

Double-Difference Tomography: The Method and Its Application to the Hayward Fault, California

by Haijiang Zhang and Clifford H. Thurber

Abstract We have developed a double-difference (DD) seismic tomography method that makes use of both absolute and relative arrival times. By reducing systematic errors using the more accurate relative arrival times, the method produces an improved velocity model. Simultaneously, it yields event locations of a quality equivalent to those of the DD location method, while avoiding simplifying assumptions of that method. We test this method on a synthetic dataset and find that it produces a more accurate velocity model and event locations than standard tomography. We also test this method on a Hayward fault, California, earthquake dataset spanning 1984–1998. The earthquakes relocated by this method collapse to a thin line along the fault trace, consistent with previous results. The DD velocity model has sharper velocity contrasts near the source region than the standard tomography model.

Introduction

Local earthquake tomography (LET) has become a relatively routine application for use in seismically active regions covered by a dense seismic network. Common LET algorithms, however, do not take advantage of the many recent developments in earthquake location techniques. Many of these developments are aimed at improving relative and/or absolute location accuracy.

The accuracy of event hypocenters is determined by several factors, including the network geometry, available phases, and arrival time accuracies (Pavlis, 1986). Due to the presence of noise, the arrival times picked either manually or automatically generally have errors (Douglas *et al.*, 1997). Recent studies have shown substantial improvements in location precision for earthquakes and explosions when waveform cross-correlation (WCC) and event-clustering techniques are used to improve arrival time estimates or determine high-precision relative arrival times (VanDecar and Crosson, 1990). Example applications include Mount St. Helens (Fremont and Malone, 1987); Hawaii (Got *et al.*, 1994; Rubin *et al.*, 1998); California (Poupinet *et al.*, 1984; Nadeau *et al.*, 1994; Shearer, 1997; Rubin *et al.*, 1999; Waldhauser *et al.*, 1999; Waldhauser and Ellsworth, 2000, 2002); the Coso geothermal field, Nevada (Lees, 1998); the Soultz geothermal field, France (Rowe *et al.*, 2002); and explosions at the Balapan test site (Phillips *et al.*, 2001; Thurber *et al.*, 2001). These studies are based on the assumption that waves generated by two similar sources, propagating along similar paths, will generate similar waveforms, and WCC can then be used to determine precise relative arrival times. These and other studies have demonstrated substantial improvement in the definition of seismogenic features and in the accuracy of relative locations of ground-truth

events that is possible using multiple-event methods with high-precision absolute or relative arrival time data. In addition, they have also provided new insight into tectonic processes, earthquake recurrence, and earthquake interaction.

We can make an important distinction between the two fundamentally different ways WCC data have been used: (1) by directly using relative arrival times to determine relative event locations (e.g., Fremont and Malone, 1987; Got *et al.*, 1994; Waldhauser and Ellsworth, 2000) or (2) by adjusting absolute arrival time picks to minimize discrepancies among relative arrival times (Dodge *et al.*, 1995, 1996; Shearer, 1997; Rowe *et al.*, 2002). The advantage of the former approach is that it incorporates all the available information contained within the multitude of relative arrival time differences with a direct measure of quality (the correlation value) associated explicitly with each datum. A disadvantage is that some simplifying assumption needs to be made to derive the locations from the arrival time differences. For example, in the methods of Fremont and Malone (1987) and Got *et al.* (1994), all events in a cluster have precisely the same takeoff angle and azimuth to each station. As a result, the derived locations are ultimately relative, not absolute, so that some assumption must be made to end up with useful event coordinates (e.g., final locations are computed relative to a catalog-based cluster centroid). Waldhauser and Ellsworth (2000) proposed a different location algorithm, in which the spatial partial derivatives for a set of events are evaluated at the current location of each event. It is assumed, however, that the path anomalies due to velocity heterogeneity are location independent. This assumption is valid for closely spaced events, but is not true for far apart events. As a result, such far apart event locations may

be biased due to velocity heterogeneity (Got *et al.*, 1994; Waldhauser and Ellsworth, 2000; Wolfe, 2002). In contrast, the latter approach uses the relative arrival times to determine a much smaller number of adjusted arrival time picks, but these picks are absolute arrival times and so can be used to determine absolute locations (in an existing velocity model, or using tomography).

We have developed a new method that combines the advantages and avoids the disadvantages of the previous approaches. It is based on the code hypoDD of Waldhauser (2001) and makes use of both absolute and relative arrival time data. The method determines a 3D velocity model jointly with the absolute and relative event locations. This approach has the advantage of including relative arrival times with their quality values along with absolute arrival times, thereby not discarding valuable information by only using adjusted picks, and at the same time dispensing with simplifying assumptions about ray path geometries or path anomalies and producing absolute locations, not just relative locations. The velocity model obtained with double-difference (DD) tomography should also be superior to that from standard tomography. With standard tomography, event locations will be somewhat scattered due to imprecise picks and correlated errors, but in DD tomography, the use of the differential arrival times (including both the WCC and catalog time difference data) removes most of these errors, which will in turn remove some fuzziness from the velocity model. To demonstrate the effectiveness of the method, we have applied it to a synthetic dataset based on the idealized velocity structure of the San Andreas fault in central California and to the Hayward fault dataset of Waldhauser and Ellsworth (2002).

DD Tomography

The body-wave arrival time T from an earthquake i to a seismic station k is expressed using ray theory as a path integral,

$$T_k^i = \tau^i + \int_i^k u \, ds, \quad (1)$$

where τ^i is the origin time of event i , u is the slowness field, and ds is an element of path length. The source coordinates (x_1, x_2, x_3) , origin times, ray paths, and the slowness field are the unknowns. The relationship between the arrival time and the event location is highly nonlinear, so a truncated Taylor series expansion is generally used to linearize equation (1). This linearly relates the misfit between the observed and predicted arrival times r_k^i to the desired perturbations to the hypocenter and velocity structure parameters:

$$r_k^i = \sum_{l=1}^3 \frac{\partial T_k^i}{\partial x_l^i} \Delta x_l^i + \Delta \tau^i + \int_i^k \delta u \, ds. \quad (2)$$

Subtracting a similar equation for event j observed at station k from equation (2), we have

$$r_k^i - r_k^j = \sum_{l=1}^3 \frac{\partial T_k^i}{\partial x_l^i} \Delta x_l^i + \Delta \tau^i + \int_i^k \delta u \, ds - \sum_{l=1}^3 \frac{\partial T_k^j}{\partial x_l^j} \Delta x_l^j - \Delta \tau^j - \int_j^k \delta u \, ds. \quad (3)$$

Assuming that these two events are near each other so that the paths from the events to a common station are almost identical and the velocity structure is known, then equation (3) can be simplified as

$$dr_k^{ij} = r_k^i - r_k^j = \sum_{l=1}^3 \frac{\partial T_k^i}{\partial x_l^i} \Delta x_l^i + \Delta \tau^i - \sum_{l=1}^3 \frac{\partial T_k^j}{\partial x_l^j} \Delta x_l^j - \Delta \tau^j, \quad (4)$$

where dr_k^{ij} is the so called double-difference (Waldhauser and Ellsworth, 2000). This term is the difference between observed and calculated differential arrival times for the two events and can also be written as

$$dr_k^{ij} = r_k^i - r_k^j = (T_k^i - T_k^j)^{\text{obs}} - (T_k^i - T_k^j)^{\text{cal}}. \quad (5)$$

The observed differential arrival times $(T_k^i - T_k^j)^{\text{obs}}$ can be calculated from both WCC techniques for similar waveforms and absolute catalog arrival times. Equation (4) is known as the DD earthquake location algorithm (Waldhauser and Ellsworth, 2000).

In this approach, earthquake locations may be biased when interevent distances exceed the scale length of velocity variations. Waldhauser and Ellsworth (2000) applied a distance-weighting factor to reduce or exclude data from event pairs that are far apart. Although the arrival difference data from such events may be excluded, they can still be linked in the inversion via a series of intermediate pairs (Got *et al.*, 1994). For example, the pair $T_k^i - T_k^j$ can be linked if the two pairs $T_k^i - T_k^m$ and $T_k^m - T_k^j$ are included.

To overcome this limitation, we use the differential arrival time data and equation (3) directly. It is known that there is a coupling effect between the event hypocenters and the velocity structure (Thurber, 1992). Our purpose is to determine not only the relative event locations, but also their absolute locations and the velocity structure. Also note that the ray paths from two nearby events will substantially overlap, meaning that the model derivatives in equation (3) will essentially cancel outside the source region. For this reason, we include the absolute arrival times in the inversion to resolve the velocity structure outside the source region. By doing this, we can jointly determine the velocity structure and the relative event locations as well as the absolute event locations accurately.

We developed a DD tomography code tomoDD based on the DD location code hypoDD (Waldhauser, 2001). In the

current version of tomoDD, we use the pseudo-bending ray-tracing algorithm (Um and Thurber, 1987) to find the rays and calculate the travel times between events and stations. The model is represented as a regular set of 3D nodes, and the velocity values are interpolated by using the trilinear interpolation method. The hypocentral partial derivatives are calculated from the direction of the ray and the local velocity at the source (Lee and Stewart, 1981). The ray path is divided into a set of segments, and the model partial derivatives (calculated in terms of fractional slowness perturbation, so that the derivatives are related to path length) are evaluated by apportioning the derivative to its eight surrounding nodes according to their interpolation weights on the segment midpoint (Thurber, 1983). For two rays observed at a common station, if the model derivatives at the common inversion node for these two rays are close enough (i.e., within 5 m), then we set the corresponding elements to zero in the model derivative matrix to make the system more stable. Distance weighting is also used in our DD tomography to control the maximum separation between event pairs and apply greater weight to the closer events, similar to Waldhauser and Ellsworth (2000).

