Geophys. J. Int. (2004) **156,** 307–328 doi: 10.1111/j.1365-246X.2003.02137.x

Three-dimensional empirical traveltimes: construction and applications

T. Nicholson,[1,*] M. Sambridge[2] and Ó. Gudmundsson[3]

[1]*Research School of Earth Sciences, Australian National University, Canberra* ACT 0200, *Australia. E-mail: todd@rses.anu.edu.au*
[2]*Center for Advanced Data Inference, Research School of Earth Sciences, Australian National University, Canberra,* ACT 0200, *Australia. E-mail: malcolm@rses.anu.edu.au*
[3]*Danish Lithospheric Centre, Øster Voldgade* 10, 1350 *Copenhagen K, Denmark. E-mail: og@dlc.ku.dk*

Accepted 2003 October 1. Received 2003 September 22; in original form 2002 August 21

SUMMARY

A new approach for predicting traveltimes is proposed which makes use of sophisticated smoothing and interpolation techniques applied to observations from large earthquake catalogues. The technique produces 3-D empirical traveltimes directly from a database of observations and does not impose any arbitrary spatial scalelength on the data, for example by smoothing observations over a pre-determined distance interval. This feature allows the spatial variability in data density, as well as multiple length scales of heterogeneity, to be handled in a natural way. After removing the empirical traveltimes from observed traveltimes, the remaining residuals are spatially uncorrelated and, in this sense, all of the heterogeneity signal in the data is accounted for.

Traveltimes predicted with the new approach compare favourably with those from 1-D and 3-D earth models. Application to nuclear blast data shows that traveltime residuals are reduced by 62 per cent, and for earthquake residuals a reduction of up to 42 per cent is achieved across a range of teleseismic phases. Our results show that the pattern of regional variability in *P*-wave traveltimes can also be recovered accurately. In the epicentre determination of 25 nuclear blasts, the use of empirical traveltimes in place of 3-D global velocity models reduced the average mislocation by over 59 per cent.

This new approach can also be used to 'de-noise' large arrival time databases and to identify outliers. The methodology is quite flexible and can equally well be applied to local and regional seismicity.

Key words: earthquake location, global seismology, lateral heterogeneity, nuclear explosions, seismology, traveltime.

1 INTRODUCTION

The accurate prediction of traveltimes is of critical importance to many applications in seismology, such as hypocentre location and body wave imaging of the mantle. The earliest 1-D traveltime tables (e.g. Jeffreys & Bullen 1940) have been gradually improved in a series of refinements (e.g. Dziewonski & Anderson 1981; Kennett & Engdahl 1991; Kennett *et al.* 1995) to the point where the major source of error in traveltime prediction comes from unmodelled lateral heterogeneity.

At present, the effect of lateral heterogeneity on traveltime estimation is usually accounted for in two ways. For short-scale lithospheric structure, methods based on station correction techniques are often used (e.g. Wolfe *et al.* 1997). At longer scalelengths, perturbations to 1-D traveltime tables in the form of 3-D earth models parametrized by either spherical harmonics or rectangular cells are popular.

20 years of body wave tomography have led to global 3-D models featuring large-scale velocity anomalies (e.g. Su *et al.* 1994; Masters *et al.* 1996); however, global tomography has only recently begun to resolve structures smaller than 500 km (e.g. van der Hilst *et al.* 1997; Bijwaard *et al.* 1998; Boschi & Dziewonski 1999). Resolution of global earth models is limited to structures larger than about 200 km. However, the influence of lateral heterogeneity on traveltime residuals is apparent over smaller scalelengths (e.g. Gudmundsson *et al.* 1990).

Most station correction techniques simply average residuals over a large (pre-determined) volume to reduce the effects of noise, and clearly any information at shorter distance scales is lost. Perhaps more importantly, no account is taken of the spatial variability in data density, and so simple averaging may over-smooth in densely sampled regions and under-smooth in regions of sparse data.

*Now at: Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, Canada.

Nevertheless, some impressive recent studies along these lines have been performed (Piromallo & Morelli 1998, 2001), providing the motivation for the present work.

We propose a new approach, which we call 3-D empirical traveltimes (ETT). The basic idea is to estimate a traveltime, for any phase between any source–receiver pair, by directly interpolating and smoothing over a corresponding database of arrival times and their inferred hypocentres. In principle, the complete set of arrival time observations recorded at a particular receiver contains all of the information available. The philosophy in this paper is to use this database directly to infer traveltimes, rather than using an earth model, which is itself derived from the observations. In this way we aim to build a flexible tool which avoids the necessary assumptions built into the parametrization of an earth model, but which perhaps more importantly accounts for the effects of lateral heterogeneity on a wide range of spatial scalelengths, and thereby improves the accuracy of traveltime prediction.

Our approach may be thought of as station corrections, which continuously vary throughout a 3-D volume, or as a 3-D traveltime table (although the full table need not be constructed), and is in that sense similar to the Kriging approach that has been applied successfully in two dimensions (e.g. Myers & Schultz 2000).

We demonstrate the potential of the new approach using examples from International Monitoring System (IMS) stations in Australia, Canada and Finland. We investigate the accuracy of ETTs using a large number of observations of events from the catalogue of Engdahl *et al.* (1998) together with a standard set of 25 teleseismically recorded explosions, and compare their accuracy with those produced by 1-D and 3-D earth models.

We investigate several sources of error that may propagate into ETTs, including the influence of errors in the assumed origin times and locations of the database events. Finally, we relocate 25 explosions using ETTs and compare our results with those found using both a 3-D earth model and a station correction approach.

2 STATION CORRECTION TECHNIQUES

The simplest station correction procedures involve static corrections, in which a single correction is applied to all observations of a given phase at a station, irrespective of the event location (e.g. Cleary & Hales 1966; Souriau & Woodhouse 1985; Robertson & Woodhouse 1997). They are usually calculated by averaging the traveltime residuals of all observations made at the station. Despite the simplicity of this approach, Antolik *et al.* (2001) found that the mislocation of sparsely sampled ground truth events could be reduced by up to 19 per cent by applying static station corrections, calculated from observations of ground truth events, even when compared with the locations found using a 3-D global reference earth model. Joint hypocentre determination (JHD) (Douglas 1967; Dewey 1972) and master event techniques (e.g. Fitch 1975) implicitly apply a form of static station correction.

Some station correction techniques involve different corrections for individual source regions (e.g. Piromallo & Morelli 1998; Richards-Dinger & Shearer 2000). The simplest of these are cap averages, which, for global studies, involve dividing the Earth up into many small regions and calculating a correction by averaging the traveltime residuals of observations made at a station from all events within a cap. Cap averaging has the advantage of being simple to apply; however, since the caps are usually large ($5^\circ \times 5^\circ$ caps are the most common) they often cannot resolve correlations in the residual pattern at length scales smaller than approximately 500 km. Consequently, the signal due to smaller-scale structure is lost. Furthermore, the size of the cap is fixed and so they do not adapt to data density and are often invariant with depth. In addition, cap average corrections are discontinuous from one cap to the next.

Richards-Dinger & Shearer (2000) proposed a different form of cap average based on averaging only the residuals of 'nearby' events. This approach can adapt to data density. Their application of cap-averaged traveltimes to event location in a regional catalogue resulted in a significant reduction of scatter in the earthquake distribution. Piromallo & Morelli (1998) proposed a 'double-ended' cap average method called empirical heterogeneity corrections (EHC), which are a form of summary rays. They average over stations as well as events, so each heterogeneity correction applies to all stations in a region. The EHC calculated from observations made at nearby stations can be used at a new station. Piromallo & Morelli (2001) showed that the application of EHCs can reduce the mislocation of explosions and earthquakes by 17 per cent and 12 per cent respectively. Hypocentre location algorithms based on double differencing (Got *et al.* 1994; Waldhauser & Ellsworth 2000) can also be viewed as station correction approaches, although they are rarely described in these terms.

Kriging has been used to smooth and interpolate traveltime residuals to produce continuous 2-D source-specific station corrections for relatively small groups of events (Schultz *et al.* 1998; Myers & Schultz 2000). The application of kriging represents a statistically more sophisticated approach to the modelling of variations in traveltime residuals, and the basic statistical tools are similar in nature to those used in the ETT approach developed in this paper. (One difference is that kriging is computationally more demanding than the technique presented here. A brief comparison appears in the Appendix).

3 THREE-DIMENSIONAL EMPIRICAL TRAVELTIMES

By examining station correction techniques we can identify the properties that an ideal traveltime estimator needs to possess. These are: (1) the ability to account for traveltime perturbations spanning all spatial scales; (2) the ability to adapt to different densities and levels of noise in the observational data (in some self-consistent way); (3) to be computationally efficient for the largest databases available, for example those with millions of observations; and (4) to be practical enough to repeat (perhaps many times) as newer (higher quality) observations become available. Each of the station correction approaches described above have some of these properties, but none has them all. We propose a new procedure in which all of these objectives are met.

The specific problem addressed here is the estimation of teleseismic traveltimes, for a given phase, to a given seismic station using global arrival time databases. Such data sets are a major information source for seismic mantle tomography; however, the events are unevenly distributed across the globe and contain noise arising from several sources, not least of which are the mislocations of database events. To obtain a continuous, and robust, predictor of traveltime residuals we must overcome noise by averaging or smoothing the observations and filling in the gaps between events by some form of interpolation.

The issue of hypocentre location errors requires further comment. It is widely accepted that hypocentre errors contain both systematic and random components. We distinguish between the two by defining random hypocentre errors as those that are spatially uncorrelated (i.e. in terms of source position), and thus lead to spatially

incoherent time residuals, and defining systematic hypocentre errors are those that are spatially correlated. Of course other factors influence residuals, for example picking errors and lateral heterogeneity. Both may give rise to either correlated or uncorrelated residuals. We treat all factors leading to uncorrelated residuals as 'noise' and all factors leading to spatially correlated residuals as 'signal'. This is a pragmatic definition, but a clear limitation, although certainly no worse than common practice in seismic body wave tomography. The hope is that a traveltime estimator removes as much of the uncorrelated error as possible leaving only the signal, which represents the influence of lateral heterogeneity and the unwanted effects of correlated picking and systematic hypocentre errors.

The influence of correlated and systematic errors may be limited by working with a database of high quality earthquakes with clear impulsive recordings and well-determined hypocentres. However, the use of such a database makes removing random errors more difficult because of the small number of data. The use of a small database also limits the resolution of the medium. Furthermore, the highest quality ground truth data are only available from near-surface events. On the other hand, a large database allows the effect of random errors to be suppressed more effectively and small-scale structure to be resolved. However, this introduces more systematic errors because the locations of small events are more vulnerable to systematic bias (Billings *et al.* 1994). We choose to concentrate on teleseismic traveltimes, and use the large database of Engdahl *et al.* (1998) (EHB) as a balance between these two extremes.

The procedure for constructing 3-D ETTs for a given phase from a particular location **x** to a station at **y** consists of four steps.

(1) Identify an appropriate set of observations surrounding the location **x**.

(2) Interpolate and smooth the chosen traveltime residuals to produce a preliminary interpolant in the vicinity of **x**.

(3) Use the preliminary interpolant to identify and remove outliers from the database.

(4) Repeat the procedure on the filtered set of observations.

In principle, the final interpolant can be evaluated anywhere in the region containing **x** to give the ETT to **y**; however, we will normally only make use of it evaluated at **x**. The main steps of the algorithm are discussed in turn below, starting with interpolation and smoothing. It is convenient to leave the event selection until last. For the moment we simply assume that a set of N observations at station **y** are available.

3.1 Interpolating and smoothing the observations

The procedures used to produce a smooth 3-D interpolant of the traveltime field into a station are thin-plate splines (TPS) and generalized cross validation (GCV). Without smoothing, TPS interpolation gives the 3-D 'surface' of minimum curvature that fits all observations (from irregularly spaced events) exactly. Smoothing is required to minimize the effects of random hypocentre errors and observational noise in the database. GCV is the technique that imposes smoothing on the TPS interpolation. It is convenient because it is an automated process in which the level of smoothing is determined by the amount of noise in the observations. Mathematical details of the two procedures are given in Section 4. Note, however, that smoothing only accounts for random, spatially uncorrelated effects and not for systematic errors in locations of events (as discussed above).

These smoothing and interpolation tools were used in Nicholson *et al.* (2002) for the hypocentre location problem. In this earlier work they were applied to the construction of a hypocentre misfit function measuring the similarity between a new event and previously well-located events. In the present study they are applied directly to traveltimes to represent the structural signal 'seen' by a receiving station.

3.2 Removal of outliers

In step 3, outliers in the traveltime residuals are filtered because they can lead to over-smoothing of the ETT function. This can occur because generalized cross validation minimizes a least-squares measure of prediction error. A large error in a single observation can lead to an unrealistically large estimate of error across the entire distribution and consequently to over-smoothing. Outliers may be due to sporadic forms of noise, such as phase misidentification, which have a large effect on a small number of observations but have no effect on the vast majority. Fortunately, the preliminary ETT function can be used to identify outliers reliably and remove them. This is done by simply rejecting all observations for which the ETT residual after step 1 is more than two standard deviations from the mean, which results in the removal of approximately 5 per cent of observations. This is reasonable because a large residual at the first stage indicates a noisy traveltime residual which does not correlate well with those of nearby events. (Further details of outlier removal in large earthquake databases appear in Nicholson 2002.)

3.3 Finding the optimal number of observations

The selection of events in the vicinity of **x** can be performed either by taking all events within a specific epicentral distance, or just the nearest N events. The latter approach guarantees a fixed number of data available for each interpolation and is used throughout this study. The choice of N will influence the accuracy of the estimated traveltimes, and some experimentation is required to find an optimum value.

We calculate ETTs for 10 000 randomly chosen P arrivals from the EHB catalogue and examine their accuracy as a function of N. In each case the ETT was based on corresponding arrivals observed at the same station, from the N events nearest the estimated hypocentre location. The spread of the differences between calculated and observed arrival times was determined for each value of N Note that spread is a robust analogue of standard deviation, and is used throughout this paper It is defined as the median of the absolute deviations from the median. A normalization factor of 1.4826 is used to yield approximately the same value as the standard deviation for data with a Gaussian distribution. For further details (see Hampel *et al.* 1986).

Fig. 1 shows the spread of the prediction errors (i.e. the difference between the ETT and the observed times) as a function of N. Note that for each of the 10 000 calculations the observed arrival time is not included in the corresponding database (i.e. it does not predict itself). Therefore this is a true cross-validation test in which the ETT performance is judged on its ability to estimate independent data. The spread of the *ak135* residuals is 1.20 s. Clearly the ETT is significantly below this value for all N. The largest reduction in spread, of 42.4 per cent over *ak135*, is found with 400 observations, although any value between 200 and 800 observations could be used without increasing the spread significantly. We choose a value of 400 for N irrespective of the epicentral distance between events. If fewer than 400 observations are available at a station we take them all.

In summary, to calculate the ETT from a location **x**, we take the nearest 400 *ak135* residuals in the database, interpolate and smooth them using thin-plate splines and generalized cross validation to

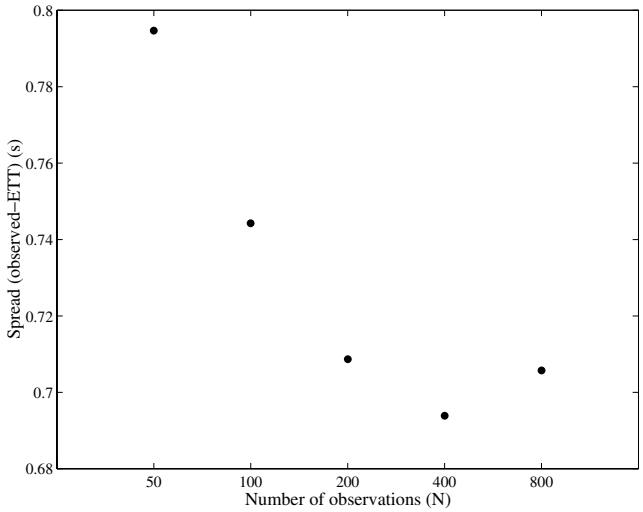

Figure 1. The spread of ETT prediction errors for 10 000 excluded P observations as a function of the number of observations used to calculate the ETT function. The spread of the *ak135* residuals was 1.20 s. All values of N reduced the residual spread significantly, but $N = 400$ produced the largest reduction.

obtain a preliminary ETT function. We then evaluate the preliminary ETT function at the locations of the 400 observations and compare these predictions with the observed values. Those observations with significantly large residuals are removed as outliers. The interpolation and smoothing are then repeated with the reduced data set to produce the final ETT.

