JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. B4, PAGES 7229–7242, APRIL 10, 1999

Experiments in migration processing of SS precursor data to image upper mantle discontinuity structure

Peter M. Shearer, Megan P. Flanagan, and Michael A. H. Hedlin

Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla

Abstract.

Long-period SS precursors result from underside reflections off upper mantle discontinuities. By grouping and stacking global seismic data by SS bounce point location it is possible to map lateral variations in depths to the 410- and 660-km discontinuities, a process analogous to common midpoint (CMP) stacking in reflection seismology. Because this method assumes horizontal reflectors, energy arriving from dipping or intermittent reflectors may not be correctly imaged. To address this possibility, we experiment with techniques based on migration processing of shallow seismic reflection data. The problem is complicated by the uneven distribution of sources and receivers for the SS precursor observations, but the data are sufficiently dense beneath the northwest Pacific Ocean that reasonably good coverage can be obtained for this region. We parameterize the model as a grid of point scatterers in latitude, longitude, and depth (from the surface to 1000 km depth) and compute travel times from each grid point to the source and receiver locations. These times are used to construct a matrix equation that yields predicted SS precursor waveforms from the assumed scatterers. To recover the model, we experiment with both simple back projection and full inversions using a conjugate gradient method. Tests on noise-free synthetic data (generated using the same source-receiver distribution as the actual data) suggest that detailed resolution of discontinuity structure is possible, at horizontal scales much smaller than the Fresnel zone. However, the real data do not produce coherent results unless some degree of horizontal smoothing is imposed, at least partially defeating the purpose of this approach. Results for the northwest Pacific find structure on the 410- and 660-km discontinuities and hints of intermittent reflectors at other depths. Random resampling tests, however, suggest that most of these features are not reliably resolved, with the exception of a depression on the 660-km discontinuity seen in the northwest Pacific. Our experiments show that it is unlikely that small-scale structure on the 660-km discontinuity near subducting slabs causes significant bias in maps of the large-scale 660-km topography derived from long-period SS precursor observations.

1. Introduction

Observations of reflected and converted phases from seismic discontinuities in the upper mantle can be used to measure the depths to these interfaces [e.g., *Vinnik*, 1977; *Revenaugh and Jordan*, 1989, 1991; *Shearer*, 1990]. Given sufficient data, it is possible to map the discontinuity topography on either regional scales using P-to-SV conversions [*Shen et al.*, 1996; *Dueker and Sheehan*, 1997; *Li et al.*, 1998], near subducting slabs using near-source reflected or converted phases [e.g., *Wicks and Richards*, 1993; *Kaneshima and Helffrich*, 1998; *Flanagan and Shearer*, 1998b], or on global scales using SS precursors [e.g., *Shearer and Masters*, 1992; *Shearer*, 1993; *Gossler and Kind*, 1996; *Gu et al.*, 1998; *Flanagan and Shearer*, 1998a]. Analyses of these data to infer discontinuity depths have many similarities to processing algorithms routinely applied to seismic reflection data from the shallow crust. First, stacking of the waveforms is typically performed to improve the signal-to-noise ratios of the discontinuity phases. This is usually done by spatially grouping the records in bins of nearby reflection/conversion points on the discontinuities, a process analogous to the common midpoint (CMP) stacking method used in reflection seismology. Differential times between the precursors and a reference phase are then measured and converted to discontinuity depths by assuming a reference velocity model.

Paper number 1998JB900119.
0148-0227/99/1998JB900119$09.00

Next, corrections are applied for lateral variations in upper mantle velocities, surface topography and crustal thickness. Finally, discontinuity topography is inferred by the differences in the discontinuity depth estimates obtained for the different bin locations.

The CMP stacking approach implicitly assumes that the discontinuity phases are generated largely by near-horizontal structures close to the predicted reflection or conversion points on the discontinuities. If significant lateral variations in discontinuity properties are present, then off-axis reflections or conversions could occur; these would not obey the simple ray theory based on Snell's law that is applicable to horizontally stratified media. As a consequence, artifacts and biases in the inferred discontinuity depths could result from simple application of the CMP method. This has long been recognized in reflection seismology and has led to the development of back projection, or migration, techniques to account for off-axis scattering. Migration processing often produces a large improvement in the quality of reflection seismic sections. In this paper, we experiment with migration methods to see whether they can produce a similar improvement in analyses of upper mantle reflectors imaged by long-period SS precursor data. The problem is complicated by the low signal-to-noise of the SS precursors and the unevenly distributed sources and receivers. We find that simple back projection methods are not effective in imaging upper mantle discontinuities with SS precursor waveforms. The discontinuities can be imaged if the full inverse problem is addressed, although a large degree of horizontal smoothing must be applied to achieve stable results with our data set. For current SS precursor data sets, our migration methods do not appear to offer a significant advantage over the CMP approach.

The Fresnel zone provides a rough measure of the effective resolution of seismic data. The width of the bins used to stack discontinuity phase data is often chosen to be roughly comparable in size to the Fresnel zone [e.g., *Shearer*, 1993; *Gu et al.*, 1998; *Flanagan and Shearer*, 1998a]. However, it is likely that discontinuity topography at scale lengths smaller than the Fresnel zone will have some effect on seismic waveforms and inferred discontinuity depths, even if the topography cannot be directly mapped in the stacked data. Recently, *Neele et al.* [1997] and *Chaljub and Tarantola* [1997] showed that short-wavelength topography can bias discontinuity depth estimates obtained from long-period SS precursors. *Neele et al.* [1997] suggested that this effect could cause significant errors in estimates of long-wavelength topography on the 660-km discontinuity in the vicinity of subducting slabs. Using synthetic seismograms, we test this idea by comparing the effect of different slab models on 660-km discontinuity depths inferred from SS precursor data by *Flanagan and Shearer* [1998a] for the northwest Pacific. We find that the simple midpoint stacking approach is not significantly biased by probable near-slab structures in this region. The topography on the 660-km discontinuity imaged by long-period SS precursors requires the presence of a depression in the 660 comparable in size to the Fresnel zone; it cannot be produced by local near-slab depressions on the interface.

2. SS Precursor Observations

We use the SS precursor dataset of *Flanagan and Shearer* [1998a], who analyzed over 13,000 long-period, transverse-component seismograms from the global seis-

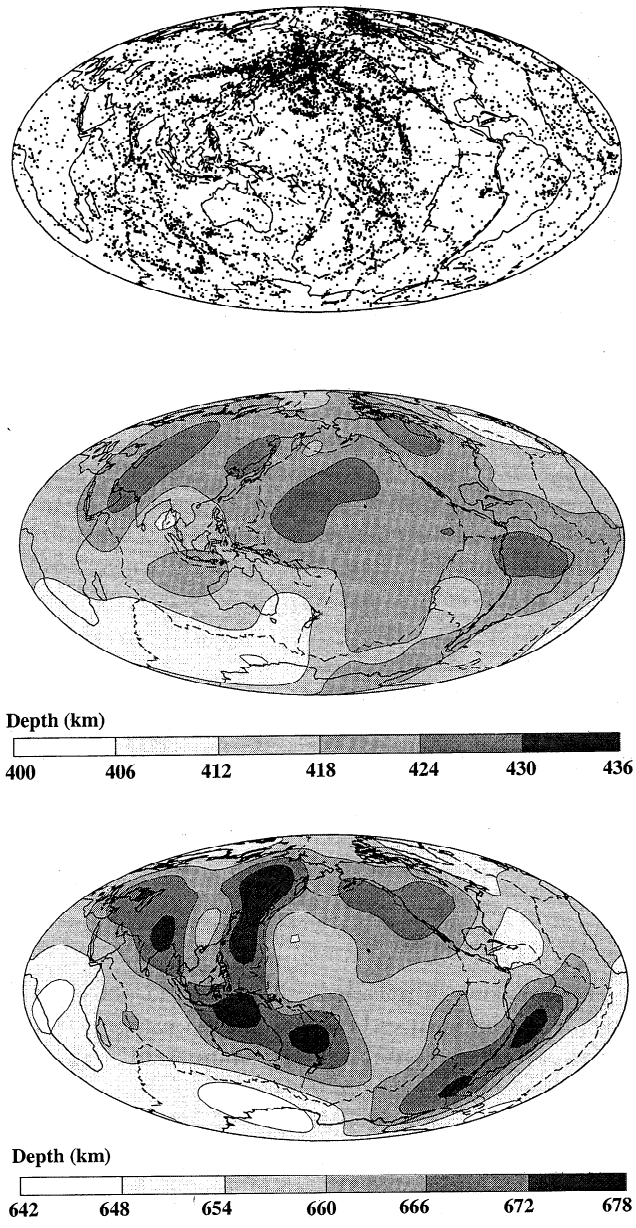

Figure 1. The SS surface bounce points for the data analyzed by *Flanagan and Shearer* [1998a] and the maps of topography on the 410- and 660-km discontinuities inferred from analysis of SS precursors. Note the extensive bounce point coverage in the northwest Pacific and the depression in the 660-km discontinuity seen in this area.

