MODELING REGIONAL SEISMOGRAMS

Don Helmberger, Sidao Ni, and Chen Ji
Geological and Planetary Sciences Division, Seismological Laboratory
California Institute of Technology

Sponsored by U. S. Department of Defense
Grant Number DSWA01-98-1-0010

ABSTRACT

Retrieving source characteristics of moderate-sized earthquakes in sparsely instrumented regions has been made possible in recent years, through the modeling of waveforms at regional distances. The techniques used in such studies model waveforms with some success at long period, using Green's functions for simple ID crustal models. For small earthquakes (M < 4), however, long-period signals are usually noisy and modeling shorter- periods requires refined Green's functions such as used in the empirical Green's functions approach. We are attempting to solve this problem by developing three approaches with varying degrees of empirical versus analytical methodologies. Approach (1), Pseudo Green's functions (Song and Helmberger, *BSSA*, **88**, 1998) uses observations from master events in conjunction with ID Green's functions as calibrations to correct for phase derivations. Approach (2) is more analytical which we call waveform tomography that generates ray responses derived from existing tomographic models. First, the ray paths from the ID layered reference model are used to localize each ray segment, where the anomalous velocities are applied by overlay, as in tomography. Next, new p_i (t_i) (p_i ray parameter, t_i travel time) are computed to satisfy Snell's law along with their numerical derivative ($\delta p/\delta t$), which can be used to construct a synthetic seismogram similar to the WKBJ method, or can be used to construct more exact Cagniard-style 2D synthetics. These individual ray responses are allowed to shift in relative timing and summed to better fit a set of observed waveforms aided by a simulated annealing code to optimize. Approach (3), inputs a profile of waveform data and produces a 2D model by direct inversion using a combination of analytical and numerical propagators (Chen *Ji et al., BSSA*, 1999, in press). This approach addresses complex structures, basins, etc., and requires data coverage. We are testing these methods against examples chosen from the TriNet (Southern California) dataset, 400 events recorded broadband with over 80 stations for some events. Preliminary applications reveal strong lateral variation near the crustal surface (± 20% in upper 3 km) with a particularly strong dependence on geology above the sources and approaching the individual stations.

Key Words: Waveform tomography; Source Inversions; 2D Synthetics; Waveform Modeling.

OBJECTIVE

We are developing methods aimed at providing a systematic means of calibrating paths for small events. The basic idea is to use large events (Masters) where locations and source parameters are reasonably well known to calibrate those particular paths to stations that observe smaller events as well, usually regional. The waveforms at a local station are then used to construct Pseudo-Green's functions which can be used to characterize still smaller events.

In this report, we will discuss efforts conducted on two tasks:

(1) Further development of Pseudo-Green's functions for use with sparce datasets (2 stations).
(2) Automate the comparison of observations with synthetic predictions from Tomographic models (Local and Global).

RESEARCH ACCOMPLISHED

In a recent paper on locating events in Pakistan, Zhu *et al.* (1997), we discuss the problems of locating events between two modern broadband stations (NIL and AAK) by using waveform modeling. The procedure worked well for deep events but became more difficult for shallow events. Here we investigate some of the propagational complexities involved in regional modeling where we have a large set of observations (TERRAscope) so that distortions can be easily assessed.

A relatively large number of regional events (500) have been recorded broadband with the implementation of modern equipment, especially in southern California (80 Stations). This data contains an immense amount of information about source processes (frequency-dependence) and crustal structure, but to unscramble these features proves very difficult. Most efforts to date have concentrated on the recovery of source parameters; namely, the Regional Centroid Moment Tensor (RCMT) solution using long-period surface wave (Ritsema and Lay, 1993). For events larger than M_w>5, it is possible to invert surface wave records for periods greater than 50 seconds assuming the PREM model (Dziewonski and Anderson, 1981) for all the western United States, essentially the standard CMT procedure. At shorter periods, the surface waves show regional variation and corresponding regionalized models are required (see for example, Patton and Zandt, 1991; Thio and Kanamori, 1992).

A second method, Dreger and Helmberger (1990), uses the relative strengths of the observed bodywaves to surface waves compared with synthetics to determine mechanisms, moment, and depth. Often, only one station is sufficient to fix the source parameters, since S and sS (SV and SH) are strongly dependent upon source orientation. By cycling through source depths, the proper timing between P and pP, S and sS, etc. allows accurate depth estimates. This approach works best at periods greater than about 5 seconds to stabilize the inversion. The biggest difficulty with this approach is that local fluctuations in arrival times, especially surface waves, requires highly localized structures. A useful modification (Zhao and Helmberger, 1994) uses a grid search and matches observed broadband seismograms against synthetics over discrete phases so that the beginning portion of records, which contains extended P-waves, can be shifted a few seconds relative to the later arriving surface waves. This approach, called the Cut and Paste method, desensitizes the crustal model used to generate the synthetics. A refinement, Zhu and Helmberger (1996), downweights nodal stations which further stabilizes the technique and allows automated source retrieval. We now determine the source characteristics within 5 minutes of an event occurrence.

