

## Development and Applications of Double-difference Seismic Tomography

HAIJIANG ZHANG,<sup>1</sup> and CLIFFORD THURBER<sup>1</sup>

*Abstract*—Double-difference (DD) tomography is a generalization of DD location; it simultaneously solves for the three-dimensional velocity structure and seismic event locations. DD tomography uses a combination of absolute and more accurate differential arrival times and hierarchically determines the velocity structure from larger scale to smaller scale. This method is able to produce more accurate event locations and velocity structure near the source region than standard tomography, which uses only absolute arrival times. We conduct a stability and uncertainty analysis of DD tomography based on a synthetic data set. Currently three versions of the DD tomography algorithms exist: *tomoDD*, *tomoFDD* and *tomoADD*. *TomoDD* assumes a flat earth model and uses a pseudo-bending ray-tracing algorithm to find rays between events and stations while *tomoFDD* uses a finite-difference travel-time algorithm and the curvature of the Earth is considered. Both codes are based on a regularly distributed inversion grid, with the former for a local scale and the latter for a regional scale. In contrast, *tomoADD* adapts the inversion mesh to match with the data distribution based on tetrahedral and Voronoi diagrams. We discuss examples of applying DD tomography to characterize fault zone structure, image high-resolution structure of subduction zones, and determine the velocity structure of volcanoes.

**Key words:** Double-difference, tomography, fault zone, subduction zone, volcano.

### *Introduction*

We have recently developed a double-difference (DD) seismic tomography method that makes use of both absolute and more accurate relative arrival times (ZHANG and THURBER, 2003). The differential times can be calculated from cross-correlation (CC) techniques for similar waveforms and by directly subtracting catalog arrival times for pairs of events at common stations (WALDHAUSER and ELLSWORTH, 2000; ZHANG and THURBER, 2003). DD tomography is a generalization of DD location (WALDHAUSER and ELLSWORTH, 2000); it simultaneously solves for the three-dimensional (3-D) velocity structure and seismic event locations. DD tomography uses an evolving weighting scheme for the absolute and differential arrival times in order to determine the velocity structure from larger scale to smaller

---

<sup>1</sup>Department of Geology and Geophysics, University of Wisconsin-Madison, 1215 W. Dayton St., Madison, WI, 53706, USA. E-mail: hjzhang@geology.wisc.edu

scale. As shown by synthetic tests (ZHANG and THURBER, 2003, 2005; ZHANG *et al.*, 2004), this method yields more accurate event locations and velocity structure near the source region than standard tomography, which uses only absolute arrival times. It has the unique ability to sharpen the velocity image near the source region because of the combination of the higher accuracy of the differential time data and the concentration of the corresponding model derivatives in the source region. The latter results from the cancellation of model derivative terms where the ray paths overlap away from the source region.

In many situations, CC data are not available due to the lack of waveform data. There have been many applications using both DD location and tomography algorithms demonstrating that using just the catalog pick differences leads to significant improvement of event locations, as well as for the velocity structure for DD tomography. As discussed in ZHANG and THURBER (2003) and ZHANG *et al.* (2005), picking errors include two parts: random and systematic. Although random errors may increase through the differencing process, the systematic errors will be reduced or cancelled. The differencing process also causes the ray-path derivatives to concentrate near the source region, rather than covering the entire model region. This in turn reduces or removes the effect of the velocity model ambiguity outside the source region on the source region model. As a result, the source region structure will be better illuminated.

The DD tomography code *tomoDD* is built upon the double-difference location code *hypoDD* written by WALDHAUSER (2001). In the original *tomoDD* algorithm, we used an approximate pseudo-bending (ART-PB) ray-tracing algorithm (UM and THURBER, 1987) to find the rays and calculate the travel times between events and stations. The model is represented by velocity values specified on a regular set of 3-D nodes and the velocity values are interpolated by using the linear B-spline 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).

*TomoDD* assumes a flat earth model and is appropriate for local scale problems (10's to 100's of kilometers). At the regional scale (100's to 1000's of kilometers), however, sphericity of the earth should be taken into account. Major velocity discontinuities such as Conrad, Moho, and subducting slab boundary should also be considered. The ART-PB approach assumes a continuous velocity model and cannot deal properly with velocity discontinuities. KOKETSU and SEKINE (1998) have developed a spherical-earth 3-D ray-tracing method (including discontinuities) using pseudo-bending, extended from a method developed in Cartesian coordinates (ZHAO *et al.*, 1992). ZHAO *et al.* (1992) and KOKETSU and

SEKINE (1998) explicitly took into account velocity discontinuities by applying Snell's law to find the intersecting point between a ray path and a discontinuity. These methods have advantages over the original ART-PB approach if the positions of velocity discontinuities are known *a priori* from other independent studies. However, these velocity discontinuities are seldom known *a priori* or are not well constrained in many subduction zones. For this reason, we developed a regional DD seismic tomography method (*tomoFDD*) that deals effectively with discontinuous velocity structures without knowing them *a priori*. *TomoFDD* uses a finite-difference method for determining travel times and ray paths, and treats the spherical Earth by embedding it (in part or in whole) within a Cartesian "box." For the local and regional-scale problems, the ray-tracing accuracy for both pseudo-bending and finite-difference algorithms is carefully controlled to match with the higher accuracy of the differential times. For example, for pseudo-bending ray-tracing algorithm we do not stop the iteration finding rays until the travel-time differences between two close iterations are smaller than 1 ms.

The current versions of *tomoDD* and *tomoFDD* use a regular inversion grid. For typical seismic tomography studies, however, the ray distribution is highly uneven due to nonuniform station geometry, uneven distribution of seismic sources, missing data, and ray bending. Some nodes or cells may have few or even no rays sampling them, while others may have very dense rays sampling them. The regular grid spacing restriction makes it difficult to adapt the model to the uneven distribution of ray paths. The mismatch between the ray distribution and the grid chosen for the tomographic inversion results in instability of the inversion. To make the inversion more stable, some regularization method is required, such as damping or smoothing, which will inevitably bias the results. Ideally the inversion grid (or mesh) should be distributed adaptively to match with the resolving power of the data. The inversion problem will then be better conditioned and smaller damping and/or reduced smoothing constraints would thus be required. Our recently developed adaptive-mesh DD seismic tomography method (*tomoADD*), based on tetrahedral and Voronoi diagrams, can automatically match the inversion mesh to the data distribution (ZHANG and THURBER, 2005).

The following material first reviews the DD tomography method and conducts a stability and uncertainty analysis based on a simple 2-D synthetic model. Then we present the regional-scale and adaptive-mesh DD tomography methods, and discuss some of their applications for a variety of regions.

### *The Double-difference Tomography Method*

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  $\tau^i$  is the origin time of event  $i$ ,  $u$  is the slowness field and  $ds$  is an element of path length. For local earthquake tomography, 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 nonlinear, thus a truncated Taylor series expansion is generally used to linearize eq. (1). If we also discretize the velocity model using a 3-D grid or mesh, we can write a linear equation relating the misfit between the observed and predicted arrival times  $r_k^i$  to the desired perturbations to the hypocenter and velocity structure parameters

$$r_k^i = \sum_{l=1}^3 \frac{\partial T_k^i}{\partial x_l^i} \Delta x_l^i + \Delta \tau^i + \sum_{m=1}^{M_{ik}} \sum_{n=1}^N w_{mn} \Delta u_n \Delta s_m, \quad (2)$$

where  $M_{ik}$  indicates the number of segments of the ray path from event  $i$  to station  $k$  and  $w_{mn}$  are interpolating weights of  $n^{\text{th}}$  mesh node on the mid-point of the  $m^{\text{th}}$  segment with length  $\Delta s_m$ . Subtracting a similar equation for event  $j$  observed at station  $k$  from equation (2), we have

$$\begin{aligned} r_k^i - r_k^j = & \sum_{l=1}^3 \frac{\partial T_k^i}{\partial x_l^i} \Delta x_l^i + \Delta \tau^i + \sum_{m=1}^{M_{ik}} \sum_{n=1}^N w_{mn} \Delta u_n \Delta s_m - \sum_{l=1}^3 \frac{\partial T_k^j}{\partial x_l^j} \Delta x_l^j - \Delta \tau^j \\ & - \sum_{m=1}^{M_{jk}} \sum_{n=1}^N w_{mn} \Delta u_n \Delta s_m, \end{aligned} \quad (3)$$

where  $r_k^i - r_k^j$  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

$$r_k^i - r_k^j = (T_k^i - T_k^j)^{\text{obs}} - (T_k^i - T_k^j)^{\text{cal}}. \quad (4)$$

The observed differential arrival times  $(T_k^i - T_k^j)^{\text{obs}}$  can be calculated from both waveform cross-correlation techniques for similar waveforms and absolute catalog arrival times.

Note that the ray paths from two nearby events will substantially overlap, meaning that the model derivative terms in eq. (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. To combine absolute and differential systems together, we apply a hierarchical weighting scheme during the inversion to apply the appropriate relative weighting between them at different stages of inversion. We start the inversion by applying higher weighting 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) in

a manner similar to *hypoDD* (WALDHAUSER, 2001). Then for later iterations 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 waveform cross-correlation data are available, they will be weighted more heavily than the catalog differential data in the final iterations 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, because the cross-correlation data are one to two orders of magnitude more precise than the manual picks (WALDHAUSER and ELLSWORTH, 2000). The relationship between arrival time residuals and slowness model parameters is more linear than the event locations (THURBER, 1992), as can be seen from eq. (1) where the integral in the equation is linear in slowness but nonlinear in the hypocentral coordinates. This indicates that the convergence of velocity structure is faster than event locations (THURBER, 1992). Consequently, the simultaneous inversion of event locations and velocity structure alternates with the inversion of only event locations in practice.

WOLFE (2002) carried out a strict and complete analysis of the use of difference operators to relocate earthquakes. Next we conduct a similar analysis on using the absolute data (i.e., standard location and tomography), the differential data (i.e., DD location and tomography), and both of them to determine event locations and/or velocity structure. Most of the notation used in this section follows WOLFE (2002).

Consider a set of  $p = 1, \dots, P$  earthquakes, with  $N_p$  arrival times for each earthquake. For each individual earthquake, eq. (2) can be represented in matrix form as follows,

$$\mathbf{A}_p \Delta \mathbf{X}_p + \mathbf{C}_p \Delta \mathbf{M} = \Delta \mathbf{T}_p, \quad (5)$$

where  $\mathbf{A}_p (N_p \times 4)$  is the partial derivative matrix corresponding to the hypocenter and origin time,  $\Delta \mathbf{X}_p (4 \times 1)$  is the perturbation vector for earthquake location and origin time,  $\mathbf{C}_p (N_p \times L)$  is the model derivative (path length times node weight) matrix corresponding to the slowness model,  $\Delta \mathbf{M} (L \times 1)$  is the vector of slowness perturbations, and  $\Delta \mathbf{T}_p (N_p \times 1)$  is the vector of arrival time residuals. A station term may be also included to take into account the velocity heterogeneity near the stations, which may not be well resolved during the inversion, as follows,

$$\mathbf{A}_p \Delta \mathbf{x}_p + \mathbf{C}_p \Delta \mathbf{M} + \mathbf{s}_p = \Delta \mathbf{T}_p, \quad (6)$$

where  $\mathbf{s}_p (N_p \times 1)$  is the station correction vector (note that it is defined as the path anomaly in WOLFE (2002)).