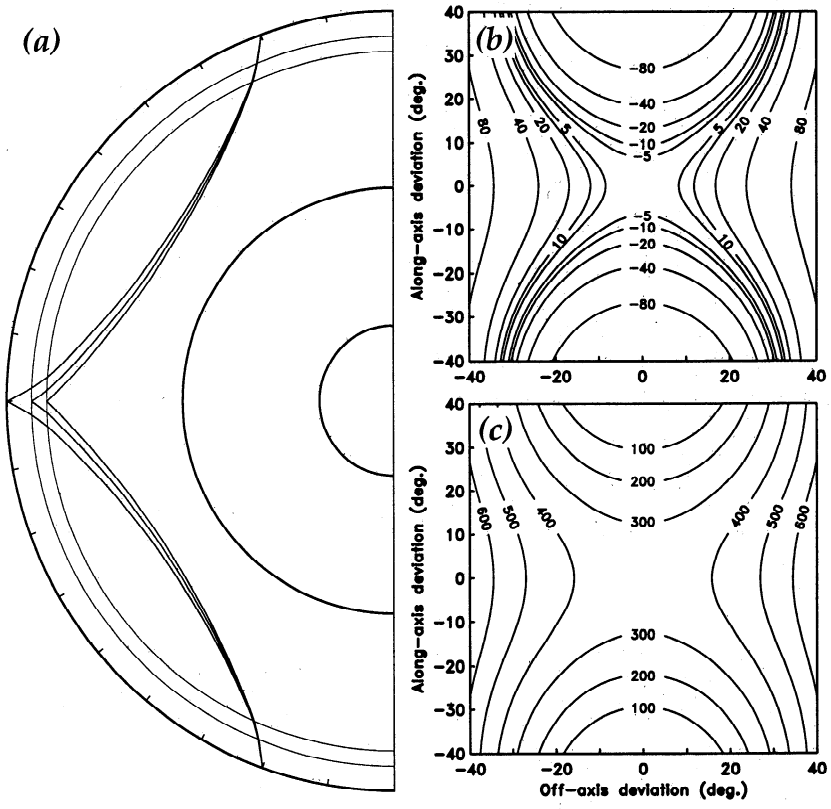

Figure 2. (a) The geometry of the SS, $S410S$, and $S660S$ ray paths, computed from the IASP91 velocity model [*Kennett*, 1991]. (b) Contours of travel time differences for lateral perturbations to the SdS bounce point at 140° source-receiver range for a discontinuity at 350 km depth. The Fresnel zone is defined by the $\pm T/4$ contours, where T is the wave period. (c) Contour of depths to hypothetical point scatterers that produce the same precursor arrival time as a midpoint reflector at 350 km depth, assuming the source-receiver range is 140°.

mic networks recorded from 1976 to 1996. The bounce point distribution of these data and maps of topography on the 410- and 660-km discontinuities produced by Flanagan and Shearer are shown in Figure 1. The discontinuity depths were obtained by stacking the data with common SS bounce points (using 416 caps of 10° radius), identifying the 410- and 660-km reflectors, converting the arrival times to depth (assuming a reference one-dimensional velocity model), and applying corrections for lateral variations in crustal and upper mantle structure. Finally, a smooth surface was fit to the individual depth estimates and their standard errors using a spherical spline approach to produce the maps shown in Figure 1. A complete description of this analysis is given by *Flanagan and Shearer* [1998a].

The geometry of the SS precursor ray paths is illustrated in Figure 2a for a source-receiver range of 140°, computed using the IASP91 velocity model [*Kennett*, 1991]. The precursors result from underside reflections off the upper mantle discontinuities and arrive 2 to 4 min before the main SS pulse. The precursor phases are denoted SdS, where d is the discontinuity depth. The amplitude of the precursors is typically <5% of SS, making the precursors hard to detect on individual seismograms. Their visibility is greatly enhanced if stacking methods are applied to reduce the level of incoherent noise in the precursor wave field. The ray theoretical bounce points for reflections from horizontal interfaces occur at the midpoints between the sources and receivers (assuming both are at the surface). The Fresnel zone for these reflections may be outlined by plotting the travel time differences for reflection points that deviate from the midpoint position. These time differences are contoured in Figure 2b for a reflector at 350 km depth. Note the mini-max behavior of the SS ray path. It is a maximum time point for bounce point perturbations along the great circle path connecting source and receiver and a minimum time point for bounce point perturbations orthogonal to the source-receiver azimuth. Thus the time differences plotted in Figure 2b form a saddle-shaped surface. The Fresnel zone is defined by the $\pm T/4$ contours, where T is the period of the wave. For $T = 30$ s (the approximate period of our data set), the Fresnel zone is 15° to 20° across (at 140° source-receiver range). The dimensions of the Fresnel zone are approximately constant for different discontinuity depths.

Considering the possibility of off-axis scattering, sin-

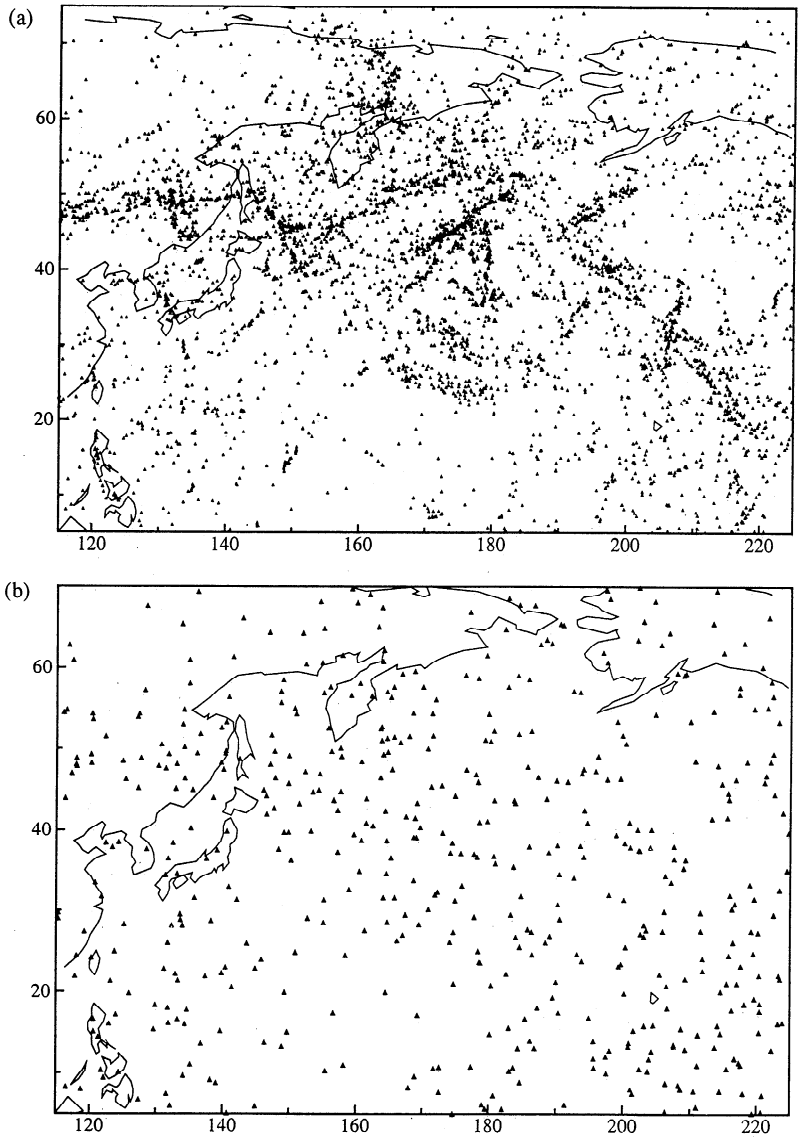

Figure 3. (a) Surface bounce points for 3988 SS phases sampling the northwest Pacific and northeast Asia from the data set of *Flanagan and Shearer* [1998]. (b) SS bounce points of 506 summary rays with similar source-receiver paths. Note the more uniform spatial coverage compared to the original data.

gle point scatterers at a variety of depths can produce the same SS precursor arrival time. This is illustrated in Figure 2c, a contour plot of depths to hypothetical point scatterers that will produce the same precursor arrival time as a midpoint reflector at 350 km depth. The assumption in the common midpoint stacking approach is that observed SS precursor energy can be mapped exclusively to the source-receiver midpoints and ignores the possibility that the energy may have originated from scattering at other depths. The effect of off-axis scattering can be accounted for with migration processing, and the surface plotted in Figure 2c is analogous to the migration hyperbola often used to explain the effect of scattering in crustal reflection data. It is the purpose of this paper to test whether off-axis scattering is a significant factor in SS precursor data and if migration methods might provide significant improvements in the imaging of upper mantle reflectors. Previous applications of migration techniques to image mantle scatterers include *Kaneshima and Helffrich* [1998] and *Castle and Creager* [1998]; these studies focused on S-to-P conversions rather than the underside S reflections examined here.

Migration methods are likely to work best in areas with the most complete data coverage; for this reason, we focus our analysis on the northwest Pacific region where the SS bounce point coverage is relatively dense. Figure 3a shows 3988 SS bounce points from the *Flanagan and Shearer* [1998a] data set for a large part of the northwest Pacific Ocean and northeast Asia. The data coverage within this region is much more dense in some areas than others; it is likely that this could

cause problems in back projection methods since only a few source-receiver paths could dominate the observations. To make the data coverage more uniform, we stack the records into summary rays from similar source and receiver locations. We divide the Earth's surface into 416 cells, spaced ~10° apart and assign a cell number to each source and receiver location. We then stack the waveforms for which both the source and receiver cells are identical, thus combining the data with similar ray paths. The stacking method, described at length by *Flanagan and Shearer* [1998a], involves both mild low-pass filtering and deconvolution of the SS reference pulse. The stacking reduces the data set to 506 summary rays whose source and receiver locations are determined by the mean values of the data in the stack. SS bounce point locations for the summary rays are plotted in Figure 3b; note the more even spatial sampling of these points compared to Figure 3a. Stacking into summary rays has the further advantage of reducing the size of the data set, thus lessening some of the computational difficulties associated with the full inversion.

