hilbert transforms. Thus, for $w_j(t)$,the field $W_j(t, \omega_l)$ is given by

$$W_j(t, \omega_l) = 2 \left| \int_0^\infty K(\omega - \omega_l) \tilde{w}_j(\omega) e^{i\omega t} d\omega \right| \tag{1}$$

where $\tilde{w}_j(\omega)$ is the Fourier transform of $w_j(t)$, i.e.,

$$\tilde{w}_j(\omega) = \int_{-\infty}^{\infty} w_j(t) e^{-i\omega t} dt$$

and $K(\omega - \omega_l)$ is the narrow-band Gaussian filter centered around center frequency ω_l

$$K(\omega - \omega_l) = e^{-\frac{(\omega - \omega_l)^2}{2\beta\omega_l^2}}. \tag{2}$$

The parameter β defines the width of the Gaussian filter and does not depend on ω. For more details see Levshin et al. (1989).

In the same way we introduce the frequency-time field $R_{ij}(t, \omega_l)$ for the cross-correlation time series $r_{ij}(t)$,

$$R_{ij}(t, \omega_l) = 2 \left| \int_0^\infty K(\omega - \omega_l) \tilde{r}_{ij}(\omega) e^{i\omega t} d\omega \right| \tag{3}$$

where $\tilde{r}_{ij}(\omega)$ is Fourier transform of $r_{ij}(t)$. Note that both W_j and R_{ij} are positive real functions of time t and central frequencies ω_l. To simplify discussion we will suppress index l in ω_l hereafter.

3.d Interpolation of frequency-time (FT) fields. For location purposes we need to interpolate the frequency-time fields $R_{ij}(t, \omega)$ to an arbitrary geographical point $\mathbf{x}$ inside the area covered by the base station deployment. The simplest way to do this is to transform $R_{ij}(t, \mathbf{x}, \omega)$ in the time domain as follows: $\tilde{t} = td/d_{ij}$, so

$$\tilde{R}_{ij}(t, \mathbf{x}, \omega) = R_{ij}(\tilde{t}, \mathbf{x}, \omega) \tag{4}$$

where $\tilde{R}_{ij}(t, \mathbf{x}, \omega)$ is the interpolated field at point $\mathbf{x}$, d is the distance between the base station i and the point $\mathbf{x}$ (or $d = \rho(\mathbf{x}_i, \mathbf{x})$), and d_{ij} is the distance between the base station i and the remote station j (or $d_{ij} = \rho(\mathbf{x}_i, \mathbf{x}_j)$).

This transformation produces the correct position for the maxima of $\tilde{R}_{ij}(t, \mathbf{x}, \omega)$ in time corresponding to the group travel times of the Rayleigh wave. More accurate interpolation can be performed by use of regional phase and group velocity maps obtained by tomographic

inversion of EGF data (e.g., Shapiro et al., 2005; Lin et al., 2008). This more rigorous approach is beyond the scope of this paper.

By adding $\tilde{R}_{ij}(t, \mathbf{x}, \omega)$ for different indices i and fixed j we obtain

$$C_j(t, \mathbf{x}, \omega) = \frac{1}{N} \sum_{i=1}^{n} \alpha_i \tilde{R}_{ij}(t, \mathbf{x}, \omega) \tag{5}$$

where $C_j(t, \mathbf{x}, \omega)$ is the FT-representation of the Composite EGF for a given point $\mathbf{x}$ within the base station deployment and remote station j, the coefficients α_i are weights that depend on the distance between each base station i and location $\mathbf{x}$, and N is the sum of the weights.

3.e Cross-correlation of FT-fields and evaluation of residuals. We assume that the two frequency-time fields, the composite EGF ($C_j(t, \mathbf{x}, \omega)$) and the event ($W_j(t, \omega)$), are similar for each remote station j but shifted in time due to the difference in position between epicenter $\mathbf{x}_0$ and a trial point $\mathbf{x}$, as well as the unknown difference between the origin time t_0 and the starting time of records t_d. The search for the unknowns $\mathbf{x}_0$ and t is performed by temporal cross-correlation of the two fields $C_j(t, \mathbf{x}, \omega)$ and $W_j(t, \omega)$ for all values of the remote station index j. Let us introduce the cross-correlation function $\mathcal{Q}$ between fields C_j and W_j as

$$\mathcal{Q}_j(\tau, \mathbf{x}) = \int_{\omega} \int_{t} C_j(t, \mathbf{x}, \omega) W_j(t + \tau, \omega) dt d\omega. \tag{6}$$

We define the value of τ at the maximum of the resulting cross-correlogram $\mathcal{Q}_j(\tau, \mathbf{x})$ as $\tilde{s}_j(\mathbf{x}) = \tau_{\mathbf{j}}^{\mathbf{max}}(\mathbf{x})$. The final expression for the residual between the observed and predicted travel times is as follows:

$$s_j(t, \mathbf{x}) = \tilde{s}_j(\mathbf{x}) - (t_d - t). \tag{7}$$

Here $t_d - t$ is the time shift due to the unknown origin time t_0. Repeating this procedure for all j, $j = 1, 2, ..., m$, we obtain the vector of travel time residuals for the given point $\mathbf{x}$. This vector should be considered as equivalent to the vector of travel time residuals in standard location procedures. The location problem is the minimization in space-time of the weighted squared norm of the residual vector given by

$$\min_{t, \mathbf{x}} \sum_{j} V_j^{-1} s_j^2(t, \mathbf{x}) \tag{8}$$

where V represents the $(m \times m)$ diagonal covariance matrix (weights) of errors in data measurements. Usually, the residuals are evaluated across the area around an event of interest, for times t limited by the lengths of both records.

3.f Grid search. In this paper we apply a simple brute force method to find the minimum of the functional (8). To do this we build a 2-D regular rectangular grid $\mathbf{x}_k, k = 1, 2, ..., K$,

in the region surrounding a preliminary solution $\mathbf{x}_0$ and compute the full residual function

$$f(\mathbf{x}_k) = \min_t \sum_j V_j^{-1} s_j^2(t, \mathbf{x}_k) \tag{9}$$

at each grid point. The minimum value of the function f at some grid point $\mathbf{x}_{k0}$ provides the solution. (Examples of full residual functions are plotted later in Figs. 10b and 11b.) Due to the ability to find the residual for any geographical point in the area covered by the base station deployment, it is possible to use more sophisticated techniques to search for the minimum, such as the conjugate gradient method (e.g., Atkinson, 1988).

3.g Confidence ellipsoid. Numerical trials demonstrate that the residuals $s_j(t, \mathbf{x})$ depend linearly on the spatial coordinates $\mathbf{x}$. Time t is also a linear term in the residual expression (7). Assuming that the linear part gives the principle contribution to the residuals and the rest is normally distributed noise, we follow the standard technique (Flinn, 1965; Jordan et al., 1981) to construct a 2-D confidence ellipsoid at the point ($\mathbf{x_{k_0}}$,t_0) with a minimum of (8). (Examples of 90% confidence ellipsoids are shown later in Figs. 10b and 11b.)

3.h Example of the Crandall Canyon Mine Collapse. The process described above is exemplified in Figure 6 for a pair of USArray Transportable Array stations, station P16A which is near the Crandall Canyon mine collapse and a remote station T14A about 300 km to the southwest. In this example, we take only a single base station, and do not stack over several EGFs for simplicity of presentation. The amplitude equalized event record after the mine collapse observed at station T14A is shown in Figure 5c, and the associated EGF for the station pair T14A-P16A is in Figure 5d. The event-station geometry is shown in Figure 6a. For demonstration purposes, we impose a small regular nine-point grid (Figures 6a,b) with a grid step of 20 km where the Ground Truth position of the mine collapse is located at the center of the grid, point 5. Grid points are numbered between 1 and 9, as shown in Figure 6b. For each grid point we transform the frequency-time representation of the EGF between stations T14A and P16A by means of equation (4). The frequency-time representations of the mine collapase and the EGF for grid point 9 are shown in Figure 6c,d and the cross-correlation between these representations (eqn. (6)) for the band between 7 and 15 sec period is shown in Figure 6e. Although the frequency-time representations are similar, they are shifted in time by about 9 seconds. Moving across the grid from points 1 to 9 changes the distances to the remote station T14A from 285 km to 340 km. The resulting time shifts in the cross-correlograms for all nine grid points as a function of the distance from the remote station T14A are shown in Figure 6f. Positions of the maxima of the cross-correlograms between the frequency-time fields correspond to time residuals. The time shift for grid point 5 has the value closest to zero, indicating that this point of the grid is the nearest to the epicenter of the mine collapse.

Note that the times of the maxima of these cross-correlograms depend linearly on the distance between the grid point and the remote station. The linearity of the residuals versus small spatial perturbations around a grid point is evidence that the squared norm of residual vector is a quadratic form of the spatial coordinates. This justifies the statement above that the task of location is reduced to searching for the vector $(t_0, \mathbf{x}_0)$ in equataion (8) that minimizes the squared sum of residuals for all remote stations.

4. Further Examples

4a. Locating Stations as Virtual Sources

Tests of the ability to locate seismic events are normally based on Ground Truth locations. However, such locations are themselves subject to uncertainties. To minimize this problem, we present a test to locate seismic stations as virtual sources. To do this, first we consider each USArray Transportable Array (TA) station in the western US to be a virtual earthquake and the estimated EGF to every other TA station to be the event records. We then attempt to locate the station based on those event records by applying EGFs between all of the other stations in the neighborhood of the virtual earthquake with the remote stations. Another advantage of this method is that it eliminates the effects of unknown source mechanism and depth on the location. We assess the effects of these variables in section 5.

