ADAPTIVE TECHNIQUES FOR ESTIMATING AND LOCATING SMALL EVENTS USING REGIONAL WAVEFORM DATA

Don V. Helmberger, Chen Ji, and Brian Savage
Seismological Laboratory
California Institute of Technology

Sponsored by The Defense Threat Reduction Agency
Arms Control Technology Center
Nuclear Treaties Branch

Grant Number DSWA01-98-1-0010

ABSTRACT

Attempts at locating and identifying small events could benefit greatly from a better appreciation of regional seismograms. Essentially, we need more information from the few regional seismograms available than just arrival times and amplitude ratios. Here we discuss a method of estimating source parameters using as few as one broadband station in conjunction with travel-time and polarities from at least one more station. The method employs an adaptive grid-search of matching three-component waveform records against synthetics to establish source location and depth. The basic matching procedure contains a trade-off of source mechanism with location. The better the constraint on mechanism and depth, the better the location. It appears that refined depths and origin times can be obtained for GT5 events if the first-motion polarities are available to constrain the focal plane. The locations of GT10 and GT25 and EHB (Engdahl, Kennett, and Buland) events can also be improved if CMT's and other independent depth estimates are available.

Preliminary results applied to the PASSCAl data recorded in Pakistan and Tibet (China) and TriNet (CAL) proved highly effective when locating events located approximately between two stations. If both stations have usable three-component data, we can determine depth and source mechanism and refine the location. If we assume the mechanism and depth are known (Master Reference Event), we can relocate the event using just one broadband station and the *P*-wave arrival time pick from the other, or we can refine the crustal model. We find good agreement when testing these results against those from the entire array.

Crustal complexities can change the separation between key phases (*Pnl*, *Snl* and surface waves) which may require some calibration for maximum sensitivity. In many situations, distortions in wave-shapes may be severe enough to require 2D and 3D Green's functions. To accomplish this objective, we are developing an adaptive method for recovering 2D structure using Waveform Tomography. Our approach uses individual generalized rays computed from a layered model. The model is divided into blocks with variable velocity perturbations such that ray responses are allowed to shift relative to each other to maximize synthetic waveform fits to data. An efficient simulated annealing algorithm is employed in this search. The technique is applied to a collection of 25 aftershocks (Landers EQ) as recorded at two stations, GSC and PFO, separated by about 200 km which bracket the event population along the Landers fault system. The events are assumed to have known mechanisms and epicenters, but both their depths and origin times are allowed to vary. The results indicate considerable variation, especially in the top layer (up to ± 13%) which mirrors surface geology. A similar experiment involving a profile of observations across Tibet indicates substantial deep crustal variation. But given such large changes in velocities requires regeneration of the ray responses and the process repeated. Presently, we are at this stage of development.

Key Words: Crustal calibration; Source estimation.

OBJECTIVES AND SIGNIFICANCE

Global seismic monitoring requires improved locations of events and reliable capabilities to discriminate nuclear explosions from earthquakes. In this section, we will review some of the fundamental problems facing this global monitoring and how our efforts will produce improvements in location and discrimination.

Several recent review meetings addressing the Comprehensive Nuclear Test Ban Treaty (CTBT) have started with the basic difficulty in location. After two years of Review Event Bulletin (REB), North reported on GSETT-3

that REB 90% confidence ellipses contain the national network event locations less than half of the time. "Thus an appropriate means of implementing path-dependent teleseismic travel times, and of transitioning from these to regional travel time curves, will need to be developed." The strategy discussed at the Oslo Workshop held in 1999 and adopted by the pIDC is to implement procedures for such calibrations. Essentially, this is a bottom-up approach where one attempts to expand coverage from global to regional. The key to this procedure is the quality of nationally located events and ground truth (GT) methodology. The ground truth phenomenology developed around the location problem is based on the first-order need for a good XY-location and developing compatibility with global estimates.

Since it has proven difficult to determine event depth and origin time independently, the methodology tends to side-step this problem by assuming surface sources. GT1 and GT2 man-made events are obviously excellent locations but generally do not apply in most regions with no previous testing experience. Thus, we need to be concerned with the REB and GT10 and GT25 events of less accurate locations. GT10 events come from local networks and are selected from existing bulletins (CEB, NEIC, and ISC) provided that at least three IMS stations report the event. They must be recorded by at least 5 stations within a distance of 3° and have coverage consisting less than a 180° azimuthal gap at a distance less than 5°. The CEB is a pIDC's Calibration Event Catalog which is growing rapidly, now at over 2200 events. Its event location accuracy is assumed to be somewhere between GT10 and GT25. This appears to be the same accuracy expected in the relocated ISC data set of about 8,000 events [Engdahl *et al*., 1998], referred as EHB.