In our simultaneous inversion for velocity structure and event locations, velocity anomalies are constrained by seeking a first-order smooth model. Smoothing regularization should provide a minimum-feature model that contains only as much as structure as can be resolved above the estimated level of noise in the data (Constable *et al.*, 1987; Sambridge, 1990; McCaughey and Singh, 1997). We apply the same smoothing weight to the horizontal and vertical directions.

Three types of data, the absolute arrival times, the catalog differential arrival times, and the WCC data, are used in the inversion. To combine these three types of data into one system, we apply a hierarchical weighting scheme during the inversion, similar to hypoDD. Waveforms normally correlate between event pairs only within the same cluster. The relative locations between two events from different clusters should be controlled by differential catalog data with larger event separation. For this reason, we start the inversion by applying greater weight to the catalog data (both differential and absolute catalog data) to establish the large-scale result (1 for absolute data, 0.1 for differential catalog data, and 0.01 for cross-correlation data), similar to hypoDD (Waldhauser, 2001). Then the catalog differential data are weighted more to refine the event locations and the velocity structure near the source regions (1 for differential catalog data, 0.1 for absolute catalog data, and 0.01 for cross-correlation data). If WCC data are available, they will then be weighted even more than the catalog differential data in a final step to further refine the event locations and the velocity structure near the source region (1 for cross-correlation data, 0.01 for differential catalog data, and 0.001 for absolute catalog data). We downweight the differential catalog data by a factor of 100 in the final stage of the inversion, as the cross-correlation data are at least an order of

magnitude more precise than the manual picks (Waldhauser and Ellsworth, 2000).

The complete system of linear equations (2) and (3), along with the smoothing constraint equations, is solved by means of the LSQR algorithm (Paige and Saunders, 1982) for the damped least-squares problem. Each equation is weighted according to the *a priori* data uncertainty, data type, distance between the event pair, and misfit during each iteration. The relative weighting for the different data types and the distance weighting are determined *a priori*, whereas the residual weighting is determined *a posteriori*, with large residuals rejected or downweighted by a biweight function (Waldhauser and Ellsworth, 2000). For LSQR, the number of iterations required to reach a certain accuracy depends strongly on the scaling of the problem (Paige and Saunders, 1982). Proper scaling of the rows or columns of the matrix makes it easier to recover the solution. For this reason, before the system is fed into the LSQR, each column is scaled so that the L^2 norm of each column is equal to 1.

Synthetic Test

To assess the effectiveness of the DD seismic tomography method, we apply the method to a synthetic dataset. We use the same synthetic dataset that was constructed based on an idealized model of the velocity structure of the San Andreas fault in central California (Kissling *et al.*, 1994). To the west of the “fault” ($x \leq 0$ km), the velocity is constant (6 km/sec); to the east, there is a very sharp gradient into a low-velocity zone of 4 km/sec from $x = 1$ to 5 km, a sharp gradient to 5 km/sec at $x = 6$ km, and then a linear increase from 5 to 6 km/sec in the region between $x = 6$ and 38 km (Fig. 1). The X - Y nodes used to represent the velocity model are shown in Figure 2; in depth, nodes are positioned at 0, 3, 7, 11, and 16 km. The events and stations used to construct the synthetic dataset are from the actual seismicity and U.S. Geological Survey stations in the Loma Prieta region (Fig. 2).

We added Gaussian random noise with zero mean and a standard deviation of 0.04 sec to all the true arrival times. In addition, we also added a constant noise term to the arrivals at each station from a uniform distribution between -0.3 and 0.3 sec. This simulates the case that the systematic errors (model errors and pick bias) associated with the arrival times are larger than the random ones. We construct the pseudo cross-correlation data directly from the absolute arrival times by differencing the synthetic arrival times at common stations for pairs of events within 20 km. As a result, the cross-correlation data are more accurate than the absolute data. We note, however, that the synthetic times were provided to us rounded off to two decimal places, so even the cross-correlation data have round-off errors up to 0.005 sec.

We first use the DD location algorithm hypoDD to relocate the events (Waldhauser, 2001). The 1D velocity model used is the Dietz and Ellsworth (1990) model, based on seismic refraction and earthquake data. We use both the

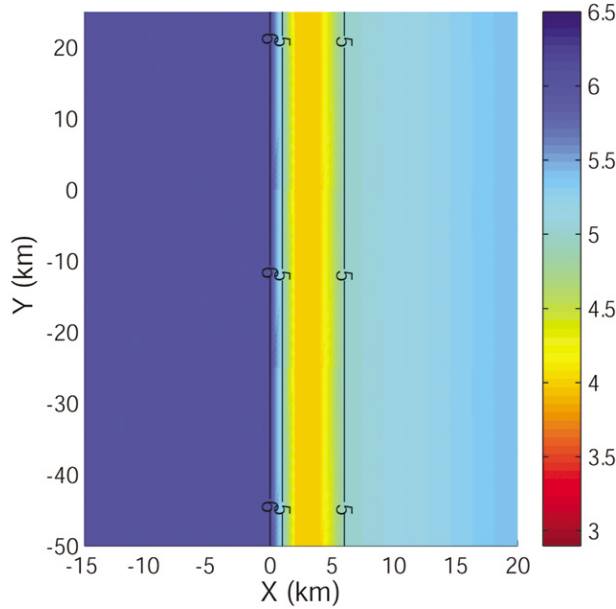


Figure 1. A horizontal slice through the true synthetic velocity model (Kissling *et al.*, 1994). The true velocity model in 3D is similar to a “vertical sandwich,” with the velocity constant with the depth.

residual weighting and distance weighting schemes in the inversion to control the large residuals, as did Waldhauser and Ellsworth (2000). The absolute differences between this set of event relocations and the true locations and their standard deviations are shown in Table 1. The results indicate that the event relocations from the DD location algorithm based on a 1D velocity model have a substantial bias (>1 km in each coordinate direction) from the true locations. This bias is caused by a difference between the velocity model (horizontal layers) used to calculate the DD locations and the true velocity model (vertical “sandwich”) used to generate the data. The heterogeneity of the true velocity model makes path anomalies different for different events. However, the DD location method assumes that the path anomalies are location independent, and this assumption introduces bias into event locations (Wolfe, 2002).

We next relocate the events by using our modified DD location method hypo3DD (hypoDD with 3D raytracing added) using the true 3D velocity model. In this case, the arrival time residuals are still only related to the perturbations to the event hypocenters and origin times, not to velocity structure. The DD location method employing the true velocity model produces much more accurate event locations than the DD location method based on a 1D velocity model, as expected (Table 1). The location errors are mainly due to the noise in the arrival times and the inaccuracies in representing the true model with sharp gradient changes by a linearly interpolated grid.

In practice, it is never the case that the true velocity model is known exactly. Recognizing the coupling effect between the event locations and the velocity structure (Thur-

ber, 1992), we next relocate the event locations and determine velocity structure simultaneously. The standard seismic tomography method uses only the absolute arrival times, whereas our method uses absolute and differential times. First, we show the results of applying the standard tomography method to the relatively noisy absolute arrival times. The computational algorithm is identical to that for DD tomography, but the differential times are excluded. Both damping and smoothing (with weight 5.0 in both the horizontal and vertical directions) are applied to the inversion to make the solution more stable. The same residual weighting as the DD location method is used. The inversion starts from the same 1D velocity model as the DD location method. The inversion grid nodes used are shown in Figure 2. Figure 3a shows horizontal slices through the velocity structure obtained from the standard tomography. The absolute difference between this velocity model and the true velocity model has a median value of 0.164 km/sec, a mean value of 0.245 km/sec, and a standard deviation of 0.249 km/sec. The main features of the true velocity model are evident in the depth slices. The event locations have smaller errors than those from the DD location method (Table 1). The standard tomography has the ability to adjust the velocity model and thus improve the locations. However, the large, variable path anomalies due to the difference between the hypoDD velocity model and the actual velocity model make the DD locations worse than the standard tomography even with more accurate differential data. This demonstrates how the coupling between the velocity structure and event locations affects the results.

DD seismic tomography uses both the noisy absolute arrival times and more accurate differential times. We apply the same damping and the smoothing constraints to stabilize the solution, with the residual weighting used to penalize the large residuals during the inversion. The DD seismic tomography starts with the same 1D velocity model and uses the same inversion grid as the standard tomography. Figure 3b shows the horizontal slices of the velocity structure from the DD tomography. The absolute difference between this velocity structure and the true velocity structure has a median value of 0.136 km/sec, a mean value of 0.178 km/sec, and a standard deviation of 0.164 km/sec. The DD tomography characterizes well the low-velocity zone of the true velocity model, especially at the depths of 3 and 7 km. Subtracting the DD tomography solution and the standard tomography solution from the true model (Fig. 3c,d), we find that the velocity model from DD tomography has a more correct value in the low-velocity trough except at the depth of 0 km, where the ray paths from event pairs almost completely overlap and the accuracy of velocity structure is mainly controlled by absolute catalog data. This indicates that DD tomography recovers better the low-velocity zone, and overall the velocity model is more accurate than the standard tomography. Compared with the standard tomography method, the DD tomography also produces more accurate event locations (Table 1).

