

EPICENTRAL LOCATION OF REGIONAL SEISMIC EVENTS BASED ON EMPIRICAL GREEN'S FUNCTIONS FROM AMBIENT NOISE

Michael H. Ritzwoller, Mikhail P. Barmin, Anatoli L. Levshin, and Yingjie Yang

University of Colorado at Boulder

Sponsored by the National Nuclear Security Administration

Award No. DE-AC52-09NA29326

Proposal No. BAA09-50

ABSTRACT

The purpose of this research is to develop and test a novel method of regional seismic event location based on exploiting Empirical Green's Functions (EGF) that are produced using ambient noise methods. Elastic EGFs between pairs of seismic stations are determined by cross-correlating long ambient noise time-series recorded at the two stations. The EGFs principally contain Rayleigh and Love wave energy and focus is placed on utilizing these signals between 5- and 15-sec periods. We have begun the fundamental R&D needed to develop and test methods to produce a geographic grid of vertical and horizontal component EGFs, which can then be used to detect and locate events accurately using waveform correlation. Because the EGFs contain the full 3-D structural response of the Earth for surface waves, the method is tantamount to using an exceptionally accurate 3-D model for detection and location. Because the method is empirical, however, location error is expected to be unbiased by 3D structure. Furthermore, ambient noise is generated at the Earth's surface, so the EGFs are tuned for surface-focus events and are, therefore, ideal for detecting and locating human-made events.

Proof-of-concept applications of a preliminary version of the algorithm are presented to estimate epicentral locations in the western U.S. using EarthScope USArray data. These include the determination of station locations (GT0) by reciprocity, thirteen previously well located events (GT1) in California, and the Crandall Mine (GT0.5) collapse. Research is ongoing to extend and generalize the method to include hypocentral location and moment tensor determination.

OBJECTIVES

The purpose of this research is to improve seismic event location accuracy and event characterization by exploiting Empirical Green's Functions (EGFs) that emerge by cross-correlating long time sequences of ambient noise observed at pairs of seismic stations. Because ambient noise EGFs are dominated by surface waves, the method uses surface wave energy for location purposes. Present efforts apply Rayleigh wave EGFs from 5 to 15 sec period to the epicentral location problem. The method of epicentral location is described here as well as proof-of-concept applications for a series of seismic events in the western US. Future work will continue to refine the method, generalize its application to Love waves, introduce phase information, and explore the ability to constrain depth and the moment tensor using ambient noise.

RESEARCH ACCOMPLISHED

Non-Technical Description of the Method

We assume first that there is a temporary dense local array, termed the **base stations**, deployed in the "region of interest" where seismic events may occur. Second, there is a more distant permanent (but potentially sparse) regional network of stations termed the **remote stations**. Using what are now well established methods (e.g., Bensen et al., 2007), the EGFs from every base to every remote station are computed. At this stage of the project, we use only vertical-to-vertical components of EGFs here, which are dominated by Rayleigh waves. The set of these EGFs can be used to calculate an EGF at any geographical point covered by the base station deployment by transforming and adding (stacking) the individual EGFs for a given remote station. Such EGFs are called Composite EGFs (CEGFs). These CEGFs then can be used to locate events within the region of interest using waveform correlation with event records at remote stations. The location technique applied here involves specific transformations of the input EGFs before stacking, which we do not describe in detail here. These transformations improve the signal/noise ratio and minimize the effect of waveform differences between event records and CEGFs.