In general, majority of local seismic networks in many seismically active regions operates exclusively with short-period vertical component instruments. Hypocentral locations are often determined by using regional and local *P*- and *S*-wave travel times picked from this type of network and focal mechanisms are determined using the first-motion polarities. Various National Data Centers (NDC) provide these locations as ground truth locations for their datasets. Most local arrays still operate short-period Wood-Anderson types of instruments (usually the vertical component only) at relatively high gain. The emphasis was on determining the locations of buried fault zone by recording a large number of small events. The recordings go off the scale for any sizeable events (M>5), and thus events recorded teleseismically are not well recorded in the faulting region except for strong-motion instruments. Only in recent events with modern equipment have we learned about the complexity of faulting. These recent studies have suggested that the location from a short-period array is usually not the location of the asperity producing the teleseismic signal. Fortunately, the mechanism of the initial break is typically the same as the larger sub-events and is reliably recorded by the direction in which the vertical record goes off-scale. Thus, the focal mechanisms are still quite useful, especially for calibration size events (3.5<M<5.5). However, a major problem is that focal mechanisms are dependent on velocity model and particularly the take-off angle and depth. Depths are usually estimated from local *S*- and *P*-wave arrival time picks. But recent studies [Savage and Helmberger, 1999] indicates that the phase routinely picked as given in the Caltech-USGS bulletin appear to be *sP*, a phase that leaves the source as the *S* wave and reflects off the free-surface and travels to the station as a *P* wave. This phase is relatively strong on the vertical component and has been studied extensively in recent years. It is particularly obvious for deep events as noted by Zhao and Helmberger [1993] in their study of Pamir-Hindu Kush events, and Saikia *et al*. [1996] in other regions. This misidentification of *S*, using short-period methodology, leads to hypocentral depths that are on average too shallow and appears to explain the disparity in results comparing the depths from the Southern California Seismic Network (SCSN) catalog with waveform inversion. In particular, Zhu and Helmberger [1996b] found an average focal depth of 5.6±4.4 km from SCSN vs. depth of 11.9±3.7 km from waveforms for a population of 355 events with magnitude >3.5. Since strike-slip events do not have nodes in the vertical plane (independent of take-off angle), they should not show a disparity in focal mechanisms between the first-motion results and the waveform results which appear to be the case for the Landers aftershocks [Jones and Helmberger, 1998]. However, Northridge aftershocks do show this difficulty for those events listed as shallow in SCSN [Song and Helmberger, 1997]. This issue is presently being addressed in Japan and is probably a global problem, and must be addressed at individual arrays usually on the scarce-side in most countries.

In the last few years, real-time systems are being developed to monitor seismic activity for rapid notification and rapid damage assessment. The TriNet system now under development in southern California is an example where the recent Hector Mine earthquake was reported on the web within minutes, including the mechanism and direction of rupture. Because of the growing global concern with loss of life and damage during catastrophic earthquakes, many countries from other tectonic regions are starting to address the issue of the earthquake hazards. Thus, we can expect a period of transition where a combination of short-period instruments with a few sparse broadband instruments will be the only resource to obtain accurate locations of seismic events, basically the situation facing monitoring in many regions.

We are contributing our efforts to this process by a series of user-friendly software package that allow a mixed

set of travel times and polarities from short-period instruments, normally included in local catalog, and waveform data to better constrain earthquake source parameters, including location, depth, origin time, and mechanism. These codes will be made available to anyone on our web site. They would be developed and tested on the present seismic networks in southern California and China (Cooperative Program starting this summer, Lupei Zhu). A demonstration of this procedure will be given at the CTBT Annual meeting (2000).

SOUTHERN CALIFORNIA – A TEST BED FOR TECHNIQUE DEVELOPMENT

Earthquake occurrence in southern California has been one of the highest in the world. Over 500 seismic stations now exist with about 150 broadband stations similar to those of the IMS. Thus, we have a natural laboratory to observe seismic wave complexity generated by different types of earthquakes in variety of tectonic settings. We will begin by displaying some interesting observations followed by our modest attempts at explaining some of the details (Figure 1).

One of the simplest profiles of data is displayed in the lower panel, on the left. The middle section displays synthetics for a simple 1D model, just the standard California model with an 8 km transition zone from crust-to-mantle, Ji and Helmberger (1999). The basic long-period *Pn*, *pPn*, and *sPn* emerge from the *PmP* (*Pg*) and depth phases. The former features are expanded in scale and replotted in velocity on the right. These depth phases are generally quite well behaved and have been used to model most of the events occurring in the western United States. Thus, the relative stability of these phases can be exploited to help constrain source properties. However, there are some complex paths such as displayed in the upper profile that requires much more effort to explain. Those arrivals with very weak *Pn*'s (SVD, DGR, PLM, GOR, and DAR) are sampling the midpoints between Death Valley and the Panamint Range. They also contain a large second arrival, about 1.5 secs back, which can be modeled as a reflected-refracted arrival from a low-velocity-zone structure, see diagram below map, Ji and Helmberger (1999). The upper-mantle structure farther to the west is dominated by the rapid development of a high-velocity lid, with a strong gradient between eastern and western California, Melbourne and Helmberger (2000). Methods for inverting observed waveforms containing such complexities for structures are under development with some progress reported on in Helmberger *et al*. (2000) as presented at the meeting.

REVIEW OF SOURCE ESTIMATION

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 [Zhu and Helmberger, 1996b]. Because of the frequent occurrence and relatively simple source functions of earthquakes in the magnitude range of 4 to 5, their waveforms are ideal for calibrating path effects and investigating regional structures.