Define  $N_T (= \sum_{p=1}^P N_p)$  as the total number of arrival time data and  $M_T$  as the total number of unknown hypocenter and origin parameters ( $= 4 \times P$ ). We can combine all the equations for  $P$  earthquakes into one linear system, as follows,

$$\mathbf{A}\Delta\mathbf{X} + \mathbf{C}\Delta\mathbf{M} + \mathbf{S} = \Delta\mathbf{T}, \quad (7)$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \bullet & \bullet & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \bullet & \bullet & \mathbf{0} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \mathbf{0} & \mathbf{0} & \bullet & \bullet & \mathbf{A}_p \end{bmatrix}; N_T \times M_T \quad (8)$$

$$\Delta\mathbf{T} = \begin{bmatrix} \Delta T_1 \\ \Delta T_2 \\ \bullet \\ \Delta T_p \end{bmatrix}; N_T \times 1 \quad (9)$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} & \bullet & \bullet & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 & \bullet & \bullet & \mathbf{0} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \mathbf{0} & \mathbf{0} & \bullet & \bullet & \mathbf{C}_p \end{bmatrix}; N_T \times L \quad (10)$$

$$\Delta\mathbf{M} = \begin{bmatrix} \Delta m_1 \\ \Delta m_2 \\ \bullet \\ \Delta m_L \end{bmatrix}; L \times 1 \quad (11)$$

$$\mathbf{S} = \begin{bmatrix} s_1 \\ s_2 \\ \bullet \\ s_p \end{bmatrix}; N_T \times 1. \quad (12)$$

For station corrections, we assume that each is a constant for all the arrivals observed at a given station. As a result, let  $\mathbf{s}_0$  stand for the station corrections for the  $K_T$  stations, which can be represented as

$$\mathbf{S} = \mathbf{B}\mathbf{s}_0, \quad (13)$$

where

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \bullet \\ \mathbf{B}_p \end{bmatrix}; N_T \times K_T, \quad (14)$$

and each  $\mathbf{B}_p$  is a  $N_p \times K_T$  matrix such that

$$[B_p]_{ij} = \begin{cases} 1 & \text{when } [\Delta T_p]_i \text{ is from station } j \\ 0 & \text{otherwise} \end{cases}. \quad (15)$$

The DD tomography eq. (3) is equivalent to the application of a difference operator  $\mathbf{Q}_{DD}$  to equation (7), where

$$\mathbf{Q}_{DD} = \begin{bmatrix} 1 & -1 & \bullet & \bullet & \bullet & 0 \\ 1 & \bullet & \bullet & -1 & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \bullet & 1 & \bullet & \bullet & -1 \end{bmatrix}. \quad (16)$$

Let  $w_k$  equal the number of earthquakes observed at station  $k$ , then the matrix  $\mathbf{Q}_{DD}$  has the dimension of  $\sum_{k=1}^{K_T} \frac{w_k(w_k-1)}{2} \times N_T$ , corresponding to the possible combinations of differences of earthquake arrival times at a station  $k$ ,  $\Delta T_k^i - \Delta T_k^j$  (for  $i < j$ , to eliminate repetition).

It can be seen that the difference operator  $\mathbf{Q}_{DD}$  will annihilate the station corrections as a result of  $\mathbf{Q}_{DD}\mathbf{B}s_0 = \mathbf{0}$ . This is an advantage when comparing DD tomography with standard tomography. The matrix form for the differential time data of DD tomography is

$$\mathbf{Q}_{DD}\mathbf{A}\Delta\mathbf{X} + \mathbf{Q}_{DD}\mathbf{C}\Delta\mathbf{M} = \mathbf{Q}_{DD}\Delta\mathbf{T}. \quad (17)$$

The difference operator  $\mathbf{Q}_{DD}$  has the rank of  $\sum_{k=1}^{K_T} (w_k - 1)$ , because the operation removes one degree of freedom from the arrival times observed at one station (WOLFE, 2002). This rank is usually considerably larger than the total number of unknown parameters, including event hypocenter and origin time parameters and model (slowness) parameters. If both matrices  $\mathbf{Q}_{DD}\mathbf{A}$  and  $\mathbf{Q}_{DD}\mathbf{C}$  have full column rank, the DD tomography system will have the ability to determine absolute event locations (MENKE and SCHAFF, 2004) and slowness (or velocity) structure uniquely. Although the hypocenter partial derivatives for closely spaced events are very similar, they will be quite different for events spaced far apart. Thus, in theory, if we include all the possible combinations of pairs of arrival times observed at common stations, the matrix  $\mathbf{Q}_{DD}\mathbf{A}$  should have full column rank.  $\mathbf{Q}_{DD}\mathbf{C}$  should also have full column rank if we remove slowness model parameters not sampled by ray paths from the inversion.

For the DD location algorithm in which the path anomaly biases between events are not taken into account explicitly, only closely spaced events should be chosen in order to reduce the path anomaly bias. Distance weighting is applied to reduce the effect of event pairs that are far apart (WALDHAUSER and ELLSWORTH, 2000). However, for closely spaced events, the partial derivatives will be very similar or equal, thus the DD location algorithm is only capable of resolving relative locations between events (WOLFE, 2002). For the DD tomography system, since the path anomaly biases between event pairs are taken into account explicitly, we do not have to limit the inversion only to the closely spaced events. Distance weighting is an option in the DD tomography algorithm and is mainly designed to downweight differential times resulting from far apart events based on the fact that the similarity

of waveforms is less for far apart events, as are the common errors associated with picks.

Equation (17) can be further transformed into a more compact form, as follows,

$$\mathbf{Q}_{DD}\mathbf{E}\mathbf{Y} = \mathbf{Q}_{DD}\Delta\mathbf{T}, \quad (18)$$

where  $\mathbf{E} = [\mathbf{A} \ \mathbf{C}]$  has the dimension of  $N_T \times (M_T + L)$ , and  $\mathbf{Y} = \begin{bmatrix} \Delta\mathbf{X} \\ \Delta\mathbf{M} \end{bmatrix}$  has the dimension of  $(M_T + L) \times 1$ . The unknown parameters including hypocenters, origin times, and slowness model parameters can be obtained, as follows,

$$\mathbf{Y} = (\mathbf{Q}_{DD}\mathbf{E})^{-1}\mathbf{Q}_{DD}\Delta\mathbf{T}. \quad (19)$$

For origin times, since the associated partial derivatives are the same (equal to 1), only relative origin times can be recovered from this system (WOLFE, 2002).

If the differential arrival times are used to determine event locations only (WALDHAUSER and ELLSWORTH, 2000), and the hypocenter partial derivatives are evaluated at different locations, the absolute event locations can potentially be resolved, as follows (WOLFE, 2002),

$$\Delta\mathbf{X} = (\mathbf{Q}_{DD}\mathbf{A})^{-1}\mathbf{Q}_{DD}\Delta\mathbf{T}. \quad (20)$$

The partial derivative operator used in DD location is  $\mathbf{Q}_{DD}\mathbf{A}$ . Note that a row of the matrix  $(\mathbf{Q}_{DD}\mathbf{A})^{-1}\mathbf{Q}_{DD}$  can be used to indicate the dependence of the corresponding model parameter (location or origin time in this case) on all the travel times.

The perturbations to event locations can also be obtained from absolute arrival times by assuming the velocity structure is known and presumably correct,

$$\Delta\mathbf{X} = \mathbf{A}^{-1}\Delta\mathbf{T}. \quad (21)$$

Similarly, the system to determine event locations and velocity structure from absolute arrival times (standard tomography) is represented as

$$\mathbf{Y} = \mathbf{E}^{-1}\Delta\mathbf{T}. \quad (22)$$

For the combined system that uses both absolute and differential data, the formula to determine event locations is

$$\Delta\mathbf{X} = \begin{pmatrix} \mathbf{Q}_{DD}\mathbf{A} \\ w\mathbf{A} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Q}_{DD} \\ w\mathbf{I} \end{pmatrix} \Delta\mathbf{T}, \quad (23)$$

where  $w$  is the relative weighting between absolute and relative arrival times, and  $\mathbf{I}$  is the identity matrix. The complete formula to solve the velocity structure simultaneously with event locations by using both absolute and differential data is

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Q}_{DD}\mathbf{E} \\ w\mathbf{E} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Q}_{DD} \\ w\mathbf{I} \end{pmatrix} \Delta\mathbf{T}. \quad (24)$$

For simplicity, we construct a 2-D synthetic model (Fig. 1) to show the properties of these different systems of equations. There are 6 clusters of earthquakes (circles) and each cluster has 4 closely spaced events with inter-event distances of less than 200 m. In the X and Y directions, the distances between clusters are 10 and 15 km, respectively. There are 36 stations around these events, with all the events observed at each station. As a result, there are 894 travel times in total for all the event and station pairs. The velocity is assumed to be constant at 5 km/s, thus the rays between events and stations are all straight lines. The region is divided uniformly into 81 cells (or blocks). The model derivatives relative to slowness are just the path lengths crossing the cells.

Since this is relatively a small data set, we can use the singular value decomposition method to calculate the singular values, model resolution and

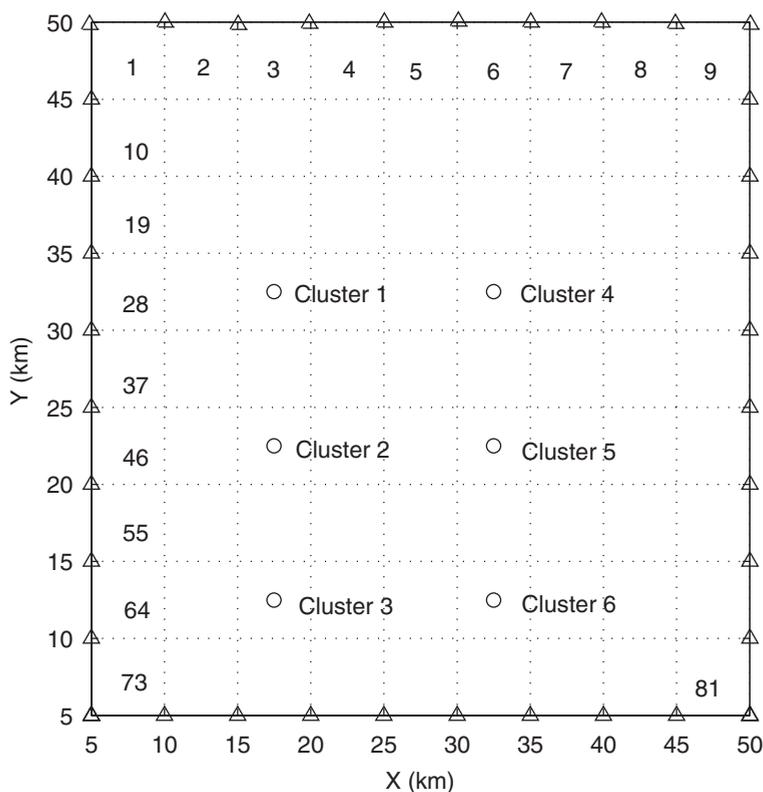


Figure 1

Map of stations (triangles) and earthquake clusters (circles) for the synthetic example. Each cluster has 4 closely spaced events, with interevent distances among them less than 200 m. The velocity model is assumed to be a constant 5 km/s. The model region is divided into 81 cells uniformly, with numbers in the cells indicating the model parameter indices.