4 THIN-PLATE SPLINE INTERPOLATION AND GENERALIZED CROSS VALIDATION

In this section we outline the statistical tools used in the 3-D ETT procedure. Some further technical details are given in the Appendix. Thin-plate spline (TPS) interpolation has been widely used in a range of disciplines including airborne geophysical surveys (Billings 1998), mapping deformations of the human heart (Sanchez-Ortiz 1996) and earthquake hypocentre determination via pattern recognition (Nicholson *et al.* 2002). Exact TPS interpolation can be used to produce the curve or, in multiple dimensions, hyper-surface, of 'minimum curvature' that fits an arbitrarily distributed set of traveltime observations exactly at the locations of the events. This is a 'least structure' approach and is similar to the Occam's inversion commonly used in geomagnetism (e.g. Constable *et al.* 1987). Note that each traveltime residual is associated with the location of the event from which it was observed.

The expansion of the thin-plate spline interpolant, s, at any point $\mathbf{x} = (x, y, z)$ has the form

$$s(\mathbf{x}) = p(\mathbf{x}) + \sum_{n=1}^{N} \lambda_n (\|\mathbf{x} - \mathbf{x}_n\|)^2 \ln(\|\mathbf{x} - \mathbf{x}_n\|), \quad (1)$$

where $\|.\|$ is the standard Euclidean norm. The set of N weights $\{\lambda_n : n = 1, \ldots, N\}$ are defined at the N nodes or centres $\{\mathbf{x}_n : n = 1, \ldots, N\}$ (these are the locations of the N observations), and the polynomial $p(\mathbf{x})$ is given by

$$p(\mathbf{x}) = a_1 + a_2 x + a_3 y + a_4 z. \quad (2)$$

This polynomial accounts for the scalar mean of the distribution and large-scale linear trends in the function across the entire data distribution, while the weights represent the smaller-scale signal. Therefore the second term in eq. (1) accounts for small-scale correlations in the data. The calculation of the parameters λ_n and a_i from the data constitute the construction phase of the procedure, the details of which are given in the Appendix. Once the polynomial coefficients and weights have been determined we can calculate the value of the TPS interpolant anywhere simply by evaluating eq. (1). Therefore, if we wanted to store all the information necessary to reconstruct any interpolant we would need only $4N + 4$ floating-point numbers (i.e. the $3N$ coordinates of the locations of the observations, the N weights and the four polynomial coefficients).

An illustration of TPS interpolation is shown in Fig. 2(a) for a simple 1-D example. Here the problem is to reconstruct the function c (the solid black/red curve) from the noisy data (shown as crosses). In Fig. 2(a), each 'observation' is generated using

$$c(x, \epsilon) = \sin(x) + 0.25\sin(10x) + \epsilon, \quad (3)$$

where the first two terms on the right represent the signal and the last term the noise in each of the 200 'observations' of the function c. In this case, ϵ is a zero-mean Gaussian random variable with a standard deviation of 0.25. Note that, in this 1-D example, no error is present in the locations of the observations. The long-period oscillation is clear in the observations. However, the short-period oscillation has a similar amplitude to the noise and is therefore more difficult to distinguish. This short-period oscillation would be very difficult to resolve simply by averaging over a series of bins, even if we had prior knowledge of the frequency of the oscillation.

Fig. 2(b) shows the exact TPS interpolant of the data (in dark grey/blue). Note the large overshoots, which stem from the minimum-curvature property of the TPS interpolation and the (present) requirement that the curve fit the data exactly. Even with the overshoots the curve is actually 'smooth', in terms of minimizing second derivatives, and so fits the long-period oscillation quite well in some parts. It is clear that the requirement of exact data fit must be relaxed in order to reduce the effects of noise in the data.

To determine a smoothed version of the TPS interpolant, for any given smoothing parameter (μ), one minimizes a new quantity H, which is a combination of data misfit and curvature:

$$H(s, \mu) = \sum_{i=1}^{N} [s(\mathbf{x}_i) - c_i]^2 + \mu J(s), \quad (4)$$

where $\mathbf{x}_i$ is the location of the ith observation, c_i is the value of that observation, $s(\mathbf{x}_i)$ is the value of the interpolant at $\mathbf{x}_i$, and $J(s)$ is a measure of the smoothness of $s(\mathbf{x})$. For details of how this function is minimized see the Appendix. $J(s)$ is the curvature measure that exact TPS interpolation minimizes. Thus to determine an appropriate level of smoothing we must consider the common trade-off between data fit and smoothness, which is regulated by the smoothing parameter μ.

An arbitrary choice of μ may result in under- or over-smoothing, both of which reduce the accuracy of the interpolant. Generalized cross validation (GCV) provides an automated way to select μ based on a quantitative measure of the noise in the data and is commonly used with TPS interpolation (e.g. Wahba 1990). The basic procedure is to cycle through a range of μ values and for each to calculate N interpolants such that the ith interpolant is constructed with the ith datum excluded. Each interpolant can then be used to predict the value of the excluded datum. Noisy observations are likely to result in large differences between predicted and observed values. A quantity, $G(\mu)$, is defined as the sum of the squares of the differences

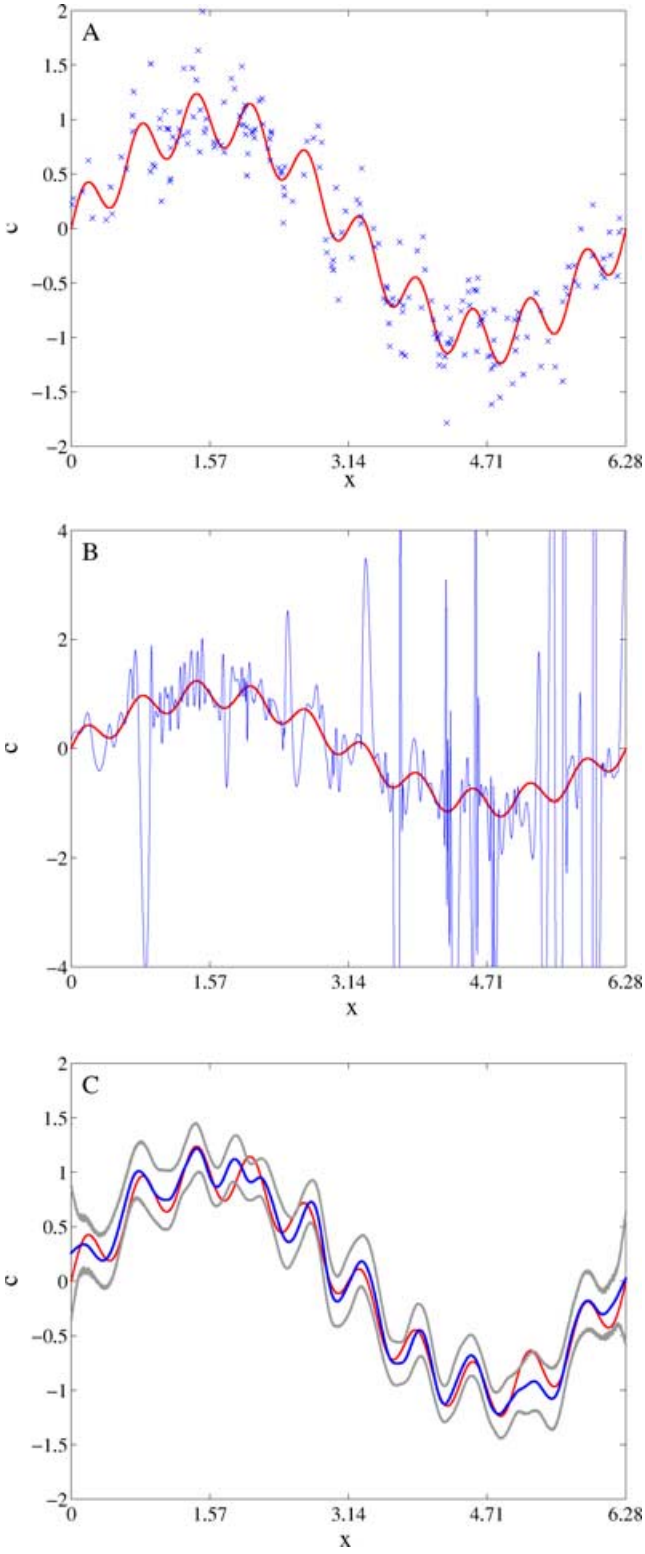

Figure 2. A simple 1-D synthetic example. We are attempting to reconstruct the black/red line, which is the noise-free function ($\epsilon = 0$) given by eq. (3). (a) The crosses show 200 'observations' with random x in the range $[0, 2\pi]$ and c given by eq. (3). (b) The exact TPS interpolant of the observations. Note the large overshoots. (c) The smoothed TPS interpolant (dark grey/blue line), with smoothing parameter chosen by GCV. The two light grey lines denote the 95 per cent confidence interval for the smoothed interpolant. The smoothed interpolant fits the true function well and the true function lies within the confidence intervals for almost all values of x.

between the observed and predicted times. Specifically, $G(\mu)$, is given by

$$G(\mu) = \sum_{i=1}^{N} (s_i[\mathbf{x}_i] - c_i)^2, \tag{5}$$

where s_i is the smoothed TPS interpolant calculated from all the data points except the ith datum. G provides a measure of the effectiveness of that value of μ. A large G implies a poor fit to the observations. For a small μ the interpolant will tend to fit the noisy data very well (as in Fig. 2) and be a poor predictor of each excluded datum, leading to high G. At the opposite extreme, a high μ will give a very smooth interpolant which will also fit the excluded data poorly, and lead to high G. At some intermediate value of μ a minimum of G will be found, which can be taken as an optimal choice of smoothing parameter. Hence one needs to minimize G with respect to μ.

It would appear then that a large number of interpolants need to be calculated to find an optimal μ, i.e. one for each value of μ and each excluded data point. Fortunately, this is not the case because an analytical expression is available to estimate an optimal μ (Craven & Wahba 1979, see the Appendix for details). (A plot of G against μ and a comparison between the analytical and exact value is shown in Fig. 14, for the traveltime problem discussed in Section 7.)

In the 1-D example the smoothed interpolant is shown in Fig. 2(c) and corresponds to a smoothing parameter of 0.578×10^{-3}. The larger wavelength sine function has been resolved across all values of x, and the smaller-wavelength oscillations are now clearly visible. Recovery of the original function is, of course, not perfect. In particular, the fourth and ninth peaks are poorly resolved, which can be attributed to the random alignment of several data points near these peaks, as shown in Fig. 2(a). Confidence intervals for the interpolated curve can also be calculated, and the 95 per cent confidence interval for the 1-D example is shown in Fig. 2(c). The ability to estimate confidence intervals can be used to give error bounds on the recovered signal. Here they are shown as an outer (light grey) pair of curves and almost everywhere the true curve lies within them.

In TPS interpolation and GCV smoothing we have a minimum-curvature approach to interpolation and an automated way of determining the optimal smoothing parameter based on the level of noise in the data. In the next section we use these tools to construct 3-D ETT 'surfaces' from a database of arrival time observations.

5 ETTS AT INTERNATIONAL MONITORING SYSTEM STATIONS

Our goal is to represent the Earth as 'seen' by a particular receiver. In this section we show several examples of the empirical technique, for P arrivals observed at the primary International Monitoring System (IMS) arrays called WRA, YKA and FINES in Australia, Canada and Finland respectively. These stations were chosen for their longevity of operation and because they are a reasonably representative, if small, sample of the overall distribution of primary IMS stations. Note that, throughout this paper, all calculations of ETTs for a particular source, receiver and phase exclude the original database observation. Hence all assessments of ETT accuracy are based on proper 'cross-validating' tests.

5.1 Regional and teleseismic *P* observations made at WRA between 1966 and 1998

Fig. 3(a) shows the observed *ak135* residuals of 49 546 *P* and *Pn* arrivals reported from WRA, plotted at the EHB location of each event. The median residual, of −0.83 s, has been removed from all the traveltime residuals shown in Fig. 3(a) for plotting purposes. Despite the noise in the residuals, several features clearly stand out in Fig. 3(a). Arrivals from shallow events in the Tonga subduction zone are fast since large portions of their ray paths lie within the subducting slab. This may also be an indication of a fast near-receiver anomaly since the arrivals from events at mid-ocean ridges south and southeast of the station, which we would expect to follow slow near-source paths, also have large negative residuals.

Most arrivals from events west of the station follow slower-than-average paths, which supports the presence of a low-velocity

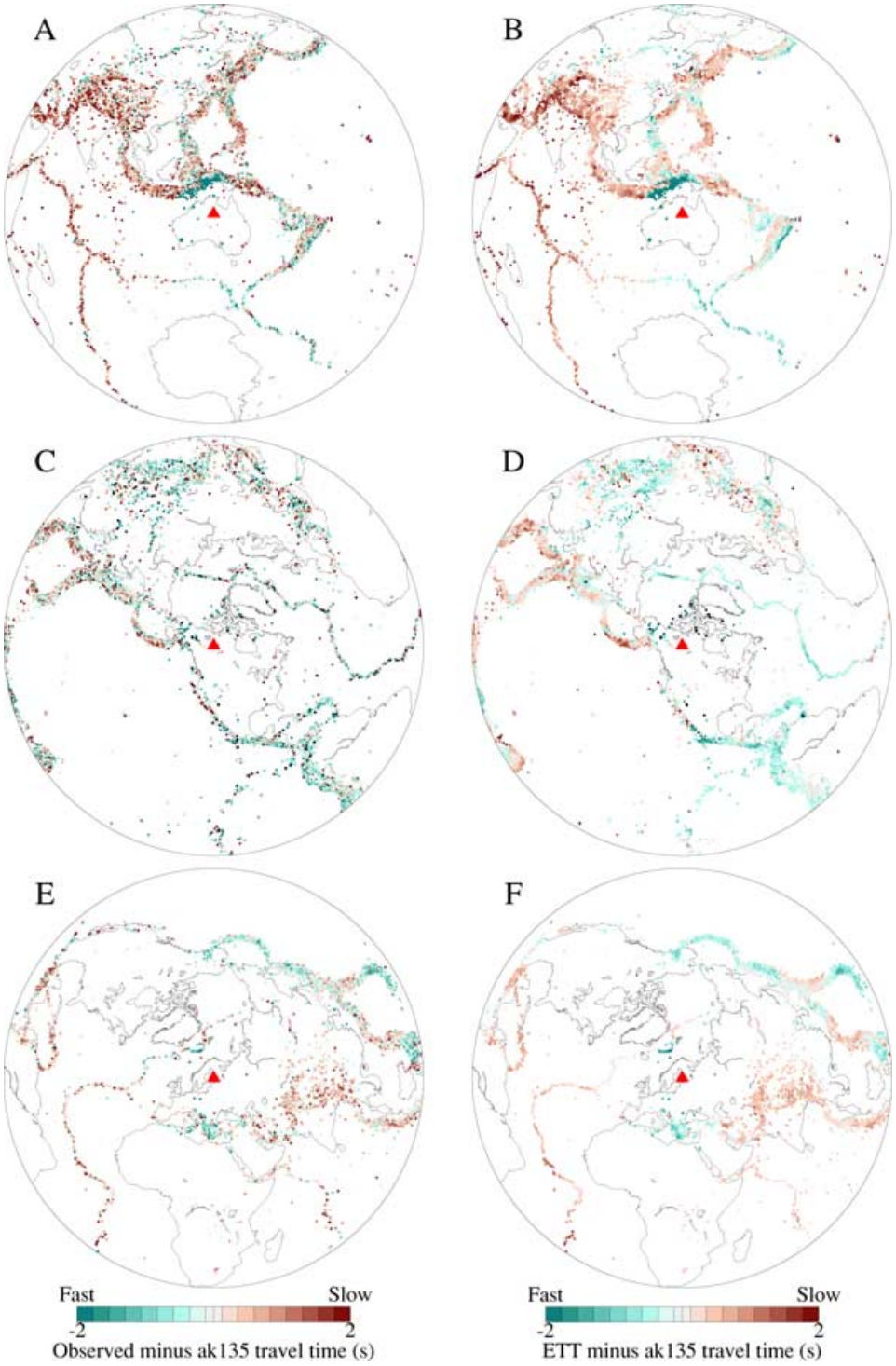

Figure 3. The observed traveltime (left) and ETT (right) minus that predicted by *ak135,* plotted at the EHB locations of the events, for three IMS stations. (a) and (b) *P* and *Pn* observations made at WRA; (c) and (d) *P* observations made at YKA, (e) and (f) *P* observations made at FINES.

