Geophysical Journal International

*Geophys. J. Int.* (2022) **231,** 1739–1749
Advance Access publication 2022 May 31
GJI Seismology

https://doi.org/10.1093/gji/ggac282

# Error estimates for seismic body wave delay times in the ISC-EHB Bulletin

Guust Nolet [1,2] and Suzan van der Lee[3]

[1] *Université de la Côte d'Azur/CNRS/OCA/IRD, Géoazur, Sophia Antipolis,* 06560*, France. E-mail: nolet@geoazur.unice.fr*
[2] *Department of Geosciences, Princeton University, Princeton, NJ* 08540*, USA*
[3] *Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL*60208*, USA*

Accepted 2022 July 21. Received 2022 July 17; in original form 2022 May 9

## SUMMARY

We analyse the consistency of the delay time data in the most recent version of the ISC-EHB bulletin published by the International Seismological Centre covering the years 1964–2018. Considering that the delays are influenced by the lateral heterogeneity in the Earth's mantle, we construct a tomographic matrix. We use singular value decomposition of the tomographic matrix for 19 707 dense clusters of earthquakes to compute objective estimates of the standard error from data that project into the null space and should be zero if there were no errors. Using a robust initial estimate of the standard deviation of the clustered delay times, we remove a small fraction of outliers before calculating the ultimate errors. We found that the errors depend on the type of body wave, depth of the earthquake (crust or mantle) and the number of decimals with which the arrival time was reported. Using these parameters, we distinguish 45 different classes of delay times for 11 different types of body waves. The errors of each class so divided generally follow a distribution that is approximately normal with a mean that ranges from 0.32 s for *PKPbc* waves from mantle earthquakes, to 2.82 s for *S* waves from shallow earthquakes bottoming in the upper mantle. The widths of the distributions of the errors themselves are small enough to serve in formal statistical quantification of the quality of fit in tomographic experiments.

**Key words:** Numerical modelling; Body waves; Seismic tomography.

**Key words:**; ISC-EHB Bulletin; Error analysis.

## 1 INTRODUCTION

Since 1964, the International Seismological Centre (ISC) has re-analysed worldwide seismic wave arrival time data to produce what is considered the most complete catalogue of earthquake parameter data—a task it inherited from the International Seismological Summary (ISS), a global catalogue started in 1918 by Milne (Adams *et al.* 1982).

As a spinoff of the epicentral determinations, ISC also publishes the residuals of the traveltimes (the misfits to the optimal hypocentre solution). Dziewonski *et al.* (1977) realized the unprecedented possibilities these data offered for seismic tomography and undertook the laborious task of retrieving residuals from the 7-track magnetic tapes (often decayed) used to store them at the time in a first attempt to produce a global 3-D image of the Earth's laterally heterogeneous internal structure. This was soon followed by many regional and global studies using the ISC residuals.

In an effort to improve the depth determination of the hypocentre locations, Engdahl *et al.* (1998) included teleseismic depth phases as well as core phases in the algorithm, to produce what is now known as the ISC-EHB Bulletin (hereafter simply referred to as 'the EHB'). The EHB constitutes a subset of the full ISC data set and includes only events that have been recorded in a minimum number of well-distributed stations. The EHB procedures have evolved somewhat over time (Storchak *et al.* 2015; Weston *et al.* 2018; Engdahl *et al.* 2020), and the full set of residuals since 1964, recently revised, is accessible from the ISC.

The residuals are with respect to the spherically symmetric isotropic Earth model AK135 (Kennett *et al.* 1995) and contain information on the deviations of AK135 with respect to the real laterally heterogeneous and anisotropic Earth. Like all experimental data the residuals contain errors, which fall into two classes: the observational errors, some of which, such as clock errors, may be systematic for some finite length of time (Röhm *et al.* 1999, 2000),

and theoretical errors due to imperfections of the hypocentre and origin time estimates. If used for seismic tomography there are additional theoretical errors associated with the inadequacy of a finite parametrization and the neglect of anisotropy to model the arrival times perfectly.

Formal statistical evaluations of tomographic images are difficult, if not impossible if the data sets and number of structural model parameters are very large. Even the simplest test, comparing the square of the misfit $\chi^2$ for different tomographic solutions, is easy to compute but useless unless one has estimates for the data errors that are sufficiently accurate. This paper presents our objective estimates of standard errors for all mantle and core phases reported in the EHB, with their uncertainty.

Morelli & Dziewonski (1987) were the first to estimate errors in the *P*, *PcP* and *PKP* delays as reported in the original ISC Bulletins, assuming that the ray paths between closely spaced sources on one end, and closely spaced stations on the other ('summary rays'), would experience the same delays caused by the Earth's heterogeneity so that any differences represent the data errors. However, data sparsity forced them to spread their net widely (cells of $5^\circ \times 5^\circ$) to ensure sufficient coverage. Velocity heterogeneity within the volume of the ray bundle did influence the scatter of the delay times, which leads to an overestimate of the errors (they found 1.4 s for *P*). In fact, when Gudmundsson *et al.* (1990) included the width of the summary ray as a factor in their estimation procedure, they found smaller errors, depending on the epicentral distance.

These early estimates did not take into account the improved depth localizations obtained with the EHB algorithm. In addition, the number of data have increased dramatically in the 30 yr since then, allowing for more phases to be analysed and allowing for a more detailed look at how errors depend on multiple factors such as hypocentres, number of decimals in the reported arrival times, depth of the earthquake and slowness of the ray paths.

We have accepted all reported arrival times for this study with the exception of those from 1808 seismic events marked as having a poor solution type (EHB variable isol set to XEQ). This data set constitutes by far the largest set available for global seismic (traveltime) tomography, and is growing at an ever more rapid pace owing to national station densifications and array installations on a global scale. The bulk of the arrival time data have been reported after the year 2000—less than 28 per cent of the *P*-wave data have been reported in the 36 yr before that epoch. Currently, some 1.5 million times are reported for teleseismic *P* waves *every year*, along with some 0.5 million regional phases (*Pn*), as well as a similar number of core phases. (Fig. 1). Note that *Pn* is not necessarily a headwave: EHB denotes with *Pn* any wave that bottoms in the uppermost mantle or an upgoing *P* wave from a source in the uppermost mantle. Excluding the 1808 events with 'poorly' constrained hypocentres, the total number of traveltimes for 11 types of mantle phases included over the 54 yr since 1964 is 427 17 476. In Table 1, we separate the counts over two epochs and the three assigned precisions of 0–2 decimals. Though shear waves are under-represented, the total of readings for *S* and *Sn* still sum up to 26 45 698 arrival times.

## 2 ERROR ESTIMATION

Errors in delay time data can be estimated objectively using the method developed by Voronin et al. (2014), which is based on singular value decomposition of the tomographic matrix $\boldsymbol{A}$ for the data that constitute a summary ray. This orthogonalizes a subset of transformed data to the range of $\boldsymbol{A}$—in other words it renders them insensitive to the heterogeneity along the ray paths, and the distribution of these transformed data gives an unbiased estimate for the probability density function (PDF) of the data errors.

Voronin et al. (2014) considered subsets of closely parallel rows from the large linear tomographic systems with model parameters in an *M*-dimensional vector $\boldsymbol{m}$ and an *N*-dimensional data vector $\boldsymbol{d}$, where $N \ll M$:

$$\boldsymbol{A}\boldsymbol{m} = \boldsymbol{d}\,. \tag{1}$$

The model $\boldsymbol{m}$ contains the velocity anomalies in the Earth's interior, and may also contain corrections to the origin time and hypocentre location. The data vector $\boldsymbol{d}$ contains the residuals, or *delays*, defined as the difference between the observed traveltime and the time predicted by a 1-D reference model.

We transform this system using an $N \times N$ orthonormal matrix $\boldsymbol{U}^T$. The columns of $\boldsymbol{U}$ are the eigenvectors of $\boldsymbol{A}\boldsymbol{A}^T$ with eigenvalues $\lambda_N^2 \geq \lambda_{N-1}^2 \geq \ldots \geq \lambda_1^2$:

$$\boldsymbol{U}^T\boldsymbol{A}\boldsymbol{m} = \boldsymbol{U}^T\boldsymbol{d} = \boldsymbol{\tau}\,, \tag{2}$$

where the *N* transformed data $\tau_i$ are distributed with a variance:

$$\sigma_{\tau_i}^2 = \sigma_m^2\lambda_i^2 + \sigma_e^2. \tag{3}$$

From eq. (3), we see that $\sigma_{\tau_i}^2 \to \sigma_e^2$ as $\lambda_i^2 \to 0$. In other words, zero (or negligible) eigenvalues are associated with transformed data $\tau_i$ that have a PDF with variance $\sigma_e^2$.

