Advancement in Source Estimation Techniques Using Broadband Regional Seismograms

by Lupei Zhu and Donald V. Helmberger

Abstract One important constraint on source retrieval from regional seismograms comes from the amplitude difference between various phases (such as *Pnl*/surface wave, *SV/SH*). Because the misfit errors used in some waveform inversions are normalized by the data and synthetics, the amplitude information in the data has not been fully utilized. In this article, we modify the "cut and paste" source estimation technique (Zhao and Helmberger, 1994) by removing this type of normalization. It is shown that the modified method increases the stability and resolution of inversion. When multiple stations at different distance ranges are used, a distance scaling factor is introduced to compensate for the amplitude decay with distance. By applying the technique to the TERRAscope data, we have determined source mechanisms and depths of 335 southern Californian events with $M_L \geq 3.5$. The amplitude decays with distance are $r^{1.13}$ for *Pnl*, $r^{0.55}$ for Love waves, and $r^{0.74}$ for Rayleigh waves. In contrast to generally shallow source depths reported by the southern California short-period network, the depth distribution from waveform inversion shows a strong peak around 12 km with few earthquakes occurring above 5 km and below 20 km.

Introduction

Significant progress has been made recently in retrieving source mechanisms from regional broadband seismograms. A consequence is that the magnitude threshold of events that can be analyzed has been lowered to less than 4. Because of the frequent occurrence and relatively simple source functions of earthquakes in the magnitude range of 3 to 5, their waveforms are ideal for investigating regional structures, which is of great importance for delineating fault structures and understanding path effects for ground motions from large earthquakes.

Two kinds of regional waveform data are typically used for source estimation: surface waves (Patton, 1980; Patton and Zandt, 1991; Thio and Kanamori, 1995) and body waves (Wallace and Helmberger, 1982; Fan and Wallace, 1991; Fan et al., 1994; Dreger and Helmberger, 1993). Generally, body waves are less affected by shallow heterogeneities and are more stable than surface waves, although they have a lower signal-to-noise ratio due to their smaller energy. There have been several inversion methods proposed recently using whole waveforms of seismograms (Walter, 1993; Ritsema and Lay, 1993; Zhao and Helmberger, 1994; Nabelek and Xia, 1995). Most of these inversions are mainly controlled by surface waves, particularly because they are performed using long-period waveforms. An exception is the "cut and paste" (CAP) method by Zhao and Helmberger (1994), which breaks broadband waveforms into *Pnl* and surface-wave segments and inverts them independently. The source mechanism is obtained by applying a direct grid search through all possible solutions to find the global minimum of misfit between the observations and synthetics, allowing time shifts between portions of seismograms and synthetics. One of the advantages of the technique is that it proves insensitive to velocity models and lateral crustal variation.

Because the misfit errors defined in Zhao and Helmberger (1994) and others (Wallace and Helmberger, 1982; Fan et al., 1994) are normalized by the data and synthetics, the amplitude information in the data has not been fully utilized to constrain the source orientation and depth. The purpose of this article is to strengthen the CAP technique by removing this normalization and allowing better use of amplitude information. As multiple stations at different distance ranges are often used in source inversion, we will also investigate the amplitude decays with distance range for different crustal phases and correct them by introducing a distance scaling factor. Some applications of the modified technique to the southern California regional events are presented.

Source Estimation Using True Amplitude Waveforms

Let $u(t)$ be the observed displacement. The corresponding synthetic displacement $s(t)$ for a double-couple source can be expressed as

$$s(t) = M_0 \sum_{i=1}^{3} A_i(\phi - \theta, \delta, \lambda) G_i(t), \qquad (1)$$

here, $i = 1, 2, 3$ corresponds to three fundamental faults, i.e., vertical strike slip, vertical dip slip, and 45° dip slip. G_i's are the Green's functions, A_i's are the radiation coefficients, and ϕ is the station azimuth. M_0 is scalar moment; θ, δ, and λ are strike, dip, and rake, respectively, of the source that we want to determine from $u(t)$. They can be estimated by solving the equation

$$u(t) = s(t). \qquad (2)$$

Since there are only limited unknown parameters and all of them are limited to a range of values ($0 \leq \theta \leq 2\pi$, $0 \leq \delta \leq \pi/2$, $0 \leq \lambda \leq 2\pi$), it is convenient to solve the above nonlinear equation by grid search method. We define an object function to measure the misfit error between u and s and search through the parameter space to find the global minimum of the object function.

In Zhao and Helmberger (1994), the misfit error is defined as the norm ($L1$ or $L2$) of the difference between u and s normalized by the norms of both u and s:

$$e = \frac{\|u - s\|}{\|u\| \cdot \|s\|}. \qquad (3)$$

Because Pnl usually has smaller amplitude than surface waves, this normalization helps to weight Pnl and surface waves equally. It also prevents the inversion from being completely dominated by the strongest station, which is usually the nearest station, if several stations at different distance ranges are used. However, the amplitude information is lost during the normalization. Some of this information, such as amplitude ratios of Pnl-to-surface waves and SV-to-SH, provide important constraints on the source orientation and depth. A more severe problem with this normalization is that it introduces singularities in the source parameter space at those points where source orientation generates nodal synthetics (where the norm of synthetics vanishes). In the case when the data include nodal records, the grid search will miss the true minimum.

