WAVEFORM MODELING OF THE CRUST AND UPPER MANTLE USING S, Sp, SsPmP AND SHEAR-COUPLED PL WAVES

Jay Pulliam, Mrinal K. Sen, Cliff Frohlich, and Steve Grand

University of Texas at Austin

Sponsored by the Defense Threat Reduction Agency

Contract No. DTRA01-00-C-0060

ABSTRACT

Phases that arrive near the direct SV phase, including Sp (converted at the base of the Moho), SsPmP, and shear-coupled PL (SPL) waves, collectively sample the Earth's crust and upper mantle at oblique angles and therefore have the potential to produce an accurate lateral average of structural properties than teleseismic P waves. SPL waves essentially mimic the propagation characteristics of regional PL phases, with the important difference that the number of events available for modeling is often greater for relatively aseismic regions, since sources are located at teleseismic distances. SPL waves are sensitive to crust and upper mantle structure, including seismic velocity gradients, Vp/Vs, impedance contrast across the Moho, and layer thicknesses. The first-arriving Sp and relatively large-amplitude SsPmP phases can often be identified and modeled simultaneously with SPL, which offers stronger overall constraints on the model.

The high frequencies and long time-series required for these phases, their deep penetration into the Earth and observation at teleseismic distances make the computation of synthetic seismograms time-consuming. We have parallelized and optimized a synthetic seismogram code based on the reflectivity method and are now able to compute complete seismograms up to 0.5 Hz in just over two minutes using eight DEC Alpha processors. The speed-up in computation time is nearly linear with the number of processors used, so a larger cluster and/or faster processors will further increase computation speed significantly.

In the past year we calibrated our reflectivity code against two widely-used codes and implemented a waveform-fitting global optimization algorithm. We are currently modeling observations of S, Sp, SsPmP, and SPL recorded for five deep earthquakes located at distances of 31°--59° from stations of the China Digital Seismographic Network (CDSN). The waveform-fitting process is controlled by a variant of Simulated Annealing (SA) that speeds up the annealing process by drawing each new model from a temperature dependent Cauchy-like distribution centered on the current model. This change with respect to SA has two fundamental effects. First, it allows for larger sampling of the model space at the early stages of the inversion and much narrower sampling in the model space as the inversion converges and the temperature decreases, while still allowing the search to escape from local minima. Second, each model parameter can have its own cooling schedule and model-space sampling scheme. In our case, each modeling run performs roughly a thousand iterations of forward calculations. Such a broad search of the model space, combined with analyses of sensitivity, resolution, and uncertainty, allows tradeoffs between model parameters to be evaluated, which helps build confidence in the final models.

OBJECTIVE

The conventional S phase is the initial, relatively sharp and pulse-like arrival that signals the beginning of a wavetrain with generally longer periods and normal dispersion. The particle motion associated with the S phase is rectilinear and all three components of motion are in phase. The dispersive wavetrain that follows S exhibits prograde elliptical particle motion that is confined to the vertical plane. Oliver (1961) named this wavetrain "shear-coupled PL" because it is analogous to the PL wavetrain, which appears between P and S arrivals at regional distances. Oliver (1961) presented a theory, based on the observed group and phase velocity of SPL that explained the phase as coupling between S and the fundamental leaking mode of Rayleigh waves in the crustal waveguide. According to this theory, shear energy generated by an earthquake (or explosion) travels through the Earth's mantle as a body wave, whereupon it impinges upon the Moho. Afterward a portion travels through the crustal waveguide as trapped P-waves and leaky SV-waves (Figure 2). The only difference between a PL phase, which is observed at regional distances from a source, and SPL phases, which are observed at teleseismic distances, is that SPL is generated by a shear wave impinging upon the Moho at regional distances from the observing station. In addition to producing SPL as it impinges on the Moho, a portion of the incident S wave converts to P as well, which then travels through the crust to arrive at the station as a precursor to S (Figure 1). This phase is called Sp and it has been used to model the crust by Jordan and Frazer (1975). Its sampling is much more localized to the station than is SPL's, making its sensitivity less representative of the broader region and more similar to that of P-coda receiver functions. SsPmP arrives at the base of the crust as a shear wave, travels upward through the crust as a shear wave, converts to a P wave at its surface reflection and bounces once off the Moho as a P wave. Langston (1996), while demonstrating that it can be highly useful for regional crustal modeling, showed that SsPmP can arrive before or after direct S, with either larger or smaller amplitude, and can also distort the S arrival pulse. We will attempt to simultaneously model S, Sp, and SsPmP which essentially isolates differences to the P structure of the crust, for data collected for SPL modeling. Because receiver function methods typically deconvolve the vertical seismogram, which is most sensitive to compressional wave energy, from the radial seismogram, constraints on P velocity are essentially sacrificed in order to obtain clean records of shear phases. The data we propose to model will provide a valuable check of receiver functions, in that they constrain the bulk properties of the crust—average P velocity and crustal thickness

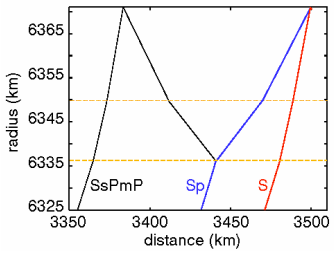
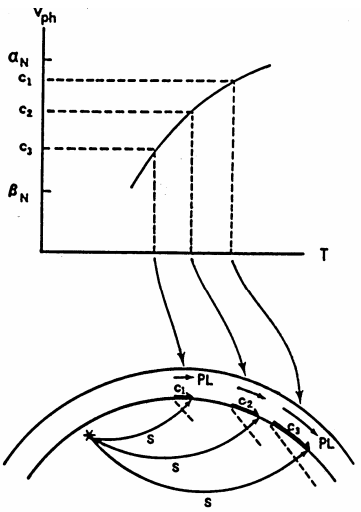


Figure 1. Typical raypaths for S, Sp, and SsPmP. These body wave phases provide constraints on crustal thickness, P and S velocities near the station.