An application of the code to Landers aftershocks was recently reported on by Jones and Helmberger (1998), see Fig. 1, where the size of the focal spheres indicates the relative magnitudes. A sample of the waveform fits (tangential component) for some of these events at PFO and GSC are displayed in Fig. 2. Note that only two events, events 24 and 31, are deep which is quite obvious based on the absence of significant surface waves. Other obvious source features are the lowest stress-drop event (19) and the highest (24), as is apparent from their relative long-period nature. Note that these source parameters were determined by the full TERRAscope array.

These synthetics were generated from the Jones model, and contain the various mechanisms and appropriate depths. Many of the fits are quite good with noticeable S_mS and sS_mS phases. The station PFO

has a particularly strong sS_mS relative to S (first arrival) for events 9 and 7, while sS_mS at station GSC is not so obvious. Thus, we need to address the issue of modeling such obvious complexity with some type of automation because of the large number of data becoming available. Developing such a technique is the main objective of this report where we will assume the sources are known and search for a 2D model compatible with observations at GSC and PFO.

Method (Ray-Shifting)

Our waveform modeling approach has been discussed in detail earlier in Song and Helmberger (1998). From the generalized ray theory (Helmberger and Engen, 1980), a synthetic seismogram consists of a series of ray responses that describe energy packets arriving at the receiver along various paths as displayed in Fig. 3. If we allow each ray response to shift in time and to vary in amplitude, synthetic fits to data can be dramatically improved. The more freedom we allow in the process, the better the synthetic fits to data. Our practice is to parameterize the problem so that the travel time of each ray response is associated with a block velocity model. Previous studies have shown that mild changes in model geometry lead to similarly shaped single ray responses (Song and Helmberger, 1996). The travel time of an individual ray is controlled by the integral slowness along its path, which allows relatively coarse model parameterization. The amplitudes of the ray responses, however, are more sensitive to the velocity perturbation and usually depend on very local changes in the velocity model. Changing the amplitude of individual rays, compared to changing their travel time, proves relatively harder to achieve with only a few blocks in the velocity model. To tackle this complexity, we adopt a two-step modeling approach. The first step is to fix the amplitude of each ray response based on the ID model and focus on the timing effect of the model parameterization. Then as we find velocity models that better explain the waveform data based on improvement to the timing of individual rays, we study the amplitude effect with ray synthetics based on the improved, now 2D, velocity models.

Our modeling involves multiple parameters. Event depth, origin time, and model velocity along an individual ray path all contribute to the timing of the ray response. With each set of parameters, individual ray responses are shifted differently in time, with their interference constructing a new synthetic seismogram. The problem is defined in terms of obtaining an optimal set of parameters that minimizes the least square error between data and synthetics. The search is conducted with a simulated annealing algorithm, as discussed in several recent seismological studies (e.g., Sen and Stoffa, 1991; Zhao and Frohlich, 1996). The reconstruction involves shifting back all the surface wave forming rays, a fraction of a second for the first ray, and gradually increasing to over a second for the last ray. The three rays forming S_mS are shifted differentially with the middle trace moving back slightly. These small adjustments produce a relatively good match of synthetic to observed waveform, although the ratio of S to S_mS is too large in the synthetic relative to the observed. The latter feature is difficult to correct without attenuating S or introducing more complex structure along the path.

Since surface multiples can easily overwhelm S_mS and sS_mS, we damp the penalty of misfit with time, thus emphasizing the fit over the time interval particular to the down-going rays. To do this, we applied to both the data and the synthetics a damping factor f (t), defined as a function of time t. $f(t) = 1$ for $t <= t_0$, and $f(t) = e ** [(t-t_0) / (t_1-t_0)]$ for $t > t_0$ ($t_0<t_l$). Time constants t_0 and t_1 are marked with vertical bars in Fig. 3. With this damping factor, the beginning of the observations is emphasized in the inversion and it places more resolution on the model recovery of the deep crust.

As demonstrated in Fig. 3, a small shift in timing of S_mS relative to direct S can usually improve fits. The flexibility provided by the interference with the two neighboring reflections is also very useful in modeling the S_mS triplication. Adjusting the timing in the surface layer multiple can likewise improve fits, although we have downweighted their contributions to concentrate on the deeper structure.

Results of Numerical Experiments

Preliminary runs revealed a tendency for faster velocities to occur towards the west. Thus, we allowed a separate set of boxes for the section west of the San Andreas although such a boundary cannot be identified precisely. We also experimented with a set of boxes for the surface layer. Results assuming 3 surface layer

sections are given in Fig. 4. Since it is difficult to judge how significant any particular simulation may be, we started by comparing inversion results as a function of the number of events used in the modeling, in particular 7, 19, and 25. The events chosen are indicated in Fig. 1 and have been picked to be representative of the population. The results proved quite stable with very rapid changes in the surface layers ranging over ± 13%. The section enclosing the upper Coachella Valley appears particularly slow which might be expected. Actually the surface variation in this section is quite compatible with local geology as can be seen by comparing with Fig. 1 where the mountainous region appear relatively fast relative to the basins. Numerous other results with different assumptions involving more blocks and allowing the sources to move are discussed in Helmberger *et al.* (1999). (Preprints available at the meeting).

