LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-CONF-467300

Mapping Diffuse Seismicity Using Empirical Matched Field Processing Techniques

J. Wang, D. C. Templeton, D. B. Harris

January 24, 2011

Thirty-Sixth Workshop on Geothermal Reservoir Engineering
Stanford, CA, United States
January 31, 2011 through February 2, 2011

Disclaimer

PROCEEDINGS, Thirty-Sixth Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, January 31 - February 2, 2011
SGP-TR-191

MAPPING DIFFUSE SEISMICITY

USING EMPIRICAL MATCHED FIELD PROCESSING TECHNIQUES

Jingbo Wang[1], Dennise C. Templeton[1], David B. Harris[2]

1. Lawrence Livermore National Laboratory, 7000 East Ave, L -231, Livermore, CA, 94550, USA
2. Deschutes Signal Processing, LLC, P.O. Box 231, Maupin, OR, 97037, USA
e-mail: wang64@llnl.gov

ABSTRACT

The objective of this project is to detect and locate more microearthquakes using the empirical matched field processing (MFP) method than can be detected using only conventional earthquake detection techniques. We propose that empirical MFP can complement existing catalogs and techniques. We test our method on continuous seismic data collected at the Salton Sea Geothermal Field during November 2009 and January 2010. In the Southern California Earthquake Data Center (SCEDC) earthquake catalog, 619 events were identified in our study area during this time frame and our MFP technique identified 1094 events. Therefore, we believe that the empirical MFP method combined with conventional methods significantly improves the network detection ability in an efficient matter.

INTRODUCTION

Accurate identification and mapping of large numbers of microearthquakes is one technique that provides diagnostic information when determining the location, orientation and length of underground crack systems for use in reservoir development and management applications. Conventional earthquake location techniques are often employed to locate microearthquakes. These techniques require picking individual seismic phase onsets across a network of sensors and work best on seismic records containing a single well-recorded event with low signal-to-noise. However, this process can be labor-intensive and poorly suited to rapid turn-around situations, such as providing feedback during hydraulic fracture operations, since even automatic picks often require assessment and correction by an analyst. Additionally, fluid injection frequently induces a large number of events with overlapping waveforms, which can complicate the picking of phases or completely obscure the onset of smaller signals.

To aid in the seismic characterization of reservoir fracture networks, we propose to complement traditional earthquake detection and location techniques with the empirical matched field processing (MFP) method. MFP, as applied in seismology, matches the spatial structure of incoming seismicity observed by a network of sensors to master templates keyed to potential event locations.

Empirical MFP develops a catalog of matching templates from a collection of representative microearthquakes that uniformly samples the study volume. The earthquakes for the empirical master templates initially will have to be located using conventional earthquake location techniques and subsequently relocated using advanced processing techniques, however all future seismicity can be mapped using the computationally efficient MFP algorithm. In this paper, we apply this technique to recent seismic swarms that occurred in the Salton Sea Geothermal Field in November 2009 and January 2010.

GEOLOGICAL BACKGROUND

The Salton Sea Geothermal Field lies on the southeastern shore of the Salton Sea. The Salton Sea is the lowest part of the Salton Trough, a tectonic depression. Near the southern end of the Salton Sea, the San Andreas Fault appears to terminate at a spreading center called the Brawley seismic zone. This zone is the most northerly in a series of spreading centers distributed along the length of the Gulf of California that forms part of the East Pacific Rise. Rifting and intrusions produce high heat flow that metamorphoses the sedimentary rocks to shallow depths (Fuis *et al*., 1984).

DATA

Seven three-component seismic stations are located around the geothermal production field (Figure 1). This array is maintained by Caltech/USGS and continuous data has been archived at the Southern California Earthquake Data Center (SCEDC) since December 2007. Earthquake catalog locations and phase data is also available at the SCEDC.

For this study, we focus on data collected in November 2009 and January 2010. These two months contain the two largest seismic swarms that have occurred in the Salton Sea Geothermal Field since data has been archived at the SCEDC. The number of events that have occurred within 10 km of station HAT in November 2009 and January 2010 is 312 and 307, respectively.