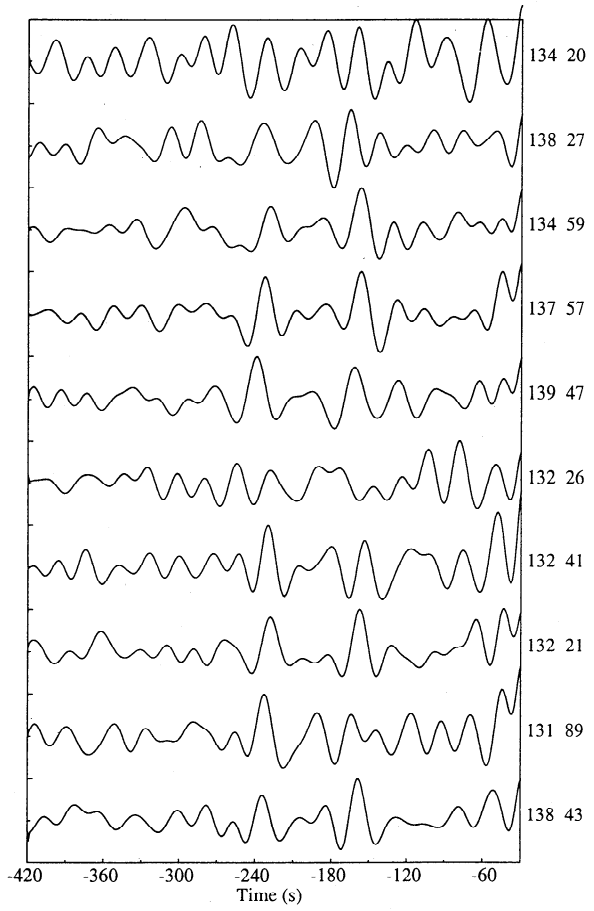

Figure 4. Examples of SS precursor waveforms obtained from stacking of long-period, transverse-component seismograms into summary rays with similar source and receiver locations. Time is relative to the main SS arrival. The numbers to the right give the average range and the number of seismograms in the stack. For example, the bottom trace is a stack of 43 traces at an average source-receiver range of 138°. The peaks in the stacked traces near -160 s and -240 s are reflections off the 410- and 660-km discontinuities; note that the visibility of these peaks varies among the different summary rays.

The summary ray stacking preserves much of the signal in the data that constrains the locations of possible off-axis scatterers because the stacking combines data only from similar source-receiver paths. In contrast to common midpoint stacking, waveforms are not grouped that have substantial differences in source and receiver locations. Many of the summary ray bounce points plotted in Figure 3b are much closer together than 10°; these represent different source-receiver ranges or ray azimuths at the bounce point. The limited nature of the summary ray stacking implies that the migration methods applied later in this paper are more analogous to prestack migration than poststack migration. Nonetheless, there is necessarily some loss of information from the stacking which may reduce the ability of migration to resolve fine-scale structure in the model. Some examples of the summary ray-stacked waveforms are shown in Figure 4. The 410- and 660-km reflections can often be identified in these stacks; differences in the positions of these peaks relative to SS result from the different source-receiver ranges in the summary rays and lateral discontinuity topography and seismic velocity variations in the upper mantle.

3. Imaging Upper Mantle Reflectors

Our goal is to recover a model of reflector (scattering) strength as a function of position in the Earth. In traditional reflection seismology experiments the data are closely spaced at equal intervals, and it is often convenient to define the model spacing to be identical to the data spacing. For our SS precursor observations the data coverage is irregular and sometimes sparse, so we define our model without regard to the exact positions of the SS bounce points. The model space extends from -10° to 86°N, 100° to 240°E, and 0 to 1000 km in depth, with a grid spacing of 2.5° in longitude, 2.0° in latitude, and 10 km in depth, for a total of 277,144 model points. For comparison, there are 197,340 data points, in 506 summary ray seismograms, digitized at 1 sample/s (using data from 420 s to 30 s prior to the SS arrival). It should be noted, however, that the effective number of data points that constrain the model is much less, since the long-period waveforms have a dominant period of about 30 s and are oversampled at 1 Hz.

3.1 Common Midpoint Stacking

To compare with our subsequent, more sophisticated, processing methods, we first perform simple common midpoint stacking on the data. For each model point we identify those summary ray seismograms with SS bounce point locations within 15° of the latitude and longitude of the model point. From the depth of the

model point, d, we compute the ray theoretical $SS - SdS$ travel time for each seismogram (based on the Preliminary Reference Earth Model (PREM) velocity model [*Dziewonski and Anderson*, 1981]), assuming a reflection from a horizontal layer midway between source and receiver. We then sum and average the data values at these points over all available seismograms. To minimize truncation artifacts from the edge of the averaging cap, we weight the data by bounce point distance from the model latitude and longitude, using a $\cos^2$ taper. This calculation naturally incorporates the travel time differences that are expressed as a "moveout correction" in traditional CMP stacking. Our stacking approach is similar to the method used by *Flanagan and Shearer* [1998a], except here we are sampling the model much more finely in horizontal position. Also, we do not apply corrections for variations in crustal thickness and upper mantle velocity structure because our emphasis in this paper is to compare the results between different methods and possibly image intermittent and/or fine-scale structure on the discontinuities, rather than obtain the most accurate discontinuity depth estimates.

Results of this processing method are displayed in Figure 5 for two W-E cross sections through the model, one at a latitude of 30°, the other at a latitude of 50°. The 410- and 660-km discontinuities are clearly visible, with hints of reflectors at other depths. The width of the 410 and 660 peaks is roughly approximated by the mapping of the SS reference pulse from time into depth; note that there are negative sidelobes surrounding each peak. These and subsequent model plots are self-scaled to unit maximum amplitude. The plot to the right of the main image shows the result of stacking over all model points at constant depth; here only the 410- and 660-km discontinuities are clearly visible (a weaker 520-km discontinuity is also present but is difficult to see at this scale). The rightmost traces showing the model stack in this and subsequent figures are self-scaled for greater visibility. The 660-km discontinuity appears depressed along the left (west) side of the model, consistent with the trough in the 660 that has been observed in previous studies [e.g., *Shearer*, 1991, 1993; *Shearer and Masters*, 1992; *Flanagan and Shearer*, 1998a].

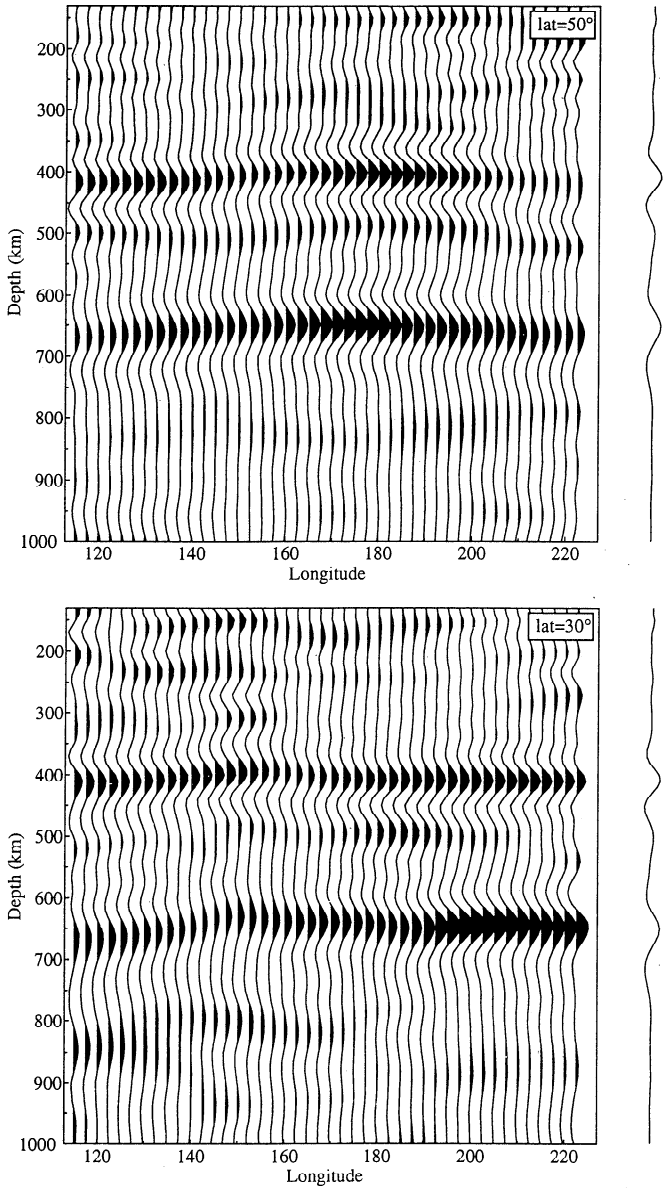

Figure 5. Model cross sections from the northwest Pacific region obtained using common midpoint (CMP) stacking, at latitudes of (bottom) 30°N and (top) 50°N. For this method, the data are stacked within caps of 15° radius, weighted by radius with a Hanning taper. A time dependent moveout correction is applied to each trace prior to stacking. Note the clear imaging of the 410- and 660-km discontinuities. The single trace to the right shows the result of stacking the entire model over latitude and longitude, illustrating the dominance of the 410- and 660-km features. The depression seen in the 660-km discontinuity at the left edge of the model is the trough in the northwest Pacific identified in previous SS precursor studies (see Figure 1).