Nolet et al. (2019) realized that it is not necessary to cluster the rays into a narrow bundle, as long as there is sufficient overlap among the rays to yield a large number of negligibly small singular values $\lambda_i^2$. Nolet *et al.* (2019) and Voronin *et al.* (2014) combined ray paths only at the source ends, using several closely spaced hypocentres.

We isolate a subset of the $\tau_i$, associated with $N_0$ eigenvalues below $\lambda_0$, to estimate $\sigma_e$ from their observed distribution. For the estimation of $\sigma_e$ we face the question what level of eigenvalue $\lambda_0$ to consider as 'negligible'. We assume here that the system (1) has been scaled such that the model variance $\sigma_m^2$ and the data error variance $\sigma_e^2$ are of order 1, in our case using prior uncertainties of 1 per cent for *P*, 2 per cent for *S*-anomalies in $\boldsymbol{m}$ and 1 s for the delays in $\boldsymbol{d}$. Thus, the threshold $\lambda_0$ must be significantly below 1. But if we choose a value that is too low, we may end up with a population $N_0$ of eigenvalues that is too small for a well-constrained estimate of the standard deviation.

An example of the distribution of the transformed data, in order of decreasing eigenvalue, is shown in Fig. 2. As predicted by eq. (3), we see that the scatter decreases with decreasing $\lambda_i$ and stabilizes towards the right of the plot where $\lambda_i \to 0$.

Fig. 3 shows how the estimated $\sigma_e$ varies as a function of the threshold $\lambda_0$ for a number of randomly selected large event clusters from two epochs and precision classes. The estimates generally converge when $\lambda_0$ is less than about 0.03 ($\lambda_0^2 < 0.009$). This is the threshold that we have adopted in this study.

## 3 DATA CHARACTERISTICS

The total data set covers 185 296 hypocentres. An initial exploration of standard errors in *P* waves showed little geographical correlations between them, except that the regions with deep earthquakes generally seemed to have smaller errors.

Among the parameters that we expect to correlate with the standard error in the ISC data is the precision with which the data are reported to ISC, given by the variable iprec. The reported precision

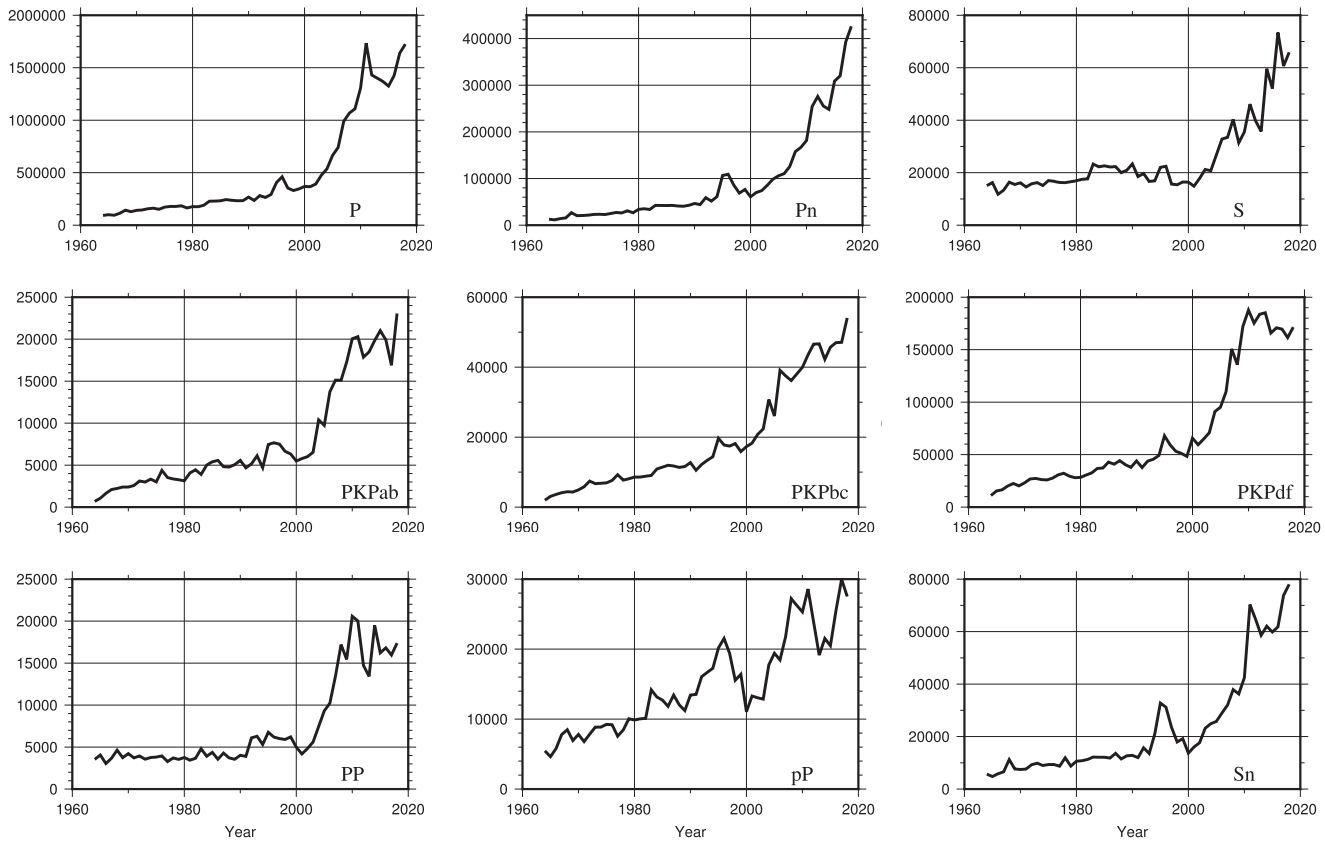


**Figure 1.** Number of traveltime readings per year included in the ISC-EHB bulletin, as a function of time, for the nine most abundant phases.

**Table 1.** Number of reported arrivals for two epochs 1964–1999 and 2000–2018, and three reported decimal precisions, 0 (1 s), 1 (0.1 s) and 2 (0.01 s).

| | 1964–1999 | | | 2000–2018 | | | |
|---|---|---|---|---|---|---|---|
| Phase | 1 s | 0.1 s | 0.01 s | 1 s | 0.1 s | 0.01 s | Total |
| *P* | 26 67 996 | 43 14 820 | 745 966 | 904 969 | 61 84 269 | 129 75 134 | 277 93 154 |
| *Pn* | 453 132 | 883 734 | 118 447 | 224 402 | 14 30 916 | 20 63 762 | 51 74 393 |
| *PKPab* | 56 946 | 84 372 | 8673 | 15 748 | 110 242 | 156 572 | 432 553 |
| *PKPbc* | 106 326 | 221 081 | 21 781 | 37 140 | 271 856 | 390 229 | 10 48 413 |
| *PKPdf* | 462 037 | 702 556 | 91 423 | 131 564 | 874 193 | 15 81 061 | 38 42 834 |
| *pP* | 195 479 | 177 559 | 39 211 | 26 220 | 160 691 | 216 131 | 815 291 |
| *pwP* | 82 738 | 71 920 | 11 855 | 8058 | 46 662 | 46 874 | 268 107 |
| *PP* | 106 923 | 43 518 | 5410 | 27 973 | 124 660 | 94 680 | 403 164 |
| *S* | 456 260 | 173 893 | 11 586 | 92 989 | 342 806 | 288 876 | 13 66 410 |
| *Sn* | 196 205 | 232 736 | 22 367 | 59 007 | 336 072 | 432 901 | 12 79 288 |
| *sP* | 91 877 | 70 052 | 11 668 | 13 938 | 70 804 | 35 530 | 293 869 |

in the arrival times is mostly one to two decimals, but a precision of 0 s—the lowest precision encountered—does occur, especially in the earlier years and for the later arriving phases. The reported iprec cannot always be trusted. ISC reports the deduced origin time with two decimals in the seconds. We noted, however, that about a quarter of the *P* arrivals that are reported with iprec=−1 (1 decimal), have an observed traveltime (obstt) with two decimals that are equal to 1.00 minus the two decimals in the origin time, indicating the arrival time was reported with an integer number for the seconds. The probability that this should occur by chance is of course only 1 per cent. We decided to correct iprec in these cases and set it to 0, thus assuming zero decimals in the reported precision. We have alerted ISC, which has contacted agencies where these anomalies appear to be generated and considers applying the same corrections for past data as we did.