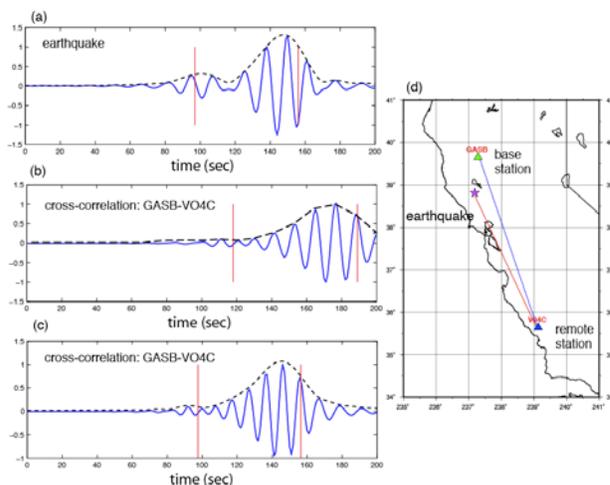


Figure 1. Comparison of an earthquake record with an EGF. (a) Earthquake record at station V04C in southern CA band-passed between 7 and 15 sec period. (Evt. #12 from Table 1.) (b) EGF between stations GASB (north of the earthquake) and V04C with the same pass-band. (c) The EGF has been shifted and deformed to correspond to the earthquake location. (d) Earthquake and station locations.

As an example that illustrates the idea of the method of epicentral location, consider the $m_b=4.5$ earthquake that occurred north of the San Francisco Bay area on May 12, 2006, Figure 1d. This is Event #12 from Table 1. The vertical component of the earthquake record and the vertical-vertical EGF from a station near the epicenter are shown in Figures 1a, and 1b. The earthquake record is from station V04C, and the inter-station EGF is for stations GASB (about 100 km from the epicenter) and V04C. Both time-series are band-pass filtered between the 7- and 15-sec period and are dominated by the Rayleigh wave. The earthquake and EGF Rayleigh waveforms are similar, but are offset 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, Figure 1c. The technical details of this transformation are not presented here. The phase content of the two records differ appreciably, because earthquakes impart an initial phase that depends on hypocentral depth and source mechanism as well as epicentral distance. For this reason, at this early stage of research, we ignore phase information and summarize both the EGFs and the earthquake records with their envelope functions, seen in Figure 1 as dashed lines.

