Bulletin of the Seismological Society of America, 90, 4, pp. 964–976, August 2000

Basin Structure Estimation by Waveform Modeling: Forward and Inverse Methods

by Chen Ji, Don V. Helmberger, and David J. Wald

Abstract We introduce a technique for using broadband seismograms recorded from earthquakes at local and regional distances to refine basin structure. For the region outside the basin, we assume a one-dimensional (1D) crustal model and analytical techniques (GRT) to propagate the energy from sources to the basin edge where the motions are then interfaced with a (2D) finite-difference algorithm (Wen and Helmberger, 1996). We parameterize the basin section by isovelocity layers with linear dipping segments between control points. The control point depths are allowed to vary to improve the modeling of waveform data of stations inside the basin. The comparison between data and synthetics is qualified by a fitness function defined by two factors; the timing shift required for best alignment and the correlation coefficient. The procedure was applied to a strong-motion waveform profile across the extended Los Angeles Basin produced by the 1992 Landers, California earthquake to refine the velocity structure using sensitivity testing and forward modeling. Only the correlation coefficient and amplitude were used because absolute timing was unknown. The procedure was extended to a direct waveform inversion by employing a conjugate gradient approach, which uses numerical derivatives. Numerical tests using the new inversion process with synthetic data demonstrate that it is possible to recover a detailed basin structure, if a sufficient amount of high-quality data exists.

Introduction

Recent earthquakes have established the damaging effects of strong shaking in basin environments (e.g., Esteva, 1988; Kawase, 1996). Unfortunately, strong-motion observations in basins are severely limited and those that do exist are not from great earthquakes. Predictions based on three-dimensional (3D) simulations have been suggested as substitutes (e.g., Olsen *et al.,* 1995) in which they suggest relatively large amplifications within Los Angeles basin. This benchmark effort has triggered intense interest in refining the basin structure and understanding such amplifications.

New models are being developed presently by the Southern California Research Center (SCEC) based on accumulating geological and geophysical data. Three models of the extended 3D Los Angeles basin structure have been recently tested against data. The early SCEC model (Magistrale *et al.,* 1996) maps geological structure into seismic velocities by assigning empirically-based vertical velocity gradients. The Graves model (Wald and Graves, 1998) is based on the same basement depth contours (McCulloh, 1960) but uses velocity profiles based on waveform-modeling earthquake ground-motion data (Scrivner and Helmberger, 1994). Another useful approach is to use *P*-wave tomography (Hauksson and Haase, 1997). Predictions based on these three models and the source description of the Landers Earthquake (Wald and Heaton, 1994) were compared against observations in Wald and Graves (1998). This study demonstrated the potential for understanding wave propagation in a complex environment. Some of these models predict the amplification and waveform complexities very well, e.g., the predictions of the Graves model for the San Fernando Valley. However, none of the models do well for the deep sediment sites within the Los Angeles Basin. Adjusting these basin models to improve their fits to data is very difficult because 3D simulations are time-consuming, and many parameters are involved.

It appears that the various broadband systems, such as TriNet (Mori *et al.,* 1998), with many basin stations, will be very useful for testing existing models and generating new ones. For example, Figure 1 shows four records from the 95/09/20 Ridgecrest event ($M_W = 5.4$), along with a topographic map of southern California. The dashed contours denote the thickness of basin sediments (McCulloh, 1960). Note that station CRN is just outside Los Angeles basin, while stations FUL, OGC, and SAN are within. The epicentral distance of station CRN is very close to that of station FUL, but the seismic records are quite different in waveform and duration (Fig. 1b). In addition, the three basin stations have

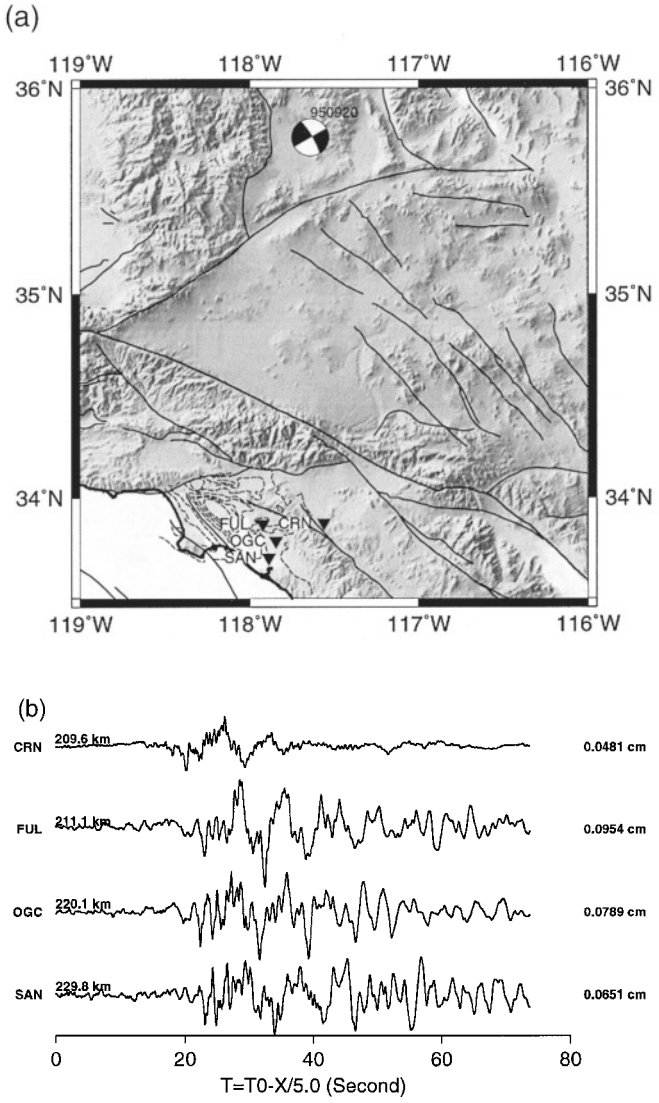


Figure 1. (a) Topography of southern California including Los Angeles basin. The dashed contour lines display the isosurface of basement rocks (McCulloh, 1960), and the black triangles show the positions of some new seismic stations. The source position and mechanism of a Ridgecrest earthquake are indicated by the focal mechanism. (b) The tangential motions as recorded at the three basin stations, FUL, OGC, and SAN along with that of a relatively hard-rock station CRN. The peak amplitudes (cm) are indicated on the right. The epicentral distances (km) are given on the left.

much larger amplitudes than the hard-rock site (CRN) as expected.

The complexity seen at these stations is also present in the other events of Ridgecrest sequence (Scrivner and Helmberger, 1999). After examining the entire TERRAscope array, they reported on considerable long-period multipath effects such that the source radiation nodes disappear. This feature can modeled to some extent by removing nodes in the radiation patterns, but amplitude levels still remain uncertain by nearly a factor of two, even for stations outside the basins.

This feature emphasizes the 3D nature of the problem and makes it difficult to isolate the basins response without a nearby hard-rock observation for direct comparison. Thus, to make effective use of weak-motion records from distant events such as in Figure 1 requires methodologies that separate the various complexities into individual operators (e.g., Wen and Helmberger, 1996). Following this approach, we break down the synthetic seismogram, sn_i (t), at station (i) into: source excitation, e(t), with various assumptions about rupture properties and depth; regional path effects, p(t), including seismic propagation from the source to the station.

$$e(t) * p_i(t) = sn_i(t) \tag{1}$$

The purpose of this article is to demonstrate the usefulness of this approach in explaining the observations from the Landers earthquake with a modified Graves model, and to introduce a waveform-inversion technique to retrieve local basin structures from weak motion of the type displayed in Figure 1.