An application of this method is displayed in Figure 2. The alignment displayed is produced by cross-correlation which is controlled by the strongest arrival, Rayleigh waves. Thus, fitting whole records automatically is determined by the ratio of Rayleigh waves to Love waves. However, there is considerable information in the extended *Pnl* in terms of *P*, *pP*, and *sP* as there is in the teleseismic waveform modeling. To use this information, we model the beginning of the *Pnl* portion separately as in Figure 2. Fortunately, they are generally compatible as indicated by the above numerical experiment. These results suggest that a great deal of information is embedded in waveforms if used properly and the structure is simple. The adaptive grid-search code is ideal for adding outside constraints as displayed in the upper right-hand corner of Figure 2. Using the *P* polarities of the TERRAscope stations and the GSC waveforms, we obtained excellent results. Generally, we cycle through depth and pick the best fit with lowest misfit error, h =11~km. This depth agrees well with the teleseismic depth phases, namely *sP* (Wald, 1992). The individual depth phases in the regional data, *pPmP*, *sPmP*, *sSmS* are quite obvious at higher frequency as pointed out by Helmberger *et al*., 1992. The depth sensitivity from regional displacement waveforms (>150 km) comes from two features. The shape of the *Pnl* is controlled by the separation of *PmP* from its depth phases, *pPmP* and *sPmP*. The other depth constraint comes from the amplitude of the *Pnl* to surface waves. Strong surface waves suggest a shallow depth. It is quite common that the result from a single station agrees quite well with the network solution [Zhao and Helmberger, 1993, 1994; Dreger and Helmberger, 1991]. It is particularly important in sparse networks to use this method, so that the *Pnl* is not overruled by the details of the Rayleigh waves.

REGIONAL EARTHQUAKE STUDIES IN SOUTHEASTERN ASIA AND RECOMMENDATION

In a study by Zhu *et al*. [1997], we established the ability of two permanent stations, NIL (Nilore) in Pakistan and AAK in Kyrgyzstan, to locate and determine source characteristics of events in the Pamir-Hindu Kush region. In the winter of 1992, a temporary PASSCAL array (PAKN) was deployed in the northern Pakistan, which is very close to NIL. To calibrate the paths of events to AAK and NIL, we selected twelve events, 6 deep and 6 shallow,

and determined a very simple two-layer crustal models which seems to fit most of the recording throughout the highlands formed by the continental collision zone including Tibet. A library of Green's functions is then computed as a function of depth and distance as in California. For crustal events, the amplitude of the surface waves are very large for the shallow events (h = 10 km) so that the amplitude ratio of the *Pnl* to surface waves proves quite diagnostic of depth as discussed earlier. As the source moves into the mantle, the high-frequency signals trapped by the crustal waveguide disappears. The waveshapes become increasingly simple as the focal depth increases, essentially reducing to just *P*-waves and *S*-waves. However, there is still important information about polarity and relative strengths.

The above Green's functions are used to find fits to observations by applying an adaptive grid-search as described earlier. An example of fit to a Tibetan event, assuming these same Green's functions, is shown in Figure 3b using data from the 1991-1992 PASSCAL experiment. The procedure was used to study 50 events occurring during this time period (Figure 3a). Details are given in Zhu [1998] showing that most events are shallower than reported by the ISC and CMT's (when reported). Many of these events are not seen teleseismically and some are badly located as displayed in Figure 4a. Some deep events also occur in southern Tibet as reported by Zhu and Helmberger (1996a].

In Figure 4, we introduce a new type of relocation process, where we use one station (LHSA, see Figure 3b) and one travel time pick at another station (TUNL) which is close to a permanent short-period station. We assume that the mechanism and depth are known (Event 222, Figure 3) and test for location resolvability. We search a relatively large area (250 km by 200 km) with a 10 km grid spacing. The waveform misfit error as a function of location is shown in gray scale. The error surface has alternating troughs (dark areas) and ridges (light areas) simply because the surface waves in the data are in and out of phase with synthetics as the epicentral distance changes. They essentially lie along circles centered at the broadband stations (LHSA in this case). There is no well-defined minimum to constrain the event location using the waveform alone. However, by using the *P* arrival time at another station (TUNL), we added additional constraint on the event location. Combining the waveform misfit and arrival time gives the "best" location, which is close to the network location. As discussed in Helmberger and Zhu [1997], this event probably qualifies as GT5 since it is very well recorded by the array (Figure 3).

In conclusion, we have reviewed some of our efforts in improving locations and depth estimation through the use of waveform modeling. We think that some of these methods will prove useful in improving monitoring in remote regions.

REFERENCES

Bent, A. L., and D. V. Helmberger, Seismic characteristics of earthquakes along the offshore extension of the Western Transverse Ranges, California, Bull. Seismol. Soc. Am., Vol. 81(2), 399-422, 1991.

Dreger, Douglas S. and D. V. Helmberger, Source parameters of the Sierra Madre earthquake from regional and local body waves, *Geophys. Res. Lett.*, **18**, No. 11, 2015-2018, 1991.