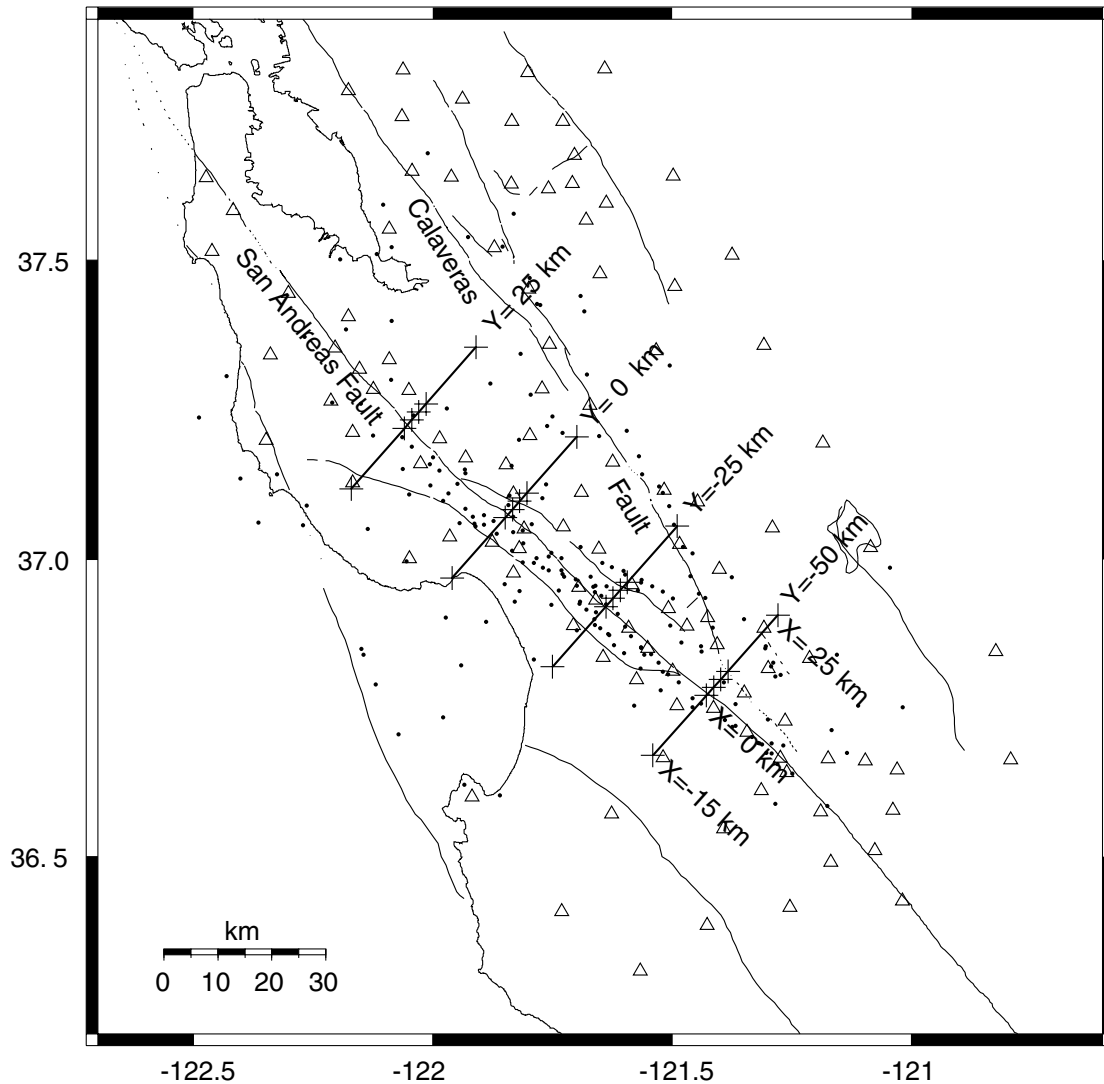


Figure 2. Event locations (filled circles) and stations (open triangles) used for the synthetic dataset. The inversion grid used in the standard tomography and DD tomography solutions is shown as the crosses. The inversion grid points are at $X = -35, -15, 0, 2, 4, 6, 20, 35$ km, at $Y = -60, -40, -20, 0, 20, 40$ km, and at $Z = 0, 3, 7, 11, 16$ km.

Table 1

The Absolute Differences between the True Locations and Those from the DD Location Method based on 1D Velocity Model, the DD Location Method based on 3D True Velocity Model, Standard Tomography, and DD Tomography

	Median Value (km)			Standard Deviation (km)		
	Latitude	Longitude	Depth	Latitude	Longitude	Depth
DD location (1D)	1.131	1.235	1.123	0.976	0.941	1.658
DD location (3D)	0.432	0.371	0.296	0.336	0.332	0.479
Standard tomography	0.320	0.295	0.460	0.399	0.342	0.575
DD tomography	0.238	0.218	0.329	0.288	0.314	0.427

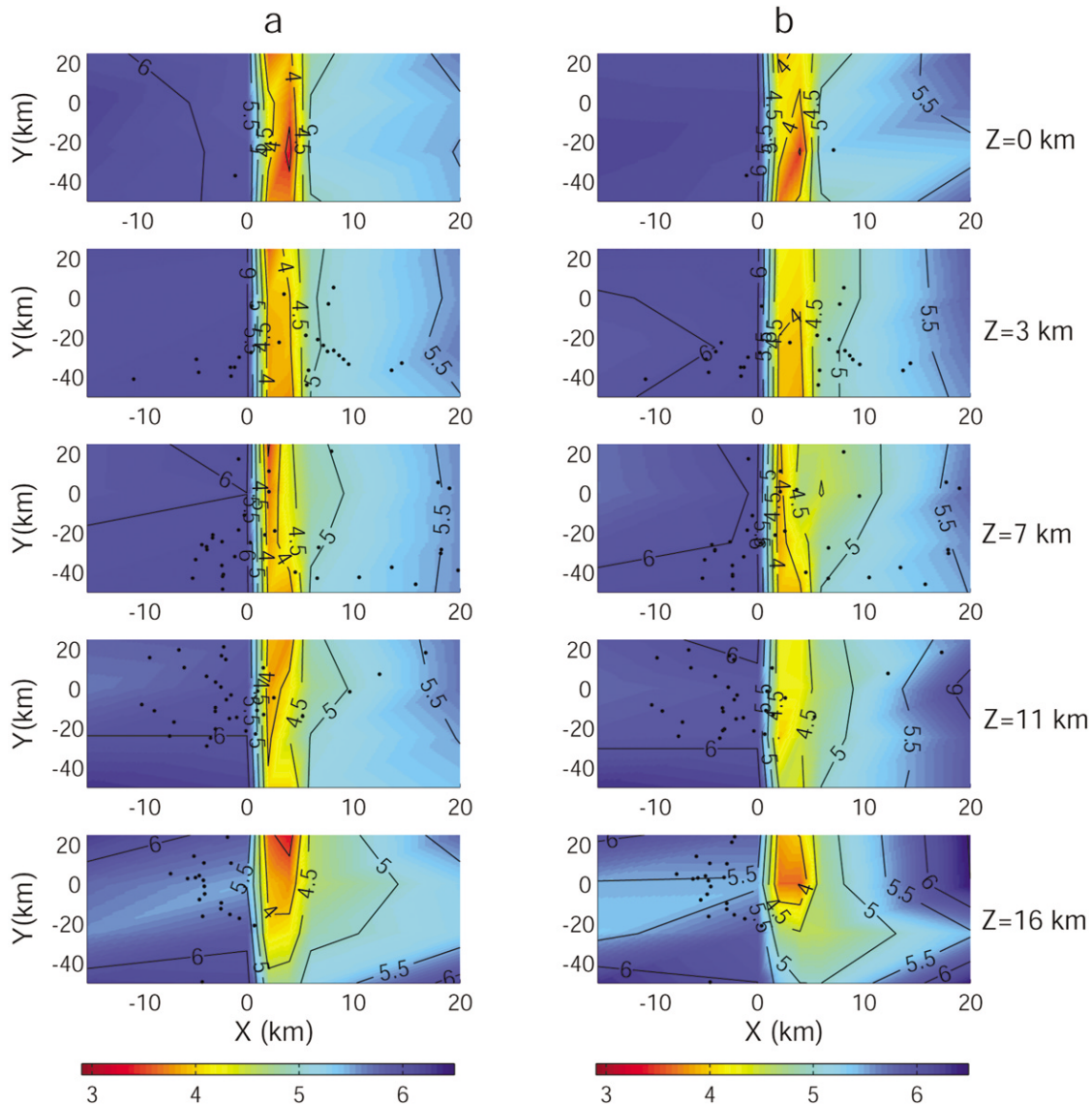


Figure 3. Horizontal slices through (a) the velocity model from standard tomography, (b) the velocity model from DD tomography. (Continued on next page.)

Application to the Hayward Fault: Event Locations

We applied the DD tomography method to earthquakes recorded between 1984 and 1998 by the Northern California Seismic Network (NCSN) on the Hayward fault, California, the same dataset as used by Waldhauser and Ellsworth (2002). The Hayward fault accommodates about 9 mm/yr of right-lateral relative plate motion between the North American plate and the Pacific plate. Franciscan rocks, consisting mainly of graywacke mélangé with smaller amounts of chert, shale, mafic volcanic rock, and limestone, form most of the upper crust to the west of the Hayward fault (Bailey *et al.*, 1964). In the upper crust to the east of the Hayward fault, there are folded and thrust shallow marine sandstones and shales of the Great Valley Sequence overlying

the Franciscan terrane (Jennings, 1977). Rocks of the Franciscan terrane are faster than the Great Valley and younger sediments (Walter, 1990; Hole *et al.*, 2000).