Landers Earthquake Data and Modeling Analysis

The Landers Earthquake shook the entire Los Angeles area and produced useful data on strong-motion accelerographs. Figure 2 shows the location of six stations that cross the extended LA basin region recording this event. The tangential components of ground-motion displacements (Fig. 3) are obtained from the original acceleration data; they were bandpass filtered between 2 and 20 seconds, then changed to displacement data by double integration. Because the instruments were triggered, the initial motions are missing and, unfortunately, only XPO and IGU had absolute timing. Thus, the data has been shifted to align with synthetic seismograms that are plotted as a reduced section. The basin effects in these data are obvious. For example, the peak displacement at station DOW (in the center of LA basin, epicentral distance is 164 km) was nearly 4 times as large as that of station XPO (outside LA basin, epicentral distance is 123 km). We attempt to explain these observed amplification by the aforementioned approach. We choose a velocity model shown in Figure 2, which can be separated into two portions, outside the basin (1D) and inside the basin (2D). We use the SoCal model (Dreger and Helmberger, 1991) for the outside portion (Table 1). Inside, we used a 2D section (Fig. 2c), modified from Graves model (Wald and Graves, 1998). Following the approach of Wen and Helmberger (1996), we use station XPO as reference station and apply the hybrid code to calculate the transfer functions of basin stations relative to it, then predict the records of stations inside the basin by convoluting the recording at station XPO with the transfer function. More precisely, we define a transfer function, T(t), as

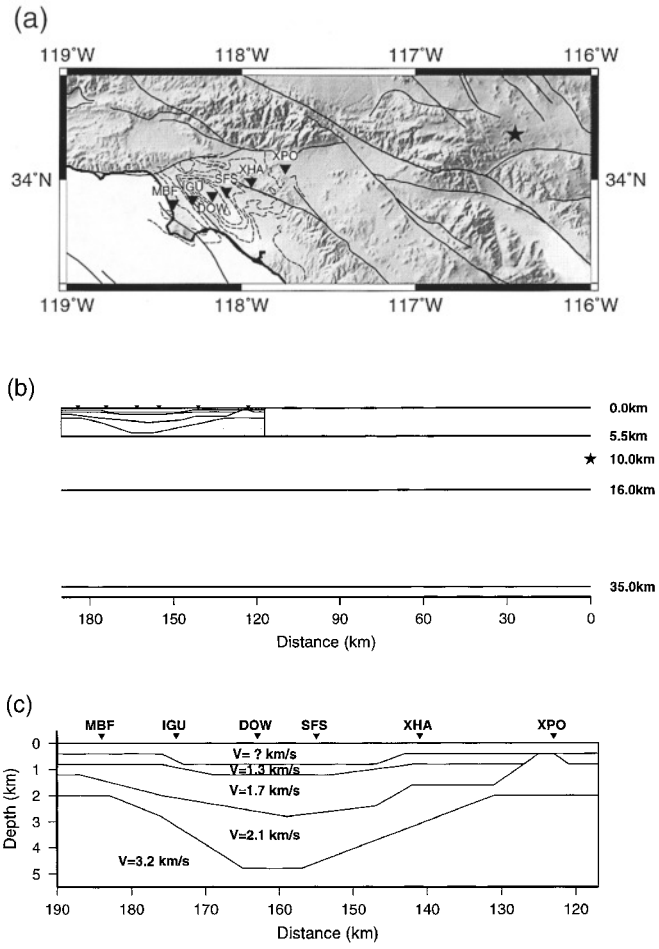


Figure 2. A strong-motion profile of the Landers earthquake. (a) Contour lines display the isosurface of basement rocks. The triangles show the positions of strong-motion stations, and the star shows the position of epicenter. (b) The shear-wave velocity structure along the profile. The shaded box shows the structure of LA basin. For the region outside basin, we assume the SoCal model (Dreger and Helmberger, 1991). (c) Enlargement of the LA basin model, modified from a cross section of the Graves model (Wald and Graves, 1998).

$$T_i(w) = p_i(w)/P_r(w) \quad (2)$$

where P_r (t) is a reference station. Thus, the predicted response at station i becomes sn_i (t) = T_i (t) * obs_r (t) where obs_r (t) is the reference observation, XPO in this situation. In many applications, researchers assume i to be a soft-rock site and r to be a hard-rock site (e.g., Borcherdt *et al.,* 1975).

Note that this approach greatly reduces the effort needed to produce the seismic field arriving at the basin edge. If we were attempting to solve the entire 3D problem, we would need to generate the response at XPO from a distribution of elements characterizing the Landers earthquake as performed by Wald and Graves (1998). Based on the Scrivner and Helmberger (1999) study, we probably cannot expect to predict XPO to within a factor of two at these distances. This situation makes it particularly difficult to isolate the true basin response from other complexities. The transfer-function approach simply assumes that everything we do not understand about the propagational path across the San Bernardino Mountains and source complexity is contained in the data at the reference site and the problem is reduced to 2D geometry. As discussed in Wen and Helmberger (1996), this assumption must be justified in each situation by examining the other components of motion to ensure that the *SH* motions are predominant throughout the profile, arriving on a Great-Circle, and so forth.

Forward Modeling

The basic idea is to start with an approximate model and make reasonable changes to the assumed parameterization, until it can explain the records inside the basin. This work can be done by the conventional trial and error method using the finite-difference method (e.g., Scrivner and Helmberger, 1994). As in this effort, we define the basin model by several two-dimensional (2D) layers (Fig. 2). The shapes of interfaces are determined by some control points that are smoothly connected by an interpolation function. Hence, we can obtain different basin models by changing the control point depths and layer velocities. The advantage of this procedure is that the number of free parameters is reduced, which greatly restricts the model parameter search.

The starting basin model originated from studies of the Whittier and Sierra Madre earthquake sequence by Scrivner and Helmberger (1994), with further modifications by Graves (1996). We made additional adjustments (using forward modeling) to arrive at the model displayed in the lower panel of Figure 2. A comparison of synthetics generated from this model against observations is given in Figure 3.

Because we do not know the trigger time of most stations, only waveforms and peak amplitudes were studied. The waveform fits are fairly good, but the peak amplitudes in the middle of the basin are underestimated by an average 25% on the left, assuming the relatively high basin velocities used by Graves (1996). However, the relative variation of peak amplitudes along the profile is similar to that of data. Note our result is similar to 3D simulation (Wald and Graves, 1998), so it is not an artifact caused by transfer function approach. Improved results are obtained by simply reducing the velocity of the top sedimentary layer. After several forward tests, we found that a value of 0.6 km/sec could explain the peak amplitude variation very well with the exception of the station MBF. Moreover, this modification also improved the waveform fit to the station XHA significantly. Thus, the amplitude is quite sensitive to the shallow velocity structure, even for energy at such long periods.

Sensitivity Testing

While this model does a reasonable job of matching the data, most researchers want to know more about the physical relationship between the model complexity and observed

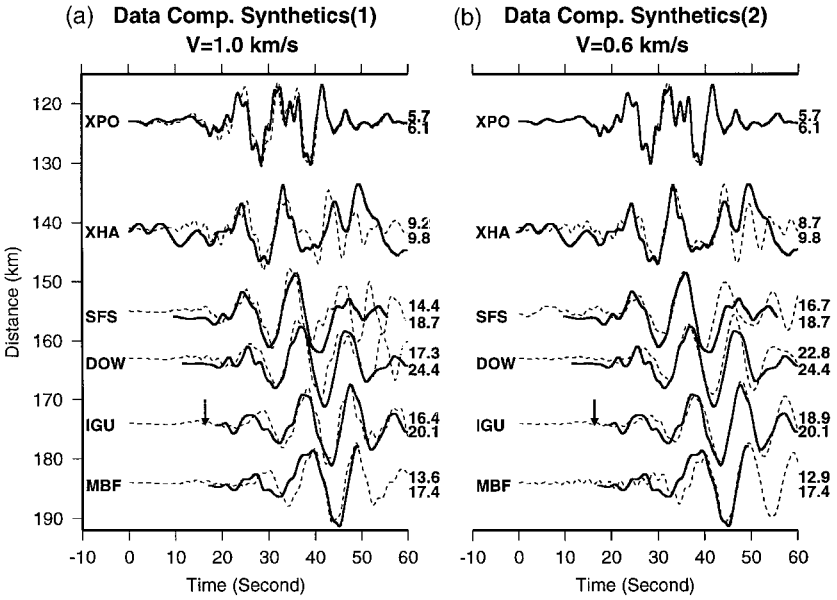


Figure 3. Comparison between the Landers earthquake data (thick lines) and synthetic waveforms (dashed lines). The velocity of the near-surface layer in the basin model (Fig. 2c) is 1 km/sec in case (a) and is 0.6 km/sec in case (b). The numbers behind each trace are peak amplitudes (cm) of data (bottom) and synthetic motions (top), respectively. The synthetic seismograms are plotted to reduced section with a reduction velocity 4.5 km/sec. The data are shifted to fit the synthetic best as discussed in the text. Note that reducing the velocity of the top layer improves the fit of the peak-amplitude variation and waveform of station XHA.