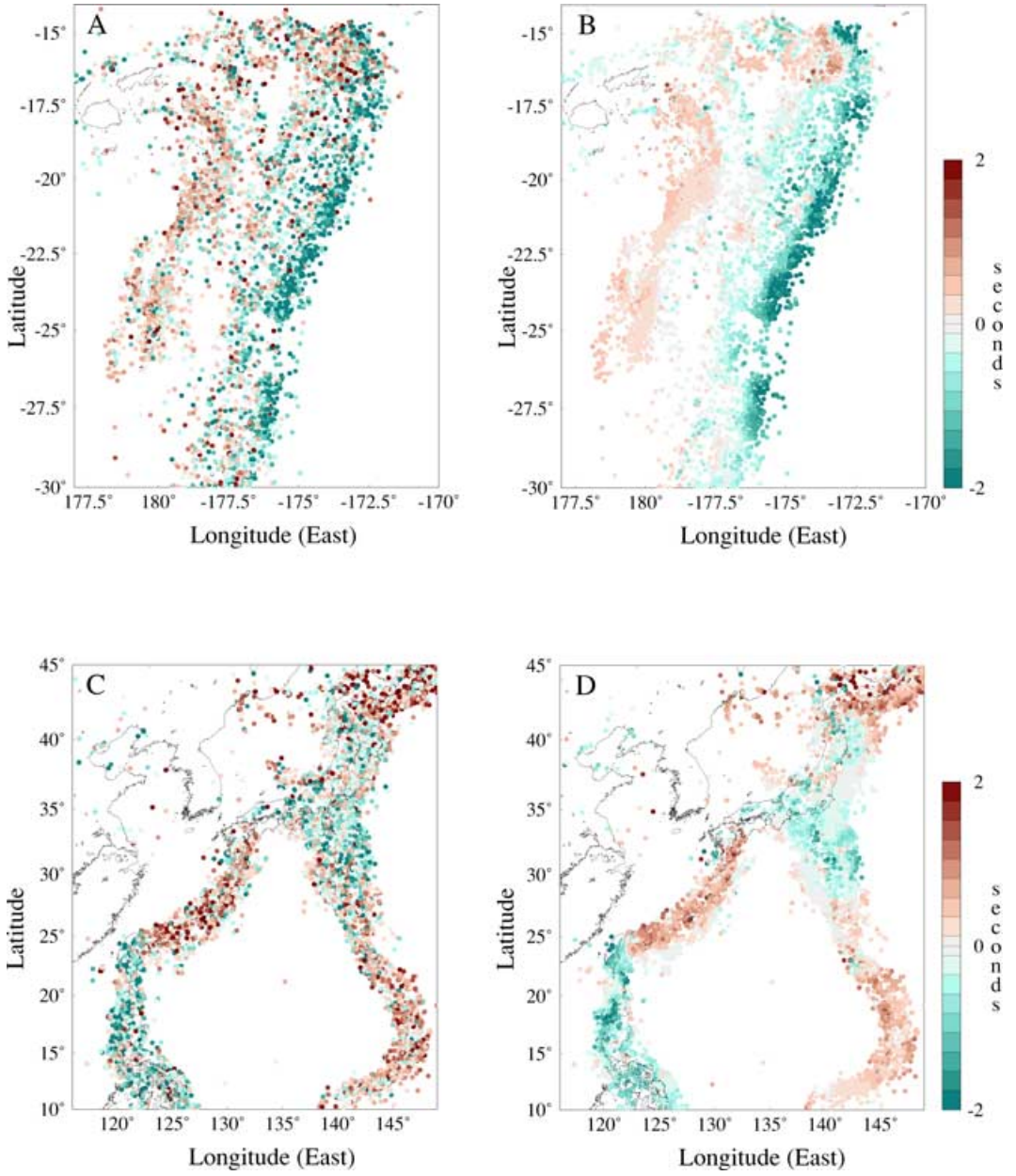

Figure 4. The observed traveltime (left) and ETT (right) minus that predicted by *ak135,* plotted at the EHB locations of the events, for *P* and *Pn* arrivals at WRA. The median residual in each region has been subtracted from each time. The ETT residuals represent the correlated portion of the *ak135* residuals. Note that the ETTs resolve the residual variation as a function of depth well. This is particularly clear in (a) and (b), where shallow events (to the east) have negative residuals while deep events have positive residuals.

anomaly, at a depth of greater than 300 km, beneath the west Australian Craton (Gorbatov, private communication). The pattern of residuals appears to be more complicated north of the station. Arrivals from events in the Indonesian, Mariana, Ryuku and Kurile subduction zones are all slow while those from the Philippine and Japan subduction zones are fast. The *Pn* arrivals from shallow events in Indonesia and New Guinea follow very fast paths to WRA, indicating that the fast anomaly known to exist beneath the station (e.g. van der Hilst *et al.* 1998), at depths shallower than 250 km, extends at least part of the way to these locations.

The ETTs to the locations of these events are shown in Fig. 3(b). All of the features that are clear in the raw residuals are also clear in the ETTs. In particular, the fast arrivals from shallow events in the Tonga and Philippine subduction zones are apparent, as well as the slow arrivals from the continental and mid-ocean-ridge seismicity to the west of the station. There is good agreement between the raw residuals and ETTs in the very sparsely sampled intraplate seismicity in east Africa, the Indian Ocean and the Hawaiian Islands. The very early *Pn* arrivals for a group of events just north of Australia stand out as by far the largest deviation from *ak135*. The amplitude of the correlated signal in the raw residuals is also well reproduced by the ETTs.

In Fig. 4 we focus on arrivals from the Tonga and Japan subduction zones. The raw residuals (4a) from events in the Tonga subduction zone show significant variation with depth as well as epicentral position and are well reproduced in the ETT map (4b). Arrivals from shallow events are likely to follow paths that pass through the subducting slab, and are therefore fast. A smaller portion of the ray

paths from deeper events lie within the slab so they follow slower-than-average paths. The same trend is seen in the Ryuku and Japan subduction zones, although the contrast is less marked.

The shorter-scale features in the residual pattern are also well resolved by the ETTs. We note that the sizes of some of these features are of the order of the distance between events and are well beyond the resolving power of summary rays or cap averages because these procedures perform simple averaging of many nearby events. Given that we are dealing with seismic waves with a wavelength of approximately 10 km, one would expect that the sensitivity of arrival times to heterogeneity would be at much larger distance scales. However, we do not impose any minimum scalelength on the procedure and so all effects are included.

5.2 Teleseismic *P* observations made at YKA between 1976 and 1998

Fig. 3(c) shows the traveltime residuals of all *P* arrivals at the YKA station in Yellowknife, Canada. The median residual, of −0.56 s, has again been subtracted from each figure. The Yellowknife array lies on the northern shore of the Great Slave Lake at the western edge of the Canadian Shield. As was the case for the WRA station there is considerable correlation in the pattern of residuals; however, the amplitude of most of the residuals at YKA appears smaller than at WRA. Most observations south, east and north of the station have negative residuals because the near-receiver portion of their ray path passes through the fast shield. However, some regions in southern Europe and continental Asia have positive (slow) residuals. Furthermore, since the Yellowknife array is on the western edge of the Canadian Shield, the majority of arrivals from the Pacific subduction zones, to the west of the station, have positive (slow) residuals. Even so, small groups of events, such as those near Vanuatu and deep events in the Kuriles, follow faster paths. Fig. 3(d) shows the ETTs from the same locations. All the major features in the *ak135* residuals are clear in the ETTs as well as numerous smaller features in the Pacific and in continental Asia. As before the ETTs have been successful in reproducing the correlated signal of heterogeneity while removing the random noise due to observational noise and random location errors.

5.3 Teleseismic *P* observations made at FINES between 1993 and 1998

The last station we consider in detail is the array beam reference point (FINES) of the FINESS array in southern Finland. This array has been in operation since an upgrade in late 1993. 13 785 *P* arrivals are reported in the EHB catalogue from this station. The array lies on the Fennoscandanavian Shield and therefore has a large negative median residual (−0.90 s). Figs 3(e) and (f) show the *ak135* residuals and ETTs to FINES, respectively. Arrivals from events west of the station, in particular in large parts of the Mid-Atlantic ridge and Central America, tend to follow slow paths. The same is true for the majority of arrivals from events east of the station, in particular those in continental Asia and the Indonesian subduction zone. As expected, most arrivals from events in the Pacific subduction zones follow fast paths. Contrary to the trend seen in most subduction zone regions, shallow events in the southern Japan and northern Mariana islands follow slow paths while deeper events are neutral. Arrivals from events in the portion of the Atlantic ridge just north of Scandinavia follow fast paths. This may be due to the fact that a large portion of their path lies in the fast keel of the Baltic Shield. Fast paths are followed by the arrivals of events near Greece since this seismicity is downdip from the Hellenic slab. The ETTs reproduce the amplitude and shape of all these features well, without being influenced by the random noise that blurs the underlying pattern in Fig. 3(e).

5.4 Residual histograms and autocorrelation functions

Fig. 5 shows histograms of the observed minus *ak135* times and observed minus ETTs for the three IMS stations shown in Fig. 3. For all three stations the distribution of ETT residuals (i.e. the

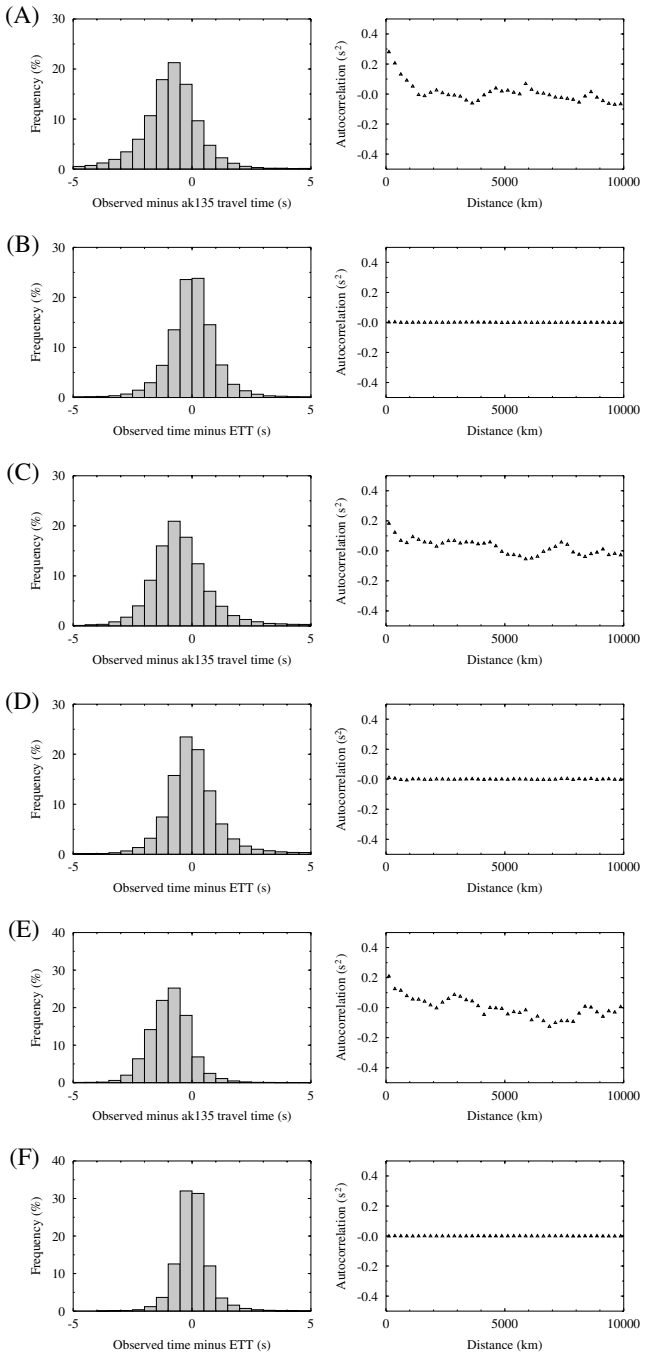

Figure 5. On the left, the *ak135* and ETT residual distributions for the observations and ETTs shown in Fig. 3. The letters used in this figure correspond to those used in Fig. 3. On the right, the autocorrelation functions for these residuals. The *ak135* residuals are correlated on many scalelengths, but the ETT residuals are totally uncorrelated.

observed time minus the ETT) has a mean of zero, and the spread of the ETT residuals is over 14.4 per cent less than the spread of the *ak135* residuals. The largest reduction is in the spread of the FINES residuals (31.2 per cent).

To measure the degree of spatial correlation in the residuals we use an autocorrelation function. The autocorrelation, a, of the jth and kth observations, $r_j = r(\mathbf{x}_j)$ and $r_k = r(\mathbf{x}_k)$, at coordinates $\mathbf{x}_j$ and $\mathbf{x}_k$ is defined as

$$a_{j,k} = a(\mathbf{x}_j, \mathbf{x}_k) = E\{r(\mathbf{x}_j)r(\mathbf{x}_k)\}. \quad (6)$$

Here $E\{\}$ represents the expectation operator. We assume that the statistics of the residuals, r, are stationary and isotropic such that the autocorrelation is only a function of the distance from $\mathbf{x}_j$ to $\mathbf{x}_k$:

$$a_{j,k} = a(\|\mathbf{x}_j - \mathbf{x}_k\|). \quad (7)$$

We can then estimate the autocorrelation function by binning the $a_{j,k}$ according to the distance $\|\mathbf{x}_j - \mathbf{x}_k\|$ and averaging the products $r_j r_k$ within each bin. The bins vary between 50 and 200 km across depending on data density.

When the mean is removed from the residuals before computing the autocorrelation, as we have done in all cases, it becomes equivalent to the autocovariance. The interpretation of the autocovariance is simple. When the distance is zero the residuals are only correlated with themselves. Therefore, $a(0)$ is the variance of the residuals, which may in part be due to structure and in part be due to noise. If the autocorrelation is greater than zero at any finite distance, then the residuals are positively correlated over that distance. That correlation must be due to structure if we assume that the noise processes in the residuals are spatially incoherent. Conversely, if the autocorrelation is negative then the residuals are anti-correlated, and if the autocorrelation is zero then the residuals are not correlated over that distance. We hope that the autocorrelation of the ETT residuals, and the residuals of 3-D earth models (see Section 7) are zero over a wide range of distances such that the signal due to heterogeneity contained in the raw residuals has been completely resolved. This does not apply for zero distance, where we expect that the ETT residuals will still contain finite noise yielding finite variance. In the autocorrelations we show in this study we do not plot the variance at zero offset.

Fig. 5 also shows the autocorrelation functions for the *ak135* residuals, once a static station correction has been applied, and the ETT residuals. The *ak135* residuals of all three stations are correlated over a range of distances. The strongest correlations are seen at very short distances (<200 km), which are well beyond the resolution of spherical harmonic models and at the limit of resolution of many block models. The autocorrelation function of the ETT residuals for all three stations is zero at all distances so the ETT functions have successfully resolved all of the correlated signal in the raw residuals and accurately represent the degree of correlation present. Note that the autocorrelation is measured in s^2 and, although the amplitude appears quite small, it corresponds to significant spatially correlated patterns across the region, as seen in Fig. 3.