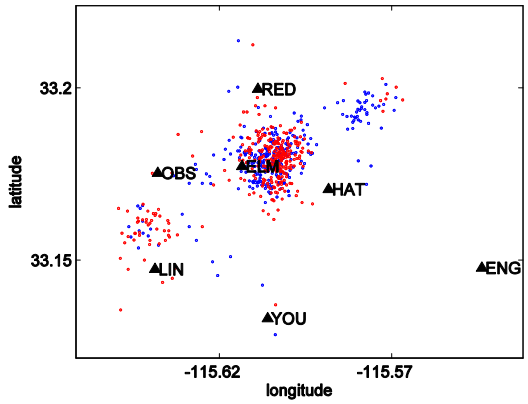

Figure1.Seven stations (black triangle with station names) map view in EN network. The dots are catalog events from SCEDC during Nov, 2009 (blue) and 2010 January (red).

METHODOLOGY

Our MFP technique is an adaptation of a signal-processing technique originally developed to locate continuous underwater acoustic sources. MFP can steer the array explicitly in the frequency domain using the complex phase and amplitude factors obtained by solving the wave equation through a propagation model. However, it is difficult to develop realistic Earth models to predict the structure of seismic wavefields at frequencies much above a tenth of a Hertz (Harris and Kvaerna 2010). An alternative to calculating the wavefield structure across an array is to estimate the structure directly from field calibration data, i.e., previous seismic events. We refer to this strategy as empirical MFP. In empirical MPF, the master templates that are created from the seismograms of previously detected micro-earthquakes thus contain contributions from direct and scattered seismic energy.

An example work flow is described in Figure 2. To determine which events we should choose for our field calibration events, we first obtain the SCEDC catalog of earthquake locations. The objective is to identify events that evenly sample the observed seismic distribution. However, since catalog locations can often have large location errors associated with them, we use the double difference method of Waldhauser and Ellsworth [2000] to relocate these microseismic events. The third step is to choose master template events as the input of the empirical match field processing methods. We visually inspect each potential master template to make sure that there are no overlapping events or noise spikes. Finally, we run the empirical MFP code on the continuous seismic data and identify events in the data stream that match the master templates.

In subsequent sections, we describe the particulars of our double-difference earthquake relocation study and the identification of master templates.

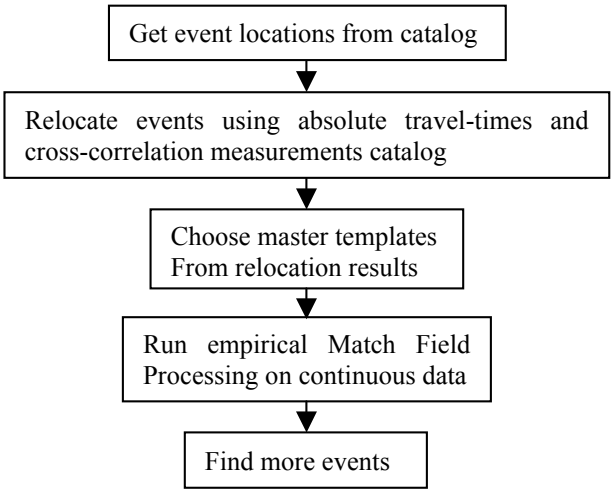

Figure2. Empirical match field processing work flow.

DOUBLE DIFFERENCE LOCATIONS

We relocate Salton Sea Geothermal Field seismicity using the double-difference earthquake relocation program hypoDD (Waldhauser and Ellsworth, 2000). This method utilizes absolute P- and S-wave travel-time measurements and cross-correlation P- and S-wave differential travel-time measurements to determine high-resolution relative hypocenter locations. For this dataset, P picks are obtained from the SCEDC catalog and S-wave travel-time measurements are picked by hand from the horizontal components. If the S pick time on the two horizontal components are different, the highest confidence pick is used.

We determine the P-wave cross-correlation differential travel-time by using a 0.5 sec window with 0.1 sec ahead of the signal on the vertical components. For the S-wave, we rotate the horizontal components and use a 0.5 sec window with 0.1 sec ahead of the signal to determine the cross-correlation coefficient on both components. The final S-wave cross-correlation measurement is then the average travel time difference calculated from the radial and transverse components.