As an example, Figure 1a shows the misfit error as defined by equation (3) as a function of dip and rake (the location of the event and the waveforms are displayed in Fig. 2). The global minimum is very obscure. The error surface is distorted by some peaks and ridges that are associated with singularities introduced by the normalization. Figure 1b is the misfit error using true amplitudes without normalization:

$$e = \|u - s\|. \qquad (4)$$

It has a well-defined global minimum at $\delta = 60°$ and $\lambda = -10°$. Note that the P-SV waves at station PFO, GSC, and ISA are close to nodal (Fig. 2), which are well matched by

the synthetics of solution b. But the synthetics of solution a have larger P-SV amplitudes compared with the data. This particular example of waveform modeling exposes the difficulties encountered at nodes. Since the P-SV motions are near nodal to the north, the radial and vertical components are easily contaminated by SH motions (see reference lines in Fig. 2). Only the PAS observations display a true Rayleigh wave with the expected lag time behind the Love waves. The cut and paste aspect of the search procedure moves the synthetic Rayleigh wave forward, attempting to capture some of this energy to constrain mechanism (see ISA). The modified method tends to suppress such spurious attempts.

Distance Range Scaling

Using true-amplitude waveforms for source inversion usually leads to the problem of the closest station dominating the inversion when stations are distributed over a large distance range. Figure 3 shows misfit errors as a function of distance. These misfit errors were obtained through a global search applying expression (4) to 335 southern Californian regional events of $M_L \geq 3.5$. Standard southern California crustal model (SC model, see Dreger and Helmberger, 1993) was used in the inversion.

The misfits of surface waves have larger scatter than body waves, which are expected because surface waves are more easily affected by shallow heterogeneity. It has been shown that Pnl at a range of 300 to 1000 km is quite stable (Helmberger and Engen, 1980) and easily inverted for source mechanisms (Wallace and Helmberger, 1982). At closer range, the details of the Moho transition plays a more important role as well as the PL waves trapped in the shallow crust (Song and Helmberger, 1996). Since both of these features show strong local variation, we should expect the large scatter displayed in Figure 3 at closer distances. However, because the Pnl's at the nearest stations play an essential role in early warning (Scrivner and Helmberger, 1995), the local velocity structure should probably be added to each station. This approach will be pursued in future efforts.

Despite the scatter of data in Figure 3, the misfit errors show a rapid decay with distance. Since radiation patterns have been taken out, this decay is related to the amplitude decay due to geometrical spreading and attenuation. To compensate for this decay, we introduce a distance range scaling factor and define the misfit error for a record at a distance r as

$$e = \left\| \left(\frac{r}{r_0} \right)^p \right\| \cdot \|u - s\|, \qquad (5)$$