Table 1 shows the reported precision for various phases considered in this study. From Fig. 4, one can clearly see that for all phases considered, the reported precision has increased substantially since the start of the catalogue. We suspect that the increase in reported precision is a consequence of the advent of array processing applied to a network's phase arrivals and of a decrease in clock errors after the widespread introduction of GPS timing in the 1990s, when clock errors were still a case of concern (Röhm *et al.* 1999, 2000).

Fig. 4 strongly suggests a major change around the year 2000: the number of readings with reported precision of 0.01 s rapidly climbs, indicating that stations with digital and array processing tools begin

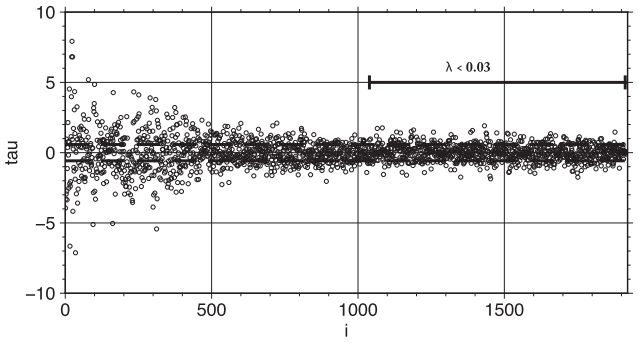


**Figure 2.** The set of 1947 transformed data $\tau_i$ for a cluster of mantle earthquakes in the Aleutians. There are 820 $\tau_i$ for which $\lambda_i < 0.03$, and their distribution allows us to estimate the error distribution of the delay times for this cluster. The dashed lines indicate the $\sigma_e$ estimate of 0.58 s from the distribution of this subset of $\tau_i$.

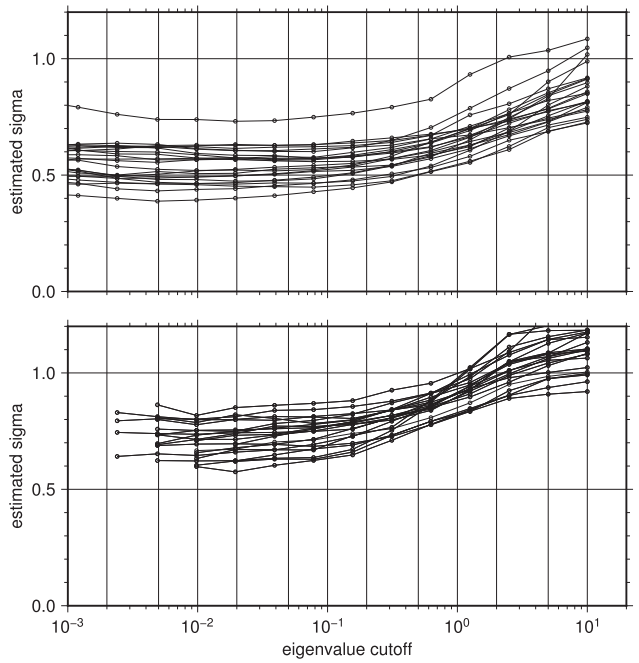


**Figure 3.** The estimated value of $\sigma_e$ as a function of the cut-off $\lambda_0$ for 22 event clusters of *P*-wave delays. The top frame shows the estimates for data reported with 2 decimal precision after 2000, the bottom one for 0 decimal precision reported between 1964 and 1999.

to dominate. For this reason we initially analysed the *P* data further into the two epochs: 1964–1999 and 2000–2018.

Most events in the catalogue have been determined with a large azimuth coverage of stations, which reduces the timing error due to mislocation of the events. For the reported *P* phases, the largest secondary azimuth gap (an azimuth range with no more than one station) is 296.9°, only 771 129 data (2.8 per cent) have a secondary azimuth >140°, 211 643 have >160° (0.8 per cent) and only 7620 or 0.03 per cent have a secondary azimuth that exceeds 180°. Some data may exceed the azimuth gap of 180° since the selection criteria were only applied early on in the analysis for each event (Engdahl, 2022, personal communication).

Given that azimuth coverage is generally satisfactory, we decided that it would not be productive to make yet another subdivision of the data, if only because it would be hard to get sufficiently large clusters with deficient azimuth coverage.

To organize the data into event clusters, we divide the Earth's surface into 10406 'patches' (Fig. S1, Supporting Information). The patches are 2 × 2° at the equator and every 2° latitude is divided in an integer number of patches such that the surface area is roughly the same (about 49 000 km$^2$) for every patch.

Not surprisingly, *most patches are empty*. Only 733 (or 7 per cent) of the patches host at least one earthquake deeper than 40 km. There are 2981 (29 per cent) patches with at least one shallow earthquake. Fig. S2 (Supporting Information) shows the distribution of the patches with sources that generate at least 3000 data. Fig. S3 (Supporting Information) does the same for *Pn*.

## 4 DATA PROCESSING AND MODELLING

To model the the delays reported by EHB, we allow for corrections to the hypocentre and the origin time in the model $\boldsymbol{m}$ (with prior uncertainties of 10 km and 1 s, respectively) and assume that the velocity anomalies are isotropic. To compute the matrix $\boldsymbol{A}$, the Earth's mantle was divided into six chunks of 128 × 128 voxels horizontally and 37 layers vertically, resulting in a 'cubic' Earth with 36 37 248 voxels as described by Charlety *et al*. (2013). This significant overparametrization is desirable so as to minimize the theoretical error introduced by limitations imposed by the finite size of voxels to model the 'true' delays. Our software did not allow to model effects of anisotropy, nor delays acquired by heterogeneity in the core. The bias introduced by this may lead to an overestimate of standard errors for *Pn*, *Sn* and the *PKP* phases. However, any mantle tomography that ignores mantle anisotropy for *Pn* or the (largely hemispheric) heterogeneity and anisotropy in the inner core will suffer the same theoretical errors and can therefore use the standard errors for *Pn*, *Sn* and *PKP* estimated in this study.

If there is more than one arrival within ±3 s of the observed time, it rejects those with an amplitude less than 2 per cent of the highest amplitude, before selecting the closest. Otherwise it selects the ray with the closest arrival time.

We then compute the matrix row for this ray, including columns associated with hypocentre and origin time corrections. Predicted arrival times are corrected for ellipticity, station elevation and crustal structure using CRUST2.0 (Bassin *et al.* 2000). Note that this differs from EHB, who apply a local 'patch correction' during the computation of the optimal source parameters, but who advise tomographers to undo that correction before inverting the data. We decided to apply the CRUST2.0 correction in a effort to at least partly remove scatter in the data caused by deviations from the crustal structure in AK135 from the crust beneath the stations within each event cluster.

We retain only patches generating clusters with sufficient ray path overlap such that $\sigma_e$ can be estimated from at least 100 negligible eigenvalues. Large clusters are split into files with about 2000 rows (some exploratory studies were initially done using larger matrices with up to 5000 rows, but this proved not really necessary and is very time-consuming since the CPU time increases as $N^3$). We start with *P* and *Pn*, which are far more abundant than the other phases.

### 4.1 Outliers

Statistical assessments of tomographic studies usually assume a Gaussian distribution, if only implicitly by solving the linear system using least-square algorithms such as LSQR. However, arrival time errors have large outliers, which are moreover biased towards