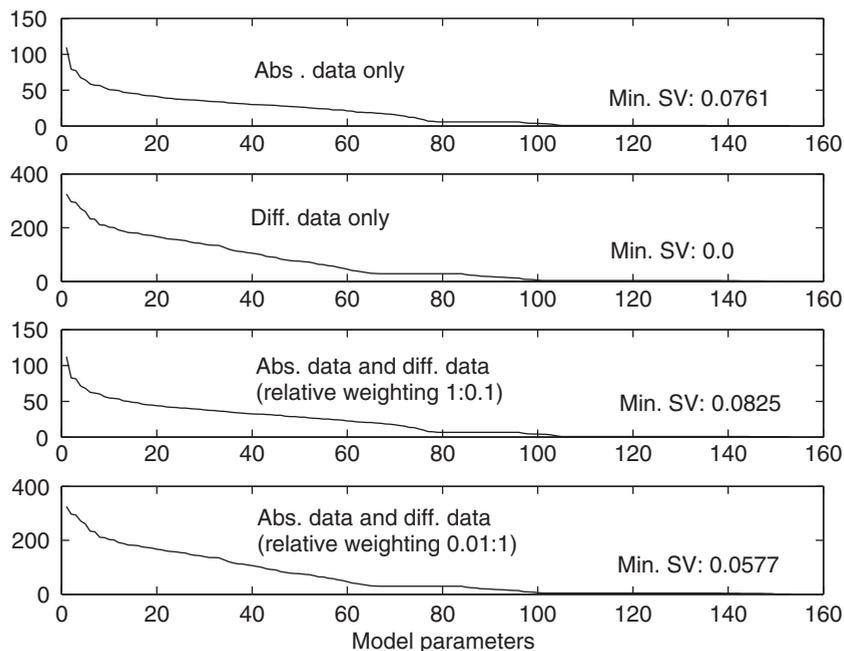


Figure 2

The singular value distribution of the partial derivative matrix from using different data types to determine event locations and velocity structure simultaneously.

uncertainties for the different systems (MENKE, 1989; ASTER *et al.*, 2005). Figure 2 shows the singular values for the operators  $\mathbf{E}$ ,  $\mathbf{Q}_{DD}\mathbf{E}$ ,  $(\mathbf{E}, 0.1\mathbf{Q}_{DD}\mathbf{E})^T$ , and  $(0.01\mathbf{E}, \mathbf{Q}_{DD}\mathbf{E})^T$  when using the absolute data only, the differential data only, and both of them (with relative weighting 1:0.1 and 0.01:1), respectively, to jointly determine event locations and slowness structure. The system using the differential data to jointly determine event locations and slowness structure has one zero singular value (Fig. 2). The zero singular value is due to the fact that the partial derivatives relative to the origin time for all events are equal and the differential operator has no ability to recover the absolute origin times, only relative ones. Including the absolute data with only a small weight is helpful to stabilize the system, by increasing the minimum singular value to 0.0577 (Fig. 2). This indicates that including the absolute data into the differential system gives it the ability to resolve the absolute origin times, as would be expected. Also notice that the largest singular value from the system having higher weighting on the differential data is about 2 times greater than that from the system having higher weighting on the absolute data (Fig. 2). To resolve these systems by using the damped least-squares algorithm, different damping values should be chosen to make their condition numbers similar. The damping value that makes the condition number of the system close to  $\sim 60$  to  $80$  is consistent with the

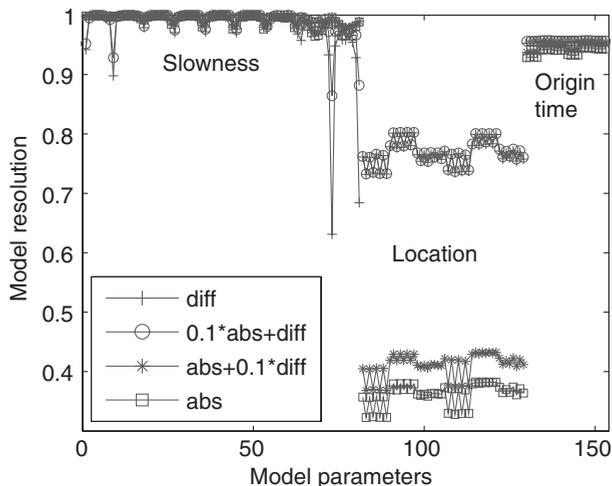


Figure 3

Model resolution for location and slowness model parameters when using different data types to determine event locations and velocity structure simultaneously by a damped least squares solution. The damping values are 1.5 for the systems of  $\text{abs}$  and  $\text{abs} + 0.1\text{diff}$ , and 4.4 for the systems of  $\text{diff}$  and  $\text{diff} + 0.01\text{abs}$ . For the model parameters, the slowness model cells are arranged at first in the sequence as shown in Figure 1, followed by the location and origin time parameters.

one chosen based on the trade-off curve between data variance and model norm. For this reason, we choose damping values of 1.5 and 4.4 for the systems having higher weighting on the absolute data and higher weighting on the differential data, which would make all the systems have condition numbers  $\sim 70$ .

Figure 3 shows model resolution for origin time, location and slowness model parameters. The systems with higher weighting on the differential data seem to better resolve event locations by having higher resolution values ( $\sim 0.7$  versus  $\sim 0.3$ ). For slowness model parameters around the edge of the model, the systems having higher weighting on the absolute data have higher model resolution. We also note that the model resolution is generally higher for the slowness parameters than location parameters. By carefully checking the singular values of the Fréchet derivative matrices for slowness and locations parameters, we found that the largest singular value for the slowness derivative matrix is considerably larger than that for the location derivative matrix. For example, the largest singular value for the slowness matrix in the case of using just differential times is 325.3, compared to 29.5 for the location matrix. Since we use the same damping value of 4.4 for slowness and location parameters, the latter ones are overdamped. This is the reason why the model resolution for slowness parameters is greater than location parameters. This overdamping for location parameters is justified since the relationship between arrival time residuals and hypocenter perturbations is nonlinear. This is helpful for

preventing large data residuals (or noise) from projecting into location parameters. We also note that events are generally relocated one more time based on the final velocity model.

We assume that the errors associated with both absolute and differential data are an uncorrelated Gaussian-distributed random process. The standard deviations for the absolute and differential data are 80 and 20 ms, respectively. This simulates the scenario that the differential data are more accurate than the absolute data. These values would be representative of absolute catalog picks and differential catalog picks; using the even more precise CC differential times would obviously make the error contrast greater. Figure 4 shows the model uncertainties (s/km for slowness parameters and km for locations) for different systems. The model uncertainties estimated from the system having higher weighting on the absolute data are substantially larger, especially for event locations and origin times. For all the systems, the model uncertainties associated with origin times are much smaller than those for hypocenters. This is caused by the different dimensionality between the origin time and hypocenter location (KLEIN, 1978). For example, a change of several km in hypocenter location is equivalent to a one second change in origin time. We also notice that the slowness model cells closer to two model region edges ( $X = 5$  and 50 km) tend to have larger model uncertainties. This phenomenon results from the relatively fewer rays crossing through those cells. The characteristics of model uncertainties having higher weighting on the differential data (diff and

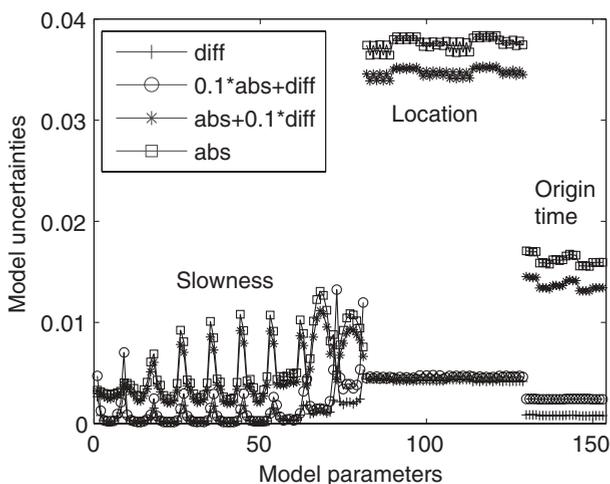


Figure 4

Model uncertainties for location and slowness model parameters when using different data types to determine event locations and velocity structure simultaneously by a damped least squares solution. The damping values for different systems are shown in Figure 3. The model parameters are arranged in the same way as Figure 3.

$\text{diff} + 0.01 \cdot \text{abs}$ ) are quite similar, with the system ( $\text{diff} + 0.01 \cdot \text{abs}$ ) using the combination of data types having slightly larger uncertainties for some slowness model cells.

These results demonstrate the advantages of using both absolute and differential data. The model resolution for location model parameters is higher in the systems having higher weighting on the differential data. The system using both the absolute and differential data with the relative weighting 0.01 to 1 between them has higher model resolution and lower model uncertainties for location parameters than the system using the absolute data only. For some model-edge slowness model cells, the systems having higher weighting ( $\text{abs}$  and  $\text{abs} + 0.1 \cdot \text{diff}$ ) on the absolute data have higher model resolution and lower model uncertainties.

### *Regional-scale Double-difference Tomography (tomoFDD)*

The regional-scale DD tomography code *tomoFDD* uses the finite-difference travel-time algorithms developed by PODVIN and LECOMTE (1991) and HOLE and ZELT (1995) (modified from VIDALE 1990) to accurately calculate travel times in the presence of severe velocity changes and discontinuities. The algorithms solve a finite-difference approximation to the Eikonal equation on a regularly gridded velocity model through a systematic application of Huygens' principle. Both algorithms repeatedly compute the travel times at the grid points that are adjacent to those with known (previously computed) travel times. This procedure can explicitly take into account different propagation modes including transmitted and diffracted body waves and head waves, in addition to direct waves.

For the finite-difference travel-time methods, the station can be regarded as a source and travel times can be calculated from this station to all the grid nodes. According to the principle of reciprocity, the travel times from the event to the station are equal to those from the station to the event. After the travel-time field for each station is calculated, the travel time from an event to this station can be interpolated by the linear B-spline interpolation method through 8 neighboring grid nodes near this event. This method is more efficient than other ray-tracing methods that trace rays from a single event to a large number of stations in the case that there are far more events than stations.