3.2 Back Projection

Next, we apply a method based on the concept of Kirchhoff migration, or diffraction stacking, in reflection seismology. We assume that each model point, i, is a hypothetical point scatterer of strength m_i and that the complete model is given by the vector $\mathbf{m}$. We consider single scattering only; it is likely that multiple scattering in long-period SS precursor data is of secondary importance since typical discontinuity reflection coefficients are <10%. By ray tracing, we compute the theoretical total source-to-scatterer-to-receiver travel time for each model point and for each summary ray. This time will correspond to a particular point on the summary ray seismogram, and a synthetic seismogram can be constructed as the sum of the contributions from assumed sources at all the model scattering points. Thus the general forward problem may be expressed as

$$\mathbf{d} = \mathbf{Gm} \tag{1}$$

where $\mathbf{d}$ is the data vector (the individual summary

ray seismograms placed end to end) and **G** is a matrix describing the linear relationship between the scattering strength in the model and the seismic observations. This equation is deceptively simple; the difficulties arise in actually computing values for the **G** matrix and solving the system for **m**. These complications include the following:

1. The ray theoretical amplitude differences given by variations in geometrical spreading for the different ray paths should be taken into account. For the ray geometries of the SS precursors, we found that these differences were sufficiently small that they could be ignored.

2. In general, the scattering strength will vary as a function of ray angle. Here we approximate the scattering as isotropic and do not include any obliquity factors, nor do we attempt to relate the absolute amplitude of the SS precursors to impedance contrasts in the model.

3. The grid points represent different scattering volumes at different latitudes due to the constant grid spacing in longitude.

4. Scattering at each grid point maps into multiple data points on each seismogram due to the duration of the source. If the source-time function is $S(t)$, then the effective source-time function to use in the integration over the different scattering points is given by $S_e = (d/dt)S(t)$, a result that follows from Kirchhoff theory. In our case, the appropriate source-time function may be estimated from the SS reference pulse, $S_{SS}(t)$. However, there is a further complication in that the SS pulse is Hilbert transformed relative to the original source-time function [e.g., *Choy and Richards*, 1975]. Thus the appropriate effective source to use at the scattering points for our problem is given by $S_e(t) = (d/dt)\mathcal{H}^{-1}[S_{SS}(t)]$, where $\mathcal{H}^{-1}$ is the inverse Hilbert transform. Remarkably, the d/dt and $\mathcal{H}^{-1}$ operators have nearly opposite effects (see Figure 6) for long-period SS waveforms, so $S_e(t)$ can be reasonably approximated as simply $S_e(t) \approx S_{ss}(t)$.

5. The large size of **G** makes explicit storage of the complete matrix impractical on most computers. The problem becomes more tractable if only the nonzero elements of **G** are saved, but even these storage requirements are significant due to the size of **m** and the fact

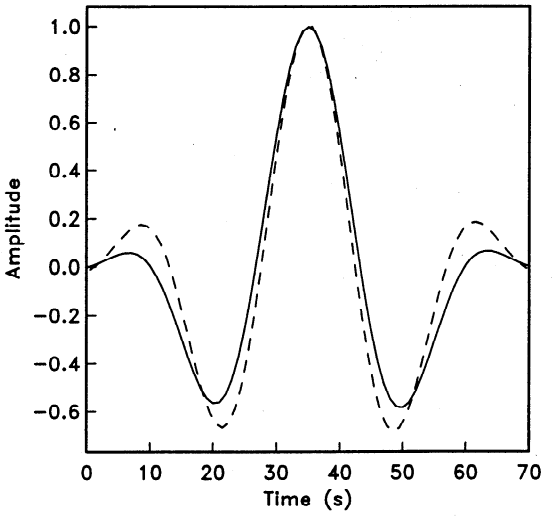

Figure 6. A stack of the SS pulses in the long-period data, $S_{SS}(t)$ (solid line) is a close approximation to the effective source-time function, $S_e(t) = (d/dt)\mathrm{H}^{-1}[S_{SS}(t)]$ (dashed line) required by the Kirchhoff calculation.

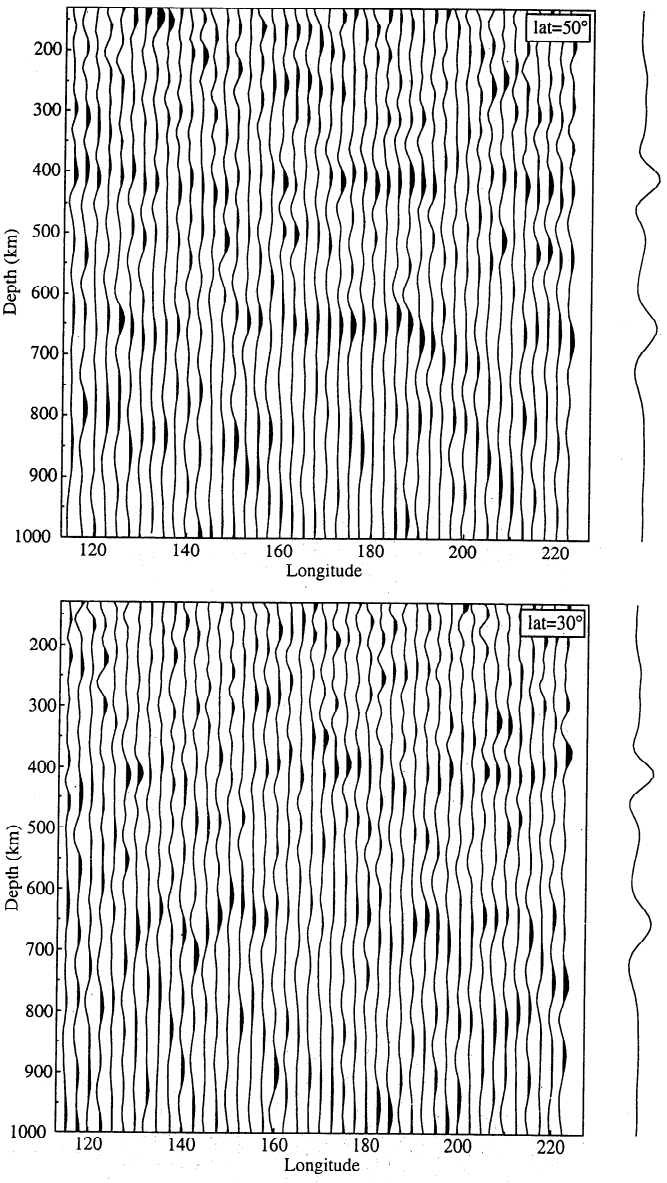

Figure 7. Model cross sections at latitudes of (bottom) 30° and (top) 50°, obtained using simple back projection of the data. Note that the results are much less spatially coherent than the CMP image shown in Figure 5 but that the 410- and 660-km reflectors are faintly visible. The single trace to the right shows the result of stacking the entire model over latitude and longitude (not just the model points shown in cross section); this emphasizes the 410- and 660-km features.

that our effective source-time function, $S_e(t)$, contains 70 points (see Figure 6).

Given the computational difficulties associated with solving for $\mathbf{m}$ in equation (1) for large data sets, migration methods have typically involved a simple back projection of the data. In this case, the solution to (1) is approximated as

$$\mathbf{m} = \mathbf{G}^T\mathbf{d} \tag{2}$$

In practice, this means that each data point of the seismograms is back projected to the possible scattering points that could have generated energy at the appropriate time.

Results of back projection applied to the SS precursors are displayed in Figure 7 for the two W-E cross sections previously examined in Figure 5. The results are disappointing; the 410- and 660-km reflectors are only intermittently visible and exhibit little spatial coherence. This presumably results from the uneven sampling in the data (see Figure 3b) which prevents the clear construction of the reflectors as a coherent feature in the back projection approach. The significance of the 410 and 660 peaks is clearly established only when the model is stacked over identical model depths (see the right-hand traces in Figure 7), in which case they are just as visible as in the CMP stack (Figure 5).

3.3 Full Inversion

Back projection provides only an approximate solution to the inverse problem of recovering the best model from the data, and the irregular coverage of our observations may be particularly problematic for this technique. Next, we experiment with more complete inversions of equation (1) to see if better results can be obtained. We use a conjugate gradient method that is suited to large inverse problems. To test our algorithm, we first apply the method to a synthetic example generated using the same distribution of sources and receivers as the actual data. We assume a model with a 220-km reflector east of 180°E, a flat 400-km discontinuity with 4° gap at 150°E, and an undulating 660-km discontinuity with 100 km of relief (see Figure 8a). The model properties are constant with latitude, varying only with longitude. Limitations in our available computer memory have so far restricted our results to models with a depth spacing of 20 km. Using equation (1), we use this model to generate a synthetic data set $\mathbf{d}$ and then invert for $\mathbf{m}$ using the conjugate gradient method. The result of the inversion is shown in Figure 8b following 10 iterations, producing a nearly perfect fit to the synthetic data; the data variance reduction is 99.5%. The model is recovered remarkably well; features smaller than the Fresnel zone are imaged, such as the gap in the 400-km reflector and the abrupt termination of the 220-km discontinuity.