Through August 2007, more than 100,000 EGFs for station pairs in California and surrounding states had become available (e.g., Moschetti et al., 2007; Lin et al., 2008). To obtain these data, 445 USArray Transportable Array (TA) and permanent stations were used as identified with colored circles in Figure 7a. The results of the station location test are presented in Figure 7a,b. The following parameters were used in this test: the grid size is a relatively coarse 500 m, $\rho_1 = 400$ km, $\rho_2 = 100$ km. That is, all stations within 100 km of each virtual event (i.e., each station) are used as base stations and the remote stations are taken from 100 km to 400 km from each virtual event. 61% of all events have location errors less than 0.5 km, and 28% have errors ranging from 0.5 to 1 km. Seven stations (1.5% of 445) are located with an error greater than 3 km, and all have exceptionally poor azimuthal coverage.

To evaluate how the number of remote stations and their azimuthal distributions influence location accuracy we focus tests on the TA station S10A in Central Nevada as a virtual earthquake (Figure 8a). Using a jack-knife procedure, we sytematically reduced the number of remote stations participating in the location from 126 to 10, while preserving a more-or-less homogeneous azimuthal coverage. This procedure included 10 runs for each given

number of remote stations, but with different ensembles of stations selected. In order to remove the effect of grid increment, we performed locations on a very fine grid (25 m). The average of the mislocations for the 12 runs is shown in Figure 8b as a function of the number of stations, presented with standard deviations. Note that the expected location error is less than 600 m even when the number of remote stations reduces to 10. Mislocation asymptotically approaches about 250 m, which we take as the practical limit for the current location algorithm when source mechanism and depth are eliminated. As discussed in section 5, this residual error appears to be determined largely by time discretization (1 sps here).

We also systematically investigated the effect of open azimuth by eliminating remote stations in azimuthal sectors that vary in size from 30° to 240°. The azimuthal position of each sector of a given size was changed 12 times by shifting the central azimuth of the sector by 30°. The average mislocations for 12 runs as a function of open azimuth plotted with the standard deviations are shown in Figure 8c. As with most location methods, open azimuth is more important than number of remote stations for location accuracy. In particular, as open azimuth grows above about 240°, location accuracy degrades rapidly. This explains the occurence of large location errors near the Pacific coast in Figure 7a. As long as open azimuth is less than 240°, however, location errors are expected to be less than 500 m and approach 250 m as open azimuth shrinks.

4.b Crandall Canyon Mine Collapse

The second test is to locate the magnitude m_b =3.9 mine collapse that occurred on Monday August 6, 2007 at 08:48:40 UTC at the Crandall Canyon mine, Utah (Pechmann et al., 2007). The US Geological Survey (USGS) and the University of Utah Seismograph Stations (UUSS) determined that this event does not have the characteristics of a typical, naturally occurring earthquake, but was caused by the collapse of the coal mine, trapping several minders underground. The schema of the mine is shown in Figure 9, which has been taken from Pechmann et al. (2007). USGS and UUSS summary information about this event is presented in Table 1.

When this event occurred, the USArray TA was situated in Utah. To locate the event we used seven stations as base stations and 66 remote stations (instead of the 33 used by the USGS) affiliated with different networks: USGS, University of Utah, and the TA (Figure 10a). The other location parameters were: grid size 100 m, ρ_1 = 400 km, ρ_2 = 100 km. The location results are as follows: 39.465N (USGS 39.465N), 111.224W (USGS 111.237W). The origin time shift is equal to -0.26 seconds. Confidence ellipsoid parameters (Figure 10b) are: a = 0.52 km (major semiaxis), and b = 0.43 km (minor semiaxis). Our

location is shifted relative to the preliminary USGS location to the east by 1.5 km, which is within the footprint of the mine and in better agreement with the location of the mine collapse and the latest UUSS location (Pechmann et al., 2008).

4.c Ground Truth Earthquakes in California

Table 2 presents Ground Truth (GT) locations and our locations for 14 California earthquakes with m_b between 4.5 and 5.1 that occurred in 2005 and 2006. The location parameters used in this test were: grid size 500 m, $\rho_1 = 400$ km, $\rho_2 = 100$ km. An example for event #1 is shown in Figure 11. The GT locations for events in southern California are determined by CalTech and for events in northern California by the Northern California Earthquake Center in Menlo Park. Both organizations use a combination of broad-band and short period instruments and 1-D models calibrated for their region. CalTech uses both P and S phases, whereas only P phases are used in northern California. The accuracy of locations for these events is believed to be better than 1 km (Egill Hauksson, James Dewey, Bob Engdahl, personal communication). If we accept GT locations as exact, differences with them are considered to be "errors" in distance and origin time. Because the GT locations are not exact, however, the errors we report here are conservative.

The location and origin time errors are listed in the two last columns of Table 2. The average error for the 14 events is 1.2 km, and the average error in time is 0.04 s. At the 90% confidence level, our locations are better than 2 km. These error estimates are larger than those reported for the station location (virtual earthquake) experiment and the location of the mine collapse because source mechanism and depth contribute group time shifts to the earthquake records that can bias the location using our method. This effect is discussed further in section 5.

4.d Smaller Ground Truth Earthquakes in Utah

To investigate whether our method is applicable to events smaller than about magnitude 4.0, we applied the same approach to the location of weaker earthquakes ($3 < m_b < 4.0$) that occurred in Utah in 2007 and 2008. The location parameters used in this test were: grid size 200 m, $\rho_1 = 400$ km, $\rho_2 = 100$ km. These events are well located and characterized by the University of Utah (UUSS) and by the Department of Earth and Atmospheric Sciences at Saint Louis University (SLU) (J. C. Pechmann, personal comminication). Information about these events is given in Table 3. The configuration of stations is very similar to that shown in Figure 10a. Average mislocation is less than 500 m, which indicates the method

appears to work well with little modification for events with magnitudes between 3.0 and 4.0. The smaller mislocation for the Utah events compared with the California events may be due to source depth, as these smaller events may be either shallower or deeper than the earthquakes in California. The effect of event depth on mislocation is discussed further in section 5.The method may be able to be improved further by tuning the band width of the method to the magnitude of the event.

5. Discussion

We have described a method to locate seismic epicenters based on cross-correlation of frequency-time representations of Composite Empirical Green's Functions with observed event seismograms in the period band between 7 and 15 s. The accuracy of the location depends on a number of factors that we discuss further here.

5.a Technical Factors Affecting Location Accuracy

There are a number of technical variables that affect the accuracy of locations. We have used EGFs and event seismograms with a time sampling rate of 1 sps, which leads to errors up to ± 0.5 s in timing of the resulting correlograms. Increasing sampling rates to 10 sps improves the location of well located events but has little effect for events mislocated by 1 km or worse. The size of the spatial grid spacing also affects accuracy, of course. The examples presented here have different spatial grids and the spacing of grid nodes is noted in each case. The current technique to interpolate the frequency-time representations of the wavefields (eq. (4)) assumes the constancy of surface wave group velocities in the frequency range considered. Significant frequency dependence of group velocities may distort the frequency-time representations. Corrections for the frequency dependence of dispersion may be introduced before the cross-correlation, however. This is particularly important in regions with strong lateral variability of crustal structure, the effects of which can be decreased by using phase and group velocity maps, which now exist in the western US (e.g., Moschetti et al., 2007; Lin et al., 2008) and many other regions.

5.b Event Source Mechanism and Depth

The most significant physical effects on location accuracy are source mechanism and depth. EGFs determined from ambient noise correspond to surface sources with a predominant source mechanism close to a vertical force. The event seismograms, however, may result

from different types of sources at different depths: e.g., earthquakes within Earth's crust, volcanic explosions, meteoritic impacts, explosions, mine collapses, or other human-related excitation processes. The radiation patterns of earthquakes depend on the source mechanism and depth as well as local structure, and are quite variable. Because the envelopes of event records and EGFs are what we compare in the frequency-time domain to determine the epicenter, it is group time shift (rather than source phase) that is the relevant effect of event mechanism and depth. Levshin et al. (1999) discuss group time shifts expected for different types of events in detail, and we only summarize briefly here.

According to Aki & Richards (1980) the azimuthal dependence of the displacement spectrum for a Rayleigh wave at a given frequency excited by a point double couple source at the depth h in a laterally homogeneous Earth is given by the complex function E:

$$\begin{aligned} E(\omega,\phi,h) &= \left.\frac{dr_2}{dz}\right|_h M_{zz} + k(\omega) r_1(h)[M_{xx}\cos^2\ \phi + 2M_{xy}\sin\ \phi\cos\ \phi + M_{yy} sin^2\ \phi] \\ &\quad + i\left[\left.\frac{dr_1}{dz}\right|_h - k(\omega) r_2(h)\right][M_{xz}\cos\ \phi + M_{yz}\sin\ \phi] \end{aligned} \tag{10}$$

where $k(\omega)$ is wave number which depends on frequency ω, $M_{zz},\ M_{xx},\ M_{yy},\ M_{xy},\ M_{xz},\ M_{yz}$ are moment tensor components, r_1 and r_2 are horizontal and vertical eigenfunctions, ϕ is azimuth taken clockwise from North. The modulus $|E|$ of the complex function E represents the source amplitude radiation pattern and the argument $\varphi = arg(E)$ represents the source phase delay. Both $|E|$ and φ are real functions that depend on ω, ϕ, and h. The imaginary part of equation (10) is proportional to the tangential component of stress and equals zero for the the surface $h = 0$. As the phase of the radiation function (source phase delay) varies with azimuth, it produces different phase shifts in the spectra of the event seismograms. However, our location method measures source group (not phase) time delays, δt_U, which are frequency derivatives of the source phase delays and will vary from station to station depending on azimuth. The group time shift may be expressed explicitly as

$$\delta t_U = \frac{\partial \varphi}{\partial \omega} = \mathrm{Im}\left\{\frac{E'_\omega E^*}{|E|^2}\right\}, \tag{11}$$

where the prime represents a frequency derivative. Substitution of E defined by equation (10) into (11) shows that the source group time delay function can be rewritten as

$$\delta t_U = \frac{A_1\cos(\phi+\alpha_1) + A_3\cos(3\phi+\alpha_3)}{1 + A_2\cos(2\phi+\alpha_2) + A_4\cos(4\phi+\alpha_4)} \tag{12}$$

where A_i and $\alpha_i(\ i = 1, ..., 4)$ are rather complicated functions of ω and h whose details are not needed here.