Table 1
Southern California Crustal Model (SoCal)

Depth (km)	V (km/sec)	p (g/cm^2)
5.5	3.18	2.4
16.0	3.64	2.67
35.0	3.87	2.80
∞	4.50	3.00

Table 2
Peak Displacements of Synthetic Seismograms

Models	XPO (cm)	XHA (cm)	SFS (cm)	DOW (cm)	IGU (cm)	MBF (cm)
Model-A	5.7	6.2	6.6	7.3	7.5	7.5
Model-B	5.7	6.0	10.0	9.3	7.6	7.0
Model-C	5.7	6.9	13.2	14.6	14.3	11.5
Model-D	5.7	8.1	15.2	18.9	17.6	13.0
Model-O	5.7	8.7	16.9	22.6	18.8	12.8

motions. This is difficult to display but can be approached by performing sensitivity testing as in Figure 4.

To proceed, we introduce a goodness-of-fit criterion. We choose an objective function (measure) that includes two parts; waveform fit and travel-time constraint. The degree to which the synthetic and observed seismograms match each other is estimated by the function,

$$f(\tau) = 1.0 - 2\,\frac{\int p(t)_{\text{obs}} p(t+\tau)_{\text{syn}}\,dt}{\int (p(t)^2_{\text{obs}} + p(t+\tau)^2_{\text{syn}})\,dt} \qquad (3)$$

Here $p(t)_{\text{obs}}$ and $p(t)_{\text{syn}}$ are the observed and synthetic seismograms, respectively, and τ is the time shift between the two. If at shift time τ, the synthetic seismogram fits the observed date perfectly, $f(\tau) = 0$. This particular form of goodness-of-fit is discussed by Sen and Stoffa (1991), they discuss the advantages of this form over the normalized cross correlation.

We seek a τ that shifts a synthetic seismogram so that it "best" matches the observed seismogram. The criterion for "best" match is defined as the travel-time residual $\Delta\tau$ that minimizes the function $f(\tau)$, i.e.

$$f(\delta\tau) = \min\,\{f(\tau) | \tau \in [-T, T]\} \qquad (4)$$

Here T is the estimated maximum travel-time difference between the observed and calculated seismograms. In our analysis, $\delta\tau$ can be seen as a travel-time residual of the surface wave since it has the strongest signal. While $f(\tau)$ provides a useful measure of shape matching, it does not strongly constrain the most important parameter, that is, the absolute amplitude. Thus, the third essential parameter is the peak amplitude ratio of synthetics to data at $\Delta\tau$.

Since the data lack timing, we will use the best fitting synthetic seismograms, on the right of Figure 3, so that we can use this fitness criterion. Thus, we can gain a better appreciation of forward modeling as well as gain insight in the use of our chosen objective function formulas by computing the corresponding fitness factors τ, $f(\tau)$, and amplitude ratio. We chose four models (Fig. 4), starting with a simple flat layer structure and adding complexity at each step.

The comparison between synthetic seismograms (thin lines) and reference data (thick lines) are shown in Figure 4. We shift synthetic seismograms to best fit the reference data and normalize them by the peak amplitude of reference data. We use three numbers to identify the values of $f(\tau)$, the peak-amplitude ratio between synthetic seismogram and datum, and travel-time residual τ. These three numbers control the modeling process and provide a good indication of how well the synthetic seismograms match the data. Note

that when $f(\tau) = 0$, a perfect fit is achieved which always occurs for the record (XPO) by the definition of transfer-function approach. The poorest fit occurs at station (IGU) for model (a) ($f(\tau) = 0.64$), which drops to 0.46 in (b) and 0.09 in (c), and 0.01 in (d). Thus, by introducing the deeper structure, we largely recover the long-period waveform characteristics. This is also the case in travel-time residual, τ, which goes from -2.76, to -2.76, to -0.88 to -0.02 at the same station. The peak amplitudes show a similar behavior where the differences between (c) and (d) are small. Note that the model in case (d) is that of Figure 2 except with the layer containing the velocity of 1.3 km/sec removed. Although the overall fits are nearly the same as the original model, the recordings are all earlier and have up to a 16% increase in amplitude (DOW).

Generally, the shapes of interfaces all affect the seismic data, but the results are different, i.e., the thin surface layer has a large effect in peak amplitude but negligible effect on timing, whereas the basin edge has a dominant effect on timing as well as the peak amplitudes; the intermediate layers have a strong effect on timing too but change the amplitude only slightly. The two stations, SFS and DOW, in the middle of the basin show the most sensitivity to the intermediate structures in all three measures.

Although the basin interface shapes are proving to be important factors in the basin amplification effect, it is not well determined, e.g., the Graves model uses the basement depth contours obtained by gravity inversion (McCulloh, 1960). Regional and local seismic data is turned out to be a good measurement in determining the basin shapes.

Full Waveform Inversion Method

Anticipating the future availability of numerous broadband data within the basins, we address the issue of model estimation in the form of the inverse problem. We represent the 2D basin model by a vector $\mathbf{x}$. The components of this vector are the parameters of the model, in this case, the velocities and positions of interface control points. From this model, we can calculate synthetic seismograms. We also define a particular function named the objective function, $\mathbf{E}(\mathbf{x})$, which measures the difference between the data and synthetic seismograms, using a combination of τ and the $f(\tau)$ described earlier. The objective function becomes:

$$E(x) = W_t * \sqrt{\frac{1}{N-1}\sum_{i=2}^{N}(\delta\tau_i - \delta\tau_l)^2} + W_f * \sqrt{\frac{1}{N}\sum_{i=1}^{N} f_i^2\,(\delta\tau_i)} \quad (5)$$

Here, W_t and W_f are weights for travel-time residuals and waveform fit, respectively, which are chosen based on data quality. N is the number of stations, and the first station is assumed to be just outside the basin. This form has been suggested by Luo and Schuster (1991), in which they point out that waveform inversion has high resolution, but it will fail if the model is too far away from the best solution. Adding travel-time information will increase the inversion robustness greatly. Note that it has a minimum if the discrepancies between data and synthetics are the smallest. We employ the conjugate gradient algorithm to find this minimum (Polak, 1971).

