

Regional and teleseismic double-difference earthquake relocation using waveform cross-correlation and global bulletin data

Felix Waldhauser¹ and David Schaff¹

Received 16 January 2007; revised 28 June 2007; accepted 4 September 2007; published 5 December 2007.

[1] We have developed a double-difference algorithm to relocate earthquakes recorded at global seismic networks, using differential arrival times for first and later arriving regional and global phases to invert for the vectors connecting the hypocenters. Differential times are formed from global seismic bulletins and are accurately measured on similar seismograms by time domain waveform cross correlation. We evaluate the performance of this spherical, multiphase double-difference algorithm using three-dimensional regional-scale synthetic data and two sets of earthquake data in different tectonic settings. The first includes 3783 intermediate depth earthquakes that occurred between 1964 and 2000 in the subducting Nazca plate beneath northern Chile, where the relocated seismicity confirms a narrowly spaced double seismic zone previously imaged with temporary local seismic data. Residual statistics and comparison with accurately known locations indicate mean relative location errors at the 90% confidence level of 2.4 km laterally and 1.8 km vertically. Later events typically constrained by cross-correlation data have errors of 1.6 km laterally and 1.4 km vertically. The second data set includes 75 crustal earthquakes in the 1999 Izmit and Düzce, Turkey, aftershock sequences, where the double-difference solutions image orientation and dip of individual fault segments that are consistent with focal mechanisms and near-surface information. Fault complexity likely causes a low level of waveform similarity in this aftershock sequence and thus generates fewer correlated events compared to the Chile earthquakes. Differences between the double-difference locations and corresponding locations in global seismicity catalogs (Earthquake Data Report, EDR; International Seismological Centre, ISC; Engdahl-Hilst-Buland, EHB) are typically greater than 10 km. We evaluate the potential of cross-correlation and double-difference methods to improve hypocenter locations on a global scale.

Citation: Waldhauser, F., and D. Schaff (2007), Regional and teleseismic double-difference earthquake relocation using waveform cross-correlation and global bulletin data, *J. Geophys. Res.*, 112, B12301, doi:10.1029/2007JB004938.

1. Introduction

[2] The increasing number and quality of permanent seismic stations over the last several decades has led to a significant yearly increase in the number of globally recorded earthquakes between 1960 and present. Previously undetected earthquakes, especially smaller magnitude events, are currently reliably recorded and located on a global scale at a rate of $\sim 250,000$ events per year. This has led to an increasing spatial density of hypocenters archived in global earthquake bulletins, especially in regions along the active boundaries of major tectonic plates.

[3] Catalogs of global seismicity, such as the Earthquake Data Report (EDR) produced by the U.S. Geological Survey's National Earthquake Information Center (NEIC), list hypocentral parameters that are routinely estimated in

near real-time from phase arrival time data. Catalogs produced by the International Seismological Centre (ISC), the most comprehensive catalog available, reanalyze and expand the NEIC catalogs using additional phase information from stations not used by the NEIC. Using statistical methods to improve phase identification and erroneous phase arrival time picks, Engdahl *et al.* [1998] (EHB) reprocessed well recorded events in the ISC catalog to produce the EHB catalog, which currently includes the most accurate global event locations.

[4] Global catalogs of seismicity are widely used as the primary data to study the physical processes controlling earthquakes and plate tectonics, to image the structure and composition of the Earth's interior, and to estimate the hazards imposed by large earthquakes. Yet, uncertainties in the hypocenter locations are generally significant. In many cases they exceed the length-scale of the tectonic units in which the earthquakes occur and in most cases the length of fault surface associated with individual earthquakes, thus inhibiting the detailed imaging of active faults at seismogenic depths. The uncertainties in global event

¹Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, USA.

locations typically arise from insufficiency in the number and distribution of available observations, errors in phase identification and phase arrival time picks, and absolute model errors associated with single-event location methods: the most commonly employed procedures that locate each earthquake individually by minimizing its residuals between observed and some predicted traveltime [Geiger, 1910].

[5] The increasing density of seismicity in global earthquake catalogs can be harnessed by reducing pick errors via waveform cross correlation and the effect of model errors via multiple event location techniques. Cross-correlation methods take advantage of the fact that two earthquakes that are close in space and have similar focal mechanisms produce similar seismograms at common stations [Poupinet *et al.*, 1984]. In these cases, cross-correlation methods can measure differential phase arrival times with subsample precision, typically resulting in more than an order of magnitude improvement over phase onset picks reported in earthquake bulletins [e.g., Poupinet *et al.*, 1984; Deichmann and Garcia-Fernandez, 1992; Schaff *et al.*, 2004]. Multievent location methods can reduce the effect of model errors inherent in single-event locations by taking advantage of traveltime information from nearby events with similar travel paths to common stations [e.g., Douglas, 1967; Jordan and Sverdrup, 1981; Pavlis and Booker, 1983; Poupinet *et al.*, 1984; Got *et al.*, 1994; Waldhauser and Ellsworth, 2000; Richards-Dinger and Shearer, 2000].

[6] In recent years the double-difference (DD) method of Waldhauser and Ellsworth [2000] has proven efficient in determining high-resolution locations of earthquakes recorded at local distances (direct Pg and Pn phases, and their S analogous) to resolve detailed seismicity structures in a wide range of tectonic settings [e.g., Waldhauser *et al.*, 1999; Schaff *et al.*, 2002; Prejean *et al.*, 2002; Rietbrock and Waldhauser, 2004; Hauksson and Shearer, 2005]. Since the algorithm naturally accepts differential traveltimes as data directly, either from high-precision cross-correlation measurements, or from ordinary arrival time picks, the simultaneous inversion of both data sets relocates correlated events to the precision of the cross-correlation data, and events that do not correlate to the precision of the phase pick data.

[7] Here we present a modification and extension of the original, local-scale double-difference algorithm hypoDD [Waldhauser, 2001] to relocate earthquakes using first and later arriving P and S phases recorded at local, regional and teleseismic distances. This spherical double-difference algorithm has been applied in specialized studies to a small number of selected earthquakes beneath the South Sandwich Islands [Zhang *et al.*, 2005] and to nuclear explosions at Lop Nor, China [Waldhauser *et al.*, 2004], producing relative source location uncertainties of less than 1 km. In this article we describe, test, and evaluate the new algorithm on a more comprehensive scale by using three different data sets: regional-scale three-dimensional (3-D) synthetic data; 4093 Wadati-Benioff earthquakes in the subducting Nazca plate beneath northern Chile; and 75 crustal aftershocks in the 1999 Izmit-Düzce, Turkey, earthquake sequence. We use both phase arrival time picks listed in the EHB catalog [Engdahl *et al.*, 1998] and differential times of regional and teleseismic phases precisely measured from digital seismo-

grams via time domain cross correlation. We discuss the potential to use these methods in a comprehensive way on a global scale.

2. Global Multiphase Double-Difference Relocation

[8] Commonly used global multiple event location algorithms such as joint hypocenter determination (JHD) [Douglas, 1967; Frohlich, 1979] and hypocentral decomposition (HDC) [Jordan and Sverdrup, 1981] attempt to correct for common path model error, due to three-dimensional Earth structure, via residual-based 1-D (static) station corrections. Such corrections reduce the effects of unmodeled structure mainly near the receiver site, to a degree that scales with the area of seismicity used to determine the correction factor.

[9] The double-difference algorithm differs from other methods of joint hypocentral location in that no station corrections are necessary, because the effects of unmodeled velocity structure is directly removed from the data pertinent to estimating the relative location of each event pair. The fundamental equation of the double-difference method relates the residual between observed and predicted phase traveltime difference, dr_k^{ij} , for pairs of earthquakes i and j observed at a common station, k , to changes in the vector connecting their hypocenters through the partial derivatives of the traveltimes, t , for each event with respect to the unknown, \mathbf{m} [Waldhauser and Ellsworth, 2000]:

$$\frac{\partial t_k^i}{\partial \mathbf{m}} \Delta \mathbf{m}^i - \frac{\partial t_k^j}{\partial \mathbf{m}} \Delta \mathbf{m}^j = dr_k^{ij} \quad (1)$$

When the earthquake location problem is linearized using the double-difference equation (1), the common mode errors cancel without the need for high accuracy of predicted traveltimes for the portion of the raypath that lies outside the focal volume. This explicit formulation of the relative location problem is particularly useful to relocate large areas of dense seismicity, where hundreds or thousands of earthquakes can be linked together through a chain of near neighbors.

[10] The double-difference approach allows each data link (equation (1)) to be directly weighted during the iterative least squares inversion, with weights defined according to a series of a priori and a posteriori criteria that include measurement type (phase pick, cross correlation) and quality (pick quality, cross-correlation coefficient), phase type (P, S, etc.), residual performance and inter-event distance (see Waldhauser and Ellsworth [2000] and Waldhauser [2001] for details). The latter two criteria are revisited after each iteration and adjusted for the new locations, thus taking into account the nonlinearity inherent in the earthquake location problem. Dynamic inter-event distance weighting is a crucial component in double-difference applications (see below), as initial hypocenter locations (i.e., those in existing catalogs) are generally mislocated by distances greater than the distances between hypocenters linked during final iterations. The importance of interevent distance weighting has not been appreciated in recent studies comparing the double-difference method with other relative location algorithms, as inversions were not iterated and therefore distance

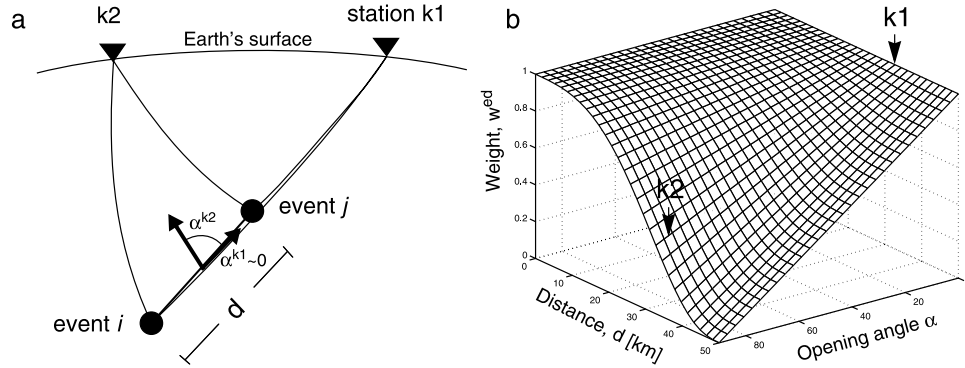


Figure 1. Illustration of the effective interevent distance weighting function. (a) Schematic description of the opening angle, α , between the relative position vector of an event pair (i, j) and the average ray takeoff direction (arrows) determined at the pair's centroid with respect to a station k . (b) Surface representing effective interevent distance weights as a function of interevent distance, d , and opening angle, α . Arrows indicate weights assigned to the phase pairs observed at station $k1$ and $k2$ in Figure 1a.

weights not readjusted [Wolfe, 2002], or distance weighting was omitted completely [Lin and Shearer, 2005]. Double-difference applications without interevent distance weighting give results similar to JHD or HDC methods, especially when all possible links between events are established.

[11] We have extended the original a priori weighting scheme described by Waldhauser and Ellsworth [2000] to accommodate global phase characteristics such as the varying phase pick quality as a function of recording distance. For example, phases recorded at crossover distances (e.g., P and PKP between 120° and 140° epicentral distance) are downweighted relative to phases that can generally be identified and picked with more confidence before and beyond this distance range. We also allow for phase specific weights, introducing the flexibility to tailor double-difference applications to tectonic environments with different wave propagation characteristics.