We utilized 1489 earthquakes with magnitudes from M 0.2 to 4.5 and 52 stations (Fig. 4). We used 17,955 P -wave differential arrival times from WCC, in addition to 767,127 P -wave arrival time differences, computed directly from the absolute catalog arrival times, and 20,257 absolute catalog P -wave arrival picks. Waldhauser and Ellsworth (2002) selected the catalog arrival time differences so that any event is linked to a maximum of 10 neighboring events by at least eight pair-wise observations. The WCC data are mainly from the events with similar waveforms along the fault and are far fewer than the differential catalog data (Waldhauser and Ellsworth, 2002).

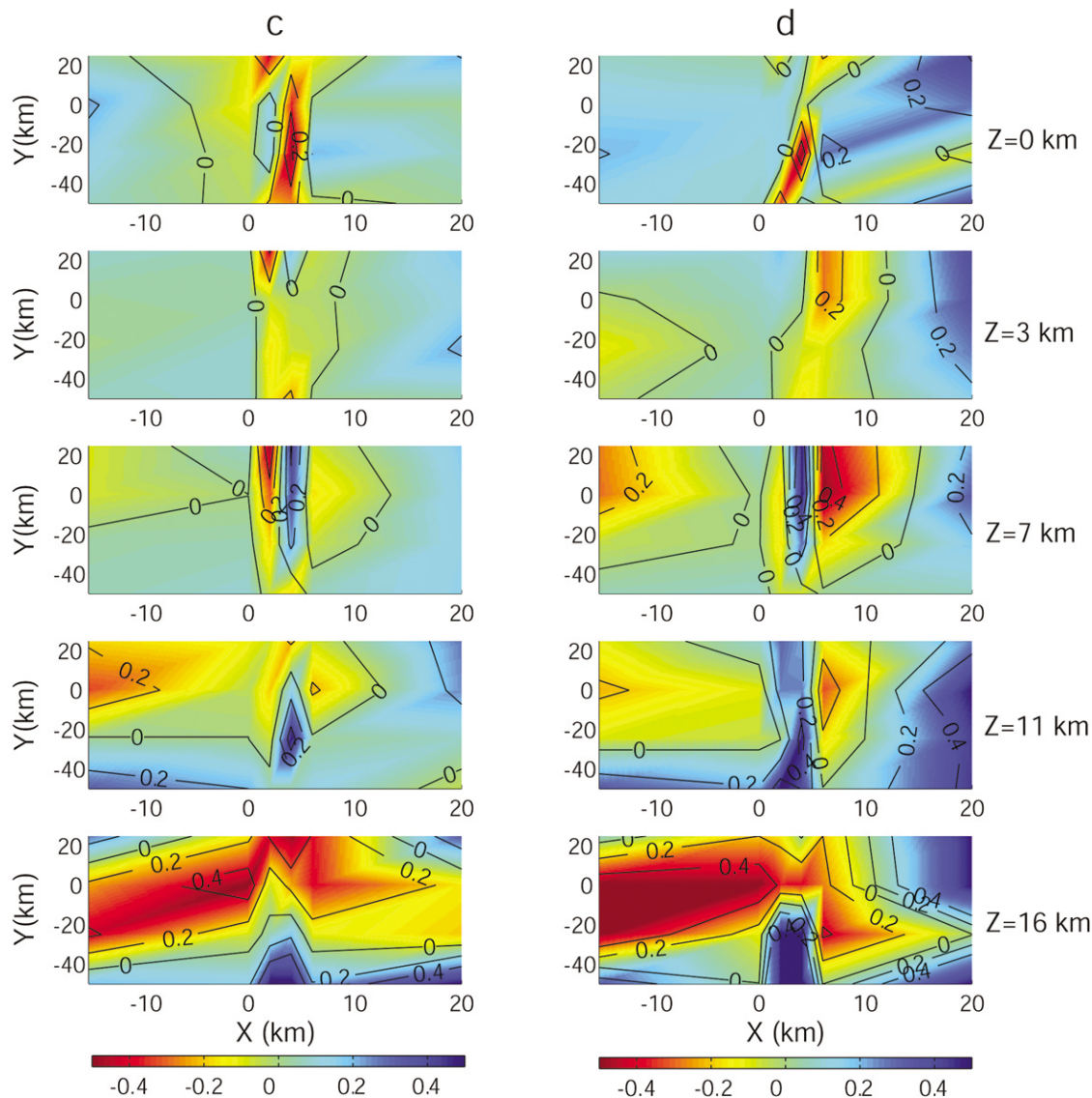


Figure 3. (continued) (c) the velocity difference between DD tomography solution and true model, and (d) the velocity difference between standard tomography solution and true model. Black dots indicate the earthquake hypocenters within half the grid size of the slice.

We use a 1D velocity model as the starting model for both the DD tomography and standard tomography. The X-Y nodes of the velocity model grid are shown in Figure 4; in depth, nodes are placed at 0, 4, 8, 12, and 20 km. Different smoothing weights of 5, 10, 15, and 30 are tested. The main features of these models are very similar in the regions with good ray coverage. In other regions, the model using the weight of 5 shows some oscillations, which are reduced by using the larger weights. Considering the trade-off between the roughness and the stabilization of the model, we choose the model using weight 10 as our preferred model.

Figure 5a shows the catalog locations, which are scattered along the fault zone. For comparison, we also relocate the events by using the standard tomography method with

only the absolute catalog data (Fig. 5b). We see that the event locations are still scattered, comparable to the catalog locations (Fig. 5a).

Figure 5c,d shows the event relocations by using the DD location method (Waldhauser and Ellsworth, 2000) and the DD tomography method, respectively. In the DD location method, the weighted root mean square (rms) residuals for cross-correlation data and catalog data decrease from 112 to 3 msec and from 159 to 26 msec, respectively. For the DD tomography method, the weighted rms residuals for cross-correlation data and catalog data decrease from 120 to 2 msec and from 176 to 24 msec, respectively (note that the DD tomography solution includes the 20,257 absolute P-wave arrival times, which is why the starting rms value is

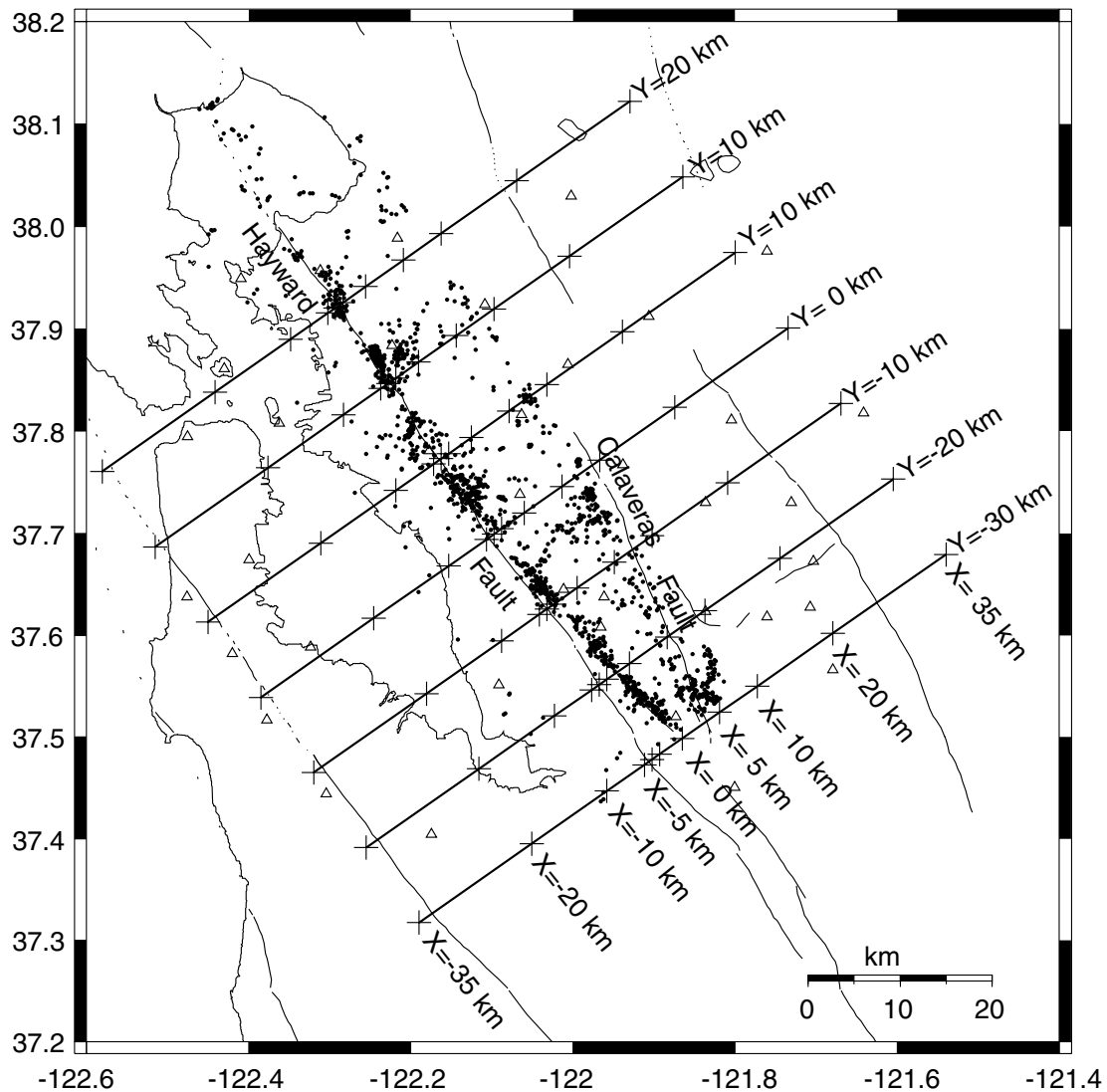


Figure 4. Stations (open triangles) and Hayward fault seismicity (filled circles) between 1984 and 1998 located by the NCSN. The inversion grid used for the standard tomography and DD tomography solutions is shown as the crosses. The inversion grid points are at $X = -35, -20, -10, -5, -4, -3, 0, 5, 10, 20, 35$ km, at $Y = -30, -20, -10, 0, 10, 20, 30$ km, and at $Z = 0, 4, 8, 12, 20$ km.

higher). We can see that, after relocation, both DD methods provide a sharp picture of the seismicity. However, the absolute event locations are different between the two methods. First, we note that the cluster centroid for the DD location method is located at latitude 37.721° , longitude -122.064° , and depth 7.467 km, whereas it is located at latitude 37.716° , longitude -122.069° , and depth 6.922 km for the DD tomography method. That is, there is a shift of about 632 m in the horizontal direction and 565 m in the vertical direction for the cluster centroid.