We investigate how two different 1D velocity models affect the relocation results. The first is a model by Holland (2002), which was determined by performing a 1-D inversion of the 1987 and 1988 data collected by LLNL and Unocal as part of the Salton Sea Scientific Drilling Project to study seismicity related to tectonics and geothermal activity. The second model is derived from the SCEDC 3D Southern California velocity model (Magistrle2000, Kohler2003). This new model was created using a combination of seismic refraction surveys, inversion of gravity observations, surface geology, and borehole data. Vp-density relationships are based on density measurements from oil well samples in the Los Angeles basin and the San Gabriel Valley, geotechnical boreholes throughout southern California, and 12 oil wells along the LARSE lines. As hypoDD can only accept 1-D velocity models, we simplify the southern California velocity model by averaging the velocity profiles under our seven seismic stations. We found that the depth thickness weighted average Vp/Vs for the SCEDC model is 1.86, which is higher than the 1.732 value used in Holland's model. In the following, we will refer to the Holland (2002) model as "Holland02" and to the SCEDC 3D model as "SoCal3D".

Table 1(a): Holland02 velocity model.

Depth (km)	0	1	2	3	4	7	10	13
Velocity (km/s)	2.02	2.46	2.98	4.26	4.97	5.28	6.05	6.45

Table 1(b): SoCal3D velocity model.

Depth (km)	0	1	2	3	4	7	10	13
Velocity (km/s)	1.80	3.25	4.59	4.80	4.93	5.80	6.44	6.77

We optimize the data weighting and re-weighting scheme in hypoDD by iterating over different combinations of *a priori* weights (WTCCP, WTCCS, WTCTP, WTCTS), re-weighting values (WRCC, WRCT, WDCC, WDCT) and dampening factors. Table 2 summarizes the weighting schemes used in this analysis. In Table 2(a) we down-weight the cross correlation data to allow the catalog data to restore the large-scale picture in the first set of iterations (1-5). In the second set of iterations (6-11), we keep the same relative *a priori* weighs, but re-weight the catalog data according to misfit and event separation to remove or down-weight outliers. In the third set of iterations (12-17), we shift the control of relocation to the catalog data. No misfit dependant weight is applied to the cross correlation data in this set of iterations, but it is applied in the fourth set of iterations (18-24). Note that the main difference between Table 2(a) and (b) are the iteration times.

Table 2(a): Weighting scheme in HypoDD using Holland02 model.

NITER	WTCCP	WTCCS	WRCC	WDCC	WTCTP	WTCTS	WRCT	WDCT	DAMP
1-5	0.01	0.01	-9	-9	1.0	0.5	-9	-9	100
6-11	0.01	0.01	-9	-9	1.0	0.5	6	2	100
12-17	1.00	0.5	6	2	0.5	0.05	-9	-9	100
18-24	1.00	0.5	6	2	0.5	0.05	6	2	100

Table 2(b): Weighting scheme in HypoDD using SoCal3D model.

NITER	WTCCP	WTCCS	WRCC	WDCC	WTCTP	WTCTS	WRCT	WDCT	DAMP
1-6	0.01	0.01	-9	-9	1.0	0.5	-9	-9	100
7-11	0.01	0.01	-9	-9	1.0	0.5	6	2	100
12-15	1.00	0.5	6	2	0.5	0.05	-9	-9	100
16-21	1.00	0.5	6	2	0.5	0.05	6	2	100

The hypoDD relocation results are shown in Figure 3. Using the Holland02 model, hypoDD is able to relocate 533 events. 555 events are relocated using the SoCal3D model. The event locations are plotted in 3D view using a time-based color scheme from blue (old) to red (new). We found that hypoDD improved the locations of the swarms. The clusters are tighter and finer structure can be observed in more detail. The two velocity models generate similar patterns of the clusters. There are three major swarms on the relocated results. The major one is in the middle and the other two smaller swarms are located at each side. In Figure 3, the original time of the earthquakes is represented by the color. It shows that events in the major center swarm occurred in both November 2009 and January 2010. But events in the northeast side-swarm mostly occurred in November 2009 while events in the southwest side-swarm mostly occurred in January 2010. There also are a few isolated events located at lower depth by both velocity models. In addition, we found that SoCal3D velocity model generates a finer structure, while relocation results using Holland02 model are more diffuse (Figure 3a and 3b).