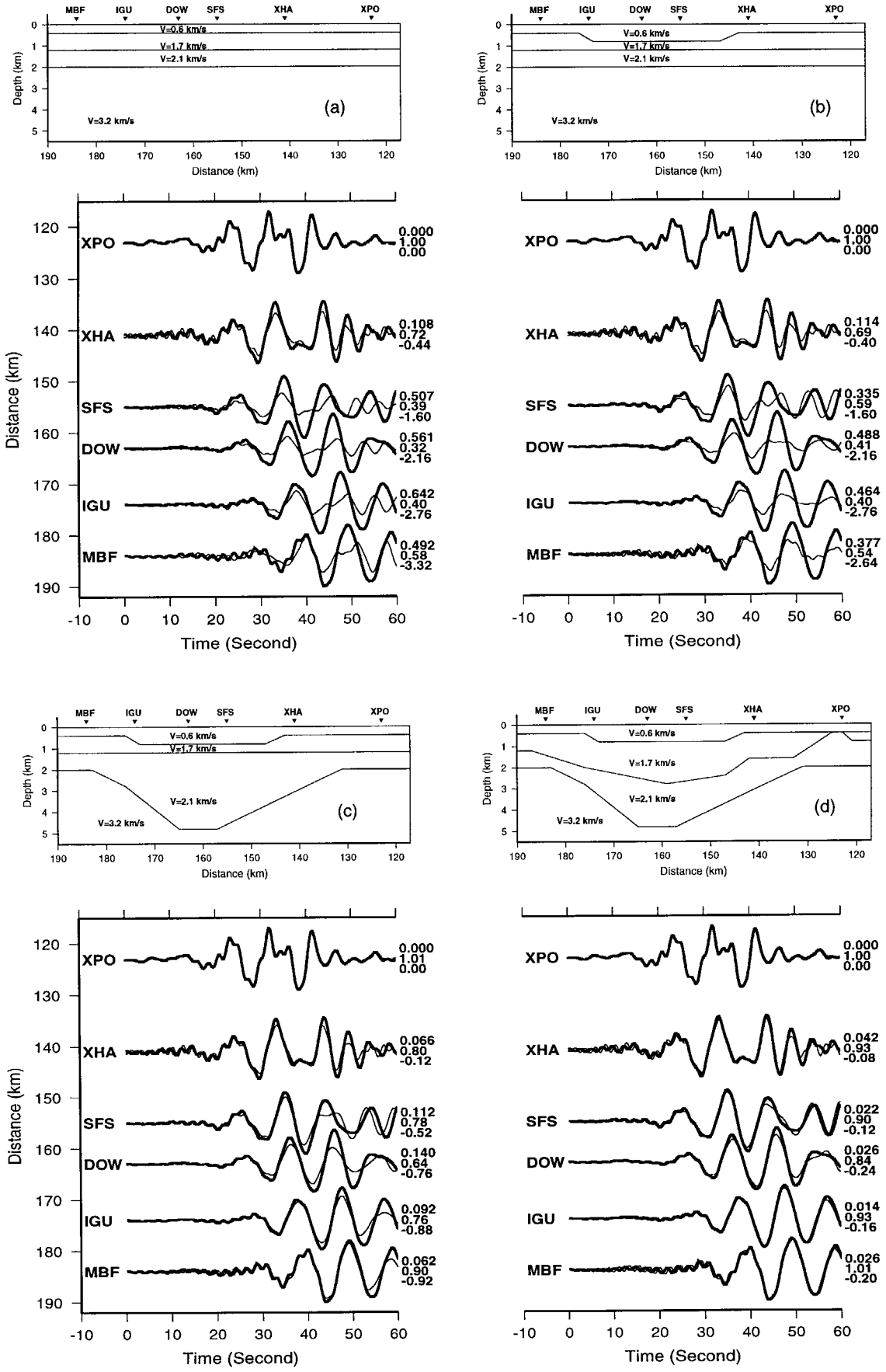
Suppose we have a starting model $\mathbf{x}_0$. The first step of a conjugate gradient algorithm is to calculate the steepest descend direction of objective function, $\mathbf{g}_0 = -\Delta E(\mathbf{x}_0)$. In addition, we define a conjugate direction $\mathbf{h}$. Starting with $\mathbf{h}_0 = \mathbf{g}_0$, we change the model along the direction $\mathbf{h}_0$ to find a local minimum of $E(\mathbf{x}_1)$. Then we perform the second iteration starting from model $\mathbf{x}_1$. In general, for j-th iteration ($j > 0$), the vector

$$\begin{aligned} g_j &= -\Delta E(x_j), \\ h_j &= g_j + \gamma_{j-1} h_{j-1}, \\ \gamma_{j-1} &= -\frac{(g_j - g_{j-1})\cdot g_{j-1}}{g_{j-1}\cdot g_{j-1}}. \end{aligned} \quad (6)$$

The algorithm will iterate until it determines an answer with an acceptable error or the default maximum iteration is reached.

Since it is very hard to obtain an analytic expression for function $E(\mathbf{x})$, the partial derivatives are calculated numerically, that is, two synthetic seismograms are generated; one for model $X_1 = X\,(x^1, x^2 \ldots x^k - \Delta x \ldots x^N)$ and the other for model $X_2 = X\,(x^1, x^2 \ldots x^k + \Delta x \ldots x^N)$. Then the partial derivative is: $\frac{\partial E}{\partial X^k} = \frac{1}{2\Delta X}\,(E(X_2) - E(X_1))$. The step

Figure 4. Results of sensitivity tests. We forward simulate the long-period responses of some basin models, which begins with a simple flat parallel layered model (a), then from (b) to (d), we gradually increase the complexity. Model (d) is similar to the model shown in Figure 2 except we remove an intermediate layer with velocity 1.3 km/sec. Synthetic seismograms shown in Figure 3b are used as reference data. The comparisons between the data (thick lines) and the synthetic waveforms (thin lines) generated from the models are shown below each model. Note the synthetic seismograms are normalized by the test data and are shifted in time in order to obtain the best fit of waveforms. The three numbers behind each trace are $f(\tau)$ (top), the ratio of peak amplitude between the synthetic motion, and the data (middle), and travel-time residual τ (bottom). Note the change from model (b) to model (c) corresponds to the largest variation in $f(\tau)$ and τ. ⟶


MBF
IGU
DOW
SFS
XHA
XPO
V=0.6 km/s
V=1.7 km/s
V=2.1 km/s
V=3.2 km/s
(a)
(b)
(c)
(d)
Depth (km)
Distance (km)
Time (Second)
0.000 1.00 0.00
0.108 0.72 -0.44
0.507 0.39 -1.60
0.561 0.32 -2.16
0.642 0.40 -2.76
0.492 0.58 -3.32
0.114 0.69 -0.40
0.335 0.59 -1.60
0.488 0.41 -2.16
0.464 0.40 -2.76
0.377 0.54 -2.64
0.000 1.01 0.00
0.066 0.80 -0.12
0.112 0.78 -0.52
0.140 0.64 -0.76
0.092 0.76 -0.88
0.062 0.90 -0.92
0.042 0.93 -0.08
0.022 0.90 -0.12
0.026 0.84 -0.24
0.014 0.93 -0.16
0.026 1.01 -0.20

Δx is chosen as one grid space of the finite difference. Furthermore, in order to make the inversion robust, we add an additional constraint to avoid the intersection of adjacent interfaces.

$$\text{if } h_j^k < 0 \text{ and } x_j^k - x_{\min}^k < \mathrm{D}x,\ h_j^k = 0$$
$$\text{if } h_{\mathrm{j}}^{\mathrm{k}} > 0 \text{ and } x_{max}^k - x_j^k < \mathrm{D}x,\ h_j^k = 0 \qquad (7)$$

Here, Δx is the grid space of the finite difference and h_j^k is the k-th component of conjugate direction in the j-th iteration.

Numerical Simulation and Testing

To investigate the resolution and uniqueness of the technique, we performed a series of inversions using synthetic data. During these tests, a strike-slip source at a depth of 10 km is chosen to generate synthetics. The magnitude is 4.4 and source-time function is a trapezoid with 0.9 sec width (0.4, 0.1, 0.4). We choose the epicentral distance of basin's left boundary as 155 km, the width of finite-difference region to be 20 km, and 10 stations along the profile (Fig. 5). The high-cut frequency of the finite difference is 1 Hz, and we use a grid space 0.06 km to guarantee the accurate of surface wave. The basin structure includes two types of unknown parameters: the velocities of sediments and shapes of interfaces between sediments. Our present effort is focused on recovering the shapes of interfaces, and we find that limiting the maximum number of iterations to five yields reasonable test results.

