Bulletin of the Seismological Society of America, Vol. 88, No. 3, pp. 722–743, June 1998

Global Teleseismic Earthquake Relocation with Improved Travel Times and Procedures for Depth Determination

by E. Robert Engdahl, Rob van der Hilst, and Raymond Buland

Abstract We relocate nearly 100,000 events that occurred during the period 1964 to 1995 and are well-constrained teleseismically by arrival-time data reported to the International Seismological Centre (ISC) and to the U.S. Geological Survey's National Earthquake Information Center (NEIC). Hypocenter determination is significantly improved by using, in addition to regional and teleseismic *P* and *S* phases, the arrival times of *PKiKP, PKPdf,* and the teleseismic depth phases *pP, pwP,* and *sP* in the relocation procedure. A global probability model developed for later-arriving phases is used to independently identify the depth phases. The relocations are compared to hypocenters reported in the ISC and NEIC catalogs and by other sources. Differences in our epicenters with respect to ISC and NEIC estimates are generally small and regionally systematic due to the combined effects of the observing station network and plate geometry regionally, differences in upper mantle travel times between the reference earth models used, and the use of later-arriving phases. Focal depths are improved substantially over most other independent estimates, demonstrating (for example) how regional structures such as downgoing slabs can severely bias depth estimation when only regional and teleseismic *P* arrivals are used to determine the hypocenter. The new data base, which is complete to about *Mw* 5.2 and includes all events for which moment-tensor solutions are available, has immediate application to high-resolution definition of Wadati–Benioff Zones (WBZs) worldwide, regional and global tomographic imaging, and other studies of earth structure.

Introduction

Although useful for seismic hazard assessment, global compilations of earthquake hypocenters and associated phase arrival times and residuals are often too inconsistent to be confidently applied to problems such as earth structure determination. The main difficulties are phase misidentification and the regionally varying level of hypocenter mislocation, particularly in focal depth, introduced by errors in the reference earth model and unmodeled effects of lateral heterogeneity. In earlier studies, we have focused our efforts toward better definition of empirical travel-time curves for later-arriving seismic phases and concurrently to developing and evaluating more suitable reference earth models (Kennett and Engdahl, 1991; Kennett *et al.*, 1995).

In this article, we describe procedures that can be used to extract a data base of teleseismically well-constrained hypocenters from global catalogs, to reduce hypocenter mislocation errors (or at least make them more regionally uniform), and to simplify the identification of later phases. The resulting data base of nearly 100,000 hypocenters for the period 1964 to 1995, which is complete to about $Mw = 5.2$ and includes all events for which moment-tensor solutions are available, has immediate application to high-resolution definition of Wadati–Benioff Zones (WBZs) worldwide, to regional (subduction zones) and global tomographic imaging, and to other studies of earth structure.

Data

Hypocentral parameters and associated phase data and residuals that are routinely reported by international agencies are generally available in digital form back to 1964. These data sets were developed in the following manner. Initially, reported arrival times of seismic phases are compiled and processed at the U.S. Geological Survey's National Earthquake Information Center (NEIC) to produce the weekly publication *Preliminary Determination of Epicenters* (PDE). Several months later, with the addition of much more data, the NEIC data set is reprocessed to produce monthly summaries in the form of the *Monthly Listing* and associated *Earthquake Data Report.* All data processed at the NEIC are eventually transferred to the International Seismological Centre (ISC) where significant numbers of additional phase

data are collected and compiled to produce an even larger data set. This data set is reprocessed at the ISC, additional hypocenters determined, and the resulting hypocentral parameters and associated phase data published in the monthly *ISC Bulletin,* usually scheduled 2 years in arrears to be as complete as possible.

ISC Bulletin data for the period January 1964 to August 1987 (excluding the time periods 19 to 31 July 1971 and 15 to 31 July 1974, which were inadvertently omitted) have been available for some time now on a CD-ROM produced by the U.S. Geological Survey in cooperation with the ISC. Data for the period September 1987 through December 1993 are currently available from the ISC on CD-ROMs. More recent data, which have not yet been processed and made available by the ISC, must be obtained in less complete form from the NEIC. For the purposes of this study, a combined ISC and NEIC data set for the period 1964 to 1995 was compiled in the standard ISC 96-byte format and reprocessed. This data set also includes arrival-time data from several temporary deployments of broadband stations [e.g., BANJO and SEDA in South America, 1994, Beck *et al.* (1994), and SKIPPY in Australia, 1993 to 1995, Van der Hilst *et al.* (1994)].

To determine hypocentral parameters, the ISC uses a standard least-squares procedure based on Jeffreys' method of uniform reduction (Jeffreys, 1932, 1939; Adams *et al.,* 1982; Bolt, 1960; Buland, 1976, 1986) and *P*-wave travel-time tables derived from the radially stratified Jeffreys–Bullen (JB) earth model (Jeffreys and Bullen, 1940). The resulting hypocentral parameters are based entirely on reported first-arriving *P*-wave times, which for most events do not include *P*-wave arrivals corresponding to upgoing ray paths. Hence, many ISC hypocenters are poorly constrained in focal depth. Note that for a substantial number of events, focal depths independently based on *pP-P* differential times are reported as well. However, reported phase travel-time residuals are based entirely on hypocenters determined solely by the *P* arrivals, which limits their application to research problems such as the evaluation of earth models. The NEIC uses a procedure, an earlier version of which is described in Engdahl and Gunst (1966), that is quite similar to that used by the ISC, except that the Bolt (1968) travel-time tables for *PKP* are used in place of the JB tables, and analysts may use *pP-P* times or other information to constrain focal depths from which the reported phase residuals are derived.

Travel-Time Tables

The standard travel-time tables used by the ISC and the NEIC are the JB tables published in 1940. Although the limitations of these tables have been known for some time, until recently no other tables could provide such a complete representation of the *P, S, PKP,* and later-arriving phases. In 1987, the International Association of Seismology and Physics of the Earth's Interior (IASPEI) initiated a major international effort to construct new global travel-time tables for earthquake location and phase identification. Two models resulted from this effort: iasp91 (Kennett and Engdahl, 1991) and SP6 (Morelli and Dziewonski, 1993). Although differences in predicted travel times between these two models were small, some effort was still required to reconcile the travel times for some important, well-observed seismic phases before either of these models could be used for routine earthquake location by NEIC and ISC.

The most significant differences between these new models and the older JB model are in the upper mantle and core. The upper mantle is highly heterogeneous, and velocities and major discontinuities in the upper mantle of recent models such as iasp91 are set at values that give an effective average representation of velocities out to 25° (cf. Kennett and Engdahl, 1991). Compared to the JB tables, the core models for iasp91 and SP6 more accurately predict the observed travel times of later-arriving core phases that bottom in the lowermost part of the outer core. Also, the baseline for the travel times of teleseismic *S* waves in both models appears to be in better agreement with *S*-wave data than the JB model.

The iasp91 and SP6 velocity models were developed to provide a means of representing the times of arrival of major seismic phases for the purpose of earthquake location and phase identification. Subsequently, new empirical travel-time curves for all the major seismic phases were derived from *ISC Bulletin* data by using the procedures described in this article and a modified iasp91 model (modified to conform to the SP6 core) as a reference to relocate a set of geographically well-distributed events. The resulting set of smoothed empirical times was then used to construct an improved model for the *P* and *S* radial velocity profile of the Earth (ak135; Kennett *et al.,* 1995). The primary means of computing travel times from such models is based on a set of algorithms (Buland and Chapman, 1983) that provide rapid calculation of the travel times and derivatives of an arbitrary set of phases for a specified source depth and epicentral distance. In the mantle, ak135 differs from iasp91 only in the velocity gradient for the D″ layer and in the baseline for *S*-wave travel times (about −0.5 sec). Significant improvement in core velocities relative to earlier model fits was also realized. Inner core anisotropy is not accounted for in the ak135 model. However, there are so few reported arrivals of *PKPdf* at large distances along the Earth's spin axis that the effects of this anisotropy in earthquake location are negligible. In the application described in this article, ak135 has proved very suitable for predicting the arrival times of a wide variety of seismic phases for use in event location and phase identification procedures.

Relocation

On a global scale, the distribution of earthquakes and seismological stations is highly heterogeneous. Most earthquakes occur in or near subduction zones where lateral variations in seismic velocities of 8 to 10% are not uncommon.

Seismological stations are mostly located in continental areas and on some islands in oceanic regions. Lateral variations in velocity, uneven spatial distribution of seismological stations, magnitude-dependent observational uncertainties, and the specific choice of which seismic data subset to use can easily combine to produce earthquake mislocations and errors in focal depth of several tens of kilometers (Engdahl *et al.* 1977, 1982; Engdahl and Gubbins, 1987; Dziewonski and Anderson, 1983; Adams, 1992; van der Hilst and Engdahl, 1992). Moreover, source mislocation is often due to bias (e.g., produced by systematic distance-dependent errors in the JB travel-time tables) that masks structural signal in the residuals (Engdahl and Gubbins, 1987). Structural signal is defined here as the portion of travel-time residuals arising from the difference between the Earth's actual velocity structure and the reference model. Source mislocation is well known to mask structural signal. For example, Davies (1992) has shown that mislocation contributes up to 35% to the travel-time variance signal at teleseismic distances for earthquakes.

We argue that the bias in hypocenter determination can be significantly reduced and at least part of the lost structural signal (having been mapped into bias) recovered by limiting the events of interest only to those that are well-constrained teleseismically, by using a reference earth model close to the true globally averaged earth structure, and by including later-arriving phases in the relocation procedure. In this article, we select for relocation only events for which the largest teleseismic open azimuth is less than 180°. That is, the largest range of epicenter-to-station azimuth with no stations at teleseismic distances (epicentral distances greater than 30°) is less than 180°. Moreover, in the relocation procedure, we use the ak135 earth model and reported arrival times for first-arriving *P* and *S* phases (excluding distances greater than 100° for *P* and 80° for *S*), *PKiKP* (at distances greater than 110°), *PKPdf* (excluding distances near the *PKP* caustic), and the depth phases *pP*, *pwP*, and *sP* (excluding distances less than 30°), which have been reidentified by procedures described later. We believe that problems in using *S* (e.g., contamination by effects of transverse isotropy cannot be avoided because there is no way of knowing whether a reported *S* arrival has been read from a horizontal or vertical component) are outweighed by the value of its constraint on depth and origin time. Thus, if we use numerous stations and seismic waves at different azimuths and distances around the source, we can attempt to average out the differences between the Earth's actual velocity structure and that of the reference model. Quantitatively, it is self-evident that constraints on the hypocenter solution can be realized by using direct phases (*P*, *PKP*, and *S*) having travel-time derivatives with respect to distance that are significantly different in magnitude and by using depth phases (*pP*, *pwP*, and *sP*) having travel-time derivatives with respect to depth that are opposite in sign to those of direct phases.

Depth phases reported for suboceanic earthquakes deserve special attention. In the relocation procedure, focal depth is largely controlled by regional *P* and *S* arrivals and depth phases. We examined residuals of reported arrivals from relocated events with depths greater than 160 km in the neighborhood of the *pP* arrival as a function of bounce-point water depth for the northwest Pacific region (Fig. 1). From the moveout of the secondary peak evident in the distribution shown, we were able to confirm that many arrivals with large positive residuals, often reported as *pP* in the ISC and NEIC data bases, were actually *pwP* phases [first identified by Mendiguren (1971); see also Yoshii (1979) and Engdahl and Billington (1986)]. In the relocation procedure described in this study, a *pwP* phase produced by a water layer 4 km thick that was misidentified as *pP* could, in the absence of regional *P* phases, result in overestimation of the focal depth by about 50 km. Although in our earlier studies (van der Hilst and Engdahl, 1991, 1992; van der Hilst *et al.*, 1991, 1993) we automatically removed these *pwP* arrivals from the hypocenter determination, reidentified *pwP* arrivals have been used to constrain depth in all subsequent studies.

Arrival times of seismic phases having absolute residuals (with travel-time corrections applied—as shown later) >7.5 sec at regional distances and >3.5 sec at teleseismic distances are not used in the relocation procedure. The larger time window for regional arrivals is necessary so that large, mostly slab-related, residuals at certain strategically located stations are not excluded. These cut-off values are somewhat subjective, but we believe that they include nearly all the expected effects of lateral heterogeneity on seismic travel times globally. In the future, compilations of individual station statistics can be used to determine whether or not a reported arrival should be used in the relocation.

Travel-Time Corrections

Ellipticity

A direct-access table for the ellipticity corrections, following the formulation of Dziewonski and Gilbert (1976) and with the procedure of Doornbos (1988), has been generated for each major phase branch (including depth phases) implemented in the Buland and Chapman (1983) software using the ak135 model (Kennett and Gudmundsson, 1996). The algorithm used to determine the ellipticity correction has been extended to include diffracted phases by simple extrapolation of the corresponding geometric rays. Linear interpolation of the table, which is set up at 5° intervals in distance and six depth levels, is used to determine the ellipticity correction based on the source-to-receiver distance and azimuth, and the source depth and co-latitude. The ellipticity correction, which can be up to 1 sec for teleseismic *P* waves, is added to the travel time computed for a spherical Earth.