To obtain better modeling results, we need to regenerate these ray responses with the 2D perturbations to move our synthetics closer to data. This approach appears possible using a new analytical method designed for generating 2D synthetics discussed in the next section, Ni *et al.* (1999) and will be pursued in future efforts. To check solutions, we can use a numerical-interfacing code outlined in Fig. 5. Here we demonstrate the method by considering the profile of stations running across the extended Los Angeles Basin starting at the basin edge, XPO, and forming a rough 2D profile across the basin. Particle motions indicate that the seismic field is predominantly SH (tangential) as displayed in the record section (black lines, lower left). For the region outside the basin, we assume a 1D crustal model, upper right) and analytical techniques to propagate the sources to the basin edge where the motions are interfaced to finite-difference technique (Wen and Helmberger). The method absorbs the uncertainties of the source and 1D propagation to fit the first record exactly; this is essentially a calibration so that the synthetic fits XPO by definition. We start by assuming a simply layered model. The three numbers accompanying each trace indicate the goodness-of-fit (synthetic to observation), the amplitude ratio, and the timing shift. For example, at DOW the misfit is (.56), the synthetic is smaller than the data by (.32) and it is late by 2.16 sec. We want to drive these numbers to (0,1,0). To improve the fit, we parameterize the layers with linear dipping segments between the inflection points which are allowed to vary. We generate numerical derivatives and apply a conjugate gradient inversion to refine the structure resulting with a new model. See Chen Ji *et al.* (1999) for details on the new method.

Preliminary Results on Global Models

New seismologic investigations of the Earth are capitalizing on the exponential growth of high quality waveform data. For example, Hendrik van Heijst, Jeroen Ritsema, and John Woodhouse (Spring AGU meeting) have teamed to produce a high order Global model. It contains data from all the networks, over 100,000 higher mode paths along with a huge collection of bodywave travel times, including complex paths (SP, S_cP, and multiples of P and S). This model is just now reaching completion and will be used to predict broadband synthetics in the near future. However, a considerable gap exists between global, large-scale seismological studies and higher resolution regional waveform analysis. For example, most global tomographic inversions do not incorporate triplication data caused by seismic discontinuities in the mantle, essentially attributing any associated travel-time anomalies to volumetric heterogeneity. Similarly, most waveform modeling uses globally averaged 1D seismic reference models focusing on isolated regions without consideration of the significant geographical variations in velocity. This disconnect between seismic inversion techniques often makes it difficult to isolate local structure from broader anomalies distributed along the ray paths. To rectify this, we have introduced physical models based on compositional states where a mapping is performed to infer temperature and the behavior of phase transitions, allowing depth-dependent behavior. Figures 6 and 7 display the application of this procedure to Grand's tomographic model. One such mantle feature in regional models is a seismic travel time triplication attributed to a sharp 2-3% seismic velocity discontinuity about 250 km above the core-mantle boundary, Lay and Helmberger (1983). The primary evidence for the triplication is an additional phase, S*cd*, arriving between the direct, *S*, and core-reflected, *ScS*, shear wave phases in approximately the 65°-83° distance range. The relative timing and amplitudes of the three phases experience significant regional variations. In some regions, the triplication is clearly displayed while becoming difficult to detect in others, suggesting an intermittent D" discontinuity, see Wysessions *et al.* (1998). Alternatively, the observed spatial intermittence of the triplication may be attributed to variations of the local seismic velocity gradients accompanying a substantially smaller (≈1%) velocity jump as displayed in Fig. 7. To determine the properties of the model, we compute the differential travel times between the various phases as in Fig. 6b and adjust the

parameterization (Clapeyron slope, percentage phase jump in velocity, etc.) to calibrate to datasets (Lay's differential times in this case). The triangles are stations and stars are events. Thousands of 2D synthetics generated for the various paths are compared against data in an automated fitting procedure, Sidorin *et al*. (1999). In most cases, this physical model predicts better fits to the seismic observations than the original 1D modeling exercises.

CONCLUSIONS AND RECOMMENDATIONS

The present plan of the PIDC is to improve locations by introducing source and stations corrections determined from ground-truth events (GTX). Although this appears to be working well, we will still be faced with particular problem events (New Year's event). In this case, being able to use the whole seismogram and interpret it in terms of travel paths and test what we would expect from the best 3D model for the particular region remains valuable. Another issue is the lack of waveform information use in definitions of ground-truth events for GTX's > 5. At least in Southern California, local modeling is forcing travel-time routines to be modified to yield source locations more compatible with waveform modeling, Luccio *et al.* (1999). In particular, the distribution of SCEC events with depth was (5.6 ± 4.4) while in waveform modeling, it was (11.9 ± 3.7 km). The latter estimates are proving to be the more accurate. Perhaps, a workshop addressing these issues would be helpful.

REFERENCES

Dreger, D., and D. V. Helmberger, Broadband Modeling of local earthquakes, *Bull Seismol Soc. Am*., **80** (5), 1162-1179, 1990.

Dziewonski, A. M., and D. L. Anderson, Preliminary Reference Earth Model, *Phys. Earth Planet. Int*, **25**, 297-356, 1981.

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, Crustal Complexity from Regional Waveform Tomography: Landers Aftershocks, submitted to *JGR,* July 1999.