The first test performed involves a single dipping layer (black line in Figure 5b). We start the inversion from an arbitrary horizontal layer model (pink line in Figure 5b). In Figure 5c, we compare the test data (black lines) and the synthetic seismograms (pink lines) calculated from the initial model. The discrepancies are quite large. We use four control points (stars) to determine the shape of this interface and apply the inversion method developed to search for the right positions of control points. Equation (6) is used to construct the objective function, and we give the same weights to both waveform fit (W_{t}) and the travel-time residual measurement (W_{f}). After 5 iterations, we obtained the inverted model indicated by the dashed line in the Figure 5b. The recovered basin structure is quite close to the test model and the waveform fits are greatly improved (Fig. 5d). Further iterations will improve the inverted basin model. Thus, the inversion of waveform data can yield strong constraints for a simplified one-layer dipping model.

High-frequency noise is a common difficulty encountered in typical travel-time inversion problems because it makes it hard to pick the arrival-time. Numerical test shows that our method is stable to given noisy data. We added 20% Gaussian noise to the test data shown in Figure 6 (black lines, the largest amplitude of noise is 20% of that of data and performed the same inversion. The model recovered is indicated by dashed line in Figure 6b, and it also fits the test model well (black line).

In the third numerical simulation, we test the ability of this method for recovering deeper structure. The basin structure is shown in Figure 7b. In this test, we assume that the top layer is known and try to determine the second interface. In Figure 7c, we compare the test data with the synthetic seismograms from the initial model; the differences are still quite large. After five iterations, the discrepancy between test model and inverted model (Fig. 6b) is larger than that in Figure 5b, the maximum difference is still less than 0.5 km, one-fourth of wavelength of 1-Hz shear wave. For this case, the synthetic seismograms fit the test data very well except for some small high-frequency differences (Fig. 7d), which indicates that it would be hard to improve the model with additional iterations.

In the general situation, we may not have such a dense station distribution. Thus, in the fourth test, we reinverted the test data with five observations (Fig. 8). The results are similar to those in Figure 5. However, in another test with only two seismograms, the program stopped at a local minimum of objective function with no useful solution. Thus, recovery detailed structure requires dense station distribution. Fortunately, one of the advantages of the transfer-function approach is that multiple events can be used in the modeling process, if the same station just outside the basin records all the events. Moreover, a mixture of mechanisms may actually help since the waveguide is more fully excited and the parameter space is better represented.

Discussion

In the previous analysis, we used synthetic data generated from an intermediate-sized earthquake. Hence, the waveforms contain energy at relatively high frequency, which are sensitive to the basin structure, as indicated in Figures 5b and 7b. In contrast, when we study the long-period strong-motion data from large events, in our example, the Landers earthquake, travel-time information becomes the more important constraint. We illustrate this effect by performing another test that is displayed in Figure 9.

In Figure 9a, we compare the test model and the initial model (dashed) where we investigate the possibility of constraining the basin edge. We choose synthetic data because the most real data do not have absolute timing. Like earlier tests, we use four control points (stars) to determine the basin shape. The comparison between test data (thick lines) and synthetic motions (thin lines) generated from initial model is shown in the Figure 9c. In order to highlight the small variation in waveform, we shift the synthetic waveforms to best fit the test data, and use a travel-time residual τ to identify the time shift. Behind each trace, we also indicated the value of $f(\tau)$ and the ratio of the peak amplitude between the synthetic motions and the test data. In general, the shifted synthetic seismograms fit the data very well, so the value of $f(\tau)$ is very small and the root mean square (rms) is only

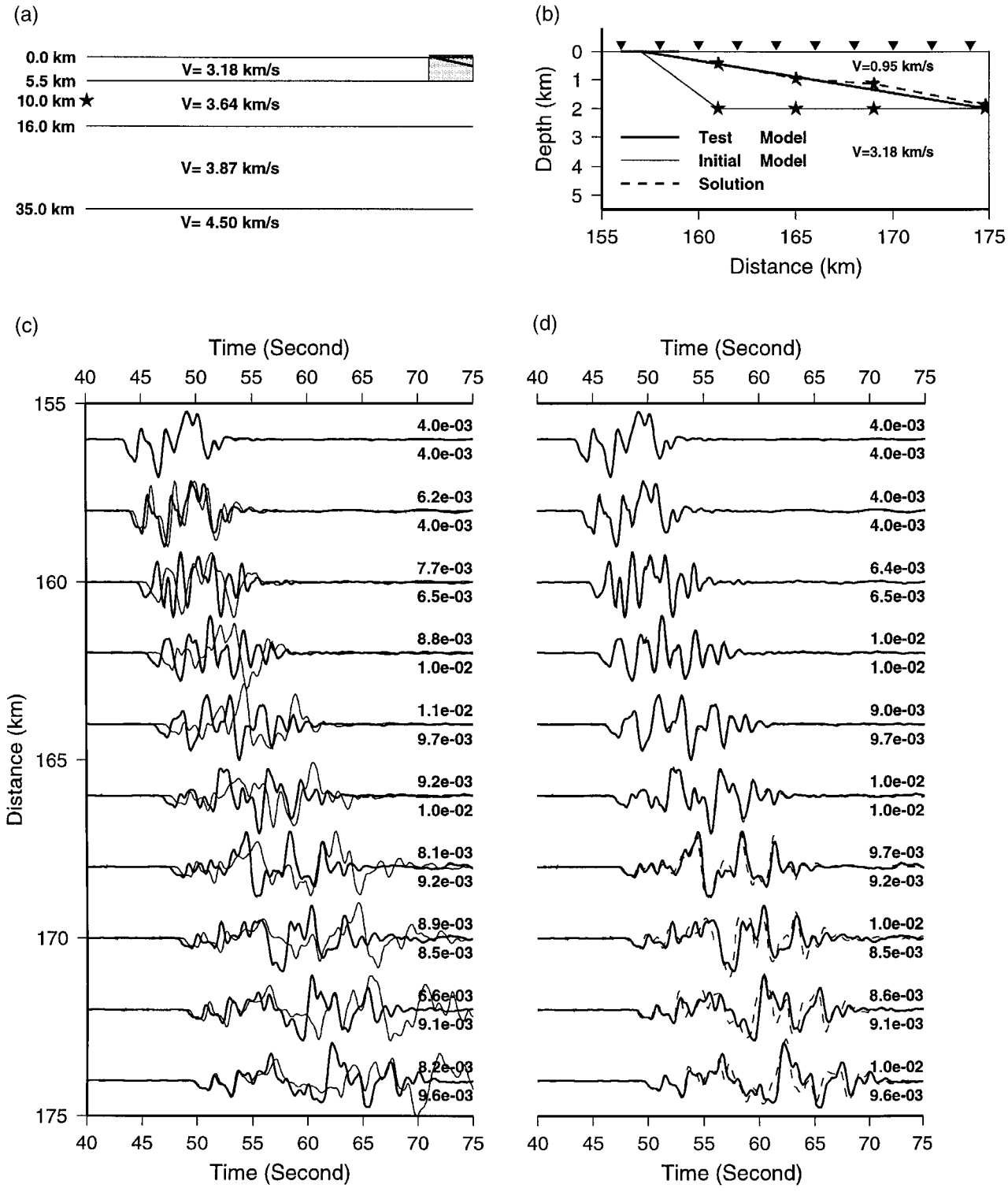


Figure 5. Inversion results of single dipping layer inversion. (a) The whole model that we used in the test; the star indicates the position of source. The shaded box indicates the basin region. (b) Sketches of Basin models in the study. The shape of the basin interface is indicated by the thick line (test model), thin line (initial model), and dashed line (inverse solution). (c) Comparison between the test data (thick lines, generated from test model) and the synthetic seismograms (thin lines, generated from initial model), the peak amplitudes (cm) are given at the end of each trace, upper (synthetic seismograms), lower (test data). (d) Comparison between the test data and the synthetic seismograms (thin lines, generated from inversion model), the peak amplitudes (cm) are given at the end of each trace, upper (synthetic seismograms), lower (test data).