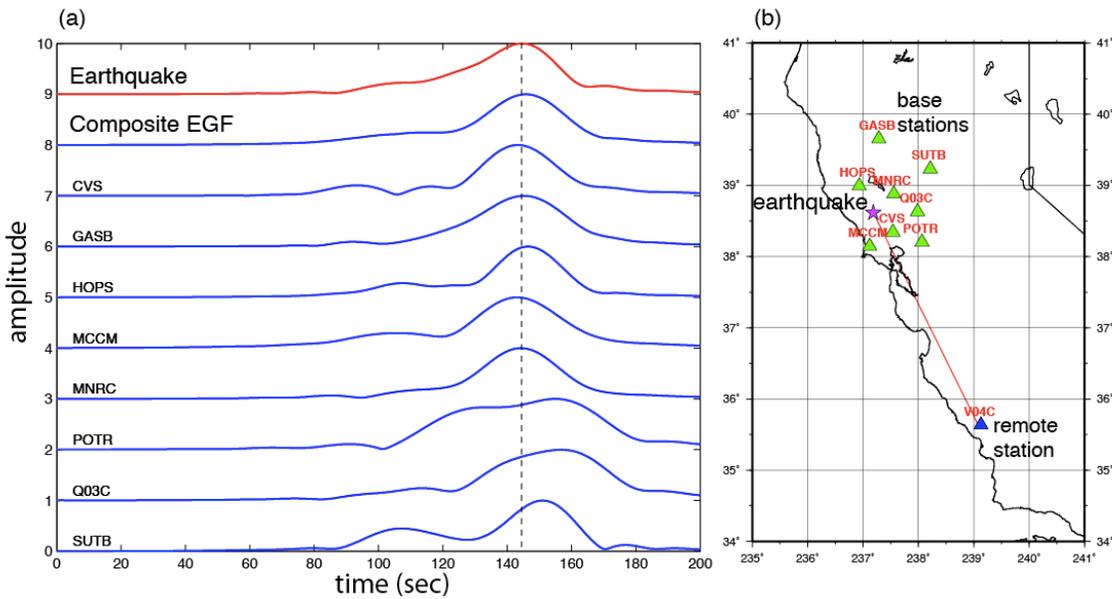


Figure 2. Comparison of EGFs for the set of base stations to a single remote station. Similar to Figure 1, but only the envelopes of the EGFs are shown for the eight base stations in (a). The envelope of the earthquake record is shown in red at top and the envelope of the Composite EGF (CEGF, sum of the individual EGFs) is on the second line. Individual EGFs and the CEGF are transformed to the epicentral location of the earthquake.

The EGFs for each of the base stations relative to the remote station can be similarly transformed to the earthquake location and stacked to produce a CEGF for a particular source location and remote station, as illustrated in Figure 2. Typically, the Composite EGF agrees with the earthquake record better than the individual EGFs from individual base stations. The Composite EGFs for a candidate event location can be produced for all remote stations, as the record section shown in Figure 3 illustrates. Inspection of fit between the Composite EGFs and the earthquake envelope functions reveals that 3-D structure similarly affects the EGFs 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 on 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.

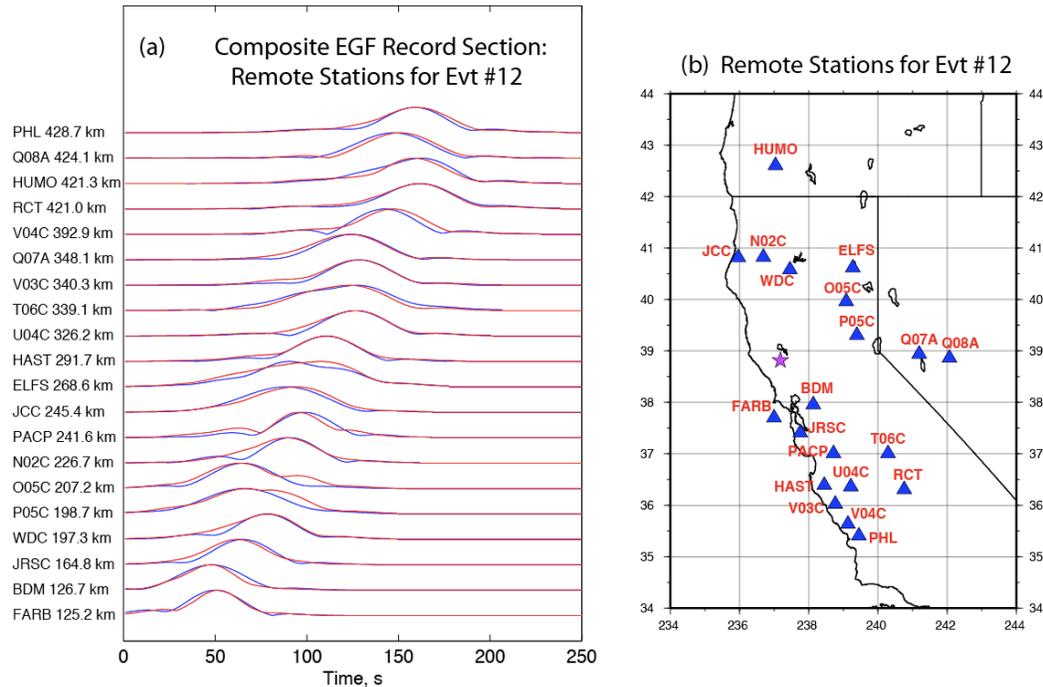


Figure 3. Record section of the CEGFs compared with the earthquake records at 20 remote stations for California Event #12. (a) Envelope functions of the earthquake observed at the remote stations (red lines) are compared with envelopes of the CEGFs (blue lines). Band-pass: 7- to 15-sec period. Epicentral distances and station names are indicated at left. (b) Locations of the remote stations (blue triangles) and earthquake (red star). The locations of the eight base stations are shown in Figure 1.