[12] We modified the dynamic distance weighting function to further reduce the effect of unmodeled velocity structure outside the focal area. The original distance weight as described by Waldhauser and Ellsworth [2000] is obtained by applying a biweight function to the hypocentral separation distance of two events in order to downweight differential times that connect hypocenters which are far apart:

$$w^d = \left\{ \max \left[0, 1 - \left(\frac{d}{c} \right)^a \right] \right\}^b \quad (2)$$

Equation (2) is typically chosen to be bicubic ($a = b = 3$). In order to more efficiently reduce model error effects between events with large separation distances, we now weight interevent distances as a function of both the distance between hypocenters and the angle, α , between the relative position vector of an event pair and the average ray takeoff direction computed at the pair's centroid with respect to the observing station (Figure 1a):

$$w^{ed} = 1 - \left\{ \left(1 - w^d \right) \cdot \frac{\alpha}{90^\circ} \right\} \quad (3)$$

where w^d is the original distance weight as stated in equation (2). This new effective interevent distance weight, w^{ed} , retains a full weight when an event pair is observed at a

common station along its relative position vector ($\alpha \sim 0$) (Figure 1b). In this case, the events share a common raypath outside the source region, regardless of the distance between their hypocenters ($k1$ in Figure 1). The effective interevent distance weight equals the original distance weight when the ray takeoff direction is perpendicular to the pair's relative position vector ($\alpha \sim 90^\circ$), and the effect of model error is greatest ($k2$, Figure 1). The effective distance weight is especially useful to minimize bias due to Earth structure in global double-difference applications where event pairs are often constrained by multiple phase pairs observed at a common station with rays leaving the sources at different angles relative to the pair's position vector.

[13] We solve the spherical forward problem of predicting partial derivatives and traveltimes of first and later arriving phases by using the standard Earth model ak135 [Kennett et al., 1995]. One-dimensional, radially symmetric Earth models are efficient tools for comprehensive double-difference applications that typically involve thousands of earthquakes, hundreds of stations, and require numerous iterations. For sufficiently dense distributed hypocenters, the double-difference algorithm effectively reduces errors that arise from deviations between the model used to predict the data and the true 3-D Earth structure. Rietbrock and Waldhauser [2004] relocated locally recorded subduction earthquakes using both a 1-D layered and a 3-D tomographic velocity model, and found differences between the two sets of double-difference solutions to be smaller than the noise level in the data. Effects of unmodeled 3-D velocity structures may become significant, however, when the shortest interevent distances exceed the length scale of velocity heterogeneities located along the raypaths of linked events. Then, the effect of model errors is likely to map into the double-difference solutions. In the following section we employ 3-D spherical ray tracing in the double-difference scheme to quantify such effects.

3. Effects of 3-D Earth Structure on 1-D Double-Difference Solutions

3.1. Synthetic Model and Data

[14] We construct a regional-scale, spherical subduction zone model that extends 2000 km in both north-south and

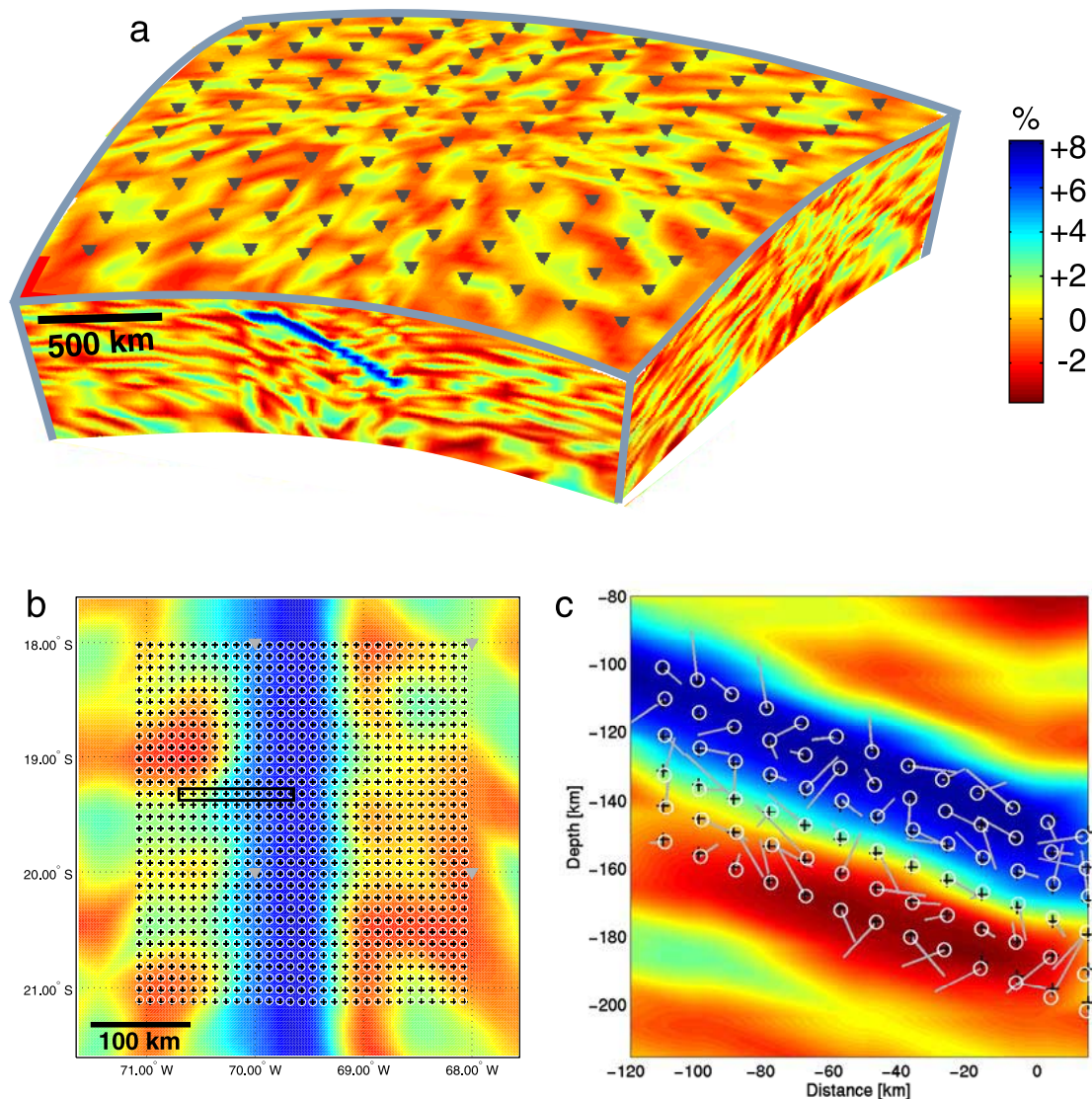


Figure 2. Synthetic test to investigate the effect of 3-D slab structures on regional double-difference solutions obtained using a standard Earth model. (a) A 3-D synthetic model representing a subducting slab with a 8% velocity increase (blue) and random $\pm 3\%$ velocity perturbations (red to green) superimposed on the IASP91 background structure. Triangles represent receiver locations. Scale is velocity change relative to the IASP91 model. (b) Horizontal slice through the 3-D model showing velocity structure at 120 km depth and (c) east-west depth section showing the slab structure in the area outlined by the black rectangle in Figure 2b. Black pluses denote true source locations, white circle denotes double-difference relocations, white lines point to the starting locations.

east-west directions, and reaches a depth of 520 km (Figure 2a). The model features a subducting, 40 km thick, cold slab with a 8% velocity increase within the slab and a $\pm 3\%$ random velocity perturbation elsewhere on top of the IASP91 background model. Traveltimes of first arriving phases were computed using the program raytrace3d [Menke, 2005] from a set of 5952 sources, placed across the slab between 120 and 290 km depth with an interevent distance of 10 km (black pluses in Figure 2b and 2c), through the 3-D structure to a grid of 121 receivers at the model's surface (black triangles in Figure 2a). Raytrace3d computes rays of both first and later arriving phases as arcs of circles within adjoining tetrahedra, a parameterization

that allows the construction of spherical models while the actual ray tracing is done in a Cartesian coordinate system. The spherical setup is necessary as the Earth's curvature has to be taken into account at regional distances and beyond [Snoko and Lahr, 2001].

[15] We form differential traveltimes links between each event and its six closest neighbors within 10 km, based on the true locations. We perturb the 5952 sources from the true locations to mimic a mean mislocation error of 15 km (gray lines in Figure 2c) and subsequently relocated the sources using the double-difference algorithm together with the ak135 model. No noise is added to the synthetic differential times in order to separate data errors from model errors. The

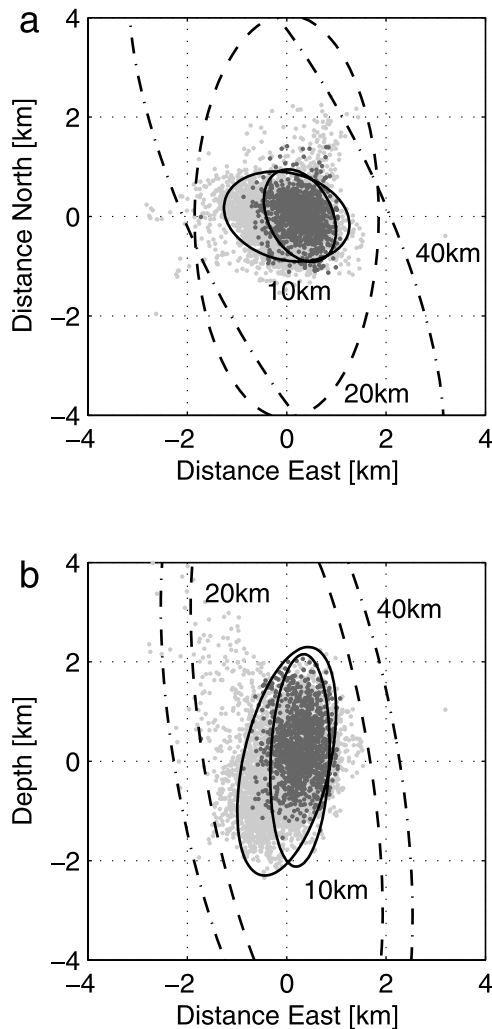


Figure 3. Difference between double-difference locations and true locations (mislocations) in (a) horizontal and (b) vertical directions. Light gray dots represent mislocations for all events with 10 km event separation, dark gray dots for those excluding the “edge events.” Large (small) solid ellipses include 90% of light gray (dark gray) samples. Dashed ellipses include 90% of the mislocations obtained from double-difference runs with events separated by 20 and 40 km.

relocated sources are shown as white circles in Figure 2b and 2c. The distribution of mislocations (final relocations–true locations) are shown in Figure 3a and 3b for the horizontal and vertical direction, respectively. After relocation the mislocation mean for the 5952 events is 0.62 km laterally and 0.68 km vertically. Large solid ellipses in Figure 3 contain 90% of the mislocation samples. If we compute the mislocation statistics without including events that locate at the edge of the band of seismicity, the mean mislocation decreases to 0.52 km horizontally and 0.57 km vertically. The smaller solid ellipses in Figure 3 contain 90% of this subset of mislocations. Events at the edge of seismicity clusters are generally less well constrained due to the limited (i.e., one-sided) spatial sampling of differential time links with neighboring events.

[16] Errors at the 90% confidence level, estimated from a bootstrap analysis of the remaining residuals [Waldhauser and Ellsworth, 2000], indicate a mean maximum horizontal error of 0.48 km, and a mean maximum vertical error of 0.57 km. These errors are comparable to the mean mislocations of 0.52 km and 0.57 km, indicating that bootstrap errors account for most of the location bias due to model errors. Note that the larger errors in relative depths result from a systematic lack of down going rays, compared to a dense sampling of shallowly leaving rays that travel to the receivers at regional distances. Relocation applications using teleseismic phases, as will be shown below, fill these gaps with differential times from deep mantle and core phases.

3.2. Three-Dimensional Effects

[17] The interevent distance over which each event is linked to its six nearest neighbors in the above inversion was a constant 10 km. In order to quantify the effect of larger interevent distances, and thus the increasing effect of model error, we have subsampled the source locations so that each event is now linked to its 6 nearest neighbors over (true) distances of 20 km. A second set of subsampled sources is produced with interevent distances of 40 km. Double-difference inversions of these two data sets result in mean mislocations of 1.6 km (1.9 km) in horizontal and 1.5 (1.7) km in vertical direction for the 20 km (40 km) interevent distance runs. The 90% coverage ellipses for the two mislocation distributions are shown in Figure 3 by dashed ellipses. The increased level of mislocation is caused by the increased effect of 3-D model errors outside the source region.