here p is a scaling factor to give the record at r the same weight as that at reference distance r_0. If we assume a spherical geometrical spreading for body waves and cylindrical geometrical spreading for surface waves, an appropriate choice of p would be $p = 1$ for body waves and $p = 0.5$
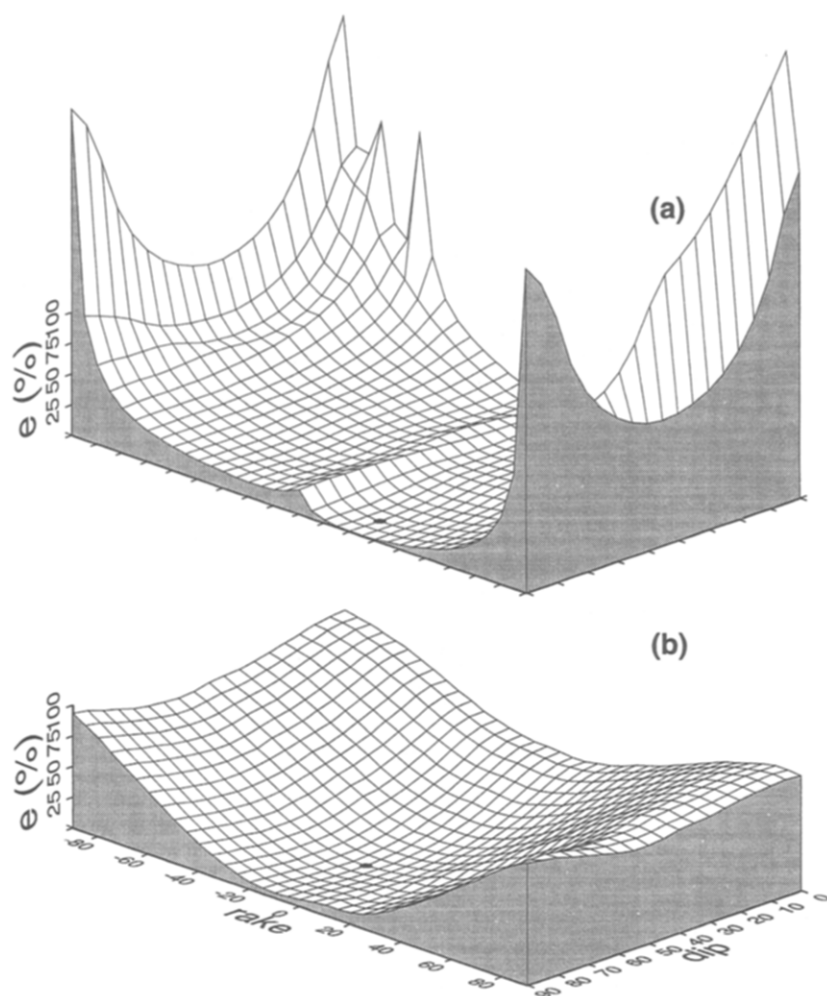
Figure 1.   Waveform misfit errors as a function of dip and rake for one of southern California events (24 May 1992, 12:22) (see Fig. 2). (a) The misfit error using normalization (expression 3 in text), solution is θ 70°, δ 80°, λ 20°, M_w 3.9, h 14 km. (b) The misfit error using the true amplitude of the recorded waveforms (expression 4), solution is θ 245°, δ 60°, λ − 10°, M_w 3.9, h 17 km. The global minima from the grid search are indicated by black dots.
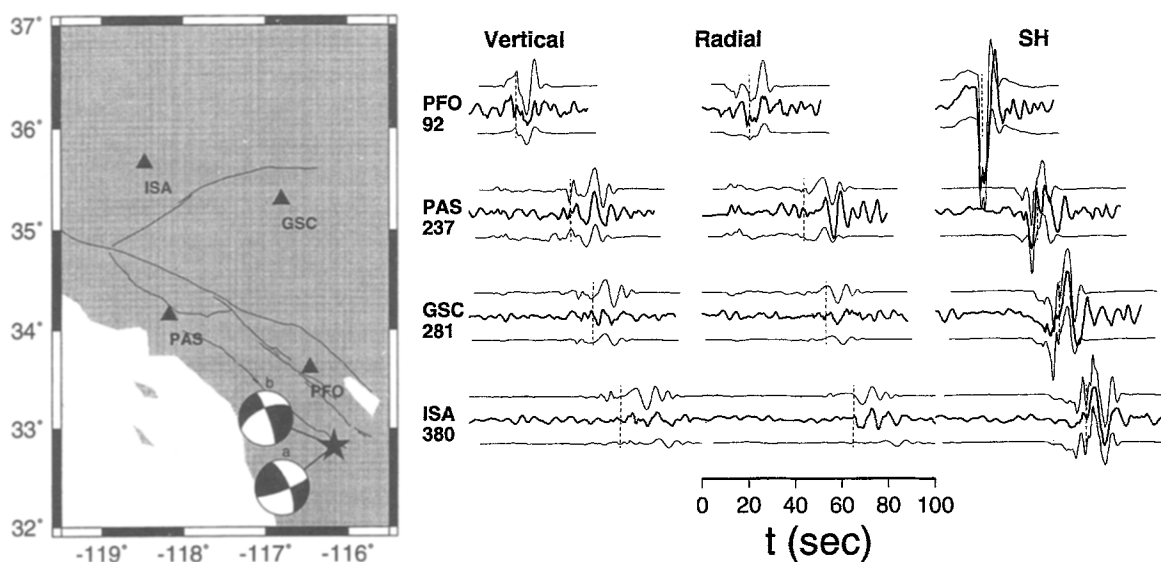
Figure 2.   The location of event (star) used in Figure 1. The focal spheres correspond to the solution (a) and (b) in Figure 1, and their synthetics are plotted above (solution a) and below (solution b) the data trace. Reference lines are arrivals of Love waves as determined from the tangential components. Numbers below station names are epicentral distances in km.
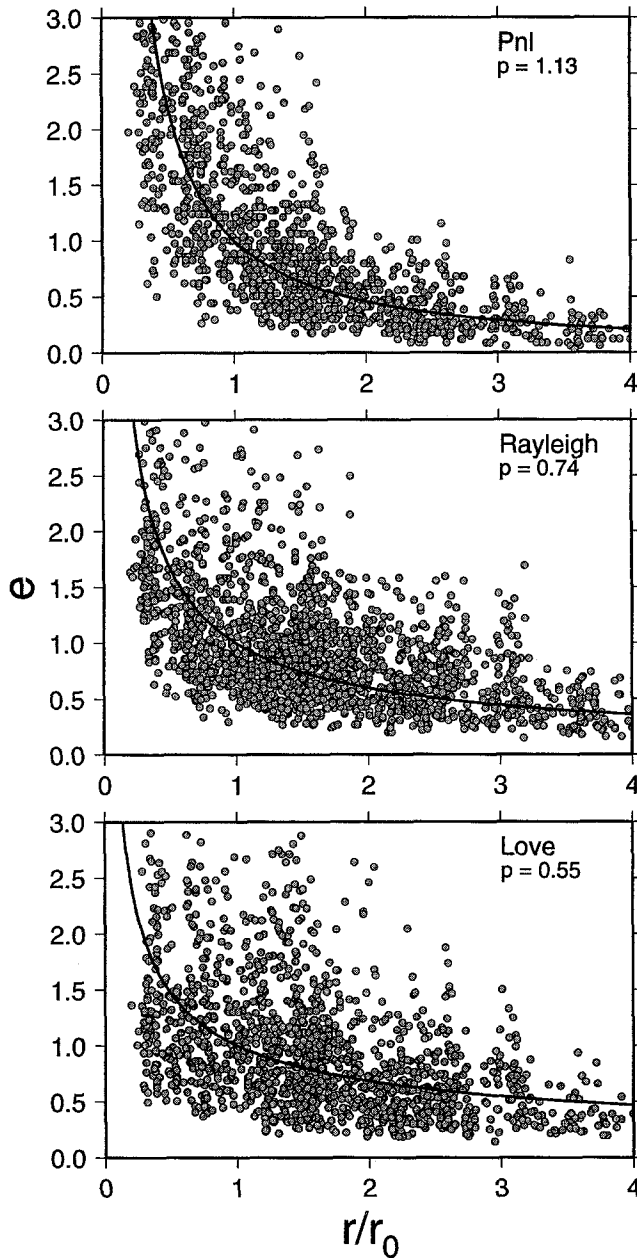
Figure 3. Misfit errors of different portions of seismograms of 335 southern California events. True amplitude waveforms are used in the inversion. We chose $r_0 = 100$ km as reference distance. The solid lines are $(r_0/r)^p$, where the p values are determined from a least-square fit.