Bounce-Point Topography and Water Depth

In order to determine *pwP* arrival times and correct all depth phases for topography or bathymetry at their reflection points on the Earth's surface, it is necessary to first determine

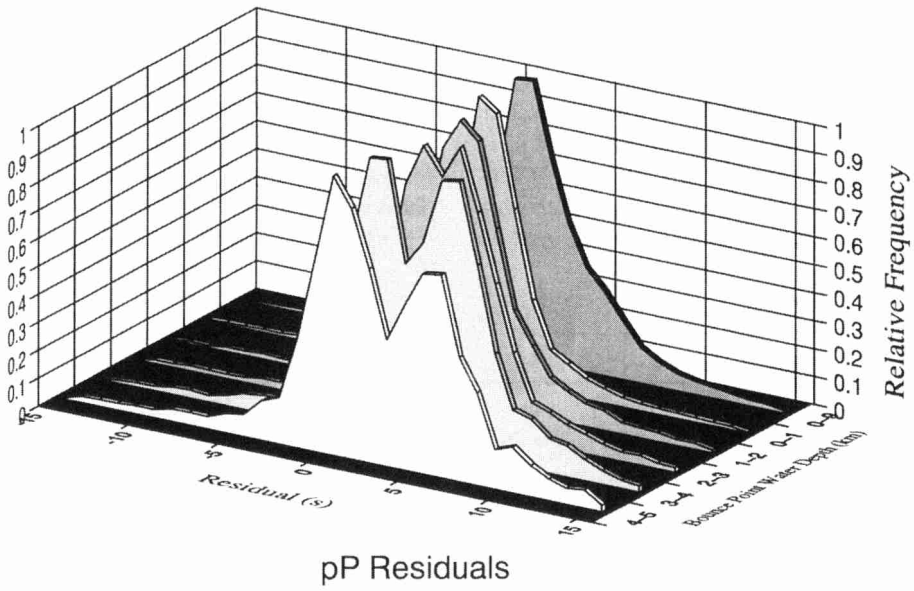


Figure 1. All possible identifications of *pP* phases over the distance range 25 to 100° (within a +/− 15-sec time window) are separated into 1-km increments of water depth at their bounce points. We further reduce the data set by excluding all *pP* phases associated with events less than 160 km in focal depth. At this depth, the *sP* phase arrives 17 sec after *pP*, which is outside the expected window for a *pwP* phase generated by a water layer 6 km or less in depth. Thus, for this data set, the problem of *pwP* phases being misidentified as *sP* is largely circumvented. However, the possibility of associating *pwP* phases as *PcP* cannot be ruled out for distances greater than about 56°, although *PcP* tends to be a weaker phase at those distances. The distribution of *pP* residuals for bounce-point water depths greater than about 2 km is clearly bimodal, as might be expected were the data set to include later-arriving *pwP* phases. For bounce-point water depths less than about 2 km, the *pP* residual distribution is slightly skewed in a positive direction.

the latitude and longitude of these bounce points and then the corresponding seafloor depth or continental elevation (see also van der Hilst and Engdahl, 1991). Bounce-point coordinates are easily computed from the distance, azimuth, and ray parameter of the depth phase (*pP* in the case of *pwP*). A version of the ETOPO5 file (National Geophysical Data Center, NOAA) was averaged over 20 × 20 minute equal-area cells and then projected on a 20 × 20 minute equi-angular cell model. The use of a smoothed version of the bathymetry is justified because the reflection of a depth phase does not take place at one single point but over a reflection zone with a size determined by the Fresnel zone of the wave. Nolet (1987) estimates the maximum half-width of a ray with a wavelength of 10 km and a ray-path length of 1000 km to be 36 km. The smoothed version of the bathymetry is interpolated bilinearly for topography or bathymetry at the bounce point. This information is used to determine the correction for bounce-point elevation or depth, which is added to the computed travel times for depth phases. Theoretical times are not computed for *pwP* phases in the case of bounce-point water depths ≦1.5 km because it is nearly impossible to separate the *pP* and *pwP* arrivals on most records (about a 2-sec separation).

Station

Because most seismic stations are in continental areas, the ak135 model was developed with a continental style for the uppermost crust and upper mantle. Of course, this form of the Earth's outermost structure is only an average, and there are significant departures from the ak135 crust and upper mantle locally. In many cases, the effects of these lateral variations in crust and upper mantle velocities can systematically displace hypocenters. To address this problem, a spatial averaging technique called patch averaging is used to compensate for coherent travel-time anomalies produced by regional differences in earth structure. In this scheme, all teleseismic rays that arrive within a given surface patch can be used to determine an average upper mantle correction in that patch with respect to the ak135 upper mantle. Patch-averaged rays should contain the coherent signal that arises due to average upper mantle structure beneath the patch, to the extent that the patch is well sampled azimuthally. Thus, patch averages can be thought of as regionally smoothed station corrections.

Corrections have been derived from *P* waves that bottom in the lower mantle for 437 five- by five-degree surface patches. In our application of the patch-averaging technique,

the median residual is first computed in each of 20° (non-overlapping) azimuthal windows (bins) for each patch. For each patch, a correction is then computed as the median of all the individual azimuthal bin medians for that patch. For a patch correction to be acceptable, at least 9 of the 18 possible azimuthal bins must have a median estimate. Otherwise, owing to uneven azimuthal sampling, the patch correction for a poorly sampled bin is set to zero. The goal of this procedure is to eliminate the effects of overweighting well-sampled azimuths in the calculation of a single azimuth-independent patch correction. Patch corrections were also independently developed for *PKP* waves. A comparison of patch medians derived from lower mantle *P* and *PKP* residuals is shown in Figure 2. Because the propagation paths of *PKP* and *P* phases in the Earth's deep interior are significantly different, the excellent correlation of the me-

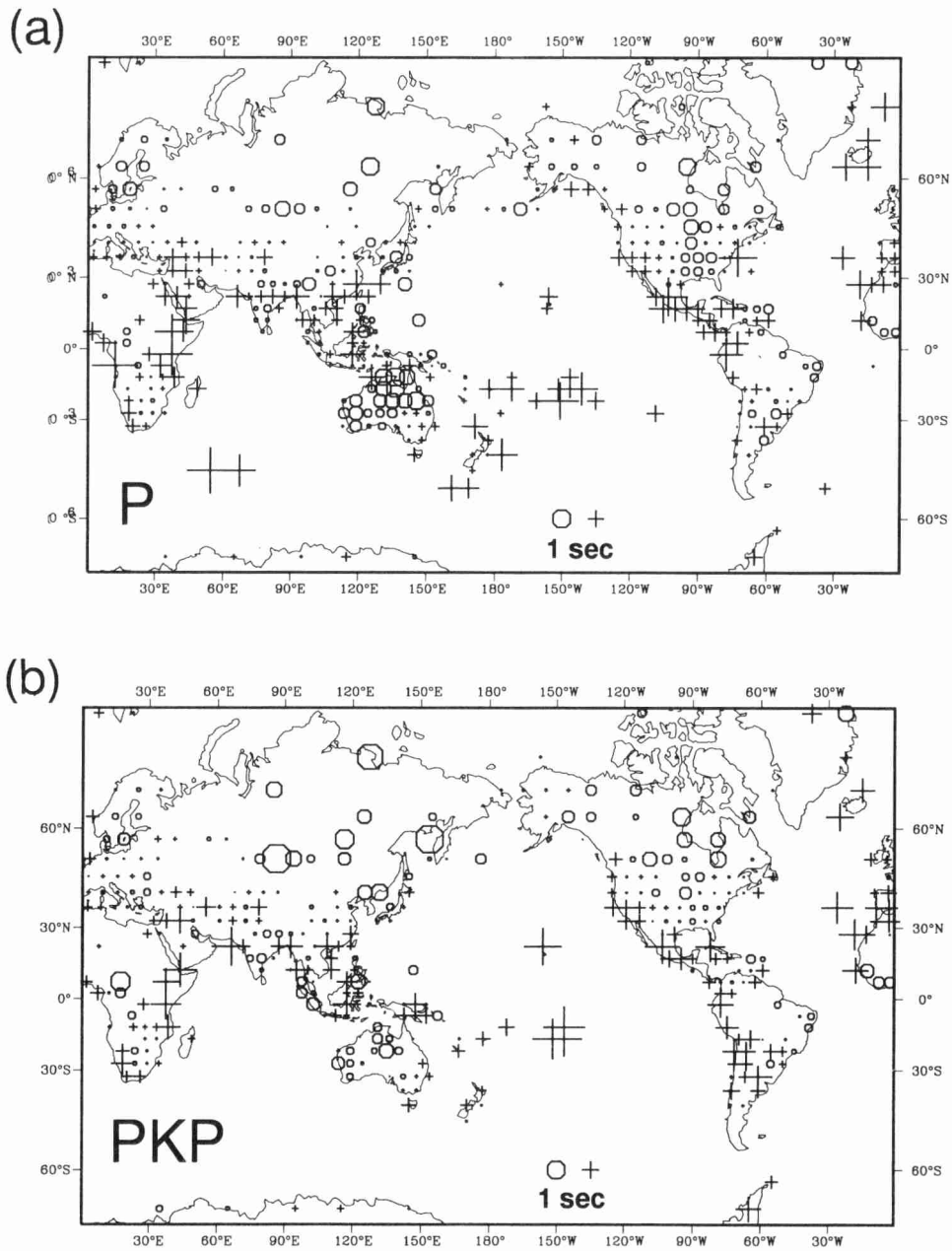


Figure 2. Upper mantle corrections found by computing patch medians (5 × 5 degree grid) of direct *P* waves bottoming in (a) the lower mantle (780 to 2740 km depths) and (b) the inner core (5153.5 to 6371 km depths). Pluses and octagons represent scaled positive and negative medians, respectively. Note that shields are generally fast (octagons) and rift zones slow (pluses). Patch medians derived independently from *P* and *PKP* data agree quite well.

dians common to both estimates indicate that they represent near-receiver structure.

The P patch medians, which range in value from -1.0 to $+2.8$ sec, are used as "station" corrections for all stations falling within a given patch. A corresponding "station" correction for S phases has been derived by assuming a Poisson's ratio of 1/4 (i.e., the S correction is the P correction times the square root of 3). Other studies of station residuals find an S-to-P correction ratio between 3 and 3.5 (e.g., Davies *et al.*, 1992). While our assumption that the ratio of time corrections due to P and S velocity heterogeneity are comparable to the ratio of P and S travel time in our 1D model may be naive, higher values depend on a degree of correlation between the P and S velocity heterogeneity that has not yet been proven in the Earth. Clearly, the relationship between P and S patch averages requires further investigation. In the relocation procedure, patch corrections are added to the theoretical travel times of *all* teleseismic phases arriving at the station as either P or S waves. This compensates for the effects of average regional differences in crust and upper mantle velocity with respect to the ak135 reference earth model at teleseismic distances.

Phase Identification

In our previous studies, phase identification has been accomplished by searching for arrivals within a time window centered on predicted phase arrival times. This approach works quite well, except in regions where there are crossing phases or the phases arrive very close in time. In particular, for shallower earthquakes, the phases *pP, sP, pwP,* and *PcP* are, without actually examining seismograms, very difficult to distinguish from one another using arrival-time data alone. Therefore, the statistical properties of these phases, as revealed in the work of Kennett *et al.* (1995) to construct the ak135 model, have been used to develop and test a new phase identification algorithm. This algorithm reduces bias in the global travel-time residual data set relative to schemes that depend on residuals alone. The method takes advantage of both the known scatter of these four phases and their relative observability. Eventually, we plan to develop a probabilistic model for identifying all later-arriving phases of interest. By providing a basis for the relative weighting of phases, the probabilistic model will also permit the use of selected later phases in calculating hypocenters.

To demonstrate the technique, we have plotted in Figure 3 the probability density functions (PDFs) for four phases centered at their theoretical relative travel times from a hypothetical deep event and arrival times for the hypothetical reported phases with unknown identifications O1, O2, and O3. Statistical studies suggest that the shapes of the *pP, sP, pwP,* and *PcP* PDFs are similar to the PDF for P. Therefore, these PDFs have been modeled as the linear combination of a Gaussian and a Cauchy distribution following Buland (1986). The parameters for the intermediate-depth distribution have been used, as they are essentially identical to the

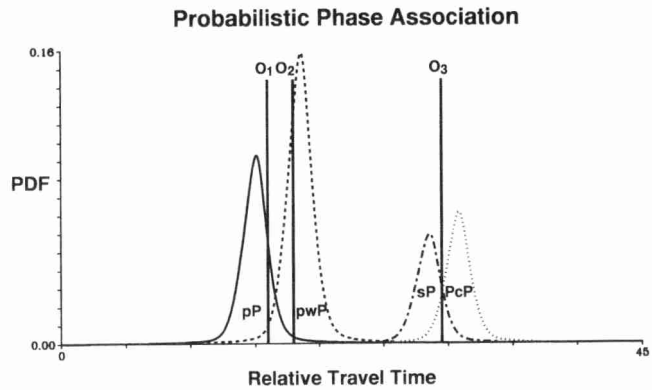


Figure 3. Probability density functions (PDFs) for four phases centered at their theoretical relative travel times from a hypothetical deep event and the observed relative arrival times for hypothetical reported phases with unknown identifications O1, O2, and O3. The relative frequencies of the phases (amplitudes of the PDFs) were developed for the Aleutian region as part of a separate study (Boyd *et al.*, 1995).

parameters for the shallow-depth distribution but without the artifact due to upper mantle triplications. The relative probability of each phase being reported was determined iteratively. For a bounce-point water depth <1.5 km, the relative probabilities were $pP = 0.60$, $sP = 0.22$, $pwP = 0$, and $PcP = 0.18$. For a bounce-point water depth $\geqq 1.5$ km, the relative probabilities were $pP = 0.26$, $sP = 0.15$, $pwP = 0.40$, and $PcP = 0.19$. The problem is to properly identify in a statistical fashion the unknown phases. Because the theoretical PDFs for the phases incorporate all known information, it is straightforward to estimate the actual relative probability of each pick being any of the theoretical phases. However, always associating each pick with the highest probability theoretical phase would be tantamount to truncating each of the theoretical probability distributions at the point where it and its neighboring phase were equally likely.

Truncation, though commonly used in the phase identification process, results in apparent distortion of the residual distributions and inevitably in bias in the global hypocenter catalog. To see how this works, consider the following thought experiment. Say that we wish to identify a number of arrivals observed over a limited distance range, all of which will be significant in locating a particular earthquake. Imagine that each of these arrivals may be identified as one of two phases whose theoretical travel times cross in this same distance range. Provided that these phases are refracted or reflected, we know observationally that the scatter of arrivals associated with them can be modeled by PDFs that are peaked near the theoretical arrival time, are approximately independent of distance, and are roughly symmetric. If we identify each of the arrivals as the phase with the nearest theoretical travel time, this is equivalent to truncating each of the phase PDFs at the point mid-way between the theoretical travel times.

Because of the truncation, each phase PDF now appears to be strongly distance dependent, becoming more and more asymmetric as the theoretical travel times converge. Summary estimates based on these phase identifications would be biased because estimators such as least squares, medians, etc., assume a degree of symmetry. For example, if these phase identifications were used to estimate "observed" travel times, the estimates for the two phases would be biased away from each other due to the apparent asymmetry of the PDFs. Similarly, the estimated location of the earthquake that generated the arrivals will be biased. This effect is truly a bias because other earthquakes in the same region located by the same set of stations would, on average, be mislocated in the same way.