[18] The distance weighting function (equation (2)) assigns unit weights to data of colocated events, and exponentially downweights data that connect events further apart. By comparison, in the above synthetic experiments unit weights were assigned to all data links as the interevent distances are a constant 10, 20, or 40 km. Thus the mislocation results shown in Figure 3 reflect the true effect of these specific interevent distances, and their associated model error, on the double-difference solutions. The generally small mislocations indicate that the effects of 3-D model errors are efficiently minimized, even when as strong a velocity contrast as ours (8%) is present in the slab model, and for relatively large interevent distances. The bias due to unmodeled velocity structure is below the typical noise level of both pick data and cross-correlation measurements (for 20 Hz seismograms). This is particularly true when the seismicity is densely distributed. Existing 1-D standard Earth models may therefore be a good approximation for comprehensive, large-scale, computer intensive double-difference inversions. In the case of areas with less dense seismicity, however, 3-D global Earth models [e.g., Ritzwoller *et al.*, 2003] may be required in order to reduce the residuals associated with relative locations of events that are farther apart.

4. Application to Subduction Earthquakes in South America

[19] Earthquakes within the Wadati-Benioff Zone (WBZ), caused by the interaction of a subducting oceanic plate with

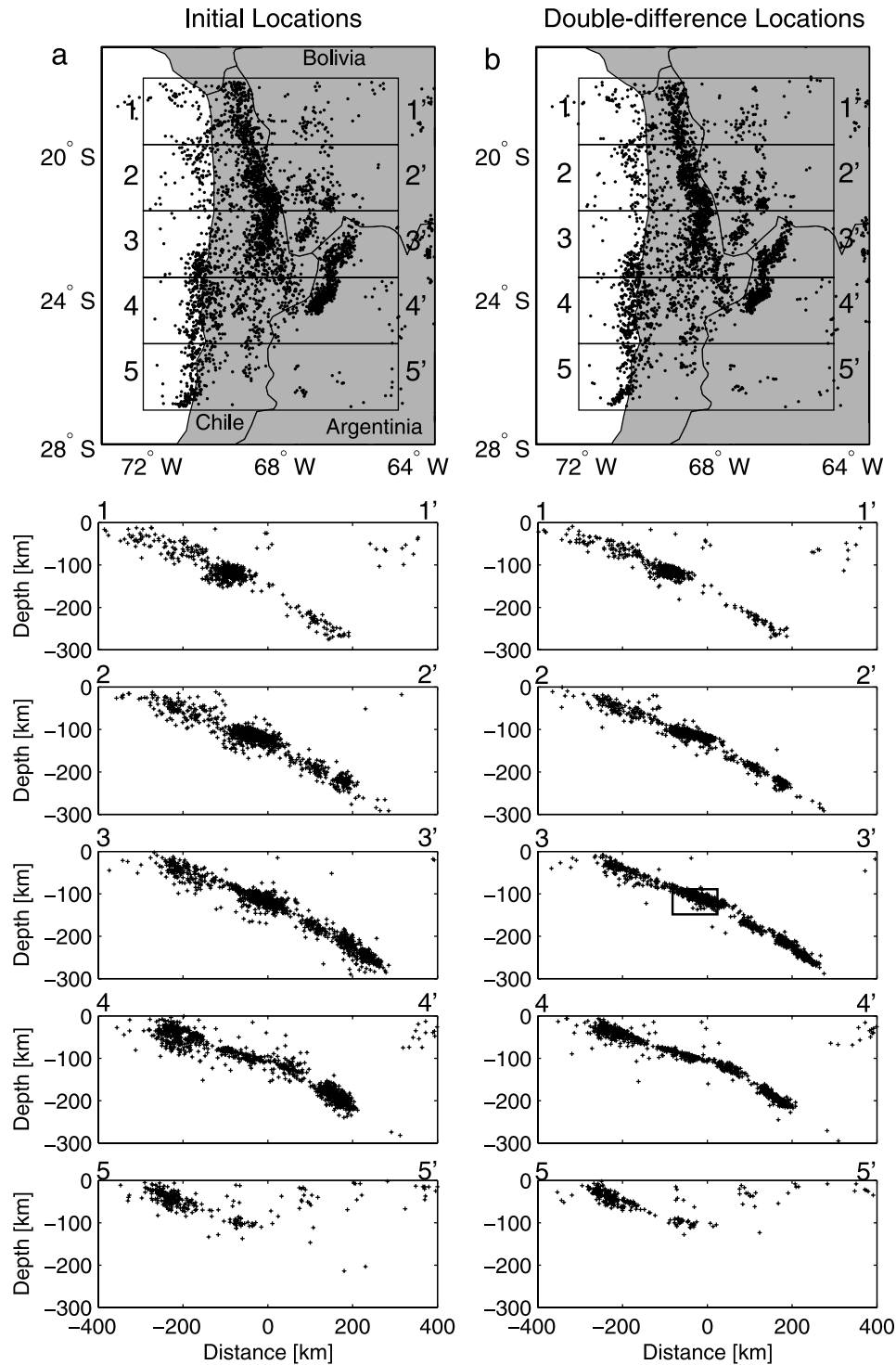


Figure 4. Map view and east-west depth sections of 3783 earthquakes in the South American subduction zone beneath northern Chile (a) before and (b) after double-difference relocation. Hypocenter locations in Figure 4a are from the EHB catalog [Engdahl *et al.*, 1998]. Boxes in map view indicate the locations of cross section. Box in cross section 3-3' includes events shown in Figure 6. Same events are shown in Figures 4a and 4b.

a continental plate, can be generated by slip along the subduction thrust fault or by slip on faults within the downgoing plate. While it is generally accepted that metamorphism in the downgoing lithosphere play an important

role in the generation of intermediate depth earthquakes [e.g., Kirby *et al.*, 1996; Hacker *et al.*, 2003], it is still not well understood how and where exactly within the subducting plate such processes may cause brittle failure. This is

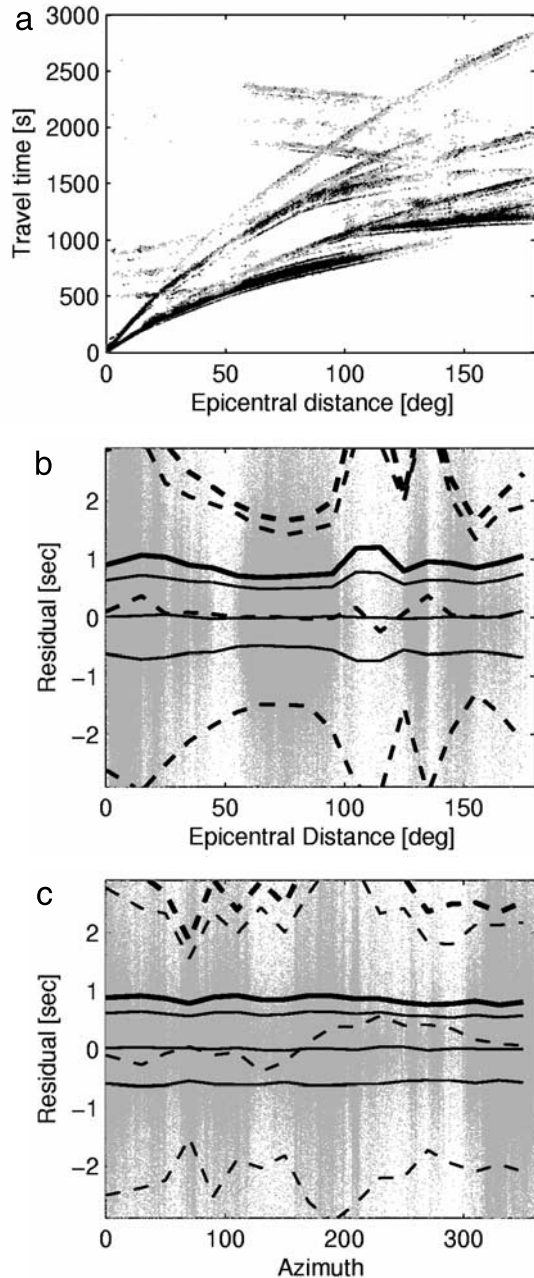


Figure 5. (a) Phase traveltimes for the 3783 events shown in Figure 4. Gray dots denote phase picks listed in the EHB bulletin, open circles those observed pair-wise and used to form the double-difference vector. Note the decrease in scatter after pair wise selection of the phases. (b) and (c) Differential time residuals after relocation (gray dots) as a function of epicentral distance (Figure 5b) and azimuth (Figure 5c). In each of the two subplots, thin solid lines indicate mean and standard deviation of the final residuals, thick lines the RMS, and dashed lines the corresponding values for the initial (EHB) residuals before relocation. RMS and standard deviations are computed within bins of 10° distance (Figure 5b) and 20° azimuth (Figure 5c). Standard deviations are found by fitting a Gaussian function to the residual distribution within each bin.

particularly true for the abundance of small and moderate size earthquakes that are believed to occur within the subducting plate, but the uncertainty in their hypocenter locations listed in existing global catalogs is generally too large to provide conclusive answers.

4.1. Relocation With Phase Pick Data

[20] We evaluate the ability of the teleseismic double-difference algorithm to resolve such details in WBZ seismicity by using 4093 globally recorded (>15 stations) earthquakes in the Nazca plate of the South American subduction zone (Figure 4a). The earthquakes occurred from 1964 to 2000 in northern Chile, covering the area between 18° – 28° S and 64° – 73° W and a depth range of 0–300 km. Event magnitudes are between M 2.8 and 6.6. We use the groomed ISC phase pick data and initial locations from the EHB bulletin [Engdahl *et al.*, 1998] to form differential traveltimes between each event and its 20 nearest neighbors that have at least 15 commonly observed phases and are within 300 km distance. In order to create a well conditioned system of double-difference equations we require a continuous chain of links with a link strength of 10 differential times per pair. The resulting cluster includes 3907 earthquakes.

[21] From an original number of 316,000 phase arrival times in the EHB bulletin, 251,763 (80%) phases have been consistently observed for at least two events at a common station (Figure 5a). The selection of pair-wise observed phases greatly reduces the scatter in the original phase pick data. During the iterative nonlinear relocation process, 33% of the original differential times were removed by the dynamic weighting scheme that revisits the quality and link strength of each differential time observation after each of the 13 iterations. The final relocated catalog includes 3783 earthquakes, about 97% of the original number of events. The relocated seismicity is shown in Figure 4b, both in map view and as a series of east-west cross sections. Visual inspection of the double-difference results show a sharper image of the seismicity associated with the downgoing slab compared to the initial (EHB) locations. Particularly, a narrower Wadati Benioff zone is imaged in virtually all cross sections. In map view, the clusters of seismicity are generally more defined.

[22] A substantial reduction in the differential time residuals is observed after relocation (Figure 5b and 5c). The unweighted (weighted) RMS decreases from 1.49 s (1 s) before relocation to 0.83 s (0.58 s) after relocation. Strong variations in initial RMS values as a function of both recording azimuth and distance (dashed thick lines in Figure 5b and 5c) are smoothed out in the final RMS residuals (solid thick lines). This is especially the case for residuals in the 100° – 150° distance range (Figure 5c), where phase arrival times are usually associated and picked with greater uncertainty due to multiple phase arrivals within a short time window.