5.5 The influence of hypocentre errors

Here we perform two simple tests to demonstrate how random errors in the database influence the resulting ETTs. Both use P observations from the station MAT and involve calculating ETTs for the 6262 observations shown in Fig. 12(a). We separately examine the influence of the random errors in picking and origin times, which directly affect the observations, and hypocentres, which indirectly contribute to residuals. In the first test we compare ETTs calculated directly from the database with those where random Gaussian noise has been added to residuals. This simulates the effects of random noise in the picking and origin times. Fig. 6(a) shows the difference between the original ETTs and the 'noise-added' ETTs as a function of the noise spread. The figure clearly shows that the influence of random observational error is minimal.

In the second test we add random Gaussian noise to the locations of the events and recalculate the ETTs. The standard deviation of the depth noise is twice that of the epicentral noise. Fig. 6(b) shows the change in the ETT spread as a function of the induced location error. Again the figure shows clearly that the hypocentre errors have only a small effect on the resulting estimated traveltime.

In both these examples the spread change is less than 0.2 s for realistic amounts of added noise—a small change compared to the size of the residuals (see Fig. 8a). In addition, as the amount of added noise increases, the effect of adding even more noise diminishes. The propagation of random errors from both picking and hypocentre sources is effectively minimized by the statistical smoothing of the ETT algorithm.

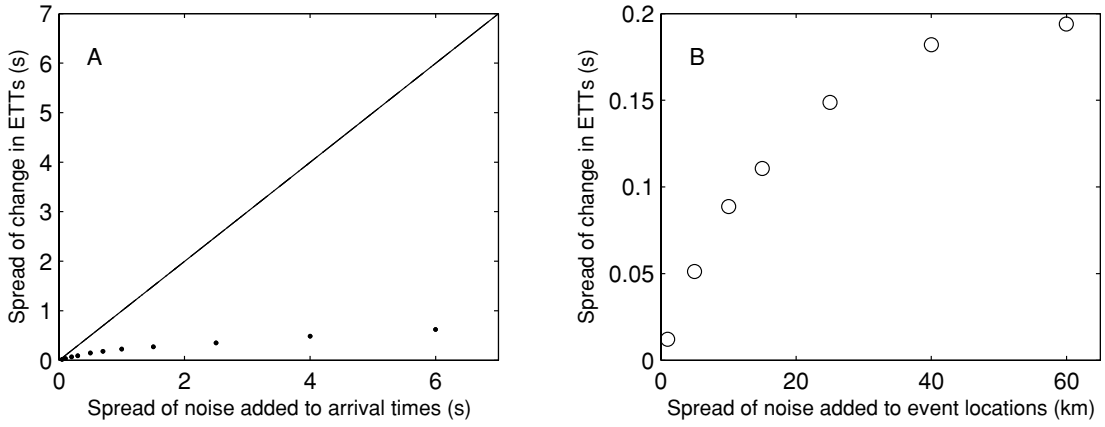

Figure 6. The change in ETTs as random errors are added to the database. (a) Random errors added to the observation residuals. (b) Random errors added to the locations of the database events. The smoothing used in the ETT approach has significantly damped the added error so that the ETTs are relatively weakly effected.

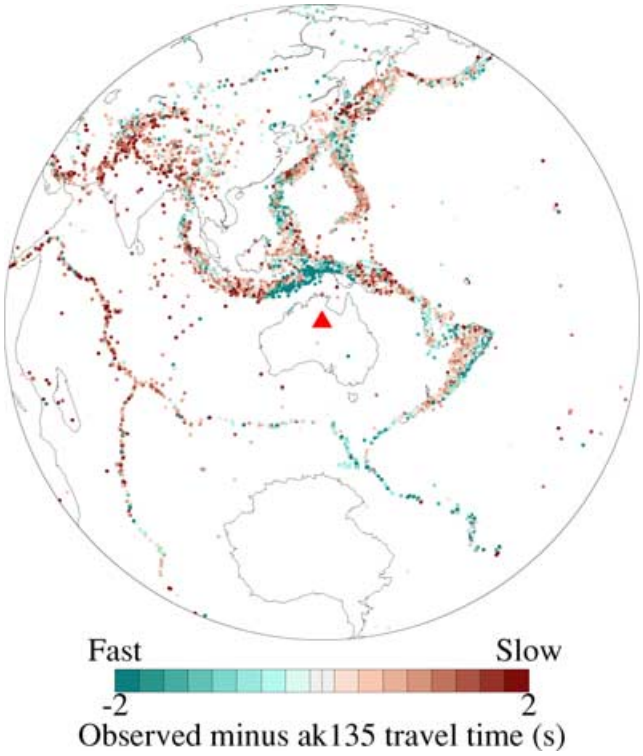

Figure 7. ETTs for *P* arrivals reported by WB2. The *ak135* time has been subtracted from each traveltime. The median time (−0.84 s) has also been subtracted for plotting purposes. Note how similar this figure is to Fig. 3(b).

5.6 The influence of station picking practices

A useful test of the ETT procedure is to compare the pattern of arrivals produced for two nearby stations. If they differ significantly there is the potential for a significant contribution from the picking routine used at either receiver. To assess the influence of picking practices we compare the ETTs for the Warramunga array (WRA) with those for WB2, a three-component station near the crossing of the two arms of WRA. These stations are at the same location, but have differing picking procedures. At WB2 the arrivals are identified manually, while at WRA an automated array processing method is used.

For each *P* or *Pn* observation made at WB2, we calculated an ETT using observations from WRA and vice versa. The station WB2 has only been in operation since 1974, and has made less than half (23, 134) of the number of observations of WRA. Fig. 3(b) shows the ETTs obtained from the WRA data. The ETTs are almost indistinguishable from those of WB2 (shown in Fig. 7). It is not surprising then that distributions of residuals are similar, with medians of −0.81 s and −0.84 s, and spreads of 0.61 s and 0.58 s, for WRA and WB2 respectively. Using WRA data to predict traveltimes observed by WB2, we get a spread of 0.75 s, and using WB2 data to predict the WRA residuals the spread becomes 0.71 s; both of which are less than the 1-D reference model *ak135* for which the residual spread is 0.95 s.

We can conclude that, for this example at least, the manual and automated picking procedures lead to similar ETTs. Hence picking procedure is not a significant factor. The use of one station to predict observations at another is in this case only used as a check of the methodology. However, where very few observations are available at a station, more reliable predictions might be possible by using the observations made at a nearby station (with greater recording length) than by using simply the observations of the actual station. This is the case for the MAJ0 (134 observations) and MAT (43 893 observations) stations in Matsushiro, Japan. Using MAT ETTs reduces the spread of the MAJ0 residuals by 24 per cent.

To summarize this section, we have shown that the ETT procedure reproduces all of the spatially correlated signal in the observations at a single station, leaving only uncorrelated noise. The influence of the random component of hypocentre and picking errors, as well as the picking practices at recording stations, have a relatively minor effect on the estimated traveltimes.

6 ACCURACY AND RESOLUTION

6.1 Teleseismic phases

In this section, we measure the accuracy of the ETT procedure for 12 of the most commonly reported teleseismic phases, including *P*, *S*, *Pn*, *Sn* and *PKP* branches, at globally distributed ISC stations. For each phase, a large number of arrivals are randomly chosen from the EHB catalogue and excluded one at a time from the database. An ETT is then estimated at the location of the excluded datum and compared with the observed traveltime to give a cross-validating measure of the accuracy.

Throughout we compare the accuracy of ETTs with that of *ak135* alone and, for comparison, *ak135* with static station, ellipticity and elevation corrections. In all cases spread is used in place of the more common variance reduction. ETTs were only calculated if more than 100 observations were available for a given phase and station. We quote the percentage of observations that satisfy this requirement in each case and assume throughout that the EHB phase identifications are correct.

Fig. 8 shows the distribution of *ak135* residuals and ETT residuals for each phase. Overall it was possible to calculate ETTs for 99 252 of 100 000 randomly chosen *P* observations from the EHB catalogue. As expected, *ak135* predicts the median of the observations well and produces a symmetrical distribution of residuals that is approximately Gaussian in the −2 s to +2 s range. The tails are broader than a Gaussian, consistent with previous studies (e.g. Buland 1986). The spread of the 99 252 *ak135* residuals is 1.20 s. When station statics, ellipticity and station elevation corrections are applied the residual spread becomes 0.98 s. The overall spread of the ETT residuals is 0.70 s, which is a reduction of 29 per cent over *ak135* plus corrections and 42 per cent over *ak135* without corrections. Note that, despite this improvement, the tails of the two distributions are almost identical, indicating that ETTs do not reduce the residuals of outliers.

Fig. 8(c) shows the distribution of *ak135* residuals for 10 000 randomly chosen *Pn* observations from the EHB catalogue. The spread of *ak135 Pn* residuals is 1.84 s. This is a considerably higher spread than for teleseismic *P* residuals, which is not surprising as *Pn* has a longer propagation path within the strongly heterogeneous near-surface layers (for further details see Nicholson 2002). The application of corrections only reduces the spread by 1 per cent to 1.82 s. In this case it was possible to calculate ETTs for 94.1 per cent of the *Pn* arrivals, and the resulting residual spread was 1.11 s, which is a reduction of 40 per cent over the reference model. The most likely explanation of this large improvement is that the bulk of the signal from heterogeneity cannot simply be attributed to the portion of the ray path directly beneath the station and therefore static station corrections are not effective. A density plot of ETT and *ak135* residuals from 117 000 randomly chosen *P* and *Pn* arrivals as a function of epicentral distance is shown in Fig. 9. Note that the

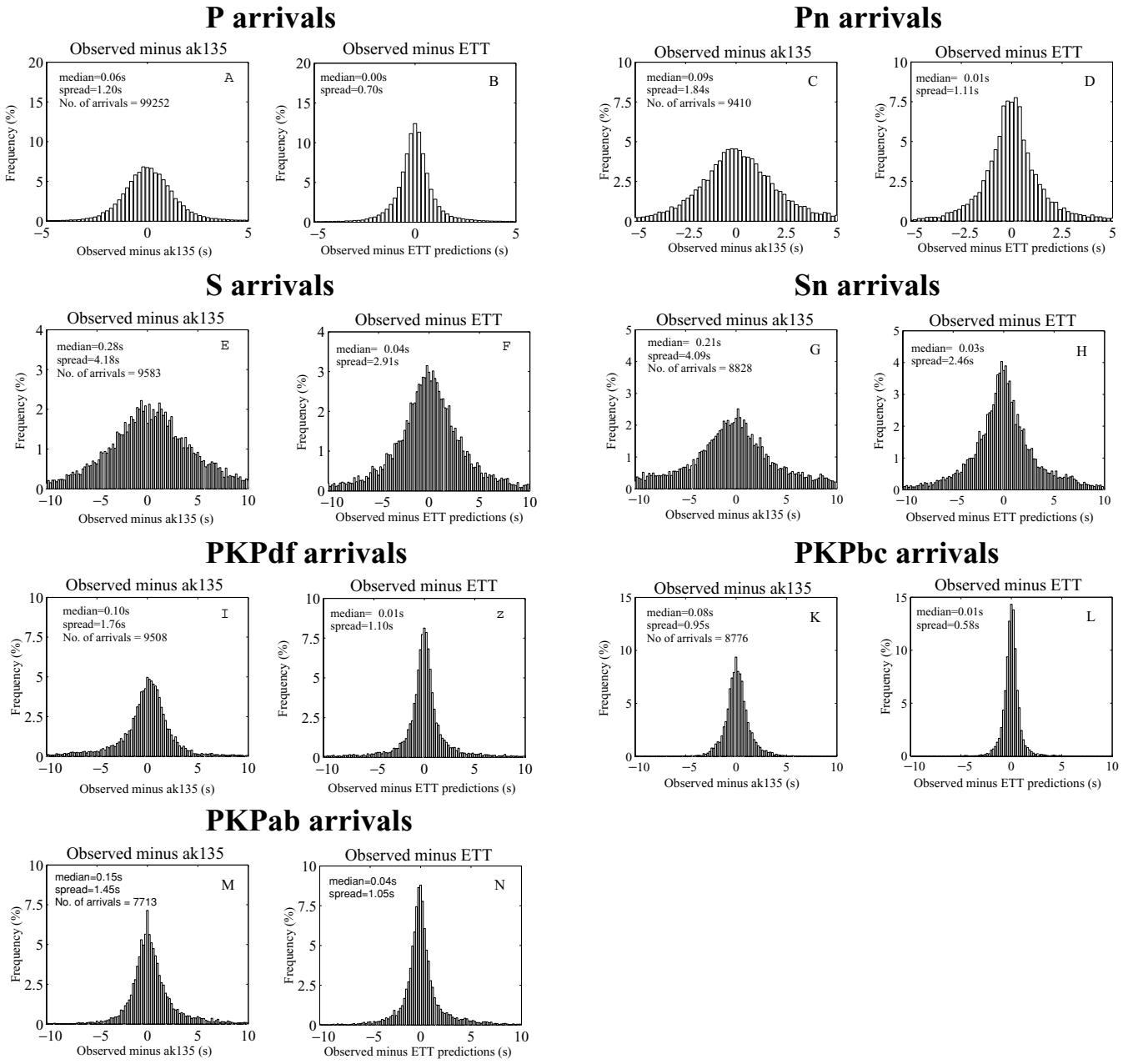

Figure 8. The *ak135* and ETT residuals for a large number of observations of seven common phases. For each ETT, the observation being predicted was excluded from the database. Using ETTs in place of *ak135* significantly reduces error for all seven phases. The number of arrivals shown in each case represents the number of observations (out of 100 000 for *P* and 10 000 for the other phases) for which ETTs were possible (see text).

spread reduction for ETT residuals lies between 35 and 50 per cent across the complete range of epicentral distances.

Fig. 8(e) shows the *ak135* residuals for 9583 *S* arrivals, which have a spread of 4.18 s. Engdahl *et al.* (1998) simply approximate static station corrections for *S*-wave arrivals as $\sqrt{3}$ times the corresponding *P* correction. However, this relies on the proportionality of *P* and *S* residuals. Although this assumption is often reasonable, the application of static station, ellipticity and elevation corrections has little effect on the residual spread, which is only reduced by 5 per cent to 3.99 s. ETTs have the potential to give a much greater reduction since they are calculated directly from *S* arrival time observations. Fig. 8(f) shows the distribution of ETT residuals for these arrivals. The spread has been significantly reduced (30.5 per cent), indicating that the EHB approximate static station corrections are inappropriate and/or that the traveltime perturbation due to near-source and deep heterogeneity is large.