The errors that arise from finite-difference ray-tracing methods mainly depend on the grid interval of the computation grid. The smaller the grid interval is, the more accurate the travel-time field will be. However, the computational efficiency is proportional to the size of the computation grid. Thus there exists a tradeoff between the computational efficiency and accuracy. In the finite-difference travel-time algorithms, it is assumed that the wavefronts are planar, which is not valid in the region near the source where the wavefronts have significant curvature. Normally the rays between the source and the grid nodes in the source region are assumed to be

straight lines, and the travel-times are calculated based on these straight lines. When the grid interval is too large, the errors associated with this simplification will be considerable. PODVIN and LECOMTE (1991) used a recursive scheme to reduce this wavefront distortion to some extent. In comparison, the finite-difference scheme

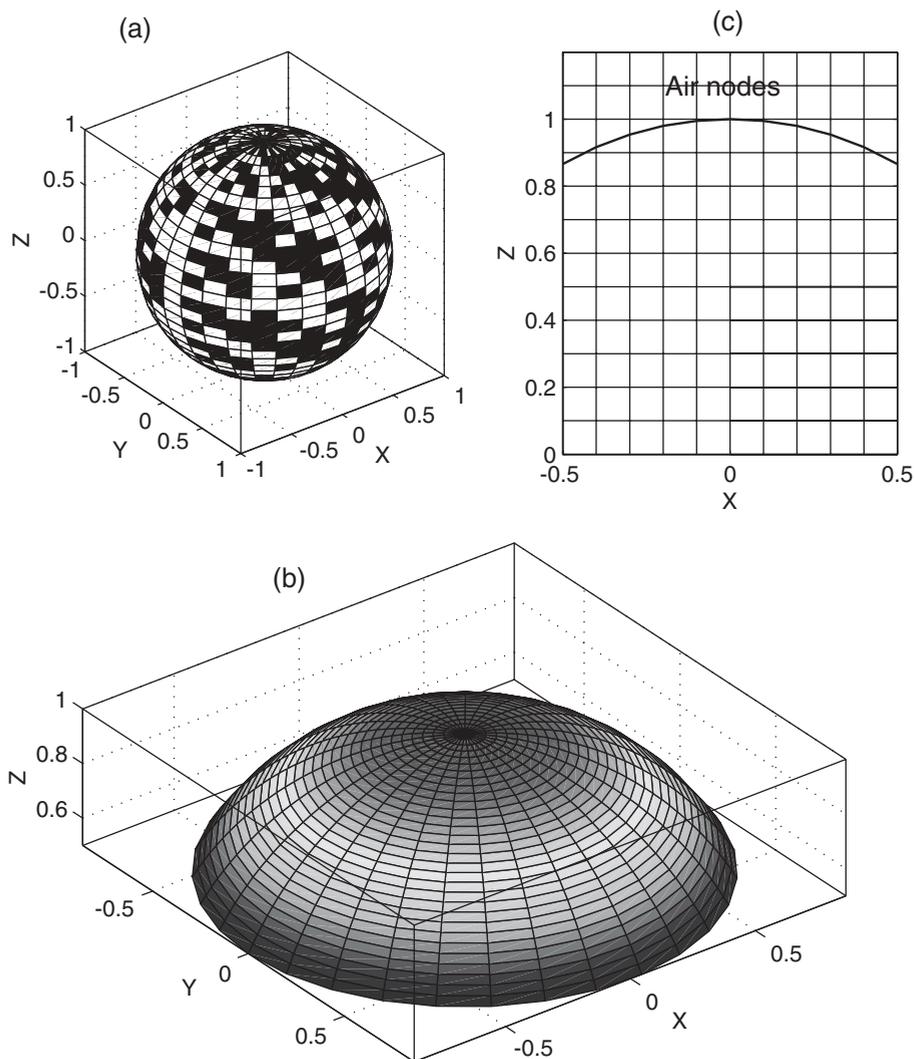


Figure 5

Grid setup for regional scale double-difference tomography. (a) The whole earth is inserted into a cubic box. (b) A rectangular box covers the region of interest. (c) The representation of (b) in the X-Z plane. The surface of the Earth is shown as a thick line. A regular grid is used for the finite-difference travel-time algorithm.

proposed by HOLE and ZELT (1995) does not have such a scheme to deal with the wavefront distortion near the source region.

Recently, ROECKER (pers. comm., 2003) proposed to calculate the travel-time field in the source region with a much finer grid, and then combine this source region travel-time field with that outside the source region. In this way, the effect of wavefront curvature in the source region will be greatly reduced and the accuracy in the travel-time field will also be improved without sacrificing too much computational time and memory space. This strategy will be incorporated in future versions of *tomoFDD*.

To adapt DD tomography to the regional scale, the curvature of the Earth must be taken into account. Following FLANAGAN *et al.* (2000), we solve this problem by parameterizing a spherical surface inside a Cartesian volume of grid nodes. Figure 5a shows the case of putting the Earth into a cube; this model can be used for global seismic tomography. For regional seismic tomography, we are only interested in a part of the Earth, so we construct a rectangular box covering the part of interest (Figs. 5b and c). The coordinate center is placed at the surface of the Earth, the positive X and Y axes point in the directions of North and West, respectively, and the positive Z axis points downward. The grid nodes above the Earth's surface (air nodes) are given the velocity for P waves traveling in air (Fig. 5c). As a result, all the rays travel inside the Earth.

There are two different grids in the algorithm: one is the inversion grid that can be nonuniform but regular and the other is the computational grid that must be uniform. The uniform grid is used to calculate the travel-time field using the finite-difference algorithms and the velocity values at its nodes are interpolated from the non-uniform inversion grid through linear B-spline interpolation. If an inversion grid node is 2 or more km above the Earth's surface (allowing for topography), then it is treated as an air node and its value is fixed during the inversion. First we treat each station as a source and calculate travel times to all velocity nodes in the volume. The travel time from a station to each earthquake is interpolated from its 8 neighboring nodes through linear B-spline interpolation. The ray path from the earthquake to the station is found iteratively with increments opposite to the travel-time gradient. After these set-ups, the ray paths and the related partial derivatives are calculated by the finite-difference method. The code has the option to choose the finite-difference method either by PODVIN and LECOMTE (1991) or by Hole and ZELT (1995). In our experience, the former method treats the source region problem better than the latter method, but it calculates the travel-time field more slowly. Our future work is to incorporate Roecker's multi-grid method (pers. Comm., 2003) into the *tomoFDD* algorithm to better deal with the source region problem.

#### *Adaptive-mesh DD Tomography (tomoADD)*

The algorithms *tomoDD* and *tomoFDD* have already proven to be valuable tools for tomography applications (see next section), but both of them are based on a

regular inversion grid. We recently developed an adaptive-mesh DD tomography method based on tetrahedral and Voronoi diagrams to automatically match the inversion mesh to the data distribution (ZHANG and THURBER, 2005). Both tetrahedral and Voronoi diagrams have been shown to be powerful tools to represent spatial relationships in three dimensions and they allow a more flexible representation of the model. For example, model volumes of widely varying sizes with complex distributions are easily implemented, and it is more convenient to build parameterizations containing particular interfaces, on which the nodes can be distributed.

Linear and natural-neighbor interpolation methods can be derived for tetrahedral and Voronoi diagrams, respectively. Linear interpolation uses 4 tetrahedron nodes to interpolate the velocity/slowness values at any point and it has the advantage of calculating the interpolating basis functions easily and quickly. However, the linear interpolation function is not continuously differentiable, which is a desired property for some applications. In comparison, the natural-neighbor interpolation method interpolates the value at any point from its  $n$  natural neighbors. The natural-neighbor interpolation function guarantees continuity in first and second derivatives except at the nodes (SAMBRIDGE *et al.*, 1995). Our algorithm allows the use of either approach.

For the linear interpolation method, we use a direct method to calculate the partial derivatives of travel times with respect to the slowness model parameters in the same way as in the regular-grid DD tomography. That is, a ray is divided into many small segments, with the mid-point of each segment located in a specific tetrahedron. Each segment length can be attributed to the relevant irregular mesh nodes in proportion to interpolating weights at the mid-point. In the case of natural neighbor interpolation, however, it is extremely slow to calculate the interpolating weights and thus the partial derivative directly, due to the large number of nodes that contribute to the interpolated value at any given point. Instead, we use an indirect way to calculate the partial derivatives of travel times with respect to the model slowness parameters, by projecting them from the regular computational grid to the irregular inversion mesh. The finer the regular grid is, the more accurate the partial derivatives are.

The inversion is started from a regular inversion grid, equivalent to that of *tomoDD*. We then construct a tetrahedral or Voronoi diagram using the Quick Hull algorithm (BARBER *et al.*, 1996) around the starting regular inversion grid that is randomly perturbed by a very small amount so that the nodes are not located on the same plane. A regular computational grid that remains fixed during the inversion is constructed to be the same as the starting regular inversion grid or finer. Ray paths between events and stations are traced using the ART-PB ray-tracing algorithm (UM and THURBER, 1987) and saved for later use in defining the adaptive mesh. In the process, we calculate the derivative weight sum (DWS) values (THURBER and EBERHART-PHILLIPS, 1999) on the inversion mesh nodes using the saved rays. Threshold DWS values are set to add or remove nodes. Currently we add nodes by inserting one additional node into the middle

of a tetrahedron when the sum of DWS values on its 4 nodes satisfies a predefined threshold value (ZHANG and THURBER, 2005). The threshold checking is repeated to determine the inversion mesh to be used for the current iteration of simultaneous inversion (ZHANG and THURBER, 2005). Once the inversion mesh is determined to be data adaptive, a new tetrahedral or Voronoi diagram is constructed and the partial derivatives of the travel times with respect to the new set of inversion mesh nodes are calculated using the saved rays for the construction of the seismic tomography equations. After each simultaneous inversion, the velocity values on the irregular inversion mesh nodes and the regular computational grid nodes are updated. For subsequent simultaneous inversions, the inversion mesh is again updated following the same procedure to better match with the ray distribution, which will change as the velocity model changes and hypocenters move.

The adaptive mesh method considers the different ray distribution between the P and S waves. Due to the relative simplification of picking P arrivals compared to S, normally there will be more data and hence more inversion mesh nodes for P waves than S waves. At each simultaneous inversion, the same procedure as discussed above is used to construct different inversion meshes for P and S waves.

The above procedure can be applied directly for absolute data with the inversion mesh built according to the absolute ray paths. In the case of using a combination of absolute and differential data, however, the model derivatives from the differential data are included in the mesh refinement process, so the inversion mesh will be distributed more finely near the source region where these derivatives are largest. In this case, we keep those irregular inversion mesh nodes outside the source region and fix their velocity values if their DWS values are smaller than a threshold value (ZHANG and THURBER, 2005). In this way, we preserve the velocity structure outside the source region that is determined by the absolute data information while refining the velocity structure near the source region.

Once the inversion mesh is set up, we can calculate the partial derivatives of travel times with respect to the event locations and slowness model parameters. In a similar way to the case of the regular inversion grid, we can then construct the linear system of equations to solve for the event locations and slowness perturbations at the irregular mesh nodes.