To avoid the bias inherent in truncation, we have used an alternative method in which each pick is associated with a theoretical phase randomly, subject to the *a priori* information about their relative likelihoods and subject to the constraint that no two picks will be associated with the same theoretical phase. This means that if a pick has a 60% chance of being *pP* and a 40% chance of being *pwP*, it will be assigned to one of them randomly, but with a 60% chance of the association being to *pP*. Although this procedure is not guaranteed to be any more accurate than truncation in identifying each individual arrival, hypocentral estimates derived from these arrivals will be unbiased in the aggregate catalog (note that a more complete description of this algorithm is being prepared for publication elsewhere). Thus, because the statistical algorithm has only been applied to the depth phases, we expect that depth estimates in the aggregate data base relocated here are unbiased (at least due to phase identification).

Until this procedure can be fully developed, all phases other than *pP*, *pwP*, *sP*, and *PcP* are simply identified on the basis of the nearest possible phase within a time window of $+/-15$ sec of the theoretical arrival time for that phase. In the case of crossing phase travel times, either the phases are excluded entirely from further processing or selection criteria based on other factors are exercised. For example, if a station operator identifies a phase arrival as an *S* and the observed travel time falls within the theoretical time window for *S*, then the phase is processed as an *S* regardless of other possible identifications.

Weighting

Phase arrivals are weighted according to the phase type and the reported precision of the arrival time (as indicated by the ISC). A weight of 1.0 is assigned to any phase arrival time reported to the nearest tenth of a second or better. Otherwise, phase arrival times reported to the nearest second are assigned a weight of 0.7, the ratio of the observed variances for these two data precision classes. An additional multiplicative factor of 0.25 (the ratio of observed teleseismic *P*-to-*S* residual variances) is applied to the weights of teleseismic *S* waves. All weights or multiplicative factors were directly determined from the reduced data base described later as part of this study.

Event Selection

The initial combined version of the ISC and NEIC data bases contains arrival-time data for many events that are not well constrained by teleseismic data (at distances >30°). For reasons previously stated, these events may be mislocated in varying unknown ways and, until a fully three-dimensional model is available for routine earthquake location, are not very useful for the applications envisaged as a result of this study. By trial and error, we found that an event epicenter is, in general, reasonably well-constrained teleseismically when there are at least 10 usable first arrivals from teleseismic stations and when the azimuthal coverage of these stations is greater than 180°. In the sense that we use it here, "reasonable" means that the effects of lateral heterogeneity at the source on mislocation are from one event to another in that source region approximately the same, regardless of size. Surprisingly, using these criteria, relatively few events were selected from the ISC data set that had not been already selected from the NEIC data set. Selected events represent only about 15% of the total number of events in the initial combined data base. However, these events are usually larger in magnitude and contain the majority of the reported data, including most of the later-arriving phases (cf. Kennett *et al.*, 1995). Before final processing, a new reduced data base containing only the selected events was constructed and used in all subsequent processing.

Event Classification

The event selection and relocation procedures did not always insure that relocated hypocenters met our acceptance criteria. High residual variance and the lack of either regional station data or depth phases warranted a classification scheme. Standard errors in the hypocentral parameters are routinely determined from the diagonal elements of the covariance matrix. A small percentage of the lower-magnitude events had standard errors in epicenter in excess of 35 km. We continue to carry events of this type in the data base but have classified them as poor solutions. For many events, no depth phases were reported, and usable regional station data were not sufficient for constraining depth to standard errors of 15 km or less. These events were also classified as poor solutions unless independent depth information, either from waveforms or from characteristic seismicity patterns, could be used for fixed-depth solutions. For the vast majority of selected events (85%), however, we were able to determine free depth solutions that met the acceptance criteria.

Magnitudes

Nearly all the events selected and relocated had associated *mb* or *Ms* magnitudes reported, and for many of these

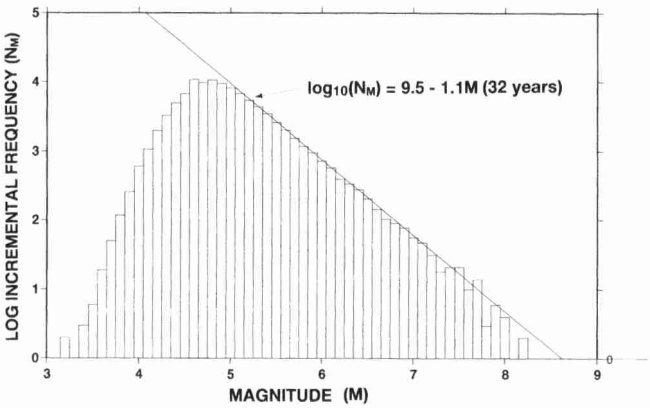


Figure 4. Log incremental frequency of events versus magnitude *M*, where *M* is, in descending order of preference, either *Mw*, *Ms* (>4.7), or *mb*, depending on availability. If no magnitude was reported, we used the relationship $M = 4.6 + 0.0047*N$, where *N* is the number or stations at teleseismic distances (i.e., >30 degrees), and the coefficients were determined by fitting *N* to reported *mb* values for lower-magnitude events. The frequency was calculated by counting events falling in magnitude interval bins that are 0.1 units wide.

events, *Mw* was available from other sources as well. To examine the completeness of the data base, we have assigned events a generic magnitude *M*, where *M* is, in descending order of preference, either *Mw*, *Ms* (>4.7) or *mb*, depending on availability. In Figure 4, we have plotted the log of the number of events falling into a series of magnitude intervals or bins. The downturn of this distribution at about *M* 5.2 suggests that the data base is complete to at least this magnitude (incompleteness at magnitudes higher than 5.2 might still be possible for remote regions of the Earth, but such incompleteness must be marginal in order not to noticeably affect the plot shown in Fig. 4).

We note that *M* is a somewhat inhomogeneous estimate of size, as *mb* can only be empirically related to *Mw* and *Ms*. However, *M* for almost all events above the completeness magnitude are based either on *Mw* or *Ms*, which are physically related. Nearly all events not meeting the acceptance criteria occurred at lower magnitudes. Events less than about 4.5 magnitude were, in most cases, deep events that have a minimal number of reporting stations at teleseismic distances but that are nevertheless well recorded due to the impulsive phase arrivals. Most events for which there were no magnitude estimates were usually either small and deeper than normal, part of an aftershock sequence or swarm, or occurred earlier in time when magnitude data was not as commonly available. Over the range where the data base is complete, the magnitudes are fitted well by the frequency–magnitude relationship $\log_{10}(N_M) = 9.5 - 1.1*M$, where N_M is the number of earthquakes in the magnitude interval ± 0.05 about *M*.

Results

Travel Times

The performance of ak135 versus the JB model in predicting travel times is best shown by comparing relocated (hereafter referred to as EHB) and ISC residual density plots for key phases. In Figures 5a and 5b, we compare ISC and EHB residuals for all first-arriving *P*-wave types (*Pg*, *Pb*, *Pn*, *P*, and *Pdiff*) from 0 to 100° distance. At all distances,

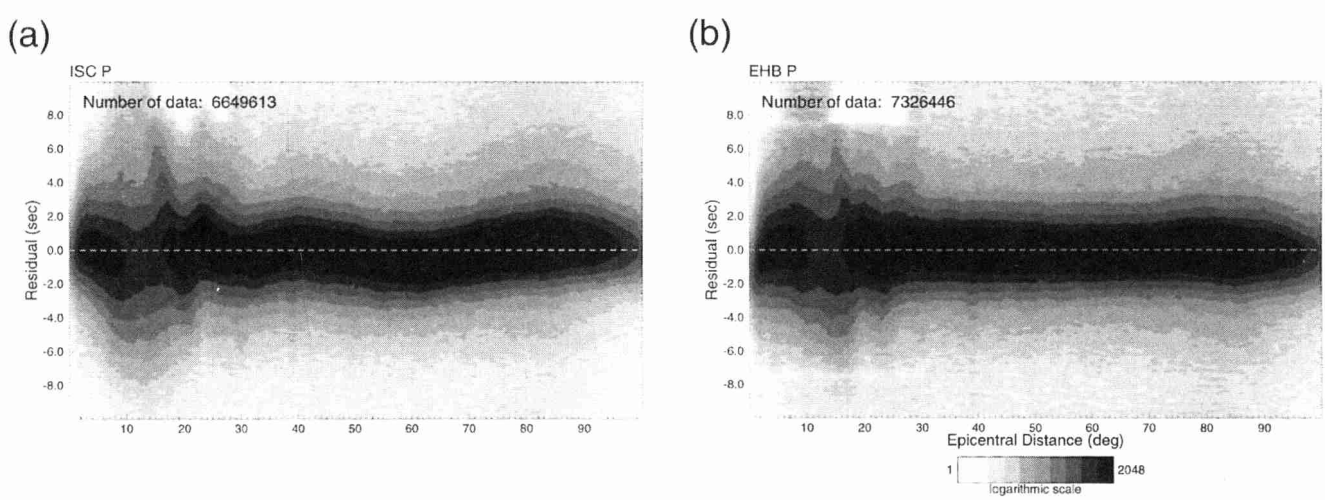


Figure 5. Residual density plots for (a) *P* residuals reported by the ISC and (b) *P* residuals determined in this study. To construct these plots, the residuals are binned in 1.0° by 0.1-sec cells and the number of hits translated into gray tones using a logarithmic scale. Obvious artifacts in the residual densities are related to the processing procedures used. For example, at distances of about 13 to 27°, EHB processing produces *P*-wave residuals that appear to be truncated at 7.5 sec because reported first arrivals have been associated with later-arriving *P* branches.

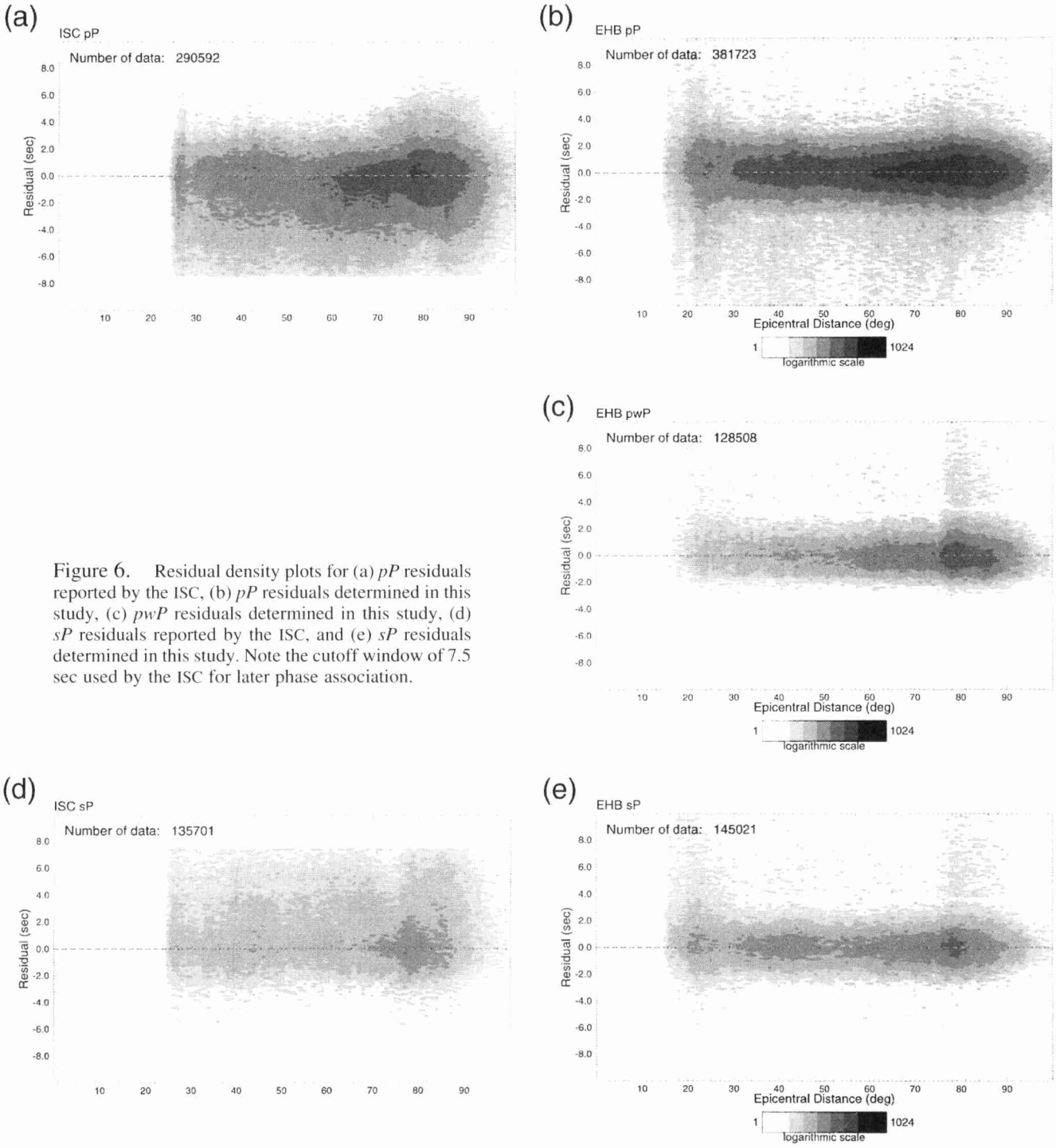


Figure 6. Residual density plots for (a) *pP* residuals reported by the ISC, (b) *pP* residuals determined in this study, (c) *pwP* residuals determined in this study, (d) *sP* residuals reported by the ISC, and (e) *sP* residuals determined in this study. Note the cutoff window of 7.5 sec used by the ISC for later phase association.

phases reported as first arrivals can, in the EHB procedure, also be associated with later-arriving *P* branches or other nearby phases provided that their residuals do not meet our criteria for first-arriving *P* (absolute value ≦3.5 sec at teleseismic distances and ≦7.5 sec at regional distances). Thus, all available data are shown in these plots unless excluded by procedures used by the ISC and EHB. Note that the statistical phase association technique that has been used here for depth phases should permit current procedural artifacts, resulting from the exclusion of data by truncation, to be elegantly removed. However, further development and extensive testing will be necessary before this methodology can be applied routinely to all phases.