However, this test is idealized because it assumes noise-free data and a perfectly known $\mathbf{G}$ matrix. Our results obtained with real data are less encouraging. Applying the method to the SS precursor observations, we obtain rapid convergence after a small number of iterations to a model (Figure 9) that produces a close fit to the data vector (Figure 10), achieving a variance reduction of better than 80% after 20 iterations. The model contains many features, but the 410- and 660-

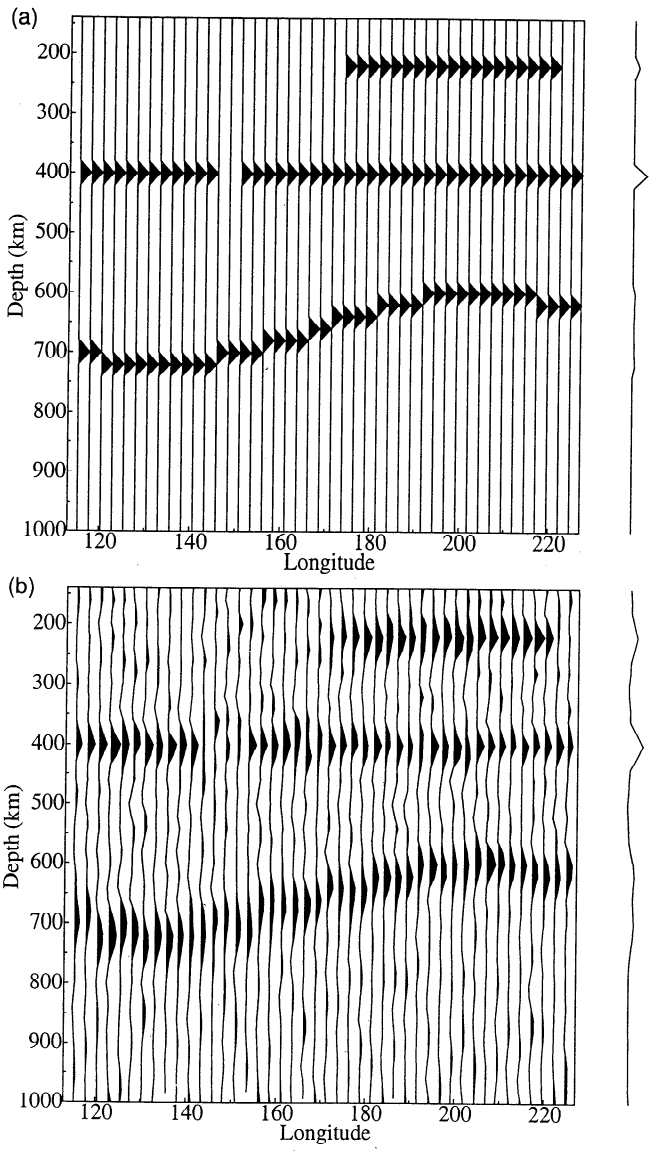

Figure 8. A test of the theoretical ability of the full inversion method to resolve small-scale features in the model along an W-E cross section at 30°N. (a) An ad hoc model containing a flat 400-km discontinuity (with a small gap), an undulating 660-km discontinuity, and a regional 220-km discontinuity. Using the same source and receiver geometries as the real data set, we generate synthetic seismograms and then invert them to recover the model, using 10 iterations of the conjugate gradient method (with no regularization). (b) Despite some artifacts, the model is recovered reasonably well, including some features (such as the gap in the 410) that are much smaller than the Fresnel zone of the SS precursors.

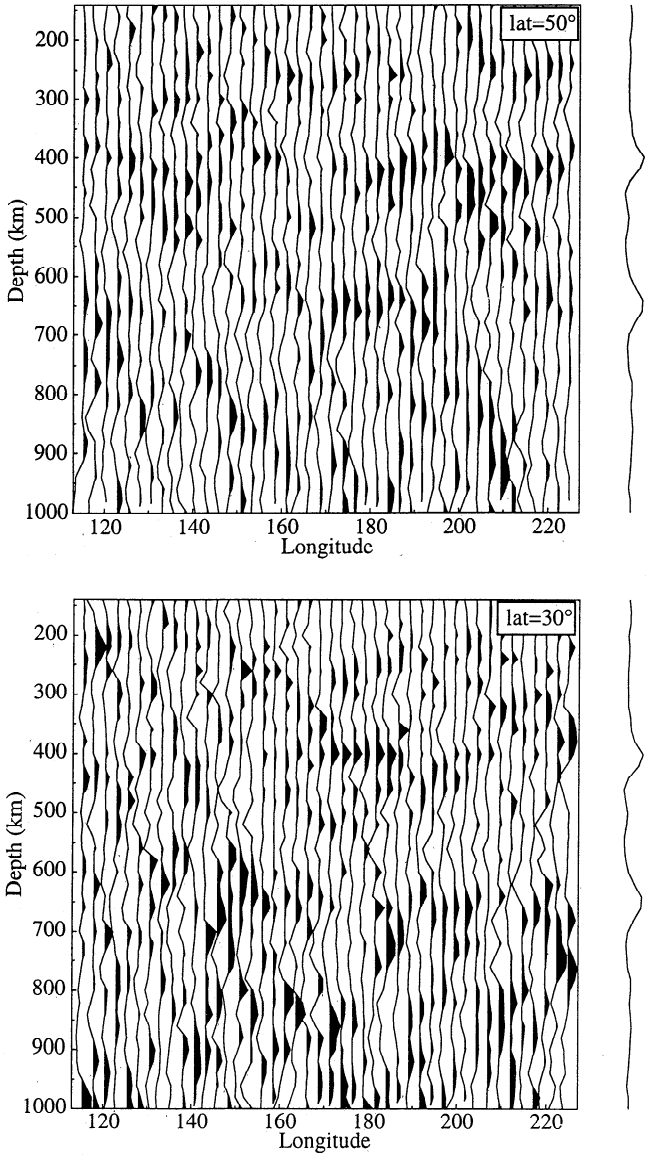

Figure 9. Model cross-sections at latitudes of (bottom) 30° and (top) 50°, obtained from a full inversion of the data, using 20 iterations of a conjugate gradient method. This model achieves a variance reduction of 80% in the precursor data (see data fit in Figure 10). The single trace to the right shows the result of stacking the entire model over latitude and longitude.

km discontinuities are only intermittently visible; other apparent reflectors are horizontally coherent for only a few model points or less. The model appears different from the simple back projection image shown in Figure 7, but it is not clear if the inversion results represent an improvement.

The incoherent, small-scale structure in the model obtained from the inversion suggests strongly that the inverse problem is overparameterized and that some form of regularization must be applied to obtain a stable solution. Thus we apply size and smoothness constraints to the model by replacing equation (1) with

$$\begin{bmatrix} \mathbf{d} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{G} \\ \lambda_1 \\ \lambda_2 \mathbf{L} \end{bmatrix} \mathbf{m} \tag{3}$$

where $\mathbf{L}$ is the finite difference approximation to the two-dimensional (horizontal) Laplacian applied over all model blocks and λ_1 and λ_2 are parameters that control the degree of damping and horizontal smoothing. We do not apply vertical smoothing because the reflectors in the upper mantle are expected to be sharp features relative to our grid spacing.

As is typical of many geophysical inverse problems, this type of regularized inversion leads to a trade-off between data fit and model amplitude and smoothness. The choice of appropriate values for λ_1 and λ_2 is subjective in our case because we lack an objective measure of the expected level of "noise" in the data, so we have no rigorous way to determine how well we should fit the data. We must rely on the prejudice that spatially coherent reflectors are present in the upper mantle, even

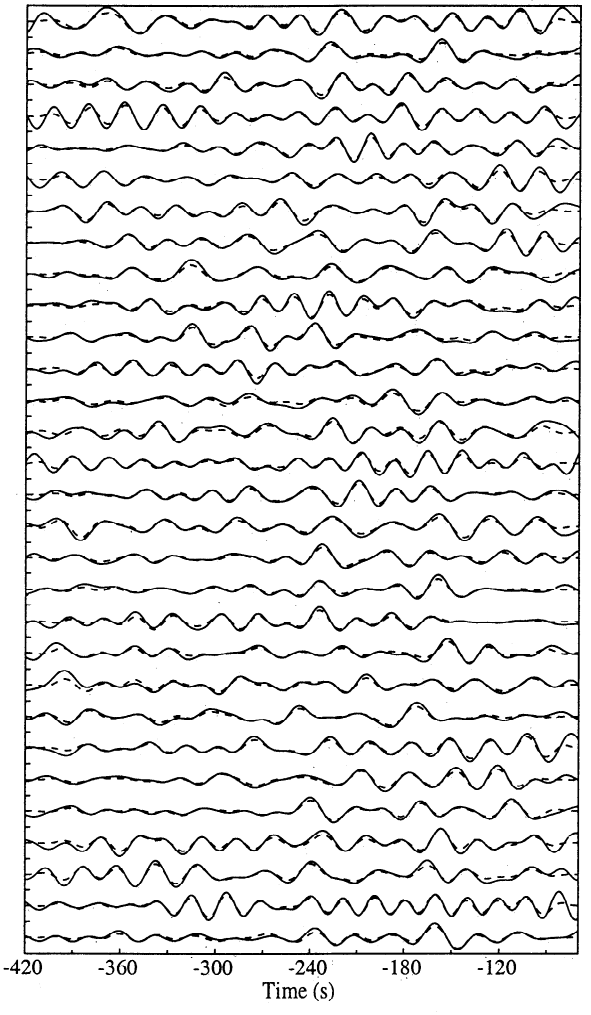

Figure 10. Example of the fit to the SS precursor data achieved by the conjugate gradient inversion method. The first 30 of the summary ray waveforms are shown as the solid lines; the dashed lines show the predictions of the model shown in Figure 9.