Applications

Because there is essentially no manual intervention required for this technique, it can be fully automated. We have implemented it using data from TERRAscope. As soon as a regional event larger than 3.5 occurred, preliminary event location is used to trigger the code to invert for source mechanism using real-time waveform data. Usually it takes 3 min on a SUN-SPARC 10 workstation for the code to complete a 10° spacing grid search while fixing the source depth. If needed, the best focal depth can be determined through an additional iteration in a range of depths.

We present the recent Ridgecrest earthquake in southern California as an example (Fig. 4). This event, which occurred on 20 September 1995, is the largest (M_L 5.8) since the 1994 Northridge earthquake. The best solution of the waveform inversion gives an almost purely strike-slip mechanism with one nodal plane striking N30°W. This mechanism is consistent with the first-motions focal mechanism (Hauksson et al., 1995). To give an estimate of uncertainty of the solution, we plot misfit error as a function of depth (Fig. 4b) and contours of misfit values around the best solution (Fig. 4c). They show that the strike is well defined (within ±10°), while the rake and dip are less well constrained (±30°). The best focal depth of 14 km determined from the inversion is deeper than the 5 km depth in the Southern California Seismographic Network (SCSN) catalog. Figure 5 shows a comparison of data with synthetics for the best solutions at the depths of 5 and 14 km. Note that all the TERRAscope stations, except NEE, show strong SH components and weak vertical and radial components because most of them are close to the P-SV nodal directions for this event (Fig. 4). The synthetics match the observations well both in shape and amplitude. This indicates that the amplitude adjustment with distance scaling is about correct and that the Green's functions are satisfactory. For all stations, except DGR, the synthetics with the deeper focal depth fit the data better. Some paths cause difficulties, such as the path to SNCC (San Nicolas Island). Some stations, also, show complexities such as the large apparent Rayleigh motions arriving on the tangential component at PAS. Fortunately, good station coverage overwhelms these types of misfits.

We have applied the technique to all available regional events of $M_L \geq 3.5$ recorded by TERRAscope back to 1990. A total of 335 focal mechanism solutions are obtained. Although a large number of them are aftershocks of the 1992 Landers and 1994 Northridge earthquakes, the data set still gives a good sampling of active source regions in southern California (Fig. 6). Figure 7 is the comparison of inverted focal depths and M_0 with the depths and M_L from the SCSN catalog. M_0 versus M_L essentially follows the relation found by Thatcher and Hanks (1973). But the waveform-inverted depths are generally deeper than the catalog depths.