Excitation of *SPL*-waves with distance and corresponding period–phase velocity curve. α_N, β_N: velocities of *P*- and *S*-waves at the Moho, *T*: period, V_{ph}: phase velocity.

Figure 2. Propagation characteristics of shear-coupled PL phases (from Baag and Langston, 1985). Note that the distance of propagation of SPL, and therefore its sampling, depends on characteristics of the velocity structure, including the slope of velocities below the Moho and attenuation. The earliest-arriving and largest-amplitude SPL waves are those that have converted from S nearest to the station, so our modeling will weight local sampling more highly than distant sampling but the wavefield still averages structure laterally.

(Sp and SsPmP)—and upper mantle (SPL). If the broad model search turns out to be too time-consuming or impractical, a fall-back strategy would be to first obtain a basic starting model by modeling Sp and SsPmP, then automating the SPL modeling using the Sp/SsPmP model as a starting point. Since reflectivity is a full-waveform method, there is no need to specify which phases should be included (nor need we identify or "pick" specific arrivals) in the waveform fitting procedure. While the waveform fitting itself will be automated, a great deal of exploration will be required to determine optimal window lengths, frequency content, and, perhaps, variable weighting functions for different portions of the seismograms.

We are evaluating the usefulness of S, Sp, SsPmP (Figure 1), and shear-coupled PL (SPL) phases (Figure 2) for modeling crustal and upper mantle structure using real and synthetic data, developing a waveform inversion technique based on a novel implementation of the reflectivity method and global optimization algorithms, and applying this method to data recorded in China. We have made substantial progress in speeding up the synthetic seismogram computations to the point where a global optimization is feasible. The reflectivity calculation involves computation of reflectivity matrices for a stack of layers as a function of ray parameter (or wavenumber) and frequency. The computation of reflectivity responses for different ray parameters and frequencies are completely independent of each other. We took advantage of this independence to develop a reflectivity code that runs on parallel computer architectures. Our code loops over ray parameters, i.e., to each node we assign a certain number of ray parameters to compute. At the end, the master node assembles the partial responses and performs the inverse transformation to generate synthetic seismograms at the required azimuths and distances. We used MPI for message passing and ran our code on a PC cluster consisting of 16 nodes; each node is a 660MHz alpha processor with 8MB cache and 1GB of RAM. A Myrinet interconnect is used to communicate between nodes.

Figure 3 shows synthetic seismograms computed using our parallelized reflectivity code for a distance of 50° and an earthquake at 600 km focal depth. Since the reflectivity algorithm is "embarrassingly parallel" in that the response for each frequency or ray parameter can be computed on independent processors, without communication between processors, on a parallel machine computation speed increases nearly linearly with the number of processors (Figure 4). In a side-by-side comparison for various distances, source depths, and model complexity, our code matched results of the Fuchs-Muller code very well.

Our modeling method retains the time and cost advantages of P-coda receiver function methods but which uses types of data that are more appropriate for CTBT purposes: Shear-coupled PL phases (SPL), Sp phases converted at the Moho, and SsPmP. SPL samples the crust and upper mantle in the vicinity of a station most broadly compared to Sp, SsPmP and P and it emulates the propagation of regional phases, which reflect at more oblique angles (or are refracted by these layers) than are the more steeply arriving body phases typically used in receiver function modeling. In short, because of the data they use, the models produced by receiver function methods may be inadequate for the purposes of CTBT monitoring. These

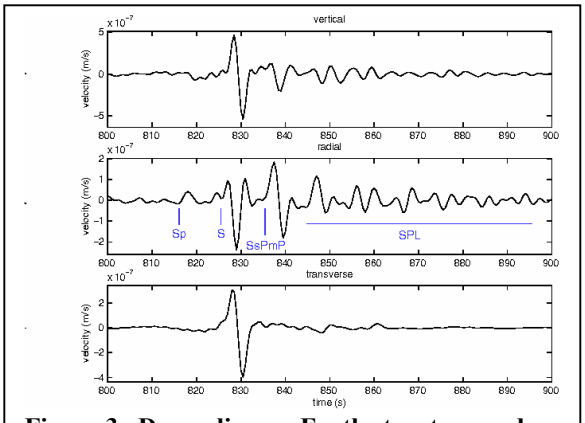


Figure 3. Depending on Earth structure and an earthquake's radiation pattern, the phases Sp, S, SsPmP, and SPL may appear prominently on the radial component seismogram at distances between 30° and 75°.

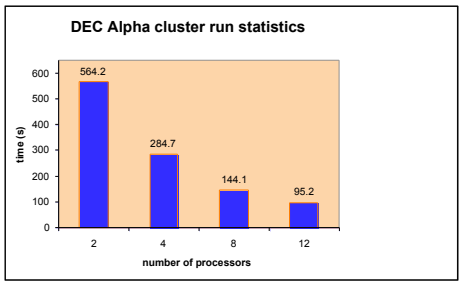


Figure 4. Statistics of an example run to compute complete seismograms for frequencies from 0 to 0.5 Hz (number of frequencies = 2000, number of layers=192, number of ray-parameters=1000) on the DEC Alpha cluster. Note the near linear speedup of the algorithm with the increase in the number of processors.

latter phases sample only a narrow cone beneath the station (e.g., Zhao and Frohlich, 1996). Modeled simultaneously (where they exist), SPL, Sp, and SsPmP offer the potential for producing azimuthally-dependent structural models.