Our location procedure minimizes the misfit between the envelopes of the CEGFs and the earthquake records by waveform cross-correlation. The minimization is performed by a straightforward grid search, as illustrated in Figure 4. This is the same earthquake as in Figures 1-3, Event #12 of Table 1. To locate this event, we used eight base stations and 55 remote stations from the Northern California Network and EarthScope's Transportable Array (Figure 4a). The grid spacing for location was about 500 m. The reference location is indicated with a red star and our location is given by a yellow star in Figure 4b. Error ellipses are estimated with a method that is based on work of Flinn (1965) and Jordan and Sverdrup (1981). The estimated error ellipse is shown with a yellow curve. The difference in locations in this example is one grid spacing, which is approximately encompassed by the estimated error ellipse.

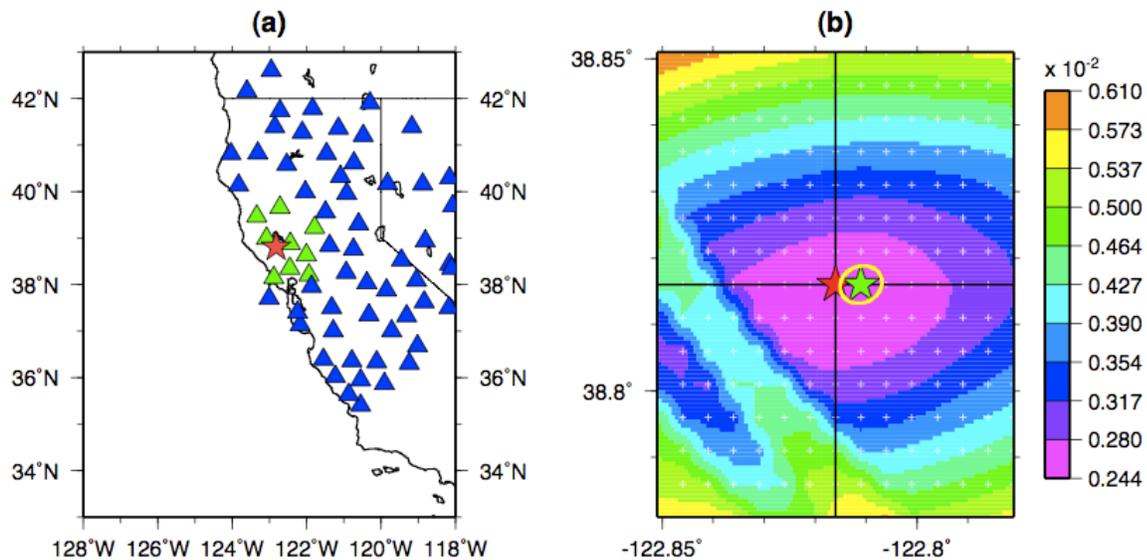


Figure 4. Example of the grid-search location of California Reference Event #12. (a) Stations used to locate this earthquake (red star): base stations (green triangles), remote stations (blue triangles). (b) Map of residuals, which is the basis for the location and error estimate: our event location (green star), reference location (red star), our 90% confidence ellipsoid (yellow line). Grid points are separated by about 500 m and are marked by white crosses.

Proof-of-Concept Applications

We present here several examples to assess the capabilities of the epicentral location method. At present, the method's development remains in preliminary stages and we consider the results presented here to be in the proof-of-concept category.

Locating stations by reciprocity (GT0). In computing EGFs from inter-station cross-correlations, either of the pair of stations can be considered to be a virtual source by reciprocity. Thus, the EGFs from a particular station to the set of other stations can be used to locate the common station. In this way we have used our epicentral location method to locate the EarthScope Transportable Array stations as well as some other local stations in the western US. The advantage of this test is that the station locations are known exactly (i.e., they are like genuine GT0 events) and, unlike earthquakes, the EGFs possess no unknown initial phase caused by event depth and mechanism or other complicated source characteristics. Thus, imperfect knowledge of real seismic events does not complicate the interpretation of the method's intrinsic accuracy. However, the method's ability in practice to locate real seismic events in all of their complexity still needs to be tested.