Several conclusions can be drawn immediately from these equations.

(1) Levshin et al. (1999) showed that group time delays are generally much smaller than the source phase delays. This is because the phase delays tend to change relatively slowly with frequency and is one of the reasons we have ignored phase information in our location method. Nevertheless, group time delays are not negligible and can bias locations.

(2) It is evident from equation (12) that group time shifts are anti-symmetric in azimuth ϕ; i.e., changing azimuth by 180° changes only the sign of the delay. Group time shifts differing in azimuth by 180°, therefore, will constructively interfere. Thus, averaging residuals over azimuth while searching for the source position only partially suppresses the effect of group time delays on location.

(3) For events at the free surface, group time delay are zero for any source mechanism due to zero amplitude of tangential stresses near the free surface. For very shallow depth events (less than 1 km) group time delays will be very small. Thus, shallow events will have very small location bias due to source mechanism. This is probably why the Crandall Canyon Mine collapse and small event #3 in Utah (Table 3) were located so well.

(4) For some mechanisms, such as a pure normal, pure thrust, or a pure strike-slip, group times delays are exactly zero. This can be seen from equation 10. For pure normal and thrust faults (45° dip and ±90° slip) the real part of E is zero. For pure strike-slip faults (90° dip and 0° slip), the imaginary part of E (traction at the surface) is zero. In these cases, $\arg(E)$ and the frequency derivative of E will be zero, so the group time delay will be zero. As discussed by Levshin et al. (1999), the source mechanism for most earthquakes is relatively close to the pure mechanisms.

These observations demonstrate that our location method will be biased minimally by events that are near the surface or that possess nearly strike-slip, thrust, or normal faulting mechanisms. It remains to be determined quantitatively how bias sets on as events deepen and mechanisms diverge from these pure types. Numerical simulations are needed.

Numerical simulations of group time delays for events between depths of 2 and 20 km show significant variability with depth, azimuth from the source, and period. Because source mechanisms at these depths typically are close to pure strike-slip, thrust, or normal fault mechanisms, they create very small group delays (from 0 to ±1 s) at periods between 6 and 15 s. More complicated mechanisms, however, can create delays up to $\pm(3-4)$ s, depending on azimuth, depth and period.

To quantify group time shifts for a variety of source mechanisms and depths for crustal events we computed simulated Green's functions for a laterally homogeneous model typical of central Nevada (Shapiro and Ritzwoller, 2002) using a surface wave 1-D synthetic code

(Herrmann, 1979; Levshin et al., 1989). The base and remote stations used are those shown in Figure 8a. The simulated Green's functions correspond to a vertical force acting at the Earth's surface. We also calculated simulated event seismograms at remote stations for four events with a fixed geographical position within the base network. These events have the source mechanisms shown in the upper part of Figure 12. These mechanisms are similar to pure normal (red) and thrust (green) faults, a vertical thrust fault (navy blue), and a strike-slip fault (light blue), but the corresponding angles characterizing the double couple mechanism (dip and rake) are $15^\circ - 20^\circ$ different from the pure mechanisms, which is sufficient to produce a significant group time shift. The polar diagrams in Figure 12a show the azimuthal distribution of group time residuals for the corresponding mechanisms with source depths of 2, 5, and 10 km. Note that our location method is based on cross-correlating frequency-time diagrams between 7 and 15 s period. Thus, the results in Fig. 12a are frequency averaged.

Applying our location technique to the simulated data we find errors in source locations as a function of source depth for the four mechanisms, as seen in Figure 12b. Shallow focus events (depths less than about 1 km) produce very small group time shifts and, therefore, minimal location bias. Similarly, events deeper than about 7 km have location errors smaller than 500 m, at least given the period band of this study (7-15 s). Generally, events with mechanisms similar to thrust or normal faults will have a larger mislocation than those similar to strike-strike. Therefore, non-strike-slip earthquakes with depths between 1 and 5 km provide the greatest challenge for our method.

Our current method is integrated over frequency and it is possible that in the location process the observation of the azimuthal distribution of the group time residuals as a function of period may be used to provide information about the accuracy of the epicentral location or provide information about source depth. Figures 13 - 15 present polar diagrams of theoretical group time shifts as a function of period for three event depths for the same mechanisms as in Figure 12. These results indicate that the azimuthal patterns of group time shifts tend not to change appreciably with period, but the amplitudes can vary strongly. Thus, in the location procedure, the observation of frequency-independent group time residuals is evidence that the epicenter of the event has not been strongly biased by source mechanism. Conversely, the observation of strong frequency-dependence is evidence that the epicenter may be biased, but that the event probably occurred between depths of about 1 and 5 km.

6. Conclusions

The method to locate the epicenter of regional seismic events described here has several features that make it a useful addition to existing location methods.

- Its accuracy does not require knowledge of Earth structure.
- It works for weak events where the detection of body wave phases may be problematic.
- The empirical Green's functions (EGFs) computed during a temporary deployment of a base network (such as the USArray Transportable Array or PASSCAL deployments) may be applied to events that occurred earlier or later than the temporary network using permanent remote stations even when the temporary stations are not present.

In addition, the method has several evident limitations.

- It requires a relatively dense base network within an area of interest for at least a few months.
- It provides the best accuracy when the frequency derivative of the source phase is relatively small; for example, when the source mechanism is a vertical force or a center of compression, when the source depth is less than 1 km or more than 5 km, and when the source mechanism is nearly purely strike-slip, thrust, or normal.

The technique described here is the first attempt at such a nonstandard approach to locate regional events using ambient noise EGFs. Among possible directions for improvement are the incorporation of phase information, considering not only vertical but also transverse components of the EGFs which are dominated by Love waves, the transformation (interpolation) of EGFs into Composite EGFs based on tomographic group and phase velocity maps for the region of interest (e.g., Moschetti et al., 2007; Lin et al., 2008), and the application of information about the frequency-dependence of the residuals. Many parameters define the technical instantiation of the method, such as the frequency band, group velocity window, base/remote station geometry, the spatial grid, and the time series sampling rate. These parameters can be varied systematically to tune the method to different observational settings and for events of different magnitude and character.

Acknowledgments

The authors gratefully acknowledge helpful converstations with Steven Myers and E. Robert Engdahl, and thank James Dewey for providing the list of Ground Truth earthquakes in

California and J.C. Pechmann for providing similar information about the smaller events in Utah. This research was supported by DoE/NNSA contract DE-AC52-09NA29326. The facilities of the IRIS Data Management System, and specifically the IRIS Data Management Center, were used to access the waveform and metadata required in this study. The IRIS DMS is funded by the National Science Foundation and specifically the GEO Directorate through the Instrumentation and Facilities Program of the National Science Foundation under Cooperative Agreement EAR-0552316.

References

Ammon, C. J., M. L. Pyle, G. E. Randall, and A. A. Velasco, 2001. Refining faulting parameters and depth estimates for earthquakes in eastern Asia, *Proceedings of the 23rd Seismic Research Review: Worldwide Monitoring of Nuclear Explosions*, October 2-5, 483-489, 2001.

Atkinson, K. A., 1988. *An introduction to numerical analysis (2nd ed.)*, Section 8.9, John Wiley and Sons.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang, 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.*, **169(3)**, 1239-1260, DOI: 10.1111/j.1365-246X.2007.03374.x.

Dziewonski, A. M., T.-A. Chou, and J. H. Woodhouse, 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, **86**, 2825 - 2852.

Ekström G., 2006. Global detection and location of seismic sources by using surface waves, *Bull. Seismol. Soc. Amer.*, **96**, 4A, 1201-1212.

Flinn, E. A., 1965. Confidence regions and error determinations for seismic event locations. *Rev. Geophys*, **3**, 157-185.

Herrmann, R. B., 1978. *Computer Programs in Earthquake Seismology*, **1**, iss. by Dept. Earth and Atmosph. Sciences, Saint Lois University.

Jordan, T. H. and K. A. Swerdrup, 1981. Teleseismic location techniques and their application to earthquake clusters in the South-Central Pacific. *Bull. Seismol. Soc. Amer.*, **71**, 4, 1105-1130.

Kennett, B. L. N., and E. R. Engdahl, 1991. Travel times for global earthquake location and phase identification, *Geophys. J. Int.*, **105**, 429-465.

Levshin, A.L., T. B. Yanovskaya, A. V. Lander, B. G. Bukchin, M. P. Barmin, L. I. Ratnikova, and E. N. Its, 1989. *Seismic Surface Waves in Laterally Inhomogeneous Earth*, (V. I. Keilis-Borok, Ed.), Kluwer Academic Publishers. Dordrecht/Boston/London.

Levshin, A.L., M. H. Ritzwoller, and J. S. Resovsky, 1999. Source effects on surface wave group travel times and group velocity maps, *Phys. Earth Planet. Inter.*, **115**, 293-312.