To compare the event locations in more detail, we zoom in on the region of latitude 37.83° to 37.91° and longitude -122.26° to -122.18° . The coordinate center is located at latitude 37.72° and longitude -122.06° . The positive Z axis is downward, and the coordinate system is rotated anticlock-

wise 35° so that the Y axis is nearly parallel to the strike of the Hayward fault, with Y increasing to the northwest and X increasing to the northeast. Figures 6, 7, 8, and 9 show map views and cross sections of event locations for catalog data, the standard tomography method, the DD location method, and the DD tomography method, respectively. The seismicity inside the box forms a northwest-striking zone associated with the Hayward fault, and that outside the box a diffuse zone of earthquakes about 2 km northeast of the fault zone. The event locations from the standard tomography method are still scattered and have very little improvement compared to the catalog locations (Figs. 6 and 7).

In map view, both DD methods collapse the on-fault seismicity to a thin line (Figs. 8a and 9a), with most of the events aligning in depth along horizontal lineations (Figs.

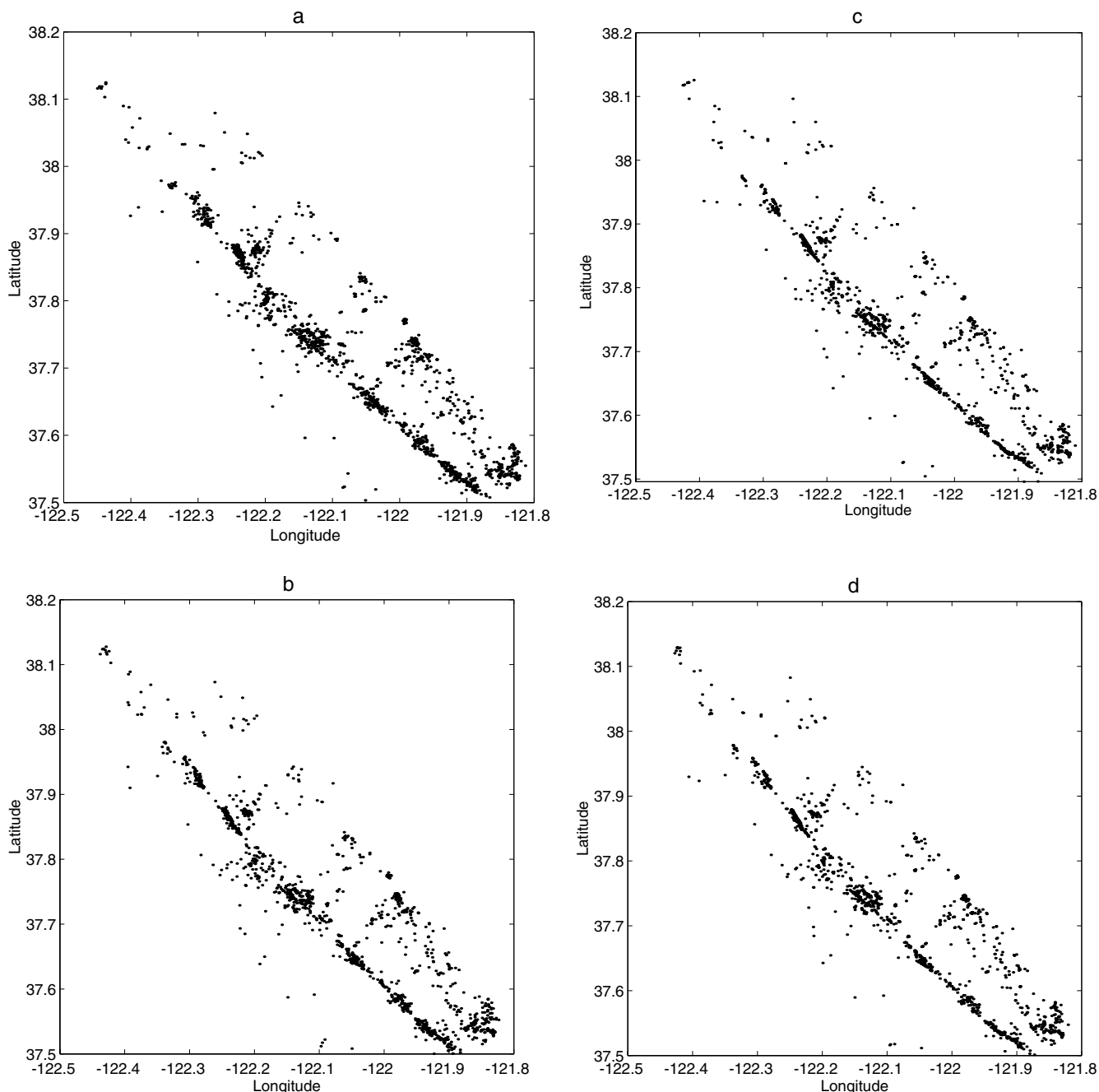


Figure 5. Hayward fault event locations. (a) NCSN catalog locations. (b) Relocations by the standard tomography method (i.e., with only absolute catalog data). (c) Relocations by the DD location method (Waldhauser and Ellsworth, 2002). (d) Relocations by the DD tomography method.

8d and 9d). Although most of the relative event locations from the two DD methods are quite similar, there are some differences between them in detail. In map view, an apparent kink in the seismicity trend at 37.86° N in the hypoDD result (Fig. 8a) becomes a simpler offset in the tomoDD result (Fig. 9a). In cross section, the event relocations define a nearly planar, vertical fault zone striking in the direction of the surface trace of the Hayward fault in the hypoDD loca-

tions (Fig. 8c). The tomoDD locations, however, show a near-vertical fault plane between 4 and 8 km depth and a slight southwest dip between 8 and 12 km depth (Fig. 9c). We project the event locations in the box into finer slices (1-km separation in the Y direction) and find that the bend at the depth of 8 km mainly happens from $Y = 20$ to 23 km. This is consistent with the velocity complexity in the same area (see next section).

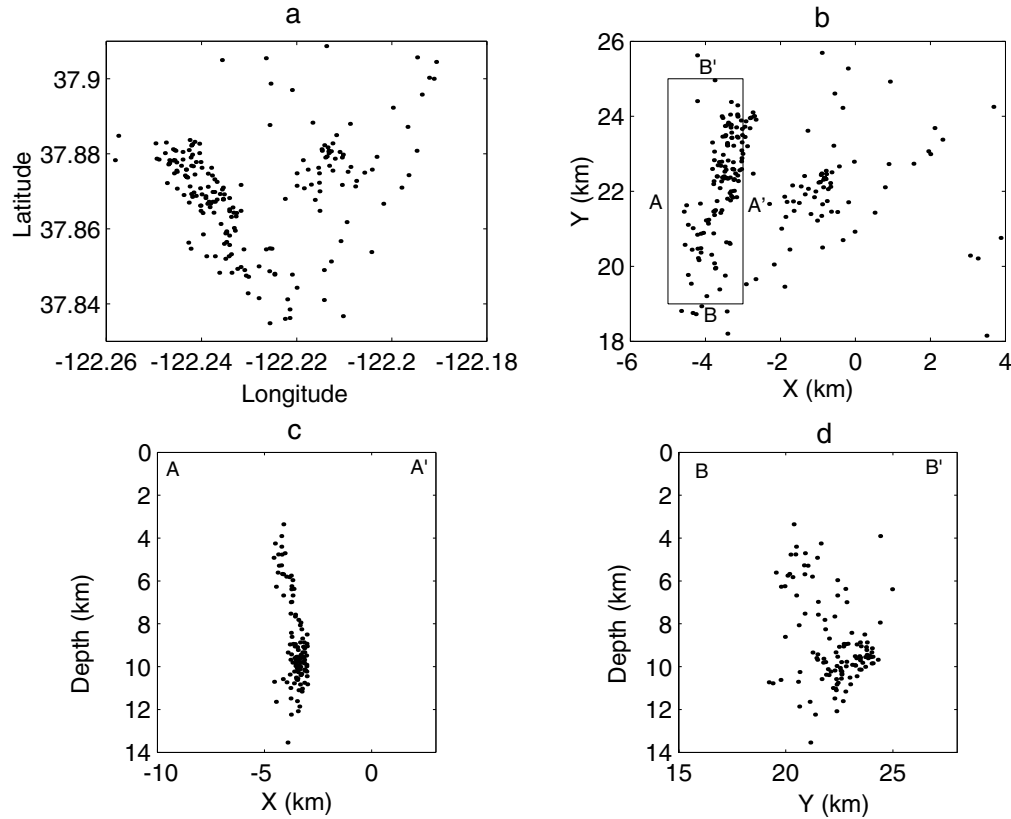


Figure 6. Catalog locations in the region of latitude 37.83° to 37.91° and longitude -122.26° to -122.18° . (a) Map view of event locations in the latitude and longitude system. (b) Map view of event locations in the coordinate system rotated anticlockwise 35° with the center located at latitude 37.72° and longitude -122.06° , positive X axis directing to the northeast, positive Y axis to the northwest, and positive Z axis downward. (c) AA' cross section. (d) BB' cross section.