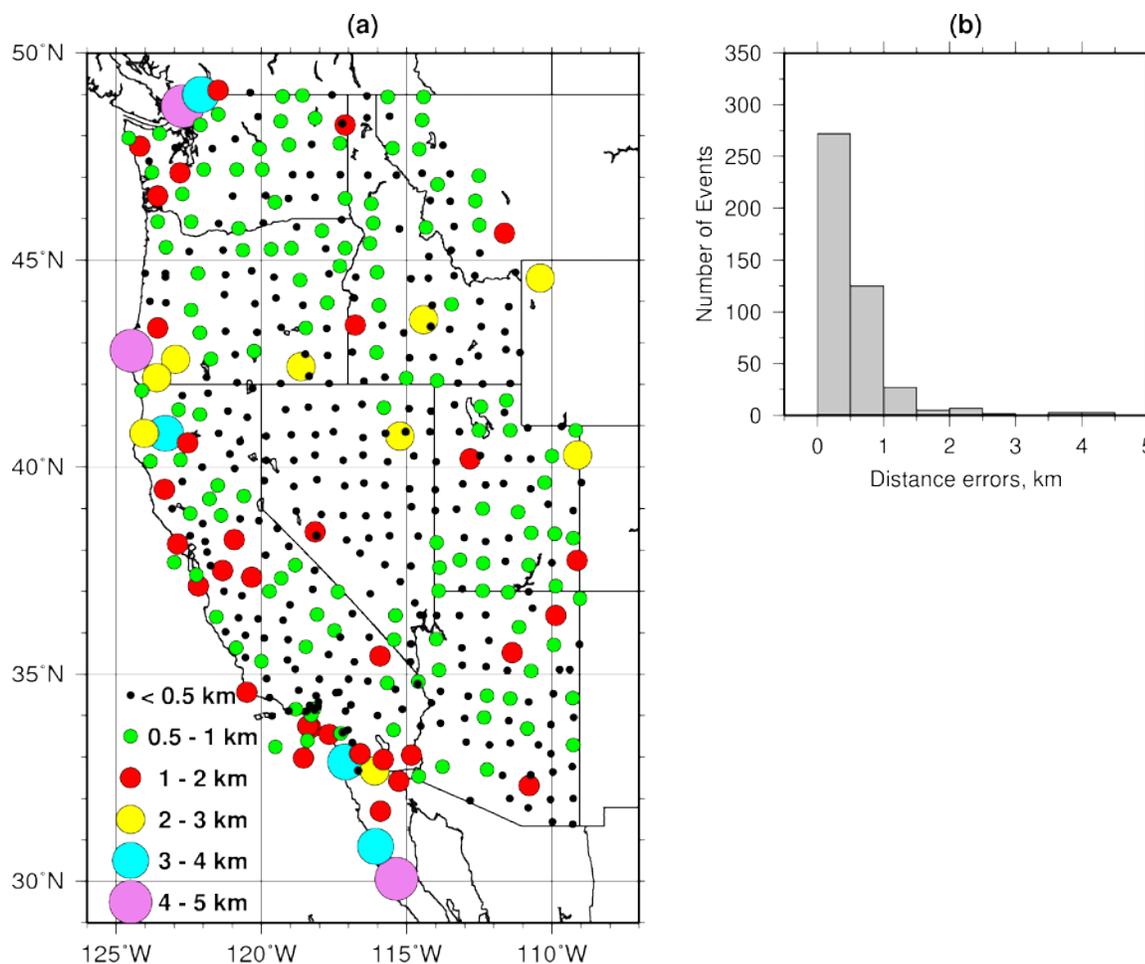


Figure 5: Location errors for 445 stations deployed in California and surrounding states located with the method described here from EGFs by reciprocity. (a) Map of location errors for existing stations. The distance between the station and the estimated location, in km, for each station position is shown with colored circles. (b) Histogram of location errors (in km).