### *Review of DD Tomography Applications*

DD tomography has been applied to a suite of problems including characterizing fault zone structure, imaging high-resolution structure of subduction zones, and determining the velocity structure of volcanoes. In these studies, we have also carried out a variety of tests of the DD tomography method using synthetic data sets. We discuss some examples of each type of application below, including references to associated synthetic tests.

### 1. Characterizing Fault Zone Structure

Our first application of DD tomography (*tomoDD*) was to study the Hayward fault in California (ZHANG and THURBER, 2003). This study used 17,955 differential arrival times calculated from waveform CC, 767,127 catalog differential times constructed directly from the absolute catalog arrival times, and 20,257 absolute catalog times for P waves (ZHANG and THURBER, 2003). In comparison to standard tomography using absolute arrival times, the DD tomography model shows a sharper velocity contrast across the fault with faster velocity to the west and slower velocity to the east of the fault. DD tomography also produces a sharp picture of event locations similar to DD location (WALDHAUSER and ELLSWORTH, 2001) but with different absolute locations. We also noticed some differences in relative locations between the two DD methods and attribute the differences to the velocity heterogeneity not considered in DD location. Without considering velocity complexity in DD location, partial derivatives of arrival times with respect to event locations may be biased and thus relative event locations may also be biased (MICHELINI and LOMAX, 2004). A synthetic test based on an idealized model of the velocity structure of the San Andreas fault in central California was also presented to show DD tomography is able to better characterize the velocity structure near the source region and the absolute event locations were more accurately located (ZHANG and THURBER, 2003). THURBER *et al.* (2004) applied *tomoDD* to study the San Andreas fault zone near Parkfield using catalog picks. Compared to the simul2000 algorithm, *tomoDD* resulted in more tightly clustered earthquake locations. Though the model results from both algorithms were comparable in most areas, the *tomoDD* model showed more detailed features near the fault zone (THURBER *et al.*, 2004).

*TomoDD* has been applied to the study of a number of fault zones in Japan. OKADA *et al.* (2004) used *tomoDD* to image the detailed structure of focal areas of the 1995 M7.3 southern Hyogo (Kobe) earthquake, 2000 M7.3 western Tottori earthquake and 2003 M6.4 northern Miyagi earthquake. They found that large slip areas (asperities) corresponded to the high velocity zones for all three earthquakes. An independent study by ENESCU and MORI (2004) also found high velocity anomalies around the 2000 M7.3 western Tottori mainshock area using *tomoDD* and waveform cross correlation. The application of *tomoDD* to the recent 2004 M6.6 mid-Niigata prefecture earthquake showed lower velocity for the hanging wall (to the west of the focal area), higher velocity for the footwall (to the east of the focal area), and aftershocks are distributed along a zone where seismic velocity changes abruptly. The results suggest the earthquake sequence occurred along faults that were normal in the Miocene and then recently reactivated as reverse faults under compression (OKADA *et al.*, 2005). This study used manually picked arrival times and associated differential times obtained from a temporary dense network composed of 54 stations around the focal area. KATO *et al.* (2005) independently applied *tomoDD* to the arrival times from 716 aftershocks collected from 14 temporary seismic stations

immediately deployed in and around the source region of the 2004 mid-Niigata prefecture earthquake. Their results also showed the aftershocks were distributed along a clear boundary between the low and high velocity structures corresponding

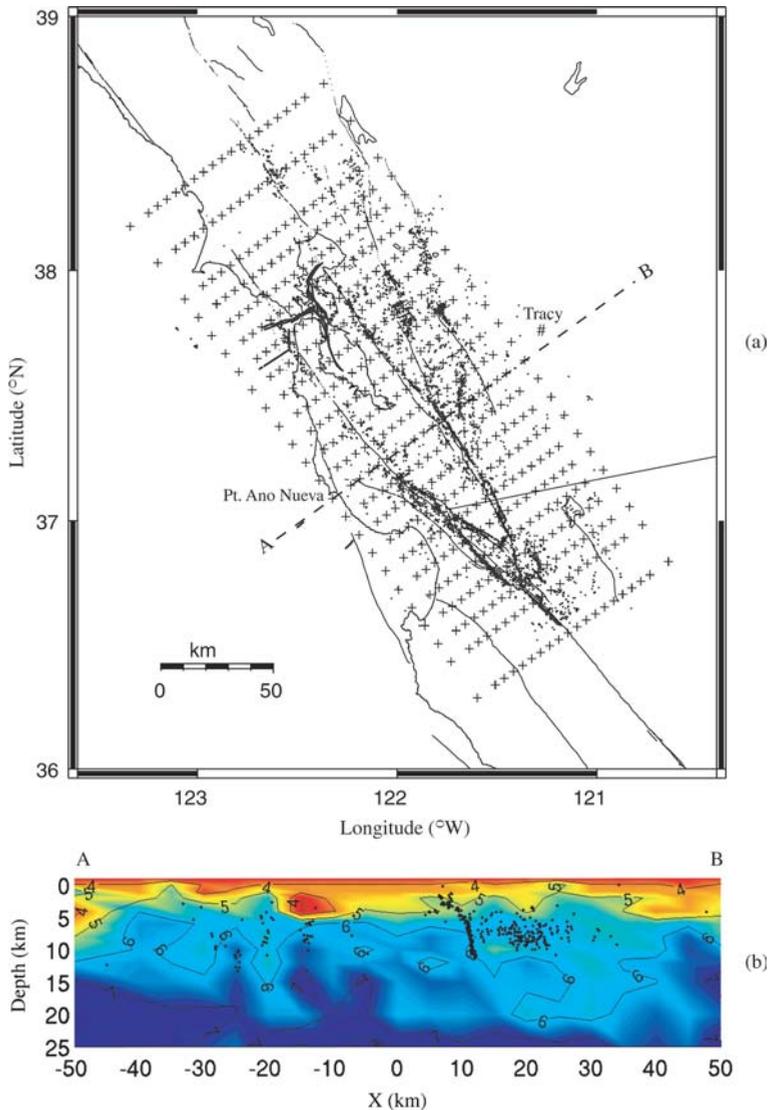


Figure 6

(a) Event and inversion grid distribution map of Northern California. Events are indicated by dots and inversion grid nodes are indicated by crosses. Thin lines indicate fault traces. The coordinate center is located at latitude 31.5° N and longitude 122° W. (b) NE-SW cross section of P-wave velocity model from Pt. Ano Nueva in the southwest to near the town of Tracy in the northeast.

to the hanging wall and footwall, respectively. TAKEDA *et al.* (2004) studied the Atotsugawa fault in central Japan using *tomoDD*. They found a low-velocity zone up to 10 km depth in the cross section along the fault and the earthquakes distributed along the upper boundary of this low-velocity zone.

We also applied the adaptive-mesh version of DD tomography (*tomoADD*) to the Parkfield, California data set (ZHANG and THURBER, 2005) using both linear and natural neighbor interpolation methods. Both methods resulted in more clustered earthquakes than standard tomography, and the events are concentrated where the velocity contrast is sharpest. The velocity models using both interpolation methods are characterized by a clear velocity contrast across the fault and they are comparable in most areas, with the velocity model using natural-neighbor interpolation being slightly smoother. We also found a velocity reversal at  $\sim 4$  km depth to the southwest of the fault when using only the absolute arrival times, but this anomaly disappears when including the differential data in the inversion. Synthetic tests indicate that this velocity reversal could be caused by the noisier absolute data. Thus including the differential data in the inversion may be helpful for removing some velocity features that could be artifacts resulting from the noisier absolute data.

Our current work involves the development of a regional crustal model for Northern California (Fig. 6a). The model will be used for improving earthquake locations and for generating synthetic waveforms and estimating ground motions for the 1906 San Francisco earthquake. Approximately 6000 earthquakes with arrival data extracted from the Northern California Earthquake Data Center (NCEDC) and about 5,500 shots or airgun blasts are used in this study. 466,368 absolute catalog times and 1,629,973 differential times were used in the inversion. The inversion grid was rotated so that the +X and +Y axes are approximately SAF-normal and SAF-parallel, respectively. The grid intervals were 5 km in the X direction and 10 km in the Y direction. In the Z direction, grid nodes were positioned at -1, 0, 1, 3, 6, 9, 12, 15, 20, 25 and 30 km. For this study, we used the *tomoFDD* algorithm considering the relatively large size of the study region. An example cross section through a preliminary model with  $5 \times 10$  km horizontal gridding is shown in Figure 6b. The section runs from Pt. Ano Nueva in the southwest to near the town of Tracy in the northeast. At  $X = -25$  km, we see the deep seismicity that is the northwestern edge of the Loma Prieta rupture zone. Just to the northeast, at  $X = -20$  km, is the diffuse, shallower seismicity representing the San Andreas fault, marked by a distinct low-velocity zone. Adjacent to that, at  $X = -15$  km, the deep, low-velocity, Cupertino Basin is evident, with an active fault situated beneath it. On the northeast side of the section, the northeast-dipping Calaveras Fault seismicity is evident at  $X = +10$  km, associated with the edge of a dipping high-velocity block. There is earthquake activity on both sides of the Calaveras here — the shallower activity to the southwest appears to mark the beginning of the Hayward fault, whereas the deeper activity to the northeast ( $X = +15$  to  $+25$  km) is on an obliquely striking fault connecting the Calaveras and Greenville faults.

These applications of DD tomography to a variety of fault zones showed that DD tomography is able to better relocate events and resolve the finer details of fault zone structure. The relationship between the event distribution and velocity contrast across the fault for some fault zones is more clearly shown. More accurate event locations and finer velocity structure will be more helpful for understanding the factors controlling the occurrence of earthquakes in the fault zone. We note that by itself, however, DD tomography is not a solution for characterizing sharp material interfaces. One can try to model secondary arrivals that provide evidence for an interface (reflections or refractions) but that is a separate issue from what differential times can provide (BEN-ZION *et al.*, 1992).