Similar results are shown in Figs 8(g) and (h) for *Sn* arrivals. The spread reduction achieved with ETTs is larger than for *S* (39.9 per cent). In this case, simple station corrections actually increase the *ak135* residual by 2 per cent. Fig. 10 shows the distribution of *ak135* and ETT residuals for 82 000 *S* and *Sn* arrivals as a function of epicentral distance. The gap in residuals near 83° is due to the difficulty in distinguishing *S* arrivals from *SKS* arrivals. As was found for *P* phases there is a greater density of ETT residuals around zero than *ak135* residuals across the whole range of epicentral distances. Engdahl *et al.* (1998) attribute the unusually large number of negative

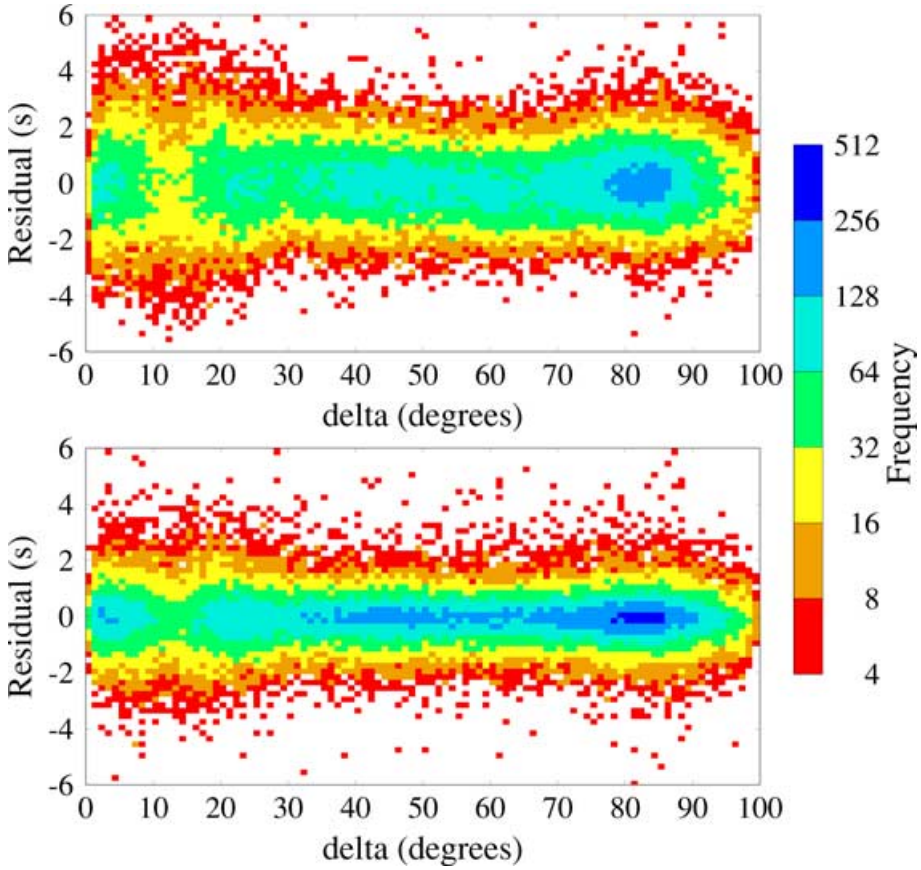

Figure 9. The distribution of *ak135* (above) and ETT (below) residuals as a function of epicentral distance for *P* and *Pn* arrivals. The use of ETTs in place of *ak135* reduces the spread of the residuals by over 30 per cent at all distances.

ak135 residuals between 5° and 10° to the presence of many paths from oceanic events to continental stations. Note that the feature is absent for the density plot of ETT residuals, indicating that this source of heterogeneity signal has been successfully reproduced in the ETTs. The shape of the density distribution is also much more symmetrical.

The *ak135* and ETT residual distributions for *PKPdf* , *PKPbc* and *PKPab* are shown in Figs 8(i)–(n). The use of simple corrections was effective in reducing the *ak135* residual spread by between 18 and 22 per cent for these phases. This is probably because the *PKP* phases travel steeply through most of the mantle and therefore have large ellipticity corrections. Using ETTs reduces the residual spread by a further 10–18 per cent.

Table 1 shows a summary of the results for all the phases discussed above and five others. In every case the ETT residuals have a smaller spread than those of *ak135* with simple corrections applied. However, the differences become smaller for the less frequent phases. This is likely to be due to the low number of observations available in these cases.

6.2 Confidence intervals for estimated traveltimes

As mentioned in Section 4, a confidence interval on ETTs can be estimated using the theory of TPS interpolation (Wahba 1990). These bounds may be useful in a variety of applications requiring traveltime estimation, for example as weights in hypocentre location or as a 'forward modelling' error in seismic tomography.

Here we test the accuracy of the estimated confidence intervals for a large number of *P* and *S* arrivals. We do this by comparing an 'observed' to a 'predicted' error histogram, for each phase. The observed error histogram is calculated from empirical minus observed arrival times. The predicted error histogram is formed from the same number of randomly generated residuals, each one being a sample from the estimated error distribution for the corresponding arrival given by Wahba (1990).

Fig. 11 shows predicted and observed error histograms for the 99 252 *P* and 9583 *S* arrivals used in Section 6.1. The match between predicted and observed histograms is very good for *P* arrivals, indicating that the estimated error distributions are statistically accurate. The only difference lies in the tails, where the removal of outliers in step 3 (see Section 3) has led to slightly higher tails in the observed histogram. (Note that the error estimation would be improved if step 3 were omitted, and outliers were allowed to remain in the database, but of course this would decrease the accuracy of the traveltime estimation itself !) For *S* arrivals, the predicted error histogram is slightly narrower than the observed one, which is also likely to be due to the removal of outliers. Overall the fit between observed and estimated histograms is good, and we can have some confidence that the error estimates are reasonably accurate, in a statistical sense.

6.3 Computational costs

The most computationally expensive part of the ETT estimator is the calculation of the smoothing parameter using generalized cross

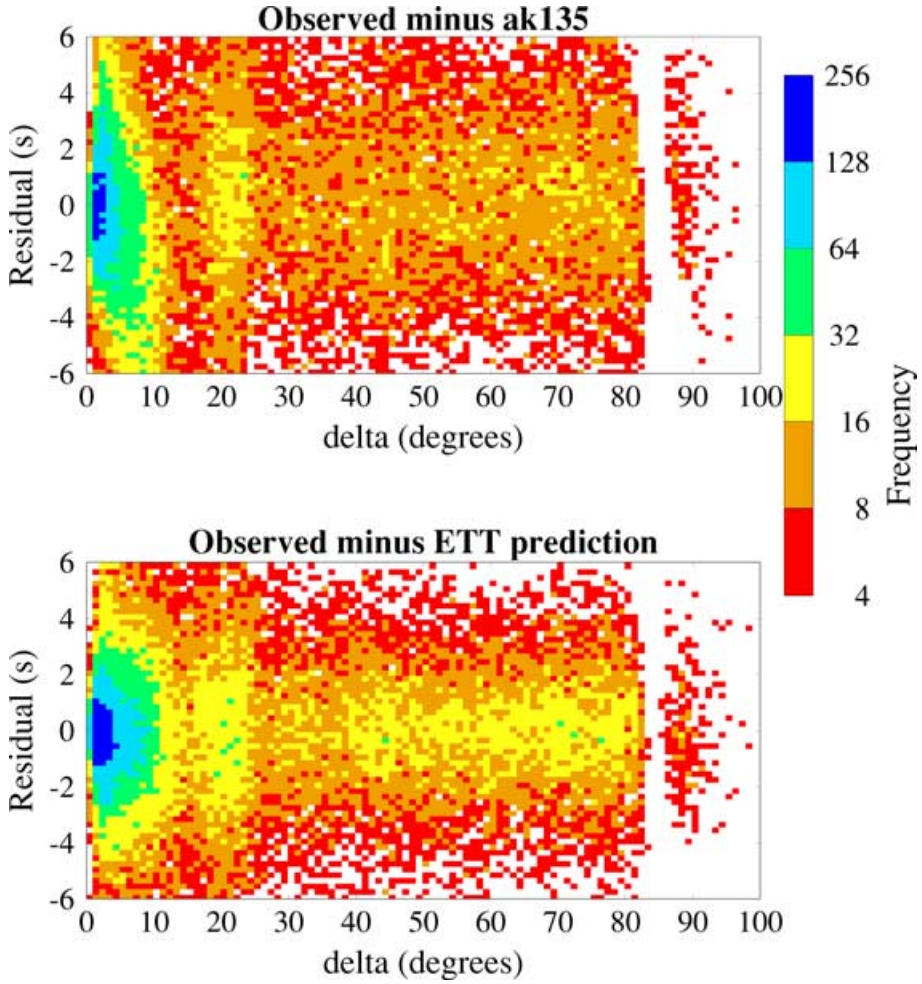

Figure 10. The distribution of *ak135* (above) and ETT (below) residuals as a function of epicentral distance for *S* and *Sn* arrivals. As for *P*-wave arrivals the spread is significantly reduced at all distances. In addition, the asymmetry at 5–10 degrees is removed.

validation. This is proportional to approximately the cube of the number of data used. As an example, the CPU time of a single traveltime estimation based on 400 nearby arrivals took approximately 6 s on a Compaq XP1000 (500 MHz, specfp95=52.2). On a more modern workstation this can be reduced to between 1 and 2 s. With 200 nearby arrivals the CPU time is reduced by almost an order of magnitude with comparatively little reduction on accuracy (see Section 3.3). A further point to note here is that the evaluation of eq. (1) is trivial once the data are selected and the coefficients of the polynomials determined. Therefore using it to give estimates of traveltimes in the 'vicinity' of an event, rather than just at the location of the event, can drastically reduce the overall computation cost. This issue is discussed further in Section 8.1 in the context of hypocentre estimation.

Table 1. Comparison of ETT prediction error with that of *ak135* and *ak135* with station corrections for a range of phases. For the *P* phase 100 000 arrivals were used. For the other phases, 10 000 arrivals were used.

Phase	Per cent for which	Spread of prediction errors (s)		
	ETTs were possible	*ak135*	*ak135* with simple corrections	ETTs
P	99.3	1.20	0.98	0.70
S	95.8	4.18	3.99	2.91
Pn	94.1	1.84	1.82	1.11
Sn	88.3	4.09	4.17	2.46
PKPdf	95.1	1.76	1.36	1.10
PKPbc	87.8	0.95	0.74	0.58
PKPab	77.1	1.45	1.17	1.05
PKiKP	56.2	2.06	1.44	1.40
pP	92.1	2.24	2.12	2.08
pwP	82.1	2.80	2.71	2.56
sP	83.5	3.31	3.31	3.25
PcP	84.6	2.65	2.48	2.42

7 COMPARISON WITH TOMOGRAPHIC EARTH MODELS

7.1 Reproducing the pattern of teleseismic residuals

In this section, we compare the accuracy of ETT estimates with those from two 3-D tomographic earth models, using earthquakes and nuclear explosions. We use the degree 12 spherical harmonic model S&P12/WM13 (Su *et al.* 1994), and the high-resolution block model of Gorbatov *et al.* (2001) (hereafter referred to as G01), which consists of $2^\circ \times 2^\circ$ surface blocks and 18 depth layers. Both models are taken as typical examples of tomographic models within each parameter class.

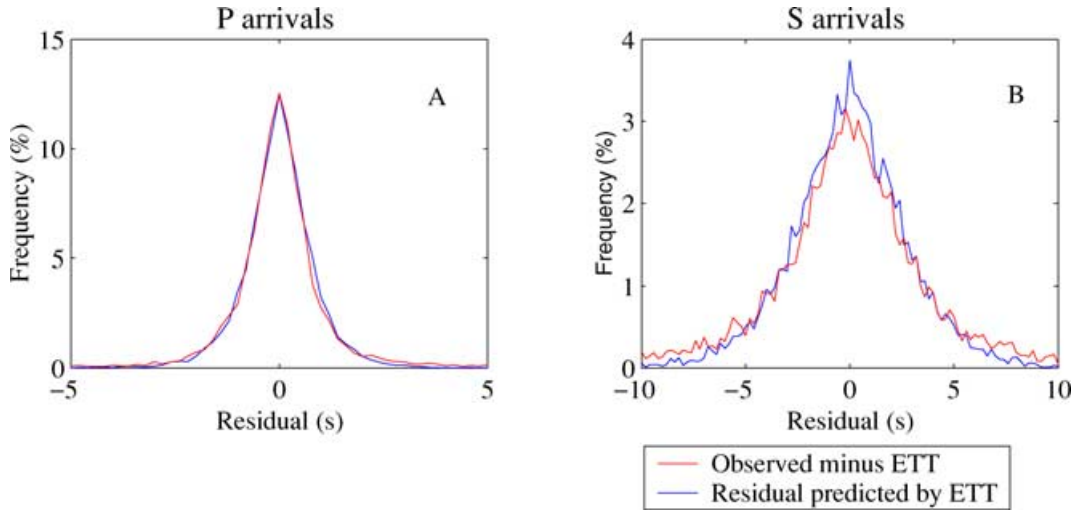

Figure 11. The observed traveltime minus the ETT for (a) 99 252 *P* arrivals and (b) 9583 *S* arrivals. The predicted residual distribution was calculated from the ETT error estimates. See text for further details. The predicted and observed residual distributions are in good agreement.

We compare traveltime estimates for *P* observations from 4870 southeast Asian events at the station MAT in Matsushiro, Japan. (Note that this station is co-located with MJAR, a primary IMS array.) Fig. 12(a) shows the observed *ak135* residuals plotted as a grey-scale/coloured dot at the EHB location. Note that the median residual has been subtracted from each observation. The seismicity south of 6°S lies in the roughly northward-dipping slab of the Indonesian subduction zone, while the seismicity north of 2°N lies in the Molucca subduction zone, which dips to the west.

Arrivals from shallow events in the Indonesian subduction zone follow fast paths at the eastern and western ends and slow paths in between. Fast paths are also followed from shallow events beneath the islands of Seram (4°S, 130°E) and Sulawesi (the large island to the west) as well as from events in a narrow band in the Molucca subduction zone. Almost all the arrivals from deep events follow slow paths. The autocorrelation function shows that the residuals are strongly (positively) correlated over distance scales shorter than 200 km and between 1100 and 1300 km, which we interpret as unmodelled heterogeneity signal. In Figs 12(b), (c) and (d) the median residual has also been removed.

Fig. 12(b) shows the traveltimes produced by S&P12/WM13 (hereafter referred to as SP12). SP12 correctly reproduces strong negative residuals in the southeast corner, weak positive residuals at the eastern end of the Indonesian subduction zone, and part of the positive residuals in areas in the Molucca subduction zone, but their predictions are poor throughout the remainder of the region shown. In particular, the positive residuals of deep events are not reproduced. These features are reasonable since SP12 is truncated at harmonic degree 12, corresponding to a minimum wavelength only slightly smaller than the dimensions of the region. The autocorrelation function shows that the SP12 residuals (i.e. the observed minus SP12 times) are correlated on the same scalelengths as the *ak135* residuals, indicating that SP12 resolves little of the correlated signal in the data. Consequently, the spread of residuals after correction for SP12 is no smaller than that of *ak135*.

Fig. 12(c) shows the traveltimes predicted by G01. The sign of the residuals is correctly reproduced throughout most of this region; however, their amplitude is underestimated. In particular, the large, correlated *ak135* residuals around Sulawesi, Seram and the Molucca subduction zone are poorly reproduced. This may be due to the damping applied in the determination of the model, or even to the spatial averaging resulting from the construction of summary rays. The autocorrelation function shows clear correlation in the G01 residuals at distances shorter than 150 km. This correlation is significantly reduced compared with the corresponding autocorrelation peak in Fig. 12(a) for the raw residuals. That is in turn primarily caused by a lower level of correlation beyond a scale of 100 km, which is close to the resolution limit of the G01 model imposed by the finite size of summary ray regions and inversion cells. The other two main features of the autocorrelation function for raw residuals (Fig. 12a), a minimum at 900 km and a maximum centred on 1200 km scales, are also clearly reduced to amplitudes that are marginally distinguishable from zero. The degree to which these features remain is an indirect indication of damping in the G01 model. The change in autocovariance at about the 1000 km scale of 0.09 s^2 in the raw residuals is reduced to about 0.025 s^2 in the G01 residuals, indicating that the G01 model explains a significant portion of the traveltime signal at that scale. The use of G01 instead of *ak135* is marginally successful in predicting the observed traveltime residuals, reducing the spread from 0.92 s to 0.88 s.

The corresponding results for ETTs are shown in Fig. 12(d). Here all the main features present in the *ak135* residual pattern are reproduced. In particular, the size, shape and amplitude of the negative residuals from events in Sulawesi, Seram and shallow events in the Indonesian and Molucca subduction zones are well reproduced, as are the positive residuals from events in the Molucca subduction zone and from deep events throughout the region. The ETTs are also able to resolve smaller-scale features, unlike models SP12 and G01. The autocorrelation function shows that the ETT residuals are uncorrelated over all distance scales. As a result, they have a significantly reduced residual spread of 0.75 s.