The results of this exercise are shown in Figure 5. Test parameters include a location grid spacing of 0.5 km, base stations are located within 100 km of each station to be located, and remote stations are situated between 100 km and 400 km from the station. About 61% of the stations are located within 0.5 km of their real location, 28% between 0.5 km and 1.0 km, and only 7 of the 445 stations are located worse than 3 km. The bad locations are mostly near the coast where azimuthal coverage is less than 180 degrees and where structural gradients are profound. The worst locations are in southern California where azimuthal coverage is particularly incomplete and near the Mendocino Transform where numerous offshore earthquakes exist and tend to vitiate the EGFs. We need to continue to tune the method to make it less sensitive to these variables. Nevertheless, at present we estimate the intrinsic accuracy of the method in this nearly ideal observational setting to be better than 0.5 km.

The station-location test will be particularly useful to test the characteristics of the location method as the base stations or remote stations decrease in number and as the EGFs decrease in quality (as they may for short station deployments or in poor signal-to-noise environments). Knowledge of the performance of the location method as station coverage degrades is important to understand its reliability in realistic monitoring situations.

Locating the Crandall Canyon mine collapse (GT0.5). The next location test moves beyond the idealization of locating stations, to the location of a shallow mine collapse in central Utah. This seismic event of magnitude 3.9 occurred on Monday August 6, 2007, at about 08:48:40 UTC, caused by the collapse of the coal mine at a depth of about 1 km that permanently trapped several miners underground. The outline of the mine is shown in Figure 6 with a yellow line. The location of the collapse and the focus of the drilling efforts is shown by a yellow push-pin to be in

the far western part of the mine several km from the mine's entrance, beneath steeply sloping surface topography. The USGS location (actually performed by the University of Utah using local seismometers as well as recently installed stations from the EarthScope Transportable Array) is shown to lie about 1.2 km from the believed location of the trapped miners, in a valley west of the contours of the mine. To locate the event, we used seven base stations and 66 remote stations (instead of 33 used by USGS) affiliated with different networks: USGS, University of Utah, and EarthScope's Transportable Array (Figure 7a). The grid spacing for location was 100 m. The location results are: 39.465N (USGS 39.465N), 111.224W (USGS 111.237W), an origin time shift is equal to -0.26 sec relative to the USGS time, and confidence ellipse parameters (Flinn, 1965; Jordan and Sverdrup, 1981) of $a=0.52$ km (semi-major axis), and $b=0.43$ km (semi-minor axis). This location is within the contours of the mine and less than 0.5 km from the best estimate of the mine collapse, which is within the error ellipse of this event.



Figure 6. Schema of the Crandall Canyon mine, outlined with the yellow line. The location of the mine collapse and trapped miners and the USGS seismic location are shown with yellow push-pins. Our location is identified by a green star and error-ellipse.