## 2. Imaging High-resolution Structure of Subduction Zones

Subduction zones are one of the most important components of the Earth's plate tectonic system. Our first application of the regional-scale version of DD tomography (*tomoFDD*) was to image high-resolution subducting-slab structure beneath northern Honshu, Japan using catalog picks (ZHANG *et al.*, 2004). Previous seismic tomography studies for the same region (ZHAO *et al.*, 1992; NAKAJIMA *et al.*, 2001)

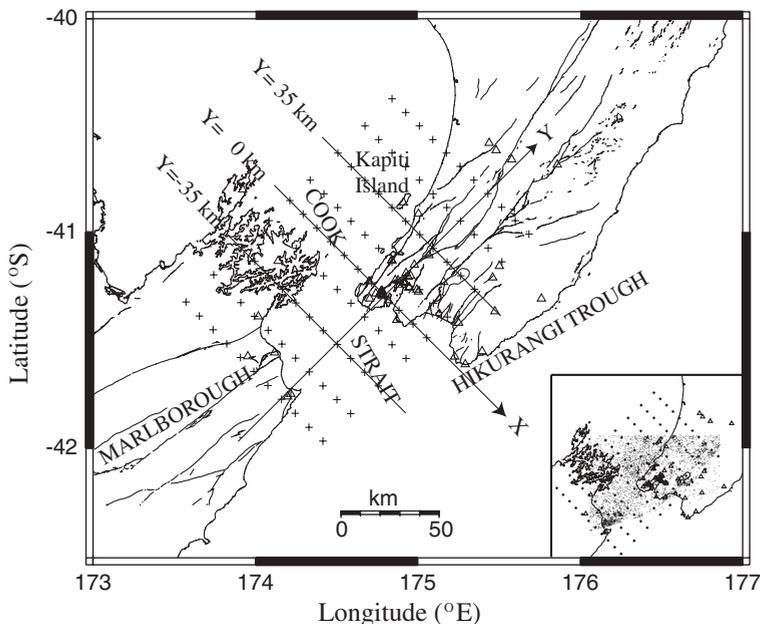


Figure 7

Event and station map for our New Zealand study. Events are indicated by dots in the inset map. Stations are indicated by triangles and inversion grid nodes are indicated by crosses. Positive Y points northeast, positive X points southeast, and positive Z points downward.

could not resolve the detailed internal structure of the down-going slab using standard tomography; their models showed the slab as a relatively homogeneous high velocity anomaly. Our model, however, showed for the first time that the interior of the slab actually has a highly heterogeneous velocity structure. The upper plane of earthquakes lies in a region with relatively low  $V_p$  and  $V_s$  values and average to high  $V_p/V_s$  ratios (1.72–1.85) that may correspond to the transformations of metabasalt and metagabbro to blueschist. The lower plane of earthquakes exists in a region of relatively low  $V_p$  and high  $V_s$  values and low  $V_p/V_s$  ratios that could be due to dehydration reactions of serpentine and other hydrous minerals. The region between the two planes of earthquakes has relatively high  $V_p/V_s$  ratios that indicate partial hydration of the normal dry mantle. Synthetic tests with and without key slab features found in this study using the same data distribution as the real data showed the model was well resolved. SHELLY *et al.* (2004) found similar results when they applied *tomoDD* and waveform cross correlation to the subduction zone near Ibaraki, Japan that is located to the south of ZHANG *et al.* (2004)'s study region.

We further applied *tomoFDD* to study the transition zone from oblique subduction to active continent-continent collision in the Wellington region, New Zealand, where a double seismic zone is also present (Fig. 7). We used 5471 earthquakes spanning the transition region between the North and South Islands that were recorded on 56 stations during the period 1990–2001. For the chosen events and stations, there are 58,082 catalog P data, 299,948 P differential catalog data, and 427,771 WCC P differential times that were verified using a bispectrum analysis method (DU *et al.*, 2004). We adopt a coordinate center of latitude  $-41.3^\circ$ , longitude  $174.8^\circ$ , and depth 0 km (with respect to sea level). The coordinate system is rotated in the horizontal plane by  $45^\circ$  clockwise to be consistent with the strike of the subducting slab. The inversion grid intervals are 10 km in the X direction and 15 to 20 km in the Y direction (Fig. 7). The node spacing in the vertical direction is 4.5 km. In addition to a low-velocity anomaly present in the upper part of the slab corresponding to the subducting oceanic crust and sediments, our model also shows varying features deeper inside the slab (Fig. 8). For the cross sections beneath the North Island, the P-wave velocity is relatively high between the two planes, and it is relatively low in the lower part of the slab, partly in association with the lower plane of seismicity (Fig. 8). Similar to the subducting slab beneath northern Honshu, Japan (ZHANG *et al.*, 2004), the low  $V_p$  region associated with the lower plane of seismicity might be caused by dehydration reactions of serpentine. However, the features are quite different between the cross sections beneath the North versus the South Island (Fig. 8). In the South Island cross sections, the velocity is relatively low and there is almost no seismicity between the two planes of seismicity (Fig. 8). The lower plane of seismicity is associated with a relatively high velocity region (Fig. 8). One possible cause of such features is that the water released from the dehydration reactions of serpentine escaped from the lower plane and entered into the region between the two planes. As a result, the minerals in the lower plane may be mainly composed of

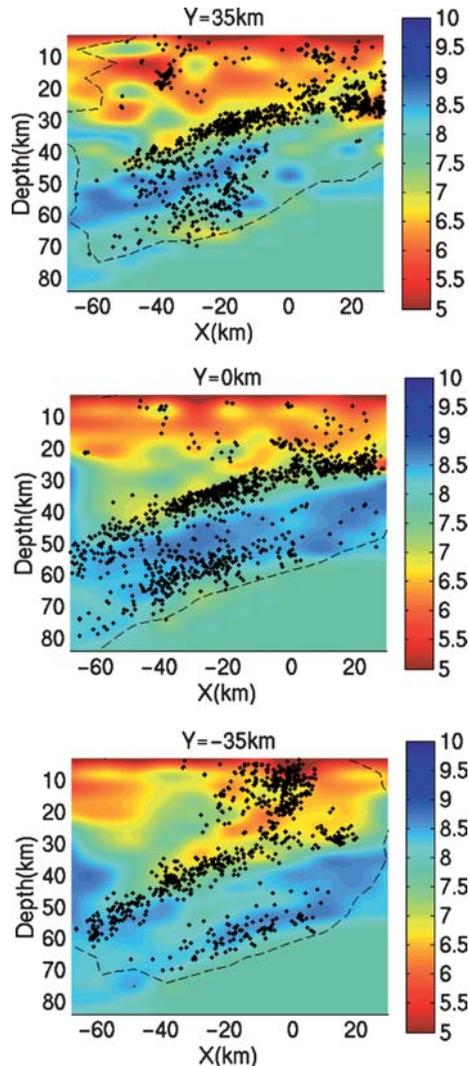


Figure 8

NW-SE cross sections of the New Zealand  $V_p$  structure at  $Y = -35, 0,$  and  $35$  km. Structure above the dashed line is meaningful, where the DWS for each node exceeds 15.

forsterite and enstatite, which have relatively high  $V_p$  (JI and WANG, 1999). The water released into the region between the two planes decreases  $V_p$ . The different features may be related to the transition from subduction to a strike-slip plate boundary from northeast to southwest, and stagnation of the subducting slab in the southwest.

We also found velocity variations in and around the subducting slab for the Kodiak, Kenai, and McKinley blocks for the Alaska subduction zone using

*tomoFDD* (ZHANG *et al.*, 2004). For the Kodiak block, a low velocity zone is present within the slab right below the high velocity zone where most of the Wadati-Benioff Zone (WBZ) earthquakes are located. In comparison, for the Kenai block, the velocity is higher in the zone below the WBZ earthquakes. The McKinley block shows a more complex slab velocity pattern, with high and low velocity zones at different depths within the slab. The high-resolution velocity structure may provide additional constraints for models of the segmentation of the Alaska subduction zone (RATCHKOVSKI and HANSEN, 2002), in addition to earthquake locations, volcanic arc geometry and composition, etc., and may provide a possible explanation for the segmentation.

Another application of *tomoFDD* is to image crust, mantle, and slab structure around the Shikoku district, Japan, and to study its relationship to the occurrence of nonvolcanic deep tremors (NAKAJIMA *et al.*, 2004). High-velocity zones are associated with the deep seismic zone beneath the western part of Shikoku and the Kii Channel that may indicate the presence of the subducting Philippine Sea plate. Low velocity zones lie immediately above them in the western part of Shikoku and the central part of the Kii Peninsula, and the  $V_p/V_s$  ratios are greater than 1.8. The nonvolcanic deep, low-frequency tremors occurring along the strike of the subducting Philippine Sea detected by OBARA (2002) occur right in and around the low-velocity and high  $V_p/V_s$  zones, lending support to the idea that fluids are involved in the tremor. This study, along with those discussed above, shows strong evidence for dehydration reactions in the subducting slab.

Thanks to the ability of DD tomography to resolve the finer structure near the source region and the seismicity distribution inside the subducting slab, it is now possible to characterize the fine details of the velocity structure and accurate earthquake locations inside the slab, as shown in studies discussed above. This will in turn be extremely helpful for understanding the constitution of the slab, the cause of the intermediate depth earthquakes inside the slab, the fluid distribution and recycling, and tremor occurrence (HACKER *et al.*, 2001; OBARA, 2002).

### 3. Determining the Velocity Structure of Volcanoes

BROWN *et al.* (2004) applied *tomoADD* to study Mt. Spurr Volcano, Alaska, located approximately 120 km west of Anchorage in the Cook Inlet, using a data set from 1991 to 2004. They found a low velocity body beneath the crater from  $\sim 1$  km above sea level to  $\sim 10$  km below sea level. This low velocity body is directly associated with the steeply dipping 1992 eruption seismicity that may reflect the magmatic conduit at depth and the presence of a shallow hydrothermal system southeast of Crater Peak. MONTEILLER *et al.* (2004) imaged a detailed structure of the Kilauea Volcano magmatic system using DD tomography based on a Bayesian approach.

Another example is the imaging of the Mt. Etna plumbing system using *tomoDD*. We chose 278 earthquakes that were recorded by 41 stations of Istituto Nazionale di

Geofisica e Vulcanologia - Sezione di Catania (INGV-CT) seismic network (Fig. 9a) during the July 12–18 eruption period, similar to the data set used by PATANÉ *et al.* (2002). A hypothesis for this eruption is that magma was continuously injected from a depth of 6 to 10 km into the shallow magma reservoir and accumulated in the upper part of Mt. Etna's plumbing system during 1994 to 2001 (PATANÉ *et al.*, 2003). This study used  $\sim 5200$  catalog P and S times and  $\sim 35,000$  catalog P and S differential times, among which S arrival times are about 1/6 the number of P arrival