for surface waves. For southern California, the p values determined from Figure 3 are 1.13 for Pnl, 0.55 for Love waves, and 0.74 for Rayleigh waves. The scatter in the figure is large, especially at the smaller distances, suggesting considerable variation in local structure and the need for regionalization. Nevertheless, it appears that the p values derived from geometrical spreading are reasonably good for the period range (5 to 100 sec) used in waveform inversion.
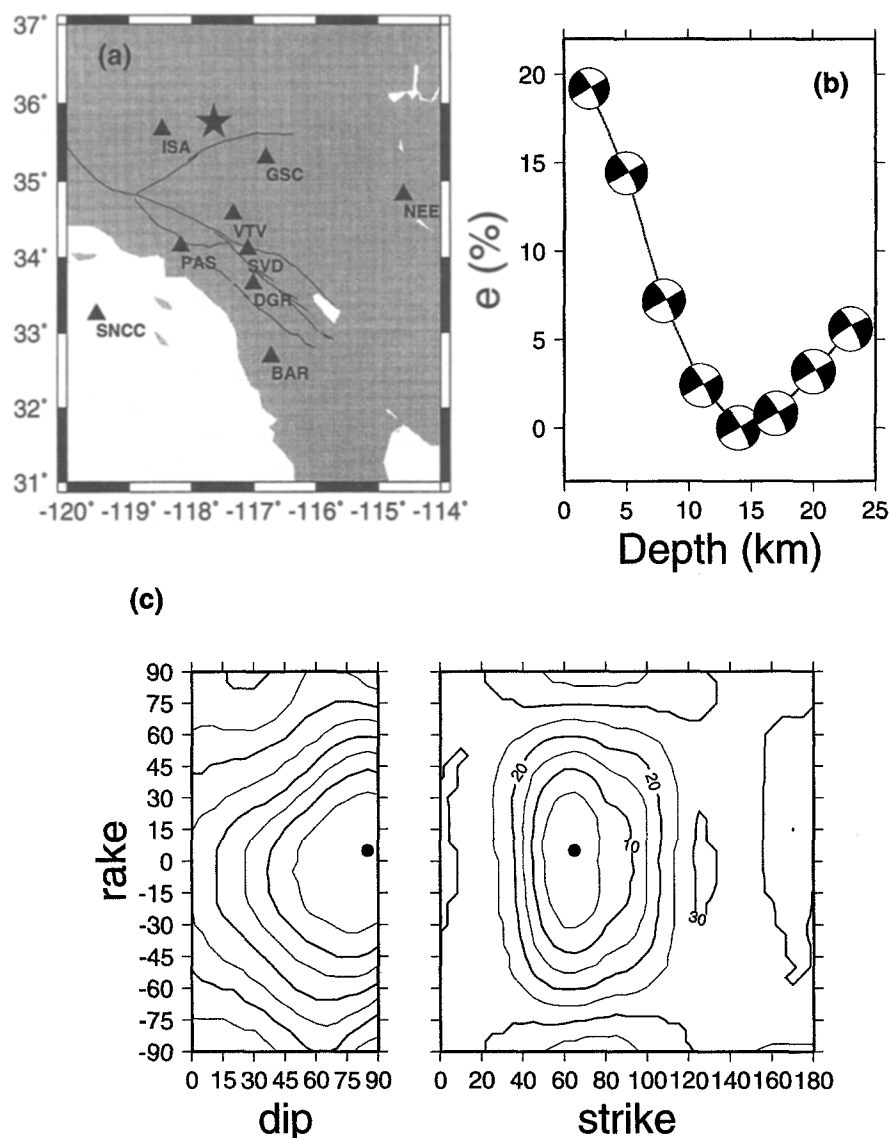
Figure 4.   (a) Map containing the location of the 20 September, 1995 Ridgecrest event (star) and the recording TERRAscope stations. (b) Misfit errors as function of depth. (c) Contours of misfit errors around the best solution (indicated by the black dot). The contour interval is 5% of the minimum value.

Focal Depth Distribution and the Seismogenic Zone in Southern California

Figure 8 shows the histograms of SCSN catalog depths and waveform inversion depths for the 335 events included in this study. The catalog depths are determined from the P arrival times as recorded by southern California short-period network. Large uncertainty exists for these routinely determined depths, mostly caused by trade-off between depth and origin time and some cases sparse station distribution. The most accurate information on depth distribution of seismicity is usually obtained from aftershock sequences during which temporary stations are deployed directly above the mainshock rupture. However, the interpretation of aftershocks

can be complicated because the mainshock greatly changes the local stress field. Preshocks and background events probably provide a more accurate picture of the evolving stress field. Since our waveform inversion technique conserves the amplitude information among various crustal phases, source depths should be well constrained, even for small-magnitude earthquakes.

The depth distribution obtained from waveform inversions shows a strong peak at about 12 km, with few events occurring above 5 km and below 20 km. This seems quite compatible with the expected seismogenic zone for a tectonically active region as southern California. It is generally thought that the lower boundary of seismogenic zone is due to a transition from brittle to ductile behavior in continental