Lin, F., M. P. Moschetti, and M. H. Ritzwoller, 2008. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.*, **173(1)**, 281-298, doi:10.1111/j1365-246X.2008.03720.x.

Moschetti, M. P., M. H. Ritzwoller, and N. M. Shapiro, 2007. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh wave group velocity maps, *Geochem., Geophys., Geosys.*, **8**, Q08010, doi:10.1029/2007GC001655.

Pechmann, J. C., W. I. Arabasz, K. L. Pankow, R. Burlacu, and M. K. McCarter, 2007. Seismological report on the 6 Aug 2007 Crandall Canyon Mine Collapse in Utah, *Seismol. Res. Lett.*, **79**,5, 620-636.

Richards, P. G., F. Waldhauser, D. Schaff, and W. Y. Kim, 2006. The applicability of modern methods of earthquake location, *Pure and Appl. Geoph.*, **163**, 2-3, 351-372.

Ritzwoller, M. H., N. M. Shapiro, A. L. Levshin, E. A. Bergman, and E. A. Engdahl, 2003. The ability of global models to locate regional events, *J. Geoph. Res.*, **108(B7)**, 2353, doi:10.1029/2002JB002167, 2003.

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

Shapiro, N. M. and M. H. Ritzwoller, 2002. Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle, *Geophys. J. Int.*, **151**, 88-105.

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

Snieder, R., 2004. Extracting the Green s function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev.* E, **69**, 046610.

Yang, X., I. Bondar, J. Bhattacharyya, M. Ritzwoller, N. Shapiro, M. Antolik, G. Ekstrom, H. Israelsson, and K. McLaughlin, 2004. Validation of regional and teleseismic travel-time models by relocation of GT events, *Bull. Seis. Soc. Am.*, **94(3)**, 897-919.

Table 1. USGS, UUSS and CU Event Parameters for the Crandall Canyon Mine Collapse on August 6, 200.7

Parameters	USGS	UUSS	CU†
Magnitude	4.2	3.9	-
Origin Time, UTC	08:48:40	8:48:40	-
Latitude, N	39.4650	39.4675	39.465
Longitude, W	111.2370	111.2248	111.224
Depth, km	1.6	0.6	-

CU – University of Colorado, this study. UUSS – University of Utah Seismograph Stations.

Table 2. California Event Information and Location Errors.

event #	date (y/m/d)	time (h:m:s)	source†	lat	lon	m_b	Error (km)	Error (sec)
1	2005/04/16	19:18:13.00	P	35.027	-119.178	4.9	0.72	-0.16
2	2005/05/16	07:24:37.50	NC	35.929	-120.477	4.3	0.90	0.02
3	2005/06/12	15:41:46.54	P	33.529	-116.572	5.1	1.08	-0.15
4	2005/06/16	20:53:26.02	P	34.058	-117.010	4.8	0.72	-0.21
5	2005/06/26	18:45:57.82	NC	39.305	-120.093	4.6	1.02	-0.03
6	2005/08/31	22:47:45.64	P	33.165	-115.635	4.5	2.18	0.20
7	2005/08/31	22:50:24.03	P	33.172	-115.610	4.9	1.78	0.07
8	2005/08/31	23:32:11.04	P	33.190	-115.602	4.5	2.86	0.02
9	2005/09/02	01:27:19.81	P	33.160	-115.637	4.9	2.17	-0.04
10	2005/09/22	20:24:48.62	P	35.043	-119.013	4.8	1.07	0.00
11	2005/10/02	13:48:09.45	NC	35.651	-121.087	4.5	0.45	-0.08
12	2006/05/12	10:37:29.31	NC	38.816	-122.816	4.5	0.43	-0.09
13	2006/06/15	12:24:51.11	NC	37.102	-121.492	4.5	0.71	0.00
14	2006/08/03	03:08:12.86	NC	38.364	-122.589	4.6	0.71	-0.14

†: P – Pasadena, NC – Northern California.

Table 3. Utah Event Information and Location Errors.

event #	date (y/m/d)	time (h:m:s)	lat	lon	depth	m_b	Error (km)	Error (sec)
1	2007/08/18	13:16:13.46	38.070	-113.323	9.0†	3.65†	0.42	0.40
2	2007/11/05	21:48:00.61	39.346	-111.648	15.0†	3.76†	0.41	0.10
3	2008/02/01	21:36:54.23	41.809	-112.218	0.18	3.59	0.40	2.0
4	2008/10/12	3:26:01.41	41.690	-111.143	8.5	3.37	0.47	0.63

†: GT location and time is from University of Utah, depth and magnitude are from Saint Louis University.

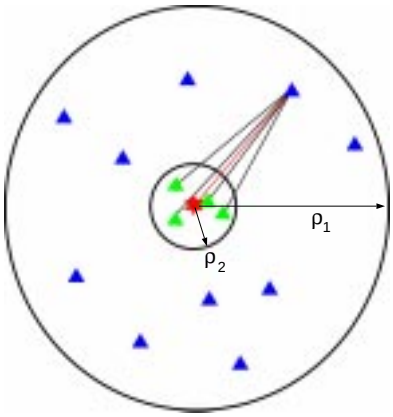

Figure 1: Schematic observational setting. A relatively dense set of temporary "base stations" (green triangles) encompasses the source region of interest (event is denoted by the red star). A sparse set of permanent "remote stations" (blue triangles) lies farther from the epicentral region.

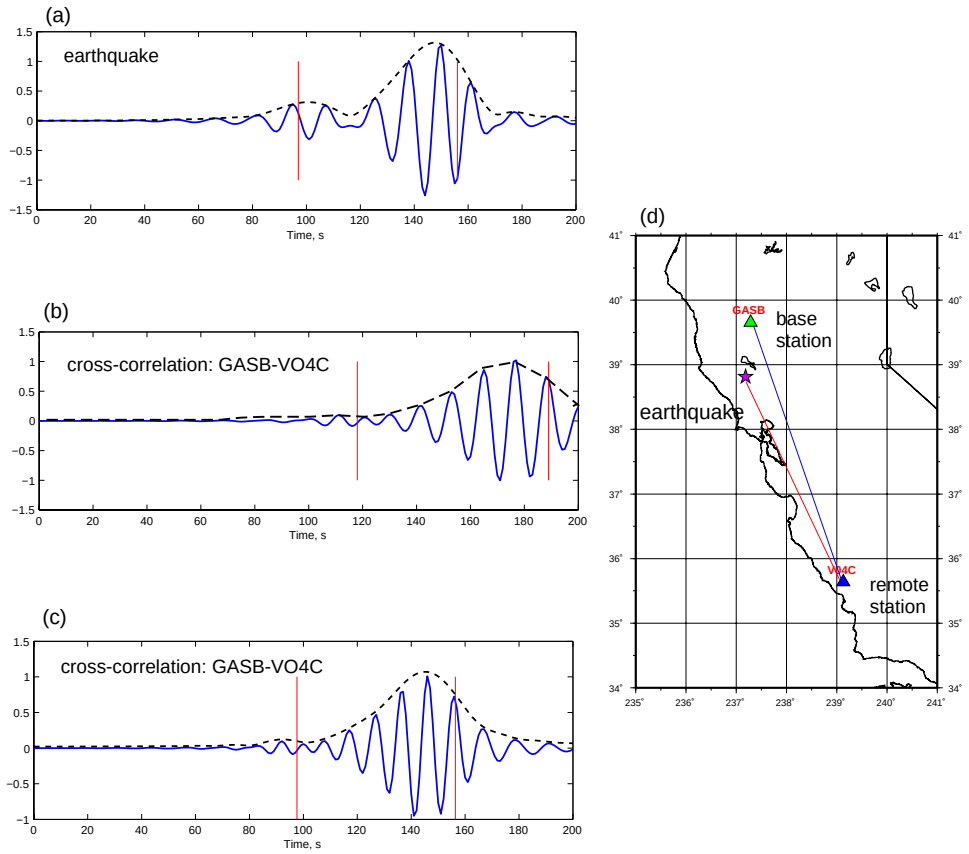

Figure 2: Comparison of an earthquake record with an Empirical Green's Function determined from ambient noise. (a) Earthquake record at station V04C in southern California band-pass filtered between 7 and 15 sec period. (Earthquake 12 from Table 3: May 12, 2006; 38.816N, 122.816E; $m_b = 4.5$, located by the Northern California Earthquake Center). (b) EGF between stations GASB (north of the earthquake) and V04C, similarly band-passed. Red lines mark times corresponding to group velocities of 2.5 and 4 km/sec. (c) Same as (b), but the EGF has been shifted and deformed to correspond to the earthquake location. Envelope functions are shown in (a), (b), and (c) with dashed lines. (d) Map showing earthquake and station locations.