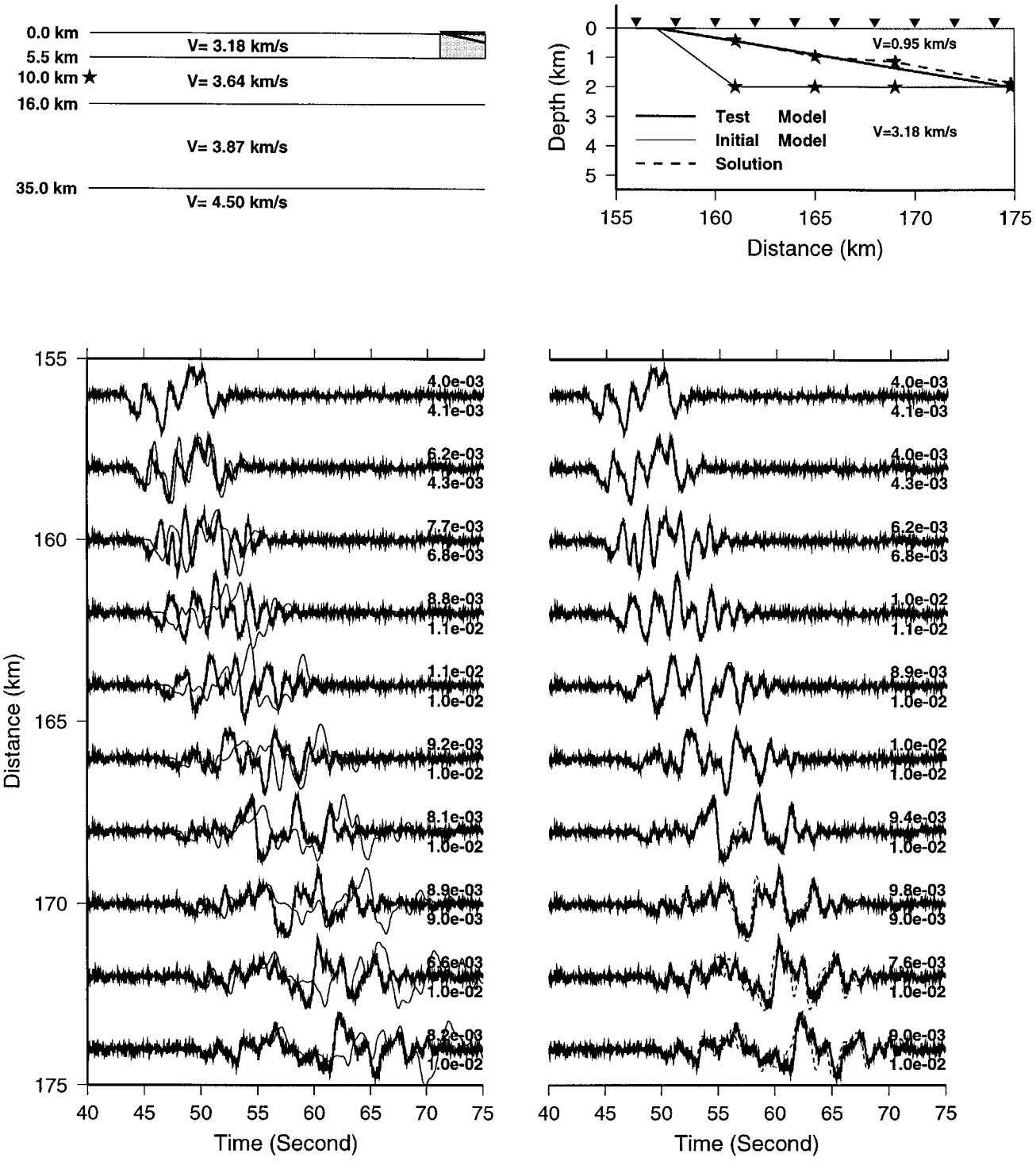


Figure 6. Inversion results of noisy data with 20% Gaussian noise added. Note the high-frequency noise does not change the inverse result. See Figure 5 caption for details.

0.04. But τ is relatively large, about 1 second on average. Finally, the sharp basin edge increases the peak amplitudes by 18% on average.

Similar to the aforementioned test, we give $f(\tau)$ and τ same weights when defining the objective function. But in this case, the amplitude of τ is 25 times as large as that of $f(\tau)$, so most of the information is provided by the travel-time residuals. After five iterations, we obtain the inverted model, which fits the test model well (Fig. 9b). The synthetic seismograms fit the data better and travel-time residual τ is