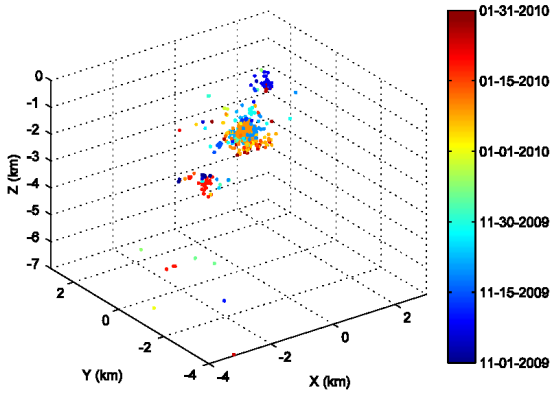

Figure 3(a). 3D view of relocation results using Holland02 velocity model. The color represents the event original time from the oldest (blue) to the newest (red).

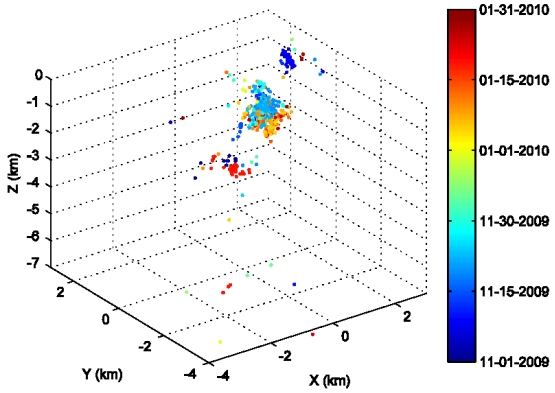

Figure 3(b). 3D view of relocation results using SoCal3D velocity model. The color represents the event original time from the oldest (blue) to the newest (red).

MASTER EVENTS SELECTION

After the events are relocated, we choose the master events as the template of the empirical MFP. We require that the master events be spatially evenly distributed so that the templates are representative of the entire study volume. We first divide the whole volume into 1km×1km×1km cubes. In each cube, the number of master events (n) is based on the number of relocated events (m) falling into this cube according to the relation ⍰=2×3m. As a consequence, when more events occur in a cube, more master events are chosen and vice versa. However, if events in one cube are concentrated in a small area, we subdivide the cube into 8 equal sub-cells. Then we use the same rule to decide how many master events should be taken from each cell. We subdivide the cube iteratively until the master events

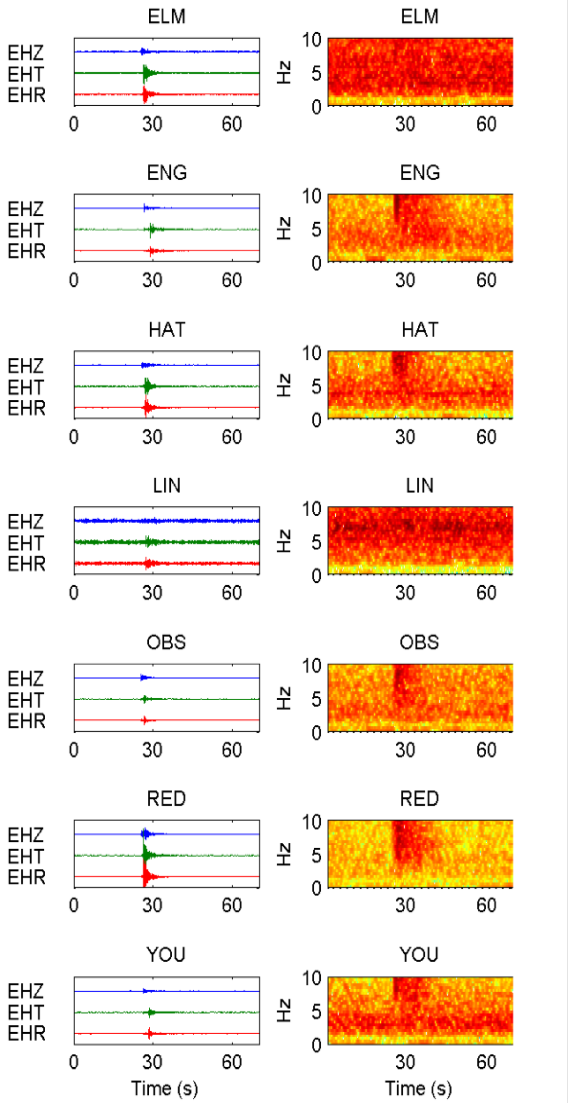

Figure 4. Example of one master event with origin time at 2009.11.23, 00:00:46.15. The left column is the waveform with the station name on top. The right column is the spectrogram with the station name on top.

are resolved in different cells to make sure the selected ones are spatially even distributed.