At teleseismic distances, *P* residuals reported by ISC

based on JB travel times (Fig. 5a) display a (well-known) undulating structure as a function of distance, indicative of deviations between the radial velocity profile of the JB model and the true spherically averaged Earth. Such undulations are not evident in Figure 5b. Van der Hilst *et al.* (1991, 1992) have shown in a regional study that the variance of *P* residuals resulting from hypocenter relocation with the iasp91 travel times is 17% less than the variance of reported ISC *P* residuals computed using the JB tables. In this study of the global data set, we find a similar variance reduction of 18% in teleseismic *P* residuals using ak135 travel times.

Density plots for depth-phase residuals (*pP, pwP,* and *sP*) over the distance range 0 to 100° are shown in Figure 6 for ISC reported data (with the exception of *pwP*) and for data produced by this study. The median value and spread (defined below) of ISC *pP* residuals in Figure 6a is $-0.7 +/- 2.7$ sec, implying that ISC depths based only on first-reported *P*-wave arrivals are slightly overestimated and may be poorly determined (see also van der Hilst and Engdahl, 1991). Thus, ISC focal depths are not generally consistent with residuals reported for later phase arrivals identified as *pP* by ISC. A better fit to reidentified *pP* data using ak135 travel times (Fig. 6b) was expected because these data were used in the hypocenter calculation. However, the spread of the *pP* residuals was significantly reduced as well (to $+/- 1.6$ sec), suggesting that focal depths for relocated hypocenters are a significant improvement over depths determined by the ISC. Note that throughout the remainder of this article, estimates of center and scatter will be quoted as median $+/-$ spread. Spread is a robust analog of standard deviation and is defined as the median of the absolute deviations of the observations from their median normalized to yield the usual standard deviation when applied to Gaussian distributed data (i.e., multiplied by 1.48258).

The *pwP* residuals (Fig. 6c) also fit reasonably well, indicating that the phase identification algorithm is operating effectively. However, near regions of high residual density in Figure 6c there appear to be significant numbers of late arrivals, which we suspect are higher order multiples of the *pwP* phase (e.g., *pwwP*) that have not been identified in our relocation procedure. The distribution of ISC *sP* residuals (Fig. 6d) demonstrates that, in addition to the negative bias identified from ISC *pP* residuals plotted in Figure 6a, there must be an even larger positive bias ($1.1 +/- 3.1$ sec) introduced along the *S* part of the path in the upper mantle. We attribute this observation to a baseline problem in the JB *S* travel times. The high spread ($+/-3.1$ sec) in Figure 6d also indicates some difficulty on the part of ISC in identifying the *sP* phase. Figure 6e shows a very good fit ($0.2 +/- 1.8$ sec) of *sP* residual data using EHB procedures. Late *sP* arrivals near regions of high residual density are probably unidentified *swP* phases.

In Figures 7a and 7b, we compare ISC and EHB residuals for all first-arriving *S*-wave types (*Sg, Sb, Sn, S,* and *Sdiff*) from 0 to 100° distance. The gap in the residuals near 83° is due to the crossing phase *SKS* where phase identification was not attempted. There are fewer data beyond the intersection of *SKS* with *S* because of ISC procedures and because the two phase arrivals are so difficult to distinguish from one another where they occur close in time. Also, residuals for regional *S* arrivals (<25°) are not reported by the ISC. Figure 7 clearly demonstrates that ak135 properly defines the *S*-wave baseline and significantly reduces the spread at teleseismic distances, whereas the JB model provides a poor fit (see also Fig. 6d).

At regional distances, it appears that the distribution of events globally would tend to favor faster *S*-wave velocities than predicted by the ak135 model. We believe this is the result of the many paths from oceanic events to continental stations (e.g., from Tonga and the Marianas). Ak135 was designed to be representative of continental paths, because most seismic stations lie on continents. A model developed

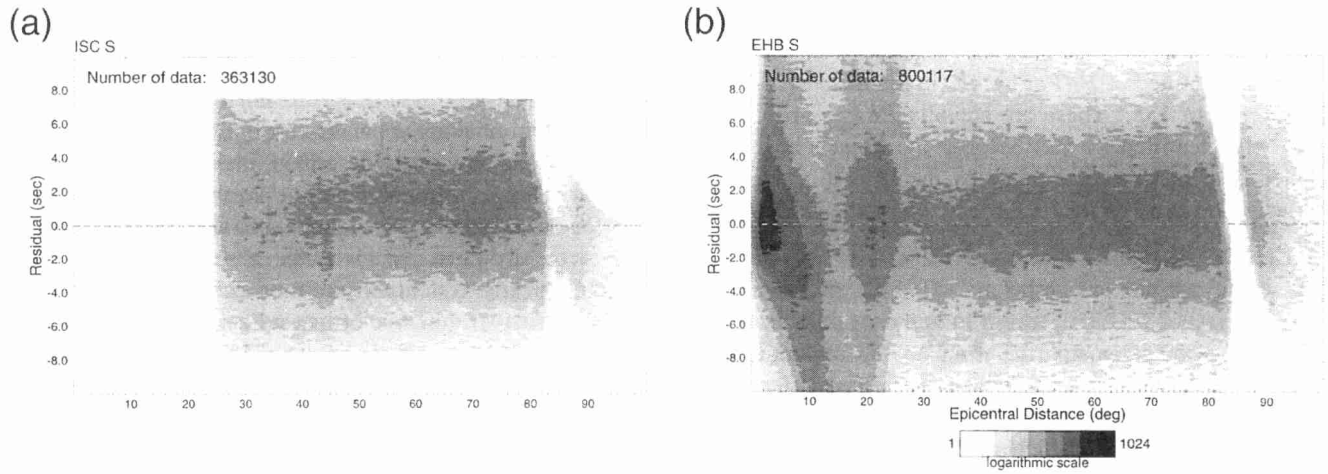


Figure 7. Residual density plots for (a) *S* residuals reported by the ISC and (b) *S* residuals determined in this study.

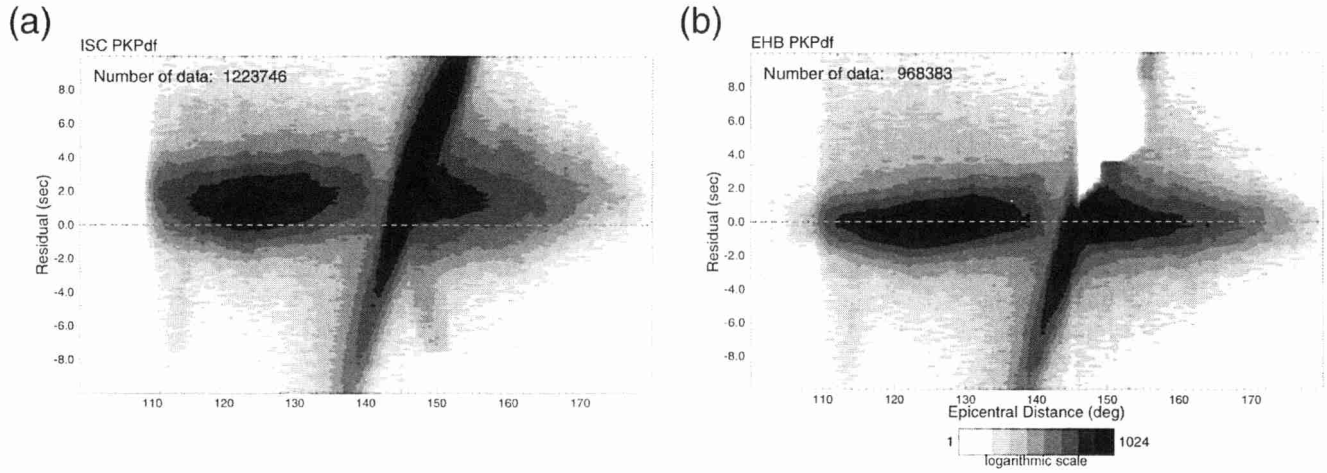


Figure 8. Residual density plots for (a) *PKPdf* residuals reported by the ISC and (b) *PKiKP* and *PKPdf* residuals determined in this study.

for purely oceanic regions (Kennett, 1992) could provide a significantly better fit to the data. However, for marginal zones and island arcs, the situation is less clear, and it is probably best to employ the ak135 model. Despite the high spread of EHB-identified *S* arrivals at regional distances, we found that because of the strong constraint provided on origin time and focal depth, it was still valuable to use these data in the event relocation.

Finally, in Figures 8a and 8b, we compare residual density plots for first-arriving *PKP* phases (*PKiKP* and *PKPdf*) over the distance range of 110 to 180°. In both plots, the effects of diffraction near the *PKP* caustic (at about 143°) and *PKP* precursors at shorter distances are evident. For the *PKP* data reported by ISC, there appears to be an offset in *PKPdf* times of about +1.7 sec, indicating that there is probably a baseline problem in JB *PKP* travel times. However, because the ISC is overestimating the focal depth of deeper earthquakes by at least 10 km on average (consistent with an ISC *pP* offset of −0.7 sec), at least part of the *PKPdf* offset (which would be opposite in sign) must be due to this effect as well. We also note the occurrence of many late arrivals identified as *PKPdf* by ISC in the distance range of about 145 to 155°. These are actually *PKPbc* arrivals improperly identified by ISC because the branch is not predicted over that distance range by the JB core model. Obviously, travel times for *PKP* predicted by the ak135 model do a much better job than JB (Fig. 8b).

Phase Identification

It is also informative to compare the reidentification characteristics of phases statistically identified in this study (*pP, pwP, sP, PcP*) to phase identifications made by the ISC for the same arrivals (Table 1). More than half of the phase identifications made for these phases in the EHB analysis were different than the identification made by the ISC, with about one-sixth of these being reidentified as *pwP*. For phases identified as *pP, sP,* and *PcP* by the ISC, about a third were reidentified in the EHB analysis.

Table 1

Comparison of the number of *pP, pwP, sP,* and *PcP* phases statistically identified in this study (EHB) to identifications for the same phases made by the ISC. Only absolute residuals for later phases of less than or equal to 7.5 sec (the ISC phase association window) over the distance range 30 to 100° were used in this comparison. "Other" includes phases for which no phase identification was made by either ISC or EHB.

		ISC			
	Phase	*PcP*	*pP*	*sP*	Other
EHB	*PcP*	86,443	4,255	2,578	34,674
	pP	16,887	193,732	28,588	78,591
	pwP	6,661	44,833	30,436	26,531
	sP	5,381	24,620	52,299	28,160
	Other	291	86	23	

Hypocenters

The median difference for all EHB epicenters relative to ISC epicenters (excluding shifts >40 km) is 7.6 +/− 5.5 km with no obvious dependence on focal depth (Table 2). Note that, as previously described, the spreads quoted here and in Table 2 represent the scatter of the data (i.e., they are analogous to standard deviations, not standard deviation of the mean). To investigate whether these epicentral shifts are geographically systematic, we plot in Figure 9a difference vectors (≦40 km) in the northwest Pacific for events greater than 70 km in focal depth that have 250 or more teleseismic reporting stations. Limiting the events plotted in Figure 9 to those that are well recorded was done simply to reduce the clutter in the figure. In fact, it makes little difference which events we choose to show, as long as they meet the event selection and classification standards previously described. In other words, all events in any given region that are well

Table 2

Median and spread of epicenter and depth shifts (in km) for EHB free-depth determinations relative to epicenters and depths determined by the ISC, NEIC (BBD, MT), and Harvard (CMT) for All, 0–70, 70–300, and >300 km EHB depth ranges. Excluded from these estimates were 835 ISC events with epicentral shifts >40 km, 2165 ISC events with depth shifts >40 km, 266 CMT events with epicentral shifts >100 km, and 200 CMT events with depth shifts >40 km. Most of these excluded events were poorly resolved estimates by ISC and Harvard. However, the largest shifts (of order several hundred kilometers) were associated with events improperly determined or simply grossly in error.

	EHB Depth Ranges			
	All	0–70	70–300	>300
Epicenter Shifts				
ISC	7.6 +/− 5.5	7.4 +/− 5.2	7.5 +/− 5.5	9.3 +/− 5.4
CMT	32.7 +/− 21.3	33.7 +/− 21.2	30.1 +/− 20.4	31.4 +/− 20.6
Depth Shifts				
ISC	−1.7 +/− 11.8	−0.3 +/− 13.1	−4.3 +/− 8.8	−3.1 +/− 7.9
BBD	1.2 +/− 5.3	−0.1 +/− 4.7	4.2 +/− 5.0	5.1 +/− 4.6
MT	1.6 +/− 11.4	1.4 +/− 11.0	2.6 +/− 12.1	1.9 +/− 9.0
CMT	1.2 +/− 13.1	2.6 +/− 12.6	−1.7 +/− 13.0	−5.5 +/− 10.7

recorded teleseismically appear to be uniformly located (or mislocated) by the EHB and ISC algorithms, regardless of size.