Application to the Hayward Fault: Velocity Structure

Figure 10a,b shows across-strike vertical sections through the velocity model resulting from DD tomography and standard tomography, respectively. We see that the Hayward fault is marked by a strong velocity contrast (Fig. 10a,b), and this contrast persists vertically beneath the surface trace of the fault to the maximum depths constrained by the model, similar to the results shown in Hole *et al.* (2000). Higher velocity rocks are located to the west, consistent with the local geological setting. The strong velocity contrast near the surface is due to the boundary between the Franciscan terrane and the Great Valley Sequence (Walter, 1990; Hole *et al.*, 2000). Comparing Figure 10a with 10b, we note that the velocity contrast is sharper in the model from the DD tomography method. Figure 10c shows the velocity difference between the DD tomography and the standard tomography. It clearly shows that the velocities from the DD tomography are faster (about 0.1–0.2 km/sec) than those from the standard tomography to the west of the fault and slower (about -0.2 km/sec) to the east of the fault. This indicates that the DD tomography produces a sharper, and

we presume more accurate, velocity contrast. Due to the massive amount of differential data, we cannot practically carry out a full resolution analysis for DD tomography. For example, estimating the resolution matrix using the method of Nolet *et al.* (1999) would require approximately a week of computation time and further would not yield a reliable result because our matrix \mathbf{A} does not satisfy the assumption that \mathbf{AA}^T is diagonally dominant. The current version of tomoDD does not use parameter separation to decouple the hypocenter and structure equations (Pavlis and Booker, 1980) and thus is not able to compute the “separated” resolution matrix. In the future, this approach will be investigated; however, given that for N observed arrival time data there will be of order N^2 differential times, the parameter separation technique will be extremely slow.

Figure 11a shows the across-strike slices through the model resolution structure (the diagonal element of the resolution matrix) from the standard tomography. The resolution for the standard tomography is estimated by inputting the velocity model inverted from DD tomography and using the solution technique in the simul2000 algorithm (Thurber and Eberhart-Phillips, 1999), using only the absolute times.

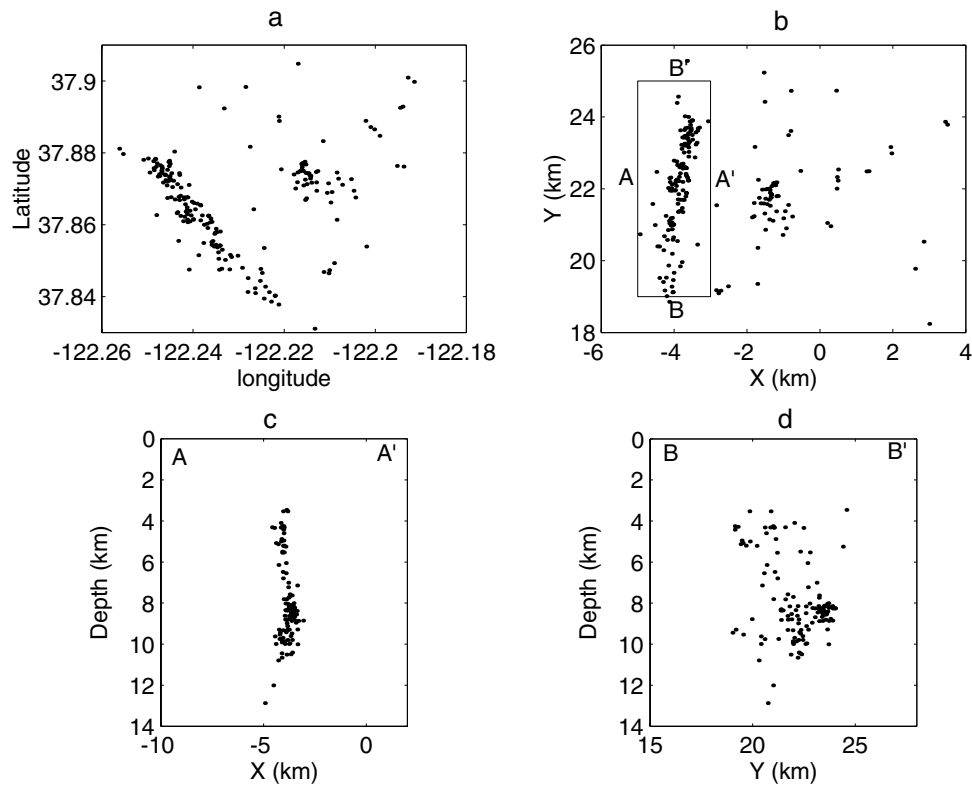


Figure 7. Same as Figure 6, but for the event locations from the standard tomography method (only absolute catalog data).

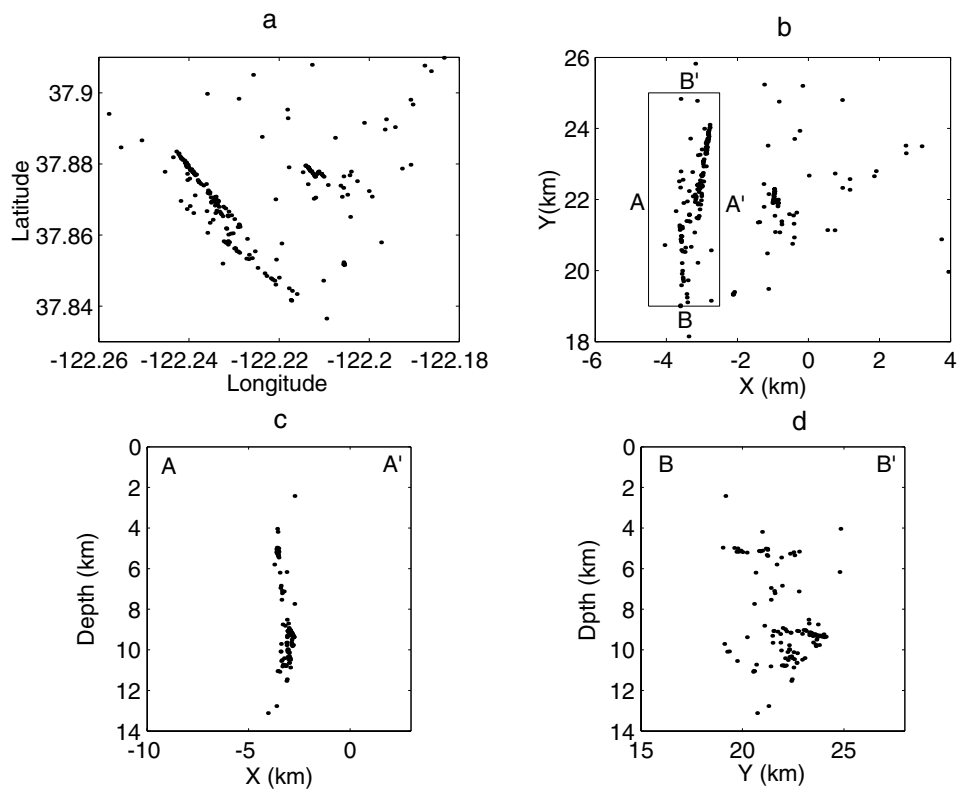


Figure 8. Same as Figure 6, but for the event locations from the DD location method (Waldhauser and Ellsworth, 2002).

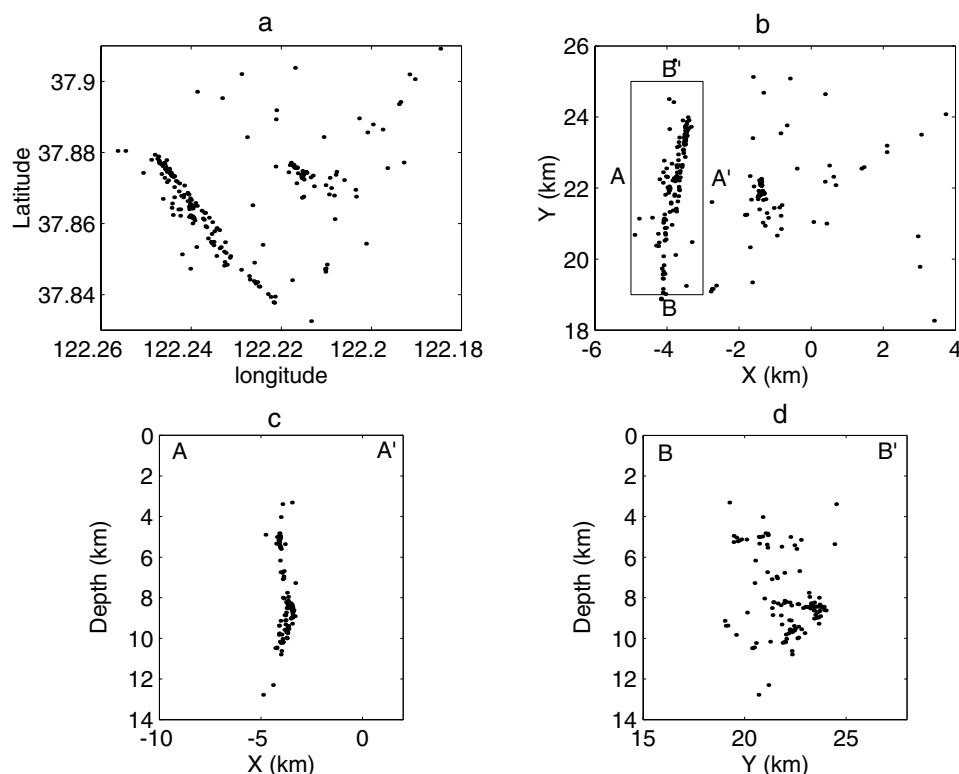


Figure 9. Same as Figure 6, but for the event locations from the DD tomography method.