Engdahl, E. R., R. van der Hilst, and R. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.*, **88** (3), 722-743, 1998.

Helmberger, D. V., and G. R. Engen, Modeling regional love waves; Imperial Valley to Pasadena, *Bull. Seismol. Soc. Am.*, **79**(4), 1194-1209, 1989.

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

Helmberger, D. V., X. J. Song, and L. Zhu (2000), Crustal Complexity from Regional Waveform Tomography: Landers Aftershocks, *JGR*, in press.

Ji, Chen, and D. V. Helmberger, Modeling regional seismograms, *Eos Trans. AGU*, **80**, Fall Meeting Suppl., 1999.

Ji, Chen, D. V. Helmberger, and D. J. Wald (2000), Basin Structure Estimation by Forward and Inversion Methods, *BSSA*, in press.

Jones, Laura E. and D. V. Helmberger, Earthquake Source Parameters and Fault Kinematics in the Eastern California Shear Zone, *Bull. Seismol. Soc. Am.*, **88**, No. 6, 1337-1352, 1998.

Langston, C. A., The SsPmp phase in regional wave propagation, *Bull. Seismol. Soc. Am.*, **86**, 133-143, 1996.

Melbourne, T. and D. V. Helmberger (2000), Lithospheric Control of Plate Boundary Deformation, *Science* (submitted).

Saikia, C. K., B. B. Woods, L. Zhu, H. K. Thio, and D. V. Helmberger, Path calibration, source estimation and regional discrimination for the Middle-East: Application to the Hindu-Kush region, *Scientific Report No. 1*, Phillips Lab, PL-TR-96-2069, 1996.

Saikia, C. K., H. K. Thio, B. B. Woods, X. Song, L. Zhu, and D. V. Helmberger, Path calibration, source estimation

and regional discrimination for the Middle-East and western Mediterranean, *Scientific Report No. 2*, Phillips Lab, PL-TR-96-2307, 1996.

Saikia, C. K. and D. V. Helmberger, A frequency wave number algorithm to compute up- and down-going wave fields from a buried seismic source, *BSSA*, **87,** 987-998, 1997.

Savage, B., and D. V. Helmberger, Fault dimensions from broadband networks (Abstract), *Eos Trans. AGU*, **80**, Fall Meeting Suppl., 1999.

Song, Xi J., L. E. Jones, and D. V. Helmberger, Source characteristics of the 17 January 1994 Northridge, California Earthquake from Regional Broadband Modeling, *Bull. Seismol. Soc. Am.* **85,** no. 6, 1591-1603, 1995.

Song, Xi J. and D. V. Helmberger, Broadband modeling of regional seismograms: the Basin and Range crustal structure, *Geophys. J. Int.* **125,** 15-29, 1996.

Song, X. and D. V. Helmberger, Source estimation of finite faults from broadband regional networks, *Bull. Seismol. Soc. Am.*, **86,** 797-804, 1996.

Song, Xi J. and D. V. Helmberger, Waveform Tomography and Pseudo Green's functions, *BSSA*, **88**, 304-312, 1997.

Song, X., and D. V. Helmberger, Northridge aftershocks, a source study with TERRAscope data, *Bull. Seismol. Soc. Am*., **87**(4), 1024-1034, 1997.

Wald, D. J., Strong motion and broadband teleseismic analysis of the 1991 Sierra Madre, California, Earthquake, *J. Geophys. Res*., **97**, 11,033-11,046, 1992.

Wallace, T. C. and D. V. Helmberger, Determining source parameters of moderate-sized earthquakes from regional waveforms, *Phys. Earth Planet. Interiors*, **30**, 185-196, 1982.

Wen, Lianxing and D. V. Helmberger, Propagational corrections for basin structure: Landers earthquake, *Bull. Seismol. Soc. Am.*, **87,** No. 4, 782-787, 1997.

Zhao, L. S., and D. V. Helmberger, Source retrieval from broadband regional seismograms: Hindu Kush region, *Phys. Earth Planet. Inter*., **78**, 69-95, 1993.

Zhao, L. S., and D. V. Helmberger, Source estimation from broadband regional seismograms, *Bull. Seismol. Soc. Am*., **84**, 91-104, 1994.

Zhao, L. S., and D. V. Helmberger, Regional moments, energy levels, and a new discriminant, *PAGEOPH*, **146**, 281-304, 1996.

Zhu, Lupei and D. V. Helmberger, Advancement in Source Estimation Techniques Using Broadband Regional Seismograms, *Bull. Seismol. Soc. Am.*, **86,** No. 5, 1634-1641, 1996.

Zhu, L. and D. V. Helmberger, Intermediate depth earthquakes beneath the India-Tibet collision zone, *GRL*, **23,** 435-438, 1996.

Zhu, L., D. V. Helmberger, C. K. Saikia, and B. B. Woods, Regional waveform calibration in the Pamir-Hindu Kush region, *J. Geophys. Res*., **102**, B10, 22,799-22,813, 1997.

Zhu, L., and D. V. Helmberger, Significant Moho relief across the northern margin of the Tibetan Plateau, *Science*, **281**, 1170-1172, 1998.