We are pursuing this strategy because efforts to determine the locations of small, regional seismic events are hampered, in most parts of the world, by insufficient knowledge of the crust and upper mantle. Also, while focal depths are often a highly useful discriminant between explosions and earthquakes, their determination is quite sensitive to crustal structure. The most encouraging approaches to determining focal depths require precise modeling of seismic waveforms, particularly for small-magnitude events, in which travel time picks are relatively more prone to errors than for larger events. Yet, a precise modeling of waveforms at regional distances requires an accurate model of the crust and upper mantle along the propagation path between source and receiver.

RESEARCH ACCOMPLISHED

Forward Modeling

In order to explore the sensitivity of each phase we computed synthetics for a variety of distances using different crust and upper mantle models, including PREM (Dziewonski and Anderson, 1981; Figure 5), SNA (Grand and Helmberger, 1986; Figure 6), TNA (Grand and Helmberger, 1986; Figure 7), ECH (Zhao et al., 1991; Figure 8), and WCH (Beckers et al., 1999; Figure 9). Differences between these results point to sensitivity on the parts of these phases to distinct parts of the models' structure. Figure 10 shows an expanded comparison of waveforms produced using the different models for a single epicentral distance.

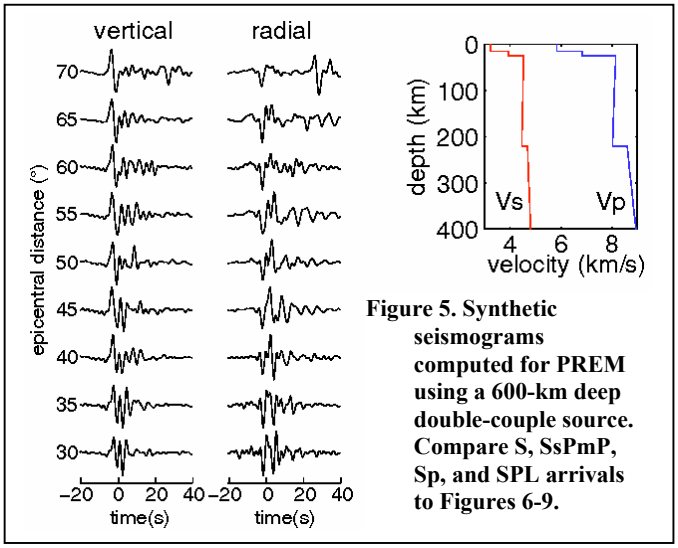


Figure 5. Synthetic seismograms computed for PREM using a 600-km deep double-couple source. Compare S, SsPmP, Sp, and SPL arrivals to Figures 6-9.

Figure 11 shows differences in waveform features to be expected from structural variations within China alone. The models at upper right were produced by Mangino et al. (1999) by modeling receiver functions for structure beneath stations of the China Digital Seismographic Network. The models differ primarily in the lower crust. One can see from the synthetics at left that these model differences are reflected by differences in the S coda, largely in SsPmP and SPL, for the distance range 30°-50°. These synthetics were produced for a 600-km deep event. This suggests that the relative timing of SsPmP-S and the phase and amplitudes of SPL phases are sensitive to the lower crust.

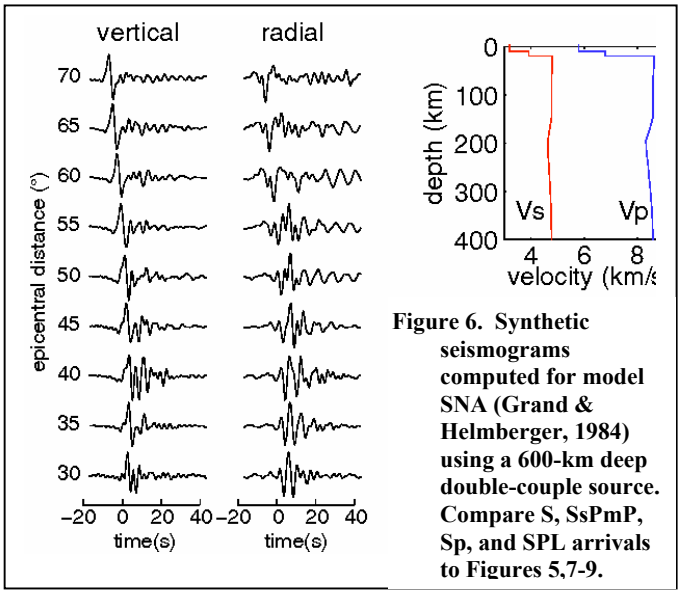


Figure 6. Synthetic seismograms computed for model SNA (Grand & Helmberger, 1984) using a 600-km deep double-couple source. Compare S, SsPmP, Sp, and SPL arrivals to Figures 5,7-9.

Inverse Modeling

Our modeling process is controlled by a global optimization algorithm called Very Fast Simulated Annealing (VFSA) (e.g., Sen and Stoffa, 1995). Simulated annealing (SA) is analogous

to the natural process of crystal annealing when a liquid gradually cools to a solid state. The SA technique starts with an initial model m_0, with associated error or energy $E(m_0)$. It draws a new model m_{new} from a flat distribution of models within the predefined limits. The associated energy $E(m_{new})$ is then computed and compared against $E(m_0)$. If the energy of the new state is less than the energy of the initial state, the new model is accepted and it replaces the initial model. However, if the energy of the new state is higher than the initial state, m_{new} is accepted with the probability of $\exp((E(\mathbf{m}_{new}) - E(\mathbf{m}_0))/T)$, where T is a control parameter called temperature. This rule of probabilistic acceptance (called the Metropolis rule) allows SA to escape local minima. The process of model generation and acceptance is repeated a large number of times with the annealing temperature gradually decreasing according to a predefined cooling schedule. A variant of SA, called Very Fast Simulated Annealing (VFSA) speeds up the annealing process by drawing new models from a temperature dependent Cauchy-like distribution centered on the current model. This change with respect to SA has two fundamental effects. First, it allows for larger sampling of the model space at the early stages of the inversion (when "temperature" is high), and much narrower sampling in the model space as the inversion converges and the temperature decreases, while still allowing the search to escape from local minima. Second, each model parameter can have its own cooling schedule and model-space sampling scheme. VFSA therefore allows for individual control of each parameter and the incorporation of a priori information. The model is parameterized in terms of layers, in which layer thickness is a free parameter, as are Vp, Vs, and density for each layer. In our modeling so far we have kept attenuation fixed, although it could be allowed to vary as well.