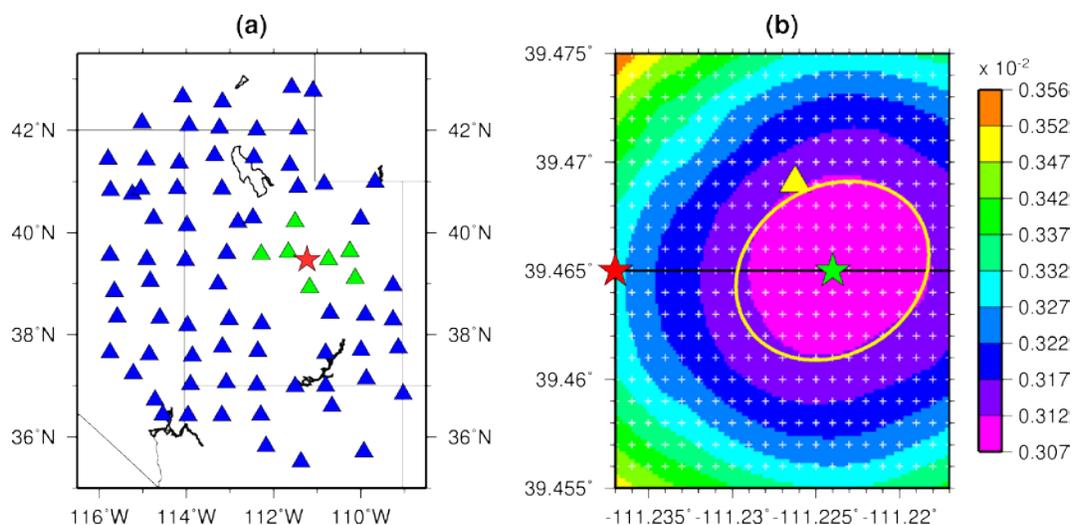


Figure 7. (a) Stations used to locate the Crandall Canyon mine collapse (red star): base stations (green triangles), remote stations (blue triangles). (b) Map of residuals, which is the basis for the location and error estimate: our event location (green star), USGS location (red star), our 90% confidence ellipsoid (yellow line), and best estimate for the location of trapped miners (yellow triangle). Grid points are separated by 100 m and are marked by white crosses.

Because this is a shallow seismic event, it provides a good test of the location method accuracy for human-made explosions. An accuracy of ~ 0.5 km within the context of a nearly ideal observational setting is a reasonable expectation. Tests of other mine collapses or historical nuclear explosions will help to calibrate these expectations. Further accuracy tests as a function of decreasing the number of observing stations are also needed.

Locating crustal earthquakes in California (GT1). Crustal events deeper than the Crandall Canyon mine collapse impart frequency dependent initial phase and group time shifts (Levshin et al., 1999) that are not known a priori. These shifts will degrade the epicentral location estimated with our method unless they are explicitly estimated or taken into account during the location process. At present, they are not modeled in our location procedure. To assess their effects, the next exercise extends the tests beyond shallow events to a set of well-located (\sim GT1) earthquakes within the California crust. Table 1 presents the locations of these reference earthquakes, which occurred in 2005 and 2006 and have mb values ranging between 4.5 and 5.1. The accuracy of these reference locations is believed to be better than 1 km in most cases. The uncertainty in the reference locations should be kept in mind in evaluating the locations from our method.

An example of the location procedure for Event #1 is shown in Figure 8. This is a southern California event where the reference locations are determined by CalTech. For events in northern California, such as those shown in Figures 1–4, the reference locations were determined by the Northern California Earthquake Center in Menlo Park. Both organizations use a combination of broad-band and short period instruments and apply 1-D models calibrated for each region. CalTech used both *P* and *S* phases, whereas only *P* phases were used in northern California. The location grid for a northern California earthquake (Event #12) is shown in Figure 4.

Table 1. Reference (GT1) events in California 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.

We located the events listed in Table 1 and the differences in the epicentral locations and origin times between our method and the reference events are summarized in the last two columns of this table. The average difference in location (called “error” in Table 1) for these events is 1.2 km, and the average “error” in origin time is 0.04 s. When the uncertainties in the reference locations are accounted for, the 90% confidence level for our locations is about 1 km. These locations, however, are degraded relative to those for the stations or the Crandall mine collapse in our earlier tests, which we estimated to be better than 0.5 km. We conclude, therefore, that un-modeled earthquake depth is contributing a bias to our locations. For our method to yield epicentral locations better than the GT1 level for shallow crustal events such as those in California the effect of earthquake depth must be incorporated in the technique. This is a direction for on-going research.