Zhu, L. (1998), Broadband waveform modeling and its application to the lithosphere structure of the Tibetan Plateau, Ph.D. Thesis, California Institute of Technology, Pasadena, California.

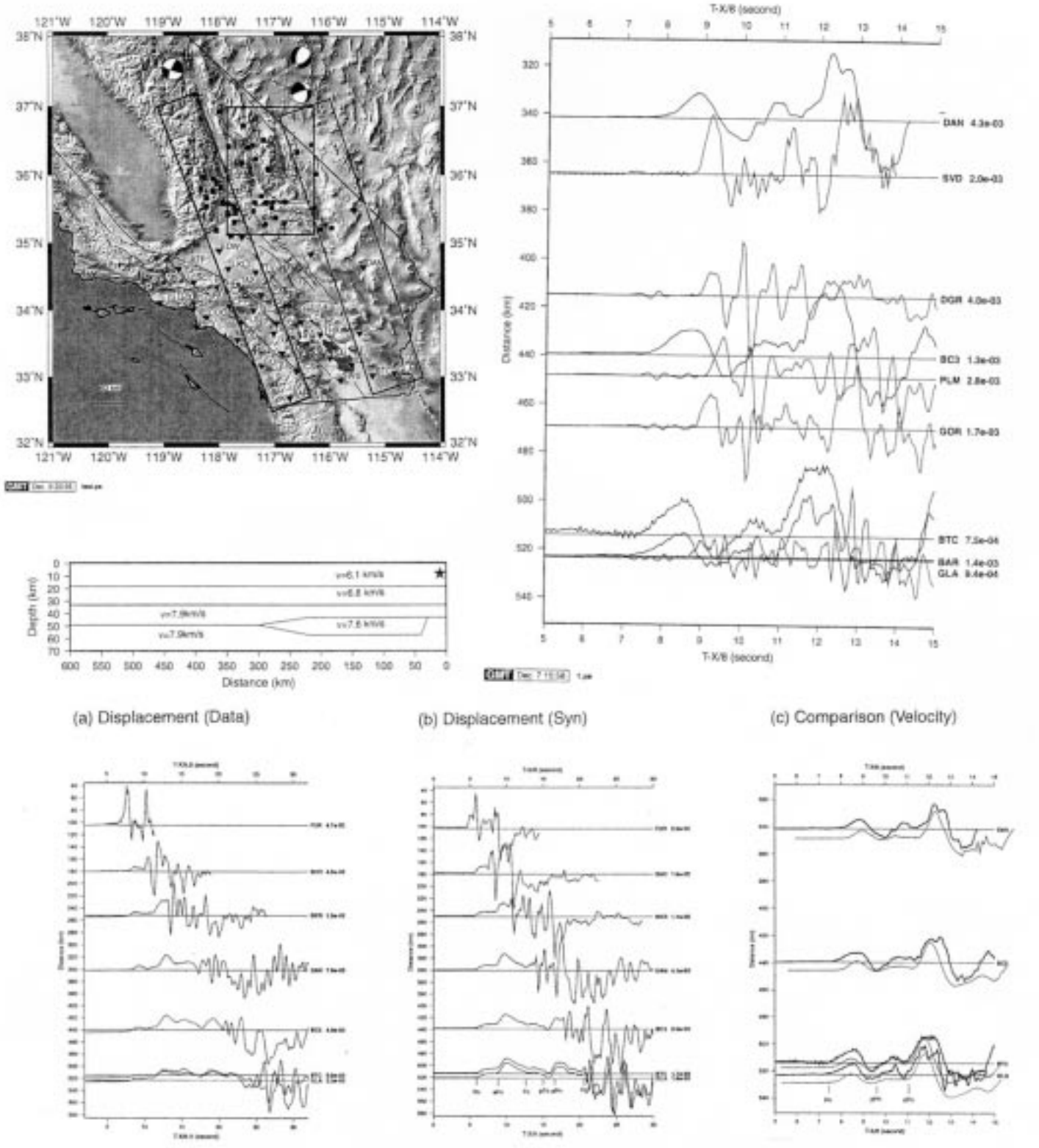

Figure 1. Upper panel (a) displays the locations of TriNet stations relative to events. The P_{nl} (extended P-waves) obtained in the easternmost box are relatively simple (black) while those sampling beneath the Panamint Range (between Death Valley and the Sierras) are late and complex (gray). The traces in gray require a complicated upper-mantle structure (see box) as discussed by Chen Ji and Helmberger (1999). Lower panel (b) displays modeling (1D) appropriate for the structure beneath the eastern block.