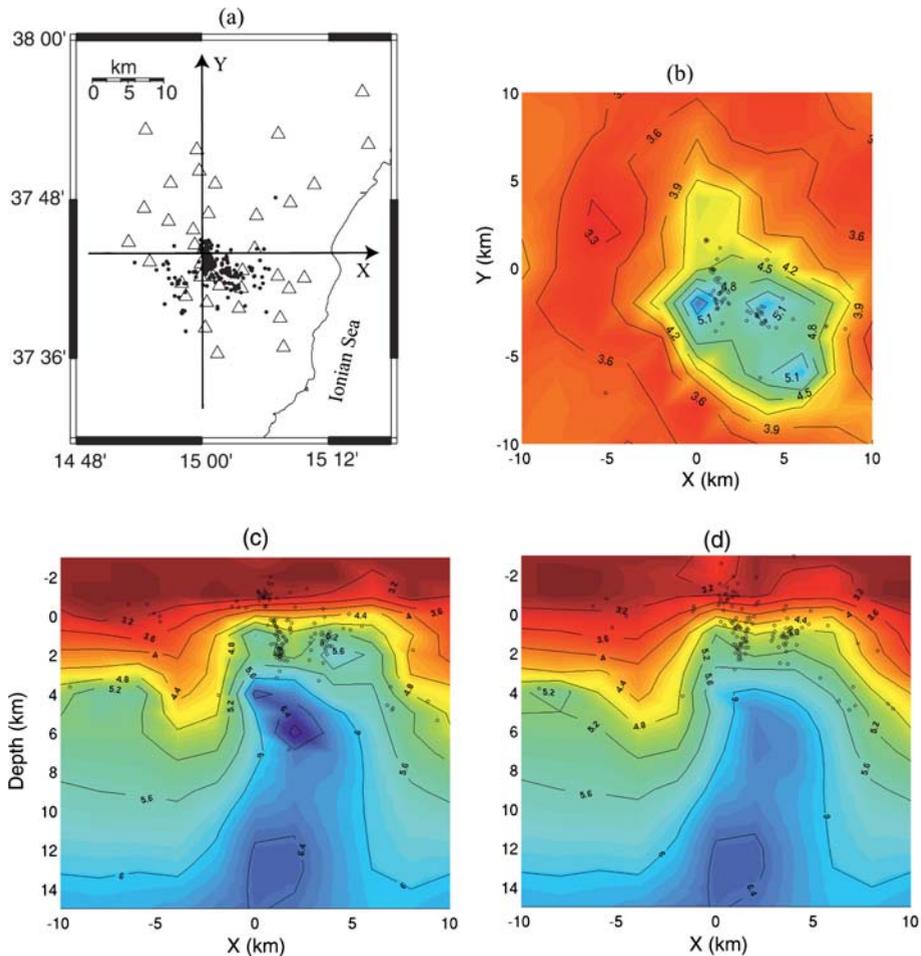


Figure 9

(a) Event and station map for Mt. Etna. The summit crater is located at latitude  $37.75^{\circ}\text{N}$  and longitude  $15.0^{\circ}\text{E}$ . (b) Horizontal section of P-wave velocity model at depth of 1 km with earthquakes within 0.5 km plotted as dots. (c) W-E cross section of P-wave velocity model at  $Y = -2$  km determined using both absolute and differential data. (d) W-E cross section of P-wave velocity model at  $Y = -2$  km using only the absolute data. For (c) and (d), earthquakes within 1 km of the cross section are plotted as dots.

times. The inversion grid intervals in the horizontal and vertical directions are 2 km and 1 km, respectively. In the horizontal section at depth of 1 km (Fig. 9b), a high P-wave velocity zone of 4.8 – 5.1 km/s exists beneath the southeastern part of the summit area, similar to the results of PATANÉ *et al.* (2002). This high- $V_p$  body extends to a depth up to 15 km as shown in the W-E cross-section of  $Y = -2$  km in the velocity model (Fig. 9c), and is interpreted as a mainly solidified intrusive body (PATANÉ *et al.*, 2003).

For comparison, Figure 9d shows the same W-E cross section of  $Y = -2$  km from the velocity model calculated using only absolute data. The velocity model is more sharply defined in the inversion using both absolute and differential times. In particular, the high velocity body shows a greater velocity contrast and more detailed features (Fig. 9c) than in the absolute-data-only model (Fig. 9d). Compared to the results using only absolute data (Fig. 9d), the event locations determined using both absolute and differential data are more clustered. Two clusters of earthquakes are present right above the high velocity body from depths of  $\sim 1$  km above sea level to  $\sim 3$  km below sea level. The western cluster is more concentrated than the southeastern cluster and is situated on the edge of a knob of high velocity (Fig. 9c), which is not evident in the velocity model using only absolute data (Fig. 9d; PATANÉ *et al.* 2002). This cluster occurs beneath a surface fractures field and shows an approximately N-S trend. Focal mechanism analysis showed that the P axes and T axes for earthquakes in the western cluster are mostly oriented approximately N-S and W-E, consistent with stress directions induced by a dike intrusion (MUSUMECI *et al.*, 2004). The southeastern cluster shows an approximately NNW-SSE alignment and might be related to an aborted intrusion or active NNW-SSE faults (MUSUMECI *et al.*, 2004).

For most volcanoes in the world, it is generally difficult to image the detailed structure of their plumbing systems, due to small numbers of stations and poor station coverage geometry. DD tomography is less affected by poor station coverage because it has higher resolution near the source region by using differential arrival times, as shown in the studies discussed above. Better characterizing the velocity structure of the volcano and relocating earthquakes would be helpful for better understanding the volcanic eruption mechanisms and thus predicting volcanic eruptions and mitigating volcanic hazards.

### *Summary and Future Work*

We conducted a resolution and uncertainty analysis based on a 2-D synthetic model for different location and tomography systems of equations, using absolute data (i.e., standard location and tomography), the differential data (i.e., DD location and tomography), and both of them to determine event locations and/or velocity structure. The study shows that the system using the differential data has lower

model uncertainties for both location and most slowness model parameters, but has lower model resolution for some boundary slowness model parameters. Including the absolute data with a relatively small weighting into the differential system is useful to stabilize the system without sacrificing the low model uncertainties and high model resolution that the differential system can provide.

Methods and applications of three versions of the DD tomography algorithm are presented: *tomoDD*, *tomoFDD* and *tomoADD*. *TomoDD* is a local-scale version that assumes a flat Earth model and uses an ART-PB ray-tracing algorithm. The regional-scale version *tomoFDD* considers the curvature of the Earth and uses a finite-difference travel-time algorithm that can better deal with sharp velocity contrasts and discontinuities. Both *tomoDD* and *tomoFDD* are based on a regularly distributed inversion grid. In contrast, *tomoADD* adapts the inversion mesh to match with the data distribution based on tetrahedral and Voronoi diagrams, such that the ray sampling densities on inversion mesh nodes are more uniform than the regular grid case. This approach provides an automated way to take full advantage of the ability to resolve fine-scale structure in the source region using the differential times. We then discuss some examples of applying DD tomography to characterize fault zone structure, image high-resolution structure of subduction zones, and determine the velocity structure of volcanoes.

There are some issues to be considered in our future work, as follows:

1. Currently we have finished the implementation of the adaptive-mesh scheme on the flat-model version of DD tomography (ZHANG and THURBER, 2005). We will implement an adaptive-mesh scheme for the regional-scale version of DD tomography.

2. All three versions of the DD tomography algorithm, *tomoDD*, *tomoFDD* and *tomoADD*, directly solve for P- and S-wave velocities using absolute and differential arrival times. Knowing 3-D Vp/Vs variations is valuable in order to have a more complete characterization of the mechanical properties and geological identity of crust and upper mantle materials (THURBER, 1993). Vp/Vs variations could be determined directly from the Vp and Vs models if they have essentially identical quality. It has been observed that in cases where S-wave arrival data are less numerous and of poorer quality than P-wave data, however, that Vs would be not as well resolved as Vp, making the interpretation of Vp/Vs variations difficult (EBERHART-PHILLIPS, 1990; WAGNER *et al.*, 2005). For this reason, it is generally not appropriate to obtain Vp/Vs ratios by dividing Vp by Vs directly. Alternatively, Vp/Vs variations can be determined by the inversion of S-P time differences (WALCK, 1988; THURBER, 1993). We plan to add this functionality to the DD tomography algorithm to directly invert Vp/Vs ratios using both absolute and differential S-P time differences.

3. The current DD tomography algorithms cannot directly estimate model resolution and uncertainty due to the substantial amount of data used for the system. In ZHANG and THURBER (2003), the model resolution and uncertainty were estimated

by means of the solution technique in the *simul2000* algorithm (THURBER and EBERHART-PHILLIPS, 1999), by inputting the velocity model inverted from DD tomography and using only the absolute times. This technique provides a realistic but conservative estimate of the DD tomography solution quality. Another way of assessing the tomography solution quality is to use a synthetic test or restoration test employing the same data distribution and inversion scheme as the real data (ZHAO *et al.*, 1992; ZHANG *et al.*, 2004). Some sensitivity tests such as a checkerboard test and statistical analysis methods such as “jackknifing” and “bootstrapping” can also be used to estimate the model resolution and uncertainty, respectively (NOLET *et al.*, 1999). However, the sensitivity tests suffer the shortcomings of measuring the sensitivity only with respect to fixed cell or grid patterns (NOLET *et al.*, 1999). The “jackknifing” and “bootstrapping” methods are computationally very expensive and of questionable use for tomographic systems (NOLET *et al.*, 1999). We are now investigating the applicability of several resolution and uncertainty estimation methods for massive tomographic systems proposed by ZHANG and MCMECHAN (1995), MINKOFF (1996), NOLET *et al.* (1999), and YAO *et al.* (1999).

4. MIYAZAWA and KATO (2004) raise questions about the differences between interpolation using velocity versus slowness in grid-based tomography algorithms. We will investigate the effect of the choice of velocity versus slowness on DD tomography both theoretically and empirically.

#### *Acknowledgements*

We are extremely grateful to our collaborators for sharing their results with us. We thank INGV-CT for allowing us to use their Mt. Etna 2001 eruption data. We especially thank Felix Waldhauser and Steve Roecker for their assistance and helpful comments in the early stage of *tomoDD* algorithm development. Megan Flanagan helped us to realize the *tomoFDD* algorithm using part of her code. Certain figures were constructed using GMT (WESSEL and SMITH, 1991). This material is based upon work supported by National Science Foundation grants EAR-9814192 (Continental Dynamics), EAR-0125164 (Geophysics), EAR-0346105 (EarthScope), and EAR-0409291 (Geophysics).

#### REFERENCES

- ASTER, R.C., BORCHERS B., and THURBER C.H., *Parameter Estimation and Inverse Problems*, (Elsevier Academic Press, Burlington, MA, 2005) 301 pp.
- BARBER, C.B., DOBKIN, D.P., and HUHDANPAA, H.T. (1996), *The Quickhull algorithm for convex hulls*, Trans. Mat. Software 22, 469–483.