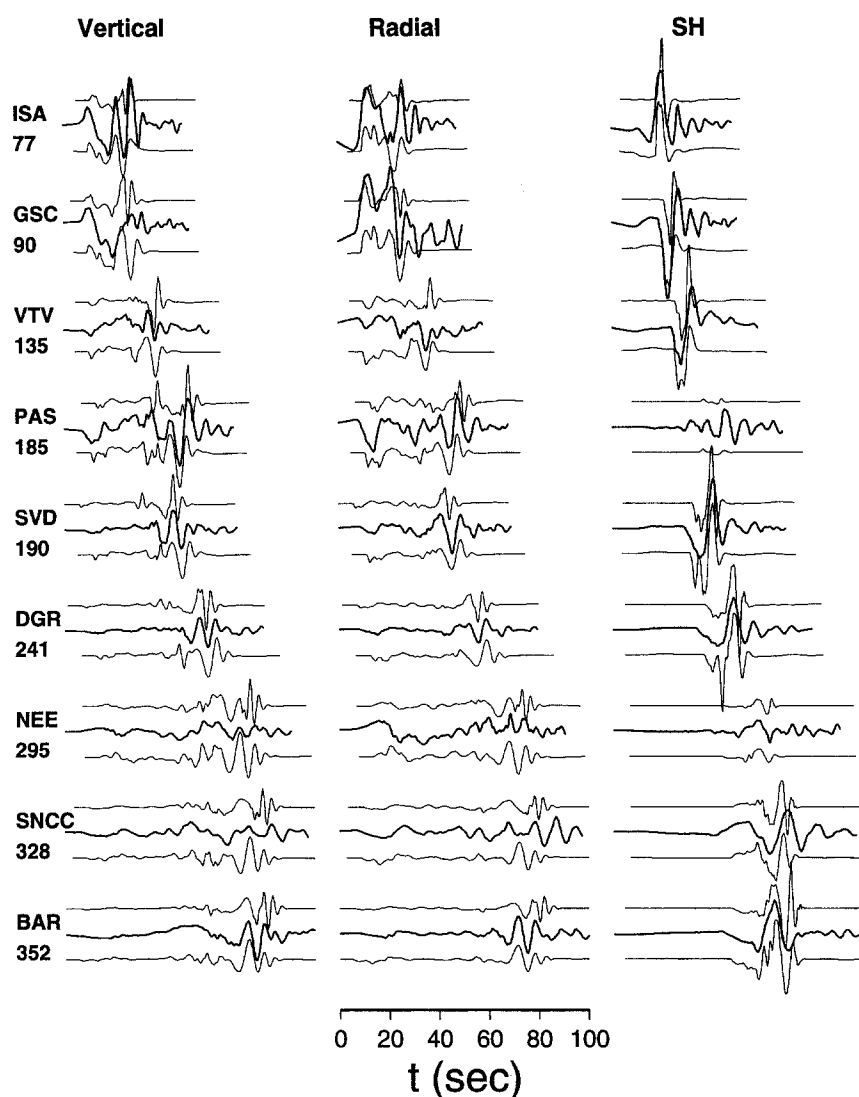**Vertical** **Radial** **SH**

0 20 40 60 80 100

t (sec)

Figure 5. Comparison of data (heavy traces) with synthetics of solution at a depth of 5 km (upper traces) and 14 km (lower traces) for the Ridgecrest event. The numbers below station names are epicentral distances in km. A distance compensation of *r* is used. Vertical and radial components are amplified by a factor of 2 with respect to the tangential for display reason.

crust (Meissner and Strehlau, 1982; Sibson, 1984; Doser and Kanamori, 1986). The top of seismogenic zone is controlled by a transition from stable sliding to stick slip, which can be attributed to the presence of unconsolidated fault gouge (Marone and Scholz, 1988). For a region overlain by sedimentary structures of unconsolidated material and with well-developed faults, they predict a minimum source depth of 3 to 5 km, which agrees well with our results.

This 1D picture of seismicity is obviously oversimplified in that there are well-known deep seismic zones (e.g., ANZA network) and shallow seismic zones (e.g., Imperial Valley). There is, also, considerable evidence for 2D and 3D structures, which is causing much of the scatter in the Figure 3. These latter features need to be incorporated in our simulation system to obtain better locations and develop the ability to recover second-order source characteristics on a routine basis.

Conclusions

In summary, we have improved the CAP technique by removing the normalization in the definition of misfit error. This modification helps to fully utilize the amplitude information of three-component records to give a better constraint on source orientations and depths. It is shown that both the stability and resolution power of inversion are increased. We also introduced a distance scaling factor to the misfit error formalism in order to compensate for the amplitude decay with distance when multiple stations at different distance ranges are used.