The model is well resolved (with resolution values > 0.8) except on the model edges, with a noticeable drop in resolution where the finer gridding is used, as expected. Figure 11b gives the $1\text{-}\sigma$ uncertainty estimates of the velocity values from the standard tomography. The errors associated with velocity values are about 0.2 km/sec for most parts of the model and are larger (about $0.3\text{--}0.4\text{ km/sec}$) near the edges of the model. We believe that the resolution and uncertainty results from the standard tomography provide a realistic and conservative estimate of the DD tomography solution quality.

Discussion

Picking errors associated with arrival times generally consist of a combination of correlated and random contributions. The correlated errors include those related to the model or structure errors and picking bias. Using the WCC techniques, both the correlated and random errors are smaller for the differential arrival times than the absolute arrival times. Using the more accurate data, the DD tomography method makes the event locations less scattered than those from the standard tomography method. As a result, this removes one component of the error in the tomography model due to the mislocations. Consequently, this makes the velocity contrast sharper (more accurate).

We also obtain the differential arrival times by directly subtracting catalog arrival times for pairs of events at the

common stations. For two events not too far apart from each other, the velocity heterogeneities along the ray paths have the same effect on the waveforms and the picking bias may thus be correlated, as are the model errors. This differencing process will normally reduce the correlated errors significantly, although it may amplify the random errors. Even with just differential catalog data, Waldhauser and Ellsworth (2000, 2002) have shown that the events are more accurately relocated. A test on the Hayward fault dataset also shows that the events are more concentrated and that the velocity structure using just absolute and differential catalog data is somewhat sharper across the fault compared to standard tomography. This result indicates that the correlated errors are generally larger than the random ones, so the differential catalog data are more accurate than the absolute data overall.

Conclusions

The DD tomography method introduced here is efficient in relocating large numbers of earthquakes accurately as well as characterizing the local velocity structure finely. This method collapses the scattered catalog locations into horizontal lineations of seismicity on the northern Hayward fault, the same as by the DD location method (Waldhauser and Ellsworth, 2002). By including the absolute arrival times in addition to the relative arrival times, our method obtains the absolute event locations without the assumptions made in the DD location method.

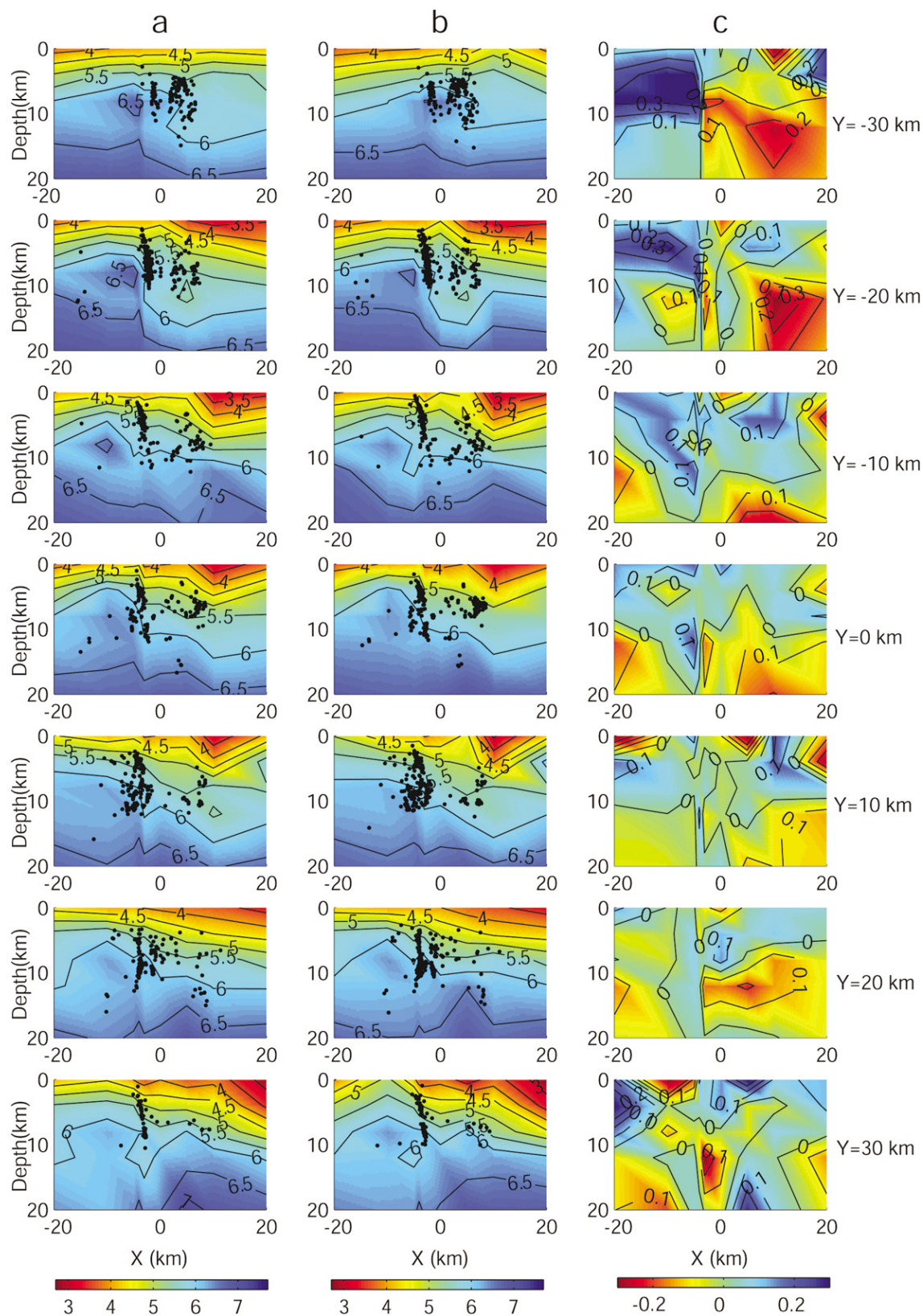


Figure 10. Across-strike vertical slices through (a) the 3D model from DD tomography, (b) the 3D model from standard tomography, (c) the velocity difference between (a) and (b). The final hypocenters within 5 km of each slice are included as black dots.

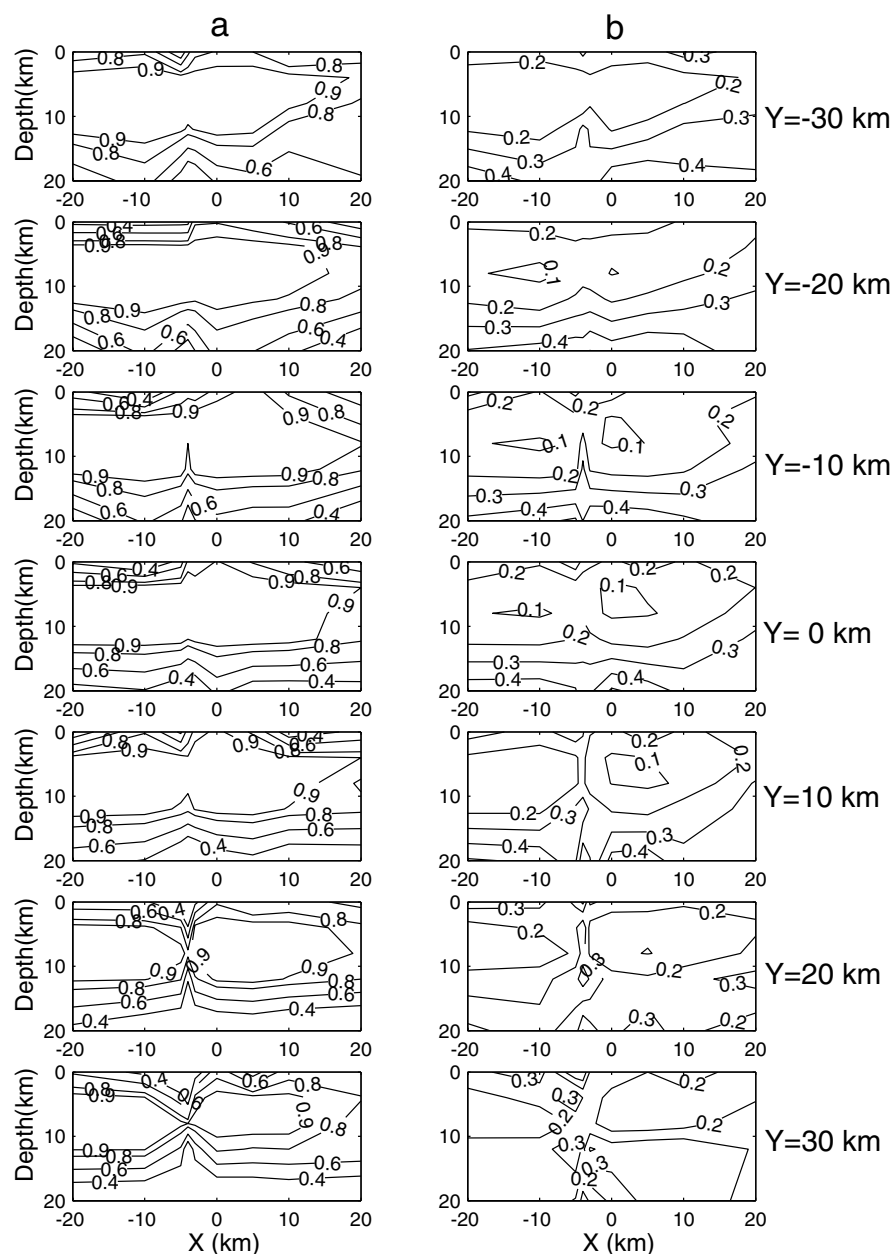


Figure 11. Across-strike vertical slices through (a) the 3D resolution structure for the standard tomography and (b) the model uncertainties.