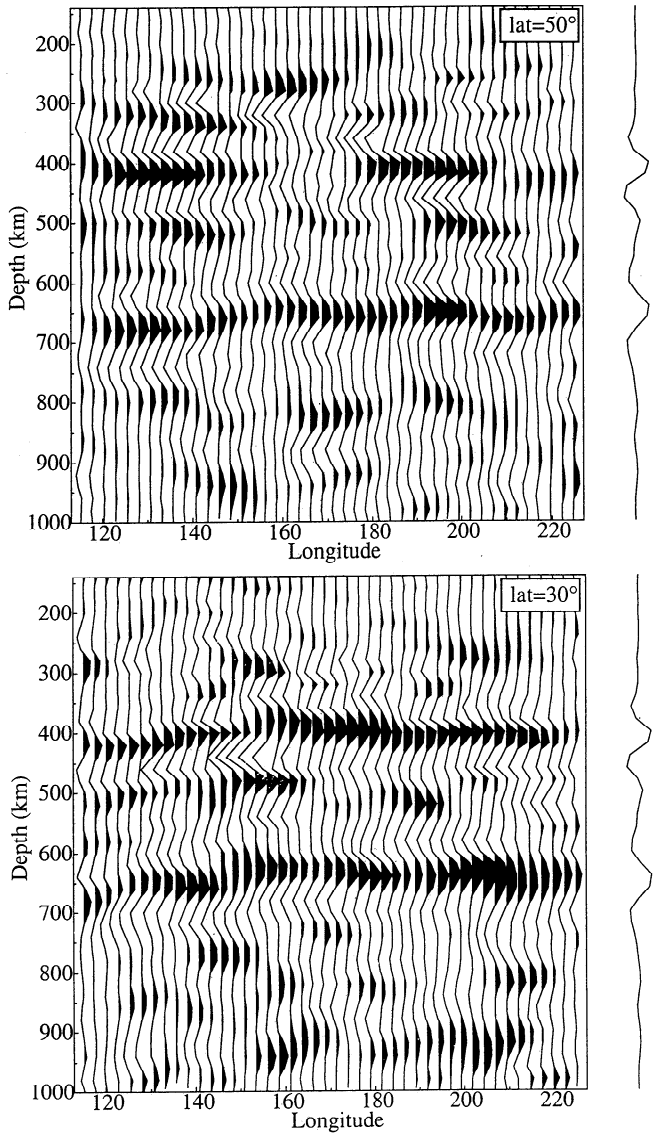

Figure 11. Model cross sections at latitudes of (bottom) 30° and (top) 50°, obtained from a full inversion of the data, using a conjugate gradient method with a horizontal smoothness constraint. Note the improvement compared to the simple back projection (Figure 7) and the undamped inversion (Figure 9) and the similarity of many features with the CMP results. This model achieves a variance reduction of 23% in the precursor data. The single trace to the right shows the result of stacking the entire model over latitude and longitude.

if this means that we will fit the data less closely than can be achieved with incoherent, small-scale structure. After some experimentation with different degrees of regularization, we obtained the model shown in Figure 11 following 20 iterations of the conjugate gradient method; it contains reasonably coherent discontinuity reflections and achieves a data variance reduction of 23%. Comparisons of model predictions to the summary ray data are shown in Figure 12. Notice that the peaks near -160 s and -240 s that result from the 410- and 660-km reflectors are often fit closely but that other peaks on individual summary rays are generally not fit.

The model plotted in Figure 11 appears similar to the CMP stacking image shown in Figure 5 but apparently reveals some features more clearly. These include a gap in the 410-km discontinuity near (30°N, 170°E) and intermittent reflectors at other depths. In addition, the relatively sharp eastern edge to the 660-km depression indicated in the CMP image at 50°N also appears in the full inversion. We note that the position of the eastern margin of the 660-km depression is very close to where the subducting slab intersects the discontinuity. This suggests that the cold thermal anomaly associated with the slab is deflected into the transition zone on the back arc side of the subduction zone, as discussed by *Shearer and Masters* [1992] and *Flanagan and Shearer* [1998a].

How reliable are the small-scale features imaged in Figure 11? A full resolution analysis is beyond the scope of this paper. However, a simple test in which we divide

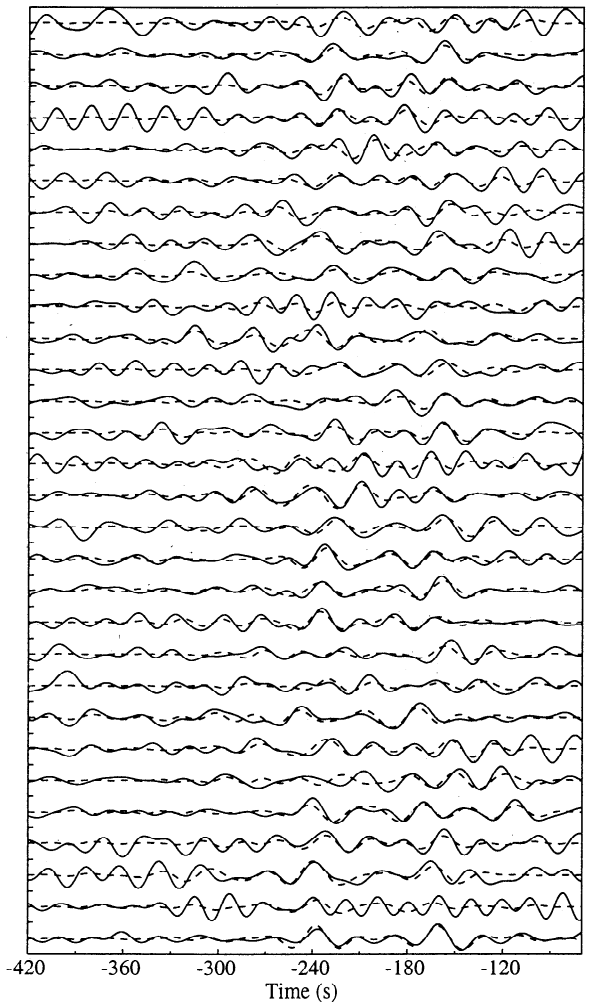

Figure 12. Example of the fit to the SS precursor data achieved by the regularized inversion. The first 30 of the summary ray waveforms are shown as the solid lines; the dashed lines describe the predictions of the model shown in Figure 9. Notice the reduction in the fit compared to that achieved for the undamped inversion in Figure 10.

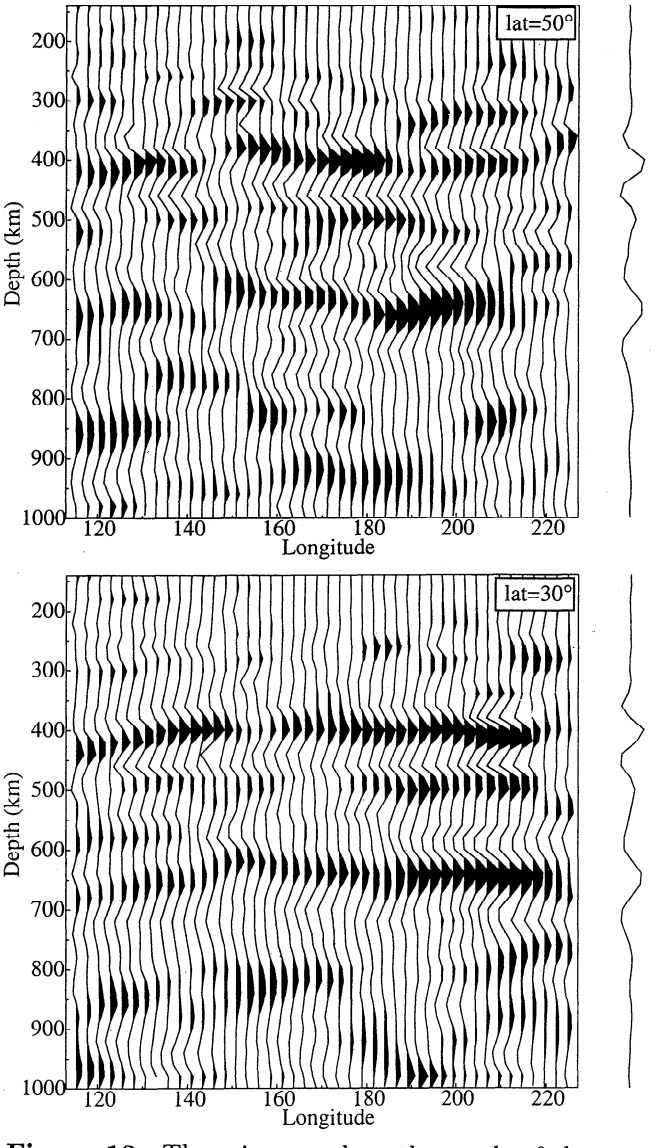

Figure 13. These images show the result of the regularized inversion applied to two random halves of the original SS precursor dataset, as plotted in cross section at a latitude of 30° (compare to the bottom part of Figure 11).

the data randomly in half and perform separate inversions (Figure 13) suggests that most of the details are not robust with respect to random changes in the data sampling. Thus the reliability of many of the features in the Figure 11 image is questionable. These include the apparent discontinuity near 250 km depth in Figure 13 (top) and the feature near 930 km in Figure 13 (bottom). The uncertainties derived from the data sampling, combined with those associated with the choice of damping parameters, suggest that the model shown in Figure 11 is not unique, and most of the details of the discontinuity structure are not reliably resolved.