Figure 4 shows one example of master events with origin time at 2009.11.23, 00:00:46.15. In general, the background noise is incoherent across the 70 sec time window at station ELM, ENG, HAT, OBS, and RED. In this particular example stations ELM and LIN display much more background noise over the frequency band in which we are searching for events. At station YOU, some narrow band noise is observable in the spectrogram. In general, LIN is consistently noisy. Therefore, we decided to remove station LIN due to its poor quality of recording. The

other stations do not consistently display such high background noise energy.

MFP RESULTS AND DISCUSSION

We run the empirical MFP code on two months of continuous data during November 2009 and January 2010. The empirical MFP code performs its calculation on the continuous data using a 70-sec window which steps forward 1 second at one time. The frequency band that we are examining is between 2 – 8 Hz. Figure 5 shows a 10-day example of the output of the empirical MFP code. This segment of data is from January 10th – 20th, 2010. The y-value at each time point indicates the normalized detection statistic. A value of 1 would indicate an exact match between the template and the incoming seismicity at that particular time. In Figure 5, Detector 10 (origin time: 2010/015/11:27:48) and 13 (origin time: 2010/016/01:24:47) detected events in a similar pattern. This indicates that these two master events are similar to each other. We can confirm that by plotting up the location of these four master events (see Figure 6). This is reasonable because these two events occurred 12 hours apart only, they might be from the same fault or at least have similar source mechanism and they are close to each other so that they have similar wavefields along the path. Detector 34 is somewhat more distant from Detectors 10 and 13, but still belongs to the major swarm. This detector is able to detect some other events that are not detected by Detectors 10 and 13. This illustrates the fact that the more spatially evenly sampled the master events are, the higher chance we will have to be able to detect more events in that area. The detection statistics using Detector 44 mostly fall below the threshold. This is because detector 44 is far away the swarm and there are not many events with similar feature occurred close to detector 44 during this particular time period.

After detection statistics are computed for each master template, we compare the matches obtained from one detector and compare it to matches at other detectors. The detector with the largest match is then determined to have detected an event.

In November 2009, the empirical MFP code detected 545 events. There were 312 events in the SCEDC catalog. 223 catalog events were detected by the code and 89 catalog events were not detected by the code. In January 2010, the empirical MFP detected 549 events. There were 307 catalog events. 242 catalog events were detected by the code and 65 catalog events were not detected by the code (see Figure 7). We believe that this may be due to the somewhat sparse coverage of our master events.

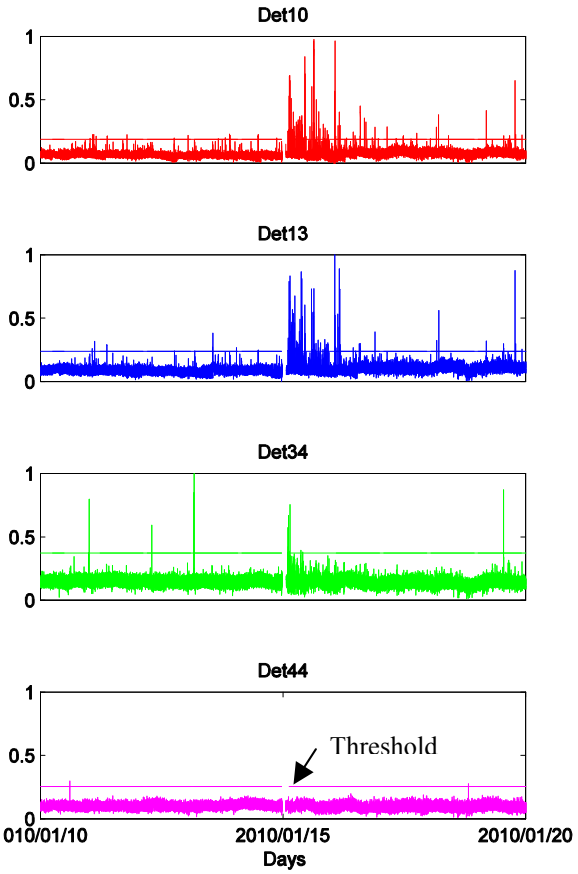

Figure 5. Empirical MFP detection using four master events over one segment during Jan, 10th – 20th, 2010. The threshold for each master event is plotted on top. Threshold is marked by the black arrow.