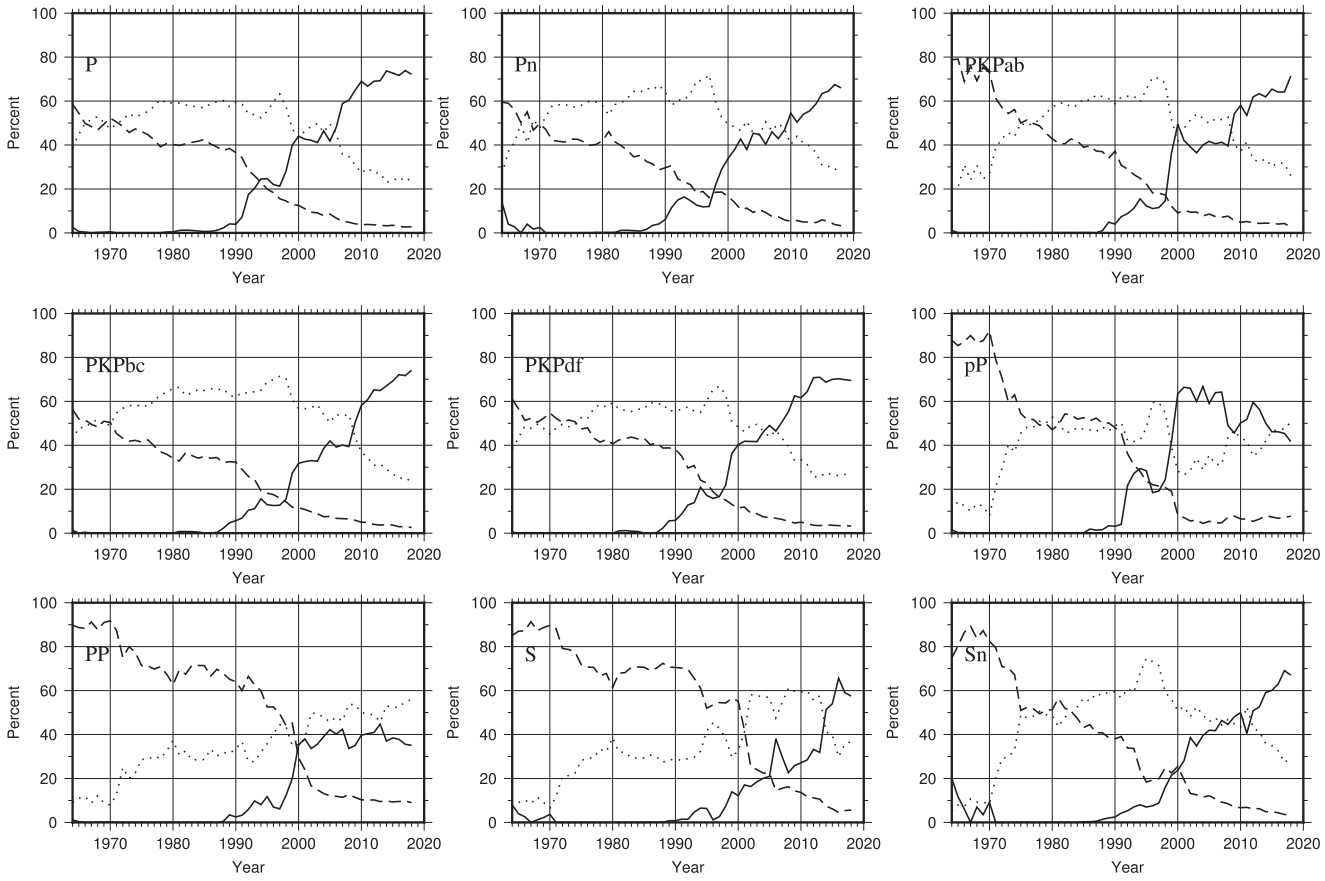


**Figure 4.** The percentage of accepted traveltimes reported with a precision of 0.01 s (solid). 0.1 s (dotted) and 1 s (dashed) as a function of time.

positive times, a fact already observed by Jeffreys (1936). Outliers can have a devastating effect on the least-squares solution.

Yang *et al.* (2019) discuss the removal of outliers in the context of machine learning and test a number of algorithms found in the literature. The best-performing algorithm in a large variety of cases, some with a large fraction of outliers, is a two-step procedure called 2T. It computes first an initial estimate $\sigma_{\text{init}}$ of the standard deviation using the full data set including outliers. It then discards any data outside of the interval $[-\gamma\sigma_{\text{init}} < d_i < \gamma\sigma_{\text{init}}]$ using a threshold $\gamma$, and recalculates $\sigma$ from the remaining data. Yang et al. deal directly with the raw data whereas in our case we deal with misfits to a linear system that changes after the removal of outliers and the corresponding rows in the matrix $\boldsymbol{A}$. But our algorithm is a straightforward adaptation of 2T. Schematically:

Cleaning step for data $\boldsymbol{d} \in \mathcal{R}^N$:

(i)Diagonalize the $N \times N$ matrix $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$.
(ii)Compute $\boldsymbol{\tau} = \boldsymbol{U}^T\boldsymbol{d}$ for the full data set $\boldsymbol{d}$, with $\boldsymbol{\tau} \in \mathcal{R}^N$.
(iii)Compute the standard deviation of $\tau_i$ belonging to the $N_0$ smallest eigenvalues. This is our initial estimate of the data error $\sigma_{\text{init}}$.
(iv)Regularize by keeping only the transformed data belonging to the $k$ largest eigenvalues: $\boldsymbol{U} \rightarrow \boldsymbol{U}_k$, and $\boldsymbol{\tau} \rightarrow \boldsymbol{\tau}_k \in \mathcal{R}^k$.
(v)Calculate residuals $\boldsymbol{r} = \boldsymbol{d} - \boldsymbol{U}_k\boldsymbol{\tau}_k$, with $\boldsymbol{r} \in \mathcal{R}^N$.

Thresholding step:

(i)Reject any data $d_i$ for which $|r_i| > \gamma\sigma_{\text{init}}$.
(ii)Recalculate $\boldsymbol{A}$, $\boldsymbol{U}$ and $\boldsymbol{\tau}$ for the thresholded data set.
(iii)Recalculate the standard deviation of the observed distribution of the $\tau_i$ for the $N_0$ smallest eigenvalues of the cleaned system. This is interpreted as a final estimate for the data error $\sigma_e$.

For the regularization in the cleaning step we truncate the matrix $\boldsymbol{U}$ at $k$ such that all $\lambda_i \geq \lambda_k \geq 1$. If our prior estimates of model and data variance are correct this means that we regularize by rejecting any (transformed) data with a signal-to-noise ratio smaller than 1 (Voronin et al., 2014). For the set of 'smallest' eigenvalues from which $\sigma_{\text{init}}$ is determined, we choose the $N_0$ eigenvalues $\lambda_i < 0.03$ (see Fig. 3).

Note that $t^2 \equiv r^2/\sigma^2_{\text{init}}$ follows a Chi-squared distribution that is close to lognormal and, for large $N_0$ approaches a normal distribution. We have only accepted clusters that yield $N_0 \geq 100$, which allows us to use classical estimates for statistical parameters that are based on a normal distribution. For a normal distribution, only 2.7 out of 1000 data exceed $\gamma = 3$ and would be falsely rejected as outliers.

We investigated the influence of the choice of $\gamma$ for a set of 80 686 *P*-wave delays in 57 clusters, each with at least 200 small eigenvalues, reported with zero precision (so as to maximize the expectation of larger errors). Rejecting data with $|r_i| > \gamma\sigma_{\text{init}}$, we find a mean $\sigma_e$ of $0.87 \pm 0.24$ if we use $\gamma = 3.0$, which is reduced to $0.81 \pm 0.22$ if we reject more outliers using a lower threshold $\gamma = 2.5$. Note that the two estimates are well within each uncertainty of one standard deviation, indicating that the choice of $\gamma$ is not very critical to our analysis. The number of outliers was 1379 and 1959, respectively, or 1.7 per cent versus 2.4 per cent outliers: a difference of 580 data that were accepted in the more permissive threshold of

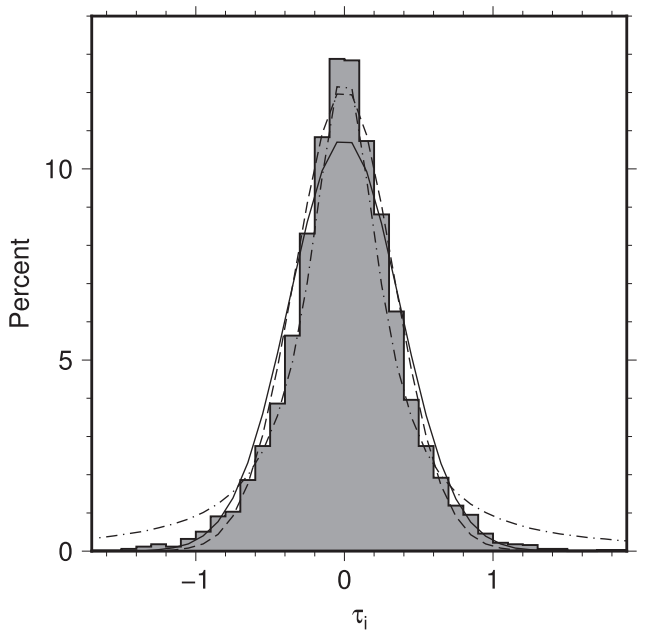


**Figure 5.** Three different distribution functions fitted to a randomly selected subset of 5053 transformed $P$ delays of slowness class A from mantle earthquakes, belonging to the smallest eigenvalues. Solid: a Gaussian curve with the standard error (eq. 4) as the standard deviation (0.37 s), dashed: idem but with the RSDR as standard deviation (0.33 s) and dotted–dashed: a Cauchy distribution with the same half-width as the histogram.

3 and may be in the grey zone of outlier/not outlier. We adopted $\gamma = 3.0$ as the threshold in this study.