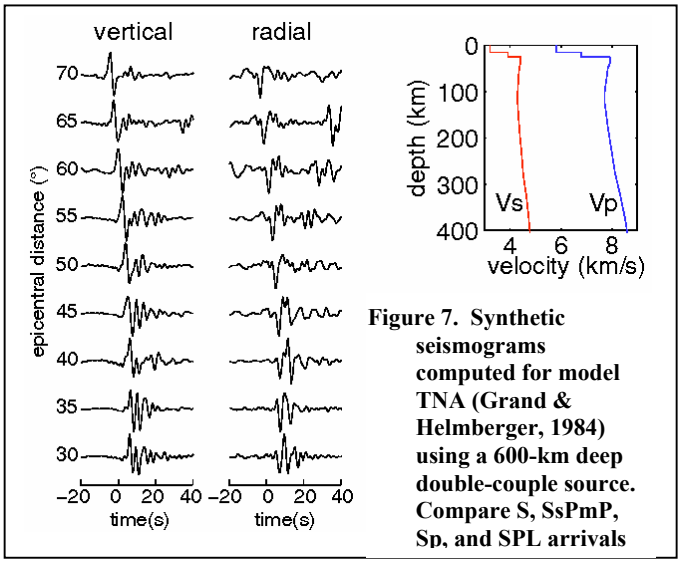


Figure 7. Synthetic seismograms computed for model TNA (Grand & Helmberger, 1984) using a 600-km deep double-couple source. Compare S, SsPmP, Sp, and SPL arrivals

Figure 12 shows the results of a synthetic inversion in which synthetic seismograms were computed for the input model shown in Figure 13 and then modeled. The result is a very good fit between output and input seismograms but the model that corresponds to the best-fitting waveforms shows some differences compared to the input model (Figure 13). However, the modeling produced accurate bulk properties for the crust, including average velocities and thickness and generally put layer boundaries close to their true locations. The differences between input and output models suggest that there is some tradeoff between parameters of different layers. These tradeoffs are the sort of information that our uncertainty estimation scheme is designed to evaluate. The goal is to produce quantitative estimates of the strengths of constraints and the uniqueness of the modeling.

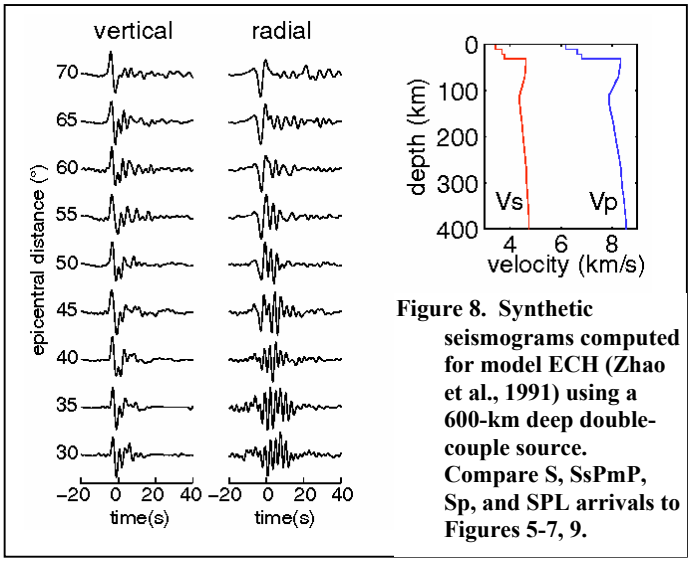


Figure 8. Synthetic seismograms computed for model ECH (Zhao et al., 1991) using a 600-km deep double-couple source. Compare S, SsPmP, Sp, and SPL arrivals to Figures 5-7, 9.

Uncertainty Estimation: PPD, Posterior Covariance and Correlation

In seismic waveform inversion, more than one model can often explain the observed data equally well and trade-offs between different model parameters are common. It is therefore important not only to find a single, best-fitting solution but also to find the uncertainty and level of uniqueness of that solution. A convenient way to address these issues is to cast the inverse problem in a Bayesian framework (e.g., Tarantola, 1987; Sen and Stoffa, 1995) in which the posterior probability density function (PPD) is the answer to the inverse problem. "Importance sampling" based on a Gibbs sampler or a Metropolis rule (Sen and Stoffa, 1998) can be used effectively to evaluate the necessary multi-dimensional integrals and to estimate PPD, posterior mean, covariance and correlation matrices. The posterior covariance and correlation matrices quantify the trade-off between different model parameters. Sen and Stoffa (1995) showed that multiple VFSA runs with different random starting models could be used to sample models from the most significant parts of the model space. This "poor man's" importance sampling, which is computationally highly efficient, results in estimates that are fairly close to the values obtained by theoretically-correct Gibbs sampling.