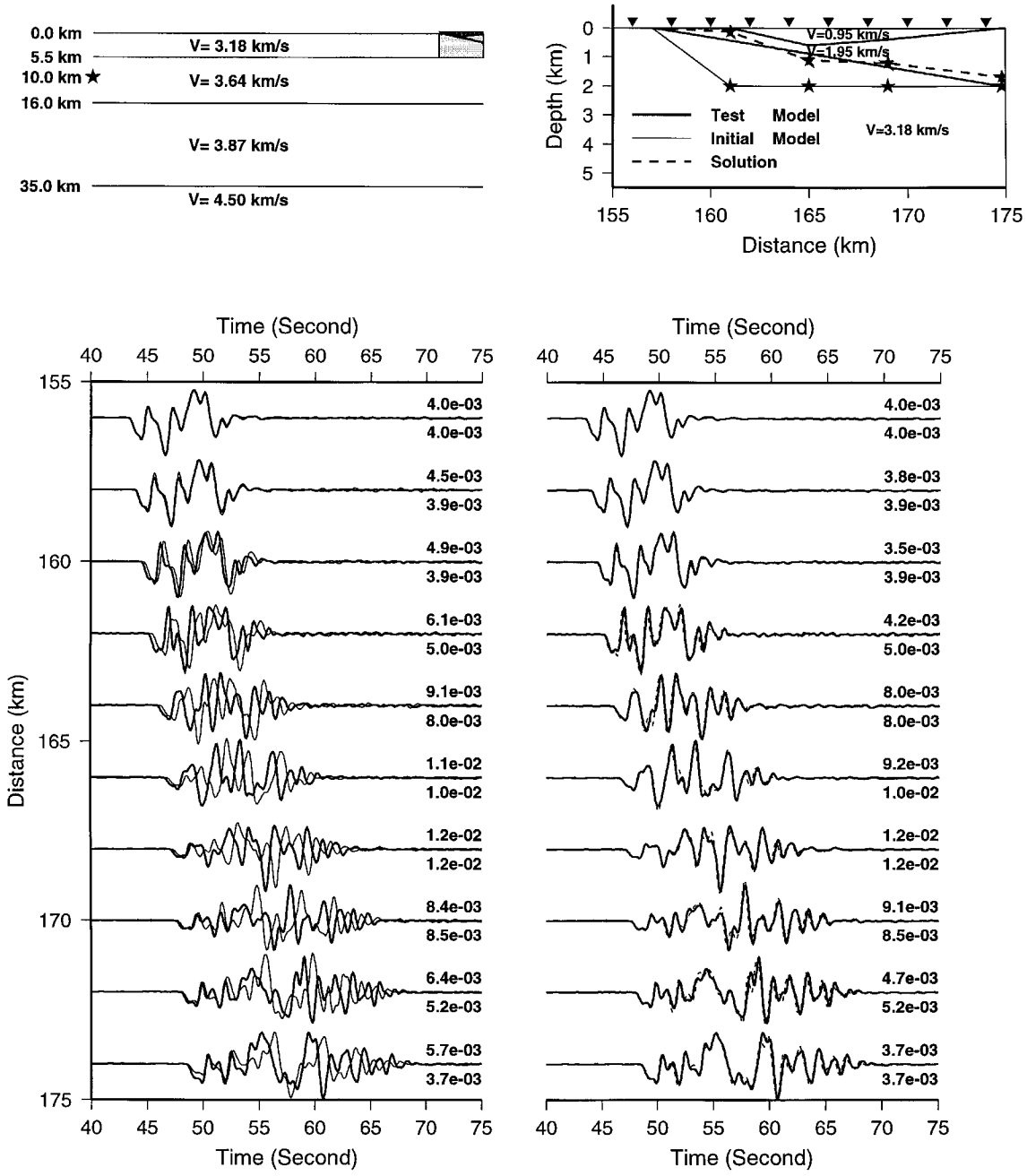


Figure 7. Inversion results for deeper structure. We assume that the first layer is known and only the shape of the second interface is to be determined. See Figure 5 caption for details.

also reduced by one order of magnitude (Fig. 9d). Note that if we did not use timing and stressed amplitudes, we could easily reduce these amplitudes by 20% by increasing the shallow velocity as indicated in Figure 3. Thus, the trade-off in structural details is not easily handled without timing even at these periods.

Unfortunately, although τ is one of the most useful parameters for determining structure, it requires absolute timing which is not available for our triggered observations. The more recent data of the type discussed early in Figure 1 will obviously solve this difficulty for future events. Moreover, the Landers dataset suggests a strategy of how we might proceed to refine the 2D sections. For example, consider broadband data along a profile similar to the Landers earthquake data set. We could begin with a sensitivity test assuming a set of models and comparing $f(\tau)$, as a function of frequency. Then, starting at the longest period available in the data, the inversion code could invert the broad structure.

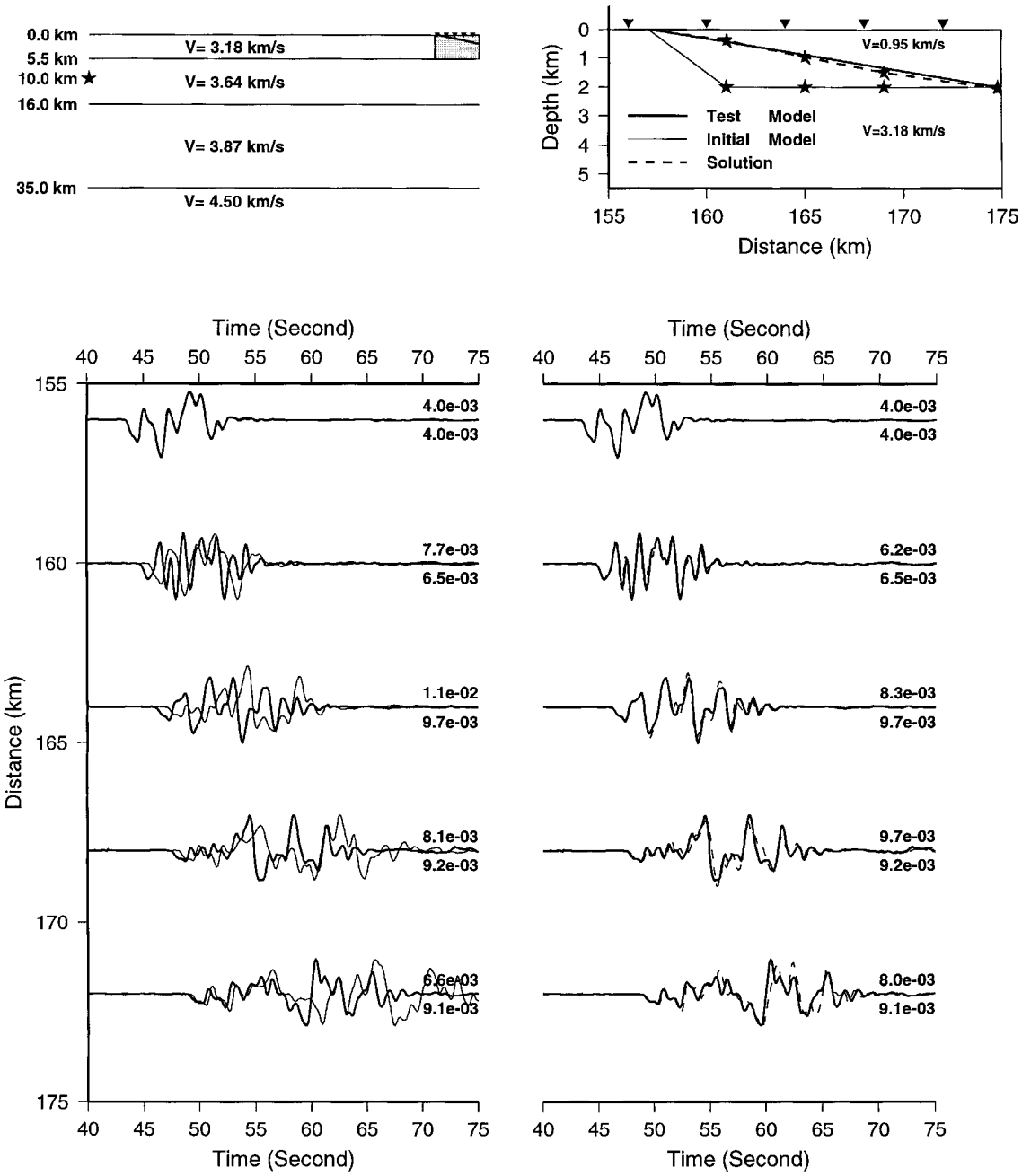


Figure 8. Inversion results using the waveforms of only five stations. See Figure 5 caption for details.

We could then allow more control points and more high-frequency signals to gain detailed knowledge about finer structure. Obviously, the station coverage, the frequency content of the data, and the complexity of basin structure will control what is possible to recover.

The methodology suggested in this article currently focuses on the inversion of basin interface shapes by *SH* wave, but extending it to invert both interface shapes and velocities is straightforward. In addition, *P-SV* energy can be used and may reduce the nonuniqueness of inversion.

In conclusion, we have demonstrated the utility of hybrid codes involving analytical-numerical interfacing and the transfer-function approach for basin-wave modeling. By localizing the basin problem, we concentrate our efforts on resolving a restricted set of parameters either by trial and error or by waveform inversion. By forward modeling the strong-motion waveforms of the Landers earthquake across an extended Los Angeles basin structure, we refined a model that requires a stronger velocity gradient with depth than indicated in the current 3D models (Wald and Graves, 1998).