We can fit a normal distribution to the $\tau_i$ with an estimated standard deviation $\sigma_{\rm init}$ equal to the standard error $s$ of the $\tau_i$:

$$s = \frac{1}{N_0}\sum_{i=1}^{N_0}(\tau_i - \bar{\tau})^2\,, \tag{4}$$

where $\bar{\tau}$ denotes the average.

Motulsky & Brown (2006) consider that the initial estimate (4) is very vulnerable to influence by the outliers. They use instead the robust standard deviation of the residuals (RSDR), which is the half-width of the central part of the distribution that contains 68.27 per cent of the $\tau_i$. In Fig. 5, we compare the Gaussians resulting from the two different estimates for $\sigma_{\rm init}$. We also attempted to fit a Cauchy distribution, known for having more data in its tail than a Gaussian. Its PDF is:

$$P(\tau) = \frac{1}{h\pi[1 + (\tau - \bar{\tau})^2/h^2]}\,, \tag{5}$$

where $h$ is the half-width, which we chose equal to the half-width of the observed distribution.

From the three curves in Fig. 5 we see that the Cauchy curve strongly overestimates the number of data present in the tail of the actual distribution, and in fact would suggest that there are no outliers. The dashed curve, based on RSDR fits best for the bulk of the data. The crossover where the theoretical curve is significantly below the observed distribution at the tails is at roughly twice the RSDR $\sigma_{\rm init}$. The solid curve, with the standard deviation $\sigma_{\rm init}$ equal to the standard error $s$, has a worse fit near the maximum, and instead attempts a better fit to the tails. We decided to use the RSDR to compute the outlier thresholds.

### 4.2 Evolution of $\sigma_e$ since 1964

In a first exploratory analysis, we ignored iprec (number of reported decimals) but tested the hypothesis that the data errors for the later years was significantly better than those for the epoch 1964–1999.

A clear indication that the quality of $P$ arrival times is actually fairly constant with time is given in Fig. 6, which summarizes the results of estimating $\sigma_e$ from 7317 matrices for event clusters. From these, 1957 event clusters had arrival times that were reported before the year 2000. Their distribution is plotted on the left. The remaining clusters represent arrival times associated with earthquakes in the 21st century and the histogram of their estimates for $\sigma_e$ is plotted on the right.

The mean $\sigma_e$ is $0.652 \pm 0.16$ s for 1968–1999, and $0.617 \pm 0.14$ for 2000–2018 (the uncertainty indicates one standard deviation in the distribution), which indicates a slight improvement over time. The medians are 0.633 and 0.612, respectively. After sorting the observations, we found that the $1\sigma$ (or 68 per cent) confidence intervals overlap largely: $\sigma_e$ is between 0.46 and 0.78 for the early data and 0.44–0.74 for the later data.

Although the difference between the two means is statistically significant (Student's $t$-test shows a probability of the order of $10^{-19}$ that this occurred by chance), the difference is so much smaller than the confidence intervals that we decided to ignore the epoch of the reported delays and focus instead on the influence of the iprec (after applying the correction mentioned earlier).

### 4.3 Dependence of $\sigma_e$ on reported precision (iprec)

The next question is how the error in $P$ arrival times depends on the reported precision, as indicated by the variable iprec (corrected, as explained earlier) for the same set of 7317 submatrices. Fig. 7 shows the distribution for a precision of 1 s (left), 0.1 s (centre) and 0.01 s (right), and here we see a larger difference than in the previous section. Table 2 lists the mean as well as the $1\sigma$ confidence interval inferred from the actual distribution. The difference between the means is significant according to Student's $t$-test. For example $\sigma_e = 0.8$ s for 0 and $\sigma_e = 0.56$ $s$ for 2 decimals gives $t = 47.8$.

The mean $\sigma_e$ and $1\sigma$ confidence intervals are listed in Table 2. When averaged over all three reported precisions, the average standard error in the $P$ phases is $0.63 \pm 0.12$ s.

We decided to subdivide the $P$ and $Pn$ data sets into those with zero decimal precision and those with one or two decimals.

### 4.4 Dependence on earthquake depth

We made a broad division into shallow quakes (<40 km) and deeper earthquakes. These populations have different errors for P delays, as is shown in the histograms of Fig. 8 and the estimates listed in Table 3, where $N_c$ denotes the number of clusters. The mean error in $P$ for deep earthquakes reported with iprec=−2 is the lowest at 0.50 s. When iprec=−1 the error rises to 0.58 s, almost the same as the 0.61 s for shallow earthquakes reported with one decimal precision. The table also lists the 68 per cent confidence interval as determined directly from the observed distribution, which differs somewhat from the $1\sigma$ interval because the distribution is not exactly normal.

### 4.5 Slowness correlation

Gudmundsson et al. (1990) observed a large scatter near an epicentral distance $\Delta \approx 20^\circ$ that is caused by triplications for rays

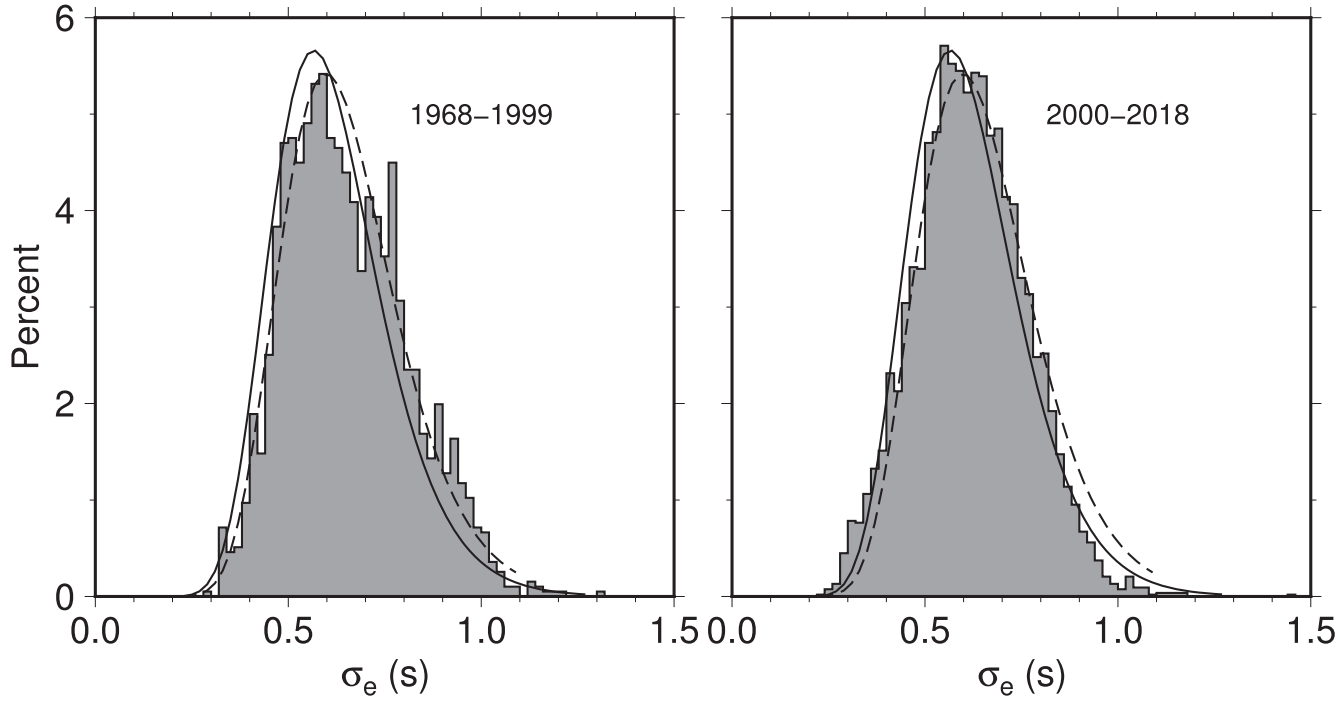


**Figure 6.** Left: histogram of $\sigma_e$ estimations for *P* data reported before the year 2000, plotted in intervals of 0.02 s. Right: for those reported in or after 2000. In both plots, the dashed curve is the best-fitting lognormal distribution for the 1968–1999 data (left-hand histogram), the solid curve for the 2000–2018 data (right-hand histogram).