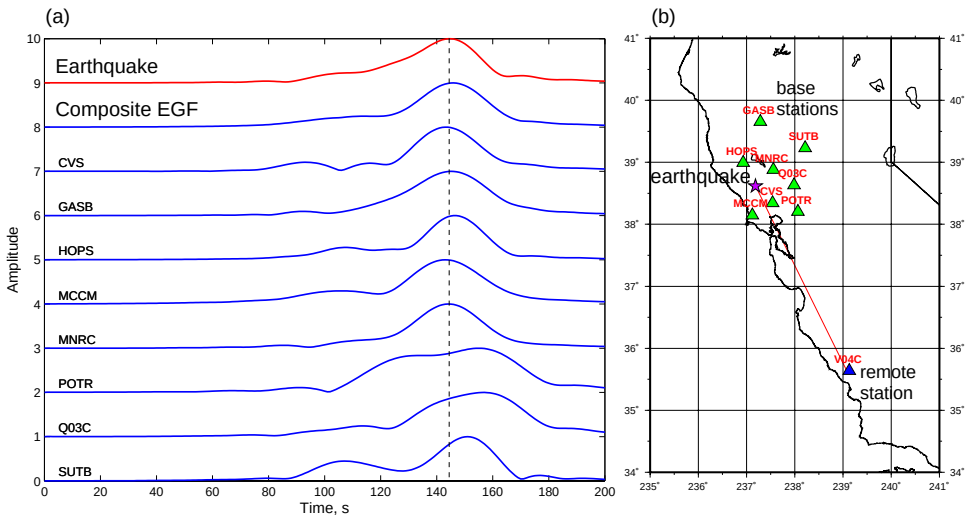

Figure 3: Comparison of ambient noise Empirical Green's Functions for the set of base stations. Similar to Fig. 2, but only the envelopes of the EGFs for the 8 base stations are shown in (a). The envelope of the earthquake record is shown in red at top and the envelopes of the Composite EGF (sum of the individual EGFs transformed to the epicentral location) is on the second line.

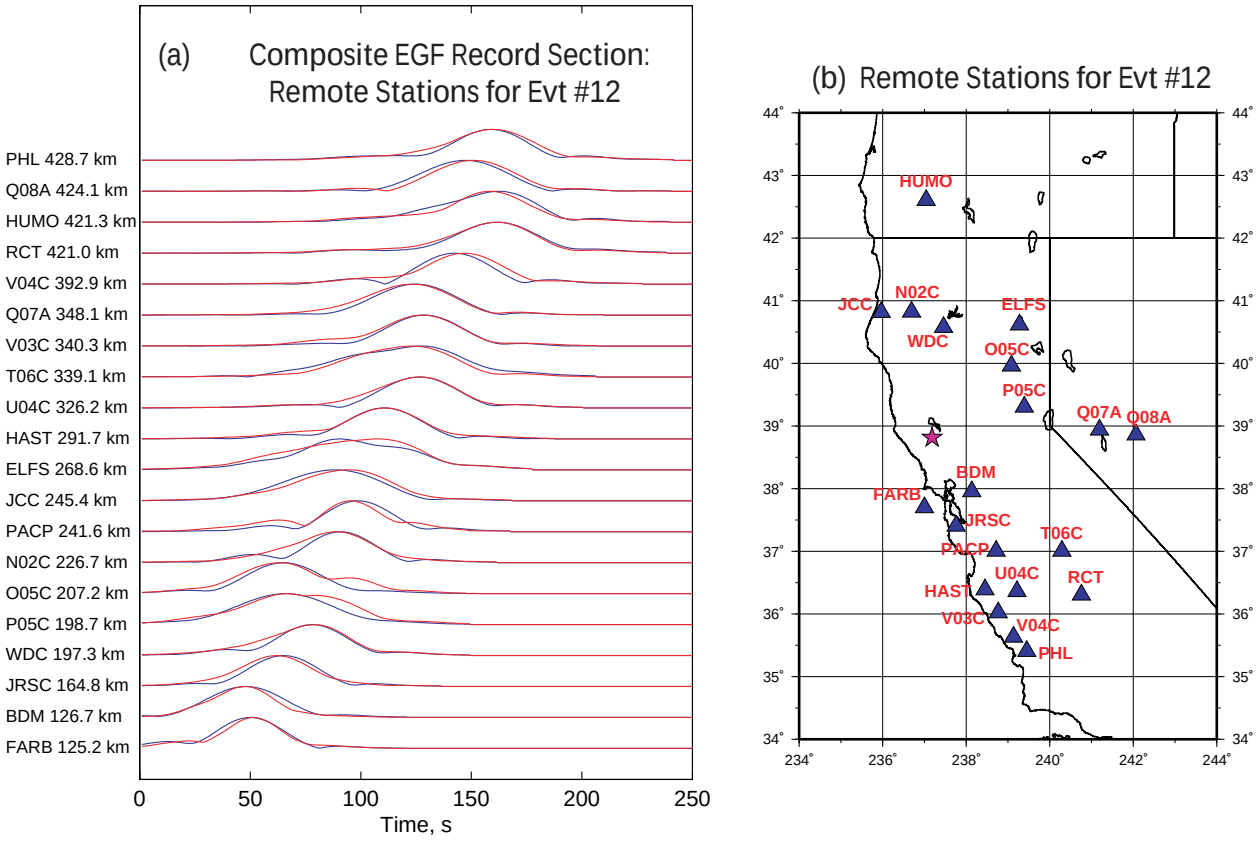

Figure 4: Record section of the Composite Empirical Green's Functions compared with the earthquake records at 20 remote stations for Event #12. (a) Envelope functions of the earthquake observed at the remote stations (red lines) are compared with envelopes of the Composite EGFs (blue lines). Band-pass: 7-15 sec period. Earthquake characteristics are listed in Fig. 2 caption and Table 3. Epicentral distances and station names are indicated at left. (b) Locations of the remote stations (blue triangles) and the earthquake (red star). The locations of the eight base stations are shown in Fig. 3.

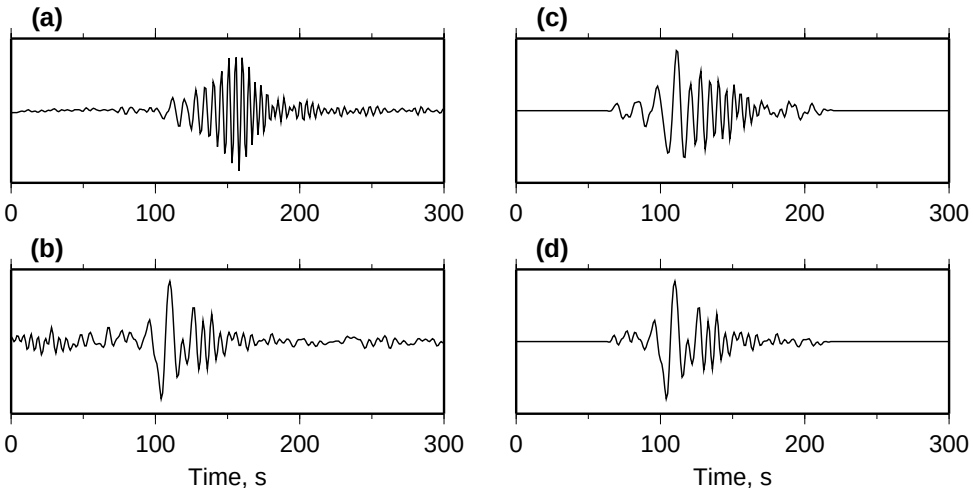

Figure 5: Normalization of amplitude spectra. (a) Event seismogram (Z-component, displacement) recorded at the station T14A (corrected for the instrument and normalized to unity) following the Crandall Canyon mine collapse on Aug 6, 2007. (b) EGF for the station pair: P16A - T14A. (c) Amplitude equalized event seismogram, cosine tapering is applied to both ends. (d) Amplitude equalized EGF, cosine tapering is applied to both ends.