Ji, Chen, D. V. Helmberger, and D. J. Wald, Basin Structure Estimation by Forward and Inversion Methods, submitted to *BSSA*, May 1999.

Jones, L. 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.

Lay, T., and D. V. Helmberger, A lower mantle S-wave triplication and the shear velocity structure of D", *Geophys. J R. astr. Soc.* **75**, 799, 1983.

Luccio, F., L. M. Jones, L. Zhu, E. Hauksson, and D. V. Helmberger, Resolution of focal depth determinations in Southern California, submitted to *BSSA,* 1999.

Ni, S., X. Ding, and D. V. Helmberger, Constructing Synthetics from Deep Earth Tomographic Models, submitted to *GJI,* May 1999.

Patton, H. J., and G. Zandt, Seismic Moment Tensors of Western U. S. Earthquakes and Implications for the Tectonic Stress Field, *J Geophys. Res.,* **96**, 18,245-18,259, 1991.

Ritsema, J., and T. Lay, Rapid Source Mechanism Determination of Large ($M_w \geq 5$) Earthquakes in the Western United States, *Geophys. Res. Lett.,* **20**, 1611-1614, 1993.

Sen, M. K., and P. L. Stoffa, Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics,* **56**, 1624-1638, 1991.

Sidorin, I., M. Gurnis, and D. V. Helmberger, Bridging the Gap Between Seismic Tomography and Waveform Modeling: Constraints on the Origin of Structures at the Base of the Mantle, submitted to *Science,* 1999.

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

Song, X. J., and D. V. Helmberger, Pseudo Green's Functions and Waveform Tomography, *Bull. Seismol. Soc. Am.*, **88**, No. 1, 304-312, 1998.

Thio, H. K., and H. Kanamori, Moment Tensor Inversions in Southern California Using Surface Waves recorded by TERRAscope (abstract), *EOS Trans. AGU*, 73,376, 1992.

Wysession, M.E. et al., in *The Core-Mantle Boundary Region*, M. Gurnis, M.E. Wysession, E. Knittle, B. Buffet, Eds., *Geodynamics Series*, Vol. 28 (American Geophysical Union, Washington, DC, 1998), pp.273-297.

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 C. Frohlich, Teleseismic body-waveforms and receiver structures beneath seismic stations, *Geophys. J Int*., **124**, 525-549, 1996.

Zhu, L., and D. V. Helmberger, Advancement in source estimation techniques using broadband regional seismograins, *Bull Seismol Soc. Am*., **86**, 1634-1641, 1996.

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

Zhu, L., and H. Kanamori, Moho Depth Variation in Southern California from Teleseismic Receiver Functions, submitted to *JGR*, March 1999.

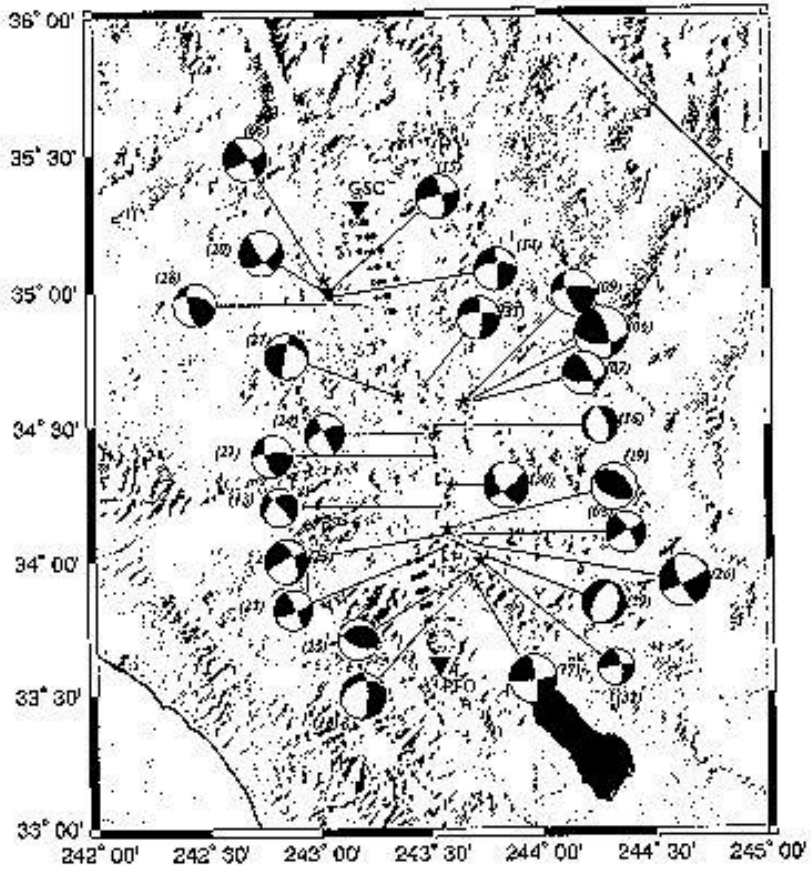

Figure 1. Topographic map of Southern California showing Landers aftershocks as black, white, and gray stars. The source mechanisms are after Jones and Helmberger (1998). The size of the focal spheres is proportional to the event magnitude. Two stations, GSC and PFO, are shown as dark triangles. The small dots indicate surface reflection points of ray paths for sS_mS phases (near events) and surface multiples of up-going S(near stations). The darkness of these dots indicates phase strength relative to a 1D-reference model.