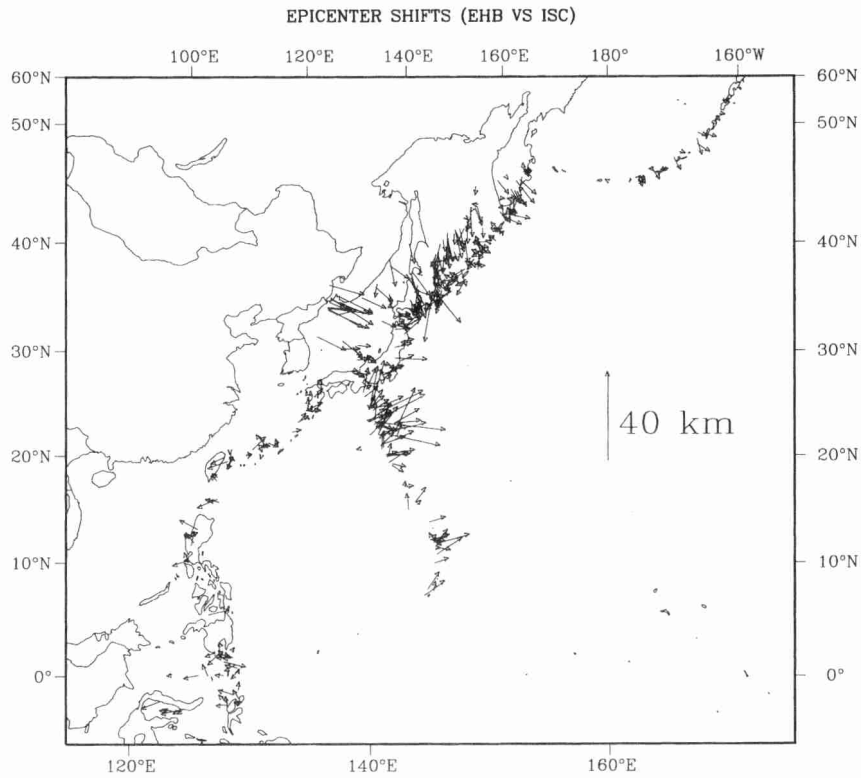
Van der Hilst and Engdahl (1992) have shown that relocations of ISC intermediate- and deep-focus events (focal depth exceeding 70 km) in northwest Pacific subduction zones yield relocation vectors with respect to ISC with lengths of order 10 km that are systematically displaced perpendicular to the strike of the seismic zone and toward the trench. The relocation thus causes the deepest part of these seismic zones to be steeper than inferred from ISC hypocenters. For shallower earthquakes seaward of the trench, they found that the epicenter relocation vectors are smaller and less systematic. These results are generally confirmed by the relocation vectors shown in Figure 9a. Relocation vectors at all depths for subduction zones in the southwest Pacific and Latin America, respectively, are plotted in Figures 9b and 9c. Though systematic shifts are evident in both of these areas, there are regions that do not conform to the northwest Pacific patterns. In particular, relocation vectors in the Vanuatu Islands and the Solomon Islands regions are actually coherently displaced away from the trench rather than toward it. In Latin America, the relocation vectors are generally coherent over smaller regions but with no clear relationship to slab geometry.

Table 2 and Figure 9d summarize the depth differences of EHB free-depth determinations relative to ISC depths. The median shift for all EHB depths relative to ISC depths (excluding depth differences >40 km) is −1.7 +/− 11.8 km. For events with depths less than 70 km, the median depth difference is small, but for larger depths, the differences are 3 to 4 km and negative. This slight overestimation of the focal depth by the ISC was also noted in the discussion of *pP* and *PKP* residual density plots. Major contributing factors to this difference are the combined effect of the ISC not using depth phases in its procedure for hypocenter determination and the negative residuals of most teleseismic *P* waves originating in high-velocity subducted slabs (van der Hilst and Engdahl, 1992).

Comparisons of EHB free-depth determinations can also be made with other independent sources. The NEIC routinely interprets broadband data from digital seismograph networks for events with m_b greater than about 5.5 (if the signal-to-noise ratio is adequate). Records that are flat to displacement between approximately 0.01 and 5.0 Hz are constructed from the digital data using methods described by Harvey and Choy (1982). If depth phases are clearly identifiable on these broadband seismograms, the differential travel times *pP-P* and *sP-P* are read using methods described by Choy and Engdahl (1987). Estimates of the focal depth are obtained by inversion of the differential times observed at several stations using the iasp91 model. The NEIC also reports depths (MT) associated with moment-tensor inversions based on long-period, vertical-component, *P* waveforms obtained from digitally recorded stations (Sipkin, 1982, 1986a,b). The inversion procedure, which until recently used the earth model 1066B, is insensitive to small errors in both epicenter and origin time. The source depth that gives the smallest normalized mean-squared error is a by-product of this algorithm.

Finally, hypocenters are estimated as part of Harvard's program for the determination of centroid moment tensors (CMT). These solutions have been determined with long-period body and mantle waveform data (low-pass filtered) using the moment-tensor inversion method described by Dziewonski *et al.* (1981), including corrections due to an aspherical earth structure of model SH8/U4L8 (Dziewonski and Woodward, 1991). Hypocentral parameters are obtained by adding perturbations resulting from the inversion to parameters reported by the NEIC. If the depth is not perturbed during the inversion, it is fixed to be consistent with the waveform matching of reconstructed broadband body waves (Ekström, 1989). The default depth is 15 km (10 km from 1981 to 1985). Table 2 indicates that the median difference

(a)



(b)

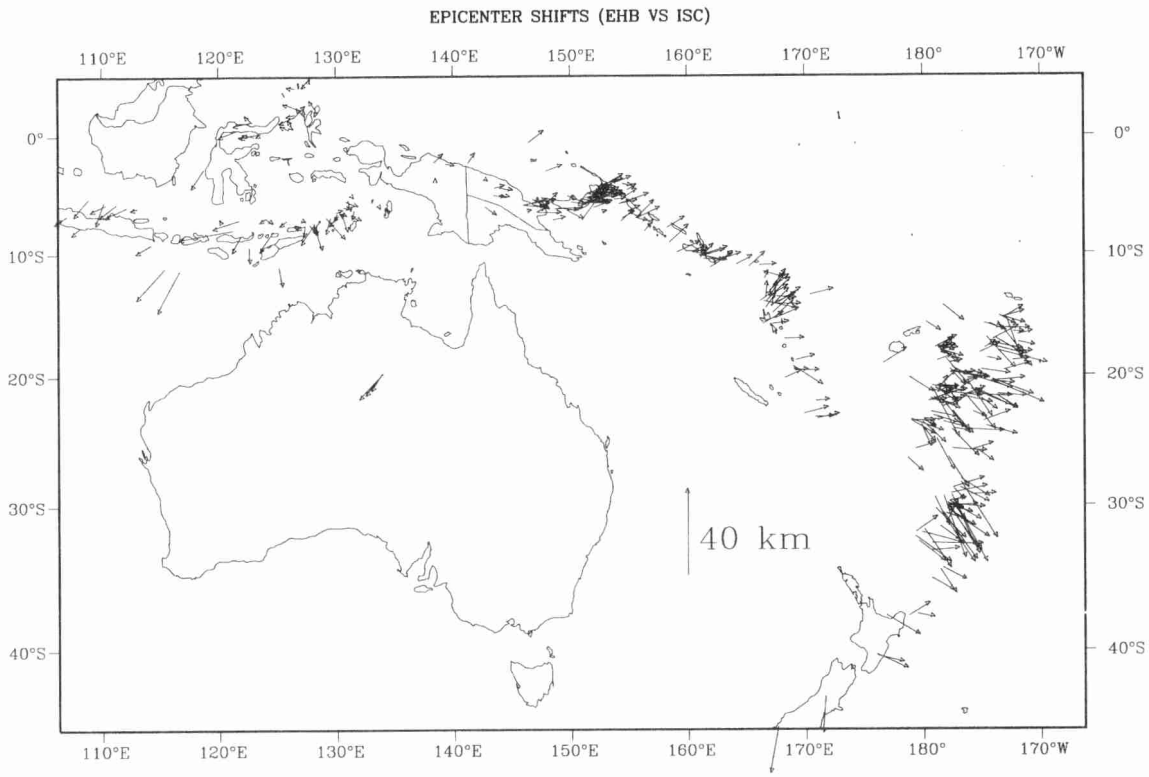
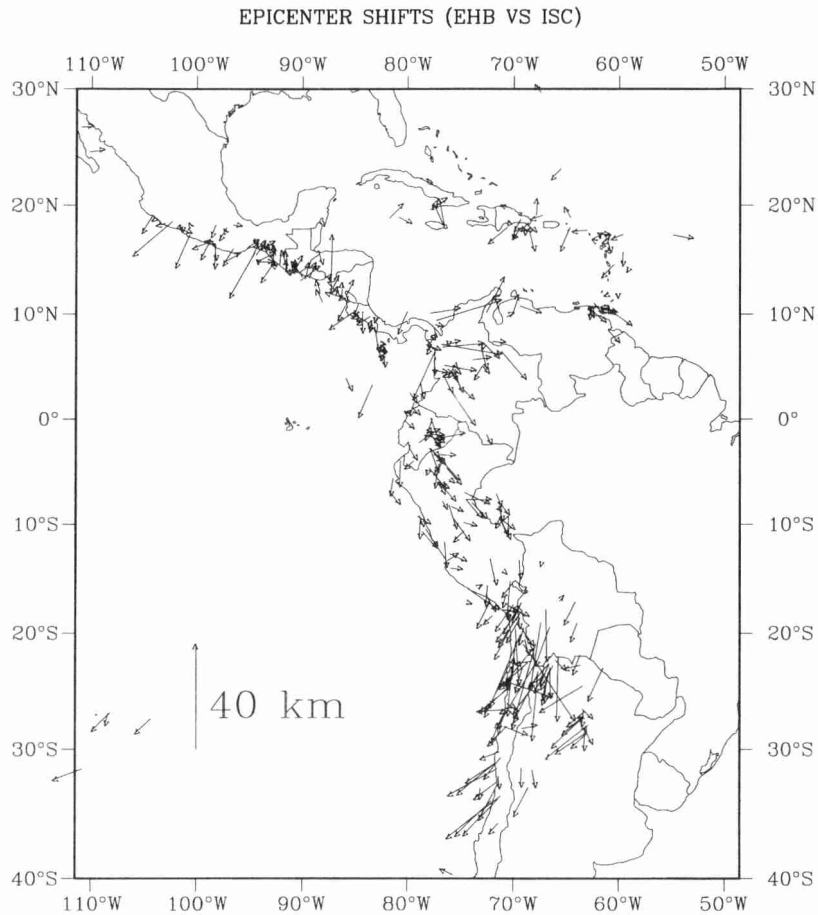


Figure 9. EHB epicenter relocation vectors (free-depth events only) relative to ISC epicenters for (a) northwest Pacific (depths >70 km only and >250 teleseismic stations), and (b) southwest Pacific (all depths and >300 teleseismic stations).

(c)



(d)

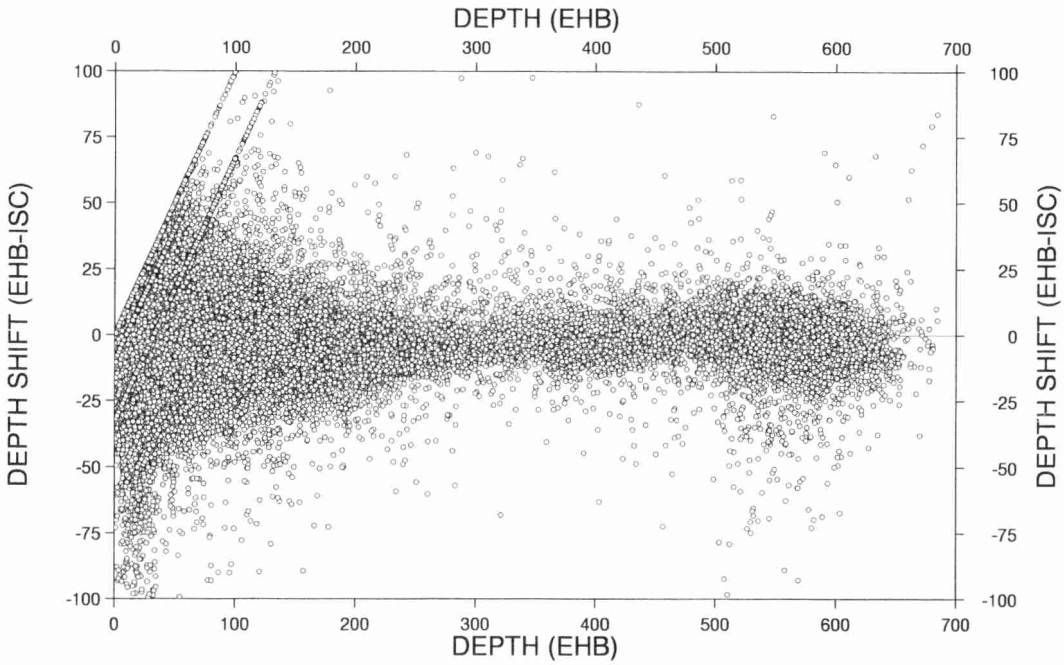


Figure 9. (*continued*) (c) South America (all depths and >250 teleseismic stations). (d) EHB depth shifts relative to ISC hypocenters. The three diagonal lines on the upper left-hand side of (d) represent fixed depths of 10, 15, and 33 km commonly set by the ISC when the depth is indeterminate based on first-arriving *P* waves alone.

for all EHB epicenters relative to CMT centroid epicenters (excluding epicentral differences >100 km) is on the order of 30 km or more, regardless of the depth range. Moreover, the CMT epicentral shift vectors relative to EHB are, for the most part, not geographically systematic, indicating a lack of epicenter resolution in the CMT centroid determinations.

Table 2 and Figure 10 summarize the depth differences of EHB free-depth determinations relative to depths independently determined by NEIC (BBD and MT) and Harvard (CMT). Except for a few outliers, EHB depths agree quite well with all other estimates over the 0- to 70-km-depth range, suggesting that for most of these events the depth phases are being properly identified both directly from long-period (MT and CMT) and broadband seismograms (BBD) and indirectly from reported arrival times (EHB) using the new statistical phase identification algorithm. For EHB depths >70 km, however, there remains a small positive difference of 4 to 5 km in EHB depths relative to BBD depths. One possible cause of this difference is that the theoretical travel times of *pwP* phases (which are based on water depth alone) are in most cases being underestimated. Because EHB depths are determined using *pwP* as well as *pP* and *sP* arrival times, this will result in a depth bias. We have examined a number of *pwP–pP* times reported for the same event by the same station and found that this may be one possible explanation. The underestimate of the *pwP* travel time may be explained by a thin layer of low-velocity sediments on the sea bottom (W. D. Mooney, personal comm.). Another possible cause of this discrepancy is the use of differential times of depth-phase arrivals (BBD) versus the absolute times of many different phases (EHB). In the former case, the effect on focal depth of lateral heterogeneity along the ray paths should be small. Finally, there are slight differences between *sP* times for iasp91 and ak135 that are of the right order (about 0.3 sec for a depth of 550) and direction to account for 1 to 2 km of the depth difference for deeper sources.