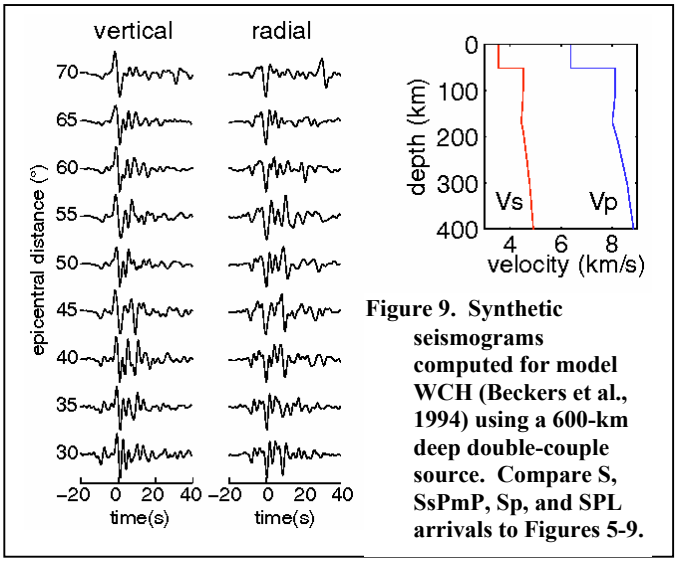


Figure 9. Synthetic seismograms computed for model WCH (Beckers et al., 1994) using a 600-km deep double-couple source. Compare S, SsPmP, Sp, and SPL arrivals to Figures 5-9.

Application to Data from the China Digital Seismographic Network

Figure 14 shows the result of modeling radial and vertical records from station BJT of the 9/28/1994 event in Indonesian (see Table 1). Subjective choices must be made in the modeling about which parts of the seismogram to try to fit. Regions with strong lateral variations or anisotropy, it may be that no single 1D model will predict all the phases adequately. Requiring the modeling to try to fit each phase similarly will result in a model that does not predict any single phase well. Given the sampling produced by the various phases, we tried to prioritize fits in the following order: SPL, SsPmP, SV, and Sp. Given the broad sampling and lateral averaging performed by SPL we believe that the single 1D model that fits SPL best will be the most useful for locating earthquakes regionally. SsPmP also has a broad sample and constrains Vp, so it's second in priority. This priority order reflects a starting point—an initial preference—and will be evaluated as a strategy.

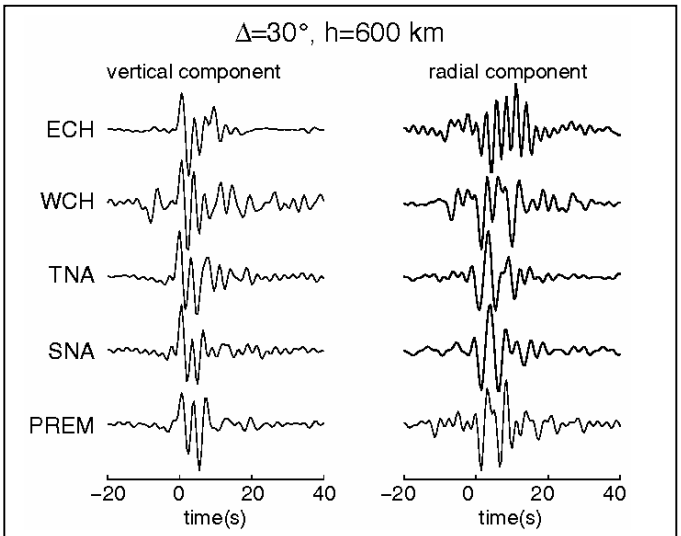


Figure 10. Comparison of synthetic seismograms computed using the models shown in Figures 5-9 for the same deep, double-couple source at an epicentral distance of 30°. Differences in the arrival times and amplitudes of S, Sp, SsPmP and SPL indicate those phases' sensitivity to structure near the receiver.

To implement this strategy we successively try to fit broader time windows of the data waveforms. Figure 14 shows the results of 600 iterations in which only the window between 815 s and 860 s is evaluated

for correlation at each iteration. The BJT receiver function model of Mangino et al. (1999) is shown for comparison, although it is not a "reference" or "starting" model in any sense. In our procedure we specify the search limits (also shown in Figure 14) and the first candidate model is chosen randomly. Figure 16 shows an attempt to fit a broader window that includes SV and Sp. Differences between the two final models shown in Figures 14 and 15 are again concentrated in the lower crust, between 30 and 50 km depth. Figure 16 shows the result of a 600-iteration run to model records from BJT of the 11/15/1994 Indonesia event (Table 1), in which the time interval fit is restricted to include SsPmP and SPL. Figure 17 shows the results of a 600-iteration run to fit a broader window, as in Figure 15. Figure 18 shows results for a third event (Table 1), also recorded at BJT. Figure 19 compares the models produced for the region containing BJT with Mangino et al.'s (1999) receiver function model, in the context of the imposed search limits. Figure 20 shows the potential applications of this technique worldwide.

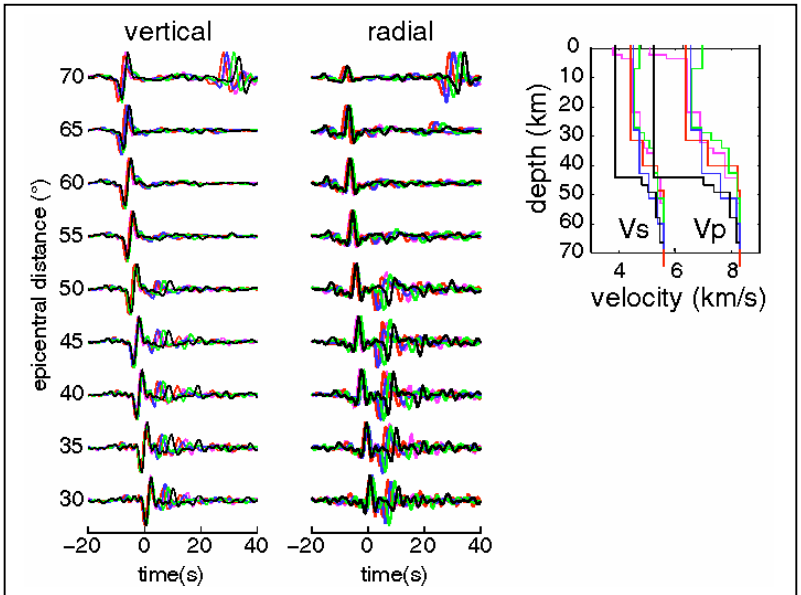


Figure 11. Synthetic seismograms computed for various receiver function models of the crust and upper mantle beneath China (models from Mangino et al., 1999). The primary differences between the seismograms are in the S coda, including SsPmP and SPL, which indicates these phases are highly sensitive to distinguishing structural details of these models.