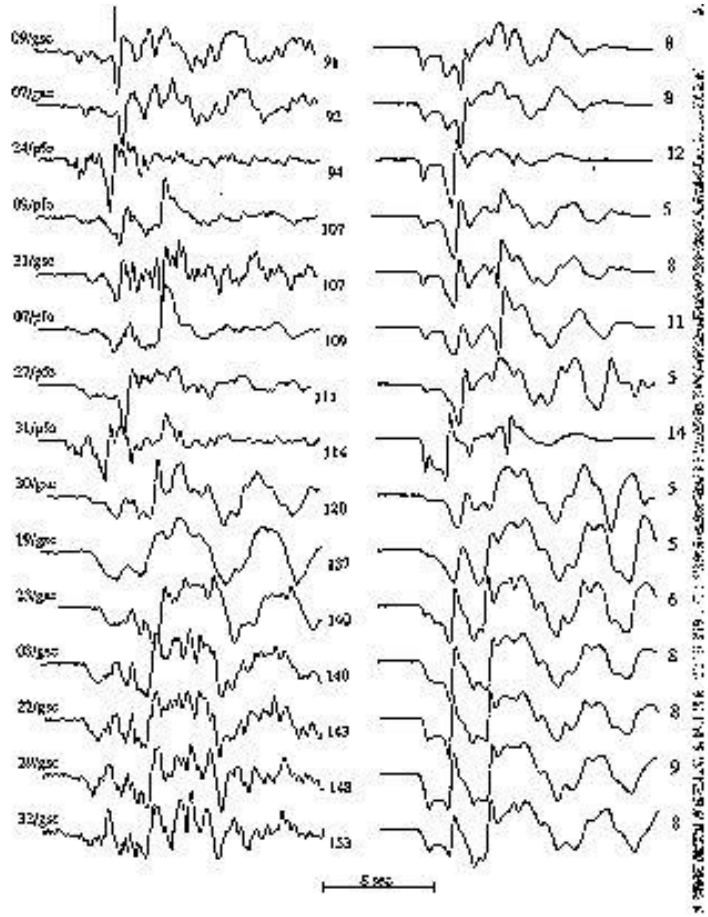

Figure 2. Comparison of observations (left) with synthetics (right) predicted from source estimation (Jones and Helmberger, 1999). Traces have been plotted on a reduced velocity section (t - Δ/3.6) along with small time shifts found by the source-shifting procedure. Note that deep events arrive early (event 31 at PFO) and shallow events late (event 9 at PFO). The range estimates and depths have been included in the middle and on the right.

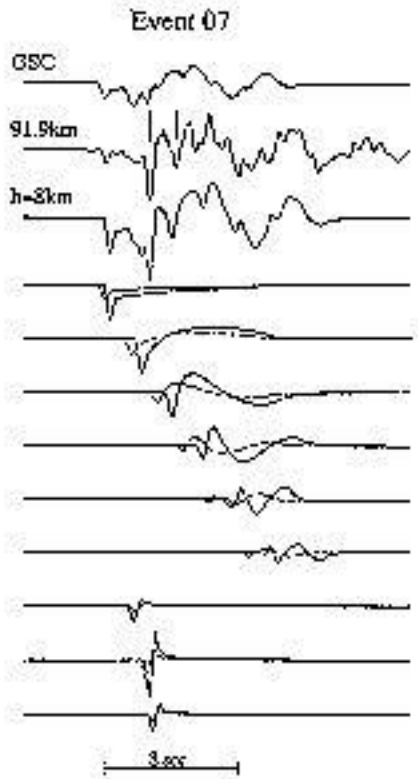

Figure 3. Pseudo-Green's function simulation procedure. The top trace displays the synthetics constructed from 1D Green's function with appropriate mechanism and moment. Middle trace displays data, and lower trace displays the Pseudo-Green's function. An amplification factor has been added to maximize the correlation with the data since the moment estimate comes from the network average and does not necessarily agree with individual observations. The lower set of traces show the shifting of the surface waves (direct plus surface multiples, 6 responses, and the bottom three include the SS rays). Note the small shifts back in time to match the data, dotted (original) and solid (shifted).