[23] While the reduction in traveltime residuals and formal location uncertainties indicate a better fit of the new solutions to the observations, these measures, on their own, are insufficient for evaluating the performance of the relocation procedure. Unlike crustal environments, where ground truth data (information on accurate location and origin time) can be readily produced through dense moni-

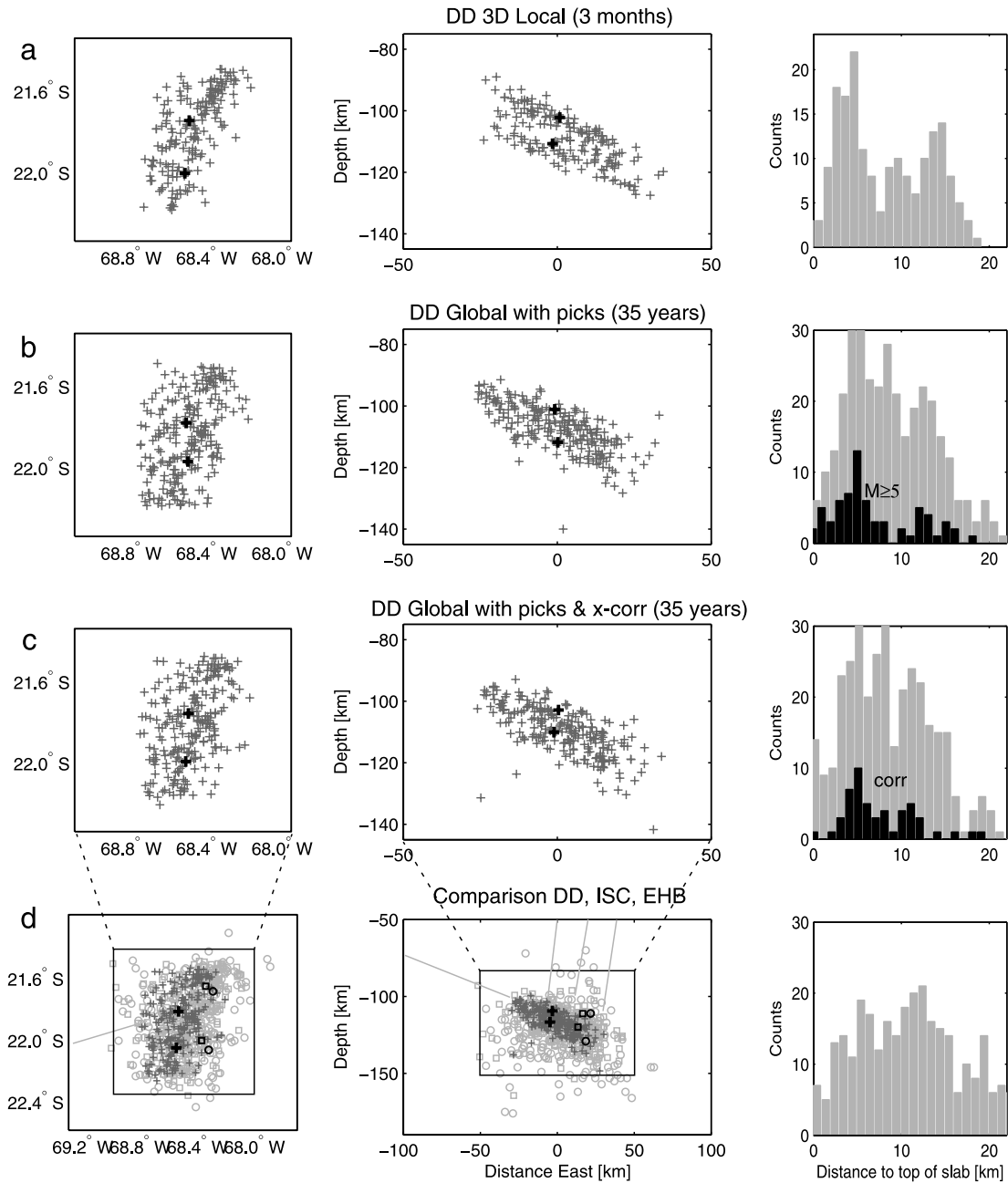


Figure 6. (left) Map view and (middle) east-west depth sections of seismicity in an undisturbed area of the subducting Nazca plate (see box in cross section 3-3 of Figure 4 for location). (a) Accurate 3-D local network double-difference solutions from *Rietbrock and Waldhauser* [2004] for 183 earthquakes recorded during a 3 months period, obtained using local P and S phase picks and cross-correlation measurements. (b) and (c) Global double-difference solutions obtained in this study for 221 earthquakes recorded during a 36 a period using Figure 6b for EHB phase picks and Figure 6c for EHB phase picks combined with cross-correlation differential times for P and PKP phases and their S wave analogs. (d) Comparison between the double-difference locations shown in Figure 6c (gray plus), ISC (light gray circles) and EHB locations (light gray squares). Events in solid black are the two events common to all data sets (plus, DD; circle, ISC; squares: EHB). Absolute locations of cluster in Figures 6b and 6c are centered on the mean of the two events common to all data sets (black plus). (right) Histograms of hypocenter depths as a function of distance from the top of the Nazca plate for the events shown in Figure 6 (middle): Figure 6a (right) local network locations [after *Rietbrock and Waldhauser*, 2004]; Figure 6b (right) global pick-based DD locations (gray bars) and a subset of $M > 5$ events (black bars); Figure 6c (right) global cross-correlation-based DD locations (gray) and a subset of strongly correlated events (black); and Figure 6d (right) EHB locations.

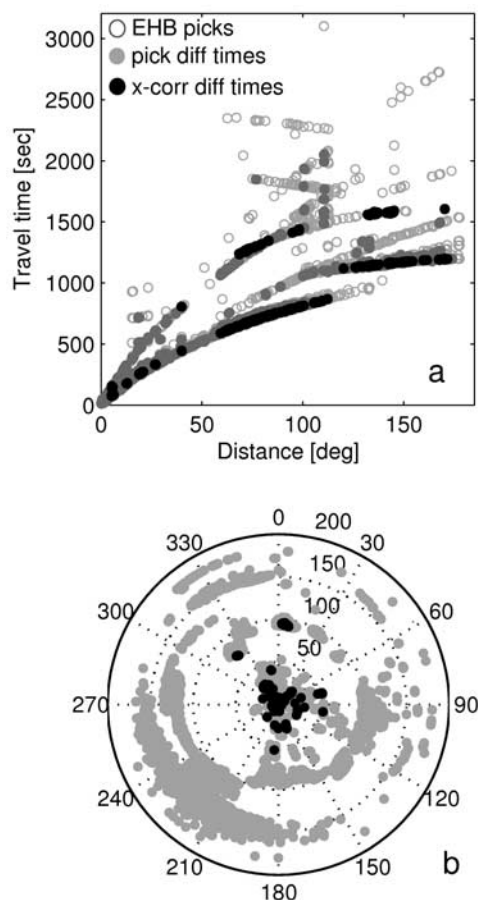


Figure 7. Observed phase traveltimes as a function of (a) distance and (b) takeoff angles as a function of azimuth for the WBZ earthquakes shown in Figure 6c. In Figure 7a gray circles denote original EHB bulletin picks, and in Figures 7a and 7b, gray dots denote pair-wise observed phase picks (two dots are plotted per linked pair), and black dots denote phases for which cross-correlation measurements were obtained.

toring networks, surface ruptures, or explosion data, ground truth data in subduction environments are difficult to obtain. In order to evaluate the accuracy of the double-difference results shown in Figure 4b, we investigate and compare our results in a selected area (see box in cross section 3-3' of Figure 4b) for which high-resolution earthquake locations are available from a dense, temporary, local seismic network.

[24] Between 20°S and 23°S latitude, covering the forearc as well as the western part of the Altiplano Plateau in northern Chile and southern Bolivia, a dense array of 37 seismic stations recorded 1050 local earthquakes ($M_L = 1.0-4.7$) from December 1996 to February 1997 [ANCORP Working Group, 1999]. Rietbrock and Waldhauser [2004] relocated 238 earthquakes in an undisturbed part of the downgoing slab by applying the double-difference method to local P and S phase pick and cross-correlation measurements while solving the forward problem by 3-D ray tracing in the high-resolution 3-D tomographic velocity model of Rietbrock and Haberland [2001]. These relocations revealed a ~ 19 km thick, clearly defined double-layered

WBZ, with the two subparallel layers separated by about 9 km (Figure 6a). Independent active source data places the upper seismic layer at the top of the oceanic crust, and the lower layer within the uppermost mantle [Rietbrock and Waldhauser, 2004].

[25] We use these high-resolution locations (absolute errors are 3–5 km and relative errors are ~ 0.3 km) to evaluate the 37 years (a) of global double-difference relocations. Figure 6b shows 359 of these relocated earthquakes that occurred between 1964 and 2000 ($M_w 3.5-6.6$) in the same $\sim 100 \times 100$ km area shown in Figure 6a. The overall structure and thickness of the two data sets are strikingly similar, indicating a spatially and temporally persistent WBZ of approximately 19 km thickness in this area. Two earthquakes in the global double-difference catalog ($M_w 3.8$ on 3 December 1996 and $M_w 3.9$ on 15 February 1997) were also recorded at the temporary array and accurately located in the Rietbrock and Waldhauser [2004] catalog (black pluses in Figure 6), allowing a direct assessment of the precision of the global relocations. The mean relative mislocation of these two events is 4.2 km horizontally and 1.1 km vertically.

4.2. Relocation With Cross-Correlation Data

[26] We employ and evaluate waveform cross-correlation methods with the goal to improve the accuracy of the differential times formed from bulletin picks, and evaluate their effect on location precision. Waveforms from all broadband stations (20Hz) available at the Data Management Center (DMC) of the Incorporated Research Institutions for Seismology (IRIS) that recorded the 359 events shown in Figure 6b were acquired to measure accurate differential times for dominant first and later arriving P and S phases. We used a time domain cross-correlation method that is described in detail by Schaff *et al.* [2004] and has been extensively applied to locally recorded phases [Schaff and Waldhauser, 2005] and regional and teleseismic phases [Waldhauser *et al.*, 2004; Zhang *et al.*, 2005] to compute differential arrival times for double-difference relocation purposes.

[27] We performed 56,621 “black box” cross correlations at 998 stations, from which 7318 P and 567 S waves have a cross correlation coefficient (CC) > 0.7 (Figure 7). Cross correlations are performed on a 10 s window surrounding the first arriving P or S wave. Lags searched over are plus and minus 5 s. In order to assure good signal energy in the correlation window, and thus enhance the robustness of the cross-correlation measurements, a STA/LTA filter is run on the traces with a trigger level set to a signal-to-noise ratio (SNR) of 3.2. The STA and LTA window lengths are chosen to be 1 s and 5 s, respectively.