This technique has been fully automated and is presently servicing the TERRAscope data stream, where it takes just a few minutes to estimate source parameters. In batch mode, it takes a few days to rerun through the complete TERRAscope data archive of local events. Preliminary results from such runs suggest deeper faulting activity, by 6
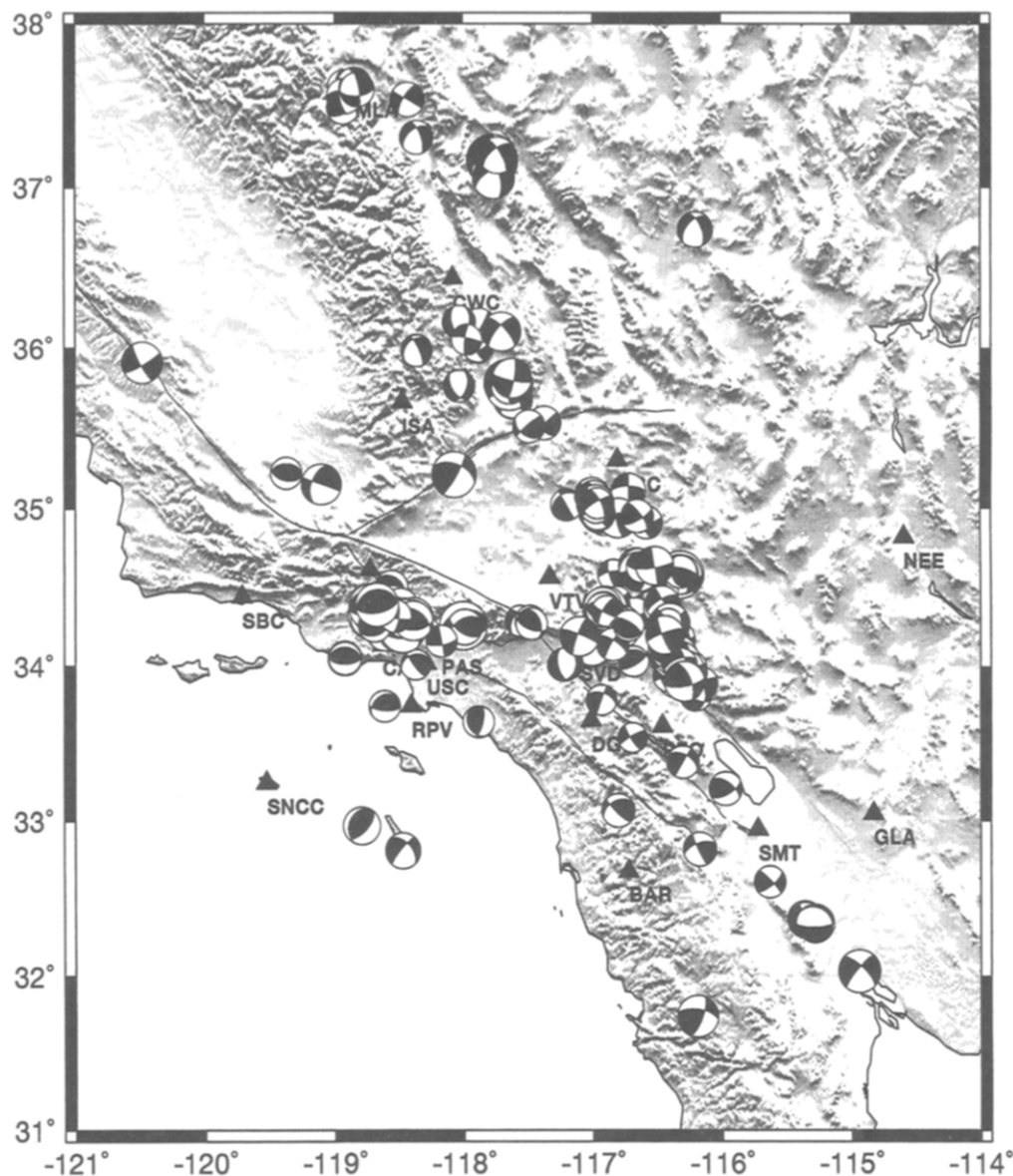
Figure 6. Focal mechanisms determined from waveform inversions of 335 regional events (see text). The sizes of the focal spheres are proportional to M_w.

km on average, than obtained from the SCSN catalog. Resolution of such disagreements and local regionalization should prove much easier with this new tool.

References

Doser, D. I. and H. Kanamori (1986). Depth of seismicity in the Imperial Valley region 1977–1983 and its relationship to heat flow, crustal structure, and the October 15, 1979 earthquake, *J. Geophys. Res.* **91,** 675–688.

Dreger, D. and D. V. Helmberger (1993). Determination of source parameters at regional distances with three-component sparse network data, *J. Geophys. Res.* **98,** 8107–8125.

Fan, G. W. and T. C. Wallace (1991). The determination of source parameter for small earthquakes from a single very broadband seismic station, *Geophys. Res. Lett.* **18,** 1385–1388.

Fan, G. W., S. L. Beck, and T. C. Wallace (1994). The seismic source parameters of the 1991 Costa Rica aftershock sequence: evidence for a transcurrent plate boundary, *J. Geophys. Res.* **98,** 15759–15778.

Hauksson, E., H. Kutton, H. Kanamori, L. Jones, J. Mori, S. Hough, and G. Roquemore (1995). Preliminary report on the 1995 Ridgecrest
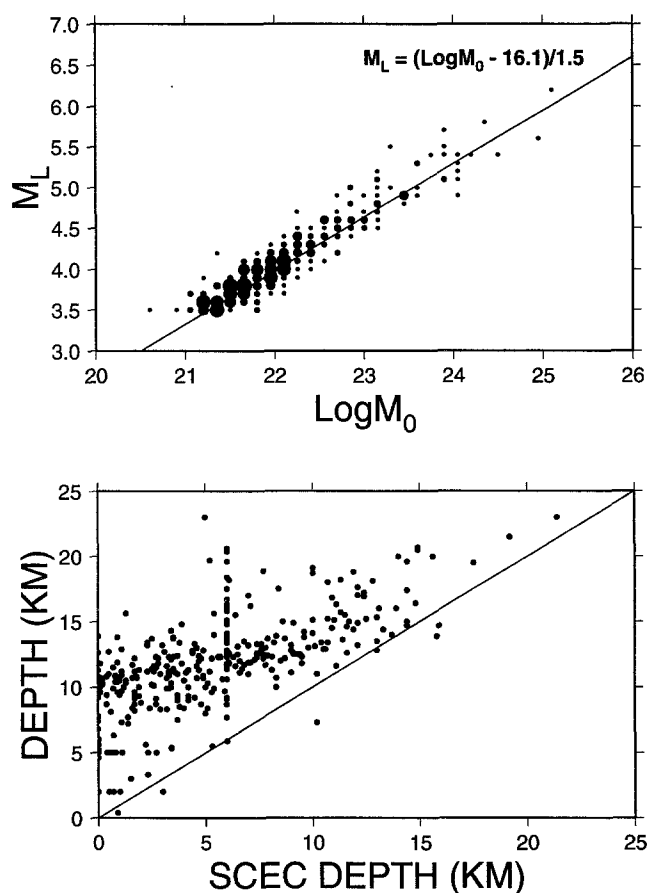
$$M_L = (LogM_0 - 16.1)/1.5$$