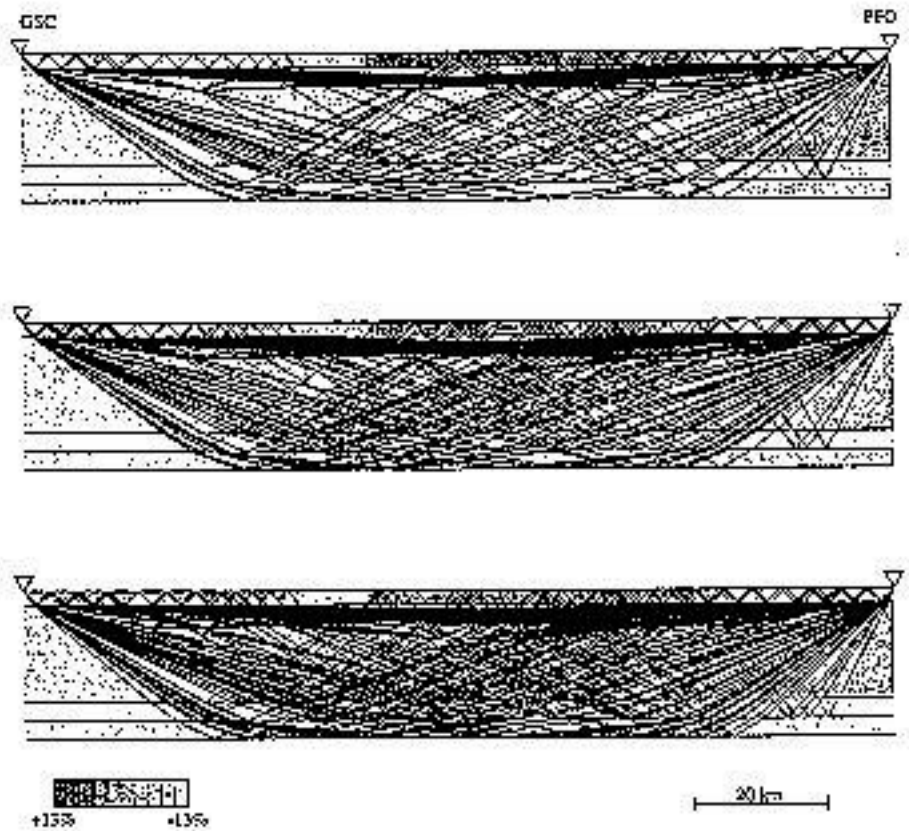

Figure 4. Comparison of results as a function of the number of the events used. Upper panel shows result of inversion using 7 events (white stars in Fig. 1); middle panel, 19 events (black and white stars); and lower panel, 25 events (all stars). The upper layer is divided into three blocks while the lower layers have only one division assumed to be at roughly the San Andreas fault boundary. In these inversions, source depth is allowed to move by up to 3 km and event origin time can vary by up to one second.

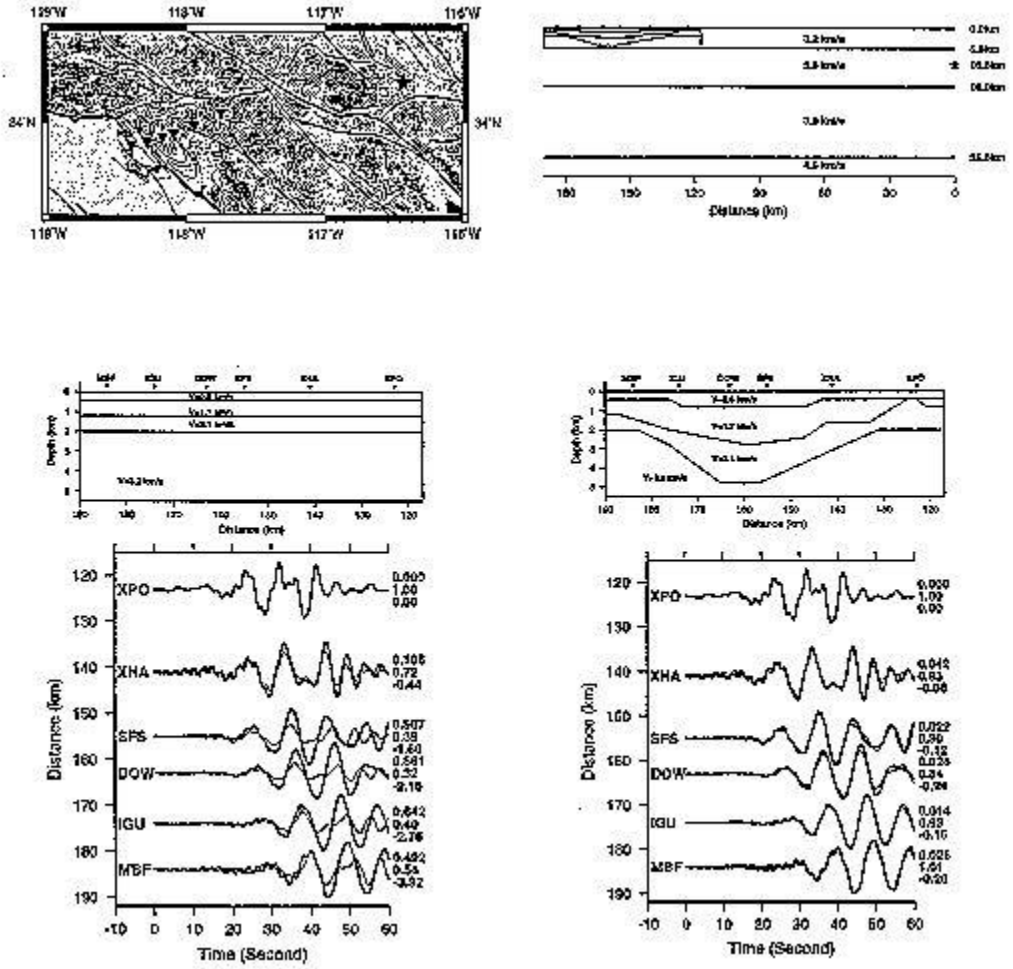

Figure 5. Layout of the waveform inversion procedure. Map of Southern California from Landers (star) to Los Angeles (contours of sediments). The full final model is shown in upper right; the analytical solution is interfaced at the shaded box edge. The initial model along with a comparison of synthetics with observation (black lines) is shown in the lower left. In the lower right, a comparison is made between observed and synthetic seismograms after adding basin structure