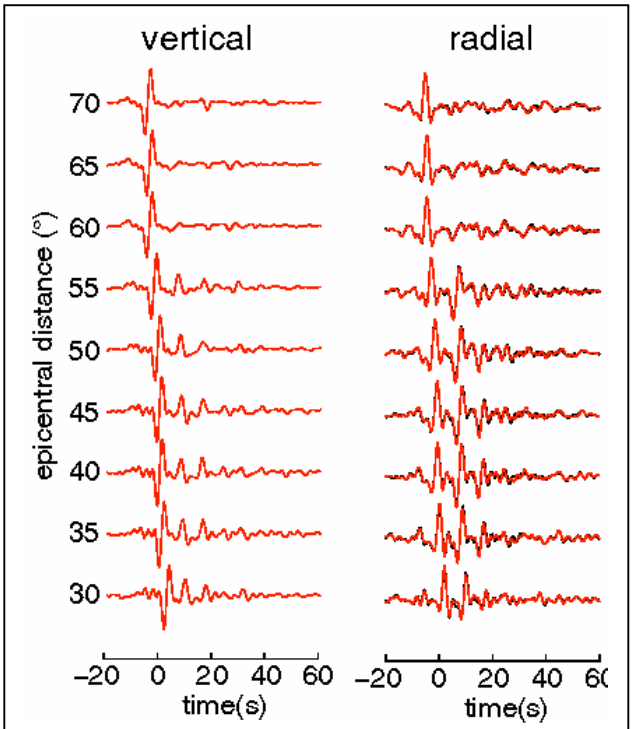


Figure 12. Waveform fits resulting from a synthetic inversion of seismograms computed for an input model shown in Figure 14 (black). Best-fitting seismograms after 400 iterations are shown in red.

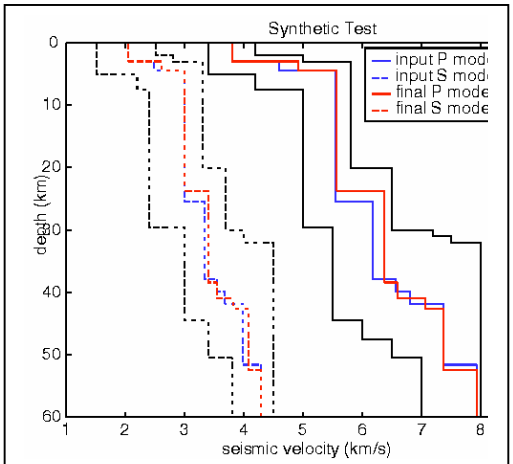


Figure 13. Input (blue) and output (red) models from the synthetic inversion test. Differences between inputs and outputs point to poor sensitivity of the phases to these parts of the model, and non-uniqueness in the inversion.

CONCLUSIONS AND RECOMMENDATIONS

We believe that modeling shear-coupled PL along with associated phases is a fruitful strategy for producing models of the crust and upper mantle that are most appropriate for locating regional seismic events. We have shown that the modeling scheme we proposed is feasible and practical but, for wide application and to evaluate uncertainties as proposed we need further increases in computational speed. We will further increase computational throughput by using additional processors (we have been unable to use all 16 processors due to minor hardware difficulties), a new cluster of faster processors, tailoring the number of ray parameters more carefully to the structure being modeled, and by pre-computing and storing reflectivity matrices for deeper parts of the mantle that will not vary in our modeling. The high frequencies needed to compute body waves accurately over the distances to which SPL is observed contribute to the time requirements for computing SPL. We plan to remove these complications by storing pre-computed products for fixed parts of the Earth model. A new cluster of processors at our institute will make our proposed approach even more practical. Evaluating objective functions and computing posterior marginal density functions and model covariance's require a negligible amount of computation time. For estimating uncertainty, we cast the inverse problem in a Bayesian framework such that, given prior information on data and model, the a posteriori probability density (PPD) function describes the answer to the inverse problem (Tarantola, 1987). Obtaining a complete description of the PPD, i.e., choosing how to sample models from the PPD, is the most challenging task. Due to the nonlinearity of the problem, an "importance sampling" approach based on Gibbs' sampler is the correct choice (Sen and Stoffa, 1998). However, this method is computationally very expensive. Sen and Stoffa (1995) have shown that the models sampled

Table 1. Events used in this study.

	Year	Month	Day	Hour	Min	Sec	Latitude (°)	Longitude (°)	Depth (km)	Mag
1	1990	5	24	20	9	23.2	-7.363	120.363	588.9	6.4
2	1994	9	28	16	39	51.7	-5.786	110.352	637.5	6.6
3	1994	11	15	20	18	11.3	-5.589	110.186	560.6	6.5

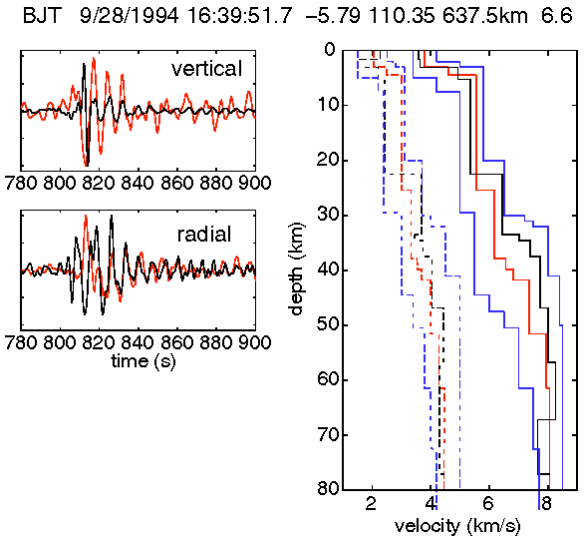


Figure 14. Optimization results after 600 iterations for fitting synthetics (black) to BJT vertical and radial records (red) of the 9/28/1994 event, using the time interval 815-860 s. Best-fitting model (black), search bounds (blue), and receiver function model by Mangino et al. (1999) (red).