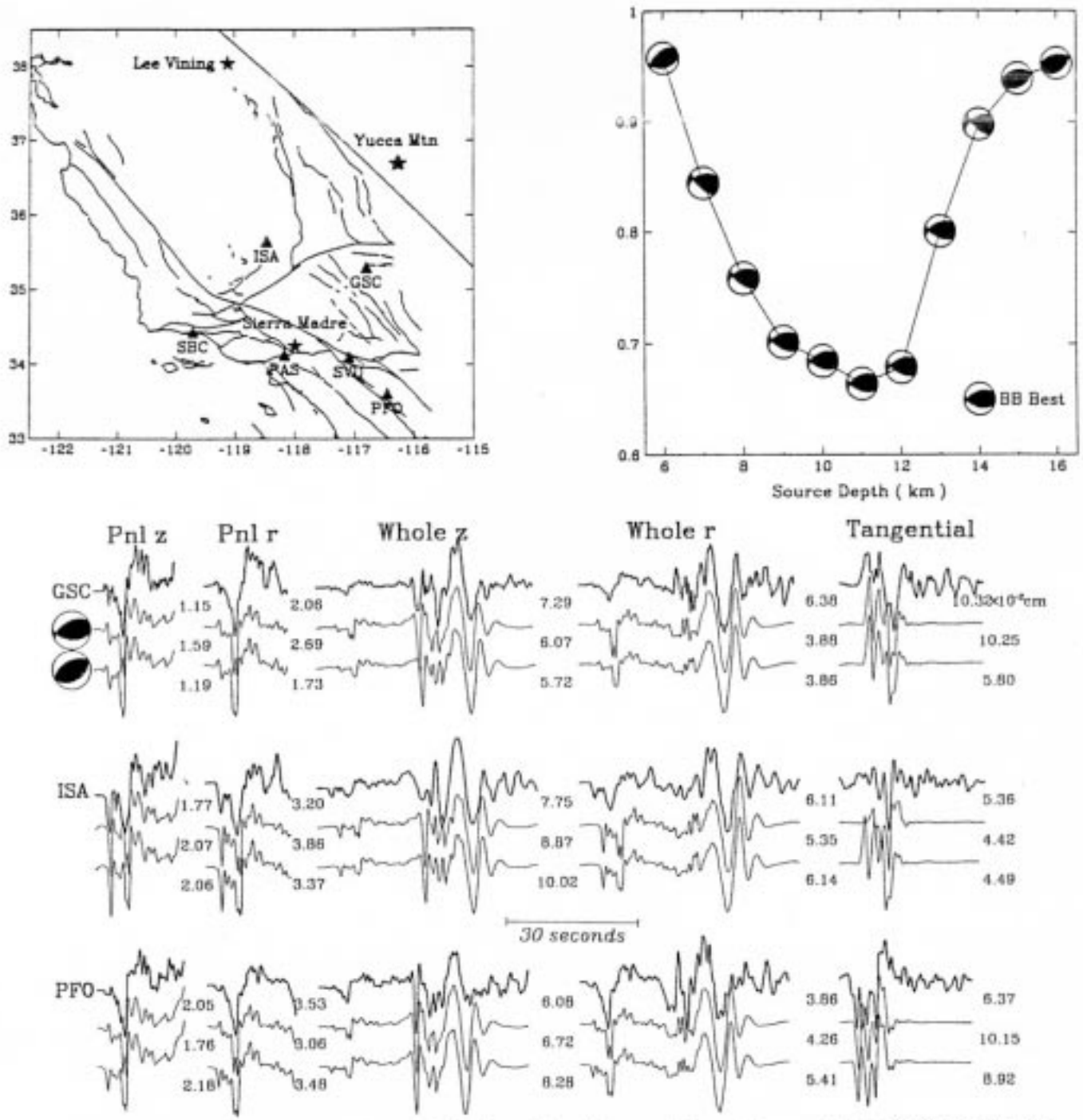

Figure 2. Comparison of Sierra Madre broadband observations at three stations (GSC, ISA, and PFO) with synthetics generated by estimating the fault parameters. Three windows of observed waveforms are used, namely, the beginning portion P_{nl} (first two columns), the Rayleigh waves or whole recording (middle two columns), and the tangential (column on the right). The corresponding Green's functions are shifted in time for best alignment and used to estimate strike, dip, and rake. The amplitude determine the seismic moment (numbers indicate peak amplitudes in cm). The bottom trace of each set corresponds to the source estimation using only the P_{nl} columns and the middle traces correspond to using all 5 columns. Both solutions yield the same moment of 3. X 10^{24} ergs. Single station solutions and various combinations of the data sets using P polarities are discussed in Zhao and Helmberger (1993).