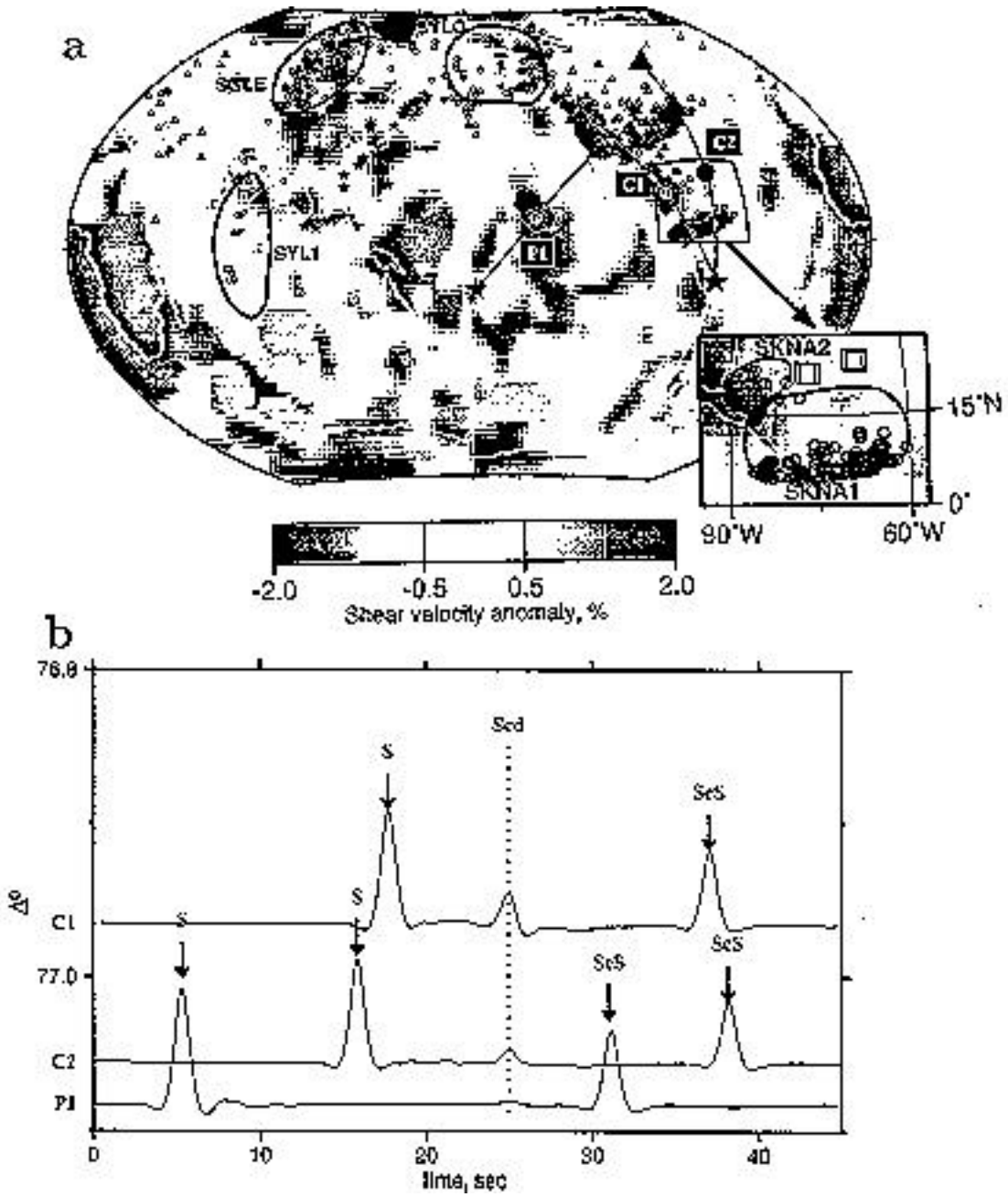

Figure 6a. Seismic observations of the D" triplication: events (stars) and stations (triangles) used in the study by Sidorin *et a1*. 1999). Circles show surface projections of *ScS* bounce points for the paths that are considered. The regions with strong observed D" triplication are shown by heavy contours and the names of the corresponding seismic 1D reference models are given next to the contours. The squares in the inset indicate *ScS* bounce points beneath Central America for ray paths that do not show any evidence for a D" triplication. The three highlighted paths are used in Fig. 7 to illustrate the sampled structure at the base of the mantle and its influence on the predicted seismic waveforms. The background color represents the shear velocity in the lowermost 240 km of the mantle, Grand (1994). **Figure 6b** indicates the relative strength of the D" triplication for central Pacific and Central America predicted by the preferred model. 2D synthetic waveforms (Green's functions from)WKM, Ni *et al.* (1999) computed for paths C1, C2, and P1 (cross-sections along these paths are shown in Fig. 7). The waveforms are normalized by the amplitude of direct S and aligned by the *Scd* peak.