The median difference of EHB depths relative to MT and CMT depths (Figs. 10b and 10c) has a spread of 9 to 13 km over all depth ranges, about the level to be expected for the depth resolution possible using long-period waveforms. There exists, however, a consistent difference in EHB–CMT depth shifts ($-5.5 +/- 10.7$ km) for EHB depths greater than about 300 km that is also evident in comparisons of CMT to BBD and MT depths (Table 2 and Fig. 10c). A likely

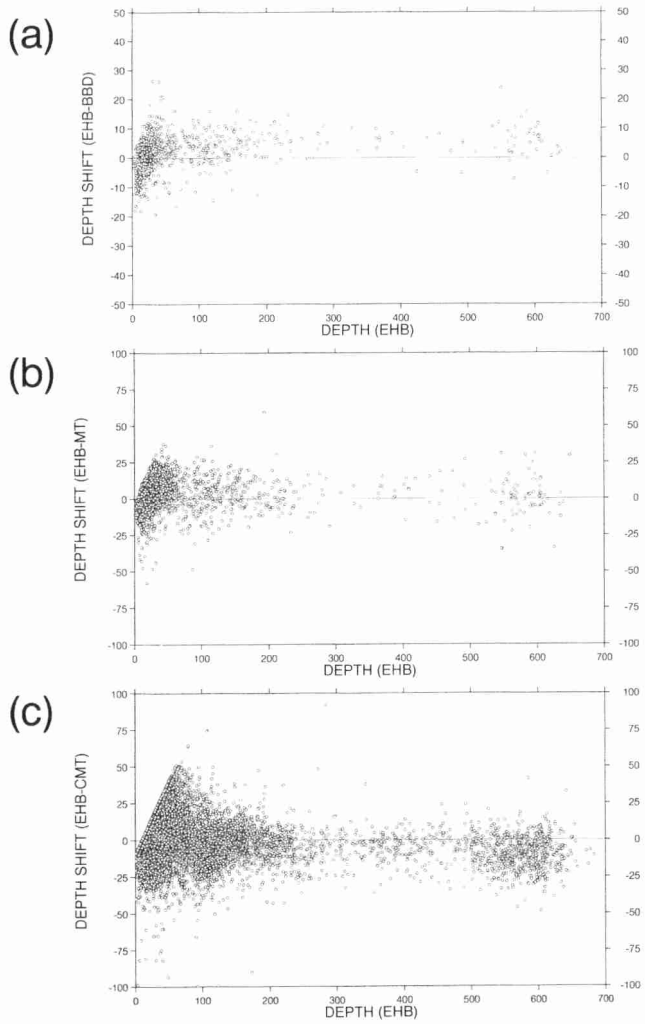


Figure 10. EHB depth shifts (free-depth events only) relative to (a) depths determined by the USGS using the differential times of depth phases picked from displacement and velocity seismograms constructed from broadband digital data (BBD). For suboceanic earthquakes, BBD depths have been corrected to sea level because observable broadband depth phases are ordinarily reflections off the sea bottom, and EHB depths are referenced to sea level; (b) depths associated with USGS moment-tensor determinations based on long-period waveforms constructed from digitally recorded data (MT); and (c) centroid moment-tensor depth determinations using low-passed seismograms constructed from digital data (CMT). The diagonal line on the upper left-hand side of this plot represents fixed depth solutions of 15 km commonly set by Harvard when the CMT depth is indeterminate.

cause for this difference is the aspherical structure used in estimating the centroid depth (Gideon Smith, personal comm.). We have, however, examined depths resulting from the simultaneous inversion of a data set of hypocenters and travel-time delays (created as part of this study) for 3D mantle structure, new hypocenters, and station corrections (Bijwaard *et al.,* 1997). Only small differences between the new depths and the starting depths were found, which generally reduce the scatter in EHB–CMT depth differences but do not change the median difference. Three-dimensional raytracing in the Harvard models has also been performed, but hardly any 3D effect on travel times was found (Wim Spakman, personal comm.). Finally, we have examined the differences in epicenter and depth between CMT centroid determinations and the EHB results as a function of moment magnitude. The differences are progressively larger with decreasing magnitude, suggesting that the largest of these differences are probably related to uncertainties other than source dimension. For the subset of earthquakes deeper than 300 km, the difference between EHB and CMT depths noted earlier appears to have no dependence on moment magnitude. Thus, the source of this difference remains enigmatic.

Example

The higher resolution of Wadati–Benioff Zone (WBZ) structure provided by free-depth EHB hypocenters is dramatically shown in Figures 11a to 11c. Subduction of the Nazca plate beneath western South America is characterized by alternating regions of normal and flat subduction and is also unusual for the presence of broad sectors of reverse curvature of the Peru–Chile trench (concave seaward). One of these sectors is located at 15 to 23° S in the Arica bend region and is bordered by two sectors of flat subduction in Peru and northern Chile. The cross sections of these three regions shown in Figure 11 compare hypocenters reported by the ISC with EHB hypocenters determined using five or more depth phases for the same events. In each case, the EHB hypocenters result in a more sharply defined WBZ, suggesting that the greater dispersion in ISC hypocenters is not a real feature. The narrowness of the WBZ is all the more remarkable in that these cross sections have been constructed for arc segments as much as 1200 km long, indicating the along-arc uniformity of geometry as well as a less diffuse definition of the WBZ itself due to reduced uncertainties in the hypocenters (cf. Kirby *et al.,* 1995, 1996).

To investigate whether the reduced dispersion in depth seen in Figure 11 also results in increased clustering of EHB relocations of ISC epicenters in map view, we compare in Figures 12a and 12b the epicentral distributions for the same events plotted in cross section in Figure 11b. There is no apparent increased clustering of epicenters as a result of EHB processing. However, there are large epicentral shifts with a general southwesterly tendency that are, over smaller regions, sometimes coherently oriented (Fig. 12c). This may be the result of the preponderance of North American stations ordinarily used to locate South American events and the distribution of patch corrections, which are generally opposite in sign in eastern and western North America (see Fig. 2).

Discussion

The systematic differences between hypocentral parameters in the ISC catalog and our new EHB data base are caused by a number of factors: (1) differences in earth models (JB versus ak135), particularly for the upper mantle; (2) procedural differences between ISC and EHB processing, such as our use of later-arriving phases; and (3) our use of station "patch" corrections. In combination with uneven station distribution, all these factors can produce hypocentral shifts. Have our (EHB) procedures generally improved hypocenter determinations for events of interest? To answer this question, let us examine some of the evidence that has been presented.

Reference Earth Models and Travel-Time Tables

It is apparent that JB travel times, which are used by both the ISC and NEIC for routine hypocenter determination, are for several reasons not optimal for that purpose. First, there is the well-known baseline error in JB *P*-wave travel times (cf. Kennett and Engdahl, 1991) as well as systematic errors in the JB lower mantle (Fig. 5a). This results in errors in origin time and mislocation that are dependent on the distribution of teleseismic stations reporting each event and how they sample the lower mantle. Second, there appears to be a significant error (on the order of at least 0.5 sec) in the baseline for JB *S*-wave teleseismic travel times (Fig. 7a). Third, the baseline for JB *PKP* travel times is also in error (on the order of at least 1.0 sec, see Fig. 8a). These biases in JB *P, S,* and *PKP* travel times can, if routinely used, result in systematic hypocenter errors. Ak135 seems to fall into a class of models that provide a better all-around fit to the global travel-time data set.

There are also differences in regional travel times predicted by JB and ak135 that result from assumptions made in deriving the two models. These differences are manifested in the epicenter shifts shown in Figure 9 that are partly model dependent. At regional distances, the ak135 model does appear to provide a better fit to *P*-wave travel times (Fig. 5). However, *S*-wave travel times predicted by the ak135 model at regional distances appear to be too slow on average (Fig. 7b), suggesting possible modifications needed in average upper mantle *S*-wave velocities to account for the contribution from faster pure oceanic paths. Thus, we conclude that the use of an improved reference earth model, by removing systematic errors in the model used, can indeed improve hypocenter determinations globally.

Use of Later-Arriving Phases

The use of later-arriving phases in routine hypocenter determination potentially provides powerful constraints on

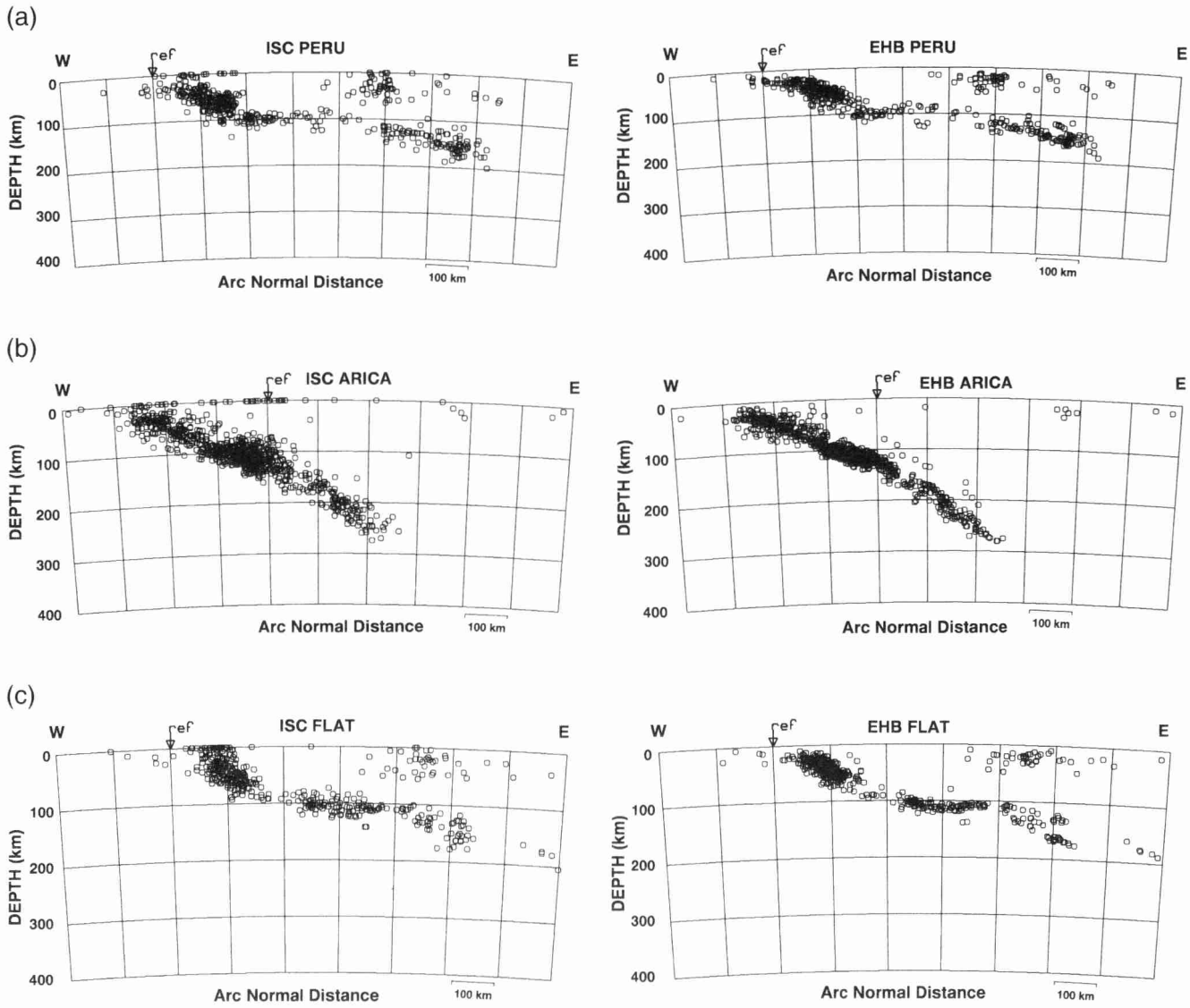


Figure 11. Cross sections of the South American subduction zone using curvilinear projections (with origin at the center of curvature of the volcanic arc or trench) that compare hypocenters reported by the ISC (left) with free-depth EHB events that have been determined using five or more depth phases (right). Cross sections are for (a) Peru (PERU), (b) Arica bend (ARICA), and (c) northern Chile (FLAT). Arc normal distance (to center of curvature) and depth are gridded in 100-km blocks. Arrow (marked "ref") points to the trench or the volcanic arc.

focal depth and reduces the effects of strong near-source lateral heterogeneities. However, both ISC and NEIC rely almost entirely on first-arriving *P* waves to locate earthquakes. Thus, without further processing, residuals for high-frequency seismic phases other than *P* that are reported by these agencies have only limited application to research problems such as source location and the evaluation of earth models. Our new phase identification algorithm, applied to the phase group immediately following the *P* wave at teleseismic distances (*pP*, *pwP*, *sP*, and *PcP*), which then allows the unambiguous identification of later phases such as *S*, has been shown to be particularly effective so that these phases can be confidently exploited in the relocation procedure. In some subduction zones (such as the northwest Pacific), the application of the relocation scheme results in shifts in earthquake epicenter that are systematic and can largely be explained by the effects of slab geometry. However, in the case of very steeply dipping slabs (Vanuatu Islands) and slabs with more complex geometries (South America), the systematics are not so obvious, as slab effects on teleseismic *P*-wave travel times are either reduced or extremely complex.

Despite the general success of the procedures described here, there remain some issues that require further investigation. For example, the relative frequency (or amplitude)

of depth-phase observations is sensitive to local structure at bounce points. Many depth phases reflect in the vicinity of plate boundaries where the slopes of surface reflectors are large (>1°). Reflections at a dipping reflection zone may lead to small asymmetries in depth-phase waveforms (Wiens, 1987, 1989) but, more importantly, may also influence their relative amplitudes and result in a greater potential for phase misidentifications. In addition, for short-period (1 sec) waves, water–sediment interfaces at the sea bottom may have small impedance contrasts. Consequently, on short-period (WWSSN) seismograms, *pwP* may have an amplitude comparable to, or larger than, the *pP* phase reflecting at the sea bottom, and *pwP* may easily be misidentified as *pP*.