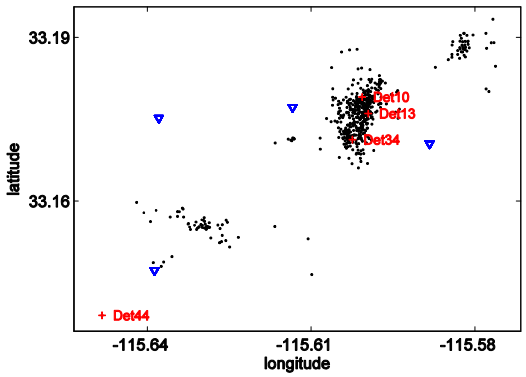

Figure 6. Location of four template detectors used in figure 5. They are marked by red symbols. The background seismicity is the block dot and the stations are blue triangles.

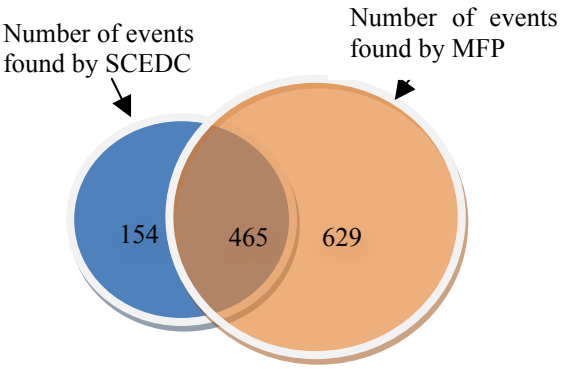

Figure 7. Number of events found by the SCEDC catalog as well as the number of more events found by empirical MFP. There are 465 events are found by both methods, and 629 events are newly detected by the empirical MFP. Another 154 events reported by SCEDC are not detected by empirical MFP.

CONCLUSIONS

Match field processing with empirically calibrated steering vectors is able to detect more events than can be detected using conventional techniques. Empirical MFP does not require a plane wave assumption as most array processing methods do. Therefore, empirical MFP has more adaptability to varying noisy environment as long as the master templates adequately cover the area where future events will possibly occur. Our test on the continuous data during November 2009 and January 2010 demonstrates the detection capability using empirical MFP. There are 1094 events detected in total by MFP, while the catalog reports 619 events. We believe that this shows that empirical MFP significantly improves seismic array detection ability.

ACKNOWLEDGEMENTS

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract ED-AC52-07NA27344. This work is funded by the American Recovery and Reinvestment Act, Pub. L. 111-5.

REFERENCES

Waldhauser, F. (2001), "hypoDD -- A program to compute double-difference hypocenter locations," *U.S. Geological Survey Open-File Report*, 01-113.

Waldhauser, F., and William L. E. (2000), "A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California," *Bulletin Seismological Society of America*, **90**, 1353-1368.

Hooland, A. A. (2002), Microearthquake Study of the Salton Sea Geothermal Field, California: Evidence of stress triggering. Idaho National Engineering and Environmental Laboratory, Geosciences Department. *Thesis*, The University of Texas at El Paso.

Magistrale, H., S. Day, R. W. Clayton, and R. Graves, (2000), The SCEC southern California reference 3D seismic velocity model Version 2, *Bulletin Seismological Society of America*, **v. 90**, no. 6B, S65-S76

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

Harries, B. D., and T. Kvaerna, (2010), Superresolution with seismic arrays using empirical matched field processing, *Geophysical Journal International*, **182**, 1455-1477.

Fuis, G. S., W. D. Mooney, J. H. Healey, G. A. McMechan, and W. J. Lutter, (1984), A seismic reflection survey of the Imperial valley, California. *Journal of Geophysical Research*, **89**, 1165-1189.