[28] Examples of cross-correlated seismograms are shown in Figure 8 for phases observed at regional (station ZOBO, Bolivia) and teleseismic (ANMO, United States; WMQ, China) distances. Superimposed on the aligned waveforms are the EHB bulletin phase picks that demonstrate the two main benefits of using waveform cross correlation: the reduction in the scatter in phase onset picks (ZOBO, WMQ), and the measurement of additional delay times of phases not picked by analysts. The standard deviation of the scatter in the analyst picks is 1.7 s at station ZOBO, 0.14 s at station WMQ, and 0.03 s at station

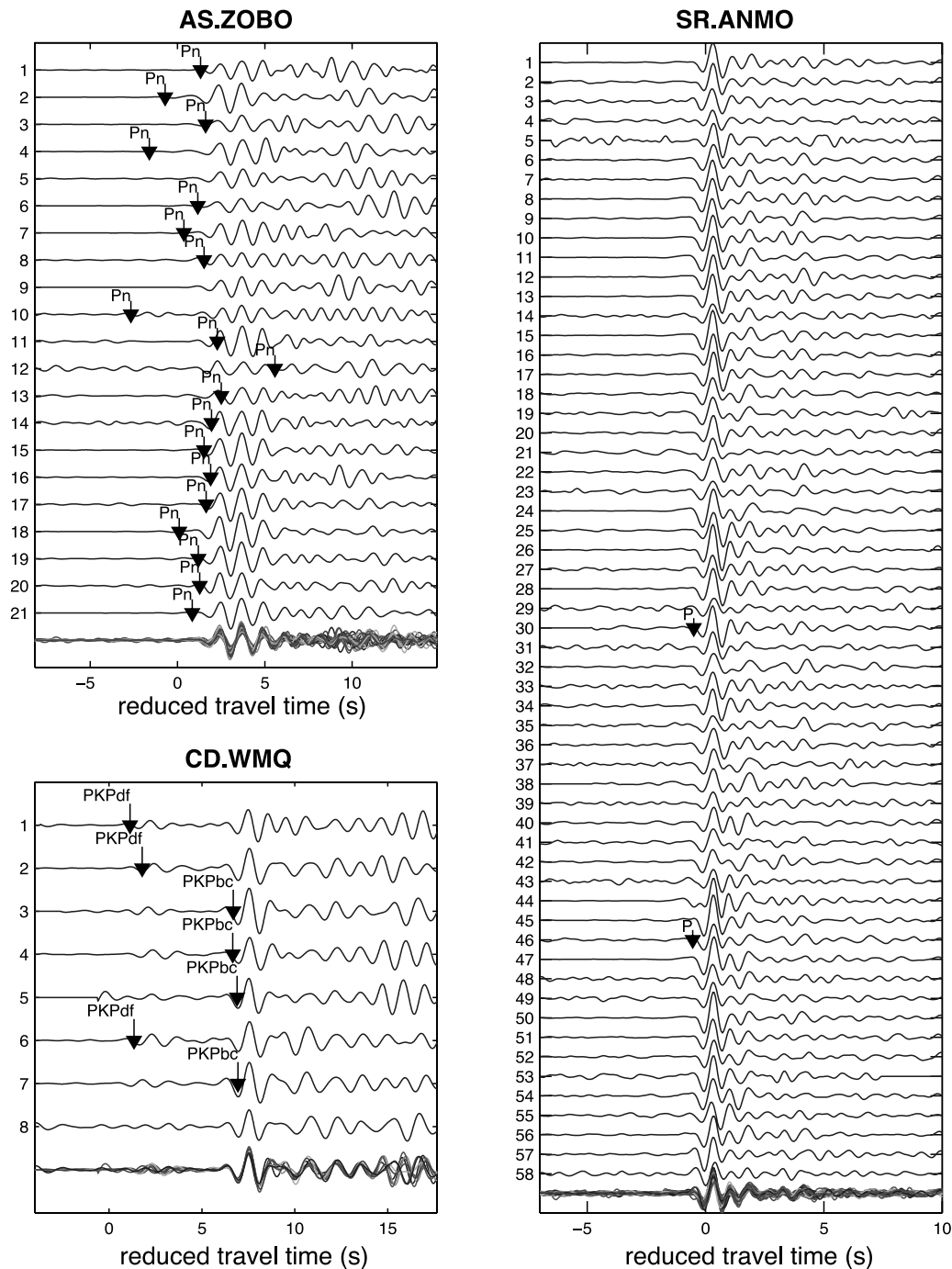


Figure 8. Filtered (0.1–2 Hz) and cross-correlation-aligned waveforms of selected events in the Chile cluster recorded at regional (ZOBO) and teleseismic (ANMO, WMQ) stations. Phases observed at ANMO are bottoming in the lower mantle, those at WMQ in the core. Bottom line in each plot superimposes traces shown above, arrows indicate arrival time picks available from the EHB bulletin.

ANMO (note that station ANMO has only two picks). Error in the correlation measurements based on an evaluation of the internal consistency is on the order of 0.01 s or less. For WMQ this represents about 1 order and for ZOBO 2 orders of magnitude improvement over differential times formed from phase picks.

[29] Figure 9a shows the difference between pick and cross-correlation traveltime differences (Δdt) as a function

of correlation coefficient for 218 event pair/station triplets common to both the pick and the cross correlation data sets. It indicates that the substantial scatter in the phase picks can be reduced via cross correlation. For comparison, the RMS of all Δdt values is 0.9 s (dashed line), similar to the RMS of the pick-based differential times after relocation (~ 1 s). The increase in RMS residuals with decreasing cross-correlation coefficients (solid line in Figure 9a) indicates

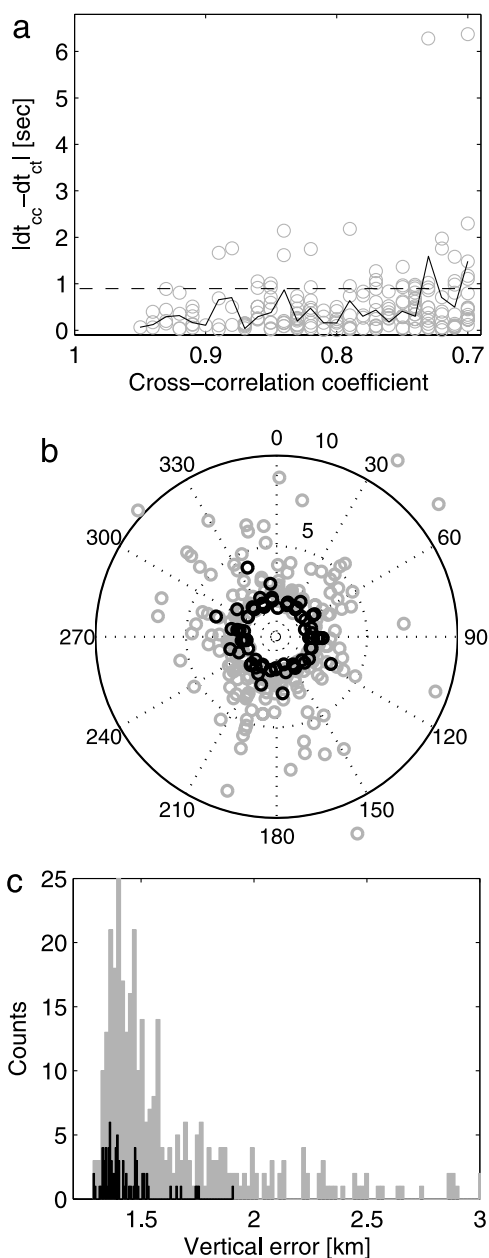


Figure 9. (a) Differences between differential times obtained from bulletin picks (dt_{ct}) and corresponding cross-correlation measurement (dt_{cc}), shown as a function of correlation coefficient. Dashed line indicates the RMS of the entire data set, solid line the RMS values within bins of 0.01 cross-correlation coefficient. (b) Polar plot of horizontal and (c) histogram of vertical bootstrap errors of all relocated events shown in Figure 6c (gray circles and bars). Black circles and bars denote errors from events constrained predominantly by cross-correlation data.

the introduction of less accurate cross-correlation measurements as waveforms become more dissimilar. Since the cross-correlation data are weighted by the squared correlation coefficient, such less accurate data is downweighted relative to the pick data during the double-difference inversions (for more details, see *Waldhauser and Ellsworth [2000]* and *Schaff et al. [2004]*).

[30] Results of the combined inversion of the 7885 cross-correlation and the 334,972 pick differential times are shown in Figure 6c. About 51% of the events have 10 or more, 32% have 50 or more, and 7% have 100 or more cross-correlation measurements. The number of cross-correlation measurements typically increases for later events, due to the increased availability of high-quality digital seismic stations. Figure 9b and 9c show the distribution of the lateral and vertical semimajor axes of the 90% confidence ellipses derived from a bootstrap analysis of the final double-difference vector. These error estimates are shown separately for all events (gray circles/bars) and for those constrained by 50 or more cross-correlation measurements (black circles/bars). The means of the horizontal and vertical semimajor axes are 3.5 km (2.0 km) and 2.1 km (1.4 km) for all events (events constrained by 50 or more cross-correlation links). Relative differences between the two ground truth events and the cross-correlation-based double-difference relocations are 1.4 km in horizontal and 0.7 km in vertical direction.

[31] Figure 6d shows the final cross-correlation-based double-difference locations superimposed on the ISC and EHB locations, in map view and cross section. Table 1 lists median differences in horizontal and vertical directions between the cross-correlation double-difference locations and corresponding locations in the ISC, the EHB, and the pick-based double-difference catalog. Differences are largest (132 km) for hypocenter depths listed in the ISC catalog. Differences between the EHB and the double-difference locations are in the range of 12 km, which is similar to the average error estimates for locations in the EHB catalog (± 10 –15 km [*Engdahl et al., 1998*]). The double-difference relocations (Figure 6c) reveal a much more focused view of the seismicity compared to the ISC and EHB locations, especially in depth where they indicate a sharp ~ 20 km thick subducting band of seismicity consistent with the seismic activity imaged by the short-term local network experiment (Figure 6a).

4.3. Resolving the Double Seismic Zone

[32] In order to investigate the existence of a double seismic zone (DSZ) in the global double-difference solutions, we show in the far right panels of Figure 6 histograms of the event depths with respect to the top of the Nazca plate. The depth distribution for the 3-D local network locations observed over a 3 months period is shown in Figure 6a as reference. This histogram shows a clear concentration of hypocenters at the top of the oceanic crust and the uppermost oceanic mantle, with some events locating in between and connecting the two layers along normal faults (see *Rietbrock and Waldhauser [2004]* for a detailed discussion). The separation between the two layers of seismicity is approximately 10 km. The depth histogram

Table 1. Horizontal (Δh) and Vertical (Δz) Shifts Between the Final Cross-Correlation-Based Double-Difference Locations for the Chile DSZ Events and the Corresponding Locations in the ISC, EHB, and Pick-Based DD Catalog

	ISC	EHB	DD_pick
Median Δh , km	15.3	12.1	2.0
Median Δz , km	132.2	12.6	0.3

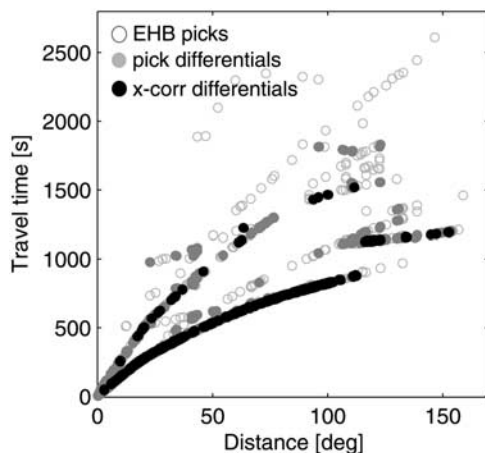


Figure 10. Phase traveltimes as a function of distance for the 75 earthquakes in the Izmit-Düzce sequence. Gray circles denote original EHB bulletin picks, gray dots pair-wise observed picks, and black dots phases for which cross-correlation measurements were obtained.

for the 37 a of pick-based double-difference locations (gray bars in Figure 6b, right) also reveals a concentration of events in the upper and lower layer, providing evidence over a few decades in support of a double seismic zone. The distribution of depths for events with $M_w \geq 5$ show the separation of ~ 10 km between the two layers even more pronounced. The selection for larger magnitudes ensures that the events are well constrained, in particular relative depths, by phases observed at teleseismic distances, thus removing events with higher location uncertainty that may obscure any underlying structure.

[33] The histogram in Figure 6c shows that the DSZ is even more clearly resolved in the cross-correlation-based relocations (gray bars). The depth distribution of strongly correlated events (black bars), however, shows a somewhat narrower separation between the two peaks in the histogram. Since this selection criteria favors pairs of events that are nearby and have similar focal mechanisms, the deeper peak in the histogram may actually represent earthquakes that occur on preexisting normal faults that connect the upper and lower layer of the DSZ [Rietbrock and Waldhauser, 2004; Bohnenstiehl et al., 2005]. Compared to the depth distribution of the ISC and EHB locations (Figure 6d), the teleseismic double-difference results show a higher-resolution image of the WBZ seismicity, confirming the existence of a narrowly spaced double-seismic zone in this area. It may therefore be possible to use double-difference methods to image currently undetected double or triple seismic zones worldwide, and to resolve the detailed structure associated with them.