- BEN-ZION, Y., KATZ, S., and LEARY, P. (1992), *Joint inversion of fault zone head waves and direct P arrivals for crustal structure near major faults*, J. Geophys. Res. 97, 1943–1951.
- BROWN, J., PREJEAN, S., ZHANG, H., POWER, J., and THURBER, C. (2004), *Double difference earthquake relocation and tomography at Mount Spurr Volcano, Alaska, 1991 to 2004*, EOS Trans. AGU 85(47), Fall Meet. Suppl., Abstract S51A–0140.
- DU, W., THURBER, C. H., and EBERHART-PHILLIPS, D. (2004), *Earthquake relocation using cross-correlation time delay estimates verified with the bispectrum method*, Bull. Seismol. Soc. Am. 94, 856–866.
- EBERHART-PHILLIPS, D. (1990), *Three-dimensional P and S velocity structure in the Coalinga region, California*, J. Geophys. Res. 95, 15343–15363.
- ENESCU, B., and MORI, J. (2004), *Relocations and 3-D velocity structure for aftershocks of the 2000 W. Tottori (Japan) earthquake and 2001 Gujarat (India) earthquake, using waveform cross correlations*, EOS Trans. AGU 85(47), Fall Meet. Suppl., Abstract S54A–05.
- FLANAGAN, M.P., MYERS, S.C., SCHULTZ, C.A., PASYANOS, M.E., and BHATTACHARYYA, J. (2000), *Three-dimensional a priori model constraints and uncertainties for improving seismic location*, Proceedings of 22nd Seismic Research Review: Technologies for Monitoring the CTBT, 2–15.
- HACKER, B.R., ABERS, G.A., and PEACOCK, S.M. (2001), *Subduction factory 1, Theoretical mineralogy, densities, seismic wave speeds, and H<sub>2</sub>O contents*, J. Geophys. Res. 108, B12030, doi:2001JB001129.
- HOLE, J.A., and ZELT, B.C. (1995), *3-D finite-difference reflection travel times*, Geophys. J. Int. 121, 427–434.
- Ji, S., and WANG, Z. (1999), *Elastic properties of forsterite-enstatite composites up to 3.0 Gpa*, Geodynamics 28, 147–174.
- KATO, A., KURASHIMO, E., HIRATA, N., SAKAI, S., IWASAKI, T., and KANAZAWA, T. (2005), *Imaging the source region of the 2004 mid-Niigata prefecture earthquake and the evolution of a seismogenic thrust-related fold*, Geophys. Res. Lett. 32, L07307, doi:10.1029/2005GL022366.
- KLEIN, F.W. (1978), *Hypocenter location program HYPOINVERSE*, U.S. Geol. Surv. Open File Rep. 78–694, 114 pp.
- KOKETSU, K., and SEKINE, S. (1998), *Pseudo-bending method for three-dimensional seismic ray tracing in a spherical earth with discontinuities*, Geophys. J. Int. 132, 339–346.
- LEE, W.H.K., and STEWART, S.W. (1981), *Principles and applications of microearthquake networks*, Adv. Geophys. Suppl. 2, 293 pp.
- MONTELLER, V., GOT, J., VIRIEUX, J., and OKUBO, P.G. (2004), *Imaging Kilauea Volcano magmatic system using an efficient algorithm for double-difference tomography and cross-spectral time delays*, EOS Trans. AGU 85(47), Fall Meet. Suppl., Abstract S11D–03.
- MENKE, W., (1989), *Geophysical Data Analysis: Discrete Inverse Theory*, revised edition, vol. 45 of International Geophysics Series (Academic Press, San Diego) 289 pp.
- MENKE, W., and SCHAFF, D. (2004), *Absolute earthquake locations with differential data*, Bull. Seismol. Soc. Am. 94, 2254–2264.
- MICHELINI, A., and LOMAX, A. (2004), *The effect of velocity structure errors on double-difference earthquake location*, Geophys. Res. Lett. 31, L09602, doi:10.1029/2004GL019682.
- MINKOFF, S.E. (1996), *A computationally feasible approximate resolution matrix for seismic inverse problems*, Geophys. J. Int. 126, 345–359.
- MIYAZAWA, M., and KATO, M. (2004), *On interpolation functions in travel-time tomography*, Geophys. J. Int. 158, 169–178.
- MUSUMECI, C., COCINA, O., GORI, P.D., and PATANÉ, D. (2004), *Seismological evidence of stress induced by dike injection during the 2001 Mt. Etna eruption*, Geophys. Res. Lett. 31, L07617, doi:10.1029/2003GL019367.
- NAKAJIMA, J., HASEGAWA, A., ZHANG, H., and THURBER, C.H. (2004), *Imaging crust, mantle and slab structure around the Shikoku district, Japan, by double-difference tomography and its relationship to the occurrence of nonvolcanic deep tremors*, EOS Trans. AGU 85(47), Fall Meet. Suppl., Abstract S51B–0161.
- NAKAJIMA, J., MATSUZAWA, T., HASEGAWA, A., and ZHAO, D. (2001), *Three-dimensional structure of V<sub>p</sub>, V<sub>s</sub>, and V<sub>p</sub>/V<sub>s</sub> beneath northeastern Japan: Implications for arc magmatism and fluids*, J. Geophys. Res. 106, 21,834–21,857.
- NOLET, G., MONTELLI, R., and VIRIEUX, J. (1999), *Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems*, Geophys. J. Int. 138, 36–44.

- OBARA, K. (2002), *Nonvolcanic deep tremor associated with subduction in southwest Japan*, *Science* 296, 1679–1681.
- OKADA, T., HASEGAWA, A., SUGANOMATA, J., ZHAO, D., ZHANG, H., and THURBER, C. (2004), *Imaging the fault plane and asperities of the 1995 southern Hyogo (Kobe) earthquake (M7.3) by double-difference tomography*, *Eos Trans. AGU* 85(47), Fall Meet. Suppl., Abstract S53C–01.
- OKADA, T., UMINO, N., MATSUZAWA, T., NAKAJIMA, J., UCHIDA, N., NAKAYAMA, T., HIRAHARA, S., SATO, T., HORI, S., KONO, T., YABE, Y., ARIYOSHI, K., GAMAGE, S., SHIMIZU, J., SUGANOMATA, J., KITA, S., YUI, S., ARAO, M., HONDO, S., MIZUKAMI, T., TSUSHIMA, H., YAGINUMA, T., HASEGAWA, A., ASANO, Y., ZHANG, H., and THURBER, C. (2005), *Aftershock distribution and seismic velocity structure in and around the focal area of the 2004 mid-Niigata prefecture earthquake obtained by applying double-difference tomography to dense temporary seismic network data*, *Earth Planets Space* 57, 435–440.
- PATANÉ, D., CHIARABBA, C., COCINA, O., and GORI, P.D. (2002), *Tomographic images and 3D earthquake locations of the seismic swarm preceding the 2001 Mt. Etna eruption: Evidence for a dyke intrusion*, *Geophys. Res. Lett.* 29, doi:10.1029/2001GL014391.
- PATANÉ, D., GORI, P.D., CHIARABBA, C., and BONACCORSO, A. (2003), *Magma ascent and the repressurization of Mount Etna's volcanic system*, *Science* 29, 2061–2063.
- PODVIN, P., and LECOMTE, I. (1991), *Finite difference computation of travel times in very contrasted velocity models: A massively parallel approach and its associated tools*, *Geophys. J. Int.* 105, 271–284.
- RATCHKOVSKI, N.A., and HANSEN, R.A. (2002), *New evidence for segmentation of the Alaska subduction zone*, *Bull. Seismol. Soc. Am.* 92, 1754–1765.
- SAMBRIDGE, M., BRAUN, J., and MCQUEEN, H. (1995), *Geophysical parameterization and interpolation of irregular data using natural neighbors*, *Geophys. J. Int.* 122, 837–857.
- SHELLY, D.R., BEROZA, G.C., ZHANG, H., THURBER, C.H., and IDE, S. (2004), *High-resolution subduction zone seismicity and velocity structure in Ibaraki, Japan*, *Eos Trans. AGU* 85(47), Fall Meet. Suppl., Abstract S43D–04.
- TAKEDA, T., KUWAHARA, Y., MIZUNO, T., IMANISHI, K., OKADA, T., ITO, K., WADA, H., and HARYU, Y. (2004), *Crustal structure and micro-seismic activity along the Atotsugawa Fault system, central Japan*, *Eos Trans. AGU* 85(47), Fall Meet. Suppl., Abstract S13D–1103.
- THURBER, C.H., *Local earthquake tomography: Velocities and Vp/Vs theory*, in *Seismic Tomography: Theory and Practice* (eds. H.M. Iyer and K. Hirahara) pp. 563–583 (Chapman and Hall, New York 1993)
- THURBER, C.H. (1992), *Hypocenter velocity structure coupling in local earthquake tomography*, *Phys. Earth. Planet. Int.* 75, 55–62.
- 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., and EBERHART-PHILLIPS, D. (1999), *Local earthquake tomography with flexible gridding*, *Computers and Geosciences* 25, 809–818.
- THURBER, C., ROECKER, S., ZHANG, H., BAHER, S., and ELLSWORTH, W. (2004), *Fine-scale structure of the San Andreas fault and location of the SAFOD target earthquakes*, *Geophys. Res. Lett.* 31, doi:10.1029/2003GL019398.
- UM, J. and THURBER, C.H. (1987), *A fast algorithm for two-point seismic ray tracing*, *Bull. Seism. Soc. Am.* 77, 972–986.
- VIDALE, J.E. (1990), *Finite-difference calculation of travel-times in three dimensions*, *Geophysics* 55, 521–526.
- WAGNER, L.S., BECK, S., and ZANDT, G. (2005), *Upper mantle structure in the south central Chilean subduction zone (30° to 36°S)*, *J. Geophys. Res.* 110, B01308, doi:10.1029/2004JB003238.
- WALCK, M.C. (1988), *Three-dimensional Vp/Vs variations for the Coso region, California*, *J. Geophys. Res.* 93, 2047–2051.
- 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 ELLSWORTH, W.L. (2000), *A double-difference earthquake location algorithm: Method and application to the Northern Hayward fault, California*, *Bull. Seismol. Soc. Am.* 90, 1353–1368.

- WESSEL, P., and SMITH, W.H.F. (1991), *Free software helps map and display data*, EOS Trans. AGU 72, 441 and 445–446.
- WOLFE, C. J. (2002), *On the mathematics of using difference operators to relocate earthquakes*, Bull. Seismol. Soc. Am. 92, 2879–2892.
- YAO, Z.S., ROBERTS, R.G., and TRYGGVASON, A. (1999), *Calculating resolution and covariance matrices for seismic tomography with the LSQR method*, Geophys. J. Int. 138, 886–894.
- ZHANG, J., and MCMECHAN, G.A. (1995), *Estimation of resolution and covariance of large matrix inversions*, Geophys. J. Int. 121, 409–426.
- ZHANG, H., and THURBER, C. (2005), *Adaptive-mesh seismic tomography based on tetrahedral and Voronoi diagrams: Application to Parkfield, California*, J. Geophys. Res. 110, B04303, doi:10.1029/2004JB003186.
- ZHANG, H., RATCHKOVSKI, N., THURBER, C., and HANSEN, R. (2004), *High-resolution seismic velocity structure of the Alaska subduction zone revealed by double-difference tomography*, EOS Trans. AGU 85(47), Fall Meet. Suppl., Abstract S51B–0168.
- ZHANG, H., THURBER, C.H., SHELLY, D., IDE, S., BEROZA, G.C., and HASEGAWA, A. (2005), *High-resolution subducting-slab structure beneath northern Honshu, Japan, revealed by double-difference tomography*, Geology 32, 361–364, doi: 10.1130/G20261.
- ZHANG, H., and THURBER, C.H. (2003), *Double-difference tomography: The method and its application to the Hayward fault, California*, Bull. Seismol. Soc. Am. 93, 1875–1889.
- Zhao, D., Hasegawa, A., and Kanamori, H. (1992), *Tomographic imaging of P- and S-wave velocity structure beneath northeastern Japan*, J. Geophys. Res. 97, 19, 909–19,928.

(Received April 1, 2005; accepted September 15, 2005)

Published Online First: February 8, 2006



To access this journal online:  
<http://www.birkhauser.ch>

---