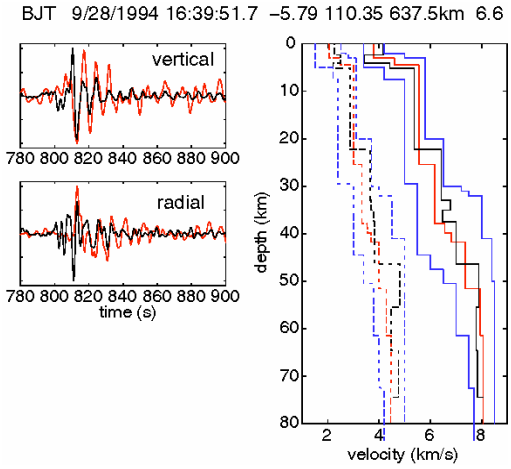


Figure 15. Same as in Figure 14 except that the interval 800-860 s was used in the waveform fitting.

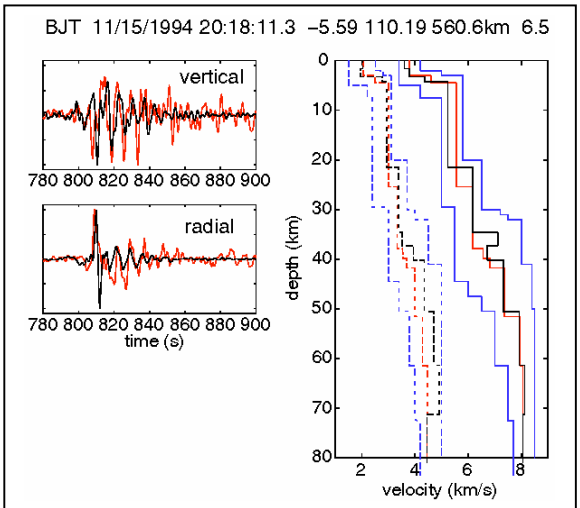


Figure 16. Optimization results after 600 iterations for fitting synthetics (black) to BJT vertical and radial records (red) of the 11/15/1994 event, using the time interval 815-860 s. Best-fitting model (black), search bounds (blue), and receiver function model by Mangino et al. (1999) (red).

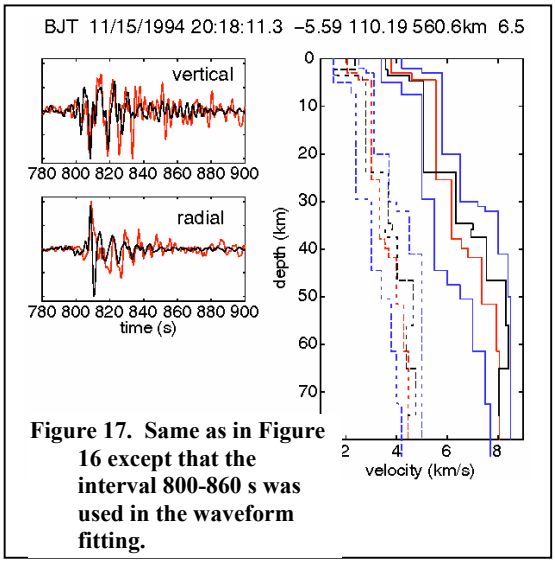


Figure 17. Same as in Figure 16 except that the interval 800-860 s was used in the waveform fitting.

by multiple modeling runs can be used to obtain quick estimates of the PPD that give fairly accurate estimates of the posterior model covariance's and correlation, although individual variances may be underestimated. We propose to use multiple VFSA runs to estimate the marginal PPD, mean and posterior covariance's, which will aid tremendously in the interpretation of our results. We will also employ the theoretically correct Gibbs' sampling approach with a narrow search range for a limited number of model layers.

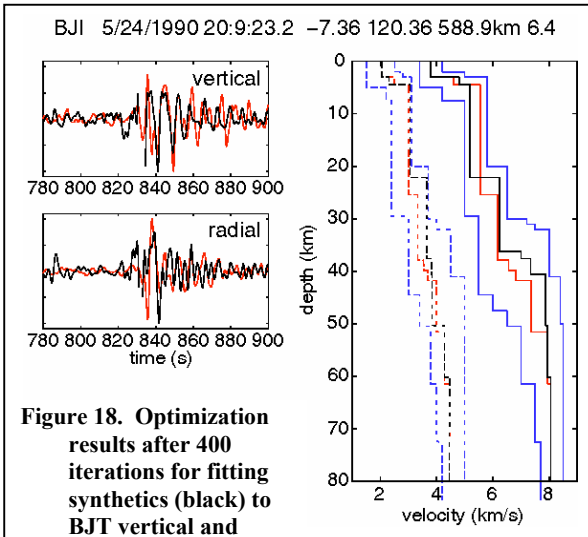


Figure 18. Optimization results after 400 iterations for fitting synthetics (black) to BJT vertical and radial records (red) of the 5/24/1994 event, using the time interval 835-870 s. See Figure 16 for other details.