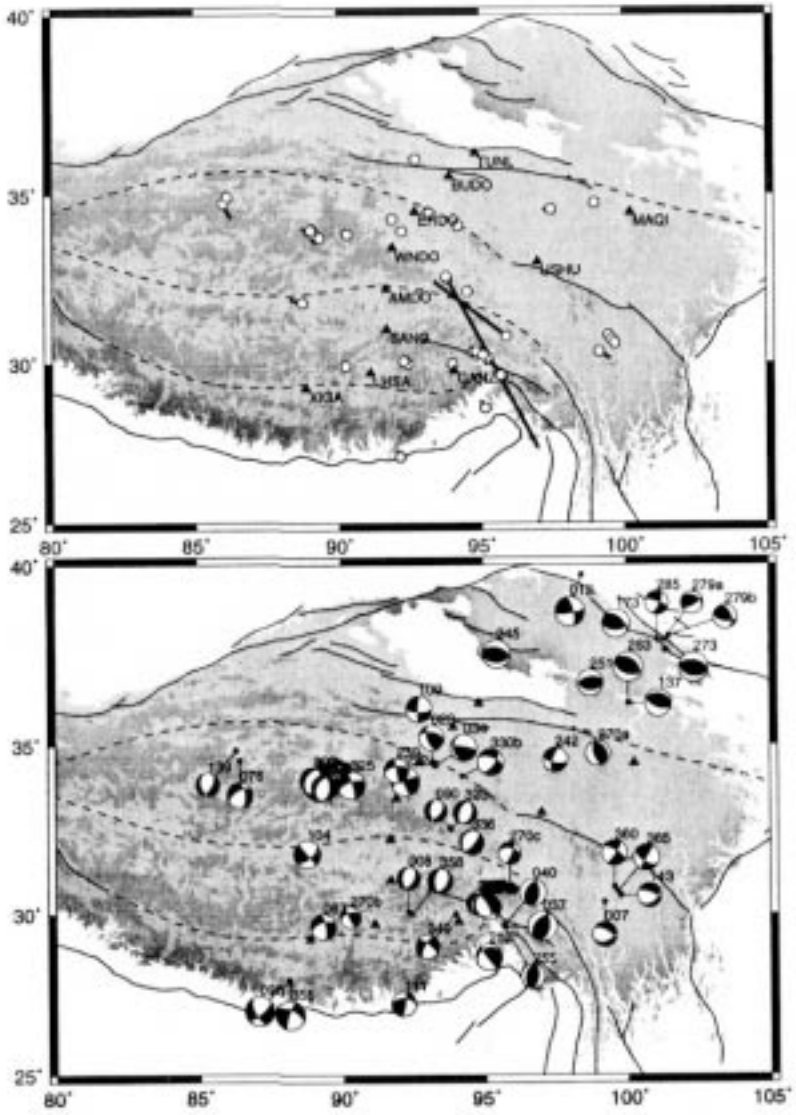

Figure 3a. Upper map displays the event mislocations relative to ISC estimates. Some of these (Mw<4) moved several degrees. Source mechanisms of earthquakes in Tibet and surrounding areas determined from inverting their broadband regional waveforms. All earthquakes are found to be located in the upper crust, between 5 and 30 km, except three events (Events 355, 095, and 067) in southern Tibet that lie at depths between 70 and 80 km.