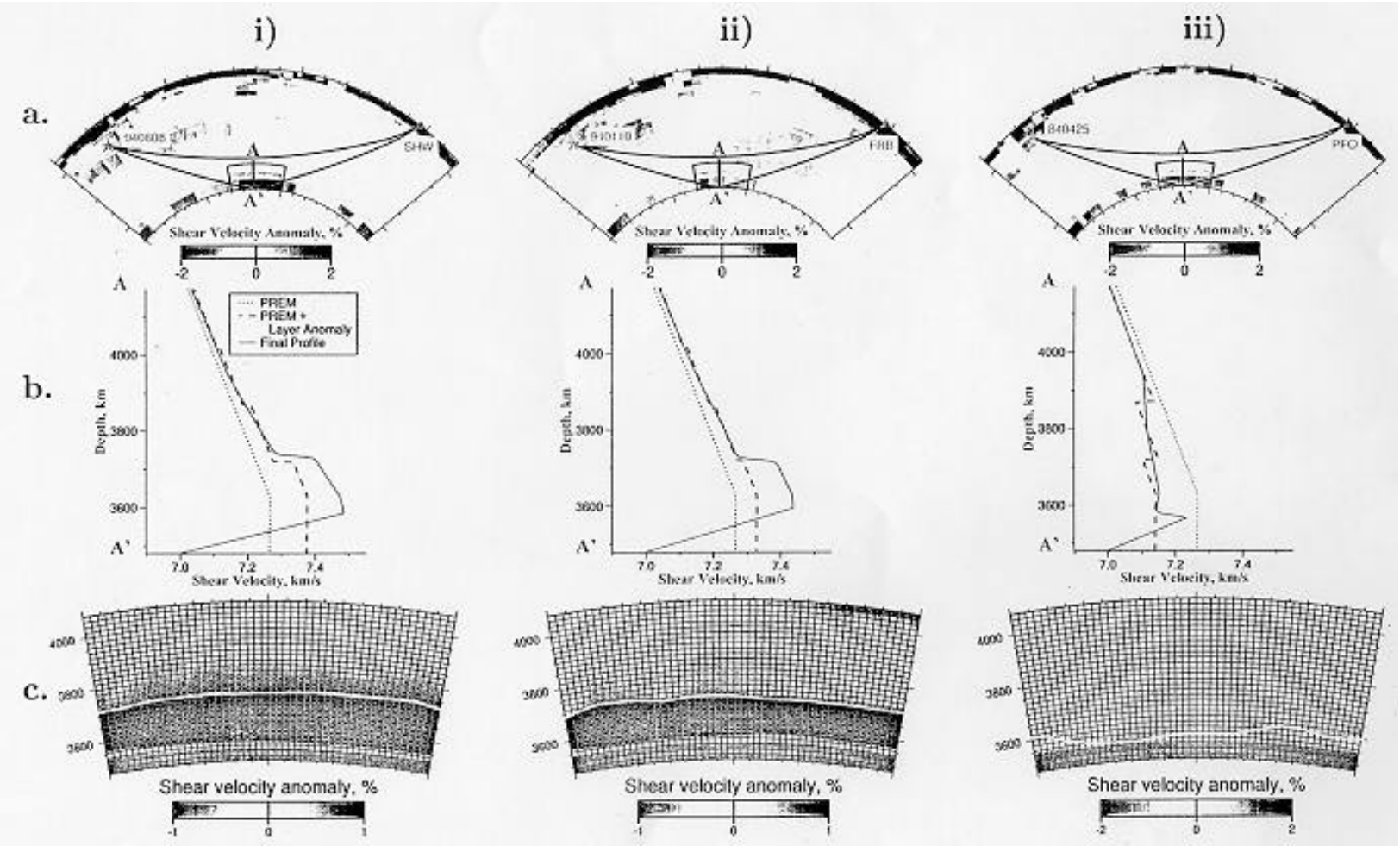

Figure 7. Shear velocity cross-sections along the ray paths highlighted in Fig. 6a. (i) Event 940808 to station SHW (path C1); (ii) event 940110 to station FRB (path C2); (iii) event 840425 to station PFO (path P1). (a) Cross-sections through Grand's tomography model. Direct (*S*) and core-reflected (*ScS*) rays are shown. (b) Peturbing a vertical column to incorporate a discontinuity using cross-section A-A' indicated in (a) as an example. The gray shade shows the layers of the tomography model. Dotted line shows PREM; dashed line shows the block anomalies (PREM values with added tomography velocity perturbations) ; the red solid line shows the final profile obtained by adding a discontinuity and a compensating negative gradient at the base. (c) Final composite model (perturbations with respect to PREM) for the region marked in (a). The grid lines of the fine mesh are shown (every other line is plotted horizontally, and one in every 10 is plotted vertically) and the phase boundary is indicated by the white line. The phase transition in (b) and (c) is characterized by h_{ph} = 200 km and γ_{ph} = 6 MPa/K