5. Application to Crustal Earthquakes in the 1999 Izmit-Düzce Sequence

5.1. Data

[34] We choose 75 events in the 1999 Izmit-Düzce earthquake sequence to investigate the performance of cross-correlation and double-difference methods to relocate hypocenters in the Earth's crust. The events span a distance

of nearly 200 km along the northernmost strand of the North Anatolian fault system, making them well suited to investigate the effect of fault structure, interevent distances, and waveform similarity on double-difference solutions. The Izmit Mw 7.4 main shock occurred on 17 August 1999 and was centered at 40.748°N , 29.864°E at a depth of 17 km (U.S. Geological Survey, USGS). It ruptured approximately 60 km of the surface in an almost pure right-lateral strike-slip fashion (USGS centroid moment tensor (CMT), strike/dip/slip = 95/81/180; Harvard CMT: 91/87/164). On 12 November 1999 a second earthquake with Mw 7.1 occurred about 100 km to the east of the Izmit event near the village of Düzce, at 40.758°N , 31.161°E and 10 km depth (USGS). Again, CMT solutions indicate almost pure right-lateral strike slip (USGS CMT: 269/73/177; Harvard CMT: 268/54/167). These two main shocks and 40 aftershocks ($3.8 < M_w < 5.8$) from the Izmit event and 33 aftershocks ($4.0 < M_w < 5.5$) from the Düzce event are used in our relocation analysis.

[35] We use phase picks and initial locations as listed in the EHB bulletin [Engdahl et al., 1998; B. Engdahl, personal communication, 2005]. From a total of 11,780 first and later arriving P and S body wave phase picks, 91,783 picks have been pair-wise observed at common stations (Figure 10). Ten stations are within local distances (< 200 km) from the cluster centroid, most of them locating south of the fault and therefore causing a primary station gap $> 180^\circ$ for most of the 75 events. Seventy-one events were recorded at one or more local stations, 37 events at 4 or more local stations. 227 stations are within regional distances (< 2000 km), with most stations located in Greece and western Europe. 383 stations recorded the events at teleseismic distances (> 2000 km). We have subsampled the station distribution for each event pair in order to avoid strong spatial clustering of partial derivatives in areas with dense seismic networks (e.g., arrays or local networks reporting to the ISC). We select only the best station (i.e., the highest quality pick) within bins of $3^\circ \times 3^\circ$ beyond a distance of 200 km from the cluster centroid.

[36] In addition to the phase pick data we compute accurate differential times by performing a total of 21,812 cross correlations on 561 filtered (0.1–2 Hz) seismograms obtained from IRIS's DMC, using the time domain method of Schaff et al. [2004]. We choose window length of 10 s around the predicted P, PKP, S, and SKS phase arrival time, and search over lags of 5 s. Similar to the Chile application a STA/LTA filter (1s/5s) is applied to the seismograms before the correlation measurements were carried out. A total of 1977 P wave and 49 S wave correlations had correlation coefficients $CC \geq 0.7$. The percentage of similar event pairs is 9% which is less than that for the Chile earthquakes (14%). Most of the events that correlate can be grouped into doublets and triplets. Careful inspection of seismograms for pairs of events with nearby hypocenters indicate that the waveforms observed at common stations are dissimilar, suggesting that variation in source mechanisms may be the reason for the low percentage of correlation measurements.

5.2. Relocation and Residual Analysis

[37] A series of 50 dynamically weighted damped least squares iterations is used to simultaneously relocate the

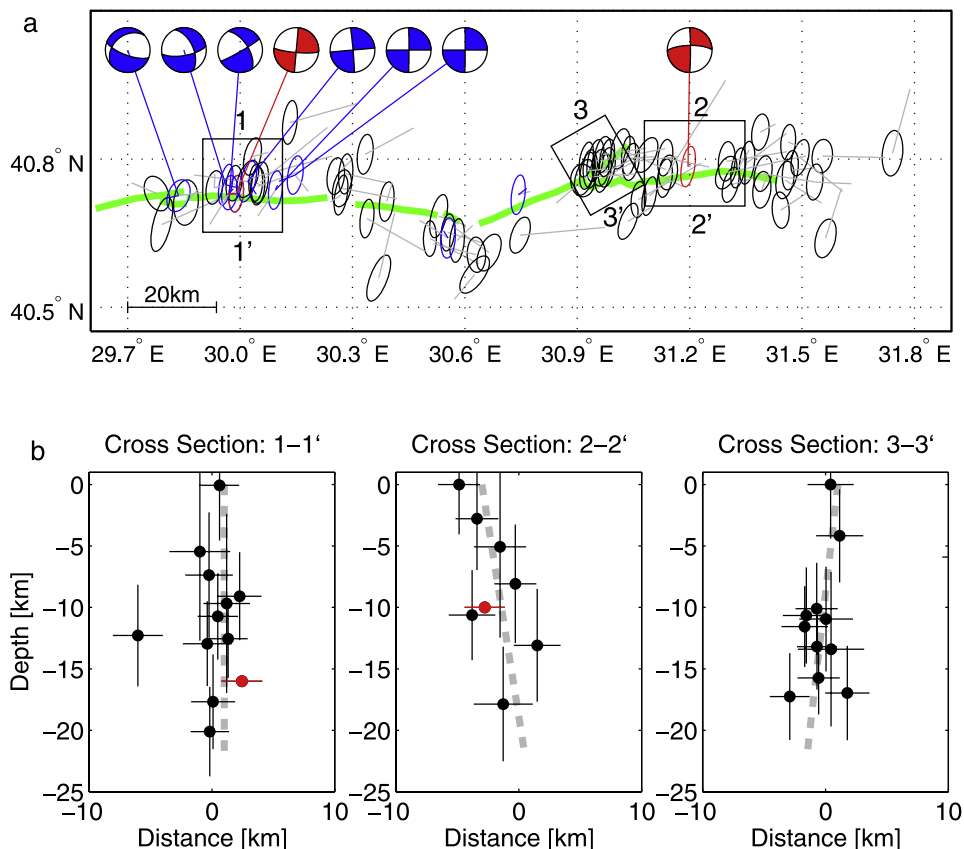


Figure 11. Double-difference locations in (a) map view and (b) fault perpendicular cross sections for the 75 events in the Izmit-Düzce sequence. Error estimates at the 90% confidence level are represented by ellipses in Figure 11a and crosses in Figure 11b. The Izmit and Düzce main shocks are shown in red; the locally recorded reference events of *Özalaybey et al.* [2002] are shown in blue. Gray lines connect DD locations to corresponding initial locations taken from the EHB catalog, red and blue lines to main shock and reference event locations, respectively. Focal mechanisms are shown for the two main shocks (red) (USGS CMT) and for six reference events (blue) [*Özalaybey et al.*, 2002]. Boxes in Figure 11a indicate location and orientation of cross sections shown in Figure 11b. Green line in Figure 11a denotes surface trace of the fault rupture [from *Pucci et al.*, 2006]. Dashed gray lines in Figure 11b represent fault dip inferred from relocated aftershocks.

75 events using the combined pick and correlation data. Interevent distance thresholds are gradually decreased from 180 km during the first iterations to 25 km during the final iterations. We invert for all hypocentral parameters (including depths) except for the depths of the two main shocks, which we fix at 17 km for the Izmit and 10 km for the Düzce shock. Note that most depths in the EHB catalog are fixed at a default value of 10 km. Before relocation, the number of links established between each event and its neighboring events in the cluster range from 759 to 4046. During relocation the data is reduced by the weighting function to between 73 and 1326 highest quality links per event. The double-difference results are shown in Figure 11 in a map view and three cross sections. Ellipses in Figure 11a and crosses in Figure 11b indicate 90% confidence levels obtained from bootstrap analysis. Error ellipses are mostly elongated in north-northeast direction, consistent with a lack of local and regional stations north of the fault. Mean location uncertainties are 3 km laterally and 4 km

vertically. Mean horizontal and vertical shifts between the initial (EHB) and relocated locations are 4.5 km and 5.1 km, respectively.

[38] The data set presented here is well suited to investigate some of the key features of the double-difference method. While JHD and HDC methods reference partial derivatives relative to the cluster centroid, their applications are limited to earthquake clusters with spatial dimensions smaller than the length scale of the velocity heterogeneity encountered by all rays between the source region and a common station (typically several kilometers). The double-difference method, on the other hand, references partial derivatives at each hypocenter with respect to a particular station (equation (1)), thus using differential times directly to solve for event separations. This approach allows large areas of seismicity to be relocated simultaneously, as long as there exists a continuous web of differential time links that connect neighboring events over distances shorter than the length scale of structural heterogeneities encountered by the

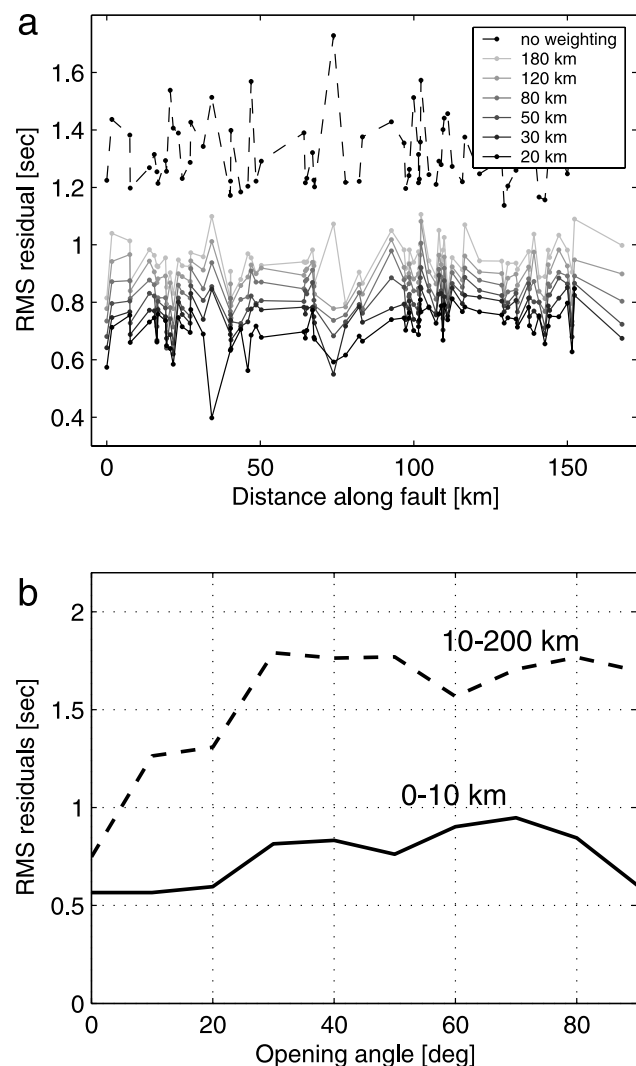


Figure 12. (a) RMS of differential time residuals shown for each event as a function of distance along the fault. Residual curves are shown for a series of double-difference inversions with different interevent distance cutoff values. (b) Binned RMS residuals as a function of the angle between an event pair’s relative position vector and the ray takeoff direction determined at the pair’s centroid, shown for two interevent distance ranges.

two rays of linked events traveling to a common station. Larger residuals and bias in double-difference solutions are therefore expected for events linked over great distances (see Figure 3).

[39] Figure 12a shows the increase in differential time residuals with increasing interevent distances for the Izmit-Düzce aftershock sequence. RMS residuals are computed for each event and its linked neighboring events for a series of double-difference runs with maximum interevent distance thresholds between 20 km and 180 km, and then plotted as a function of distance along the fault. A constant increase in the average RMS residual from about 0.7 s (20 km) to 1 s (180 km) is observed, as increasingly different raypaths introduce model errors that originate from outside the source region. Note that these results are

obtained with the distance weighting function still activated, thus data that link together events over larger distances are downweighted. If no distance weighting is used, and data links of up to 180 km are allowed, then the average RMS residual jumps to 1.3 s (Figure 12a). A small trend toward higher residuals for the Düzce aftershocks (90–160 km distance in Figure 12a) compared to the Izmit aftershocks (0–90 km) may indicate the higher complexity of fault structures along this segment of the North Anatolian fault.