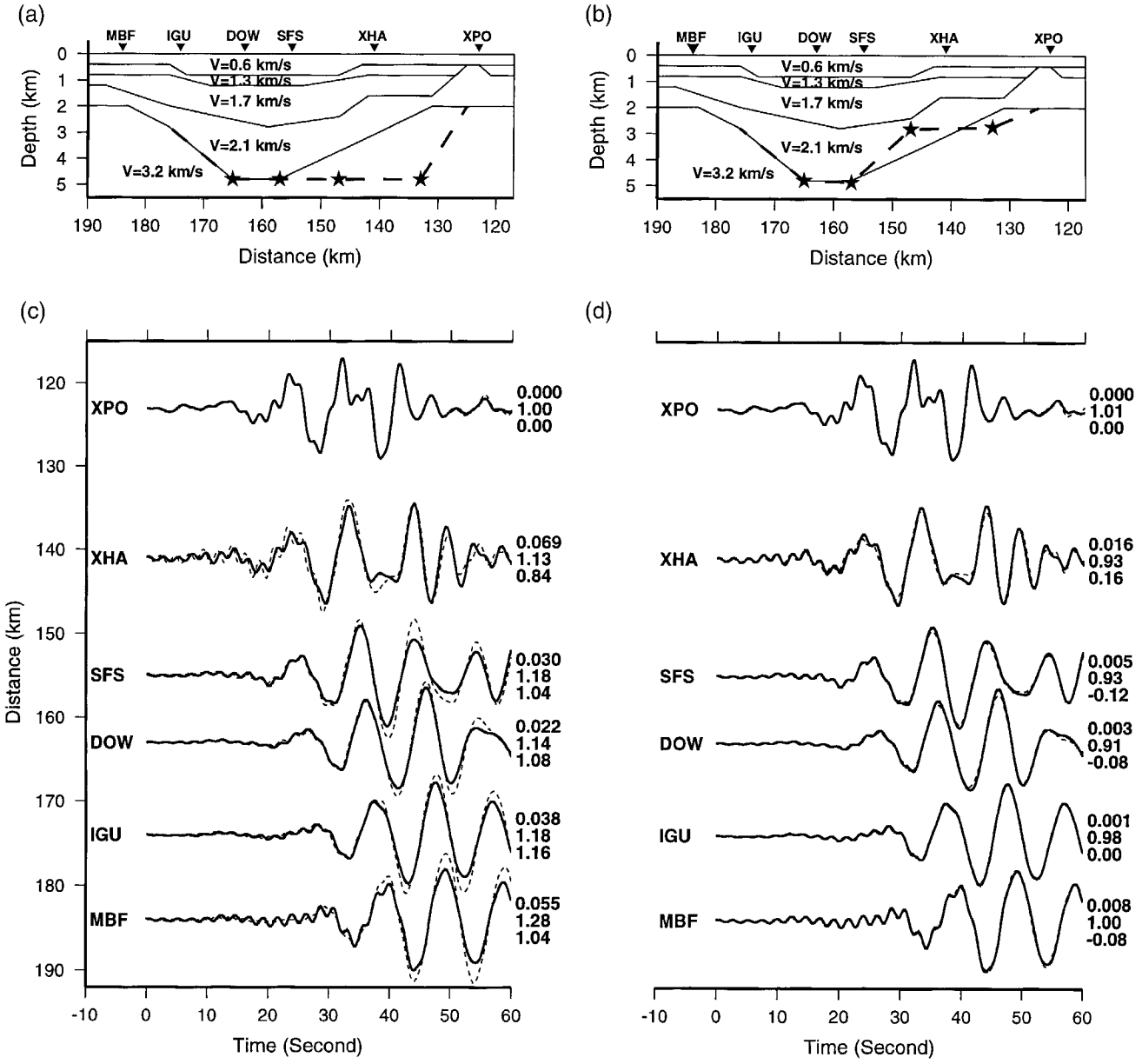


Figure 9. Inversion results assuming lower-frequency data. (a) The basin models, where the black lines indicate the test model. The dashed line indicates the difference between the initial model and the test model. Stars show the positions of control points. (b) Comparison between inversion model and test model. The dashed line indicates the difference between two. (c) Comparison between the test data (thick lines) and the synthetic waveforms (thin lines, generated from initial model). Note the synthetic seismograms are normalized by the test data and are shifted in order to best fit the data. The three numbers behind each trace are $f(\tau)$ (top), the ratio of peak amplitudes between the synthetic motion and the data (middle), and the travel-time residual τ (bottom). (d) Comparison between the test data and the synthetic waveforms are generated from the inversion model.

The model proved quite satisfactory in explaining the amplification from hard rock to the center of the basin (about a factor of 4 in displacement).

We also developed a waveform-inversion approach to study the shape of basins, in terms of dipping interfaces. A series of numerical tests using synthetic data were performed, which indicated that the waveform data and travel-time residuals can provide strong constraints on the shape of the basin. Moreover, it appears possible to obtain a detailed basin structure if we have sufficient high-quality data.

Acknowledgments

We thank Lianxing Wen for his help in using the interfacing code. This study was supported by a SCEC contract funded by NSF EAR 8920136 and by the USGS 99HQGR0038. This is Contribution Number 8629 of the Division of Geological and Planetary Sciences, California Institute of Technology.

References

Borcherdt, R. D., W. B. Joyner, R. E. Warrick, and J. F. Gibbs (1975). Response of local geologic units to ground shaking, in *Studies for Seismic Zonation of San Francisco Bay Region, U.S. Geol. Surv. Profess. Pap. 941A,* 52–67.

Esteva, L. (1988). The Mexico earthquake of September 19, 1985: consequences, lessons, and impact on research and practice, *Earthquake Spectra* **4,** 413–426.

Dreger, D. S., and D. V. Helmberger (1991). Source parameters of Sierra Madre earthquake from regional and local bodywaves, *Geophys. Res. Lett.* **18,** 2015–2018.

Graves, R. W. (1996). Simulating realistic earthquake ground motions in the regions of deep sedimentary basins, in *Proc. of the 11th World Conference on Earthquake Engineering,* Acapulco, Mexico.

Hauksson, H., and J. Haase (1997). Three-dimensional v_p and V_p/V_s in the Los Angeles basin and central Transverse Ranges, California, *J. Geophys. Res.* **102,** 5423–5454.

Kawase, H. (1996). The cause of damage belt in Kobe: "The basin edge effect", constructive interference of the direct S-wave with the basin-induced diffracted/Rayleigh waves, *Seism. Res. Lett.* **67,** no. 5, 25–34.

Luo, Y., and G. T. Schuster (1991). Wave-equation travel-time inversion, *Rev. Geophys.* **56,** 645–653.

McCulloh, T. H. (1960). Gravity variations and the geology of the Los Angeles Basin, *U. S. Geol. Surv. Profess. Pap. 400-B,* 320–325.

Magistrale, H., K. McLaughlin, and S. Day (1996). A geology-based 3D velocity model of the Los Angeles Basin sediments, *Bull. Seism. Soc. Am.* **86,** 1161–1166.

Mori, J., H. Kanamori, J. Davis, E. Hauksson, R. Clayton, T. Heaton, L. Jones, and A. Shakal (1998). Major improvements in progress for Southern California earthquake monitoring, *EOS* **79,** 217–221.

Olsen K. B., R. J. Archuleta, and J. R. Matarese (1995). Magnitude 7.75 earthquake on the San Andreas fault: three-dimensional ground motion in Los Angeles, *Science* **270,** 1628–1632.

Polak, E. (1971). Computational Methods in Optimization, Academic Press, New York, 44–65.

Scrivner, C., and D. V. Helmberger (1994). Seismic waveform modeling in the Los Angeles Basin, *Bull. Seism. Soc. Am.* **84,** 1310–1326.

Scrivner, C., and D. V. Helmberger (1999). Variability of Ground Motions in Southern California-Data from the 1995 to 1996 Ridgecrest Sequence, *Bull. Seism. Soc. Am.* **89,** no. 3, 626–639.

Sen, M. K., and P. L. Stoffa (1991). Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics* **56,** no. 10, 1624–1638.

Wald, D. J., and R. W. Graves (1998). The seismic response of Los Angeles Basin, California, *Bull. Seism. Soc. Am.* **88,** 337–356.

Wald, D. J., and T. H. Heaton (1994). Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake, *Bull. Seism. Soc. Am.* **84,** 668–691.

Wen, L., and D. V. Helmberger (1997). Propagation corrections for basin structure: Landers Earthquake, *Bull. Seism. Soc. Am.* **87,** 782–787.

Seismological Laboratory
California Institute of Technology
Pasadena, California
(C. J., D. V. H.)

U.S. Geological Survey
525 South Wilson Avenue
Pasadena, California
(D. J. W.)

Manuscript received 15 June 1999.