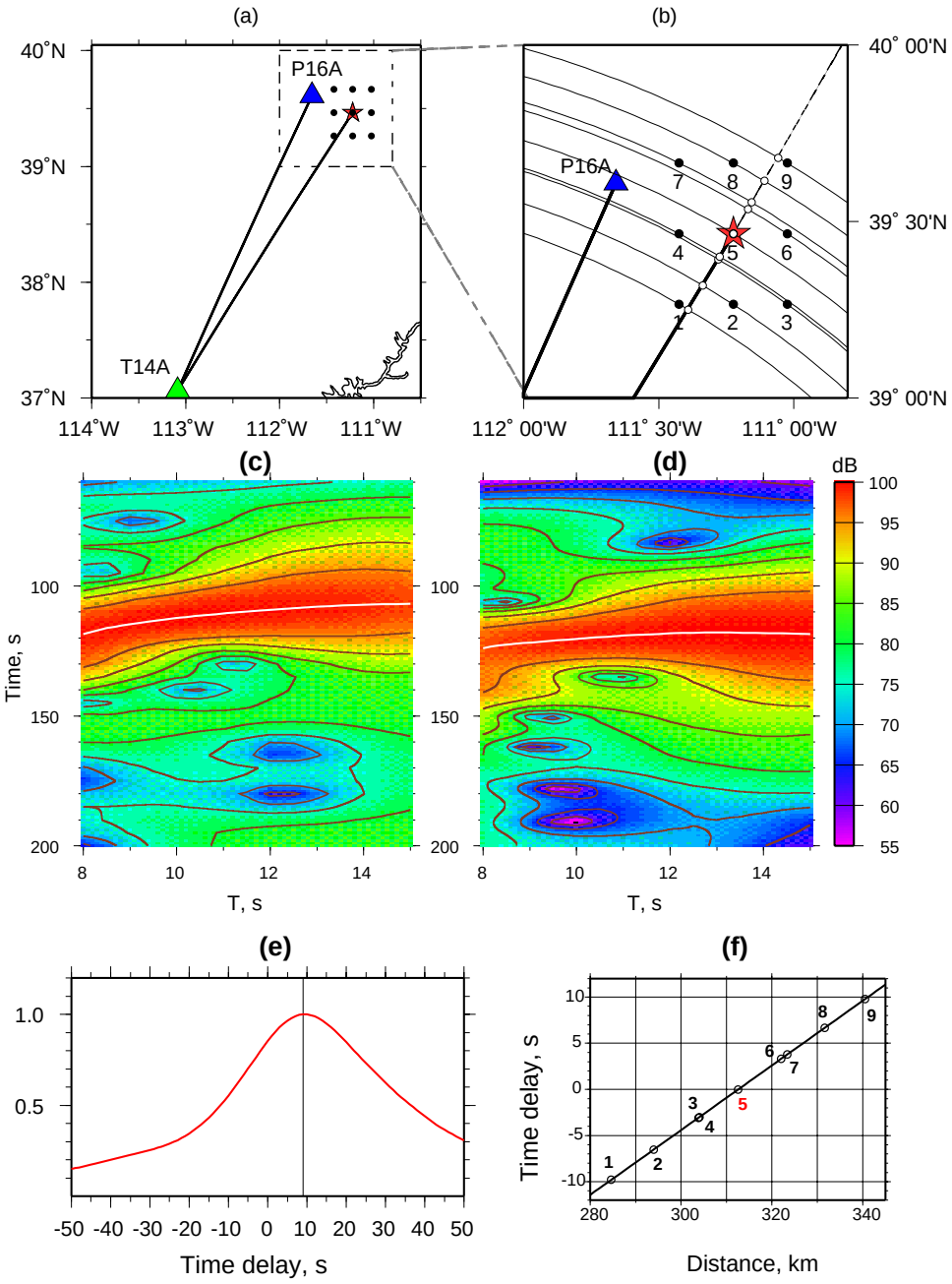

Figure 6: Illustration of location procedure based on cross-correlation of frequency-time (FT) representations of an event record and an EGF. (a) Event-station geometry. Event location (star) of the Crandall Canyon mine collapse and position of the base station P16A and the remote station T14A (triangles) are shown. The epicenter is surrounded by a 9-point grid (bold dots). (b) More detailed picture of the grid with grid points numbered from 1 to 9. The curved lines are circles centered on the remote station T14A. (c) FT-diagram of the event record for station T14A following the Crandall Canyon Mine mine collapse. (d) FT-diagram of the EGF for the station-pair T14A and P16A (transformed to the distance between grid point 9 and T14A). (e) Cross-correlogram (eqn. (6)) of the two FT-representations in (c) and (d) with the time delay of the maximum around 10 s. (f) Time delays of maxima of cross-correllograms as a function of distance from remote station T14A for all nine grid points. The minimum time delay corresponds to grid point 5, near the known epicenter of the mine collapse. Linear regression time-distance is shown by a solid line.

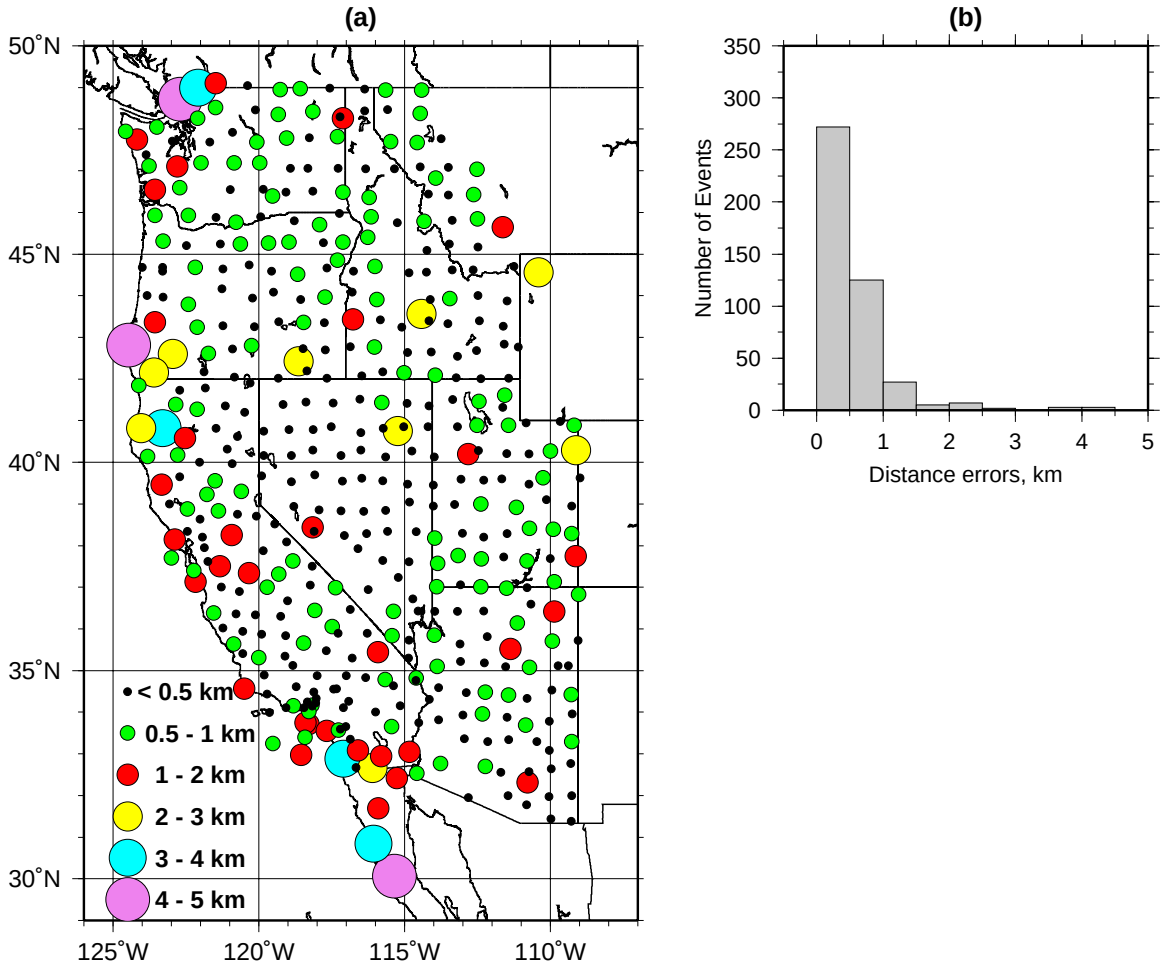

Figure 7: Location errors of 445 "virtual sources" (stations) in California and surrounding states. (a) Map of location errors for the stations. The distance between the known station location and the estimated location, in km, for each station position is shown by the colored circles. (b) Histogram of location errors (in km).

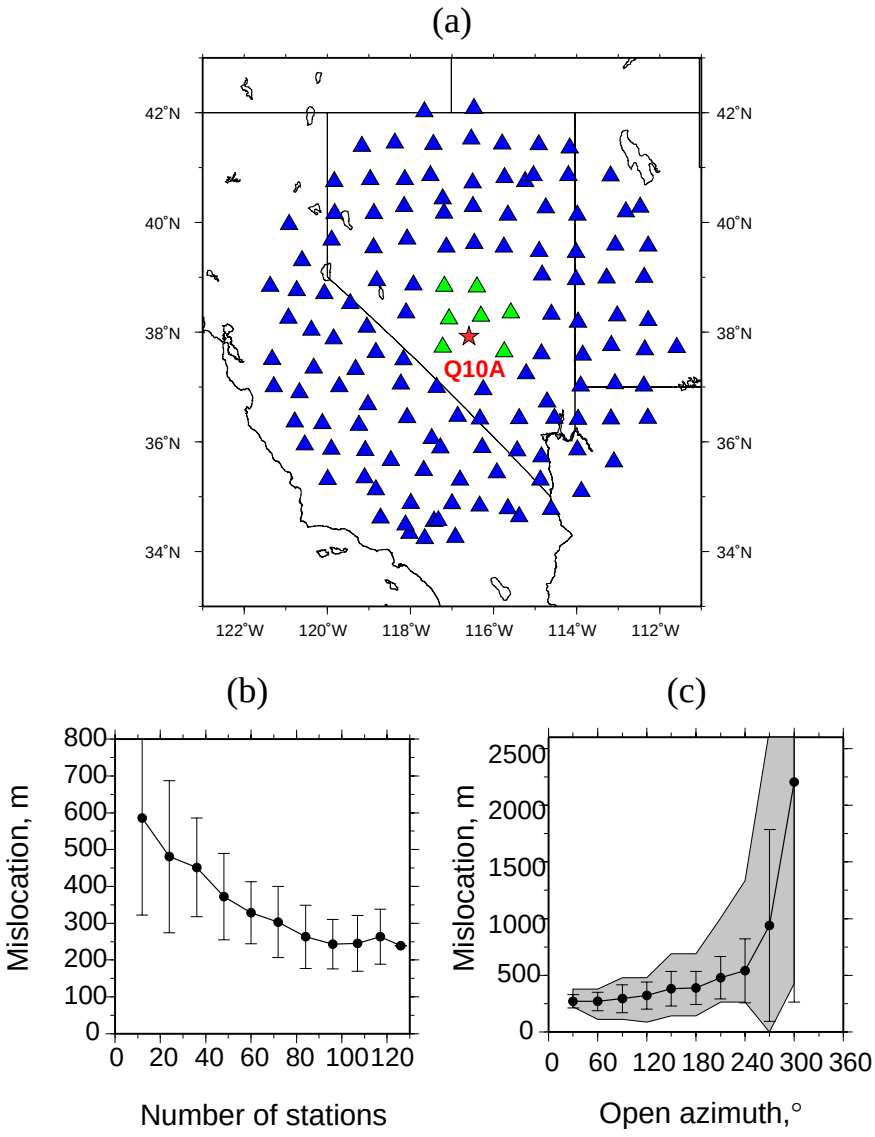

Figure 8: Location error as a function of the number of remote stations and azimuthal coverage. (a) Location of stations and virtual event (TA station Q10A): base stations (green triangles), remote stations (blue triangles), red star is the virtual source. (b) Effect of varying the number of remote stations on location accuracy. (c) Effect of the azimuthal coverage on location accuracy. Shaded domain covers all solutions for ten realizations of each open azimuth. Error bars are 1 standard deviation.