One feature that does appear robust is the 20 to 30 km depression in the 660-km discontinuity along the western edge of the model, which is seen consistently in all of our inversions, as well as the model of *Flanagan and Shearer* [1998a]. This depression appears associated with the subduction zones in the northwest Pacific [e.g., *Shearer and Masters*, 1992]. We had hoped that the inversion method might resolve details along the eastern boundary of this depression. Unfortunately, Figure 13 shows that this part of the image is highly sensitive to the data sampling; Figure 13 (top) shows an abrupt change in the 660 depths near 145° longitude whereas Figure 13 (bottom) shows a more gradual change. Finally, we note that our model includes a region near Izu-Bonin for which *Niu and Kawakatsu* [1997] found evidence for a discontinuity near 970 km depth; we cannot clearly identify such a feature in our images.

4. A Test of Possible Bias From Near-Slab Structure

The depression in the 660-km discontinuity behind the subducting slabs in the northwest Pacific (see Figure 1) identified in previous SS precursor studies [e.g., *Shearer*, 1991, 1993; *Gossler and Kind*, 1996; *Gu et al.*, 1998] is one of the most robust features in maps of 660 topography. However, *Neele et al.* [1997] proposed that these maps may be biased by the effect of small-scale structure. In particular, Neele et al. suggested that narrow (4° wide) discontinuity depth anomalies along the subduction zones in the northwest Pacific could be falsely mapped into broad regional anomalies similar to those observed. *Chaljub and Tarantola* [1997] further explored this effect using finite difference modeling ap-

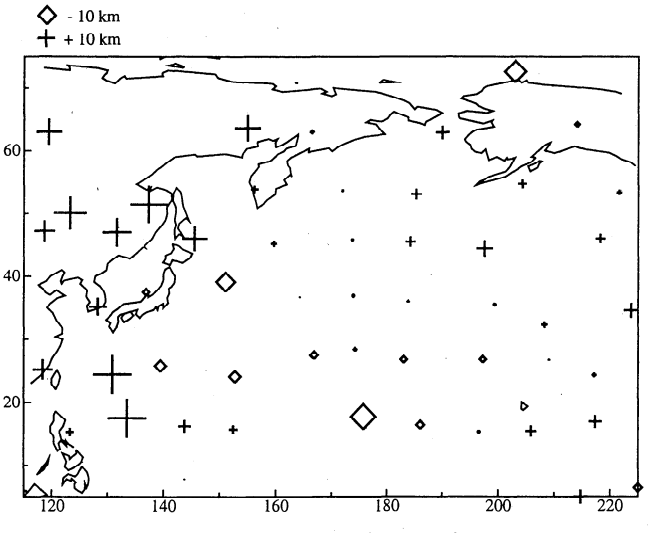

Figure 14. Topography on the 660-km discontinuity measured by stacking SS precursor data within caps of 10° radius as determined by *Flanagan and Shearer* [1998a]. Results have been corrected for biases due to lateral differences in crust and upper mantle velocities. Note the ~20 km depression in the 660-km discontinuity along the left side of this plot (large crosses). This is a major feature in the global 660-km topography (see Figure 1) that appears to be related to subduction along the western Pacific margin.

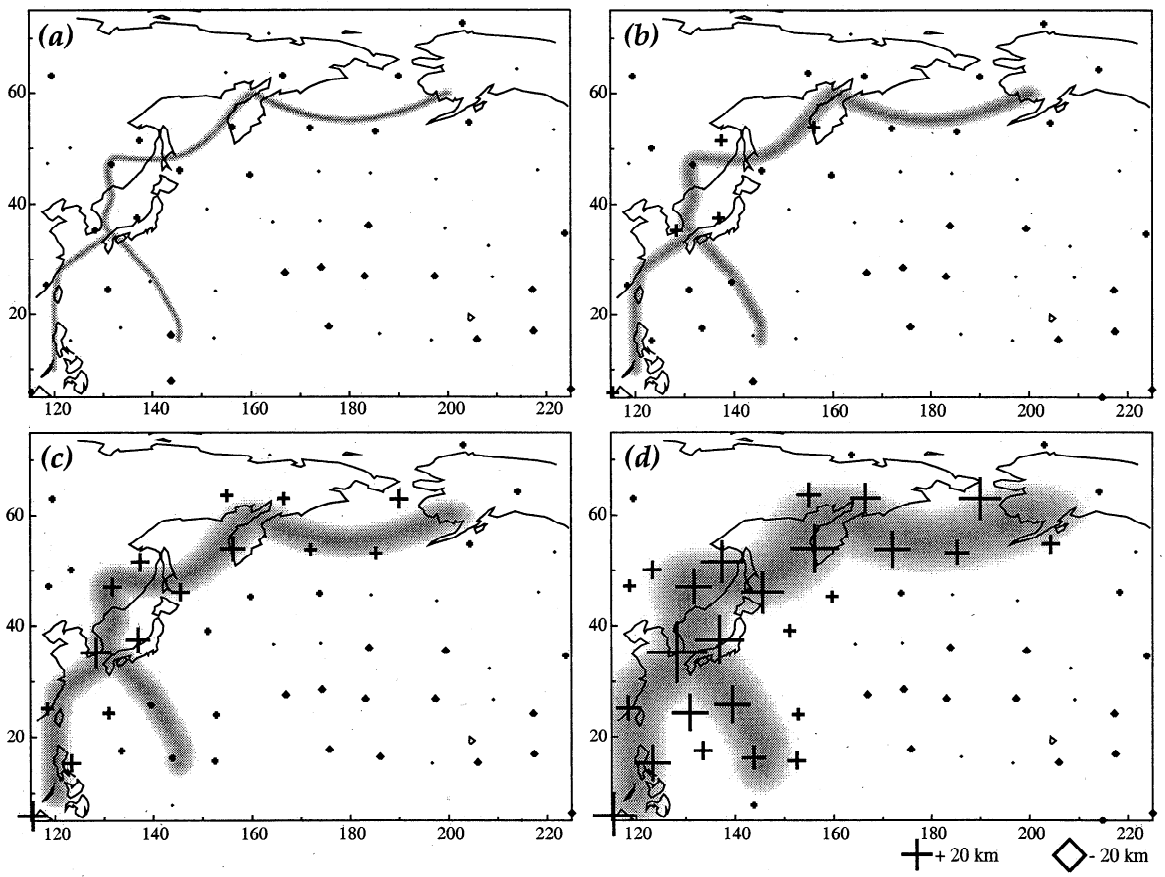

Figure 15. Maps showing models of near-slab troughs in the 660-km discontinuity (shaded regions), compared to discontinuity topography inferred from synthetic seismograms generated for these models (symbols). These troughs are described by a $\cos^2$ function of depth (50 km total depth) and widths of (a) 2°, (b) 4°, (c) 8°, and (d) 16° (1° = 111 km). The discontinuity depth estimates are obtained from synthetic seismograms using the same distribution of sources and receivers and the same analysis method as *Flanagan and Shearer* [1998a]. Note that the inferred discontinuity depths approach the size of the anomalies seen in the real data (see Figure 14) only when the 660-km depressions are of significant width (8° or greater).

plied to discontinuity topography with different scale lengths and also considered the possible biasing effects of high-velocity anomalies below the 660-km discontinuity.

We test the Neele et al. hypothesis by computing synthetic seismograms for a suite of discontinuity topography models, using the source-receiver geometries of the data of *Flanagan and Shearer* [1998a]. We compute the synthetics using a method similar to that described in section 3, except that we fully implement the Kirchhoff method [e.g., *Scott and Helmberger*, 1983; *Kampfmann and Müller*, 1989] for an undulating interface by computing the appropriate obliquity factors and reflection coefficents. We assume velocity and density values from the PREM model at 670 km depth in computing the reflection coefficients. The ~4000 SS bounce points are shown in Figure 1, and the 660 topography obtained through stacking the data in bounce point caps of 10° radius is plotted in Figure 14 (see *Flanagan and Shearer* [1998a] for details concerning the stacking method). Note the ~20 km depression in the 660-km discontinuity along the left side of the plot.