[40] The increase in differential time residuals with increasing angle between the relative position vector of an event pair and the ray takeoff directions (see Figure 1a) is shown in Figure 12b, using stations out to 1000 km. For events separated between 0 and 10 km the RMS residuals generally increase from 0.6 s when the raypaths leave the source area in direction of the relative position vector of the event pairs to about 0.9 s when the raypaths leave at an angle perpendicular to the relative position vectors. The increase in residuals with increasing angle becomes significant when events are linked over distances greater than 10 km, ranging from 0.75 s when the rays leave in direction of the event pairs’ relative position vectors to about 1.75 s, well above the noise level of the differential time data, when they leave perpendicular to them (Figure 12b). The new interevent distance weighting scheme described above (Figure 1b) suppresses the influence of these lower quality differential data during double-difference inversions, thus increasing location precision while at the same time maintaining the link coverage between the events.

5.3. Evaluation Using Local Network Data and Surface Information

[41] We evaluate the accuracy of the relocation results with respect to locations of events determined in special studies using dense local networks [Özalaybey *et al.*, 2002] and with respect to near-surface geologic information [Pucci *et al.*, 2006]. A detailed study of the Izmit aftershock sequence by Özalaybey *et al.* [2002] produced an aftershock catalog (RMS = 0.16 s) with average horizontal and vertical errors of 1.7 km and 2.3 km, respectively. From a list of 27 well-located larger magnitude aftershocks [Özalaybey *et al.*, 2002, Table 2], 7 events, located near the main shock between 29.8°E and 30.2°E, are among the 75 events studied here. Accurate locations of an additional two events determined from recordings at a local temporary seismic network (N. Seeber and J. Armbruster, personal communication, 2005) are also used for comparison. The first event is an aftershock of the Izmit main shock and locates at approximately 30.55°E in Figure 11a slightly south of the surface trace. The second is an aftershock of the Düzce main shock and occurred at 30.75°E close to the fault. A comparison between these nine accurately located “reference” events and their corresponding double-difference solutions indicates mean horizontal and vertical differences of 2.4 and 1.9 km, respectively. Except for one event, all corresponding error ellipses overlap.

[42] Differences between the double-difference solutions and locations listed in the catalogs of the EDR, ISC, and EHB are shown in Figure 13a and 12b. Median (mean) horizontal/vertical differences are 7.2/4.5 (8.6/5.3) km for EDR, 4.8/5.4 (6.4/6.1) km for ISC, and 8.2/4.5 (8.1/5.6) km for EHB locations. Note that most events in the EHB and

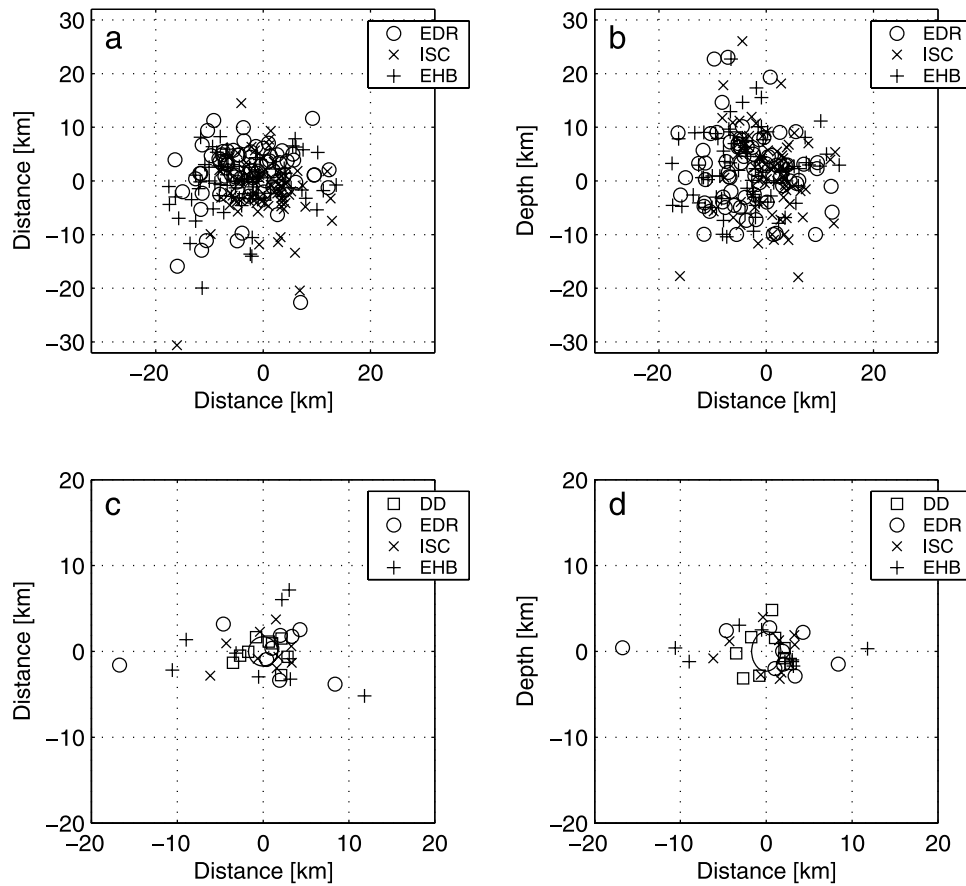


Figure 13. (a) Horizontal and (b) vertical distribution of differences between the 75 DD locations and their corresponding locations listed in the EDR, ISC, and EHB catalogs. (c) Horizontal and (d) vertical differences between the nine reference locations of *Özalaybey et al.* [2002] and the DD, EDR, EHB, and ISC locations. Ellipses indicate the approximate average error of the reference events.

EDR catalogs have depths fixed at default values. Figure 13c and 13d show the horizontal and vertical distribution of mislocations for the 9 events relative to the local network solutions of *Özalaybey et al.* [2002]. The median (mean) horizontal mislocation are 3.9 (5.5) km for the EDR, 3.3 (3.3) km for the ISC, and 6.4 (6.8) km for the EHB locations. These values are significantly greater than the median (mean) horizontal mislocation of 2.6 (2.4) km for the double-difference solutions.

[43] The overall pattern of the relocated aftershocks of both the Izmit and Düzce main shocks correlate with the general trend of the surface trace (Figure 11a). The scatter in the epicenter distribution likely reflects the complex rupture of both events, as evident from surface expressions of the main ruptures [*Barka et al.*, 2002; *Akyüz et al.*, 2002; *Pucci et al.*, 2006] and the variation of focal mechanisms (Figure 11a). Relocated aftershocks near the Izmit hypocenter indicate a 20 km deep, near vertical fault (cross section 1-1', Figure 11b) consistent with the focal mechanisms of the main shock (red, Figure 11a) and those of three aftershocks east of the main shock (blue, Figure 11a) [*Özalaybey et al.*, 2002]. The width of the seismically imaged fault is not resolvably different from zero. Large variations in focal mechanisms are observed for three aftershocks west of the main shock.

[44] In a recent study, *Pucci et al.* [2006], on the basis of detailed geologic mapping, inferred a rather complex structure of the fault associated with the Düzce earthquake sequence. Of particular interest to this study is their finding of a south dipping fault of the Cinarli segment on which the main shock occurred, and a north dipping fault for the adjacent Yenikoy segment to the west. Their results are consistent with fault perpendicular cross sections of our aftershock locations along these segments. Aftershocks along the Cinarli segment (Figure 11b, cross section 2-2') indicate a 10° south dipping fault, while aftershocks along the Yenikoy segment show an approximately 5° north dipping fault (cross section 3-3'). The situation along the Yenikoy segment is complicated as some of the events may have reactivated a strand that ruptured during the Izmit event slightly to the north of the Yenikoy segment.

[45] The general complexity of the fault structure, expressed by the aftershock locations, the variation in focal mechanisms and the complex surface ruptures, appears to be the main reason for the low number of earthquakes with similar waveforms at common stations, and thus the relatively low number of cross-correlation differential time measurements. However, the combination of phase pick and available cross-correlation data is sufficient to resolve relative depths to the extent that we can image strike and dip

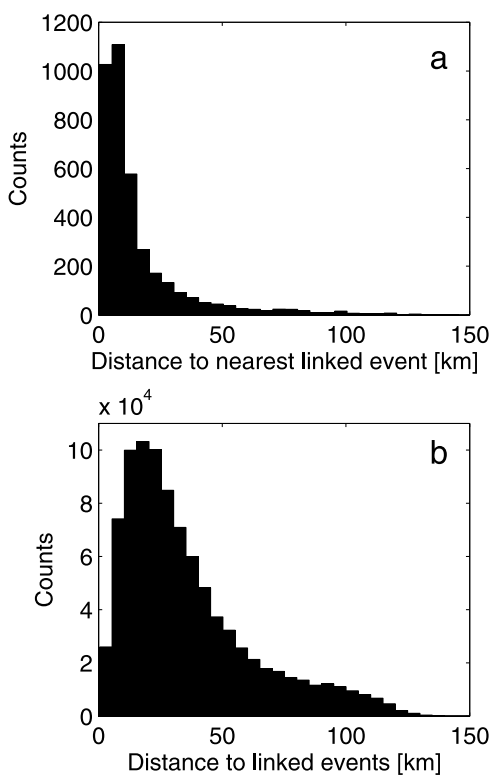


Figure 14. Distribution of distances from each of the 3783 hypocenters in the South American subduction zone shown in Figure 4b (a) to their nearest linked neighbor and (b) to all events linked during the final inversion (interevent distance cutoff is 150 km).

of active fault planes from aftershock data recorded at global seismic networks.

6. Discussion and Conclusion

[46] In this study we have chosen two earthquake data sets that are different in terms of both the type of earthquakes (long term seismicity versus aftershocks) and the tectonic environment in which they occur (subduction versus crustal strike slip), to demonstrate improvements in global event locations that can be achieved by the double-difference method. It is evident from these and previous local-scale applications that some of the most crucial parameters of the double-difference method are the quality and strength of differential time links that connect hypocenters, and the dynamically assigned weights that control these links during the least squares inversions. While the quality of the differential times is mainly controlled by the accuracy of the delay time measurements (e.g., automated picks, hand picks, cross-correlation measurements), the link strength is mainly controlled by the number of stations and the effective distance over which hypocenters are linked to their nearest neighbors.

[47] In the South American subduction zone (Figure 4), for example, the dominant links that connect each event to its single nearest neighbor feature a median distance of only 9 km, and the distribution of interevent distances for all linked events has a median of 28 km (Figure 14). Our synthetic experiment shows that such a hypocenter density

is well suited for the double-difference method to remove most of the model error associated with standard Earth models. The density of earthquakes and their occurrence on preexisting faults in this area cause waveforms to be similar at common stations, and therefore enable substantial improvements in bulletin-based relative arrival times by waveform cross correlation (Figure 8).

[48] In comparison, the complexity of the North Anatolian fault system and the resulting diversity of waveforms for events in the Izmit/Düzce sequence hamper the improvement of the bulletin data via the application of cross-correlation methods. However, the median distance between each event and its nearest neighbor is only about 5 km, therefore enabling substantial reduction of model error inherent in single-event locations via application of the double-difference method. This example indicates the importance of continuing to generate and archive high-quality picks of phase onsets, and to improve existing bulletins of phase picks using novel methods [e.g., *Engdahl et al.*, 1998]. The routine measurement of surface wave differential times [e.g., *Schaff and Richards*, 2004; *Ammon*, 2004]

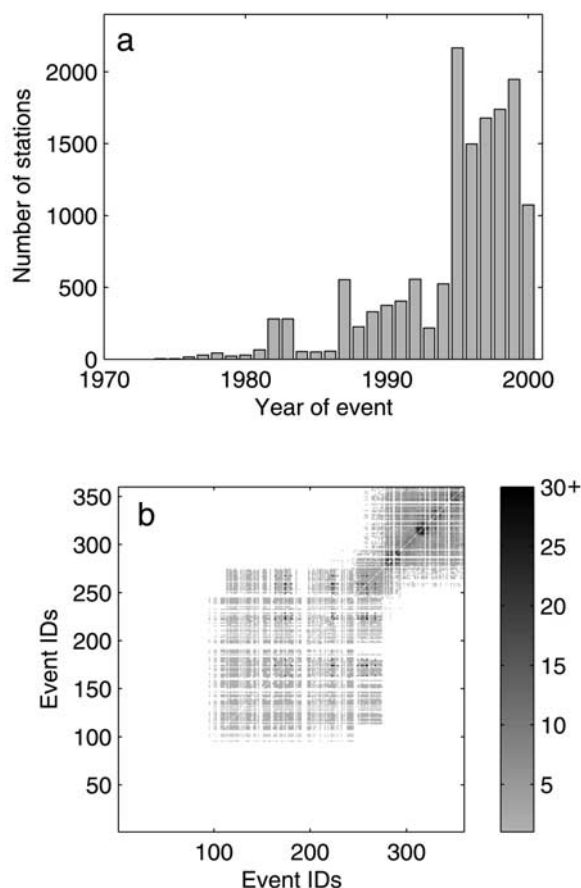


Figure 15. (a) Yearly increase in the number of digital seismograms obtained from IRIS for the 359 events (1964–2000) in northern Chile analyzed in this study. (b) Number (gray scale) and distribution of cross-correlation measurements for pairs of events with similar seismograms recorded at common stations. Event IDs are ordered by date of occurrence. Note the increase in connectivity between later events.

and their inclusion in the double-difference procedure is likely to improve the location of epicenters of crustal earthquakes, especially for events recorded at regional distances in areas where Lg waves, for example, propagate efficiently.