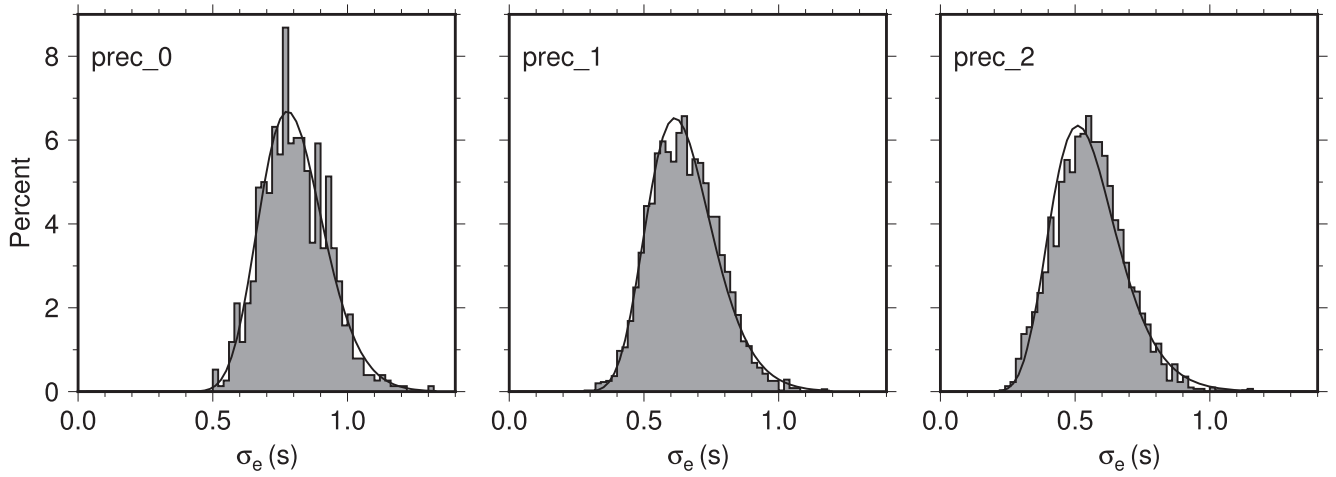


**Figure 7.** Histogram of $\sigma_e$ for the reported precision, from left to right: 0, 1 and 2 decimals. The smooth curve is the best-fitting lognormal distribution function.

**Table 2.** Confidence intervals for $\sigma_e$ for *P* delays as a function of reported precision

| iprec | *N* | Mean (s) | 68 per cent confidence (s) |
|---|---|---|---|
| 0 | 760 | 0.80 ± 0.12 | 0.66–0.90 |
| −1 | 3500 | 0.65 ± 0.12 | 0.51–0.76 |
| −2 | 3057 | 0.56 ± 0.13 | 0.40–0.66 |

bottoming near 400 km depth, resulting in a standard error of 1.4 s, which decreased substantially at teleseismic distances before increasing again as the ray approaches the core–mantle boundary. However, because of the variations in earthquake depth, the epicentral distance Δ does not uniquely predict where the ray bottoms and whether it will be triplicated or not. We therefore prefer to examine how the errors change as a function of ray slowness.

Using AK135 for ray tracing, we consider five slowness classes shown in Fig. 9. A sixth slowness class F was added to account for *P* waves that depart upward from a mantle source (Table 4).

We select slowness ranges such that the rays in each class encounter at most one significant discontinuity (410, 660 and D″).

In a small number of cases, our ray tracing failed to converge within the imposed number of iterations, notably for rays involved in upper mantle triplications (range C and D). We have made no effort to retrieve these rays.

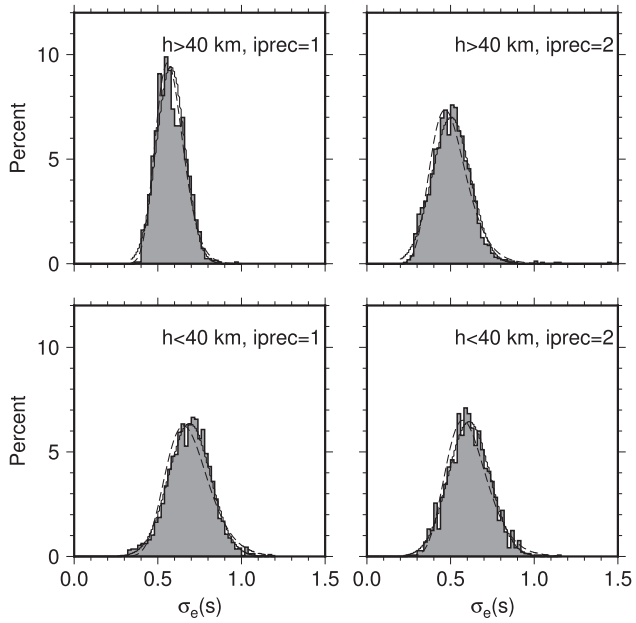


**Figure 8.** Histograms of $\sigma_e$ for mantle earthquakes (more than 40 km deep, top) and shallow quakes (bottom). Data with iprec=−1 are shown on the left, those with iprec=−2 on the right. The solid curve shows the best-fitting normal distribution function, to be compared to the lognormal function (dashed).

**Table 3.** Confidence intervals for $\sigma_e$ for $P$ delays as a function of depth

| $h$ | iprec | $N_c$ | Mean (s) | 68 per cent conf. (s) |
|---|---|---|---|---|
| <40 km | −1 | 2256 | 0.69 ± 0.13 | 0.52–0.80 |
| <40 km | −2 | 1450 | 0.61 ± 0.12 | 0.46–0.72 |
| Mantle | −1 | 1244 | 0.58 ± 0.08 | 0.46–0.64 |
| Mantle | −2 | 1607 | 0.50 ± 0.11 | 0.36–0.58 |

**Table 4.** Slowness ranges and turning point depths in AK135.

| | Depth range (km) | $P$ slowness (s rad$^{-1}$) | $S$ slowness (s rad$^{-1}$) |
|---|---|---|---|
| A | 2591–2891 | 255–280 | 478–526 |
| B | 1106–2591 | 280–453 | 526–816 |
| C | 635–1106 | 453–567 | 816–1032 |
| D | 385–635 | 567–670 | 1032–1240 |
| E | 35–385 | 670–788 | 1240–1414 |
| F | Upward | 0–788 | 0–1414 |

### 4.6 Later phases

As is clear from Table 1, the other phases present in EHB are far less abundant. To obtain a sufficient number of clusters $N_c$ with $N_0 \geq 100$ small eigenvalues, we ignore the subdivision for the number of reported decimals. We still are able to separate shallow and mantle quakes, but for the slowness, we divide the rays into those that may be influenced by the 20° discontinuity near 400 km depth (classes D and E in Table 4) and deeper rays.

For each phase, we extracted about 5000 transformed data $\tau_i$ from a few large clusters and plotted their histograms (shown in Figs S4–S7, Supporting Information). Most of the $\tau_i$ have a distribution of which the central part is well modelled by a Gaussian, using the RSDR as its width. For some later phases with few data, we noted an excess of zero $\tau_i$. We traced these back to duplicate entries in the EHB (their zero difference is a transformed datum from the null space of $\boldsymbol{A}$). We assume that these are identical readings from two different instruments in the same station, and suspect that stations with multiple sensors are also more likely to report later phases such as *PP*, which explains that their relative contribution is larger than for a phase such as *P* (an alternative explanation is that a same reading reached the ISC through different communication paths). We assumed that both readings are valid—in other words we accepted them as supporting a high precision of the arrival time in that station.

### 4.7 Procedure

The preliminary analyses described in this section motivated us to distinguish between 45 different classes of data divided over all mantle phases present in EHB. For each of these, we isolate as many clusters as we can ($N_c$) that have $N_0 \geq 100$ negligible eigenvalues. We remove outliers using the 2T algorithm with $\gamma = 3.0$. For each class of data, this give $N_c$ estimates of $\sigma_e$. These we plot in a histogram.

Each estimate $\sigma_e$ is drawn from a population that follows a Chi-square distribution with $N_0 - 1$ degrees of freedom. But since $N_0$ differs for each estimate, there is not one unique Chi-square distribution that predicts our observed histogram. Fortunately, since all $N_0$ are large, we may assume that the observed histograms are approximately normally distributed and we expect then the same for their joint PDF. $N_c$ is large, so that we can estimate $\sigma_e$ using the usual estimate for standard errors, as we did for $s$ in eq. (4).

## 5 FINAL RESULTS AND DISCUSSION

The results of our analysis are summarized in Table 5. Individual histograms are shown in Figs S8–S11 (Supporting Information), where the codes identify the classes listed in the table.

The difference between the mean and the median is rarely more than 0.1 s. But for the smaller populations ($N_c$ in Table 5) we see that the PDF of the $\sigma_e$'s may be far from normal, so for these the uncertainty in $\sigma_e$ should be viewed with caution.