Although erroneously identified as other phase types by

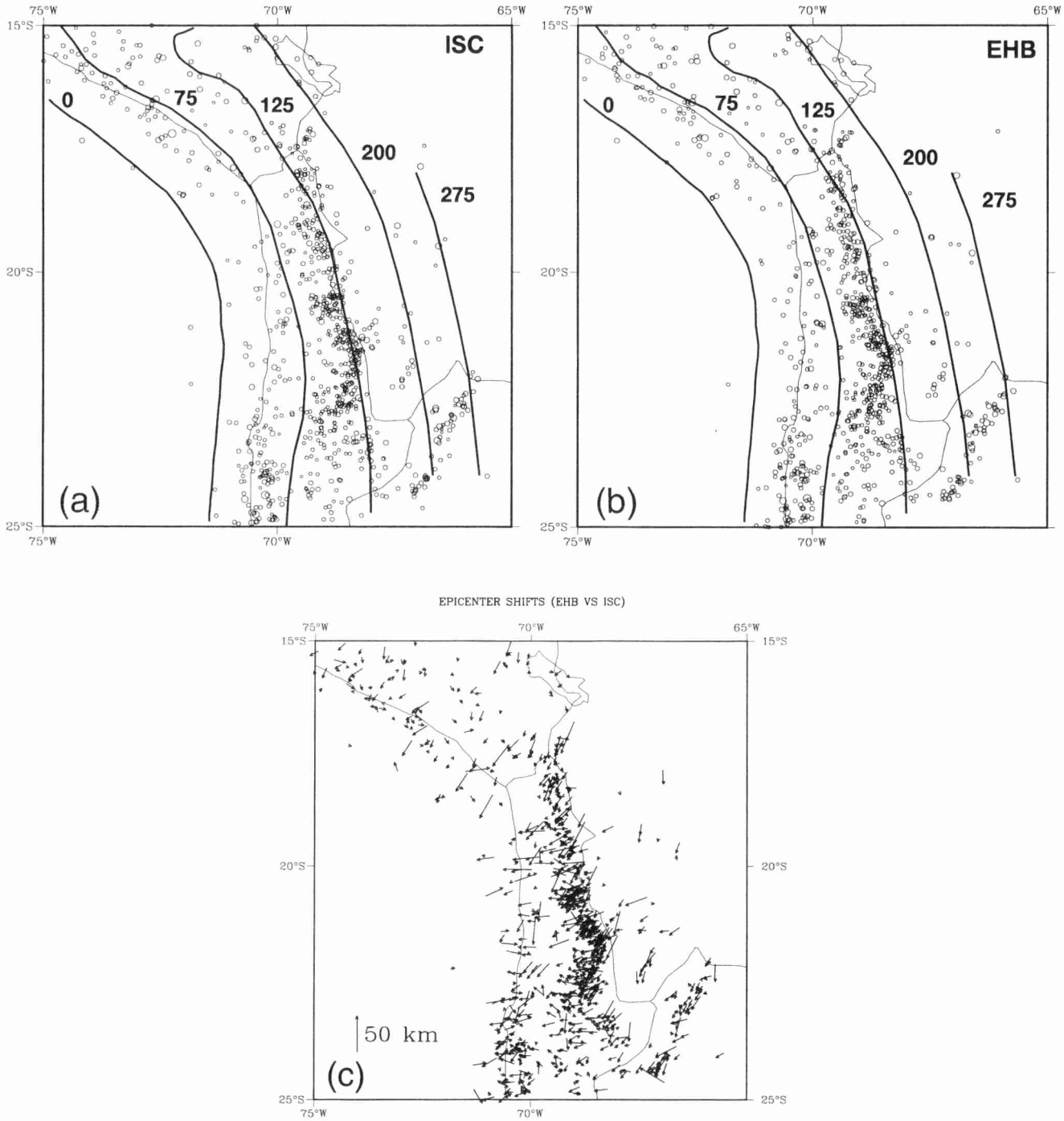


Figure 12. Epicenters for the same set of events plotted in Figure 11b: (a) ISC, (b) EHB, and (c) vectors showing EHB epicenter shifts relative to ISC locations. Tails of the vectors are the ISC epicenters. Wadati–Benioff Zone contours are from Kirby *et al.* (1995).

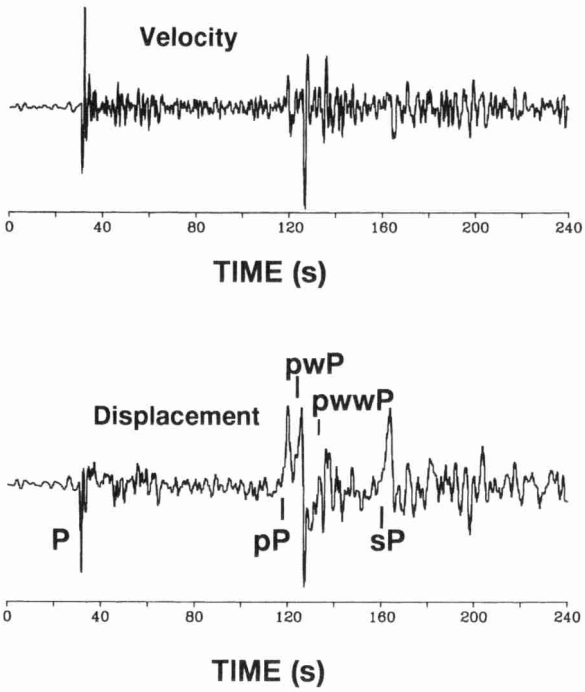


Figure 13. Broadband displacement and velocity seismograms recorded at the broadband station OBN (73°) from a deep earthquake (395 km depth) south of Honshu, Japan. Onsets of observed phases are labeled. The water depth at the bounce point for the *pP* and *pwP* phases observed at this station is about 4.4 km. Horizontal scale is time in seconds and vertical scale is normalized amplitude.

station operators as well as by the ISC, identifications of the *pwP* phase are pervasive in our data base. The *pwP* phase is sometimes also observed in longer-period records. Figure 13 shows a teleseism recorded by a broadband station in Russia from a suboceanic earthquake near Japan. This is one of the rare observations of a *pwP* arrival with significant amplitude on a broadband displacement seismogram, and it was widely reported as *pP*. There also remain effects of the water layer that have not been accounted for, such as the possibility of *swP* and *pwwP* phases being routinely observed, especially in the case of an enhanced *pwP* signal. For the seismograms shown in Figure 13, it is quite easy to read the onset times of *pP*, *pwP*, and *sP*, as well as the multiple *pwwP*.

Another outstanding problem is that for large shallow-focus complex earthquakes, *pP* often arrives in the source-time function of *P* that may consist of one or more subevents. The gross features of the source-time functions of *P* and *pP*, however, remain discernible in broadband displacement records, and the exact onset times of depth phases can be further refined by examination of velocity seismograms that are sensitive to small changes in displacement (Choy and Engdahl, 1987). In this study, however, we have relied primarily on reported data, usually read from short-period seismograms. Hence, larger complex events always require critical review for the possibility of phase misidentification by our statistical algorithm.

Effects of Aspherical Earth Structure

The travel times predicted by the earth model ak135 are extremely valuable for earthquake location and phase identification using a radially symmetric model. Nevertheless, most deeper-than-normal earthquakes occur in or near subducted lithosphere where aspherical variations in seismic wave velocities are large (i.e., on the order of 8 to 10%; Engdahl *et al.*, 1977). Such lateral variations in seismic velocity, the uneven spatial distribution of seismological stations, and the specific choice of seismic data used to determine the earthquake hypocenter can easily combine to produce bias in earthquake locations of several tens of kilometers (Engdahl *et al.*, 1977, 1982; Engdahl and Gubbins, 1987; Dziewonski and Anderson, 1983; Adams, 1992; van der Hilst and Engdahl, 1992).

Kennett and Engdahl (1991) have shown that a set of "test events" (events for which we have well-constrained hypocenters, such as nuclear explosions or earthquakes located within a local network) are on average mislocated by about 14 km using standard procedures. The set of test events used by Kennett and Engdahl was enhanced by additional nuclear explosions globally (there are 1166 explosions in our data base). The enhanced test-event data set was relocated using the procedures described in this article, and the mislocation vectors are plotted in Figure 14. The mean length of the mislocation vectors is not a strong function of the 1D model used to locate these events, as Kennett and Engdahl found almost the same epicenter mislocation using the JB model. The influence of the higher-velocity slab for events in most subduction zones still shows up clearly, because most mislocation vectors in these regions point in the direction of subduction. Our average test-event mislocation of 9.4 +/− 5.7 km is only slightly larger than the best rms misfit to known locations (7.2 km) recently found by using a 3D Harvard large-scale mantle model (S&P12/WM13) to relocate the explosions in the Kennett and Engdahl data set (Smith and Ekström, 1996). Obviously, this quantitative comparison between rms mislocations achieved using nearly the same data set but two different location procedures may have limited meaning. However, it is clear that where slabs with strong small-scale heterogeneities occur, modeling only large-scale heterogeneity (Smith and Ekström, 1996, 1997) cannot significantly improve teleseismic event location.

We have tried to reduce the effects of aspherical structure, but until we can account for lateral variations in velocity in the earthquake location procedure, there is not much hope of significantly reducing the remaining bias. Nevertheless, the use of numerous stations and seismic waves at different azimuths and distances around the source in the relocation procedure does seem to reduce mislocation errors introduced by lateral heterogeneity (e.g., Fig. 11) and does enhance the structural signal to the extent that higher-resolution imaging of 3D global *P*-wave structure obtained by

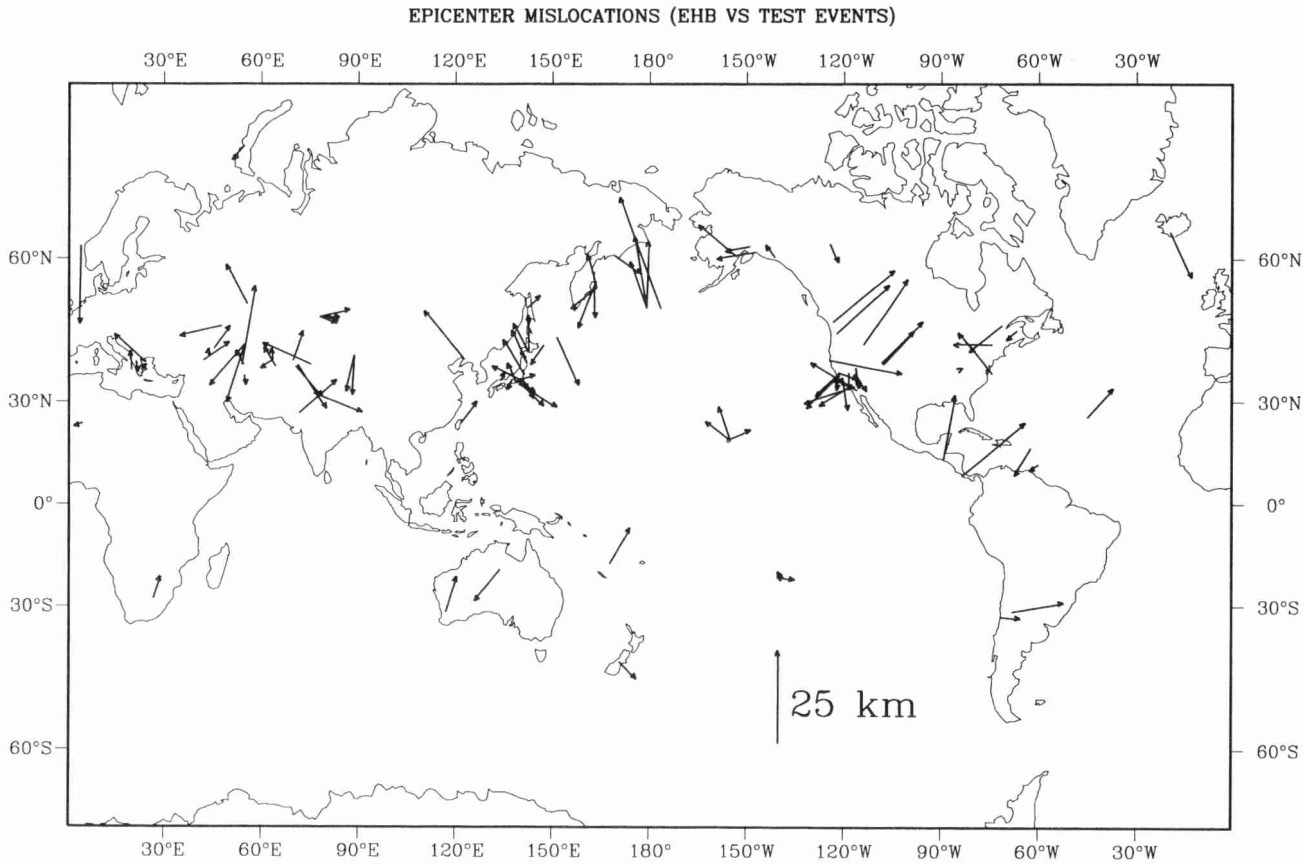


Figure 14. Mislocation vectors for test events using EHB procedures (at fixed depth). Vectors show EHB mislocation. The tail of the vector is the actual test-event location.

tomographic inversions of EHB hypocenters and phase data (van der Hilst *et al.*, 1997) has been made possible.

The use of *S* waves, particularly at regional distances, provides not only additional strong constraints on focal depth but a surprisingly valuable *S* residual data set for the tomographic imaging of *S*-wave velocities in the lower mantle (e.g., Vasco *et al.*, 1994, 1995; Kennett *et al.*, 1998). Similar improvements to core velocities may also be possible from the inversion of *PKP* residual data resulting from the EHB processing.