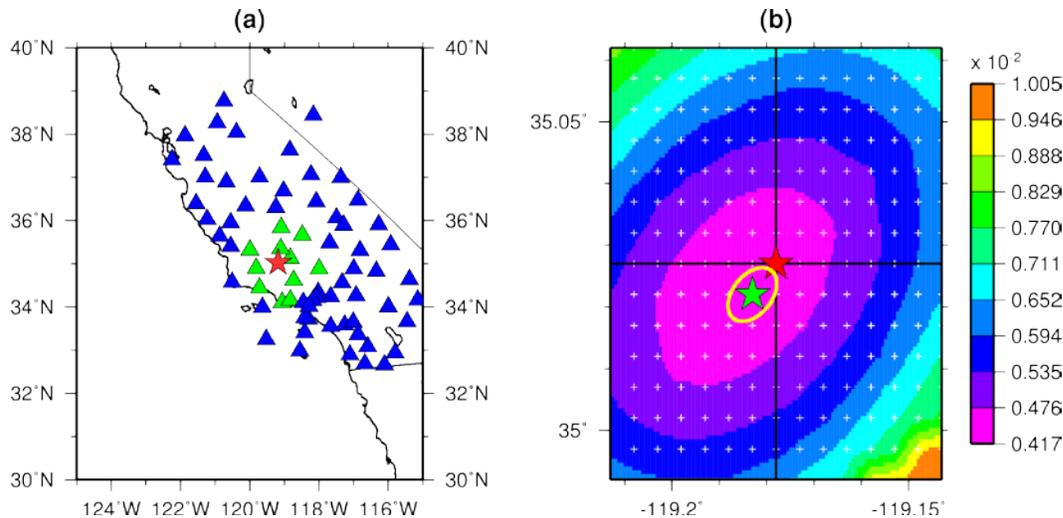


Figure 8. Epicentral location test for California Event #1. (a) Location of the stations used: base stations (green triangles), remote stations (blue triangles). Reference event location is marked with a red star. (b) Map of the grid of residuals which our location is based on: our event location (green star), reference location (red star), and the 90% confidence ellipsoid (yellow line). The ~0.5 km grid points are marked with white crosses.

CONCLUSIONS AND RECOMMENDATIONS

The work presented here has been performed at the very early stages of the contract, and remains preliminary in nature. Nevertheless, work to date demonstrates that absolute locations based on the use of ambient noise empirical Green's functions can, for stations (located by reciprocity) or shallow seismic events, yield epicentral locations with an intrinsic error of about 500 m, on average. Crustal earthquakes, however, impart phase and group time shifts that can bias the location, and we show how earthquakes in California are located with our method with an accuracy of about 1 km. In addition, the results presented here are for a nearly optimal observational setting, in which there are more than five base stations within 100 km of each event and dozens of high-quality remote stations within 400 km. Future work will generalize the method to use more information from the Empirical Green's function (e.g., Love waves, phase information, frequency dependence of misfit), explicitly estimate or correct for the biasing effect of non-surface event depth, and quantitatively assess how the reduction in base and remote stations degrades the method.

ACKNOWLEDGMENTS

The data used in this research were obtained from the IRIS Data Management Center and originate predominantly from the Transportable Array component of EarthScope/USArray. We would like to thank Egill Hauksson, James Dewey, and Bob Engdahl for valuable information gleaned from personal communications.

REFERENCES

- 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: 1239–1260.
- Flinn, E.A. (1965). Confidence regions and error determinations for seismic event locations, *Rev. Geophys* 3: 157–185.
- Jordan, T.H., Swerdrup K.A. (1981). Teleseismic location techniques and their application to earthquake clusters in the South-Central Pacific, *Bull. Seismol. Soc. Amer.* 71: 4, 1105–1130.
- Levshin, A.L., Ritzwoller, M.H., Joe S. Resovsky (1999). Source effects on surface wave group travel times and group velocity maps, *Phys. Earth Planet. Inter.* 115: 293–312.