[49] With the increase in availability and coverage of digital seismic data worldwide, and the resulting increase in the number of earthquakes recorded, located, and archived each year in catalogs such as the ISC (~250,000 per year), consistent improvement in hypocenter locations can be expected by applying double-difference methods to earthquakes in most seismically active areas of the world. The median of all nearest neighbor distances in the ISC catalog (1964–2003) is 2.4 km, much smaller compared to 9 km in the Chile seismicity. For each event in the ISC catalog there are an average of 3868 neighbors within 100 km distance, at an average distance of 63 km. The continuing reduction in link distances to neighboring events increases the potential to reduce effects of model errors and measure accurate differential times via cross correlation to improve upon differential times formed from bulletin data. Thus the increased availability and accessibility of digital waveform data from existing and new stations through organizations like IRIS and their contributing network operators may lead to substantial improvement in existing event locations. As an example, Figure 15a shows the increase, over the last several decades, in the number of digital seismograms that are available through IRIS for the 359 events in northern Chile analyzed in this study. As a result both the number of cross-correlation measurements and the level of connectivity between events increased substantially for pairs of events where both events occurred in recent years (Figure 15b).

[50] The simultaneous use of several decades of parametric and digital seismic data relocates newer events with correlated seismograms to the accuracy of the cross-correlation data while events that do not correlate are simultaneously determined to the accuracy of the phase pick data. The operation of global, long running, high-quality digital seismic networks and maintenance of accessible waveform archives is therefore of great importance for the generation of high-resolution global seismicity catalogs. With pick and waveform data available in near-real time, waveform-based double-difference applications may become an important tool to estimate, for example, dip and orientation of active fault planes shortly after the occurrence of new events worldwide. Such information is crucial for the modeling of strong ground motions and the estimation of seismic hazard.

[51] **Acknowledgments.** We thank the personnel at the ISC for their dedicated effort in producing the ISC catalog, Bob Engdahl for producing and making available his EHB catalog, the various network operators for the waveform data used in this study, and IRIS for enabling its easy access. We thank Bill Menke for his RAYTRACE code, Mike West for help with setting up the synthetic model, and Stefano Pucci for a copy of the Izmit-Düzce fault data. We thank Mike Ritzwoller, Shamita Das, Paul Richards, and an anonymous reviewer for comments that helped to improve the manuscript. Funding for this research was supported by NSF/EAR grant 02-29832 and USGS/NEHRP grant 05HQGR0175. This is Lamont-Doherty Earth Observatory contribution 7098.

References

Akyüz, H. S., R. Hartleb, A. Barka, E. Altunel, G. Sunal, B. Meyer, and R. Armijo (2002), Surface rupture and slip distribution of the 12 November 1999 Düzce earthquake (M 7.1), North Anatolian Fault, Bolu, Turkey, *Bull. Seismol. Soc. Am.*, *92*, 61–66.

- Ammon, C. J. (2004), Relative earthquake location using surface waves, *Eos Trans. AGU*, *85*(47), Fall Meet. Suppl., Abstract S11B-1015.
- ANCORP Working Group (1999), Seismic reflection image revealing offset of Andean subduction-zone earthquake locations into oceanic mantle, *Nature*, *397*, 341–344.
- Barka, A., et al. (2002), The surface rupture and slip distribution of the 17 August 1999 Izmit earthquake (M 7.4), North Anatolian Fault, *Bull. Seismol. Soc. Am.*, *92*, 43–60.
- Bohnenstiehl, D. R., F. Waldhauser, M. Nedimovic, and A. Rietbrock (2005), Outer-Rise Faulting, Abyssal Fabric and the Structure of Double Seismic Zones, *Eos Trans. AGU*, *86*(52), Fall Meet. Suppl., Abstract T13B-0476.
- Deichmann, N., and M. Garcia-Fernandez (1992), Rupture geometry from high-precision relative hypocenter locations of microearthquake clusters, *Geophys. J. Int.*, *110*, 501–517.
- Douglas, A. (1967), Joint epicenter determination, *Nature*, *215*, 47–48.
- Engdahl, E. R., R. van der Hilst, and R. Buland (1998), Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.*, *88*, 722–743.
- Frohlich, C. (1979), An efficient method for joint hypocenter determination for large groups of earthquakes, *Comput. Geosci.*, *5*(3–4), 387–389.
- Geiger, L. (1910), Herdbestimmung bei Erdbeben aus den Ankunftszeiten, *K. Ges. Wiss. Göttingen*, *4*, 331–349.
- Got, J.-L., J. Fréchet, 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.
- Hacker, B. R., S. M. Peacock, G. A. Abers, and S. D. Holloway (2003), Subduction factory 2. Are intermediate-depth earthquakes in subducting slabs linked to metamorphic dehydration reactions?, *J. Geophys. Res.*, *108*(B1), 2030, doi:10.1029/2001JB001129.
- Hauksson, E., and P. Shearer (2005), Southern California hypocenter relocation with waveform cross-correlation, part 1: Results using the double-difference method, *Bull. Seismol. Soc. Am.*, *95*, 896–903, doi:10.1785/0120040167.
- Jordan, T. H., and K. A. Sverdrup (1981), Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, *Bull. Seismol. Soc. Am.*, *71*, 1105–1130.
- Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.*, *122*, 108–124.
- Kirby, S. H., E. R. Engdahl, and R. Denlinger (1996), Intermediate-depth intraslab earthquakes and arc volcanisms as physical expression of crustal and uppermost mantle metamorphism in subducting slabs, in *Subduction: Top to Bottom*, *Geophys. Monogr. Ser.*, vol. 96, edited by G. E. Bebout et al., pp. 195–214, AGU, Washington, D. C.
- Lin, G., and P. Shearer (2005), Tests of relative earthquake location techniques using synthetic data, *J. Geophys. Res.*, *110*, B04304, doi:10.1029/2004JB003380.
- Menke, W. (2005), Case studies of seismic tomography and earthquake location in a regional context, in *Seismic Earth: Array Analysis of Broadband Seismograms*, *Geophys. Monogr. Ser.*, vol. 157, edited by A. Levander and G. Nolet, pp. 7–36, AGU, Washington, D. C.
- Özalaybey, S., M. Ergin, M. Aktar, C. Tapirdamaz, F. Biçmen, and A. Yörük (2002), The 1999 Izmit earthquake sequence in Turkey: Seismological and tectonic aspects, *Bull. Seismol. Soc. Am.*, *92*, 376–386.
- Pavlis, G. L., and J. R. Booker (1983), Progressive multiple event location (PMEL), *Bull. Seismol. Soc. Am.*, *73*, 1753–1777.
- Poupinet, G., W. L. Ellsworth, and J. Fréchet (1984), Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras fault, California, *J. Geophys. Res.*, *89*, 5719–5731.
- Prejean, S., W. Ellsworth, M. Zoback, and F. Waldhauser (2002), Fault structure and kinematics of the Long Valley Caldera region, California, revealed by high-accuracy earthquake hypocenters and focal mechanism stress inversions, *J. Geophys. Res.*, *107*(B12), 2355, doi:10.1029/2001JB001168.
- Pucci, S., N. Palyvos, C. Zabczi, D. Pantosti, and M. Barchi (2006), Co-seismic ruptures and tectonic landforms along the Düzce segment of the North Anatolian Fault Zone (M , 7.1, November 1999), *J. Geophys. Res.*, *111*, B06312, doi:10.1029/2004JB003578.
- Richards-Dinger, K. B., and P. M. Shearer (2000), Earthquake locations in southern California obtained using source-specific station terms, *J. Geophys. Res.*, *105*, 10,939–10,960.
- Rietbrock, A., and C. Haberland (2001), A tear in the subducting Nazca slab: Evidence from local earthquake tomography and high precision hypocenters, *Eos Trans. AGU*, *82*(47), Fall. Meet. Suppl., Abstract T31A-0822.
- Rietbrock, A., and F. Waldhauser (2004), A narrowly spaced double-seismic zone in the subducting Nazca plate, *Geophys. Res. Lett.*, *31*, L10608, doi:10.1029/2004GL019610.

- Ritzwoller, M. H., N. M. Shapiro, A. L. Levshin, E. A. Bergman, and E. R. Engdahl (2003), Ability of a global three-dimensional model to locate regional events, *J. Geophys. Res.*, *108*(B7), 2353, doi:10.1029/2002JB002167.
- Schaff, D. P., and P. G. Richards (2004), Lg-wave cross correlation and double-difference location: Application to the 1999 Xiuyan, China, sequence, *Bull. Seismol. Soc. Am.*, *94*, 867–879.
- Schaff, D. P., and F. Waldhauser (2005), Waveform cross-correlation-based differential travel-time measurements at the Northern California Seismic Network, *Bull. Seismol. Soc. Am.*, *95*, 2446–2461.
- Schaff, D. P., G. H. R. Bokelmann, G. C. Beroza, F. Waldhauser, and W. L. Ellsworth (2002), High-resolution image of Calaveras Fault seismicity, *J. Geophys. Res.*, *107*(B9), 2186, doi:10.1029/2001JB000633.
- Schaff, D. P., G. H. R. Bokelmann, W. L. Ellsworth, E. Zankerka, F. Waldhauser, and G. C. Beroza (2004), Optimizing correlation techniques for improved earthquake location, *Bull. Seismol. Soc. Am.*, *94*, 705–721.
- Snoke, J. A., and J. C. Lahr (2001), Locating earthquakes: At what distance can the Earth no longer be treated as flat?, *Seismol. Res. Lett.*, *72*(5), 538–541.
- Waldhauser, F. (2001), HypoDD: A computer program to compute double-difference earthquake locations, *U.S. Geol. Surv. Open File Rep.*, *01-113*.
- Waldhauser, F., and W. L. Ellsworth (2000), A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, *Bull. Seismol. Soc. Am.*, *90*, 1353–1368.
- 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.
- Waldhauser, F., D. P. Schaff, P. G. Richards, and W.-Y. Kim (2004), Lop Nor revisited: Nuclear explosion locations, 1976–1996, from double-difference analysis of regional and teleseismic data, *Bull. Seismol. Soc. Am.*, *94*, 1879–1889.
- Wolfe, C. J. (2002), On the mathematics of using difference operators to relocate earthquakes, *Bull. Seismol. Soc. Am.*, *92*, 2879–2892.
- Zhang, J., X. Song, Y. Li, P. G. Richards, X. Sun, and F. Waldhauser (2005), Inner core differential motion confirmed by earthquake waveform doublets, *Science*, *309*, 1357–1360.

D. Schaff and F. Waldhauser, Lamont-Doherty Earth Observatory, Columbia University, P.O. Box 1000, Palisades, NY 10964, USA. (felixw@ldeo.columbia.edu)