Comparison with Independent Hypocenter Estimates

Comparison of EHB depths with ISC depths on the one hand and with (presumably superior) depth estimates from waveform modeling on the other suggest that about half of the difference between ISC and waveform depths is recovered by the EHB processing. However, the comparison also presents some interesting unanswered questions. In particular, comparisons of EHB hypocenters to centroid moment-tensor inversions (CMT) performed by the Harvard group warrant some comment. The CMT data set provides an extremely valuable resource of source mechanism data but apparently lacks the hypocenter (centroid) resolution at lower moment magnitudes to make meaningful comparisons with our data. We find differences of EHB hypocenters to CMT centroids to be not only large but nonsystematic. Epicenter and depth shifts greater than about 25 km increase with decreasing scalar moment, suggesting that the larger discrepancies are probably a CMT resolution problem (see also Smith and Ekström, 1997). The observed difference between EHB and CMT depths for deeper-than-normal earthquakes is more difficult to explain but is probably related to the differences between EHB and CMT processing techniques. However, there is an obvious danger in making these comparisons for the largest events, as we are locating the nucleation point with high-frequency first arrivals whereas CMT solutions represent the centroid of fault slip. Some preliminary work using EHB hypocenters with CMT source mechanism data reveals fine-scale structure in the orientation of stress axes regionally that was not apparent using centroid locations.

Conclusions

Nearly 100,000 events that occurred during the period 1964 to 1995 and that are well-constrained teleseismically

by arrival-time data reported to the International Seismological Centre (ISC) and to the U.S. Geological Survey's National Earthquake Information Center (NEIC) have been relocated. Hypocenter determination is significantly improved by using, in addition to regional and teleseismic *P* and *S* phases, the arrival times of *PKiKP* and *PKPdf*, and the depth phases *pP*, *pwP*, and *sP* in the relocation procedure. The use of a global probability model developed for later-arriving phases significantly improves the identification of depth phases and the reliability of depth estimates.

A comparison of EHB relocations to hypocenters reported in the ISC and NEIC catalogs reveals systematic epicenter and depth differences. These differences appear to be regionally dependent on the combined effects of the observing station network, the plate geometry, the use of later phases by EHB, and the differences in upper mantle travel times between the reference model ak135 and the Jeffreys–Bullen model used by these agencies. Focal depths, as revealed by a reduction of scatter in Wadati–Benioff Zone seismicity, are significantly improved, demonstrating how regional structures such as downgoing slabs can severely bias depth estimation when only regional and teleseismic *P* arrivals are used to routinely locate earthquakes. Differences in EHB depths with respect to depths determined independently from the processing of waveform data are small compared to the inherent uncertainty in EHB depth estimates ($+/-10$ to 15 km, based on the variances of depth phases).

The resulting EHB data base, which is complete to about *Mw* 5.2 and includes all events for which moment-tensor solutions are available, has immediate application to high-resolution definition of Wadati–Benioff Zones worldwide, regional (subduction zones) and global tomographic imaging, and other studies of deep earth structure. Most importantly, the depth distribution of global seismicity can be accurately portrayed and, in particular, the depth distributions of intermediate-depth earthquakes within major subduction zones worldwide can be compared. The use of a proper reference model and unbiased starting hypocenters is especially important for tomographic imaging. An inadequate reference model and mislocated hypocenters can often result in loss of structural signal in the data and mapping of nonlinear effects into the tomographic images.

We anticipate the introduction of many of the features of EHB processing into the routine determination of hypocenters by NEIC. Ak135 or a closely related model should be adopted as the 1D reference model. The statistical phase identification algorithm should be extended to all later-arriving phases. This will permit selected later phases, including depth phases, *S* arrivals, and the branches of *PKP*, to be confidently included with first-arriving *P* waves in all hypocenter determinations. A smoothed version of global topography and bathymetry should be used to determine bouncepoint corrections for surface-reflected phases and for the computation of theoretical *pwP* arrival times. Patch corrections should be replaced or augmented with individual station statistics that should permit weighting by observable station characteristics such as noise level, typical reading errors, etc. The ultimate objectives will be to introduce a realistic 3D earth model, with high resolution of subduction zone structures, and to use fast nonlinear raytracing techniques now under development for routine earthquake location.

Acknowledgments

We thank E. A. Bergman, J. W. Dewey, and S. H. Kirby for critical reviews; W. Spakman and H. Bijwaard for helpful discussions; G. Choy for construction of Figure 13; and two anonymous reviewers who addressed a number of important points and generally helped to improve the article. E. R. Engdahl received partial financial support for this study from the NWO/Pioneer Project PGS 76-144 "Detailed Structure and Dynamics of the Upper Mantle." A compressed hypocenter data file (EHB.HDF) with worldwide coverage for the period 1964–1995 and a format description (FORMAT.HDF) can be retrieved by anonymous ftp to machine address 136.177.20.1 and then change directory (cd) to /misc/engdahl/EHB.

References

Adams, R. D. (1992). Effects of heterogeneity on earthquake location at ISC, *Phys. Earth Planet. Interiors* **75,** 1–5.

Adams, R. D., A. A. Hughes, and D. M. McGregor (1982). Analysis procedures at the International Seismological Centre, *Phys. Earth Planet. Interiors* **30,** 85–93.

Beck, S. L. *et al.* (1994). Across the Andes and along the Antiplano: a passive seismic experiment, *IRIS Newslett.* **13,** 1–3.

Bijwaard, H., W. Spakman, and E. R. Engdahl (1997). Closing the gap between regional and global travel time tomography, *J. Geophysical Res.* (submitted).

Bolt, B. A. (1960). The revision of earthquake epicenter, focal depths, and origin times using a high-speed computer, *J. R. Astr. Soc.* **3,** 433–440.

Bolt, B. A. (1968). Estimation of PKP travel times, *Bull. Seism. Soc. Am.* **58,** 1305–1324.

Boyd, T. M., E. R. Engdahl, and W. Spence (1995). Seismic cycles along the Aleutian Arc-Analysis of seismicity from 1957 through 1991, *J. Geophys. Res.* **100,** 621–644.

Buland, R. (1976). The mechanics of locating earthquakes, *Bull. Seism. Soc. Am.* **66,** 173–187.

Buland, R. (1986). Uniform reduction error analysis, *Bull. Seism. Soc. Am.* **76,** 217–230.

Buland, R. and C. Chapman (1983). The computation of seismic travel times, *Bull. Seism. Soc. Am.* **73,** 1271–1302.

Choy, G. L. and E. R. Engdahl (1987). Analysis of broadband seismograms from selected IASPEI events, *Phys. Earth Planet. Interiors* **47,** 80–92.

Davies, J. H. (1992). Lower bound estimate of average earthquake mislocation from variance of travel-time residuals, *Phys. Earth Planet. Interiors* **75,** 89–101.

Davies, J. H., O. Gudmundsson, and R. Clayton (1992). Spectra of mantle shear wave velocity structure, *Geophys. J. Int.* **108,** 865–882.

Doornbos, D. J. (1988). Asphericity and ellipticity corrections, in *Seismological Algorithms,* D. J. Doornbos (Editor), Academic Press, London, 77–85.

Dziewonski, A. M. and D. L. Anderson (1983). Travel times and station corrections for P waves at teleseismic distances, *J. Geophys. Res.* **88,** 3295–3314.

Dziewonski, A. M. and F. Gilbert (1976). The effect of small, aspherical perturbations on travel times and a re-examination of the correction for ellipticity, *J. R. Astr. Soc.* **44,** 7–17.

Dziewonski, A. M. and R. L. Woodward (1991). Acoustic imaging at the

planetary scale, in *Acoustical Imaging* E. Ermert and H.-P. Harjes (Editors), Vol. 19, Plenum, New York.

Dziewonski, A. M., T. Chou, and J. H. Woodhouse (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.* **86,** 2825–2852.

Ekström, G. (1989). A very broad band inversion method for the recovery of earthquake source parameters, *Tectonophysics* **166,** 73–100.

Engdahl, E. R. and S. Billington (1986). Focal depth determination of central Aleutian earthquakes, *Bull. Seism. Soc. Am.* **76,** 77–93.

Engdahl, E. R. and D. Gubbins (1987). Simultaneous travel-time inversion for earthquake location and subduction zone structure in the central Aleutian Islands, *J. Geophys. Res.* **92,** 13855–13862.

Engdahl, E. R. and R. H. Gunst (1966). Use of a high-speed computer for the preliminary determination of earthquake hypocenters, *Bull. Seism. Soc. Am.* **56,** 325–336.

Engdahl, E. R., N. H. Sleep, and M.-T. Lin (1977). Plate effects in North Pacific subduction zones, *Tectonophysics* **37,** 95–116.

Engdahl, E. R., J. W. Dewey, and K. Fujita (1982). Earthquake location in island arcs, *Phys. Earth Planet. Interiors* **30,** 145–156.

Harvey, D. and G. L. Choy (1982). Broadband deconvolution of GDSN data, *Geophys. J. Int.* **69,** 659–668.

Jeffreys, H. (1932). An alternative to the rejection of observations, *Proc. Roy. Soc. London A* **137,** 78–87.

Jeffreys, H. (1939). *Theory of Probability,* Oxford Univ. Press, Oxford.

Jeffreys, H. and K. E. Bullen (1940). *Seismological Tables,* British Association for the Advancement of Science, London.

Kennett, B. L. N. (1992). Locating oceanic earthquakes—the influence of regional models and location criteria, *Geophys. J. Int.* **108,** 848–854.

Kennett, B. L. N. and E. R. Engdahl (1991). Travel times for global earthquake location and phase identification, *Geophys. J. Int.* **105,** 429–465.

Kennett, B. L. N. and O. Gudmundsson (1996). Ellipticity corrections for seismic phases, *Geophys. J. Int.* **127,** 40–48.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.* **122,** 108–124.

Kennett, B. L. N., S. Widiyantoro, and R. D. van der Hilst (1998). Joint seismic tomography for bulk sound and shear wavespeed in Earth's mantle, *J. Geophys. Res.* (in press).

Kirby, S. H., E. A. Okal, and E. R. Engdahl (1995). The June 9 Bolivian deep earthquake: an exceptional event in an extraordinary subduction zone, *Geophys. Res. Lett.* **22,** 2233–2236.

Kirby, S. H., E. R. Engdahl, and R. Denlinger (1996). Intermediate-depth intraslab earthquakes and arc volcanism as physical expressions of crustal and uppermost mantle metamorphism in subducting slabs, in *Subduction: Top to Bottom,* G. Bebout, D. Scholl, S. Kirby, and J. Platt (Editors), Geophysical Monograph 96, American Geophysical Union, Washington, D.C., 195–214.

Mendiguren, J. A. (1971). Focal mechanism of a shock in the middle of the Nazca plate, *J. Geophys. Res.* **76,** 3861–3879.

Morelli, A. and A. M. Dziewonski (1993). Body wave traveltimes and a spherically symmetric P- and S-wave velocity model, *Geophys. J. Int.* **112,** 178–194.

Nolet, G. (1987). Seismic wave propagation and seismic tomography, in *Seismic Tomography,* G. Nolet, (Editor), Reidel, Dordrecht, 1–23.

Sipkin, S. A. (1982). Estimation of earthquake source parameters by the inversion of waveform data, synthetic waveforms, *Phys. Earth Planet. Interiors* **30,** 242–259.

Sipkin, S. A. (1986a). Interpretation of non-double-couple earthquake source mechanisms derived from moment tensor inversion, *J. Geophys. Res.* **91,** 531–547.

Sipkin, S. A. (1986b). Estimation of earthquake source parameters by the inversion of waveform data: global seismicity, 1981–1983, *Bull. Seism. Soc. Am.* **76,** 1515–1541.

Smith, G. P. and G. Ekström (1996). Improving teleseismic event locations using a three-dimensional Earth model, *Bull. Seism. Soc. Am.* **86,** 123–132.

Smith, G. P. and G. Ekström (1997). Interpretation of earthquake epicenter and CMT centroid locations, in terms of rupture length and direction, *Phys. Earth Planet. Interiors* **102,** 788–796.

van der Hilst, R. D. and E. R. Engdahl (1991). On ISC PP and pP data and their use in delay-time tomography of the Caribbean region, *Geophys. J. Int.* **106,** 169–188.

van der Hilst, R. D. and E. R. Engdahl (1992). Step-wise relocation of ISC earthquake hypocenters for linearized tomographic imaging of slab structure, *Phys. Earth Planet. Interiors* **75,** 39–53.

van der Hilst, R. D., E. R. Engdahl, W. Spakman, and G. Nolet (1991). Tomographic imaging of subducted lithosphere below northwest Pacific island arcs, *Nature* **353,** 37–43.

van der Hilst, R. D., E. R. Engdahl, and W. Spakman (1993). Tomographic inversion of P and pP data for aspherical mantle structure below the northwest Pacific region, *Geophys. J. Int.* **115,** 264–302.

van der Hilst, R. D., B. L. N. Kennett, D. Christie, and J. Grant (1994). SKIPPY: mobile broad-arrays to study the seismic structure of the lithosphere and mantle beneath Australia, *EOS* **75,** 177–181.

van der Hilst, R. D., S. Widiyantoro, and E. R. Engdahl (1997). Evidence for deep mantle circulation from global tomography, *Nature* **386,** 578–584.

Vasco, D. W., L. R. Johnson, R. J. Pulliam, and P. S. Earle (1994). Robust inversion of IASP91 travel time residuals for mantle P and S velocity structure, earthquake mislocations, and station corrections, *J. Geophys. Res.* **99,** 24037–24059.

Vasco, D. W., L. R. Johnson, and R. J. Pulliam (1995). Lateral variations in mantle velocity structure and discontinuities determined from P, PP, S, SS, and SS-SdS travel time residuals, *J. Geophys. Res.* **100,** 13727–13755.

Wiens, D. A. (1987). Effects of near source bathymetry on teleseismic P waveforms, *Geophys. Res. Lett.* **14,** 761–764.

Wiens, D. A. (1989). Bathymetric effects on body waveforms from shallow subduction zone earthquakes and application to seismic processes in the Kurile trench, *J. Geophys. Res.* **94,** 2955–2972.

Yoshii, T. (1979). A detailed cross section of the deep seismic zone beneath northeastern Honshu, Japan, *Tectonophysics* **55,** 349–360.

NEIC
U.S. Geological Survey
Box 25046, Mail Stop 966
Denver Federal Center
Denver, Colorado 80225
engdahl@gldfs.cr.usgs.gov
(E.R.E., R.B.)

Department of Earth, Atmospheric and Planetary Sciences
Massachusetts Institute of Technology
Rm 54-514
Cambridge, Massachusetts 02139
(R.D.V.)

Manuscript received 3 April 1997.