The fraction of outliers hovers around 2 per cent for most cases, with a notable exception of 8.1 per cent *PKPbc* from mantle quakes, which could be due to frequent misidentification of *PKPab* as *PKPbc*. This may not be the only cause, however: the *PKPbc* from mantle quakes have a low RSDR which leads to a more restrictive outlier threshold and, once the outliers are removed, *PKPbc* has the smallest standard error of all phases. The error in PKPdf may be influenced by frequent misidentification of *PKPbc*, which has a large amplitude, as *PKPdf*.

Traveltimes from mantle earthquakes stand out by having significantly smaller errors than those from shallow ones, often only half of the error of shallow earthquakes in the same slowness and precision class. This is most likely due to the sharper onset of deep earthquakes. Indeed, the difference is far less pronounced in the errors for *PP*, a wave that never has a sharp onset because its waveform is the Hilbert transform of the P wave.

The error estimates in Table 5 are for the EHB data set, and cannot be uncritically adopted for the full ISC data set since EHB contains only a carefully selected subset of delay for re-analysed high-quality event locations. Tomographers using the $\sigma_e$ estimates from Table 5 should follow the same protocol as described in the previous section, notably estimating RSDR from a first regularized inversion and using a threshold of $3\sigma_e$ for the rejection of outliers.

**Table 5.** Standard error estimate $\sigma_e$ for phases reported in the EHB catalogue

| Phase | Deci-mals | Ray depth km | *h* | Label in Figs S4–11 (Supporting Information) | $N_c$ | Outliers per cent | $\sigma_e$ s | Median s |
|---|---|---|---|---|---|---|---|---|
| *P* | 1,2 | 2591–2891 | Mantle | Ad | 521 | 2.2 | 0.45 ± 0.14 | 0.43 |
| *P* | 0 | 2591–2891 | Mantle | Ad0 | 88 | 2.0 | 0.85 ± 0.24 | 0.82 |
| *P* | 1,2 | 2591–2891 | <40 | As | 453 | 2.9 | 0.85 ± 0.40 | 0.77 |
| *P* | 0 | 2591–2891 | <40 | As0 | 95 | 2.4 | 1.65 ± 0.44 | 1.59 |
| *P* | 1,2 | 1106–2591 | Mantle | Bd | 2540 | 2.6 | 0.43 ± 0.12 | 0.41 |
| *P* | 0 | 1106–2591 | Mantle | Bd0 | 372 | 2.1 | 0.79 ± 0.18 | 0.78 |
| *P* | 1,2 | 1106–2591 | <40 | Bs | 3667 | 2.6 | 0.69 ± 0.25 | 0.64 |
| *P* | 0 | 1106–2591 | <40 | Bs0 | 667 | 2.6 | 1.31 ± 0.35 | 1.25 |
| *P* | 1,2 | 635–1106 | Mantle | Cd | 1215 | 2.2 | 0.56 ± 0.13 | 0.56 |
| *P* | 0 | 635–1106 | Mantle | Cd0 | 195 | 1.8 | 0.87 ± 0.14 | 0.87 |
| *P* | 1,2 | 635–1106 | <40 | Cs | 1898 | 2.3 | 0.76 ± 0.19 | 0.75 |
| *P* | 0 | 635–1106 | <40 | Cs0 | 344 | 2.1 | 1.28 ± 0.27 | 1.25 |
| *P* | 1,2 | 385–635 | Mantle | Dd | 393 | 1.6 | 0.70 ± 0.16 | 0.69 |
| *P* | 0 | 385–635 | Mantle | Dd0 | 26 | 1.3 | 0.87 ± 0.21 | 0.80 |
| *P* | 1,2 | 385–635 | <40 | Ds | 532 | 1.4 | 0.81 ± 0.16 | 0.80 |
| *P* | 0 | 385–635 | <40 | Ds0 | 42 | 1.2 | 1.05 ± 0.16 | 1.04 |
| *Pn* | 1,2 | 35–385 | Mantle | Ed | 772 | 2.5 | 0.82 ± 0.25 | 0.79 |
| *Pn* | 0 | 35–385 | Mantle | Ed0 | 111 | 1.7 | 1.40 ± 0.49 | 1.28 |
| *Pn* | 1,2 | 35–385 | <40 | Es | 1433 | 2.9 | 0.94 ± 0.26 | 0.92 |
| *Pn* | 0 | 35–385 | <40 | Es0 | 211 | 2.3 | 1.53 ± 0.44 | 1.51 |
| *Pup* | 1,2 | 0–h | Mantle | Fd | 65 | 2.0 | 0.55 ± 0.15 | 0.52 |
| *Pup* | 0 | 0–h | Mantle | Fd0 | 5 | 1.6 | 0.68 ± 0.08 | 0.63 |
| *PKPab* | 0–2 | OC | Mantle | Kad | 107 | 3.0 | 0.59 ± 0.33 | 0.46 |
| *PKPab* | 0–2 | OC | <40 | Kas | 75 | 3.0 | 1.14 ± 0.49 | 0.95 |
| *PKPbc* | 0–2 | OC | Mantle | Kbd | 273 | 8.1 | 0.32 ± 0.10 | 0.29 |
| *PKPbc* | 0–2 | OC | <40 | Kbs | 239 | 3.9 | 0.53 ± 0.20 | 0.49 |
| *PKPdf* | 0–2 | IC | Mantle | Kdd | 892 | 3.1 | 0.88 ± 0.35 | 0.83 |
| *PKPdf* | 0–2 | IC | <40 | Kds | 995 | 2.8 | 1.15 ± 0.44 | 1.09 |
| *pP* | 0–2 | >635 | Mantle | Ldd | 217 | 3.0 | 1.18 ± 0.52 | 1.05 |
| *pP* | 0–2 | <635 | Mantle | Ldu | 68 | 2.8 | 1.51 ± 0.56 | 1.50 |
| *pP* | 0–2 | >635 | <40 | Lsd | 47 | 3.0 | 1.33 ± 0.58 | 1.17 |
| *pP* | 0–2 | <635 | <40 | Lsu | 15 | 2.2 | 1.68 ± 0.78 | 1.23 |
| *pwP* | 0–2 | >635 | Mantle | Wd | 36 | 3.7 | 1.43 ± 0.67 | 1.20 |
| *PP* | 0–2 | >635 | Mantle | PPdd | 21 | 1.7 | 2.22 ± 0.42 | 2.07 |
| *PP* | 0–2 | <635 | Mantle | PPdu | 11 | 1.6 | 2.28 ± 0.59 | 2.08 |
| *PP* | 0–2 | >635 | <40 | PPsd | 27 | 1.4 | 2.69 ± 0.43 | 2.79 |
| *PP* | 0–2 | <635 | <40 | PPsu | 20 | 1.6 | 2.54 ± 0.72 | 2.41 |
| *S* | 0–2 | >635 | Mantle | Sdd | 52 | 1.0 | 1.88 ± 0.77 | 2.04 |
| *S* | 0–2 | <635 | Mantle | Sdu | 145 | 1.1 | 1.74 ± 0.65 | 1.64 |
| *S* | 0–2 | >635 | <40 | Ssd | 70 | 0.7 | 2.60 ± 0.67 | 2.70 |
| *S* | 0–2 | <635 | <40 | Ssu | 78 | 0.4 | 2.82 ± 0.59 | 2.87 |
| *Sup* | 0–2 | 0-h | Mantle | FSd | 58 | 1.5 | 1.16 ± 0.48 | 1.02 |
| *Sn* | 0–2 | <635 | Mantle | Sndu | 274 | 2.0 | 1.55 ± 0.68 | 1.41 |
| *Sn* | 0–2 | <635 | <40 | Snsu | 346 | 2.1 | 1.77 ± 0.67 | 1.65 |
| *sP* | 0–2 | >635 | Mantle | sPd | 54 | 3.1 | 1.85 ± 0.54 | 1.86 |

The error estimates we report here now allow us for the first time to assess the reliability of a tomographic solution with an accurate and formal statistical measure of the fit. Assuming the data in the tomographic inversion are scaled with $\sigma_e$, the statistical quantity that describes the fit is $\chi^2$, or the *reduced*$\chi^2$, which is scaled by the number of data:

$$\chi^2_{\rm red} \equiv \frac{\chi^2}{N}\,. \tag{6}$$

Strictly speaking one should divide by $N - K$ instead of $N$, where $K$ is the rank of the tomographic matrix. Unfortunately, for large tomographic inversions it is computationally too expensive, or downright unrealistic, to compute $K$. But such large-scale tomographic problems generally have $N \gg M$. We know that $K \leq M$, and usually $K \ll M$. Therefore $K \ll N$ and thus one can usually accept (6) as approximate but sufficiently accurate for our purpose.