Figure 9: Schema of the Crandall Canyon min and location of mine collapse. Our location of the event (green star) and corresponding 90% confidence ellipse. The left yellow push-pin marks the USGS event location, and the right push-pin is the approximate location of the mine collapse and trapped miners. The map is made with Google Earth.

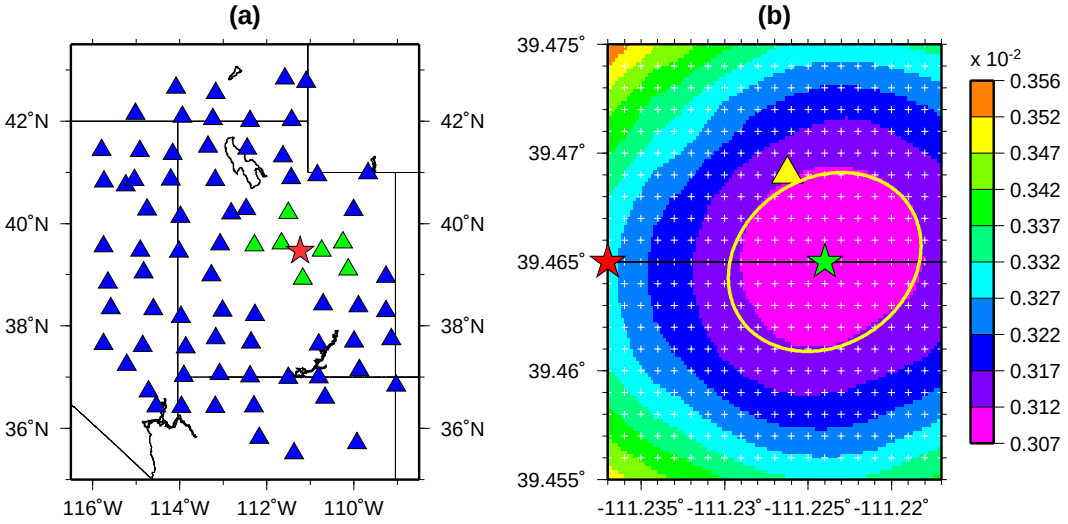

Figure 10: Station configuration and misfit surface for the location of the Crandall Canyon mine collapse. (a) Symbols: base stations (green triangles), remote stations (blue triangles), and event location (red star). (b) Map of residuals: our event location (green star), USGS location (red star),the 90% confidence ellipsoid (yellow line), and the place believed to be where miners were trapped (yellow triangle). Grid points are marked with white crosses separated by about 100 m.

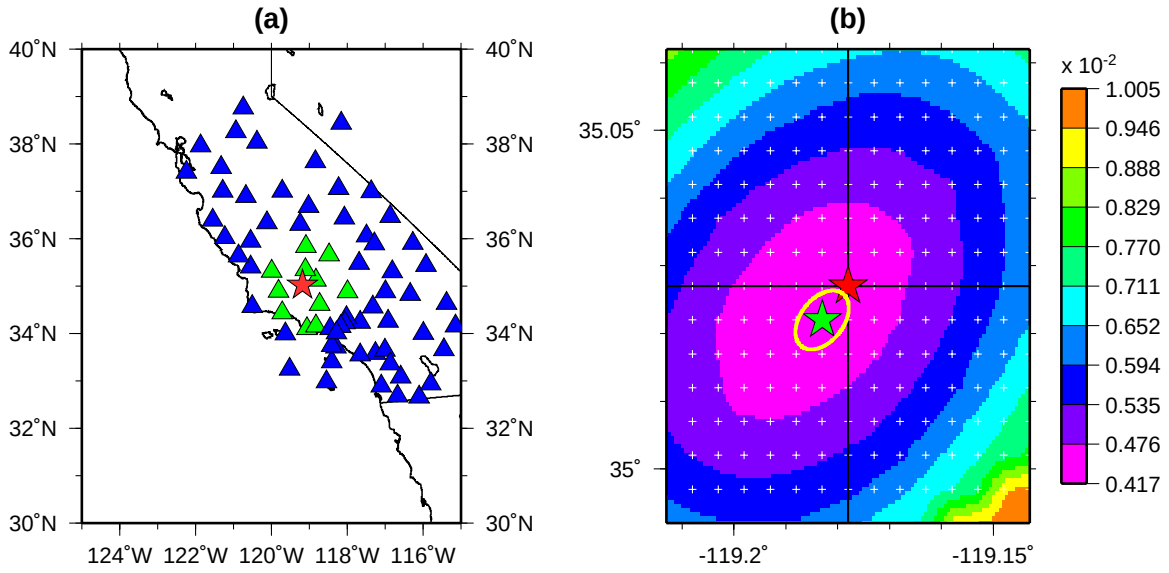

Figure 11: California event #1 station configuration and misfit surface. (a) Symbols: base stations (green triangles), remote stations (blue triangles), and event location (red star). (b) Map of residuals: our event location (green star), USGS location (red star), 90% confidence ellipsoid (yellow line). Grid points are marked with white crosses and are separated by about 500 m.

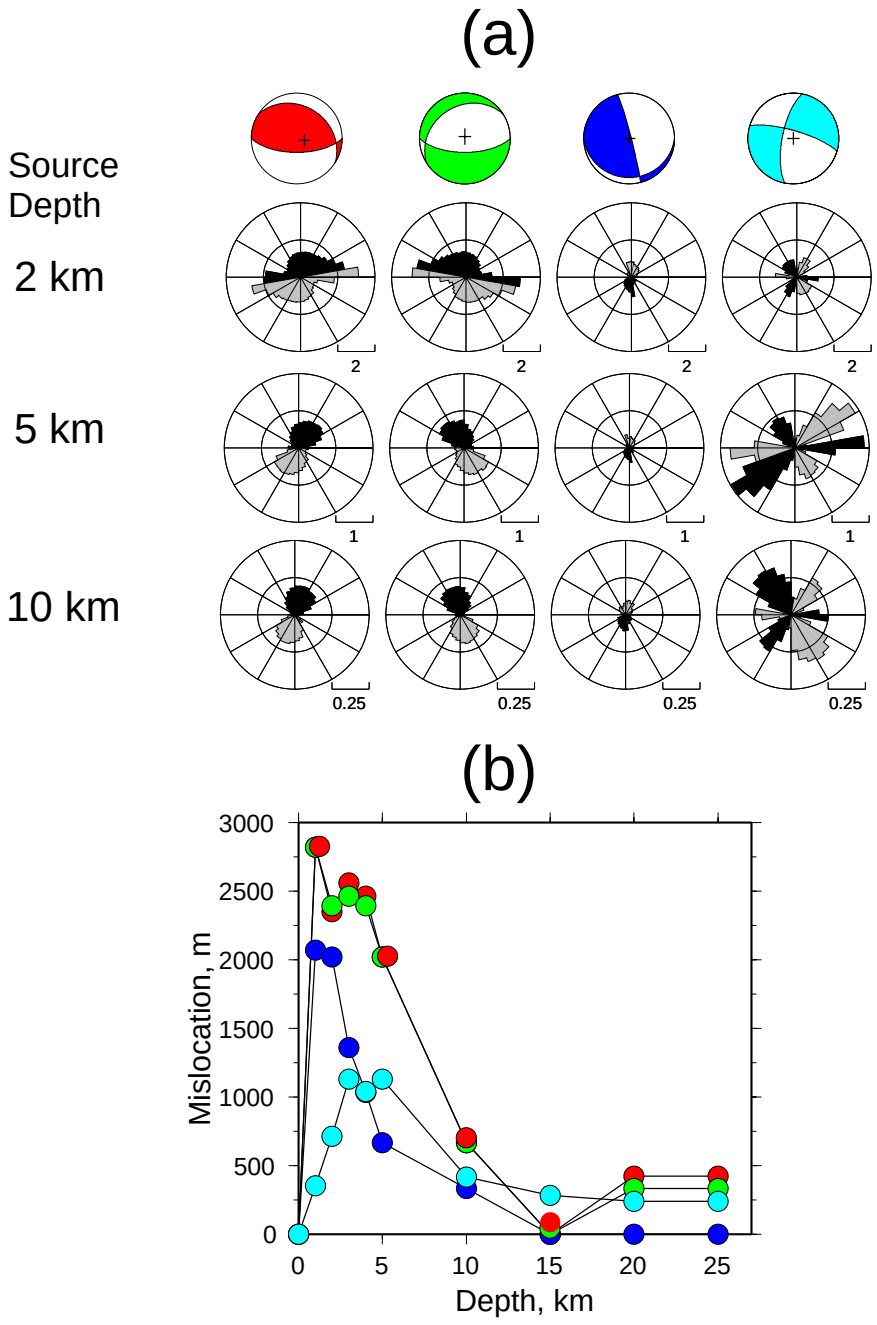

Figure 12: Simulation of the effect of source mechanism and depth on group time shifts and mislocation. Combination of base-remote stations used in the numerical simulation is shown in Figure 8a. (a) Source mechanisms (upper) and corresponding polar diagrams of group time residuals (lower) for the indicated source depths. Mechanism: red, near normal; green, near thrust; navy, near vertical thrust; light blue, near strike-slip. (b) Event mislocation for the four different source mechanisms as a function of depth. Symbol color correspond to the source mechanisms shown in (a).