By using the accurate WCC-derived relative arrival times directly, the DD tomography method is able to sharpen the image of the velocity structure, especially near the source region. In particular, we obtain a sharper velocity contrast along the Hayward fault compared to the standard tomography method.

Acknowledgments

We are extremely grateful to Felix Waldhauser for sharing the Hayward data and for valuable comments on early drafts of this manuscript. Bill Ellsworth provided us with the synthetic data set and suggested the name *tomoDD*. We thank Steve Roecker, Andy Michael, and an anonymous

reviewer for their constructive comments. This research has resulted from a project supported in part by Defense Threat Reduction Agency (contract number DTRA01-01-C-0085), U.S. Department of Defense; the content does not necessarily reflect the position or the policy of the U.S. government, and no official endorsement should be inferred. This material is based upon work supported in part by the National Science Foundation under grant no. EAR-0125164.

References

- Bailey, E. H., W. D. Irwin, and D. L. Jones (1964). Franciscan and related rocks, and their significance in the geology of western California, *Bull. Calif. Div. Mines. Geol.* **183**, 177 pp.
- Constable, S. C., R. L. Parker, and C. G. Constable (1987). Occam's in-

- version: a peractical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics* **52**, 289–300.
- Dietz, L. D., and W. L. Ellsworth (1990). The October, 17, 1989 Loma Prieta, California, earthquake and its aftershocks: geometry of the sequence from high-resolution locations, *Geophys. Res. Lett.* **17**, 1417–1420.
- Dodge, D. A., G. C. Beroza, and W. L. Ellsworth (1995). Foreshock sequence of the 1992 Landers, California, earthquake and its implications for earthquake nucleation, *J. Geophys. Res.* **100**, 9865–9880.
- Dodge, D. A., G. C. Beroza, and W. L. Ellsworth (1996). Detailed observations of California foreshock sequences: implications for the earthquake initiation process, *J. Geophys. Res.* **101**, 22,371–22,392.
- Douglas, A., D. Browsers, and J. B. Young (1997). On the onset of *P* seismograms, *Geophys. J. Int.* **129**, 681–690.
- Fremont, M.-J., and S. D. Malone (1987). High precision relative locations of earthquakes at Mount St. Helens, Washington, *J. Geophys. Res.* **92**, 10,223–10,236.
- Got, J.-L., J. Frechet, and F. W. Klein (1994). Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea, *J. Geophys. Res.* **99**, 15,375–15,386.
- Hole, J. A., T. M. Brocher, S. L. Klemperer, T. Parsons, H. M. Benz, and K. P. Furlong, Three-dimensional seismic velocity structure of the San Francisco Bay area, *J. Geophys. Res.* **105**, 13,859–13,874.
- Jennings, C. W. (1977). Geologic map of California, Calif. Div. of Mines and Geol., Sacramento.
- Kissling, E., W. L. Ellsworth, D. Eberhart-Phillips, and U. Kradolfer (1994). Initial reference models in local earthquake tomography, *J. Geophys. Res.* **99**, 19,635–19,646.
- Lee, W. H. K., and S. W. Stewart (1981). Principles and applications of microearthquake networks, *Adv. Geophys. Suppl.* **2**, 293 pp.
- Lees, J. M. (1998). Multiplet analysis at Coso geothermal, *Bull. Seism. Soc. Am.* **88**, 1127–1143.
- McCaughy, M., and S. C. Singh (1997). Simultaneous velocity and interface tomography of normal-incidence and wide-aperture seismic traveltime data, *Geophys. J. Int.* **131**, 87–99.
- Nadeau, R., M. Antolik, P. A. Johnson, W. Foxall, and T. V. McEvilly (1994). Seismological studies at Parkfield. III. Microearthquake clusters in the study of fault-zone dynamics, *Bull. Seism. Soc. Am.* **84**, 247–263.
- Nolet, G., R. Montelli, and J. Virieux (1999). Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems, *Geophys. J. Int.* **138**, 36–44.
- Paige, C. C., and M. A. Saunders (1982). LSQR: sparse linear equations and least squares problems, *ACM Trans. Math. Softw.* **8**, no. 2, 195–209.
- Pavlis, G. L. (1986). Appraising earthquake hypocenter location errors: a complete practical approach for single-event locations, *Bull. Seism. Soc. Am.* **76**, 1699–1717.
- Pavlis, G. L., and J. R. Booker (1980). The mixed discrete continuous inverse problem: application to the simultaneous determination of earthquake hypocenters and velocity structure, *J. Geophys. Res.* **85**, 4801–4810.
- Phillips, W. S., H. E. Hartse, and L. K. Steck (2001). Precise relative location of 25 ton chemical explosions at Balapan using IMS stations, *Pure Appl. Geophys.* **158**, 173–192.
- Poupinet, G., W. L. Ellsworth, and J. Frechet (1984). Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *J. Geophys. Res.* **89**, 5719–5713.
- Rowe, C. A., R. C. Aster, W. S. Phillips, R. H. Jones, B. Borchers, and M. C. Fehler (2002). Using automated, high-precision repicking to improve delineation of microseismic structures at the Soultz geothermal reservoir, *Pure Appl. Geophys.* **159**, 563–596.
- Rubin, A. M., D. Gillard, and J. L. Got (1998). A reinterpretation of seismicity associated with the January 1983 dike intrusion at Kilauea Volcano, Hawaii, *J. Geophys. Res.* **103**, 10,003–10,015.
- Rubin, A. M., D. Gillard, and J. L. Got (1999). Streaks of microearthquakes along creeping faults, *Nature* (London) **400**, 635–641.
- Sambridge, M. S. (1990). Nonlinear arrival time inversion: constraining velocity anomalies by seeking smooth models in 3-D, *Geophys. J. Int.* **102**, 653–677.
- Shearer, P. M. (1997). Improving local earthquake locations using the L1 norm and waveform cross correlation: application to the Whittier Narrows, California, aftershock sequence, *J. Geophys. Res.* **102**, 8269–8283.
- Thurber, C. H. (1983). Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, *J. Geophys. Res.* **88**, 8226–8236.
- Thurber, C. H. (1992). Hypocenter-velocity structure coupling in local earthquake tomography, *Phys. Earth Planet. Interiors* **75**, 55–62.
- Thurber, C. H., and D. Eberhart-Phillips (1999). Local earthquake tomography with flexible gridding, *Comput. Geosci.* **25**, 809–818.
- Thurber, C. H., C. Trabant, F. Haslinger, and R. Hartog (2001). Nuclear explosion locations at the Balapan, Kazakhstan, nuclear test site: the effects of high-precision arrival times and three-dimensional structure, *Phys. Earth Planet. Interiors* **123**, 283–301.
- Um, J., and C. H. Thurber (1987). A fast algorithm for two-point seismic ray tracing, *Bull. Seism. Soc. Am.* **77**, 972–986.
- VanDecar, J. C., and R. S. Crosson (1990). Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seism. Soc. Am.* **80**, 1548–1560.
- Waldhauser, F. (2001). hypoDD: a computer program to compute double-difference hypocenter locations, *U.S. Geol. Surv. Open File Rep.* **01-113**, 25 pp.
- Waldhauser, F., and W. L. Ellsworth (2000). A double-difference earthquake location algorithm: method and application to the northern Hayward Fault, California, *Bull. Seism. Soc. Am.* **90**, 1353–1368.
- Waldhauser, F., and W. L. Ellsworth (2002). Fault structure and mechanics of the Hayward Fault, California, from double-difference earthquake locations, *J. Geophys. Res.* **107**, ESE 3-1–3-15.
- Waldhauser, F., W. L. Ellsworth, and A. Cole (1999). Slip-parallel seismic lineations on the northern Hayward Fault, California, *Geophys. Res. Lett.* **26**, 3525–3528.
- Walter, A. W. (1990). Upper-crust velocity structure near Coalinga, as determined from seismic-refraction data, in *The Coalinga, California, Earthquake of May 2*, M. J. Rymer and W. L. Ellsworth (Editors), U.S. Geological Survey, Reston, Virginia, 23–29.
- Wolfe, C. J. (2002). On the mathematics of using difference operators to relocate earthquakes, *Bull. Seism. Soc. Am.* **92**, 2879–2892.

Department of Geology and Geophysics
University of Wisconsin–Madison
1215 W. Dayton Street
Madison, Wisconsin 53706

Manuscript received 13 September 2002.