From these experiments we conclude that ETTs predict traveltimes for events in this region to the MAT station significantly better than the spherical harmonic and block models. Unlike the parametrized earth models, ETTs reproduce all of the spatially correlated signal in the pattern of residuals. These results suggest that current tomographic models leave a significant part of the heterogeneity of the Earth unmodelled, primarily at small scales (<200 km).

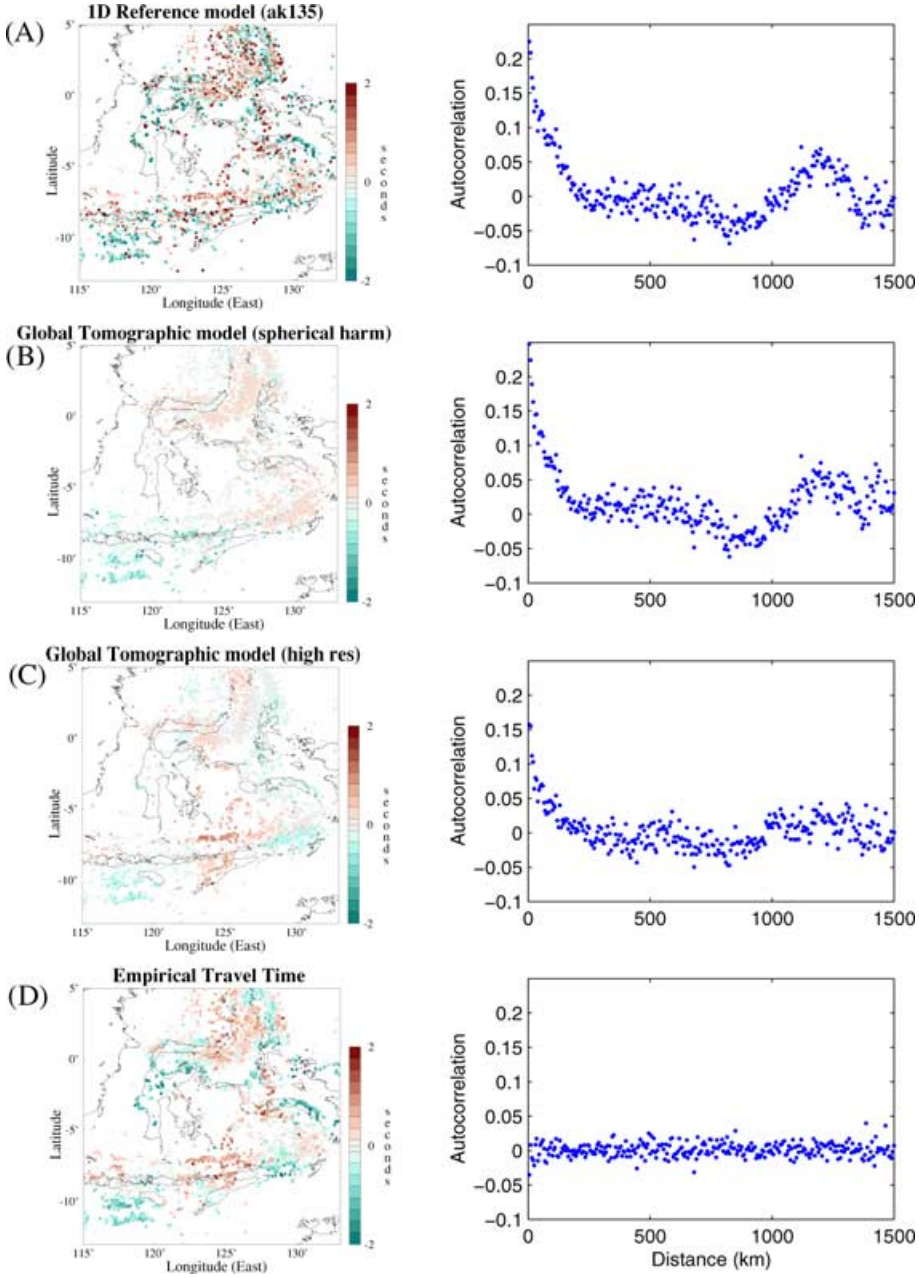

Figure 12. 4870 *P* observations made at MAT. In each case, the median observed (a) or predicted (b, c, d) residual has been subtracted from each value. On the left, (a) the observed *ak135* residual, (b) the S&P12/WM13 time minus *ak135,* (c) the G01 time minus *ak135,* (d) the ETT minus *ak135*. On the right, the autocorrelation functions for the observed traveltimes minus those predicted by (a) *ak135,* (b) S&P12/WM13, (c) G01 and (d) ETTs.

7.2 The influence of the reference model

Another issue for the proposed traveltime estimator is the influence of the background 1-D earth model. Recall that the interpolation and smoothing techniques are applied to traveltime residuals with respect to the 1-D reference model, *ak135*. To test the influence of the choice of reference model we repeated the calculation for the observations shown in Fig. 12, but this time applied the method directly to traveltimes rather than residuals. The result was that the average accuracy of the ETTs was reduced by 0.03 s (equivalent to less than 4 per cent of the average *ak135* residual). This suggests that, while a reference model can be beneficial in improving the accuracy of the estimated traveltime, the results are not significantly biased by the choice. We note here that just as in many other studies there is no particular restriction on the nature of the reference model, and so a laterally varying one could equally well be used.

7.3 ETTs for nuclear and chemical explosions

In the previous section, the ETTs from earthquakes were compared with the traveltimes determined from the EHB catalogue, using estimated hypocentres. As discussed above, the influence of the random component of hypocentre errors is minimized by the smoothing processes. However, high-precision ground truth events provide a useful cross-check on the accuracy of the traveltime estimator. Here we use

Table 2. Comparison of *ak135,* BDP98 and ETT traveltime prediction errors based on 25 teleseismically recorded explosions. The explosions are those used by Smith & Ekstrom (1996). Elevation, ellipticity and static station corrections were used for the BDP98 traveltimes.

Id	Latitude	Longitude	No. of	Median residual (s)			Residual spread (s)		
	(°)	(°)	arrivals	*ak135*	BDP98	ETT	*ak135*	BDP98	ETT
A	37.197	−74.352	60	1.21	0.19	**0.44**	1.15	0.86	**1.03**
B	37.130	−116.064	127	0.06	−1.10	**− 0.30**	1.04	0.87	**0.43**
C	38.634	−116.215	146	0.24	−0.92	**− 0.03**	0.86	0.68	**0.72**
D	37.295	−116.456	165	−0.11	−1.33	**− 0.30**	1.25	0.73	**0.43**
E	39.406	−107.948	49	0.07	−1.08	**− 0.08**	1.08	0.67	**0.78**
F	49.924	78.956	189	−1.11	0.14	**− 0.66**	1.01	0.71	**0.37**
G	49.769	78.034	210	−0.70	−0.87	**0.07**	0.96	0.60	**0.26**
H	51.472	179.107	342	−2.83	−2.58	**− 2.02**	1.63	1.43	**1.34**
I	49.765	78.059	91	−0.80	−0.81	**0.26**	0.83	0.37	**0.33**
J	49.927	78.758	242	−1.02	−1.02	**− 0.68**	0.84	0.58	**0.38**
K	39.793	−108.366	74	−0.30	−1.20	**− 0.12**	0.78	0.50	**0.69**
L	50.956	110.983	81	−0.10	−0.03	**0.21**	0.76	0.52	**0.48**
M	51.362	53.306	120	−1.27	−1.94	**− 0.14**	0.74	0.70	**0.22**
N	51.367	53.327	119	−1.22	−1.92	**− 0.17**	0.78	0.65	**0.15**
O	51.380	53.340	113	−1.21	−1.89	**− 0.06**	0.77	0.70	**0.21**
P	46.783	48.315	61	−1.30	−2.94	**− 0.20**	0.61	0.52	**0.12**
Q	46.788	48.297	60	−1.23	−2.96	**− 0.15**	0.60	0.50	**0.22**
R	46.767	48.311	60	−1.44	−2.97	**− 0.23**	0.56	0.41	**0.16**
S	46.749	48.303	61	−1.50	−3.08	**− 0.29**	0.53	0.34	**0.19**
T	46.754	48.289	64	−1.33	−2.88	**− 0.13**	0.56	0.37	**0.23**
U	46.766	48.274	66	−1.49	−2.98	**− 0.28**	0.64	0.59	**0.20**
V	51.358	53.319	132	−1.20	−1.99	**− 0.12**	0.78	0.61	**0.24**
W	51.390	53.351	136	−1.25	−1.97	**− 0.18**	0.75	0.65	**0.29**
X	51.371	53.337	130	−1.31	−2.03	**− 0.18**	0.77	0.61	**0.26**
Y	37.252	−116.377	239	0.16	−1.26	**− 0.04**	0.77	0.62	**0.32**
Median over all explosions (s)				−1.20	−1.89	**− 0.15**	0.77	0.61	**0.29**

P arrivals from the 25 nuclear and chemical explosions studied by Smith & Ekstrom (1996). (See Table 2 for a summary.) Residuals are calculated for *ak135*, a high-resolution 3-D block model, BDP98, (Boschi & Dziewonski 1999), and ETTs. In all cases residuals are determined using ground truth hypocentral coordinates. The traveltimes from the model BDP98 also include ellipticity and station elevation corrections taken from the EHB catalogue, and a correction for the local crustal structure beneath the station based on the CRUST5.1 model of Mooney *et al.* (1998).

For the ETT calculations we use the EHB catalogue as the reference database (together with its erroneous hypocentres), but combine this with GT0–GT5 ground truth events from the Prototype International Data Centre. Only ground truth events that also appear in the EHB catalogue are included. For ground truth events, traveltimes were determined from EHB arrival times and ground truth hypocentres. As previously, to avoid biasing ETTs towards the 'correct' values, the traveltime corresponding to each explosion was omitted from the database when predicting itself.

In total, ETTs were possible for 3157 of the 3170 *P* arrivals. A summary of results appear in Table 2 and Fig. 13. Events A–K were used as test events for the construction of *ak135*, and so their median *ak135* traveltime residuals are significantly lower than those of events L–Y. Overall, the median of the ETT residual is more than a second smaller than those of *ak135* and BDP98, while the median spread of residuals is about 62 per cent smaller for ETT compared with *ak135* and 52 per cent compared with BDP98. Comparing event by event, we notice that the spread of the ETT residuals is smaller than those of *ak135* for all the events used and less than the BDP98 residuals for all but four of the events. In general, accuracy appears to correlate positively with the local density of ground truth

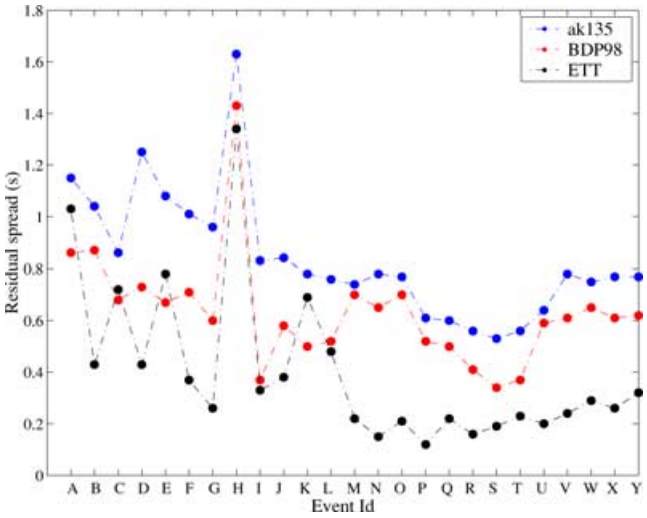

Figure 13. The spread in prediction errors for the observed traveltimes of the 25 explosions shown in Table 2. ETT residuals have the smallest spread for 21 events. The median, over all these events, of the ETT spread is less than half of the BDP98 spread.

events. This is to be expected because a high density of accurate database observations in the vicinity of a test point will allow the interpolation and smoothing procedures to work effectively.

In all of these examples the ETT estimator clearly gives more accurate results than either the spherical harmonic or block model. It can also resolve much smaller-scale correlations in the pattern of residuals than the longer-wavelength parametrized models. The

Table 3. Epicentre mislocations for the 25 explosions used by Smith & Ekstrom (1996) when SP12, EHCs and ETTs are used for the calculation of predicted traveltimes.

Id	Year	Month	Day	Time	Latitude (°)	Longitude (°)	Mislocations (km) SP12	EHC	ETT
A	1965	7	15	14:16:08.10	37.197	−74.352	10.3	11.20	**9.43**
B	1967	5	20	15:00:00.20	37.130	−116.064	4.4	7.61	**5.69**
C	1968	1	19	18:15:00.10	38.634	−116.215	5.4	5.81	**7.54**
D	1968	4	26	15:00:00.10	37.295	−116.456	3.1	4.91	**1.91**
E	1969	9	10	21:00:00.10	39.406	−107.948	9.3	4.91	**9.07**
F	1969	11	30	3:32:59.70	49.924	78.956	6.2	7.28	**6.73**
G	1971	4	25	3:32:59.90	49.769	78.034	5.4	9.82	**1.73**
H	1971	11	6	22:00:00.06	51.472	179.107	6.2	19.27	**10.49**
I	1972	8	16	3:16:59.80	49.765	78.059	6.2	10.98	**1.52**
J	1972	11	2	1:27:00.20	49.927	78.758	6.2	10.98	**3.74**
K	1973	5	17	16:00:00.00	39.793	−108.366	7.3	16.58	**11.45**
L	1977	8	10	22:00:00.10	50.956	110.983	6.2	6.94	**7.89**
M	1983	7	10	3:59:59.99	51.362	53.306	13.6	——	**1.34**
N	1983	7	10	4:04:59.94	51.367	53.327	6.2	9.57	**2.50**
O	1983	7	10	4:09:59.85	51.380	53.340	0.0	8.50	**2.38**
P	1983	9	24	5:00:00.03	46.783	48.315	13.6	——	**2.47**
Q	1983	9	24	5:05:00.03	46.788	48.297	7.0	3.11	**0.40**
R	1983	9	24	5:10:00.08	46.767	48.311	5.4	2.20	**1.25**
S	1983	9	24	5:15:00.14	46.749	48.303	8.5	4.39	**7.22**
T	1983	9	24	5:19:59.93	46.754	48.289	7.6	4.91	**1.69**
U	1983	9	24	5:25:00.00	46.766	48.274	8.5	3.11	**6.25**
V	1984	7	21	2:59:59.81	51.358	53.319	6.6	6.21	**1.42**
W	1984	7	21	3:04:59.71	51.390	53.351	2.2	7.28	**2.02**
X	1984	7	21	3:09:59.84	51.371	53.337	4.4	6.21	**1.31**
Y	1988	7	7	15:05:30.07	37.252	−116.377	3.8	9.05	**5.47**
Root Mean Square Mislocation (km)							7.21	8.90	**5.60**
Median Mislocation (km)							6.20	6.94	**2.50**

fact that residuals from the empirical procedure are uncorrelated over all distance scales suggests that no coherent effects of seismic heterogeneity remains unaccounted for.

8 HYPOCENTRE DETERMINATION USING ETTS

In the context of the Comprehensive Test Ban Treaty, and more generally, the teleseismic location of nuclear explosions has received considerable attention recently (e.g. Antolik *et al.* 2001; Piromallo & Morelli 2001; Chen & Willeman 2001). Although 1-D traveltime tables combined with simple station corrections are commonly employed for hypocentre location (e.g. Engdahl *et al.* 1998), it has long been known that the failure to take account of the Earth's lateral heterogeneity is a significant factor in limiting improvements. Smith & Ekstrom (1996) relocated a set of 25 nuclear and chemical explosions using a 3-D laterally heterogeneous earth model together with crustal and station corrections and observed a reduction in the rms mislocation of up to 40 per cent. Piromallo & Morelli (2001) used the same set of 25 explosions to show that improvements of up to 17 per cent can be achieved when a form of source-specific station correction called empirical heterogeneity corrections (EHC) is applied.