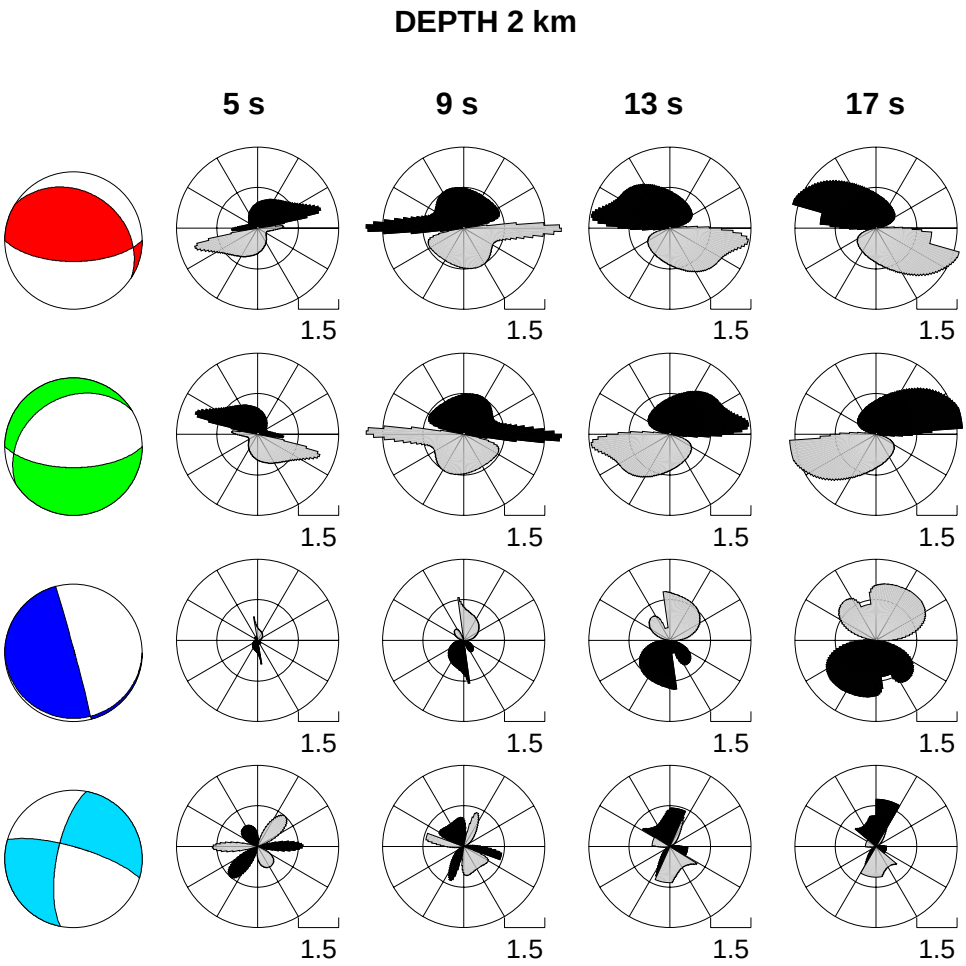

Figure 13: Theoretical prediction of group time delays at different periods. Azimuthal dependence of group time delays are presented for four periods and four different mechanisms (the same as in Fig. 12). The size of the time delays scale according to the bar at the lower right-hand side of each component of the figure, in seconds.

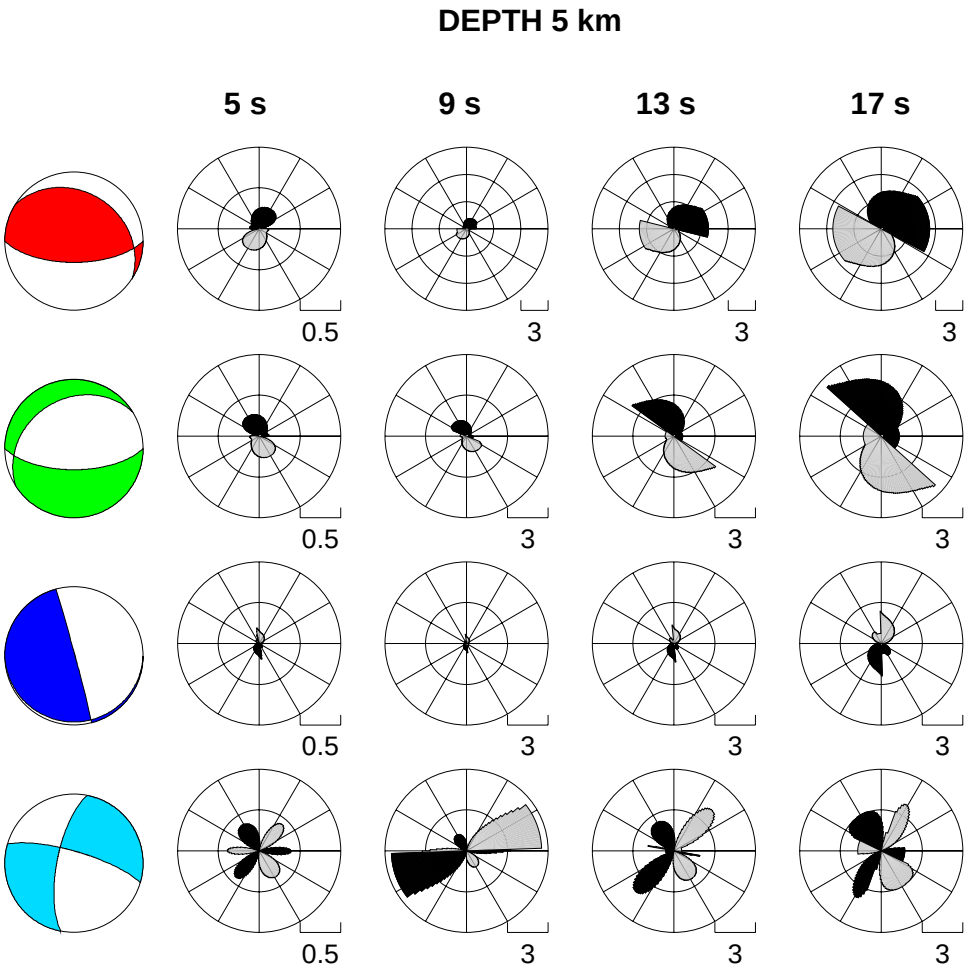

Figure 14: Theoretical prediction of group time delays at different periods. Same as Fig. 13, but for events at 5 km depth.

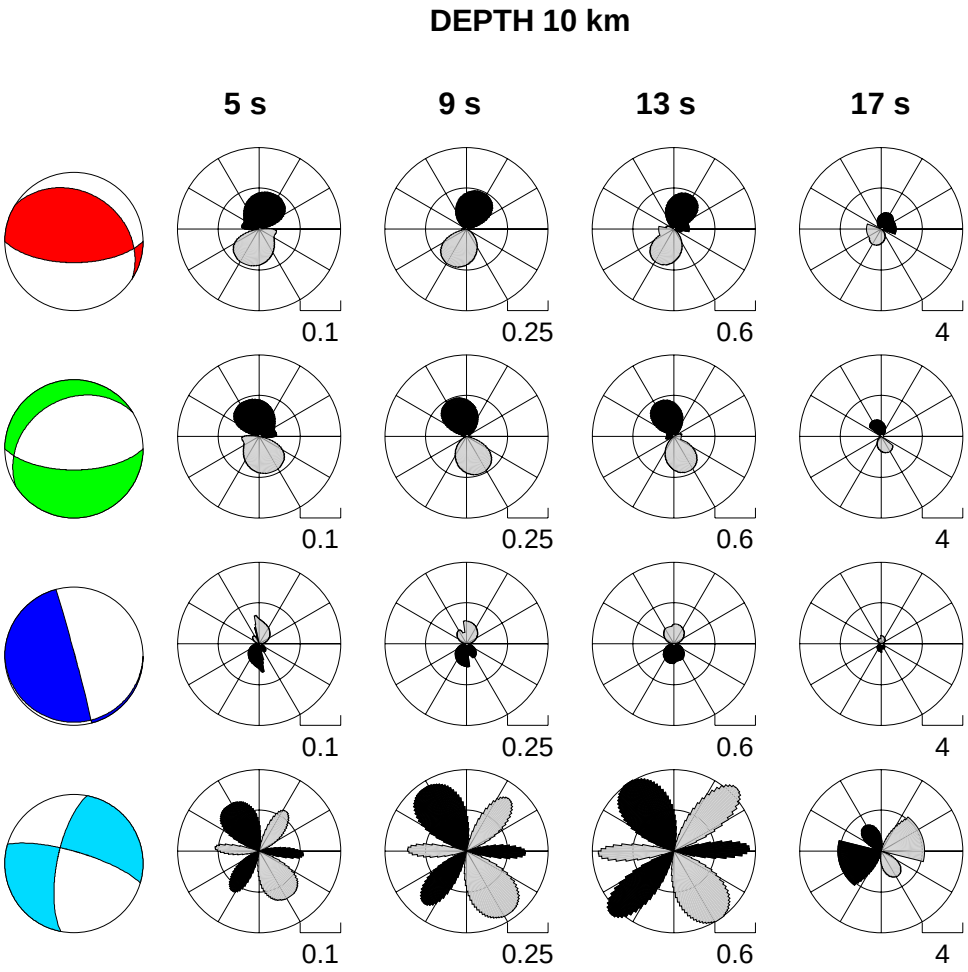

Figure 15: Theoretical prediction of group time delays at different periods. Same as Fig. 13, but for events at 10 km depth.