Statistical theory tells us that $\chi^2_{\rm red}$ has a distribution around a mean of 1 with a variance of $2/N$ (Nolet 2008, chapter 14). Since the variance is extremely small for large $N$, this seems to suggest that we must aim for a data fit with $\chi^2_{\rm red}$ exactly equal to 1. This, however, ignores the 'uncertainty in the uncertainty', that is, the uncertainty in $\sigma_e$.

The standard deviations for $\sigma_e$ listed in Table 5 give important new insight into the range of possible standard errors in EHB that permit us to reject or accept solutions depending on the fit to the data. The absolute uncertainties in $\sigma_e$ range between 0.08 and 0.78 s, but for the purpose of $\chi^2_{\rm red}$ we must consider the relative uncertainty in $\sigma_e$.

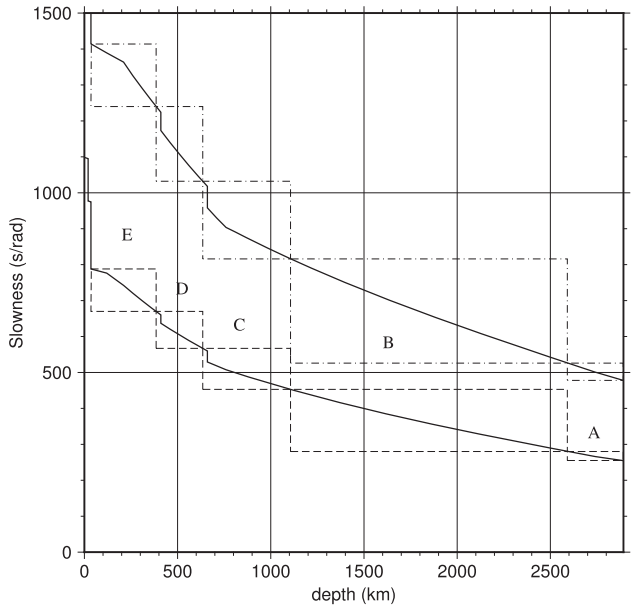


**Figure 9.** Slowness as a function of turning point depth in model AK135 for *P* and *S* waves (solid curves). The five slowness classes A–E (Table 4) are indicated by dashed lines for *P* and dotted–dashed lines for *S*.

Table 5 lists 22 error estimates for *P* and *Pn* (from 15 645 different clusters). The average relative uncertainty in these estimates, weighted with the number of clusters $N_c$ as a proxy for their frequency, is 23 per cent. But $\chi^2_{\rm red}$ is a quadratic function of the errors, and if we consider a solution acceptable if it is within a $1\sigma$ interval of possible values, we should aim for solutions within $(1 \pm 0.23)^2$ or $0.59 < \chi^2_{\rm red} < 1.52$.

If the tomography is done with all phases, the average relative errors are somewhat larger. Averaged over all 45 classes (from 19 707 different clusters), the relative error is 31 per cent and we shall wish that $0.48 < \chi^2_{\rm red} < 1.72$. If $\chi^2_{\rm red} > 1.72$ the damping or smoothing of the solution must be reduced. Conversely, if $\chi^2_{\rm red} < 0.48$, regularization must be strengthened to avoid overfitting. If a satisfactory $\chi^2_{\rm red}$ is not within reach even with little damping, it indicates a shortcoming of the model parametrization, and one needs to decrease voxel size, introduce anisotropy, or both (assuming source corrections were already part of the model parameters).

## ACKNOWLEDGMENTS

We thank Bob Engdahl for his advice and for commenting on a first draft of this paper. This research was started while SvdL was on sabbatical at Geoazur when both authors received support from the ERC, advanced grant 226837 'Globalseis'. We thank Eric Bergman and an anonymous reviewer for offering constructive comments.

## DATA AVAILABILITY

All data are available from the International Seismological Centre (http://www.isc.ac.uk/isc-ehb/).

## REFERENCES

Adams, R., Hughes, A. & McGregor, D., 1982. Analysis procedures at the International Seismological Centre, *Phys. Earth Planet. Int.,* **30,** 85–93.

Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS Trans. Am. geophys. Un.,* **81,** F897.

Charlety, J., Voronin, J.S., Nolet, G., Loris, I., Simons, F.J. & Daubechies, I.C., 2013. Global seismic tomography with sparsity constraints: Comparison with smoothing, damping regularization, *J. geophys. Res.,* **118,** 1–13, doi:10.1002/jgrb.50326.

Dziewonski, A., Hager, B. & O'Connell, R., 1977. Large-scale heterogeneities in the lower mantle, *J. geophys. Res.,* **82,** 239–255.

Engdahl, E., Giacomo, D.D., Sakarya, B., Gkarlaouni, C., Harris, J., J. & Storchak, D., 2020. ISC EHB1964-2016, an improved data set for studies of earth structure and global seismicity, *Earth Space Sci.,* **7,** e2019EA000897.

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

Gudmundsson, O., Davies, J. & Clayton, R., 1990. Stochastic analysis of global travel time data: mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.,* **102,** 25–44.

Jeffreys, H., 1936. On travel times in seismology, *Bur. Centr. Seism. Trav. S.,* **14,** 3–36 (reprinted in The Collected Papers of Sir Harold Jeffreys, Vol. 2, Gordon and Breach 1973).

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

Morelli, A. & Dziewonski, A., 1987. Topography of the core-mantle boundary and lateral homogeneity of the liquid core, *Nature,* **325,** 678–683.

Motulsky, H. & Brown, R., 2006. Detecting outliers when fitting data with nonlinear regression: a new method based on robust nonlinear regression and the false discovery rate, *BMC Bioinform.,* **7,** 123.

Nolet, G. *et al.*, 2019. Seismic tomography with floating seismometers: a first application to Galàpagos, *Sci. Rep.,* **9,** 1326, doi:10.1038/s41598-018-36835-w.

Nolet, G., 2008. *A Breviary of Seismic Tomography,* Cambridge Univ. Press, Cambridge, UK.

Röhm, A., Bijwaard, H., Spakman, W. & Trampert, J., 2000. Effects of arrival time errors on traveltime tomography, *Geophys. J. Int.,* **142,** 270–276.

Röhm, A., Trampert, J., Paulssen, H. & Snieder, R., 1999. Bias in arrival times deduced from ISC Bulletin, *Geophys. J. Int.,* **137,** 163–174.

Storchak, D., Giacomo, D.D., Engdahl, E., Harris, J., Bondar, I., Lee, W., Bormann, P. & Villasenor, A., 2015. The ISC-GEM global instrumental earthquake catalogue (1900-2009): introduction, *Phys. Earth Planet. Inter.,* **239,** 48–63.

Voronin, S., Mikesell, D., Slezak, I. & Nolet, G., 2014. Solving large tomographic linear systems: size reduction and error estimation, *Geophys. J. Int.,* **199,** 276–285.

Weston, J., Engdahl, R., Harris, J., Giacomo, D.D. & Storchak, D., 2018. ISC-EHB: reconstruction of a robust earthquake dataset, *Geophys. J. Int.,* **214,** 474–484.

Yang, J., Rahardja, S. & Fränti, P., 2019. Outlier detection: how to threshold outlier scores?, in *AIIPCC19: Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing, December 2019,* pp. 1–6.

## SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

**Figure S1**. The division of the Earth's surface into 10 406 patches of about $2^\circ \times 2^\circ$ size.

**Figure S2**. The location of patches generating at least 3000 matrix rows for *P*. The colour scale runs from 3000 (violet) to 30 000 (red). Of the 2011 patches shown here, 48 have more than 30 000 rows.

**Figure S3**. The location of 341 patches generating at least 3000 matrix rows for *Pn*. The colour scale runs from 3000 (violet) to 30 000 (red). Only one patch has more than 30 000 (located at 21°E, 37°N).

**Figure S4**. The distribution of a subset of about 5000 $\tau_i$ in the null space of $A$ for $P$ waves in slowness classes A, B and C. The codes in the top right corner correspond to the codes in Table 5. The solid curve denotes a Gaussian distribution with the RSDR computed from the observed histogram.
**Figure S5**. As Fig. S4, but for slowness classes D, E and F.
**Figure S6**. As Fig. S4, but for core phases and *pP*, *pwP* and *sP*.
**Figure S7**. As Fig. S4, but for *PP*, *S* and *Sn*.
**Figure S8**. The distribution of $\sigma_e$ estimates from patches that generate at least 100 negligible eigenvalues ($\lambda_i < 0.03$), for $P$ waves in slowness classes A, B and C. The solid curve denotes a Gaussian distribution with the RSDR taken from the observed distribution.
**Figure S9**. As Fig. S8, but for slowness classes D, E and F.
**Figure S10**. As Fig. S8, but for core phases and *pP*.
**Figure S11**. As Fig. S8 but for *PP*, *S* and *Sn*.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.