The success of the empirical approach in estimating traveltimes of ground truth events lends considerable encouragement for their application to hypocentre location. Here we examine the usefulness of ETTs in location by performing similar experiments to Smith & Ekstrom (1996), Antolik *et al.* (2001) and Piromallo & Morelli (2001) in relocating 25 explosions (largely from the Nevada and Semiplatinsk test sites) and comparing results. This is done first using all available *P* arrivals, and then with a reduced set of traveltimes, as was done by Antolik *et al.* (2001).

8.1 Relocation of chemical and nuclear explosions

A minor difference between the relocations performed here and those of Smith & Ekstrom (1996) and Antolik *et al.* (2001) is that we use a robust grid search location algorithm rather than an iterative linearized method. Another difference is that in the earlier studies the 3-D velocity model was used to produce a single 'region to station' correction for each ray path. This correction was held fixed while the hypocentre was iteratively updated, and hence repeated ray tracing through 3-D earth models was avoided. This is a reasonable approximation when the length scale of velocity variation is much larger than perturbations in the hypocentres. For the ETT estimator it is straightforward to evaluate the traveltime estimation through eq. (1) and hence take account of the changes in traveltimes caused by movements in the hypocentres. However, it would become computationally burdensome to repeat the four steps (Section 3) for every station and every point in a dense grid search. Therefore we introduce a simplifying approximation. To explain this, we first note that, in calculating a single ETT between a particular source, $\mathbf{x}_s$, and receiver, $\mathbf{x}_r$, we actually obtain a complete 3-D function of traveltimes centred on $\mathbf{x}_s$. Specifically, when we solve for the coefficients λ_n and a_i in eq. (1) (see Appendix) we obtain $s(\mathbf{x}; \mathbf{x}_s, \mathbf{x}_r, N)$. This is the ETT 'surface', evaluated at $\mathbf{x}$, for a receiver at $\mathbf{x}_r$, using the N nearest database events to $\mathbf{x}_s$. To date we have only used the single value $s(\mathbf{x}_s; \mathbf{x}_s, \mathbf{x}_r, N)$, but in principle we could equally well use it to give ETTs from the location $(\mathbf{x}_s + \delta\,\mathbf{x})$ to the receiver $\mathbf{x}_r$. In this case no recalculating of the coefficients is required and we simply need

Table 4. ETT epicentre mislocations for the 25 explosions used by Smith & Ekstrom (1996) when only a limited number of observations are used.

Id	Latitude	Longitude	8 observations		30 observations	
	(°)	(°)	Median mislocation (km)	per cent within 17.84 km	Median mislocation (km)	per cent within 17.84 km
A	37.197	−74.352	18.85	42	11.96	74
B	37.130	−116.064	12.06	63	6.68	84
C	38.634	−116.215	20.66	41	9.15	88
D	37.295	−116.456	16.81	52	7.00	87
E	39.406	−107.948	25.69	29	11.44	84
F	49.924	78.956	9.40	71	5.99	90
G	49.769	78.034	8.63	70	4.75	90
H	51.472	179.107	23.00	35	17.59	52
I	49.765	78.059	7.28	74	4.92	81
J	49.927	78.758	9.22	75	7.02	91
K	39.793	−108.366	35.53	30	16.83	53
L	50.956	110.983	16.08	57	8.22	100
M	51.362	53.306	6.06	87	2.64	100
N	51.367	53.327	5.24	92	2.92	100
O	51.380	53.340	5.74	87	2.91	100
P	46.783	48.315	4.98	97	3.47	100
Q	46.788	48.297	3.76	97	1.58	100
R	46.767	48.311	5.51	85	2.42	100
S	46.749	48.303	5.83	99	4.80	100
T	46.754	48.289	4.32	97	2.02	100
U	46.766	48.274	10.26	71	3.59	97
V	51.358	53.319	5.78	83	3.70	98
W	51.390	53.351	6.59	90	4.42	100
X	51.371	53.337	7.20	82	3.56	100
Y	37.252	−116.377	10.64	70	6.27	96
Median			8.63	74	4.8	97
Mean			11.40	70.1	6.23	91.0

to evaluate the traveltime approximation for each source location, which is trivial. Note that the approximation being used here is

$$s(\mathbf{x}_s + \delta\mathbf{x}; \mathbf{x}_s + \delta\mathbf{x}, \mathbf{x}_r, N) \approx s(\mathbf{x}_s + \delta\mathbf{x}; \mathbf{x}_s, \mathbf{x}_r, N). \quad (8)$$

Note also that this would be exact if the N database events closest to $\mathbf{x}_s + \delta\,\mathbf{x}$ and $\mathbf{x}_s$ were the same. For $\delta x \sim$tens of km, this is often likely to be the case. To proceed, we use eq. (8) in all relocation experiments, while removing the arrivals of the event being relocated from the database in each case.

The results of relocations are shown in Table 3, along with a summary from Smith & Ekstrom (1996) and Piromallo & Morelli (2001). The maximum epicentral mislocation of ground truth events found with ETTs is 11.5 km, which is within the allowable tolerance specified by the Comprehensive Nuclear Test Ban Treaty (1000 km^2). With ETTs the rms and median mislocations are 5.6 and 2.5 km, which represent a reduction of 22 and 60 per cent compared with model SP12 with azimuthally varying station corrections, and a reduction of 37 and 64 per cent compared with empirical heterogeneity corrections (EHCs). The ETT procedure also results in more accurate locations for 18 of the 25 explosions compared to SP12 with corrections applied, and 20 out of 25 compared to locations using EHCs.

8.2 Relocation with sparse data sets

In Section 8.1, nuclear tests, with more than 50 *P*-wave arrivals, were relocated using all available data from stations at teleseismic distances. These events were observed at more stations than might be available for any new nuclear test, particularly given the currently limited number of primary and secondary IMS stations. Antolik *et al.* (2001) argued that it is important to know how location procedures are affected by having relatively few observations. To examine this we relocate each of the explosions in Table 3 100 times using only a limited number of randomly chosen arrivals. We follow Antolik *et al.* (2001) and use first eight and then 30 randomly chosen teleseismic *P* arrivals.

Antolik *et al.* (2001) found that with the 3-D tomographic model BDP98 (Boschi & Dziewonski 1999) and ground truth station corrections, 58.3 per cent of the relocations made using only eight arrivals, and 79.6 per cent of those made using only 30 arrivals were within 17.84 km of the true locations. They also found mean mislocations of 28.6 and 10.8 km using model BDP98 with eight and 30 arrivals, respectively. The results for three other global tomographic models, SP12, the block model of van der Hilst *et al.* (1997), and MK12 (Su *et al.* 1997), were slightly worse. Relocations using ETTs are summarized in Table 4. In this case 70.1 and 91.0 per cent of the relocations fall within the required 17.84 km radius from the true value, for eight and 30 arrivals, respectively. Average mislocations of 11.40 and 6.23 km were achieved for eight and 30 arrivals respectively using our ETT approach. Note that our results are not directly comparable with those of Antolik *et al.* (2001) because they also included five Chinese explosions whose data were not available for the present study. Nevertheless, the use of ETTs appears to give a clear improvement in location quality. Although our experiments are encouraging the data set is clearly limited. Improvements in location may also result from further refinements, for example by using an L1 misfit norm (as in Shearer 1997), or by incorporating regional *P*, *S* and *PKP* arrivals.

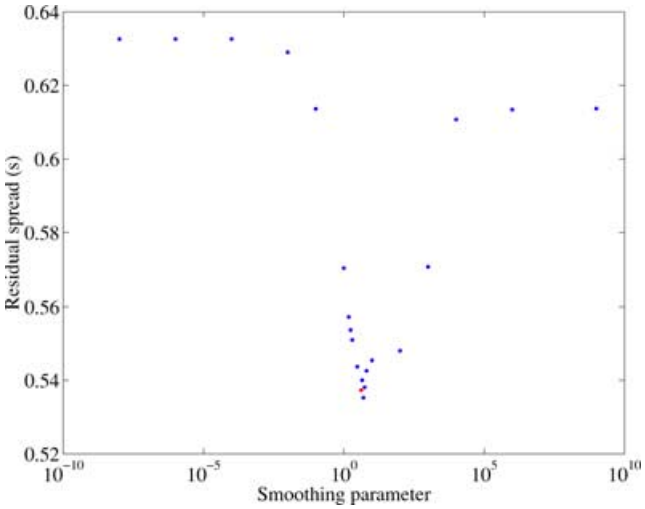

Figure 14. A test of how well GCV chooses the optimum smoothing parameter (see text for details). The smoothing parameter chosen by GCV is shown as a red dot. Clearly, the choice of smoothing parameter is vital and the GCV smoothing parameter is very close to the optimum.

9 CONCLUSIONS

A new empirical approach to traveltime estimation has been presented based on direct interpolation and smoothing of a large arrival time database. By avoiding parametrization of the Earth, the technique is able to reproduce the pattern of teleseismic residuals observed at seismic receivers over a wide range of distance scales. Furthermore, it can extract all spatially coherent signal in travel (arrival) times caused by, for example lateral heterogeneity, Earth's ellipticity of figure and topography. Spatially incoherent contributions to the residual pattern are regarded as noise and automatically eliminated. It is also efficient enough to be used 'on the fly', without extensive pre-processing of local, regional or global databases.

Extensive numerical tests show that the new approach performs well in reproducing traveltimes of globally distributed *P S*, *Pn*, *Sn*, *PKPdf* and depth phases as well as those from tightly constrained nuclear events and earthquakes. Relocation of ground truth and nuclear explosion data shows that the empirical approach produces smaller mislocations as compared with 1-D traveltime tables and 3-D earth models combined with azimuthally and spatially varying station corrections.

The influence of hypocentre errors in the arrival time database has also been investigated. Results show that the propagation of random hypocentre errors into traveltime estimations is significantly damped by the smoothing processes in the algorithm. Systematic hypocentre errors are spatially correlated and cannot be distinguished from heterogeneity signal. As with most other studies a judicious choice of database events is required to limit the influence of systematic hypocentre errors.

In addition to a traveltime estimation itself the procedure also provides formal measures of traveltime error. Tests show that these error estimates are statistically accurate, and hence may be useful in applications such as hypocentre location or seismic tomography.

The traveltime estimator may be used in any situation where a sufficient database of previously recorded events is available for a region. In cases where data are sparse or irregularly distributed, one could introduce synthetic traveltime data corresponding to a preferred earth model, and hence estimated times would smoothly change from being data-dependent to model-dependent within the volume of interest.

In addition to event relocation the ETT estimator may be useful in structural studies. The algorithm automatically filters spatially incoherent noise and could therefore be used to de-noise a data set prior to tomographic imaging. This, in turn, could reduce the need for damping in the tomographic inversion and lead to improved resolution.

ACKNOWLEDGMENTS

We are indebted to Brian Kennett and Nick Rawlinson for providing comments on an earlier draft of this paper. Alexei Gorbatov and Michael Antolik provided valuable feedback and also software used to calculate traveltimes through their 3-D seismic models. Two software packages for interpolating and smoothing thin-plate splines have been used in this study. One is by Hutchinson (2000) (available from http://cres.anu.edu.au/outputs/anusplin.html) and the other is by Bates *et al.* (1986) (available from http://franz.stat.wisc.edu/∼bates/ GCVPACK.html). Authors of all software are duly acknowledged. We are also grateful to Steve Billings who introduced us to the theory of thin-plate splines and GCV. Comments from two anonymous reviewers were helpful in improving an earlier draft of this manuscript.

REFERENCES

Antolik, M., Ekstrom, G. & Dziewonski, A.M., 2001. Global event location with full and sparse data sets using three-dimensional models of mantle *P*-wave velocity, *Pure appl. Geophys.,* **158,** 291–317.

Bates, D., Lindstrom, M., Wahba, G. & Yandell, B., 1986. *GCVPACK-Routines for Generalized Cross Validation*, UW-Madison Statistics Dept TR, 775(REV).

Bijwaard, H., Spakman, W. & Engdahl, E. R., 1998. Closing the gap between regional and global travel-time tomography, *J. geophys. Res.,* **103,** 30 055–30 078.

Billings, S.D., 1998. Geophysical aspects of soil mapping using airborne Gamma-Ray Spectrometry, *PhD thesis*, Fac. Agriculture, Univ of Sydney, Australia.

Billings, S.D., Sambridge, M.S. & Kennett, B.L.N., 1994. Errors in hypocenter location—picking, model, and magnitude dependence, *Bull. seism. Soc. Am.,* **84,** 1978–1990.

Boschi, L. & Dziewonski, A.M., 1999. High and low resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling, *J. geophys. Res.,* **104,** 25 567–25 594.

Buland, R., 1986. Uniform reduction error analysis, *Bull. seism. Am.,* **76,** 217–230.

Chen, Q.F., Willemann, R.J., 2001. Global test of seismic event locations using three-dimensional earth models, *Bull. seism. Soc. Am.,* **91** (6), 1704–1716.

Cleary, J. & Hales, A.L., 1966. An analysis of the traveltimes of *P* waves to North American stations, in the distance range 30° to 100°, *Bull. seism. Soc. Am.,* **56,** 467–489.

Constable, S.C., Parker, R.L. & Constable, C.G., 1987. Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics,* **52,** 289–300.

Craven, P. & Wahba, G., 1979. Smoothing noisy data with spline functions, *Numer. Math.,* **31,** 377–403.

Dewey, J., 1972. Seismicity and tectonics of Western Venezuela, *Bull. seism. Soc. Am.,* **62,** 1711–1751.

Douglas, A., 1967. Joint Epicentre Determination, *Nature,* **215,** 47–48.

Duchon, J., 1976. Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces, *R. A. I. R. O. Anal. Num.,* **10,** 5–12.

Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary Reference Earth Model, *Phys. Earth planet. Inter.,* **25,** 297–356.

Elden L., 1984. A note on the computation of the generalised cross-validation function for ill-conditioned least squares problems, *BIT,* **24,** 467–472.

Engdahl, E.R., van der Hilst, R.D. & Bullen, R.P., 1998. Global teleseismic earthquake relocation with improved traveltimes and procedures for depth determination, *Bull. seism. Soc. Am.,* **88,** 722–743.

Fitch, T.J., 1975. Compressional velocity in source regions of deep earthquakes: An application of the Master Event Technique, *Earth planet. Sci. Lett.,* **26,** 156–166.

Golub, G.H. & Van Loan, C.F., 1996. *Matrix Computations,* 3rd edn, Johns Hopkins Univ. Press, Baltimore, MD, USA.

Gorbatov, A., Fukoa, Y., Widiyantoro, S. & Gordeev, E., 2001. Seismic evidence for a mantle plume oceanwards of the Kamchatka-Aleutian trench junction, *Geophys. J. Int.,* **146,** 282–288.

Got, J.L., Frechet, J. & Klein, F.W., 1994. Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea, *J. geophys. Res.,* **99,** 15 375–15 386.

Gudmundsson, O., Davies, J.H. & Clayton, R.W., 1990. Stochastic analysis of global traveltime data: mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.,* **102,** 25–43.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. & Stahel, W.A., 1986. *Robust Statistics: The Approach based on Influence Functions,* John Wiley, New York.

Hutchinson, M.F., 1993. On thin-plate splines and kriging, in *Computing and Science in Statistics: Interface Foundation of North America,* pp. 55–62, eds Tarter, M.E. & Lock, M.E., University of California, Berkeley.

Hutchinson, M.F., 2000. *ANUSPLIN version 4.1 Users Guide,* CRES, Australian National University, Canberra.

Hutchinson, M.F. & de Hoog, F.R., 1985. Smoothing noisy data with spline functions, *Numerische Mathematik,* **47,** 99–106.

Hutchinson, M.F. & Gessler, P.E., 1994. Splines—more than just a smooth interpolator, *Geoderma,* **62,** 45–67.

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

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

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

Masters, G., Johnson, S., Laske, G. & Bolton, H., 1996. A shear-velocity model of the mantle, *Phil. Trans. R. Soc. Lond., A.,* **354,** 1385–1411.