The discontinuity topography models that we use to generate our synthetics have a 660-km discontinuity at constant depth, except for a narrow 50-km-deep valley centered along lines where we estimate that the subducting slabs intersect the discontinuity. This "valley" is defined as a $\cos^2$ function with widths of 2°, 4°, 8°, and 16° (111 km = 1°). For each model we process our 4000 synthetic seismograms in exactly the same way as the data examined by *Flanagan and Shearer* [1998a]. The results are displayed in Figure 15 and show that discontinuity valleys up to 4° wide are not resolvable in the long-period data and do not lead to biases in the regional depth estimates. Larger depressions become progressively more visible in the stacked results. These results are consistent with what is expected based on the 10° to 20° width of the long-period SS Fresnel zones. This conclusion is supported by the observation that the depression in the 660-km discontinuity is not

centered about the slab intersection points, as would be expected if near-slab structure were falsely mapped into the long-wavelength maps.

We conclude that the depression in the 660-km discontinuity imaged in long-period SS precursor studies is not an artifact of small-scale structure near the slabs. The data require the presence of a large regional depression in the 660-km discontinuity that is displaced westward of where subducting slabs intersect the discontinuity. We do not fully understand why our results appear to disagree with those of *Neele et al.* [1997]. They modeled a smaller data set of SS precursors [*Shearer*, 1993] than that of *Flanagan and Shearer* [1998a], but it is unlikely that this is the source of the discrepancy. Other differences include the fact that their modeling was of PdP rather than SdS, and they used synthetic reference pulses with periods significantly shorter than those of the actual SS precursor data.

A direct comparison of our results with the *Chaljub and Tarantola* [1997] study is difficult due to differences in the model geometry. However, a comparison between Figure 15 of this paper and Figure 5 of *Chaljub and Tarantola* [1997] suggests rough agreement between the methods. In our synthetic experiment a 50-km-deep, 8°-wide (i.e., ~900 km wide) slab-induced trough on the 660-km discontinuity appears at long periods to be a wider feature that is about 20 km deep. A synthetic trough of similar width in the *Chaljub and Tarantola* [1997] experiments yields an apparent depth of about 18 km. In both cases, apparent depressions of 20 km or more are only obtained when the trough on the 660-km discontinuity is at least 1000 km across. Chaljub and Tarantola also explored the effect of high-velocity anomalies in the midmantle on inferred discontinuity topography. They showed that an apparent 20 to 25 km depression on the 660-km discontinuity (with a width of about 20°) could result if S velocities were 5% fast in a 1000 km square box below the 660-km interface. In this case, the bias in the inferred topography is a direct result of the large velocity anomaly assumed in the mantle. A midmantle 5% S velocity anomaly with these dimensions is much larger than the velocity perturbations imaged in current tomography models [e.g., *Masters et al.*, 1996; *Grand et al.*, 1997]; it is unlikely that such a feature could be responsible for the observed depression on the 660-km discontinuity seen in the northwest Pacific.

5. Conclusions

Kirchhoff migration techniques are applicable to the analysis of long-period SS precursor data. Synthetic experiments suggest that these techniques might prove very effective; in theory they can resolve features that are substantially smaller than the Fresnel zones. However, for the SS precursor observations that we examine, significant lateral smoothing constraints must be applied to obtain stable results, decreasing the advantages of the technique (which is very computationally intensive) over simple common midpoint (CMP) stacking methods.

Synthetic experiments show that CMP stacking of SS precursors produces discontinuity depth estimates that are not significantly biased by the presence of small-scale topographic features. In particular, the large depression in the 660-km discontinuity observed westward of the subducting slabs in the northwest Pacific is unlikely to be an artifact of near-slab structure, as hypothesized by *Neele et al.* [1997]. Rather it appears to be a significant, well-resolved, regional feature at least 10° across.

Similar methods could be applied to receiver function analyses of P-to-SV conversions from upper mantle discontinuities [e.g., *Shen et al.*, 1996; *Dueker and Sheehan*, 1997; *Li et al.*, 1998]; it is possible that improvements in model resolution could result.

Acknowledgments. The conjugate gradient inversion scheme is based on notes distributed by Jon Claerbout. Justin Revenaugh and Rob van der Hilst provided useful and constructive reviews. This work was supported by National Science Foundation grants EAR93-15060, EAR95-07994, EAR96-14350, and EAR96-28020. M.P. Flanagan is supported by a NSF Postdoctoral Fellowship and the Cecil H. and Ida M. Green Foundation.

References

Castle, J.C., and K.C. Creager, Topography of the 660-km seismic discontinuity: Implications for tectonic history and slab deformation, *J. Geophys. Res.*, *103*, 12,511–12,527, 1998.

Chaljub, E., and A. Tarantola, Sensitivity of SS precursors to topography on the upper-mantle 660-km discontinuity, *Geophys. Res. Lett.*, *24*, 2613–2616, 1997.

Choy, G.L., and P.G. Richards, Pulse distortion and Hilbert transformation in multiply reflected and refracted body waves, *Bull. Seismol. Soc. Am.*, *65*, 55–70, 1975.

Dueker, K.G., and A.F. Sheehan, Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hotspot track, *J. Geophys. Res.*, *102*, 8313–8327, 1997.

Dziewonski, A.M., and D.L. Anderson, Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, *25*, 297–356, 1981.

Flanagan, M.P., and P.M. Shearer, Global mapping of topography on transition velocity discontinuities by stacking SS precursors, *J. Geophys. Res.*, *103*, 2673–2692, 1998a.

Flanagan, M.P., and P.M. Shearer, Topography on the 410-km seismic velocity discontinuity near subduction zones from stacking of sS, sP, and pP precursors, *J. Geophys. Res.*, *103*, 21,165–21,182, 1998b.

Gossler, J., and R. Kind, Seismic evidence for very deep roots of continents, *Earth Planet. Sci. Lett.*, *138*, 1–13, 1996.

Grand, S.P., R. van der Hilst, and S. Widiyantoro, A snapshot of convection in the Earth, *GSA Today*, *7*, 1–7, 1997.

Gu, Y., A.M. Dziewonski, and C.B. Agee, Global de-correlation of the topography of transition zone discontinuities, *Earth Planet. Sci. Lett.*, *157*, 57–67, 1998.

Kampfmann, W., and G. Müller, PcP amplitude calcula-

tions for a core-mantle boundary with topography, *Geophys. Res. Lett.*, *16*, 653–656, 1989.

Kaneshima, S., and G. Helffrich, Detection of lower mantle scatterers northeast of the Marianna subduction zone using short-period array data, *J. Geophys. Res.*, *103*, 4825–4838, 1998.

Kennett, B.L.N. (Ed.), *IASPEI 1991 Seismological Tables*, Res. Sch. of Earth Sci., Aust. Natl. Univ., Canberra, 1991.

Li, A., K.M. Fischer, M.E. Wysession, and T.J. Clarke, Mantle discontinuities and temperature under the North American continental keel, *Nature*, *395*, 119–120, 1998.

Masters, G., S. Johnson, G. Laske, and H. Bolton, A shear-velocity model of the mantle, *Philos. Trans. R. Soc. London*, *354*, 1385–1411, 1996.

Neele, F., H. de Regt, and J. VanDecar, Gross errors in upper-mantle discontinuity topography from underside reflection data, *Geophys. J. Int.*, *129*, 194–204, 1997.

Niu, F., and H. Kawakatsu, Depth variation of the mid-mantle seismic discontinuity, *Geophys. Res. Lett.*, *24*, 429–432, 1997.

Revenaugh, J., and T.H. Jordan, A study of mantle layering beneath the western Pacific, *J. Geophys. Res.*, *94*, 5787–5813, 1989.

Revenaugh, J., and T.H. Jordan, Mantle layering from *ScS* reverberations, 2, The transition zone, *J. Geophys. Res.*, *96*, 19,763–19,780, 1991.

Scott, P., and D. Helmberger, Applications of the Kirchhoff-Helmholtz integral to problems in seismology, *Geophys. J. R. Astron. Soc.*, *72*, 237–254, 1983.

Shearer, P.M., Seismic imaging of upper-mantle structure with new evidence for a 520-km discontinuity, *Nature*, *344*, 121–126, 1990.

Shearer, P.M., Constraints on upper mantle discontinuities from observations of long-period reflected and converted phases, *J. Geophys. Res.*, *96*, 18,147–18,182, 1991.

Shearer, P.M., Global mapping of upper mantle reflectors from long-period *SS* precursors, *Geophys. J. Int.*, *115*, 878–904, 1993.

Shearer, P.M., and G. Masters, Global mapping of topography on the 660-km discontinuity, *Nature*, *355*, 791–796, 1992.

Shen, Y., S.C. Solomon, I.T. Bjarnason, and G.M. Purdy, Hot mantle transition zone beneath Iceland and the adjacent Mid-Atlantic Ridge inferred from *P*-to-*S* conversions and the 410- and 660-km discontinuities, *Geophys. Res. Lett.*, *23*, 3527–3530, 1996.

Vinnik, L.P., Detection of waves converted from *P* to *SV* in the mantle, *Phys. Earth Planet. Inter.*, *15*, 39–45, 1977.

Wicks, C.W., and M.A. Richards, A detailed map of the 660-kilometer discontinuity beneath the Izu-Bonin subduction zone, *Science*, *261*, 1424–1427, 1993.

M.P. Flanagan, M.A.H. Hedlin, and P.M. Shearer, IGPP 0225, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093-0225 (pshearer@ucsd.edu).

(Received May 20, 1998; revised December 8, 1998; accepted December 16, 1998.)