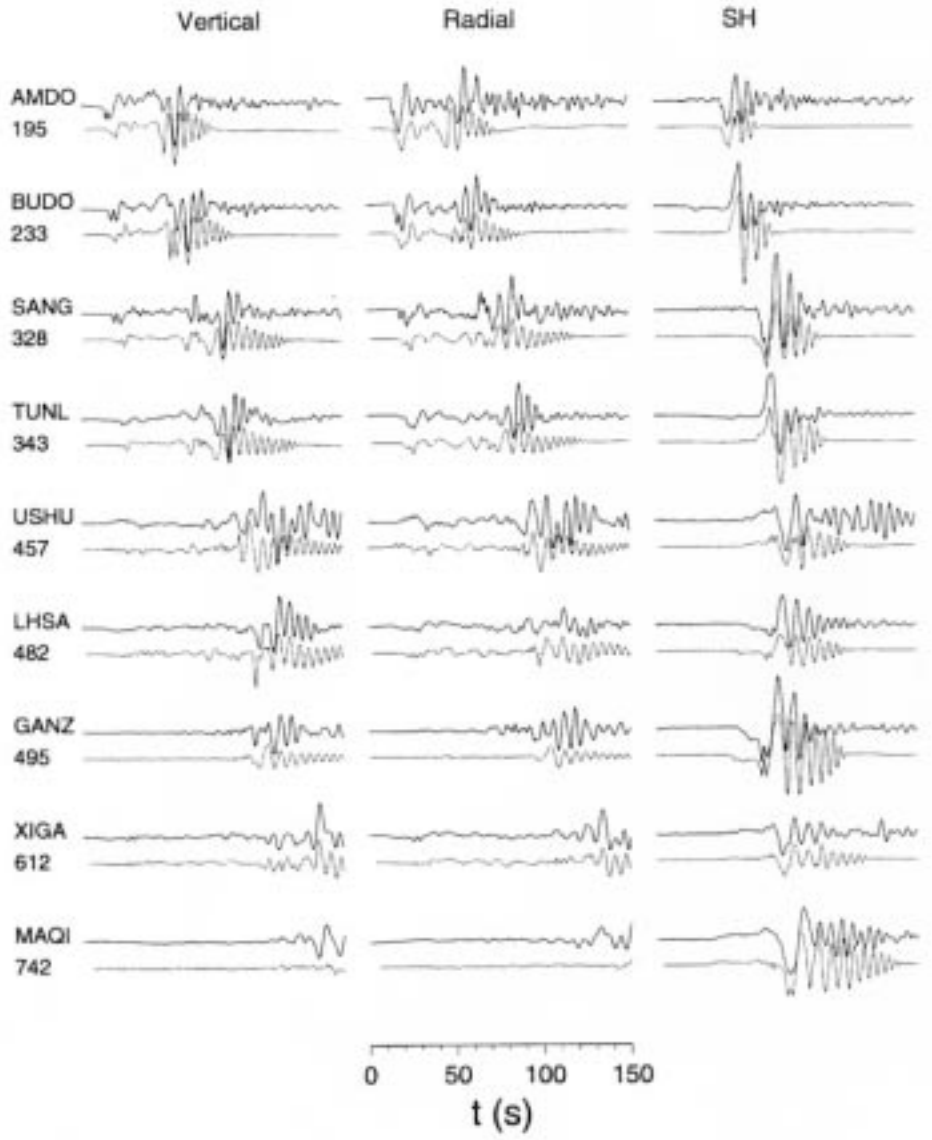

Figure 3b. Displacement waveform fit of a crustal event (Event 222), see map on the left produced by the waveform inversion. Each trace is multiplied by the square root of its epicentral distance (indicated by the number below the station name) to account for the surface wave amplitude decay with distance. This event had a Mw = 4.7 and a depth of 10 km.

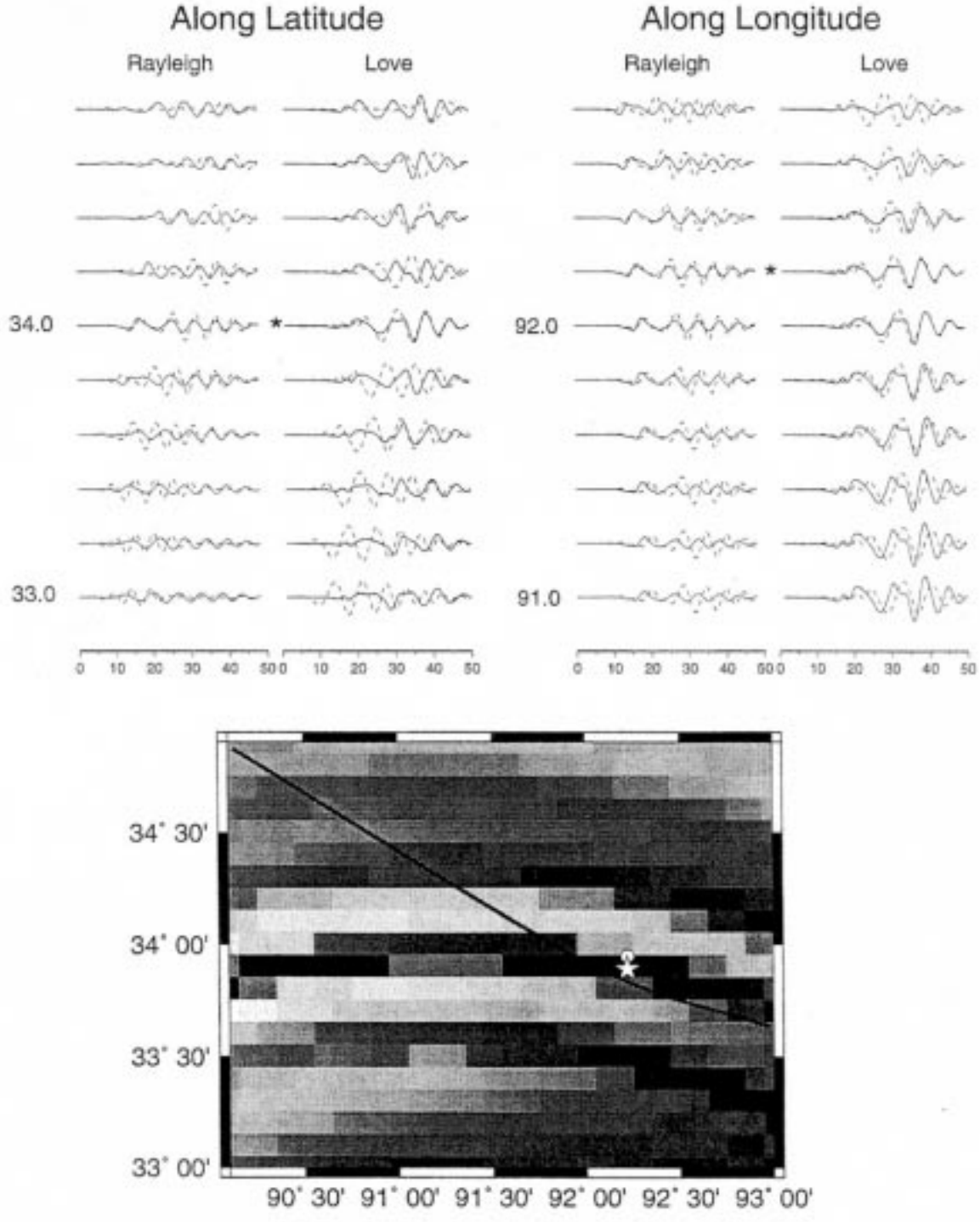

Figure 4. The bottom panel shows the waveform misfit error as a function of event location in gray scale (dark represents the minimum and white represents the maximum). The broadband station (LHSA) which provides the waveform in velocity is located in the south (see Figure 3a). Solid line represents possible event location to satisfy arrival time constraints at LHSA and another station to the north (TUNL).