Matheron, G., 1980. Splines and kriging; their formal equivalence, *Syracuse Uni. Geol. Contrib.,* **8,** 77–95.

Mitasova, H. & Mitas, L., 1993. Interpolation by regularized spline with tension: I. Theory and implementation, *Math. Geol.,* **6,** 641–655.

Mooney, W.D., Laske, G. & Masters, T.G., 1998. CRUST5.1: A Global Crustal Model at $5^\circ \times 5^\circ$, *J. geophys. Res.,* **103,** 727–747.

Myers, D.E., 1988. Interpolation with positive definite function, *Sci. Terre,* **28,** 251–265.

Myers, S.C. & Schultz, C.A., 2000. Improving sparse network location with Bayesian kriging and teleseismically constrained calibration events, *Bull. seism. Soc. Am.,* **90,** 199–211.

Nicholson, T.A., 2002. The development and application of new techniques for global travel times and source location, *PhD thesis*, Research School of Earth Sciences, Australian National University, Australia.

Nicholson, T., Sambridge, M. & Gudmundsson, O., 2002. Hypocentre location by pattern recognition, *J. geophys. Res.,* **107** (B6), doi:10.1029/2000JB000035.

Nychka, D., 1988. Bayesian confidence intervals for smoothing splines, *J. Am. Statist. Assoc.,* **83,** 1134–1143.

Piromallo, C. & Morelli, A., 1998. P-wave propagation heterogeneity and earthquake location in the Mediterranean Region, *Geophys. J. Int.,* **135,** 232–254.

Piromallo, C. & Morelli, A., 2001. Improving seismic event location: an alternative to three-dimensional structural models, *Pure appl. Geophys.,* **158,** 319–348.

Powell, M.J.D., 1992. The theory of radial basis function approximation in 1990, in *Advances in Numerical Analysis II: Wavelets, Subdivision Algorithms and Radial Functions,* pp. 105–210, Oxford University Press, Oxford.

Richards-Dinger, K.B. & Shearer, P.M., 2000. Earthquake locations in southern California obtained using source-specific station terms, *J. geophys. Res.,* **105,** 10 939–10 960.

Robertson, G.S. & Woodhouse, J.H., 1997. Comparison of *P*- and *S*-station corrections and their relationship to upper-mantle structure, *J. geophys. Res.,* **102,** 27 355–27 366.

Sanchez-Ortiz, G.I., Rueckert, D. & Burger, P., 1996. Motion and deformation of the heart using thin-plate splines and density and velocity encoded MR images, in *16th Leeds Ann. Stat. Res. Workshop: Image Fusion and Shape Variability Techniques,* pp. 71–78, eds Mardia, K.V., Gill, C.A. & Dryden, I.L.

Schultz, C.A., Myers, S.C., Hipp, J. & Young, C.J., 1998. Nonstationary Bayesian kriging: A predictive technique to generate spatial corrections for seismic detection, location, and identification, *Bull. seism. Soc. Am.,* **88,** 1275–1288.

Shearer, P.M., 1997. Improving local earthquake locations using the L1-norm and waveform cross-correlation: Application to the Whittier Narrows, California, aftershock sequence, *J. geophys. Res.,* **102,** 8269–8283.

Sidje, R.B. & Williams, A.B., 1997. *Fast Generalized Cross Validation*, Research report, Mathematics Department, University of Queensland, Australia.

Silverman, B.W., 1986. *Density Estimation for Statistics and Data Analysis,* Monographs on statistics and applied probability, Chapman and Hall, London.

Smith, G.P. & Ekstrom, G., 1996. Improving teleseismic event location using a three-dimensional earth model, *Bull. seism. Soc. Am.,* **86,** 788–796.

Souriau, A. & Woodhouse, J.H., 1985. A worldwide comparison of predicted S-wave delays from a three-dimensional upper mantle model with P-wave station corrections, *Phys. Earth planet. Int.,* **39,** 75–88.

Su, W.J. & Dziewonski, A.M., 1997. Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle, *Phys. Earth planet. Inter.,* **100,** 135–156.

Su, W.J., Woodward, R.L. & Dziewonski, A.M., 1994. Degree 12 model of shear velocity heterogeneity in the mantle, *J. geophys. Res.,* **99,** 6945–6980.

van der Hilst, R.D., Kennett, B.L.N. & Shibutani, T., 1998. Upper mantle structure beneath Australia from portable array deployments, in *Structure and Evolution of the Australian Continent,* Vol. 26, pp. 39–58, eds Braun, J., Dooley, J., Goleby, B., van der Hilst, R. & Klootwijk, C., AGU Geodynamics Series.

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

Wahba, G., 1983. Bayesian confidence intervals for the cross-validated smoothing spline, *J. R. Stat. Soc. Ser. B,* **45,** 133–150.

Wahba, G., 1990. Spline models for observational data, *CBMS-NSF Regional Conference Series in Applied Mathematics,* **59,** Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

Waldhauser, F. & Ellsworth, W.L., 2000. A double-difference location algorithm: Method and application to the northern Hayward fault, *Bull. seism. Soc. Am.,* **90,** 1353–1368.

Wolfe, C.J., Bjarnason, I.T., VanDecar, J.C. & Solomon, S.C., 1997. Seismic structure of the Iceland mantle plume, *Nature,* **385,** 245–247.

APPENDIX A: THIN-PLATE SPLINE INTERPOLATION AND GENERALIZED CROSS VALIDATION

A1 Thin-plate spline interpolation

The theoretical basis of thin-plate spline interpolation has been developed by many authors (Duchon 1976; Wahba 1990; Nychka

1988; Hutchinson & Gessler 1994) in conjunction with significant improvements in computational efficiency (Elden 1984; Hutchinson & de Hoog 1985; Sidje & Williams 1997). The thin-plate spline expansion has the form

$$s(\mathbf{x}) = p(\mathbf{x}) + \sum_{n=1}^{N} \lambda_n (\|\mathbf{x} - \mathbf{x}_n\|)^2 \ln(\|\mathbf{x} - \mathbf{x}_n\|), \tag{A1}$$

where $\|.\|$ is the Euclidean norm. The set of N weights $\{\lambda_n : n = 1, \ldots, N\}$ are defined at the N data points $\{\mathbf{x}_n : n = 1, \ldots, N\}$, and the polynomial $p(\mathbf{x})$ is given by

$$p(\mathbf{x}) = a_1 + a_2 x + a_3 y + a_4 z, \tag{A2}$$

where $\mathbf{x}$ is the position of any point in the medium and has coordinates (x, y, z). This polynomial models the scalar mean of the distribution and large-scale linear trends in the function across the entire data distribution while the weights model the smaller-scale signal. Therefore the second term in eq. (A1) models the small-scale correlations in the data values. This formulation, in which the influence of the nth weight actually increases with distance from the nth datum, seems counterintuitive but has been shown to be very useful in the interpolation of discretely sampled data (e.g. Powell 1992).

There are $N + 4$ unknowns in eq. (A1), namely the (λ_n)s $(n = 1, 2, \ldots, N)$ and the four coefficients in the polynomial p. Together the weights and polynomial coefficients are called the function coefficients. N constraints are obtained by requiring that $s(\mathbf{x})$ fit the data exactly at the N data points:

$$s(\mathbf{x}_n) = c_n, \quad \text{for} \quad n = 1, \ldots, N, \tag{A3}$$

where c_n is the observation at the nth data point. A unique interpolant is produced by requiring that the weights satisfy the following conditions:

$$\sum_{n=1}^{N} \lambda_n p_i(\mathbf{x}_n) = 0, \quad \text{forall} \quad i = 1, \ldots, 4, \tag{A4}$$

where $p_1(\mathbf{x}) = 1$, $p_2(\mathbf{x}) = x$, $p_3(\mathbf{x}) = y$, $p_4(\mathbf{x}) = z$ (see eq. A2).

Together (A3) and (A4) can be written as the linear system

$$\begin{bmatrix} \mathbf{A} & \mathbf{P} \\ \mathbf{P}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \bar{\lambda} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{0}, \end{bmatrix} \tag{A5}$$

where $A_{mn} = \|\mathbf{x}_m - \mathbf{x}_n\|^2 \ln(\|\mathbf{x}_m - \mathbf{x}_n\|)$, $P_{mk} = p_k(\mathbf{x}_m)$, $\bar{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)^{\mathrm{T}}$, $\mathbf{a} = (a_1, a_2, a_3, a_4)^{\mathrm{T}}$ and $\mathbf{C} = (c_1, c_2, \ldots, c_N)^{\mathrm{T}}$. In eq. (A5) $\mathbf{A}$ is an NxN matrix and $\mathbf{P}$ is an Nx4 matrix, while $\mathbf{0}$ represents a matrix of zeros. If $\bar{\lambda}^{\mathrm{T}}\mathbf{A}\bar{\lambda} > 0$ (or $\bar{\lambda}^{\mathrm{T}}\mathbf{A}\bar{\lambda} < 0$) for all $\bar{\lambda}$ such that $\mathbf{P}^{\mathrm{T}}\bar{\lambda} = \mathbf{0}$, then eq. (A5) can be transformed into a positive-definite form (e.g. Golub & Van Loan 1996). This is also a requirement for a valid semi-variogram or covariance function in kriging (Myers 1988). The positive-definite form can then be solved by LU decomposition (e.g. Golub & Van Loan 1996). In practice, however, eq. (A5) for the unsmoothed TPS is not ever solved.

Duchon (1976) showed that when we solve eq. (A5) to obtain $\bar{\lambda}$ and $\mathbf{a}$ and use them in eq. (A1) we obtain the thin-plate spline interpolant, s, which fits the data and in three dimensions minimizes

$$J(s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\left[\frac{\partial^2 s}{\partial x^2}\right]^2 + \left[\frac{\partial^2 s}{\partial y^2}\right]^2 + \left[\frac{\partial^2 s}{\partial z^2}\right]^2 + 2\left[\frac{\partial^2 s}{\partial x \partial y}\right]^2 + 2\left[\frac{\partial^2 s}{\partial x \partial z}\right]^2 + 2\left[\frac{\partial^2 s}{\partial y \partial z}\right]^2 \right) dx\, dy\, dz. \tag{A6}$$

Here J is a measure of curvature and therefore the solution, s, is the function with minimum curvature that fits the data exactly. The minimum-curvature property ensures that the interpolated function has a relatively smooth appearance even in the presence of noise. However, it can also result in large overshooting of data values when there are two or more very nearby noisy data points and the interpolant struggles to pass through all the data in a smooth manner. This problem will be alleviated when smoothing is applied.

A2 Smoothing the interpolant and generalized cross validation

Generally there are two sources of noise in the data: errors in the locations of the data points and errors in the value of the function at these points. The interpolated function is smoothed to reduce the effects of both sources of noise. The unsmoothed TPS interpolant minimizes the curvature function, $J(s)$, under the constraint that it must fit the data exactly. This constraint is relaxed in order to increase the smoothness. The smoothed interpolant for any given smoothing parameter (μ), minimizes H, which is a combination of data misfit and curvature:

$$H(s, \mu) = \sum_{i=1}^{N} [s(\mathbf{x}_i) - c_i]^2 + \mu J(s), \tag{A7}$$

where the parameter μ governs the trade-off between the goodness of fit and the smoothness.

We address the problem of determining the smoothed interpolating function, s, for a given smoothing parameter before dealing with the determination of the optimal smoothing parameter. One obvious way of at least approximately solving eq. (A7) is to substitute eq. (A1) into eq. (A7). This gives the discrete regularized least squares problem,

$$\min(\bar{\lambda}, \mathbf{a}) : (s - \mathbf{P}\mathbf{a} - \mathbf{A}\mu)^{\mathrm{T}}(s - \mathbf{P}\mathbf{a} - \mathbf{A}\mu) + \mu\bar{\lambda}^{\mathrm{T}}\Omega\bar{\lambda}, \tag{A8}$$

where $\mathbf{\Omega}$ is a positive-definite matrix. Indeed it can be shown that the solution to eq. (A8) is also the exact solution to eq. (A7). Moreover, the matrix $\mathbf{\Omega}$ is identical to the matrix $\mathbf{A}$ (Duchon 1976; Mitasova & Mitas 1993). Taking derivatives of eq. (A8) and rearranging terms, the solution of eq. (A8) satisfies the matrix system (e.g. Wahba 1990)

$$\begin{bmatrix} (\mu\mathbf{I} + \mathbf{A}) & \mathbf{P} \\ \mathbf{P}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \bar{\lambda} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{0} \end{bmatrix}. \tag{A9}$$

Thus apart from the addition of the positive constant μ to the diagonal, the matrix system is identical to that for exact interpolation given by eq. (A5). This system can be converted to a positive-definite form (Mitasova & Mitas 1993; Billings 1998) and therefore always has a solution.

Eq. (A9) is an expression for the coefficients of s for any given μ; however, the optimal μ must still be found. GCV is essentially a bootstrap method for determining the predicted error of the function. The GCV measure of data fit is calculated by removing each datum in turn, taking the difference at the node between an interpolant fitted to the remaining data points and the observed value, and then summing the squares of the residuals. Thus GCV minimizes the following least-squares quantity:

$$G(\mu) = \sum_{i=1}^{N} [s_i(\mathbf{x}_i) - c_i]^2, \tag{A10}$$

where s_i is the smoothed interpolant calculated from all the data points except the ith one, using a smoothed TPS interpolant. Note that the dependence of G on μ is via each 'sub'-interpolant $s_i(\mathbf{x}_i)$. Craven & Wahba (1979) showed that an implicit GCV measure can

be used in place of eq. (A10),

$$G(\mu) = \frac{(\mathbf{C} - \mathbf{AC})^{\mathrm{T}}(\mathbf{C} - \mathbf{AC})}{tr(\mathbf{I} - \mathbf{A})}, \tag{A11}$$

where $tr(\mathbf{I} - \mathbf{A})$ is the trace of the matrix $\mathbf{I} - \mathbf{A}$. G can then be minimized as a function of μ by a number of search algorithms. Here we use a simple grid search approach.

To test that the appropriate amount of smoothing is chosen by GCV, we randomly excluded half the data shown in Fig. 12(a) and predicted their residuals using ETT functions. The ETT functions were calculated from the other half of the data and a range of smoothing parameter values were used. Fig. 14 shows the resulting spread of ETT residuals as a function of smoothing parameter. Clearly, a good choice of smoothing parameter is vital to obtain an accurate interpolant. In this case, GCV choose a smoothing parameter slightly smaller than the optimum smoothing parameter. However, the GCV smoothing parameter is very close to the optimum. This result has been verified for many different cases.

Wahba (1983) and Silverman (1986) showed that the 95 per cent confidence interval, $CI_{.95}$, for the interpolated function is given by

$$CI_{.95} = \pm 1.96\sqrt{\mathbf{b}_x^{\mathrm{T}}\mathbf{V}\mathbf{b}_x}, \tag{A12}$$

where $b_{i,x} = (\| \mathbf{x} - \mathbf{x}_i \|)^2 \ln(\| \mathbf{x} - \mathbf{x}_i \|)$ and V is the error covariance matrix of the function coefficients. These confidence intervals must be interpreted 'across the function', as opposed to point-wise. If the interpolation and smoothing were repeated for the same f with new random noise from the same noise distribution then 95 per cent of the true values at the nodes would lie within the confidence intervals..

A3 Kriging

The formal equivalence of splines and kriging is well known (Matheron 1980; Wahba 1990); however, their formulations are quite different. In principle, if the data is well modelled as a spatial random field by a variogram and if the parameters in this variogram have been estimated successfully, then the kriging function is more accurate than a spline interpolation because the kriging function is the best linear predictor that can be obtained from the data. In practice, there is usually little difference between the fit of the kriging function and that of the thin-plate spline function. Hutchinson (1993) showed that the same is true of fits to simulated data, even when the same assumptions are used in the simulation as are used in the variogram of the kriging function. Therefore, as Billings (1998) and Hutchinson (1993) observed, since it is difficult to decide whether the kriging assumptions are indeed met, and since the use of a thin-plate spline requires significantly less calculation and fitting, there is much to be said for using thin-plate splines in place of kriging.