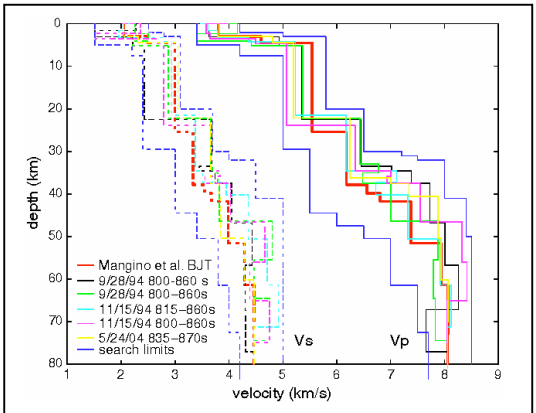


Figure 19. Comparison of models for the crust and upper mantle in the vicinity of station BJT resulting from waveform fits to various data (see legend), Mangino et al.'s (1999) receiver function model (red), and the simulated annealing search limits (blue).

REFERENCES

Baag, C.-E. and C.A. Langston, Shear-coupled PL, *Geophys. J. R. Astr. Soc.,* **80**, 363-385, 1985.

Beckers, J., S. Schwartz, and T. Lay, *Geophys. J. Int*., **119**, 574-594, 1994

Chander, R., Alsop, L.E., and Oliver, J., On the synthesis of shear coupled PL waves, *Bull. Seism. Soc. Am.,* **58**, 1849-1877, 1968.

Crotwell, H.P., T.J. Owens, and J. Ritsema, The TauP toolkit; flexible seismic travel-time and ray-path utilities, Seism. Res. Lett., **70**, 154-160, 1999.

Dziewonski, A. M., Anderson, D. L., Preliminary reference Earth model, *Phys. Earth. Planet. Int.,* **25**, p. 297-356, 1981.

Frazer, L.N., Synthesis of shear coupled PL. *Ph.D. thesis*, Princeton University, 54 pp., 1977.

Grand, S.P., and D. Helmberger, Geophys. *J. R. Astr. Soc.,* **76**, 399-438, 1984

Helmberger, D.V., and G.R. Engen, Upper mantle shear structure, *J. Geophys. Res.*, **79**, 4017-4028, 1974.

Jordan, T.H., and L.N. Frazer, Crustal and upper mantle structure from Sp phases, J. Geophys. Res., **80**, 1504-1518, 1975

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

Mangino, S., K. Priestley, and J. Ebel, The receiver structure beneath the China Digital Seismograph Network Station, *Bull. Seism. Soc. Am.,* **89**, 1053-1076, 1999

Oliver, J., On the long period character of shear waves, *Bull. Seism. Soc. Am*., **51**, 1-12,1961.

Sen, M., and P. L. Stoffa, 1995, **Global Optimization Methods in Geophysical Inversion**, Elsevier Science Publishing Company, The Netherlands.

Swenson, J.L., Beck, S.L., Zandt, G., Regional distance shear-coupled PL propagation within the northern Altiplano, Central Andes, *Geophys. J. Int.,* **139**, p. 743-753, 1999.

Tarantola, A., **Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation**, 613 pp., Elsevier, Amsterdam, 1987.

Zhao, L.-S., D. Helmberger, and D.V. Harkrider, *Geophys. J. Int.*, **105**, 713-730, 1991

Zhao, L.-S. and C. Frohlich. Teleseismic body waveforms and receiver structures beneath seismic stations, *Geophys. J. Int.,* **124**, 525-540, 1996.

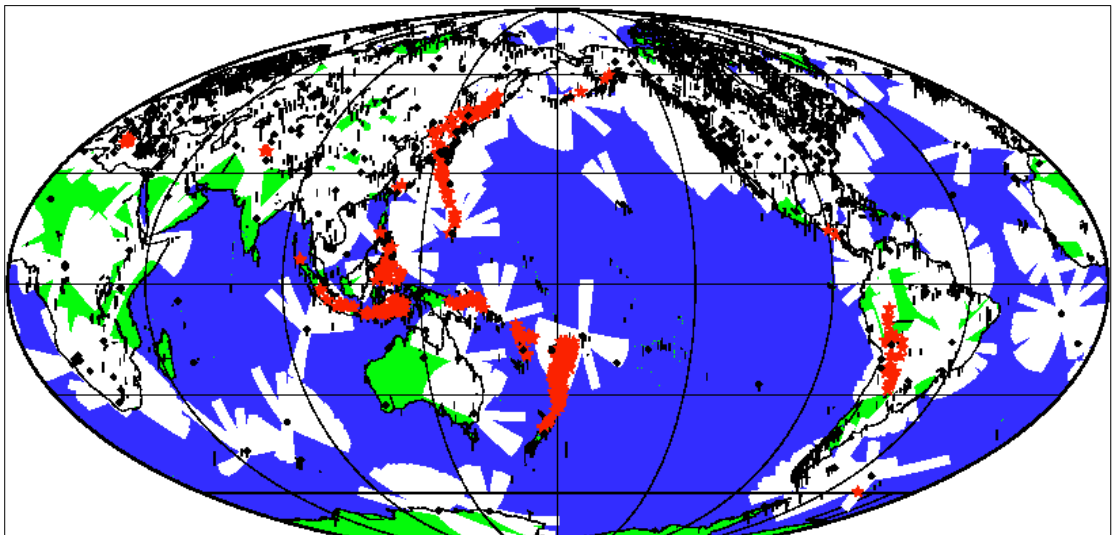


Figure 20. Potential global sampling of SPL waves using earthquakes (red stars) with source depths greater than 200 km for the years 1970-2002. Black triangles indicate broadband seismographic stations for which data are generally available. This testifies to the attractive potential for using SPL to model crustal and upper mantle structure beneath the continents. However, not all of these events will produce useful SPL phases. The amplitude of SPL depends on structural characteristics (e.g., a low-