Figure 7. Upper panel displays the SCSN catalog M_L versus M_0 determined by the waveform inversion. These results can be compared against the relation determined by Thatcher and Hanks (1973). Lower panel contains the depths estimated from the waveform inversion versus the SCSN catalog depths.

earthquake sequence in eastern California, *Seism. Res. Lett.* **66**, 54–60.

Helmberger, D. V. and G. Engen (1980). Modeling the long-period body waves from shallow earthquakes at regional ranges, *Bull. Seism. Soc. Am.* **70**, 1699–1714.

Marone, C. and C. H. Scholz (1988). The depth of seismic faulting and the upper transition from stable to unstable slip regimes, *Geophys. Res. Lett.* **15**, 621–624.

Meissner, R. and J. Strehlau (1982). Limits of stresses in continental crust and their relationship to the depth-frequence distribution of shallow earthquakes, *Tectonics* **1**, 73–89.

Nabelek, J. and G. Xia (1995). Moment-tensor analysis using regional data: application to the 25 March, 1993, Scotts Mills, Oregon, earthquake, *Geophys. Res. Lett.* **22**, 13–16.

Patton, H. J. (1980). Reference point method for determining the source and path effects of surface waves, *J. Geophys. Res.* **85**, 821–848.

Patton, H. J. and G. Zandt (1991). Seismic moment tensors of western U.S. earthquakes and implication for the tectonic stress field, *J. Geophys. Res.* **96**, 18245–18259.

Ritsema, J. and T. Lay (1993). Rapid source mechanism determination of large ($M_w \geqq 5$) earthquakes in the western United States, *Geophys. Res. Lett.* **20**, 1611–1614.

Scrivner, C. W. and D. V. Helmberger (1995). Preliminary work on an early warning and rapid response program for moderate earthquakes, *Bull. Seism. Soc. Am.* **85**, 1257–1265.

Sibson, R. (1984). Roughness at the base of the seismogenic zone: contributing factors, *J. Geophys. Res.* **89**, 5791–5799.

Song, X. J. and D. V. Helmberger (1996). Broadband modeling of regional seismograms, the Basin and Ranges crustal structure, *Geophys. J. Int.* (in press).

Thatcher, W. and T. C. Hanks (1973). Source parameters of southern California earthquakes, *J. Geophys. Res.* **78**, 8547–8576.

Thio, H. K. and H. Kanamori (1995). Moment-tensor inversions for local earthquakes using surface waves recorded at TERRAscope, *Bull. Seism. Soc. Am.* **85**, 1021–1038.

Wallace, T. C. and D. V. Helmberger (1982). Determination of seismic parameters of moderate-size earthquake from regional waveforms, *Phys. E-Plan.* **30**, 185–196.

Walter, W. R. (1993). Source parameters of the June 29 Little Skull Mountain earthquake from complete regional waveforms at a single station, *Geophys. Res. Lett.* **20**, 403–406.

Zhao, L. S. and D. V. Helmberger (1994). Source estimation from broadband regional seismograms, *Bull. Seism. Soc. Am.* **84**, 92–104.

Seismological Laboratory
California Institute of Technology
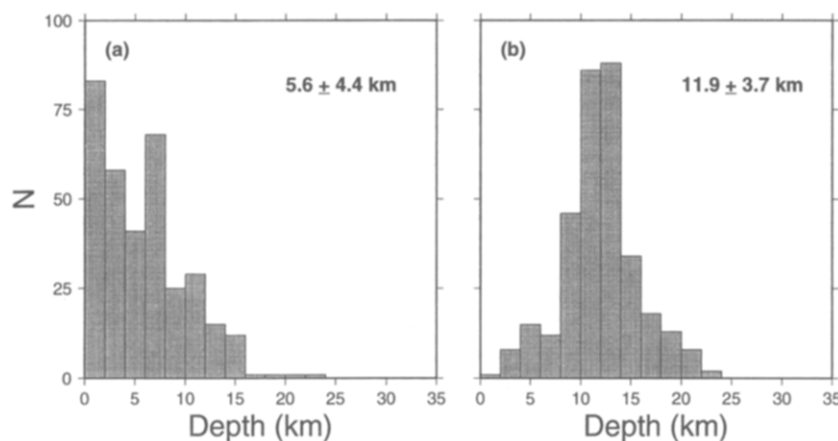Pasadena, California 91125

Figure 8. Histograms of depth distribution in southern California: (a) SCSN catalog depths and (b) waveform inverted depths. The numbers in the figure are the average focal depths and the standard deviation.